Aspects of Matching and Power in Group Randomized Trials

Andrew J. Dunning

A dissertation submitted in partial fulfillment
of the requirements for the degree of

Doctor of Philosophy

University of Washington

2001

Program Authorized to Offer Degree: Department of Biostatistics

UMI Number: 3013953

UMI®

UMI Microform 3013953

Bell & Howell Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

University of Washington

Graduate School

This is to certify that I have examined this copy of a doctoral dissertation by

Andrew J. Dunning

and have found that it is complete and satisfactory in all respects, and that any and all revisions required by the final examining committee have been made.

Chair of Supervisory Committee:

Ziding D. Feng

Reading Committee:

Ziding D. Feng

Paula K. Diehr

Arthur V. Peterson

Date: May 25, 2001

University of Washington

Abstract

Aspects of Matching and Power in Group Randomized Trials

by Andrew J. Dunning

Chair of Supervisory Committee:

Professor Ziding D. Feng
Department of Biostatistics

A group randomized trial is an experiment in which groups, such as communities, schools or workplaces, are randomized to receive treatment or act as controls. Frequently, the number of groups is small. Because it is the groups that are randomized, the statistical analysis may often be conducted using summary statistics for each group.

A key issue in the design of group randomized trials is whether a matched or unmatched design should be used. Because the number of groups is often small, any imbalance between the treatment and control groups with respect to some important characteristic is more apparent, and can lead to the suggestion that an observed treatment effect is due to imbalance rather than to the treatment. Matching can deflect that suggestion.

Matching has countervailing effects on the power of the trial. It potentially increases power by reducing variability, but reduces power because of the loss of degrees of freedom when the number of groups is small. The question therefore arises as to whether a matched design mandates a matched analysis, or whether the matching may be ignored in the analysis, and the power lost through loss of degrees of freedom recovered.

This dissertation examines the consequences of ignoring the matching in the analysis. Group summary statistics in a matched trial may be modeled by a correlated bivariate normal distribution, the correlation representing the effectiveness of the matching. We show that when the matching induces a negative correlation, ignoring the matching leads

to an anti-conservative test, and thus under such circumstances may not be ignored. When the matching induces a positive correlation, the matching may be ignored.

If the matching induces a positive correlation, ignoring the matching increases power when the correlation is small; retaining the effect of matching in the analysis increases power when the correlation is large.

We advance a theorem that, when the matching is ignored, the power of the experiment increases with the correlation when the correlation is positive and the power is more than 50%.

We derive and evaluate certain adaptive tests that adjust for the estimated correlation.

TABLE OF CONTENTS

List of Figures v

List of Tables vii

Chapter 1: Introduction 1

1.1 Group Randomized Trials 1

1.2 Matching in Group Randomized Trials 3

1.3 Models and Parameterizations for Matched Designs 11

1.4 This Dissertation 14

Chapter 2: Notation 21

2.1 Observations 21

2.2 Statistics 21

2.3 Tests 22

2.4 Functions and Symbols 22

2.5 Pochhammer's Symbol 23

Chapter 3: Bias of the MU Procedure 25

3.1 The Distribution of $T_2|\rho$ under the null in computable form 25

3.2 Evaluation of the bias 30

Chapter 4: A Theorem concerning the Power of the MU procedure 31

4.1 Illustration of the Theorem 31

4.2 Statement of the Theorem 31

4.3 General Comments 32

4.4 Asymptotic Proof 34

4.5 Proof for small δ when $\beta=\frac{1}{2}$ 35

4.6 Proof when $c=0$ 38

4.7 Induction 40

4.8 Attempted proofs at $\rho=0$ 41

4.9 Proof when $\rho=0$ and $\delta=0$ 42

4.10 Total Positivity 42

4.11 When $\rho=0$: Remove alternating signs 46

4.12 When $\rho=0$: Contiguous Horn functions 47

4.13 When $\rho=0$: Parallel to the asymptotic proof 48

4.14 Demonstration of the Theorem for random parameter values 53

Chapter 5: A Test based on $\bar{\bar{Y}}|$V **55**

Chapter 6: Tests based on $T_2|\rho$ **58**

6.1 ρ known 58

6.2 Specifying $\rho \neq 0$ 58

6.3 Specifying r for ρ 59

Chapter 7: Tests based on $T_2|r$ **61**

7.1 Generalizing to $\rho\neq 0$. 62

7.2 Specifying r for ρ 63

Chapter 8: Other conditional tests **65**

8.1 Test based on $T_2|r_2$ 65

8.2 Test based on $T_2|r_3$ 66

Chapter 9: Kenward and Roger's Method **70**

9.1 Background 70

9.2 Application to our problem 72

Chapter 10: Flexible Conditional Tests 79

10.1 Introduction 79

10.2 Choice of c_r 82

10.3 Linear function of r 83

10.4 c_r as a second and third order power series in r 86

10.5 Two flexible functions for c_r 88

Chapter 11: A Two-stage Procedure 92

11.1 Background 92

11.2 Derivation 93

11.3 Evaluation 95

Chapter 12: Discussion 98

12.1 The MU procedure 98

12.2 Other Tests for Matched Designs 99

12.3 Matching and Group Randomized Trials 104

12.4 Further work 106

Bibliography 108

Appendix A: The Distribution of $T_2|\rho$ 113

A.1 The Distribution Function of $T_2|\rho$ as a Power Series 113

A.2 The distribution function of $T_2|\rho$ as a guaranteed convergent power series suitable for computing 121

A.3 The Density of $T_2|\rho$ 124

A.4 The Asymptotic Distribution of $T_2|\rho$ 126

A.5 The Derivative with respect to ρ of the distribution function of $T_2|\rho$ 130

A.6 Confirmation of the Distribution of $T_2|\rho$ for negative ρ by simulation 131

A.7 Splus code for the Distribution Function of $T_2|\rho$ 132

A.8 Splus code for the Density Function of $T_2|\rho$ 135

Appendix B: The Distribution of $T_2|r$ **137**

B.1 The Distribution Function of $T_2|r$ 137

B.2 A convergent form of the distribution function 143

B.3 The density function of $T_2|r$ 143

Appendix C: The Distribution of $T_p|r$ **144**

C.1 The Distribution Function of $T_p|r$ 144

C.2 A convergent form of the distribution function 152

Appendix D: Negative Correlation from Treatment×matching Interaction153

Appendix E: Equating Coefficients **155**

LIST OF FIGURES

1.1 The relative power of matched and unmatched experiments. 9

1.2 The power of the paired t test and the two-sample t test when the data come from a bivariate normal distribution. 16

4.1 Power curves for UU and MU procedures, for various ρ, $n=7$ pairs, nominal test size .05. 32

4.2 The value of the power and derivative functions over a range of δ. 34

4.3 The approximately linear relationship of c and δ when $\beta=\frac{1}{2}$. 38

4.4 The value of the function discussed in 4.13, by n. 50

5.1 Illustration of the joint conditional density of $\bar{y}_{\cdot 1}$ and $\bar{y}_{\cdot 2}$ given v on the region of integration. 57

6.1 Bias of a test based on $T_2|\rho$, $n=5$ pairs, nominal size .05. 59

7.1 Bias of a test based on $T_2|r$, $n=5$ pairs, nominal size .05. 63

8.1 Estimated .975 quantiles of the distribution of $T_2|r_2$, $n=5$ pairs. 66

9.1 Power of a test based on Kenward & Roger's modified method, compared to the power of a two sample test and a paired test, for various ρ, $n=5$ pairs, nominal size .05, effect size 1.5. 75

9.2 Power of a test based on Kenward & Roger's standard method, compared to the power of a two sample test and a paired test, for various ρ, $n=5$ pairs, nominal size .05, effect size 1.5. 78

10.1 Critical values of tests based on $T_2|r$ and $T_2|\rho$, $n=5$ pairs, nominal size .05. The potential critical values illustrate a set of c_r to be evaluated. 81

10.2 Eight sets of linear critical values for a test based on $T_2|r$, $n=5$ pairs, nominal size .05, unbiased at $\rho=0$. 84

10.3 Power of a test using the $c_{r(3)}$ linear critical values, compared to the power of a two sample test and a paired test, for various ρ, $n=5$ pairs, nominal size .05, effect size 1.5. 86

10.4 Critical values for a test based on $T_2|r$ using a third order quadratic function, $n=5$ pairs, nominal size .05. 87

10.5 Critical values for a test based on $T_2|r$ using a smoothed step function, $n=5$ pairs, nominal size .05. 89

10.6 Power of a test with smoothed step critical values, compared to the power of a two sample test and a paired test, for various ρ, $n=5$ pairs, nominal size .05, effect size 1.5. 90

10.7 Critical values for a test based on $T_2|r$ using a flexible function, $n=5$ pairs, nominal size .05. 91

11.1 Power of a two stage procedure, compared to the power of a two sample test and a paired test, for various ρ, $n=5$ pairs, nominal size .05. 96

LIST OF TABLES

1.1 Estimated matching correlation between group level outcomes for seven pair matched group randomized trials. 7

3.1 Bias of the MU procedure for various ρ and n, nominal size .05. 30

6.1 Bias of a test based on the distribution of $T_2|\rho$ using r as a 'plug-in' estimator of ρ, for various 'true' ρ, $n=5$ pairs, nominal size .05. 60

7.1 Bias of a test based on the distribution of $T_2|r,\rho$ using r as a 'plug-in' estimator of ρ, for various 'true' ρ, $n=5$ pairs, nominal size .05. 64

9.1 Bias of a test based on Kenward & Roger's modified method, for various ρ, $n=5$ pairs, nominal size .05. 74

9.2 Bias of a test based on Kenward & Roger's standard method, for various ρ, $n=5$ pairs, nominal size .05. 77

10.1 Bias of tests based on $T_2|r$ and linear critical values, $n = 5$ pairs, nominal size .05. 85

10.2 Bias of a test with smoothed step critical values, for various ρ, $n=5$ pairs, nominal size .05. 89

11.1 Bias of a two stage procedure based on r, for various ρ, $n=5$ pairs, nominal size .05. 96

12.1 Comparison of four tests, each unbiased at $\rho=0$; $n=5$ pairs, nominal size .05, effect size 1.5. 101

A.1 $P(T_2|\rho<t)$ for negative ρ by simulation and by computation. 132

Chapter 1

INTRODUCTION

1.1 Group Randomized Trials

A group randomized trial is an experiment in which groups, such as communities, schools and workplaces, rather than individuals, are randomized to receive treatment or act as controls. This is often because of the nature of the treatment. For example, in the Working Well Trial [2], interventions designed to reduce smoking and improve nutrition, including promotional and educational events, information and education materials, and campaigns and contests, were implemented on a workplace-wide basis. In the Kaiser Family Foundation's Community Health Promotion Grant Program (the Kaiser CHPGP) [41], community health intervention programs and funding directed at reducing cardiovascular disease, cancer, substance abuse, adolescent pregnancy and injuries, were provided on a community-wide basis. In the Hutchinson Smoking Prevention Project trial [30], interventions included teacher-led curriculum units, teacher training and motivation and self-help tobacco use cessation materials; school districts were randomized to treatment or control.

Frequently, the number of groups in a group randomized trial is small, because the cost of school-, workplace- or community-wide interventions, and the cost of collecting data from all the members of the group, or at least a representative cross-section, puts pressure on the resources available. The Community Intervention Trial for Smoking Cessation [17] (COMMIT) consisted of 11 pairs of communities. The Minnesota Heart Health Program [21] comprised three pairs of communities. The Kaiser CHPGP involved 18 communities.

1.1.1 Analysis of Group Randomized Trials

A distinguishing feature of group randomized trials is that it is the groups which are randomized to treatment or control, rather than the individuals. The statistical analysis of a trial must take this randomization by group into account; failing to do so can generate misleading results. Cornfield [8], in a much quoted precept, stated "Randomization by cluster accompanied by an analysis appropriate to randomization by individual is an exercise in self-deception". ('Clusters' here is a another term for 'groups').

One way of taking account in the analysis of the randomization by group is intra-class correlation. The members of each group are considered to have characteristics in common deriving from their membership of the group, and their responses are consequently similar. In statistical terms, such similarity is measured by intra-class correlation, the correlation between responses from members within the same group. Omitting to include this correlation in the statistical analysis can lead to overstated results. Donner et al. [10] found that in 8 of 16 trials examined, no account was taken of the intra-class correlation, and the significance of the results was thus seriously overstated.

Another approach to the analysis which takes into account the randomization by group is to base the analysis on measures which summarize the responses for each group. Such summary statistics might be group means, proportions or rates. This approach accords with Fisher's admonition to "analyze as you randomize" [15]. The units of randomization are the groups; the analysis is conducted based on group summary statistics.

If the number of individuals in each group is large (perhaps 50 or more), such group summary statistics may validly be modeled as normally distributed by virtue of the central limit theorem.

Koepsell [27] showed that an analysis that treats group means as the unit of analysis is equivalent to a mixed model analysis based on individual responses if the group effects are modeled as random effects and the group sizes are equal.

Group effects are normally treated as a random effects in the analysis of group randomized trials, because it is usually intended that the results of a trial should be generalizable to other groups in addition to those in the trial. That is, it is intended that the inference from the trial should extend to some 'population of groups' from which these particular ones are considered a sample, and to achieve this, group effects are treated as random effects in the analysis.

1.2 Matching in Group Randomized Trials

A key design issue in group randomized trials is whether a matched or unmatched design should be used. Because the number of groups in each arm is often small, any imbalance between the treatment and control groups with respect to some important characteristic is more apparent. Such imbalance can lead to the suggestion that an observed treatment effect is due to the imbalance rather than to the treatment. For example, if 10 communities are randomized to treatment or control, and it happens that 4 predominantly urban communities are randomized to the treatment arm and 4 predominantly rural communities to the control arm, the suggestion might arise that any change in behavior observed is due to the urbanness of the treatment groups rather than to the treatment itself.

Imbalance does not necessarily mean that an observed treatment effect is is spurious. An observed treatment effect can only be the result of the imbalance if there is some association between the characteristic which is out of balance and the outcome being observed.

However, because there is a multiplicity of characteristics by which the groups in a trial may be characterized, countering the suggestion that an observed result is due to an imbalance can be problematic, particularly when the number of groups is small. For example, communities may be characterized by their urban or rural character, their population size, average income, average educational level, age distribution, geographical location or a number of other criteria. In any randomization to treatment and control when the number of communities is small, it is quite likely that the treatment groups and the control groups will be out of balance with respect to one of these characteristics.

A pair matched design is therefore frequently used. Groups having similar characteristics are paired together. Then one member of each pair is randomly assigned to treatment or control. The other member of the pair takes the other assignment.

In practice, evidence that there is an association between the matching variable or variables and the outcome variable is often weak at best. Dunning and Diehr (1998) [11] found some evidence of associations between changes in various health behaviors and some potential matching variables, but the strength of the associations was generally weak. Further, matching according to one characteristic, such as the rural/urban character of the communities, does not guard against imbalance on some other characteristic, such as average age of the population.

Nevertheless, in practice matching does deflect the suggestion that an observed treatment effect is due to imbalance; for this reason, matching to avoid the suggestion of invalidity has sometimes been described as 'defensive matching'.

Just as the similarity between responses from individuals within groups is measured by the intra-class correlation, so the similarity between groups in the same matched pair is measured by the matching correlation, or the correlation induced by matching. When the matching correlation is high, the matching is said to be 'effective'.

Freedman et. al. (1990) [16] found strong gains in the power of a trial could be achieved in a matched pair design when the matching correlation was high. He showed that in the COMMIT trial, a gain in efficiency of the order of 50% (that is, an increase in the power of the trial equivalent to a 50% increase in the number of communities) could have been achieved if the communities had been paired according to the quit smoking rates for the previous five years, and the quit smoking rate over the course of the trial (one of the primary endpoints of the trial) had been strongly correlated with the rate for the previous five years.

Gail et. al. (1996) [18] examined randomization based inference for community intervention trials. He found that under a variety of assumptions, randomization tests of pair matched designs were almost always unbiased, as were paired t tests when the number of groups was so small that randomization tests were not possible. They examined the performance of various matched and other designs and analyses against both the 'weak' null hypothesis that the 'average' effect of treatment was zero and the 'strong' null hypothesis that treatment had no effect on any of the treatment groups.

Klar and Donner (1997) [25] found a number of limitations arising from pair matching in group randomized trials. They note the loss of power arising from reduced degrees of freedom in matched pair designs, as is discussed below. They recommend a two strata design, that is, a design where the communities are segregated into two strata according to chosen matching criteria, and randomized to treatment or control within each strata. They argue that this involves minimal loss of degrees of freedom, and that allowing a stratum effect in the model resolves other statistical and analysis difficulties.

Design and Analysis of Group Randomized Trials (Murray, 1998) [33] is a comprehensive and detailed book bringing together much of the work on Group Randomized Trials published over the last several years and setting it in an ordered context. Planning of the trial, including an extensive section on sources of bias, and different research designs are covered. In the chapter 'Planning the Analysis', statistical issues and methods for the analysis of group randomized trials are presented, including Generalized Linear Mixed Models and General Estimating Equations. There are several chapters on the analysis of different types of trials, stratified, matched, cohort, cross-sectional, pre- and post-test, and extensive examples of analysis methods are given for each type.

Murray does not find the limitations on matched designs mentioned by Klar and Donner, although he does note that it is not possible to distinguish non-homogenous intervention effects (treatment×pair interaction effects) from group effects in pair matched designs. He reports the findings of Martin and Diehr concerning loss of degrees of freedom in matched designs and the need to weigh the resulting loss of power against the gain in power from the reduction of variability by matching.

Murray also notes the desirability of matching or stratifying to strengthen the design against common sources of bias, and highly recommends this procedure for group randomized trials with a limited number of groups, including any study with fewer than 20 groups per condition. He reports Martin's finding that matching will increase power of the analysis when the correlation between the matching and outcome variables is more than .30 or there are more than 10 groups per condition; then, the increased efficiency of the matched analysis will usually offset the reduced degrees of freedom, as further discussed below.

Feng et. al. (1999) [14] examine the data from four large community based group randomized trials and surveys. Variance component methods are used to estimate the community level variance components, and to assess the extent to which between-community variance could be reduced by inclusion of covariates in the analysis. When the community effect is properly treated as a random effect in the analysis, the size of the estimated community level variance is used to assess the significance of the intervention effect, measured at the group level. Thus, reduction of community level variance by inclusion of community level covariates increases the power of the experiment.

Feng found that matching or blocking by covariates associated with the outcome of interest improved power when the outcome of interest is measured by a single cross sectional survey at the conclusion of the trial. On the other hand, when the outcome of interest is measured by a change between two time points, at the beginning and end of the trial, that is, a pre-test post-test design, matching was less effective in increasing power. It appeared that taking the difference in outcome between the two time points eliminated the effect of the covariates on which matching was based. In fact, in several instances, power was lost, apparently because of loss of degrees of freedom. Matching might still be considered as a means to increase power in a two-timepoint study if covariates known to affect the time trend of the outcome variable could be found.

The first three chapters of the book *Design and Analysis of Cluster Randomization Trials in Health Research* (Donner and Klar, 2000) [26] provide a broad and thorough review of group randomized trials. They note 'the well accepted design principle (e.g. Pocock, 1983) that stratification is most effective in small studies, a principle that is particularly relevant to cluster randomization trials'.

They discuss the work of several authors on the gains in efficiency that may be achieved by matching, and provide a table showing the estimated matching correlation between group level outcomes for seven pair matched group randomized trials. The table is reproduced at table 1.1. The matching correlations shown are estimates and subject to the wide variability of empirical correlation estimates when the number of pairs is small.

Several of the analysis methods for group randomized trials they propose rely on estimates of intra-class correlation to calculate variance inflation factors. This is a limitation in pair-matched trials where the treatment$\times$pair interaction effects cannot be distinguished from group effects, and thus an estimate of their variance must be the variance of both effects together. Since the estimate of the intra-class correlation is made from the estimate of the variance of the group effects, it cannot be separately estimated in a pair matched

Table 1.1: Estimated matching correlation between group level outcomes for seven pair matched group randomized trials (from Donner and Klar, *Design and Analysis of Cluster Randomization Trials in Health Research* [26]).

Source	Unit of randomization	Number of pairs	Outcome variable	Matching correlation
Stanton & Clemens (1987)	Cluster of families	25	Childhood rate of diarrhea	0.49
Bass et. al. (1986)	Physician practice	17	Death rate	0.41
Thompson et. al. (1997)	Physician practice	13	Levels of coronary risk factors	0.13
Ray et. al. (1997)	Nursing home	7	Rate of recurrent falling	0.63
Haggerty et. al. (1994)	Community	9	Childhood rate of diarrhea	-0.32
Grosskurth et. al. (1995)	Community	6	HIV rate	0.94
COMMIT Research Group (1995)	Community	11	Smoking quit rate	0.21

design, and hence statistical methods which rely on variance inflation factors based on estimated intra-class correlation cannot be used. However. other methods of analysis, such as mixed models and generalized linear mixed models can avoid this limitation for estimation of and inference on treatment effects.

Murray et. al. (2000) [32] published a detailed examination of the design and analysis of the REACT trial, a trial to assess the effects of interventions on the time from onset of symptoms of myocardial infarction to arrival at a hospital emergency department. The trial was a matched pair design with 20 communities. 10 in each arm. It was noted that effective matching can control for the potential confounding effect of the matching factors and improve the precision of the estimate of the intervention effect; ineffective matching can reduce power through the loss of degrees of freedom. They evaluated the trial using both a matched and an unmatched analysis, and found virtually no difference in the estimates of the primary endpoint.

Bellamy et. al. (2000) [4] examined alternative methods for analyzing the dichotomous data from the Cancer Action in Rural Towns program, a group randomized trial of 20 Australian rural towns. The design was pair matched, but the analysis chose to ignore the effect of matching, arguing that since the purpose of the matching had been to reduce the variability in the responses and gain power, an as the actual gain turned out to be very small, the matching could be ignored.

A number of recent publications have emphasized the ascertainment and consequences of non-homogenous intervention effects in the analysis of group randomized trials. Non-homogenous intervention effects are when the effect of the intervention (or treatment) is not the same for all treatment groups; or, in a stratified design. for all strata. The 'strong' null hypothesis is that treatment has no effect on any community mean. However, as Gail [18] points out, rejecting the 'strong' null and concluding that some of the treatment community means were not zero, is of little scientific interest if the overall effect of treatment was not significantly positive. "One would not ordinarily want to adopt a treatment that produced no improvement in average community response. even though it might influence some communities favorably and others unfavorably."

1.2.1 Power of Matched vs. Unmatched Designs

When the number of observations in an experiment is large. many test statistics have a normal distribution, but as the number of observations declines the distribution moves away from normal. The critical values of the test get larger. The change in the distribution is said to result from the loss of 'degrees of freedom' of the test statistic.

In a matched pairs experiment, the number of degrees of freedom is half that of an unmatched design. This has little effect when the number of experimental units is large. However, when the number of experimental units is small. as is often the case in a group randomized trial, the less degrees of freedom of the matched design reduces the power of the experiment to detect a real treatment effect.

Thus in group randomized trials, the effect of matching. that is, the effect of a matched design compared to an unmatched, is twofold. The matching potentially increases the power of the experiment by reducing the variability of the observations, but power is reduced because of loss of degrees of freedom.

Martin et. al (1993) [31] quantified these effects. He showed that in the case of normally distributed data, if the matching was only mildly effective, that is, if the correlation induced by matching was small, the power of a matched experiment was less than could be obtained with an unmatched experiment for the same number of groups. When the matching was more effective, the power gained from reduction of variability offset the loss from reduced degrees of freedom.

Figure 1.1 shows the break even matching correlation for an experiment with nominal test size 0.05 and 80% power. In each case, the test is the usual test for the design, a paired t test for the matched experiment and a two sample t test for the unmatched experiment. The power is the probability that the test will conclude that a treatment effect is real.

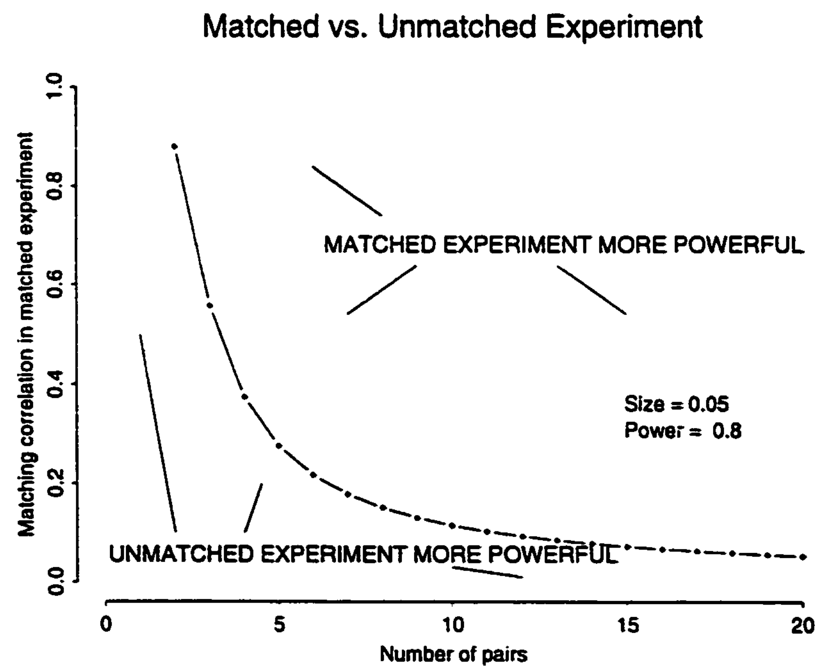


Figure 1.1: The relative power of matched and unmatched experiments (after Martin et. al (1993) [31]).

Martin provided figures for other test sizes and power percentages.

1.2.2 Matched Design with Unmatched Analysis

Diehr et. al. [9] considered whether a matched experiment mandates a matched analysis. If an investigator matched for 'defensive' reasons - to deflect criticism that the groups appeared unbalanced - yet believed that in fact the effectiveness of the matching was low, did he have to perform an analysis appropriate for a matched design, or could the matching be ignored and an analysis appropriate for an unmatched design conducted, thus recovering the loss of degrees of freedom?

Using Martin's bivariate normal model, Diehr evaluated the consequences of using an unmatched analysis when the data came from a matched design. She used the following notation to denote combinations of matched and unmatched experimental designs with matched (paired t test) and unmatched (two-sample t test) statistical analyses:

	Two sample t test	Paired t test
Matched experiment	MU procedure	MM procedure
Unmatched experiment	UU procedure	

Diehr evaluated the bias and power of the MU procedure, that is a two-sample test when the data came from a correlated bivariate normal distribution, by simulation. She found that the two-sample test (MU procedure) is conservative when the matching correlation is positive (that is, that a test with nominal size 0.05 rejects less than 5% of the time when the null is true). The MU procedure was sometimes more powerful than a UU procedure and sometimes less. In other words, if an unmatched analysis was performed, the effect of matching the experiment was to sometimes increased power and sometimes not. Further, if the experiment was matched, an unmatched analysis (MU) was sometimes more powerful than a matched analysis (MM) and sometimes less.

It may be asked why the matching correlation may not be estimated from the data, and the more powerful analysis then used. Such an approach is poor practice for two reasons.

Firstly, estimates of correlation coefficients are widely variable for a given true correlation ρ when the number of observations is small. Secondly, multiple comparison issues arise. A procedure which tests the data twice, once to decide which test to use and once to carry out the test, can be anti-conservative. That is, it can reject the null hypothesis more than 5% of the time when the hypothesis is in fact true, (for a nominal 0.05 test size).

Diehr confirmed that a procedure which first tested whether the sample correlation coefficient was significantly different from zero, then performed a paired test if it was, or a two sample test if it was not, was anti-conservative.

Analytical evaluation of the various options was hindered because the distribution of the two-sample t statistic when the data came from a bivariate normal distribution was not known. Proschan [34] derived that distribution in 1996.

1.3 Models and Parameterizations for Matched Designs

In *Testing Statistical Hypotheses* [29], p. 265, Lehmann models matched and unmatched designs as additive random effects models. For the matched design, with n pairs, $i=1,\ldots,n$,

$$\begin{aligned} Y_{i1} &= U_i + V_i \quad \text{for the controls} \\ Y_{i2} &= U_i + V_{n+i} \quad \text{for the treatment group.} \end{aligned}$$

For an unmatched design,

$$\begin{aligned} Y_{i1} &= U_i + V_i \quad \text{for the controls} \\ Y_{i2} &= U_{n+i} + V_{n+i} \quad \text{for the treatment group,} \end{aligned}$$

where

a) $U \sim N(\mu, \sigma_1^2)$ represents the unit (or matching) effects;

b) $V \sim N(\xi, \sigma^2)$ represents the experimental effects in the control group; and

c) $V \sim N(\eta, \sigma^2)$ represents the experimental effects in the treatment group.

Lehmann's models may be restated as bivariate normal models; for the matched case

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} \sim N_2\left(\begin{matrix} \mu+\xi \\ \mu+\eta \end{matrix} , \begin{bmatrix} \sigma_1^2+\sigma^2 & \sigma_1^2 \\ \sigma_1^2 & \sigma_1^2+\sigma^2 \end{bmatrix} \right) .$$

In the unmatched case

$$\begin{pmatrix} Y_1' \\ Y_2' \end{pmatrix} \sim N_2\left(\begin{matrix} \mu+\xi \\ \mu+\eta \end{matrix}, \begin{bmatrix} \sigma_1^2+\sigma^2 & 0 \\ 0 & \sigma_1^2+\sigma^2 \end{bmatrix} \right).$$

In the matched model, the matching correlation (the correlation between the treatment and control observations) $\rho = \frac{\sigma_1^2}{\sigma_1^2+\sigma^2}$ is constrained to be positive by the requirement that the variance of the unit or matching effects σ_1^2 must be positive.

Martin and Diehr similarly modeled matched pairs as a bivariate normal distribution. They hypothesize a continuous matching variable X correlated ρ_{xy} with the outcome variables Y_1 and Y_2; the outcome variables are correlated $\rho_{yy} = \rho_{xy}^2$ with each other, and thus with an assumption of normality may be modeled with a bivariate normal distribution with positive correlation.

Smith & Murray [37] give an example of a matched experiment where the correlation between the pairs was estimated to be to be negative. Forty pairs of twin calves were weighed at the time of weaning to ascertain whether there was a 'mother' effect, whether some mother cows were better mothers to their calves than others. Mother effects were treated as random effects. However, partitioning the sums of squares yielded a negative estimate for the variance of the mother effect.

Smith & Murray showed that the model could be reparameterized as a bivariate model. The negative variance estimate then became a negative correlation between the pairs of calves, suggesting that within pairs, one calf gained weight at the expense of its twin, possibly because they competed for their mother's nutrition. Further, reparameterizing the model this way allows for a two-sided F test to be performed, with in their case a significant result.

Hocking [20] showed that any additive random or mixed effects model could be reparameterized as a multivariate model. Under such reparameterizations, negative estimates of variance components become negative correlations. Searle, Casella & McCulloch [36] question the generality of such reparameterizations, noting that an estimated negative correlation is not unrestrained, but is limited by the requirement that the estimated variance-covariance matrix be positive definite. Pukelsheim & Styan [35] showed that the positivity of the sums of squares in the variance decomposition guaranteed the positive definiteness of the estimated variance-covariance matrix.

An additive random effects model (without reparameterization or negative variances) may also be shown to allow a negative correlation between the observations in the treatment and control groups when the treatment×matching interaction effect is greater than, and in the opposite direction from, the main treatment effect, as follows.

Let X be the matching variable, Y be the response. Suppose

$$Y = \beta_0 + \beta_1 X + \beta_2 I_{(\text{treatment})} + \beta_3 X I_{(\text{treatment})} + \epsilon$$

where $X \sim (0, \sigma_x^2)$, $\epsilon \sim (0, \sigma^2)$, $X \perp \epsilon$, and

a) β_0 represents the overall mean;

b) β_1 represents the matching effect;

c) β_2 represents the treatment effect;

d) β_3 represents the treatment×matching interaction;

Put $Y_1 \equiv Y|I_{(\text{treatment})} = 0$, $Y_2 \equiv Y|I_{(\text{treatment})} = 1$. Then it may be readily shown that

$$\begin{aligned}
Var(Y_1) &= \beta_1^2\sigma_x^2 + \sigma^2 \\
Var(Y_2) &= (\beta_1 + \beta_3)^2\sigma_x^2 + \sigma^2 \\
Cov(Y_1, Y_2) &= \beta_1(\beta_1 + \beta_3)\sigma_x^2 \\
Corr(Y_1, Y_2) &= \frac{\beta_1(\beta_1 + \beta_3)\sigma_x^2}{\sqrt{(\beta_1^2\sigma_x^2 + \sigma^2)((\beta_1 + \beta_3)^2\sigma_x^2 + \sigma^2)}}
\end{aligned}$$

$$\begin{aligned}
Corr(Y_1, Y_2) < 0 &\Leftrightarrow \beta_1(\beta_1 + \beta_3)\sigma_x^2 < 0 \\
&\Leftrightarrow \beta_1(\beta_1 + \beta_3) < 0 \\
&\Leftrightarrow |\beta_3| > |\beta_1| \text{ AND } \beta_1\beta_3 < 0
\end{aligned}$$

What this means is that the correlation between two outcome variables can be negative if the treatment×matching interaction effect is greater than, and in the opposite direction from, the main treatment effect.

Details of the derivations are given in Appendix D.

Wacholder and Weinberg [40] give a similar example of when a negative matching correlation is possible. Consider a clinical trial comparing one treatment with another, and suppose the trial is matched according to age. If one treatments is more effective in older subjects than in younger, while the other treatment is more effective in younger subjects, then the matching correlation will be negative.

1.4 This Dissertation

This dissertation first sets out to answer the questions posed, and partially answered, by Diehr.

1. Should a group randomized trial be matched? Under what circumstances is a matched design preferable to an unmatched design?

2. If a trial is matched, should a matched or an unmatched analysis be used? What are the consequences of conducting an analysis appropriate for an unmatched experiment if the experiment is matched?

We use the bivariate normal distribution to model the group means. A bivariate normal model allows the matching correlation to be either positive or negative. There is an immediate translation between normally distributed group mean models and mixed models when the group sizes are balanced, and for unbalanced data. methods such as REML, which are known to give the same results as ANOVA when the data are balanced (Searle, Casella & McCulloch [36]), may be expected to depart only slowly from the balanced results as the degree of imbalance increases. For moderate sized groups, the central limit theorem allows group means to be modeled as normally distributed whether the individual level observations are binary responses, continuous, or counts. Thus, findings applicable to bivariate normal distributed group means may readily be extended to other models and methods.

The criteria by which different procedures are assessed are bias and power. A test is biased conservative if its type 1 error rate is less than its nominal size α. That is, a test with nominal size 0.05 is biased conservative if it rejects less than 5% of the time when the null is true, and anti-conservative if it rejects more than 5% of the time. Anti-conservative tests are generally inadmissible. The power of a test is its ability to detect a real treatment effect and is measured by the probability that it will reject when the null is not true.

Some basic results are well known. Referring to an extension of Diehr's schema

	Two sample t test	Paired t test
Matched experiment	MU procedure	MM procedure
Unmatched experiment	UU procedure	UM procedure

a) The paired t test is an unbiased test of the null hypothesis for bivariate normal data (Lehmann [29] 4.4 and 5.2). (An unbiased test is a test whose type 1 error rate is equal to its nominal size). That is, the MM procedure is unbiased;

b) The two-sample t test is an unbiased test of the null hypothesis for uncorrelated bivariate normal data (Lehmann [29] 5.3); the UU procedure is unbiased;

c) The UM procedure, that is, performing a paired t test on data from an unmatched experiment, is always less powerful than a UU procedure because of reduced degrees of freedom, and given that the latter is unbiased, the former may be discarded.

We use the distribution of the two-sample t statistic provided by Proschan to evaluate the properties of the MU procedure.

3. We show that the MU procedure is biased conservative when the matching correlation is positive. and anti-conservative when the matching correlation is negative.

Power curves for the MU procedure (figure 4.1. chapter 4) suggest that the power of the procedure increases as the matching correlation ρ increases when the power is greater than 50% and ρ is positive. If this is true, this means that an investigator who is trying maximize power, who designs the experiment for power greater than 50% and who can match effectively so as to induce a positive matching correlation. may discard the UU option. He need only consider an MU or an MM procedure; that is, between a matched experimental design and unmatched analysis, or a matched experimental design and matched analysis. Whichever analysis he chooses, his power will increase with effective matching.

4. This assertion - that the power of the MU procedure increases as the matching correlation ρ increases when the power is at least 50% and ρ is not negative - is stated as a theorem and a proof attempted.

As we have noted, a matching correlation ρ may under certain circumstances be negative. In general, however, most investigators would consider this possibility to be slight. If that is so, and the theorem is true, a matched experiment designed for power greater than 50% will, under either method of analysis, be unbiased or conservative, and more powerful than an unmatched experiment.

As between the two methods of analysis for the matched experiment, sometimes the two-sample test is more powerful, sometimes the paired test. The paired test is less powerful when the matching correlation ρ is small, because of loss of degrees of freedom; it is more powerful when the matching correlation ρ is larger, because of the reduction of variance. Figure 1.2 illustrates the power of these two options, for $n = 5$ pairs, nominal test size 0.05 and an effect size of 1.5 standard deviations.

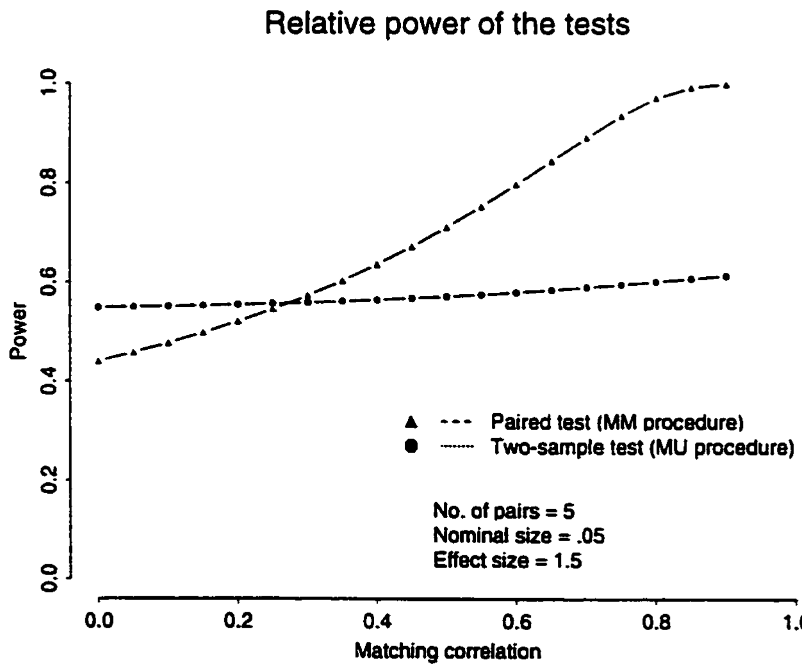


Figure 1.2: The power of the paired t test and the two-sample t test when the data come from a bivariate normal distribution.

The question then arises as to whether there are other options besides a straightforward matched analysis or unmatched analysis for the matched experiment; in terms of the bivariate normal group means model, whether there are other options besides the two-sample t test and the paired t test.

5. We seek to derive other tests for bivariate normal data.

1.4.1 Tests

The group-level statistics of a pair matched group randomized trial are to be modeled by a bivariate normal distribution

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} \sim N_2\left(\begin{matrix} \mu_1 \\ \mu_2 \end{matrix} , \sigma^2 \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right)$$

ρ represents the matching correlation. The hypothesis of interest is that $\mu_1 = \mu_2$.

The paired t test has a number of optimal properties for this problem. It is the uniformly most powerful unbiased test (Lehmann [29] 4.4 and 5.2). That is, it is the most powerful test of $\mu_1 = \mu_2$ against any alternative hypothesis, among all tests whose type 1 error rate equals their nominal size. It is the uniformly most powerful invariant test (Lehmann 6.6 and 6.14). We will also show in chapter 5 that it is equivalent to a conditional test based on the distribution of $\bar{y}_{\cdot 1} - \bar{y}_{\cdot 2}$ conditioned on $\mathbf{v} \equiv \sum_{i=1}^{n} \vec{y}_i \vec{y}_i'$ the sufficient statistic for the variance-covariance matrix under the null. Also, we will show in section 8.2 that it is equivalent to a conditional test based on the two sample t statistic conditioned on the maximum likelihood estimator of ρ under the assumption of equality of the bivariate normal means and variances.

Conditioning on sufficient statistics for the nuisance parameters in a model is a well recognized way of deriving tests. By the definition of sufficiency, the joint distribution of the data conditioned on the sufficient statistics for the nuisance parameters is independent of the nuisance parameters. Sufficient statistics may readily be found (if they exist) by factoring the probability density function; statistics which are inseparable from parameters in the factorization are sufficient for those parameters.

In our case, the parameters of interest are μ_1 and μ_2, and the nuisance parameters are σ^2 and ρ. In the paired test, the distribution of both the numerator of the paired t statistic

$\sqrt{n}(\bar{y}_{\cdot 1}-\bar{y}_{\cdot 2})$ and the denominator $\sqrt{\frac{1}{n-1}(s_1^2+s_2^2-2s_{12})}$ are scaled by the common factor $\sigma\sqrt{2(1-\rho)}$.

$$\frac{\sqrt{n}(\bar{Y}_{\cdot 1}-\bar{Y}_{\cdot 2})}{\sigma\sqrt{2(1-\rho)}} \sim N\left(\frac{\sqrt{n}(\mu_1-\mu_2)}{\sigma\sqrt{2(1-\rho)}},1\right)$$
$$\frac{S_1^2+S_2^2-2S_{12}}{2\sigma^2(1-\rho)} \sim \chi^2_{n-1}$$

(See chapter 2 for details of notation). Hence, the factor disappears in forming the test statistic,

$$\frac{\frac{\sqrt{n}(\bar{Y}_{\cdot 1}-\bar{Y}_{\cdot 2})}{\sigma\sqrt{2(1-\rho)}}}{\sqrt{\frac{1}{n-1}\frac{S_1^2+S_2^2-2S_{12}}{2\sigma^2(1-\rho)}}} = \frac{\sqrt{n}(\bar{Y}_{\cdot 1}-\bar{Y}_{\cdot 2})}{\sqrt{\frac{1}{n-1}\left(S_1^2+S_2^2-2S_{12}\right)}} = T_p$$

and its distribution under the null $\mu_1=\mu_2$ is independent of both σ^2 and ρ.

However, the two-sample t statistic does not possess the same nice property of having a distribution independent of the nuisance parameters. The distribution of its numerator is still scaled by $\sigma\sqrt{2(1-\rho)}$, and the denominator $\sqrt{\frac{1}{n-1}(s_1^2+s_2^2)}$ is scaled by σ, since

$$\frac{S_1^2}{\sigma^2} \sim \chi^2_{n-1} \quad \text{and} \quad \frac{S_2^2}{\sigma^2} \sim \chi^2_{n-1}\,.$$

but $\frac{S_1^2+S_2^2}{\sigma^2(1-\rho)}$ does not have a distribution that is independent of ρ. Hence the distribution of the quotient

$$\frac{\sqrt{n}(\bar{Y}_{\cdot 1}-\bar{Y}_{\cdot 2})}{\sqrt{\frac{1}{n-1}\left(S_1^2+S_2^2\right)}} = T_2$$

is independent of σ but not of ρ. ρ remains as a nuisance parameter.

There are no sufficient statistics in the bivariate normal distribution for the parameter ρ alone. However, for testing $\mu_1=\mu_2$ one may without loss of generality set $\mu_1=\mu_2=0$ for the null. Then the density function becomes, for n pairs

$$f_Y(y) = \left(\frac{1}{2\pi\sigma^2\sqrt{1-\rho^2}}\right)^n \exp\left\{-\frac{1}{2\sigma^2(1-\rho^2)}\left(\sum_{i=1}^n y_{i1}^2 - 2\rho\sum_{i=1}^n y_{i1}\,y_{i2} + \sum_{i=1}^n y_{i2}^2\right)\right\}$$

and the statistics $\sum_{i=1}^n y_{i1}^2$, $\sum_{i=1}^n y_{i1}\,y_{i2}$, $\sum_{i=1}^n y_{i2}^2$ are together sufficient for ρ and σ^2 the parameters of the variance-covariance matrix.

The above gives the rationale for a test based on $\bar{y}_{i1} - \bar{y}_{i2}$ conditioned on the sufficient statistics for the covariance matrix. Under the null, the distribution of $\bar{y}_{i1} - \bar{y}_{i2}$ depends only on ρ and σ^2; conditioning on the sufficient statistics for ρ and σ^2 gives a distribution independent of nuisance parameters. However, the resulting test will be found to be equivalent to a paired t test.

When $\rho = 0$, the two sample t test is optimal. It is the uniformly most powerful unbiased test and the uniformly most powerful invariant test of $\mu_1 = \mu_2$ when $\rho = 0$. (Lehmann 5.3).

Given the above, what can one hope for from other tests? Can one retain the benefit of the power of the two sample test when ρ is small, yet utilize some of the power of the paired test when ρ is large? Are there tests which can avoid the anti-conservative bias of the two-sample test if ρ is less than zero?

Given the optimality properties of the paired and two sample tests, clearly such gains cannot be achieved without some costs elsewhere. What properties of the optimal tests might be foregone to acquire such advantages? It may be plausibly argued that the high power of the paired test for large ρ may be foregone, since most investigators would not rely on a high matching correlation to achieve their required power. The conservative bias of the two sample test for $\rho > 0$ may clearly be dispensed with. Few researchers would argue that it was essential that a test for a defensively matched group randomized trial must be unbiased when ρ was highly negative; some might consider some anti-conservativeness acceptable for high negative ρ.

In general, tests are sought which use available information about the effectiveness of the matching to enhance power in parameter regions considered likely to be found in group randomized trials, that is, when ρ is small and positive.

1.4.2 Arrangement of Chapters

In chapter 2 we set out the notation to be used.

In chapter 3 we show that the MU procedure is biased conservative when the matching correlation ρ is positive, and anti-conservative when the matching correlation is negative, using distribution functions derived in Appendix A from the distribution derived by Proschan.

In chapter 4 we attempt to prove the theorem about the power of the MU procedure, using distribution, density and derivative functions derived in Appendix A.

Chapters 5 through 11 examine various alternative tests. Chapter 5 derives the conditional test based on $\bar{y}_{i1} - \bar{y}_{i2}$ conditioned on the sufficient statistics for the covariance matrix. Chapter 6 considers tests based on the distribution of $T_2|\rho$. Chapter 7 considers tests based on the distribution of $T_2|r_1$. Chapter 8 evaluates tests conditioning on r_2 and r_3, other estimators of ρ. In in 1997, Michael Kenward and James Roger propounded a method for estimating parameters of Gaussian linear models with small sample sizes; chapter 9 examines their method as applied to our problem. Chapter 10 develops and evaluates conditional tests based on flexible critical values. A procedure similar to the two stage procedure considered by Diehr is evaluated at chapter 11.

A discussion and summary is at chapter 12.

Chapter 2

NOTATION

Generally, the notation given below is used throughout this dissertation. However, in Appendix B, the statistics s_1, s_2 and s_{12} are redefined so as to be consistent with the authors cited there. In addition, in some places where there is no risk of confusion, r is used to represent the Pearson product-moment sample correlation coefficient; where it is necessary to distinguish Pearson's r from other estimates of correlation, r_1 is used.

2.1 *Observations*

$\begin{pmatrix} y_{11} \\ y_{12} \end{pmatrix} \cdots \begin{pmatrix} y_{n1} \\ y_{n2} \end{pmatrix}$ represents n pairs of observations which are modeled as realizations of random variables having the bivariate normal distribution $\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} \sim N_2\left(\begin{matrix} \mu_1 \\ \mu_2 \end{matrix}, \sigma^2 \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right)$.

2.2 *Statistics*

$s_1^2 \equiv \sum_{i=1}^{n} (y_{i1} - \bar{y}_{\cdot 1})^2$;

$s_2^2 \equiv \sum_{i=1}^{n} (y_{i2} - \bar{y}_{\cdot 2})^2$;

$s_{12} \equiv \sum_{i=1}^{n} (y_{i1} - \bar{y}_{\cdot 1})(y_{i2} - \bar{y}_{\cdot 2})$;

(In Appendices B and C, s_1, s_2 and s_{12} have the meanings given by the authors whose work is drawn on there, which differ from the above).

$t_2 \equiv \dfrac{\sqrt{n}\,(\bar{y}_{\cdot 1} - \bar{y}_{\cdot 2})}{\sqrt{\frac{1}{n-1}\,(s_1^2 + s_2^2)}}$, the two sample t-statistic;

$t_p \equiv \dfrac{\sqrt{n}\,(\bar{y}_{\cdot 1} - \bar{y}_{\cdot 2})}{\sqrt{\frac{1}{n-1}\,(s_1^2 + s_2^2 - 2s_{12})}}$, the paired t-statistic;

$r_1 \equiv \frac{s_{12}}{\sqrt{s_1^2 s_2^2}}$, the Pearson product-moment sample correlation coefficient; r_1 is the maximum likelihood estimator (MLE) of ρ when no assumption is made about the equality of means or variances; in places where there is no chance of confusion, r_1 is simply denoted by r;

$r_2 \equiv \frac{s_{12}}{\frac{1}{2}\left(s_1^2+s_2^2\right)}$; r_2 is the MLE of ρ when the variances of Y_1 and Y_2 are assumed equal;

$r_3 = \frac{\sum_{i=1}^{n}(y_{i1}-\bar{y}_{\cdot\cdot})(y_{i2}-\bar{y}_{\cdot\cdot})}{\frac{1}{2}\left(\sum_{i=1}^{n}(y_{i1}-\bar{y}_{\cdot\cdot})^2+\sum_{i=1}^{n}(y_{i2}-\bar{y}_{\cdot\cdot})^2\right)}$; r_3 is the MLE of ρ when the means and variances are assumed equal, i.e. when $\mu_1=\mu_2$.

2.3 *Tests*

α represents the size of a test, that is, the probability that the test will reject when the null hypothesis is true. In general, we use *size* to mean the nominal size used to determine the critical value of the test, and *type 1 error rate* to denote the frequency with which a test rejects under the null as assessed by simulation or other evaluations of the test;

c the critical value of a test;

$\delta \equiv \left(\frac{\mu_2-\mu_1}{\sigma}\right)\sqrt{\frac{n}{2}}$;

$\delta_\rho \equiv \frac{\delta}{\sqrt{1-\rho}}$;

β the power of a test, that is, the probability that the test will reject when a specified alternative hypothesis is true;

β_{MU} the power of a two sample t test when the data come from a bivariate normal distribution.

2.4 *Functions and Symbols*

$\Gamma(\cdot)$ the Gamma function, (see *Handbook of Mathematical Functions*, Abramowitz and Stegun [1]);

$H(\alpha,\beta,\gamma;x) \equiv \frac{\Gamma(\gamma)}{\Gamma(\alpha)\,\Gamma(\beta)}\sum_{i=0}^{\infty}\frac{\Gamma(\alpha+i)\,\Gamma(\beta+i)}{\Gamma(\gamma+i)}\frac{x^i}{i!}$ the (Gauss) Hypergeometric Function; Erdélyi [12] gives an excellent treatment;

$M(\alpha,\gamma;x) \equiv \frac{\Gamma(\gamma)}{\Gamma(\alpha)} \sum_{i=0}^{\infty} \frac{\Gamma(\alpha+i)}{\Gamma(\gamma+i)} \frac{x^i}{i!}$ the Confluent Hypergeometric Function;

$\Psi_1(\alpha_1,\alpha_2,\gamma_1,\gamma_2;x,y) \equiv \frac{\Gamma(\gamma_1)\,\Gamma(\gamma_2)}{\Gamma(\alpha_1)\,\Gamma(\alpha_2)} \sum_{i,j=0}^{\infty} \frac{\Gamma(\alpha_1+i+j)\,\Gamma(\alpha_2+i)}{\Gamma(\gamma_1+i)\,\Gamma(\gamma_2+j)} \frac{x^i}{i!} \frac{y^j}{j!}$, the Horn function given by Erdélyi [12], p. 225, equation (23).

$(\alpha)_i$ Pochhammer's symbol, see below.

2.5 *Pochhammer's Symbol*

Pochhammer's symbol $(\alpha)_i$ is most easily understood as

$$(\alpha)_i = \frac{\Gamma(\alpha+i)}{\Gamma(\alpha)} \quad \text{or} \quad \alpha(\alpha+1)\cdots(\alpha+i-1)$$

whichever works. It is most commonly encountered in situations where $\alpha > 0$ and $i = 0, 1, 2, 3, \ldots$, when the expressions are equivalent. It simplifies the writing of hypergeometric functions, for example, the (Gauss) Hypergeometric Function may be written

$$H(\alpha,\beta,\gamma;x) \equiv \sum_{i=0}^{\infty} \frac{(\alpha)_i\,(\beta)_i}{(\gamma)_i} \frac{x^i}{i!}$$

Exton [13] and Srivastava [38] both give definitions, though they are not the same. We use the definition, for i an integer

$$(\alpha)_i = \begin{cases} 1 & \text{if } i=0 \\ \alpha(\alpha+1)\cdots(\alpha+i-1) & \text{if } i=1,2,\ldots \\ \dfrac{(-1)^{-i}}{(1-\alpha)_{-i}} & \text{if } i=-1,-2,\ldots \end{cases} \tag{2.1}$$

This is Exton's definition with the additional clarification that $(0)_0, (-1)_0, (-2)_0, \ldots = 1$, or Srivastava's definition extended to negative i.

2.5.1 i not an integer

One sometimes comes across situations such as $\sum_{j=0}^{\infty} \frac{\Gamma\left(\alpha+\frac{j}{2}\right)}{\Gamma(\alpha)}$ which it would be convenient to write as $\sum_{j}(\alpha)_{\frac{j}{2}}$, and there is no problem with

$$(\alpha)_i = \begin{cases} 1 & \text{if } i=0 \\ \dfrac{\Gamma(\alpha+i)}{\Gamma(\alpha)} & \text{if neither } \alpha \text{ nor } \alpha+i \text{ is a non-positive integer.} \end{cases}$$

Chapter 3

BIAS OF THE MU PROCEDURE

An MU procedure is a test of the equality of the means of observations from a correlated bivariate normal distribution using a two sample t test. The two sample test assumes the samples are independent; thus using this test when the samples are correlated amounts to ignoring the matching in the analysis.

Since the two sample t-statistic is

$$t_2 \equiv \frac{\sqrt{n}\,(\bar{y}_{\cdot 1}-\bar{y}_{\cdot 2})}{\sqrt{\frac{1}{n-1}\left(s_1^2+s_2^2\right)}}\,,$$

but the unbiased estimate of the variance of $\bar{y}_{\cdot 1}-\bar{y}_{\cdot 2}$ is

$$\tfrac{1}{n-1}\left(s_1^2+s_2^2-2s_{12}\right)\,,$$

and since s_{12} will tend to increase with the matching correlation ρ, the denominator of the two-sample t statistic may be expected to to overestimate the variance of $\bar{y}_{\cdot 1}-\bar{y}_{\cdot 2}$ when ρ is positive, and result in a conservative test. When ρ is negative an anti-conservative test is to be expected.

To more precisely evaluate the bias of the procedure, the distribution under the null of the two sample statistic when the samples come from a correlated bivariate normal distribution is derived in computable form.

3.1 *The Distribution of $T_2|\rho$ under the null in computable form*

Proschan [34] showed (at equation (9)) that

$$F_{T_2}(t) = \int_{w=0}^{\infty}\left[G_{2n-2}\left(t;\delta_\rho,\frac{2\rho w}{1-\rho}\right)\right] f_{n-1}(w)\,dw$$

where

i) $T_2 \equiv \dfrac{\sqrt{n}\,(\bar{Y}_{\cdot 1}-\bar{Y}_{\cdot 2})}{\sqrt{\frac{1}{n-1}\,(S_1^2+S_2^2)}}$, the two-sample t-statistic;

ii) $\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} \sim N_2\left(\begin{matrix} \mu_1 \\ \mu_2 \end{matrix}, \sigma^2 \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right), \quad \rho > 0;$

iii) $G_\nu(\cdot\ ;\delta,\lambda)$ is the doubly non-central t distribution function with ν degrees of freedom and non-centrality parameters δ and λ;

iv) $\delta_\rho \equiv \sqrt{\dfrac{n}{2(1-\rho)}}\left(\dfrac{\mu_2-\mu_1}{\sigma}\right)$; and

v) $f_\nu(\cdot)$ is the $\mathcal{X}^2$ density function with ν degrees of freedom.

For the doubly non-central t distribution function, Krishnan [28] (at equation (11)) gives

$$\begin{aligned} G_\nu(t;\delta,\lambda) \;=\; & G(0;\delta) + \sum_{i,j=0}^{\infty} \frac{E(\lambda,i)\,E(\delta^2,j)\,a^{j+1/2}}{2\,\Gamma(\nu/2+i)\,\Gamma(j+3/2)} \\ & \times \Bigg[\Gamma(i+j+\nu/2+1/2)\,H(j+1/2,1-i-\nu/2,j+3/2;a) \\ & \quad + \frac{\delta}{j+1}\sqrt{\frac{a}{2}}\,\Gamma(i+j+\nu/2+1)\,H(j+1,1-i-\nu/2,j+2;a)\Bigg] \end{aligned}$$

where

i) $E(\alpha,k) \equiv \dfrac{e^{-\alpha/2}\left(\frac{\alpha}{2}\right)^k}{\Gamma(k+1)}$

ii) H is the Gauss hypergeometric function; see 2.4;

iii) $a \equiv \dfrac{t^2}{2n-2+t^2}$.

The $\mathcal{X}^2$ density function is well known

$$f_\nu(u) \;=\; \frac{u^{\nu/2-1}\exp\left(-\frac{u}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)2^{\nu/2}}\,.$$

Hence, expanding $G_{2n-2}(0;\delta_\rho) = \frac{1}{2} - \frac{\delta_\rho}{\sqrt{2\pi}}M\left(\frac{1}{2},\frac{3}{2};-\delta_\rho^2/2\right)$,

$$\begin{aligned} F_{T_2}(t) \;=\; & \int_{w=0}^{\infty}\Bigg[\frac{1}{2}-\frac{\delta_\rho}{\sqrt{2\pi}}M\left(\tfrac{1}{2},\tfrac{3}{2};-\delta_\rho^2/2\right)+\sum_{i,j=0}^{\infty}E\left(\frac{2\rho w}{1-\rho},i\right)\frac{E(\delta_\rho^2,j)\;a^{j+1/2}}{2\,\Gamma(n+i-1)\;\Gamma(j+\frac{3}{2})} \\ & \times\Bigg(\Gamma(n+i+j-\tfrac{1}{2})\;H(j+\tfrac{1}{2},2-n-i,j+\tfrac{3}{2};a) \\ & + \frac{\delta_\rho}{j+1}\sqrt{\frac{a}{2}}\,\Gamma(n+i+j)\;H(j+1,2-n-i,j+2;a)\Bigg)\Bigg]\frac{w^{(n-3)/2}\exp\left(-\frac{w}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)\;2^{(n-1)/2}}\,dw \end{aligned}$$

where M is the confluent hypergeometric function; see 2.4.

Setting $\delta=0$ gives

$$F_{T_2}(t) = \int_{w=0}^{\infty}\left[\frac{1}{2}+\sum_{i=0}^{\infty}E\left(\frac{2\rho w}{1-\rho},i\right)\frac{a^{\frac{1}{2}}\,\Gamma(n+i-\frac{1}{2})}{\sqrt{\pi}\,\Gamma(n+i-1)}H(\tfrac{1}{2},2-n-i,\tfrac{3}{2};a)\right]\frac{w^{\frac{n-3}{2}}\exp\left(-\frac{w}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)\,2^{\frac{n-1}{2}}}\,dw$$

Integrating term by term

$$F_{T_2}(t) = \frac{1}{2}+\sum_{i=0}^{\infty}\frac{a^{\frac{1}{2}}\,\Gamma(n+i-\frac{1}{2})\,H(\frac{1}{2},2-n-i,\frac{3}{2};a)}{\sqrt{\pi}\,2^{\frac{n-1}{2}}\,\Gamma\left(\frac{n-1}{2}\right)\Gamma(n+i-1)}\int_{w=0}^{\infty}E\left(\frac{2\rho w}{1-\rho},i\right)w^{\frac{n-3}{2}}\exp\left(-\frac{w}{2}\right)\,dw$$

Evaluating the integral

$$\begin{aligned}
&\int_{w=0}^{\infty}E\left(\frac{2\rho w}{1-\rho},i\right)w^{\frac{n-3}{2}}\exp\left(-\frac{w}{2}\right)\,dw\\
&= \int_{w=0}^{\infty}\frac{e^{-\frac{\rho w}{1-\rho}}\left(\frac{\rho w}{1-\rho}\right)^{i}}{\Gamma(i+1)}w^{\frac{n-3}{2}}\exp\left(-\frac{w}{2}\right)\,dw\\
&= \frac{1}{\Gamma(i+1)}\times\int_{w=0}^{\infty}e^{-\frac{w(1+\rho)}{2(1-\rho)}}\left(\frac{\rho}{1-\rho}\right)^{i}w^{\frac{n+2i-3}{2}}\left(\frac{1+\rho}{2(1-\rho)}\right)^{\frac{n+2i-3}{2}}\left(\frac{2(1-\rho)}{1+\rho}\right)^{\frac{n+2i-3}{2}}\,dw\\
&= \frac{\left(\frac{\rho}{1-\rho}\right)^{i}\left(\frac{2(1-\rho)}{1+\rho}\right)^{\frac{n+2i-3}{2}}}{\Gamma(i+1)}\int_{w=0}^{\infty}e^{-\frac{w(1+\rho)}{2(1-\rho)}}\left(\frac{w(1+\rho)}{2(1-\rho)}\right)^{\frac{n+2i-3}{2}}\,dw
\end{aligned}$$

put $u=\frac{w(1+\rho)}{2(1-\rho)}$, so $dw=\frac{2(1-\rho)}{1+\rho}du$,

$$\begin{aligned}
&= \frac{\left(\frac{\rho}{1-\rho}\right)^{i}\left(\frac{2(1-\rho)}{1+\rho}\right)^{\frac{n+2i-1}{2}}}{\Gamma(i+1)}\int_{u=0}^{\infty}e^{-u}u^{\frac{n+2i-3}{2}}\,du\\
&= \frac{\left(\frac{\rho}{1-\rho}\right)^{i}\left(\frac{2(1-\rho)}{1+\rho}\right)^{\frac{n+2i-1}{2}}}{\Gamma(i+1)}\Gamma\left(\frac{n+2i-1}{2}\right).
\end{aligned}$$

Then

$$\begin{aligned}
F_{T_2}(t) &= \frac{1}{2}+\sum_{i=0}^{\infty}\frac{a^{\frac{1}{2}}\,\Gamma(n+i-\frac{1}{2})\,H(\frac{1}{2},2-n-i,\frac{3}{2};a)}{\sqrt{\pi}\,2^{\frac{n-1}{2}}\,\Gamma\left(\frac{n-1}{2}\right)\Gamma(n+i-1)}\frac{\left(\frac{\rho}{1-\rho}\right)^{i}\left(\frac{2(1-\rho)}{1+\rho}\right)^{\frac{n+2i-1}{2}}}{\Gamma(i+1)}\Gamma\left(\frac{n+2i-1}{2}\right)\\
&= \frac{1}{2}+\frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{\sqrt{\pi}\,\Gamma\left(\frac{n-1}{2}\right)}\sum_{i=0}^{\infty}\frac{\Gamma\left(\frac{n-1}{2}+i\right)\Gamma(n+i-\frac{1}{2})}{\Gamma(n+i-1)\,\Gamma(i+1)}a^{\frac{1}{2}}\left(\frac{2\rho}{1+\rho}\right)^{i}H(\tfrac{1}{2},2-n-i,\tfrac{3}{2};a)
\end{aligned}$$

Putting $a \equiv \frac{t^2}{2n-2+t^2}$ and using $H(\alpha,\beta,\gamma;x) \equiv (1-x)^{-\alpha}H(\alpha,\gamma-\beta,\gamma;\frac{x}{x-1})$ (Erdélyi [12], p.109) gives

$$F_{T_2}(t)$$
$$= \frac{1}{2} + \frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{\sqrt{\pi}\,\Gamma(\frac{n-1}{2})}\sum_{i=0}^{\infty}\frac{\Gamma(\frac{n-1}{2}+i)\,\Gamma(n+i-\frac{1}{2})}{\Gamma(n+i-1)\,\Gamma(i+1)}\left(\frac{t}{\sqrt{2n-2}}\right)\left(\frac{2\rho}{1+\rho}\right)^{i} H\left(\tfrac{1}{2},n+i-\tfrac{1}{2},\tfrac{3}{2};\tfrac{-t^2}{2n-2}\right)$$

Expanding the Gauss hypergeometric function (see 2.4) gives

$$F_{T_2}(t)$$
$$= \frac{1}{2} + \frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{\sqrt{\pi}\,\Gamma(\frac{n-1}{2})}\sum_{i,j=0}^{\infty}\frac{\Gamma(\frac{n-1}{2}+i)\,\Gamma(n+i-\frac{1}{2})}{\Gamma(n+i-1)\,\Gamma(i+1)}\left(\frac{t}{\sqrt{2n-2}}\right)\left(\frac{2\rho}{1+\rho}\right)^{i}$$
$$\times \frac{\Gamma(j+\frac{1}{2})\,\Gamma(n+i-\frac{1}{2}+j)\Gamma(\frac{3}{2})}{\Gamma(\frac{1}{2})\,\Gamma(n+i-\frac{1}{2})\Gamma(j+\frac{3}{2})\,\Gamma(j+1)}\left(\frac{-t^2}{2n-2}\right)^{j}$$
$$= \frac{1}{2} + \frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{2\sqrt{\pi}\,\Gamma(\frac{n-1}{2})}\sum_{i,j=0}^{\infty}\frac{\Gamma(\frac{n-1}{2}+i)\,\Gamma(n+i-\frac{1}{2}+j)\,\Gamma(j+\frac{1}{2})}{\Gamma(n+i-1)\,\Gamma(j+\frac{3}{2})\,\Gamma(i+1)\,\Gamma(j+1)}\left(\frac{t}{\sqrt{2n-2}}\right)\left(\frac{2\rho}{1+\rho}\right)^{i}\left(\frac{-t^2}{2n-2}\right)^{j}$$

This is a hypergeometric function in two variables. the F_2 function by Appell's definition (see [3] or [12]). In this form, however, it will not converge for $\rho \le -\frac{1}{3}$ or $t^2 \ge 2n-2$; the infinite sums are therefore expressed as hypergeometric functions in single variables and known transformations applied to yield a convergent expression.

Express the i series as a Gauss hypergeometric function (see 2.4)

$$\sum_{i=0}^{\infty}\frac{\Gamma(\frac{n-1}{2}+i)\,\Gamma(n+i-\frac{1}{2}+j)}{\Gamma(n+i-1)\,\Gamma(i+1)}\left(\frac{2\rho}{1+\rho}\right)^{i} = \frac{\Gamma(\frac{n-1}{2})\,\Gamma(n-\frac{1}{2}+j)}{\Gamma(n-1)} H\left(\tfrac{n-1}{2},n-\tfrac{1}{2}+j,n-1;\tfrac{2\rho}{1+\rho}\right)$$

Apply Erdélyi's [12] transformation(4) p.111; express the resulting Gauss hypergeometric function as a power series

$$= \frac{\Gamma(\frac{n-1}{2})\,\Gamma(n-\frac{1}{2}+j)}{\Gamma(n-1)}\left(1-\frac{\rho}{1+\rho}\right)^{-(n-\frac{1}{2}+j)} H\left(\frac{n-\frac{1}{2}+j}{2},\frac{n-\frac{1}{2}+j}{2}+\frac{1}{2},\frac{n}{2};\rho^2\right)$$
$$= \frac{\Gamma(\frac{n-1}{2})\,\Gamma(n-\frac{1}{2}+j)}{\Gamma(n-1)}(1+\rho)^{n-\frac{1}{2}+j}\frac{\Gamma(\frac{n}{2})}{\Gamma\left(\frac{n-\frac{1}{2}+j}{2}\right)\Gamma\left(\frac{n-\frac{1}{2}+j}{2}+\frac{1}{2}\right)}$$
$$\times \sum_{i=0}^{\infty}\frac{\Gamma\left(\frac{n-\frac{1}{2}+j}{2}+i\right)\Gamma\left(\frac{n-\frac{1}{2}+j}{2}+\frac{1}{2}+i\right)}{\Gamma(\frac{n}{2}+i)\,\Gamma(i+1)}\left(\rho^2\right)^{i}.$$

Since $\Gamma(x)\,\Gamma(x+\frac{1}{2}) = \dfrac{\Gamma(2x)\,\Gamma(\frac{1}{2})}{2^{2x-1}}$

$$\begin{aligned} &= \frac{\Gamma(n-1)\,\Gamma(\frac{1}{2})\,\Gamma(n-\frac{1}{2}+j)}{2^{n-1-1}\Gamma(n-1)}\,(1+\rho)^{n-\frac{1}{2}+j}\,\frac{2^{\,n-\frac{3}{2}+j}}{\sqrt{\pi}\,\Gamma(n-\frac{1}{2}+j)} \\ &\quad\times\sum_{i=0}^{\infty}\frac{\sqrt{\pi}\,\Gamma(n-\frac{1}{2}+j+2i)}{2^{\,n-\frac{3}{2}+j+2i}\,\Gamma(\frac{n}{2}+i)\,\Gamma(i+1)}\left(\rho^2\right)^i \\ &= 2\sqrt{\pi}\left(\tfrac{1+\rho}{2}\right)^{n-1}(1+\rho)^{\frac{1}{2}+j}\sum_{i=0}^{\infty}\frac{\Gamma(n-\frac{1}{2}+j+2i)}{\Gamma(\frac{n}{2}+i)\,\Gamma(i+1)}\left(\tfrac{\rho^2}{4}\right)^i . \end{aligned}$$

Thus

$$F_{T_2}(t) = \frac{1}{2}+\frac{\left(\frac{1-\rho^2}{4}\right)^{\frac{n-1}{2}}}{\Gamma(\frac{n-1}{2})}\sum_{i,j=0}^{\infty}\frac{\Gamma(n-\frac{1}{2}+j+2i)\;\Gamma(j+\frac{1}{2})}{\Gamma(\frac{n}{2}+i)\,\Gamma(j+\frac{3}{2})\,\Gamma(i+1)\,\Gamma(j+1)}\left(\tfrac{\rho^2}{4}\right)^i\left(\tfrac{-t^2(1+\rho)}{2n-2}\right)^j\left(\tfrac{t\sqrt{1+\rho}}{\sqrt{2n-2}}\right).$$

Transform the j series

$$\begin{aligned} &\sum_{j=0}^{\infty}\frac{\Gamma(j+\frac{1}{2})\,\Gamma(n-\frac{1}{2}+j+2i)}{\Gamma(j+\frac{3}{2})\,\Gamma(j+1)}\left(\tfrac{-t^2(1+\rho)}{2n-2}\right)^j \\ &= \frac{\Gamma(\frac{1}{2})\;\Gamma(n-\frac{1}{2}+2i)}{\Gamma(\frac{3}{2})}\,H\!\left(\tfrac{1}{2},n-\tfrac{1}{2}+2i,\tfrac{3}{2};\tfrac{-t^2(1+\rho)}{2(n-1)}\right) \\ &= \frac{\Gamma(\frac{1}{2})\;\Gamma(n-\frac{1}{2}+2i)}{\Gamma(\frac{3}{2})}\left(1-\tfrac{-t^2(1+\rho)}{2(n-1)}\right)^{-(n-\frac{1}{2}+2i)}H\!\left(1,n-\tfrac{1}{2}+2i,\tfrac{3}{2};\frac{\frac{-t^2(1+\rho)}{2(n-1)}}{\frac{-t^2(1+\rho)}{2(n-1)}-1}\right) \end{aligned}$$

(see Erdélyi [12], transformation(4) p.105):

$$\begin{aligned} &= \frac{\Gamma(\frac{1}{2})\;\Gamma(n-\frac{1}{2}+2i)}{\Gamma(\frac{3}{2})}\left(\tfrac{2(n-1)}{2(n-1)+t^2(1+\rho)}\right)^{n-\frac{1}{2}+2i}H\!\left(1,n-\tfrac{1}{2}+2i,\tfrac{3}{2};\tfrac{t^2(1+\rho)}{t^2(1+\rho)+2(n-1)}\right) \\ &= \Gamma(\tfrac{1}{2})\left(\tfrac{2(n-1)}{2(n-1)+t^2(1+\rho)}\right)^{n-\frac{1}{2}+2i}\sum_{j=0}^{\infty}\frac{\Gamma(n-\frac{1}{2}+2i+j)}{\Gamma(\frac{3}{2}+j)}\left(\tfrac{t^2(1+\rho)}{t^2(1+\rho)+2(n-1)}\right)^j . \end{aligned}$$

Thus

$$\begin{aligned} F_{T_2}(t) &= \frac{1}{2}+\frac{\sqrt{\pi}}{\Gamma(\frac{n-1}{2})}\left(\tfrac{(n-1)\sqrt{1-\rho^2}}{2(n-1)+t^2(1+\rho)}\right)^{n-1}\times \\ &\sum_{i,j=0}^{\infty}\frac{\Gamma(n-\frac{1}{2}+2i+j)}{\Gamma(\frac{n}{2}+i)\,\Gamma(\frac{3}{2}+j)\,\Gamma(i+1)}\left(\tfrac{\rho(n-1)}{2(n-1)+t^2(1+\rho)}\right)^{2i}\left(\tfrac{t\sqrt{1+\rho}}{\sqrt{2(n-1)+t^2(1+\rho)}}\right)\left(\tfrac{t^2(1+\rho)}{2(n-1)+t^2(1+\rho)}\right)^j . \end{aligned}$$

This function will always converge and hence is suitable for computing purposes.

Table 3.1: Bias of the MU procedure for various ρ and n, nominal size .05.

	Type 1 error rate					
ρ	-0.5	-0.25	0	0.25	0.5	0.75
$n=5$	0.10350	0.07461	0.05	0.02968	0.01402	0.00375
$n=8$	0.10645	0.07676	0.05	0.02735	0.01058	0.00164
$n=\infty$	0.10953	0.07959	0.05	0.02363	0.00557	0.00009

3.2 Evaluation of the bias

Using the above form of the distribution function, the type 1 error rate of the MU procedure under the null was computed for various values of ρ and n. The results are shown in table 3.1 for a nominal test size .05.

The procedure is biased conservative when the matching correlation is positive, and anti-conservative when the correlation is negative.

The results for $\rho \geq 0$ are consistent with Diehr's findings by simulation.

Chapter 4

A THEOREM CONCERNING THE POWER OF THE MU PROCEDURE

4.1 Illustration of the Theorem

The theorem, which is sometimes referred to as Theorem 1, asserts that if an MU procedure is performed, that is, a two-sample t test is conducted on observations from a matched experiment, which are modeled as data from a bivariate normal distribution with matching correlation ρ, then the power of the test increases with ρ when the power is at least 50% and $\rho \geq 0$. This means, among other things, that under the conditions stated, a) an MU procedure is more powerful than a UU procedure. and b) the power of the MU procedure is increased by effective matching.

Figure 4.1 illustrates the theorem. It shows five power curves for the UU and MU procedures for different values of ρ, $n = 7$ pairs and test size .05. It can be seen that the theorem is correct for these values.

4.2 Statement of the Theorem

Given

i) $\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} \sim N_2\left(\begin{matrix} \mu_1 \\ \mu_2 \end{matrix}, \sigma^2 \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right)$;

ii) $T_2(Y) \equiv \dfrac{\sqrt{n}\,(\bar{Y}_{2\cdot} - \bar{Y}_{1\cdot})}{\sqrt{\frac{1}{n-1}\,(S_1^2 + S_2^2)}}$, the two-sample t-statistic;

iii) $\beta \equiv P(T_2(Y) > c \mid \delta)$, the power of a two-sample t test for some critical value $c > 0$, effect size δ, and $n = 2, 3, 4, \ldots$,

then

$$\frac{d\beta}{d\rho} > 0$$

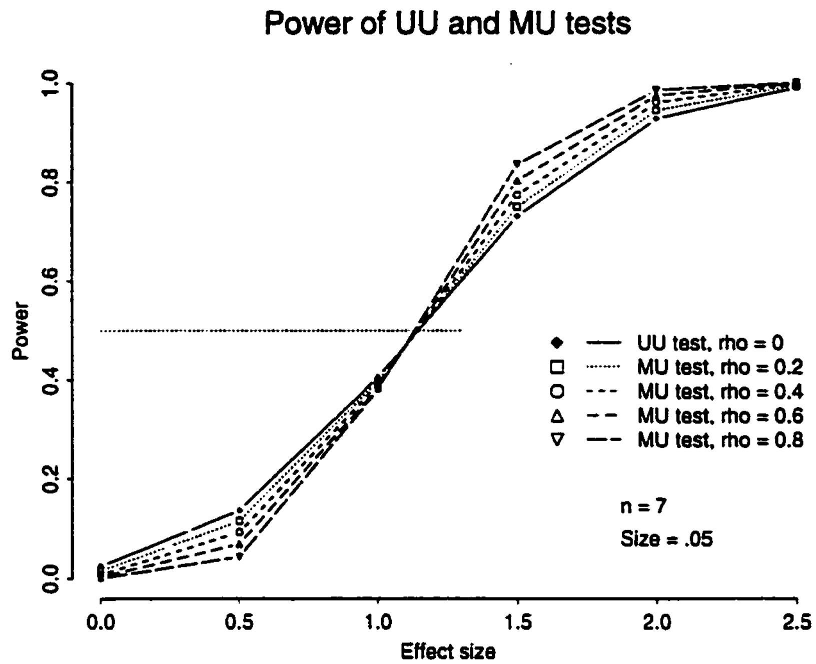


Figure 4.1: Power curves for UU and MU procedures, for various ρ. $n = 7$ pairs, nominal test size .05.

whenever

$$\beta \geq \tfrac{1}{2} \quad \text{and} \quad \rho \geq 0 \, .$$

4.3 General Comments

Note that the theorem is independent of the size of the test α. The condition $\beta \geq \frac{1}{2}$ is determined by the probability distribution under the alternative, which is a function of the critical value c and the effect size δ (and n and ρ). The critical value is determined under the null, by reference to α. Thus the theorem is 'about' c and δ. For any c, there will always be some δ large enough to make the power greater than $\frac{1}{2}$ (for given n and ρ). The theorem asserts that for values of c and δ which make $\beta \geq \frac{1}{2}$, the conclusion holds.

In general terms, the proof was approached by deriving the distribution function for the two sample t statistic under the alternative when the data come from a correlated bivariate normal distribution, and hence obtaining the power function for the MU procedure; and by deriving the derivative of the power function with respect to ρ. The derivations are given in Appendix A. There are a number of different ways of expressing the power function and the derivative. For the power function, any of the expressions in section A.1 may be used, each of which suggests possible approaches to proving the theorem. The expressions at (A.3) and (A.7)/(A.8) are most frequently used.

The theorem also appears to be true mathematically in another case, namely when $c \geq 0$ and $\beta > \frac{1}{2}$, though a critical value of 0 makes less sense statistically. In fact, it can be seen from section 4.6 that $c=0$ and $\beta=\frac{1}{2}$ implies $\frac{d\beta}{d\rho}=0$; in order for $\frac{d\beta}{d\rho}>0$ requires either $c>0$ or $\beta>\frac{1}{2}$.

A general proof was not accomplished. Proofs were achieved for a number of special cases and boundary values, for large n, for $c=0$, and when δ is small and $\beta=\frac{1}{2}$ (sections 4.4 through 4.6); these were accomplished by algebraic methods. The method of induction would appear to offer promise, particularly since the theorem is true asymptotically; an approach by induction is examined (section 4.7). At $\rho=0$. the power and derivative functions are closely similar: several approaches when $\rho=0$ are presented (sections 4.8 through 4.13). A demonstration for random parameter values is given at section 4.14.

A common difficulty was encountered in a number of approaches. A natural way to approach the problem

$$\begin{array}{lrcl} \text{Given :} & \beta-\frac{1}{2} & \geq & 0 \\ \text{show that :} & \dfrac{d\beta}{d\rho} & > & 0 \end{array}$$

is to show that for some positive constant or positive bounded function k

$$\frac{d\beta}{d\rho} > k\left(\beta-\tfrac{1}{2}\right) \tag{4.1}$$

at least when $\beta-\frac{1}{2}\geq 0$. However, figure 4.2, which illustrates the two functions, suggests that the derivative $\frac{d\beta}{d\rho}$ tends to zero as δ gets large. Asymptotically, this is clearly so, since

$$\frac{d\beta}{d\rho} \overset{n\to\infty}{\longrightarrow} \frac{\delta-c}{2\sqrt{2\pi}(1-\rho)^{\frac{3}{2}}} \exp\left\{-\frac{(\delta-c)^2}{2(1-\rho)}\right\} \overset{\delta\to\infty}{\longrightarrow} 0 \,,$$

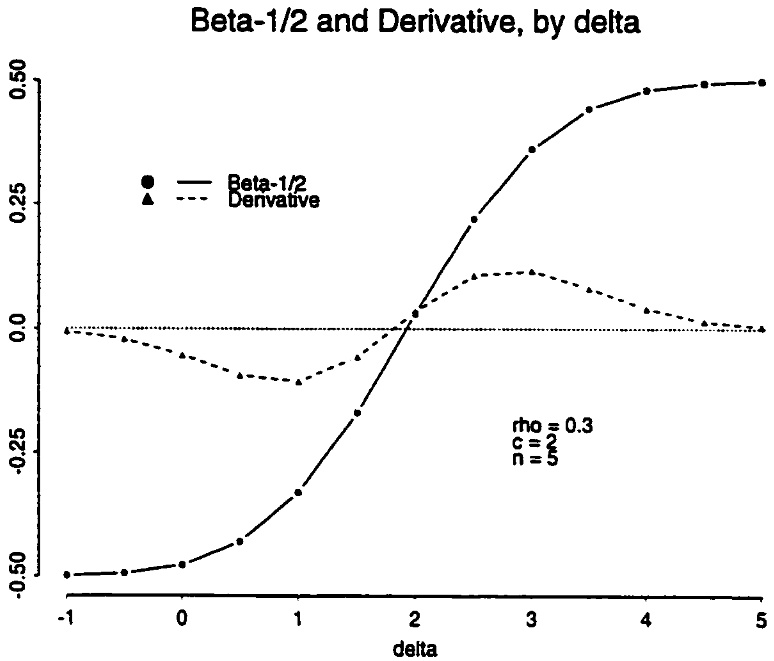


Figure 4.2: The value of the power and derivative functions over a range of δ.

see (4.3). Since $\beta-\frac{1}{2} \stackrel{\delta\to\infty}{\longrightarrow} \frac{1}{2}$ (it being a power function), there can be no k small enough to satisfy (4.1). What this means is that an approach to a proof based on showing that

$$\frac{d\beta}{d\rho} > k\left(\beta-\tfrac{1}{2}\right)$$

cannot be successful.

Figure 4.2 also illustrates that the theorem is true for the values of the parameters used.

4.4 Asymptotic Proof

The theorem is true asymptotically. From (A.6), we have

$$F_{T_2}(t) \stackrel{n\to\infty}{\longrightarrow} \frac{1}{2}+\frac{1}{\pi\sqrt{2}}\sum_{k=0}^{\infty}\frac{\Gamma\left(k+\frac{1}{2}\right)\,(-2)^k\,\Gamma\left(\frac{1}{2}\right)}{\Gamma(k+1)\,\Gamma\left(k+\frac{3}{2}\right)2^{2k+1}}\left(\frac{t}{\sqrt{1-\rho}}-\delta_\rho\right)^{2k+1}.$$

Thus

$$\beta \to \frac{1}{2}+\frac{1}{\sqrt{2\pi}}\sum_{k=0}^{\infty}\frac{\Gamma\left(k+\frac{1}{2}\right)\,(-1)^k}{\Gamma(k+1)\,\Gamma\left(k+\frac{3}{2}\right)2^{k+1}}\left(\frac{\delta-c}{\sqrt{1-\rho}}\right)^{2k+1} \qquad (4.2)$$

$$= \frac{1}{2} + \frac{\delta - c}{\sqrt{2\pi(1-\rho)}} \frac{\Gamma(\frac{3}{2})}{\Gamma(\frac{1}{2})} \sum_{k=0}^{\infty} \frac{\Gamma(k+\frac{1}{2})}{\Gamma(k+\frac{3}{2})\,\Gamma(k+1)} \left(-\frac{(\delta-c)^2}{2(1-\rho)}\right)^k$$
$$= \frac{1}{2} + \frac{\delta - c}{\sqrt{2\pi(1-\rho)}} M\left(\tfrac{1}{2}, \tfrac{3}{2}; -\tfrac{(\delta-c)^2}{2(1-\rho)}\right) .$$

(This expression may also be obtained by noting that asymptotically the two sample t statistic has a normal distribution).

Kummer's transformation $M(\alpha, \beta; x) = e^x\, M(\beta - \alpha, \beta; -x)$ (see *Handbook of Mathematical Functions* [1]) may be applied to the confluent hypergeometric function. Thus $\beta > \frac{1}{2}$ implies

$$0 < \frac{\delta - c}{\sqrt{2\pi(1-\rho)}} \exp\left\{-\tfrac{(\delta-c)^2}{2(1-\rho)}\right\} M\left(1, \tfrac{3}{2}; \tfrac{(\delta-c)^2}{2(1-\rho)}\right) .$$

Note that $\exp\left\{-\frac{(\delta-c)^2}{2(1-\rho)}\right\}$ will always be positive. and $M\left(1, \frac{3}{2}; \frac{(\delta-c)^2}{2(1-\rho)}\right)$, being a sum of positive terms, will always be positive, so $\beta > \frac{1}{2}$ implies $\delta > c$.

Differentiating (4.2) with respect to ρ gives

$$\frac{d\beta}{d\rho} \to \frac{1}{\sqrt{2\pi}} \sum_{k=0}^{\infty} \frac{\Gamma(k+\frac{1}{2})\,(-1)^k}{\Gamma(k+1)\,\Gamma(k+\frac{3}{2})\,2^{k+1}} \left(\frac{\delta-c}{\sqrt{1-\rho}}\right)^{2k+1} \frac{2k+1}{2(1-\rho)}$$
$$= \frac{\delta - c}{2\sqrt{2\pi}(1-\rho)^{\frac{3}{2}}} \sum_{k=0}^{\infty} \frac{\Gamma(k+\frac{1}{2})\,(2k+1)}{\Gamma(k+1)\,(k+\frac{1}{2})\,\Gamma(k+\frac{1}{2})\,2^{k+1}} (-1)^k \left(\frac{(\delta-c)^2}{1-\rho}\right)^k$$
$$= \frac{\delta - c}{2\sqrt{2\pi}(1-\rho)^{\frac{3}{2}}} \exp\left\{-\tfrac{(\delta-c)^2}{2(1-\rho)}\right\} \tag{4.3}$$

which must therefore be positive. Similarly. it can be seen that $\beta = \frac{1}{2}$ implies $\frac{d\beta}{d\rho} = 0$. □

4.5 *Proof for small δ when $\beta = \frac{1}{2}$*

Since the distribution of $T_2|\rho$ will be approximately centered around δ, power will equal approximately $\frac{1}{2}$ when c is close to δ. Thus when $\beta = \frac{1}{2}$, it might be expected that c could be closely approximated by a linear function of δ.

To derive such an approximation, c may be expressed as a power series in δ and coefficients of δ equated at $\beta = \frac{1}{2}$.

The derivation is given in Appendix E. In outline. the distribution function $F_{T_2}|\rho$ derived at (A.3) is restated as a power function and set to $\frac{1}{2}$, and the critical value c is restated as a power series in δ

$$c = \sum_{l=0}^{\infty} c_l \delta^l = c_0 + c_1\delta + c_2\delta^2 + c_3\delta^3 + \cdots ,$$

giving

$$0 = \frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{\sqrt{2}\,\pi\,\Gamma\left(\frac{n-1}{2}\right)} \sum_{i,j=0}^{\infty} \sum_{k=0}^{2j+1} \frac{\Gamma\left(\frac{n-1}{2}+i\right)\Gamma\left(n+i-1+\frac{k}{2}\right)\Gamma\left(j+\frac{1}{2}\right)(-2)^j}{\Gamma(n+i-1)\,\Gamma(2j-k+2)\,\Gamma(i+1)\,\Gamma(k+1)} \times$$
$$\left(\frac{2\rho}{1+\rho}\right)^i \left(-\delta_\rho\right)^{2j-k+1} \left(\frac{1}{\sqrt{n-1}}\right)^k \left(c_0^k + k\,c_0^{k-1}\left(c_1\delta + c_2\delta^2\right) + \frac{k(k-1)}{2} c_0^{k-2}\left(c_1\delta\right)^2 + O(\delta^3)\right) .$$

Equating coefficients of δ gives

$$\begin{aligned} c_0 &= 0 \\ c_1 &= \frac{\sqrt{n-1}\,\Gamma(n-1)}{\Gamma\left(n-\frac{1}{2}\right)\, H\left(\frac{1}{4}, -\frac{1}{4}, \frac{n}{2}; \rho^2\right)} \\ c_2 &= 0 . \end{aligned}$$

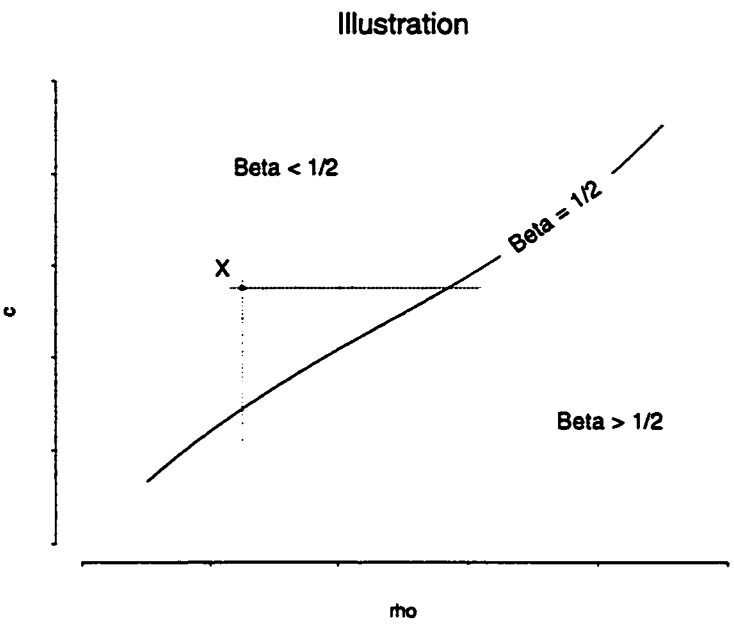
An approximation to the reciprocal of the Gauss hypergeometric function in the expression for c_1 may be derived by equating coefficients of ρ. giving finally

$$c = \delta\sqrt{n-1}\,\frac{\Gamma(n-1)}{\Gamma\left(n-\frac{1}{2}\right)} \left(1 + \frac{1}{8n}\rho^2 + \frac{17n+4}{128n^2\,(n+2)}\rho^4 + O(\rho^6)\right) + O(\delta^3) \qquad (4.4)$$

Hence

$$\frac{dc}{d\rho} = \delta\sqrt{n-1}\,\frac{\Gamma(n-1)}{\Gamma\left(n-\frac{1}{2}\right)} \left(\frac{1}{4n}\rho + \frac{17n+4}{32n^2\,(n+2)}\rho^3 + O(\rho^5)\right) + O(\delta^3)$$

This derivative $\frac{dc}{d\rho}$ will be positive for small positive δ and positive ρ. Consider a contour of constant power ($\beta = \frac{1}{2}$) in the c-ρ space (with δ and n fixed), as illustrated below.



The slope of this contour will be monotone increasing because $\frac{dc}{d\rho}$ is positive. Consider the point X; for a given ρ (and n and δ), power here will be less than $\frac{1}{2}$ because c is larger here than at $\beta=\frac{1}{2}$, and power decreases with increasing c. Now, fixing c, returning to the contour corresponds to an increase in ρ and an increase in β: hence power increases with ρ. Thus when $\beta=\frac{1}{2}$ with fixed c, n and δ, it follows that $\frac{d\beta}{d\rho}$ will be positive, at least up to a first order approximation in δ. □

Assessing the closeness of the approximation for c at (4.4) by computing c_3 the coefficient of δ^3 proved intractable. However, the closely linear relationship may be illustrated graphically. Figure 4.3 shows that when power equals $\frac{1}{2}$, then c increases approximately linearly with δ.

Figure 4.3 also illustrates the correctness of the theorem for the values of the parameters shown, $n=7$ pairs and $\rho=.4$. The dashed line shows where the derivative $\frac{d\beta}{d\rho}=0$. Both the power and the derivative increase towards the bottom right corner. In the c-δ region where $\beta \geq \frac{1}{2}$, it can be seen that $\frac{d\beta}{d\rho}>0$.

The method of equating coefficients cannot be applied when $\beta \neq \frac{1}{2}$, since, equating the

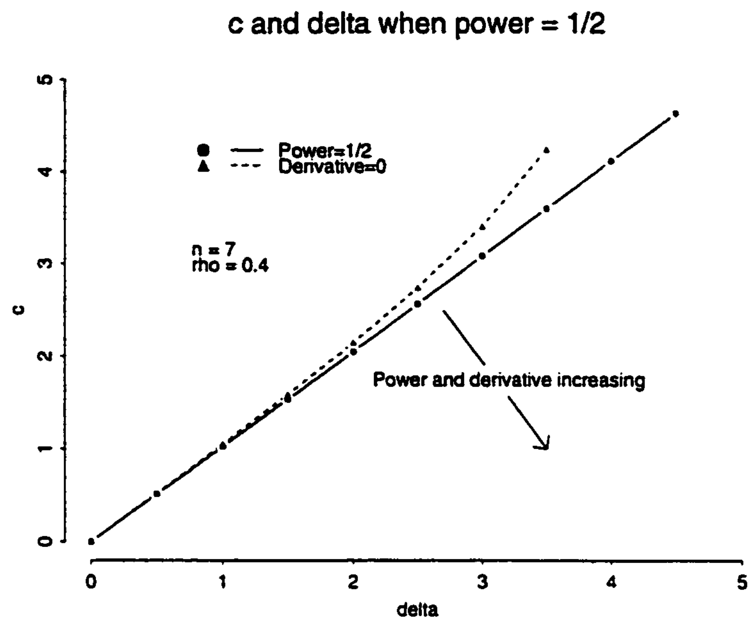


Figure 4.3: The approximately linear relationship of c and δ when $\beta = \frac{1}{2}$.

coefficients of δ^0, one arrives at

$$\beta = \frac{1}{2} - \frac{\left(\frac{c_0}{\sqrt{n-1}}\right)\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{2\sqrt{2\pi}\,\Gamma\left(\frac{n-1}{2}\right)} \sum_{i,j=0}^{\infty} \frac{\Gamma\left(\frac{n-1}{2}+i\right)\Gamma\left(n+i-1+j+\frac{1}{2}\right)\Gamma\left(j+\frac{1}{2}\right)}{\Gamma(n+i-1)\,\Gamma\left(j+\frac{3}{2}\right)\Gamma(i+1)\,\Gamma(j+1)} \left(\frac{2\rho}{1+\rho}\right)^{i} \left(\frac{-c_0^2}{2n-2}\right)^{j}$$

and although the infinite sums may be expressed as a Horn function (see Erdélyi [12], p. 224, equation(7)), the function cannot readily be inverted to express c_0 as a function of β.

4.6 *Proof when* $c=0$

From (A.2) we have

$$F_{T_2}(t) = \frac{1}{2} + \frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{2\,\Gamma\left(\frac{n-1}{2}\right)} \sum_{i=0}^{\infty} \frac{\Gamma\left(\frac{n-1}{2}+i\right)\left(\frac{2\rho}{1+\rho}\right)^{i}}{\Gamma(i+1)} \sum_{j,k=0}^{\infty}$$
$$\left[\frac{\Gamma\left(n+i+k-\frac{1}{2}\right)\Gamma\left(j+k+\frac{1}{2}\right)}{\Gamma(n+i-1)\,\Gamma\left(j+\frac{1}{2}\right)\Gamma\left(k+\frac{3}{2}\right)\Gamma(j+1)\,\Gamma(k+1)} \left(\frac{t}{\sqrt{2n-2}}\right)\left(\frac{-\delta_\rho^2}{2}\right)^{j}\left(\frac{-t^2}{2n-2}\right)^{k}\right.$$

$$-\frac{\Gamma(n+i+k-1)\,\Gamma\left(j+k+\frac{1}{2}\right)}{\Gamma(n+i-1)\,\Gamma\left(j+\frac{3}{2}\right)\Gamma\left(k+\frac{1}{2}\right)\Gamma(j+1)\,\Gamma(k+1)}\left(\frac{\delta_\rho}{\sqrt{2}}\right)\left(\frac{-\delta_\rho^2}{2}\right)^j\left(\frac{-t^2}{2n-2}\right)^k\Bigg]$$

Thus when $c=0$, $\beta>\frac{1}{2}$ implies

$$0 \;<\; \frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{2\sqrt{\pi}\,\Gamma\left(\frac{n-1}{2}\right)}\sum_{i,j=0}^{\infty}\frac{\Gamma\left(\frac{n-1}{2}+i\right)\Gamma(n+i-1)\,\Gamma\left(j+\frac{1}{2}\right)}{\Gamma(n+i-1)\,\Gamma\left(j+\frac{3}{2}\right)\Gamma(i+1)\,\Gamma(j+1)}\left(\frac{2\rho}{1+\rho}\right)^i\left(\frac{\delta_\rho}{\sqrt{2}}\right)\left(\frac{-\delta_\rho^2}{2}\right)^j$$

The i series collapses, since $\frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{\Gamma\left(\frac{n-1}{2}\right)}\sum_{i=0}^{\infty}\frac{\Gamma\left(\frac{n-1}{2}+i\right)}{\Gamma(i+1)}\left(\frac{2\rho}{1+\rho}\right)^i = 1$. The j series may be expressed as a confluent hypergeometric function, and Kummer's transformation $M(\alpha,\beta;x)=e^x\,M(\beta-\alpha,\beta;-x)$ (see [1]) applied, implying

$$\begin{aligned}
0 \;&<\; \frac{1}{2\sqrt{\pi}}\sum_{j=0}^{\infty}\frac{\Gamma\left(j+\frac{1}{2}\right)}{\Gamma\left(j+\frac{3}{2}\right)\Gamma(j+1)}\left(\frac{\delta_\rho}{\sqrt{2}}\right)\left(\frac{-\delta_\rho^2}{2}\right)^j \\
&=\; \frac{1}{2\sqrt{\pi}}\frac{\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{3}{2}\right)}\left(\frac{\delta_\rho}{\sqrt{2}}\right)M\left(\tfrac{1}{2},\tfrac{3}{2},\tfrac{-\delta_\rho^2}{2}\right) \\
&=\; \frac{\delta_\rho}{\sqrt{2\pi}}\,e^{-\frac{\delta_\rho^2}{2}}\,M\left(1,\tfrac{3}{2},\tfrac{\delta_\rho^2}{2}\right)\;.
\end{aligned}$$

$e^{-\frac{\delta_\rho^2}{2}}$ is positive. $M\left(1,\frac{3}{2},\frac{\delta_\rho^2}{2}\right)$, being a sum of positive terms, is positive. Hence

$$\delta_\rho \;>\; 0. \tag{4.5}$$

From (A.7) we have

$$\begin{aligned}
\frac{dF_{T_2|\rho}}{d\rho} \;=\; & \frac{e^{-\frac{\delta_\rho^2}{2}}\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{2\sqrt{2\pi}\,(1-\rho)\,\Gamma\left(\frac{n-1}{2}\right)}\times \\
& \Bigg[\frac{2}{1+\rho}\sum_{i,j,k=0}^{\infty}\frac{\Gamma\left(\frac{n+1}{2}+i\right)\Gamma\left(n+i+\frac{j-1}{2}+k\right)\left(-\frac{1}{2}\right)^k}{\Gamma(n+i)\,\Gamma(i+1)\,\Gamma(j+1)\,\Gamma(k+1)}\left(\frac{2\rho}{1+\rho}\right)^i\left(\delta_\rho\right)^j\left(\frac{t}{\sqrt{n-1}}\right)^{j+2k+1} \\
& -\sum_{i,j,k=0}^{\infty}\frac{\Gamma\left(\frac{n-1}{2}+i\right)\Gamma\left(n+i+\frac{j-2}{2}+k\right)\left(-\frac{1}{2}\right)^k}{\Gamma(n+i-1)\,\Gamma(i+1)\,\Gamma(j+1)\,\Gamma(k+1)}\left(\frac{2\rho}{1+\rho}\right)^i\left(\delta_\rho\right)^{j+1}\left(\frac{t}{\sqrt{n-1}}\right)^{j+2k}\Bigg]
\end{aligned}$$

Thus when $c=0$

$$\frac{d\beta}{d\rho} \;=\; \frac{e^{-\frac{\delta_\rho^2}{2}}\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{2\sqrt{2\pi}\,(1-\rho)\,\Gamma\left(\frac{n-1}{2}\right)}\sum_{i=0}^{\infty}\frac{\Gamma\left(\frac{n-1}{2}+i\right)\Gamma(n+i-1)}{\Gamma(n+i-1)\,\Gamma(i+1)}\left(\frac{2\rho}{1+\rho}\right)^i\left(\delta_\rho\right)$$

$$= \frac{\delta_\rho \, e^{-\frac{\delta_\rho^2}{2}}}{2\sqrt{2\pi}\,(1-\rho)}$$

and in view of (4.5)

$$\frac{d\beta}{d\rho} > 0 \, .$$

Similarly, it can be seen that when $\beta = \frac{1}{2}$ then $\frac{d\beta}{d\rho} = 0$. □

4.7 Induction

Since the theorem has been shown to be true asymptotically, and also considerably simplified, induction would appear to be a promising approach to a proof for finite n. The normal direction of the induction would be reversed; that is, instead of showing that if the theorem is true for n then it is true for $n+1$, one would show that if it is true for n then it is true for $n-1$; then knowing that it was true as $n \to \infty$ would yield the conclusion. Schematically, the question may be stated

$$\begin{array}{rlcl} \text{Given} & \beta_n \geq \frac{1}{2} & \Rightarrow & \frac{d\beta_n}{d\rho} > 0, \\ \text{show that} & \beta_{n-1} \geq \frac{1}{2} & \Rightarrow & \frac{d\beta_{n-1}}{d\rho} > 0 \end{array}$$

where $\Rightarrow$ means "implies". A straightforward approach is to attempt to show that

$$\begin{array}{ccc} \beta_n \geq \frac{1}{2} & & \frac{d\beta_n}{d\rho} > 0 \\ \Uparrow & \text{and} & \Downarrow \\ \beta_{n-1} \geq \frac{1}{2} & & \frac{d\beta_{n-1}}{d\rho} > 0 \end{array}$$

which yields, with the condition

$$\begin{array}{ccc} \beta_n \geq \frac{1}{2} & \Rightarrow & \frac{d\beta_n}{d\rho} > 0 \\ \Uparrow & & \Downarrow \\ \beta_{n-1} \geq \frac{1}{2} & & \frac{d\beta_{n-1}}{d\rho} > 0 \end{array}$$

and the conclusion follows by the chain of implications.

Showing that $\begin{array}{c} \frac{d\beta_n}{d\rho} > 0 \\ \Downarrow \\ \frac{d\beta_{n-1}}{d\rho} > 0 \end{array}$ simply requires showing that $\frac{d\beta_{n-1}}{d\rho} > \frac{d\beta_n}{d\rho}$, which appears from numerical analysis to be true.

However, unfortunately $\begin{matrix}\beta_n \geq \frac{1}{2} \\ \not\Downarrow \\ \beta_{n-1} \geq \frac{1}{2}\end{matrix}$, because $\beta_n \not\geq \beta_{n-1}$. Numerical analysis suggests in fact that β decreases with increasing n. (Note that here, δ is defined as the 'contiguous effect size', that is, $\delta = \frac{\mu_2 - \mu_1}{\sigma}\sqrt{\frac{\pi}{2}}$, so when '$\delta$ is fixed' and n varies, it is assumed that $\mu_2 - \mu_1$ varies so as to keep δ constant. Hence the previous statement is not inconsistent with the known fact that power increases with n).

4.8 *Attempted proofs at* $\rho = 0$

The theorem is considerably simplified at $\rho = 0$ since then the power and derivative functions are both simpler and moie similar. Several attempts were therefore made to prove the theorem at $\rho = 0$.

When $\rho = 0$, the distribution function $F_{T_2}(t)$ derived at (A.3) may be re-expressed as a power function

$$\beta - \tfrac{1}{2} = \frac{1}{\sqrt{2}\,\pi\,\Gamma(n-1)} \sum_{j=0}^{\infty} \sum_{k=0}^{2j+1} \frac{\Gamma(n-1+\frac{k}{2})\,\Gamma(j+\frac{1}{2})}{\Gamma(2j-k+2)\,\Gamma(k+1)} \left(\delta\right)\left(-2\delta^2\right)^j \left(\frac{-c}{\delta\sqrt{n-1}}\right)^k \quad (4.6)$$

Similarly at $\rho = 0$, the derivative at (A.7) becomes

$$\frac{d\beta}{d\rho} = \frac{e^{-\frac{\delta_\rho^2}{2}}}{2\sqrt{2\pi}} \left[\sum_{j,k=0}^{\infty} \frac{\Gamma\left(n+\frac{j-2}{2}+k\right)\left(-\frac{1}{2}\right)^k}{\Gamma(n-1)\,\Gamma(j+1)\,\Gamma(k+1)} \left(\delta\right)^{j+1} \left(\frac{c}{\sqrt{n-1}}\right)^{j+2k} - \sum_{j,k=0}^{\infty} \frac{\Gamma\left(n+\frac{j-1}{2}+k\right)\left(-\frac{1}{2}\right)^k}{\Gamma(n-1)\,\Gamma(j+1)\,\Gamma(k+1)} \left(\delta_\rho\right)^j \left(\frac{c}{\sqrt{n-1}}\right)^{j+2k+1} \right]$$

which may be shown to be equivalent to

$$\frac{d\beta}{d\rho} = \frac{1}{\sqrt{2}\,\pi\,\Gamma(n-1)} \sum_{j=0}^{\infty} \sum_{k=0}^{2j+1} \frac{\Gamma(n-1+\frac{k}{2})\,\Gamma(j+\frac{3}{2})}{\Gamma(2j-k+2)\,\Gamma(k+1)} \left(\delta\right)\left(-2\delta^2\right)^j \left(\frac{-c}{\delta\sqrt{n-1}}\right)^k \quad (4.7)$$

Note that the only respect in which the expression for the derivative differs from the power function is that the former has $\Gamma(j+\frac{3}{2})$ where the second has $\Gamma(j+\frac{1}{2})$.

Several methods were tried for comparing the two functions; the following sections describe them.

4.9 *Proof when* $\rho=0$ *and* $\delta=0$

When $\rho=0$ and $\delta=0$, the proof is straightforward. (4.6) becomes

$$\beta-\tfrac{1}{2} = \frac{1}{\sqrt{2}\,\pi\,\Gamma(n-1)}\sum_{j=0}^{\infty}\frac{\Gamma\left(n-1+\frac{2j+1}{2}\right)\Gamma(j+\frac{1}{2})}{\Gamma(2j-(2j+1)+2)\,\Gamma(2j+1+1)}\left(\delta\right)\left(-2\delta^2\right)^j\left(\frac{-c}{\delta\sqrt{n-1}}\right)^{2j+1}$$

Some algebra yields

$$\beta-\tfrac{1}{2} = \frac{-c\,\Gamma(n-\frac{1}{2})}{\sqrt{2\pi(n-1)}\,\Gamma(n-1)}\,H\left(n-\tfrac{1}{2},\tfrac{1}{2},\tfrac{3}{2};\tfrac{-c^2}{2(n-1)}\right)$$

The transformation $H(\alpha,\beta,\gamma;x) \equiv (1-x)^{-\alpha}H(\alpha,\gamma-\beta,\gamma;\frac{x}{x-1})$ (Erdélyi [12], p.109) may be applied, giving

$$\beta-\tfrac{1}{2} = \frac{-c\,\Gamma(n-\frac{1}{2})}{\sqrt{2\pi(n-1)}\,\Gamma(n-1)}\left(\tfrac{2(n-1)}{c^2+2(n-1)}\right)^{n-\frac{1}{2}} H\left(n-\tfrac{1}{2},1,\tfrac{3}{2};\tfrac{c^2}{c^2+2(n-1)}\right)$$

The Gauss hypergeometric function is a sum of positive terms and hence is positive. Thus $\beta\geq\frac{1}{2}$ implies $c\leq 0$.

For the derivative, when $\delta=0$, (4.7) becomes

$$\begin{aligned}\frac{d\beta}{d\rho} &= \frac{-c}{2\sqrt{2\pi(n-1)}\,\Gamma(n-1)}\sum_{j=0}^{\infty}\frac{\Gamma(n-\frac{1}{2}+j)}{\Gamma(j+1)}\left(\tfrac{-c^2}{2(n-1)}\right)^j \\ &= \frac{-c\,\Gamma(n-\frac{1}{2})}{2\sqrt{2\pi(n-1)}\,\Gamma(n-1)}\left(\tfrac{2(n-1)}{c^2+2(n-1)}\right)^{n-\frac{1}{2}}\end{aligned}$$

Thus $\frac{d\beta}{d\rho}\geq 0$. □

4.10 *Total Positivity*

4.10.1 Background

Total positivity is a characteristic of certain mathematical functions. Totally positive functions figure prominently in problems involving convexity and moment spaces. Many probability density functions are totally positive. They are dealt with extensively in *Total Positivity* by Samuel Karlin [23].

Closely related to the concept of total positivity are the concepts of sign regularity and sign consistency. For our purposes we will use the concept of sign consistency. A function

$F(x,y)$ is sign consistent of order 2 (SC_2) on $\mathcal{X}\times\mathcal{Y}$ if there is an ε, either +1 or −1, such that

$$\varepsilon \begin{vmatrix} F(x_1,y_1) & F(x_1,y_2) \\ F(x_2,y_1) & F(x_2,y_2) \end{vmatrix} \geq 0$$
$$\forall \ \{x_1,x_2 \,|\, x_1<x_2;\, x_i\in\mathcal{X}\} \text{ and } \{y_1,y_2 \,|\, y_1<y_2;\, y_i\in\mathcal{Y}\}.$$

The order of the sign consistency refers to the number of points $x_1, x_2, x_3, \cdots, y_1, y_2, y_3, \cdots$, for which the sign consistency holds. Thus for a function that is SC_4, the determinant of the 4×4 array of functions on $\{x_1,x_2,x_3,x_4 \,|\, x_1<x_2<x_3<x_4;\, x_i\in\mathcal{X}\} \times \{y_1,y_2,y_3,y_4 \,|\, y_1<y_2<y_3<y_4;\, y_i\in\mathcal{Y}\}$ has a consistent sign. For our purposes, we will only be concerned with sign consistent functions of order 2 (SC_2 functions).

By way of illustration, the Poisson density function

$$\frac{\exp\{-\lambda\}\,\lambda^x}{x!}, \quad x=0,1,2,\cdots \quad \lambda>0$$

is SC_2 with $\varepsilon=+1$, since

$$\begin{aligned} &\begin{vmatrix} \frac{\exp\{-\lambda_1\}\,\lambda_1^{x_1}}{x_1!} & \frac{\exp\{-\lambda_2\}\,\lambda_2^{x_1}}{x_1!} \\ \frac{\exp\{-\lambda_1\}\,\lambda_1^{x_2}}{x_2!} & \frac{\exp\{-\lambda_2\}\,\lambda_2^{x_2}}{x_2!} \end{vmatrix} \\ = \; & \frac{\exp\{-\lambda_1\}\,\lambda_1^{x_1}}{x_1!}\frac{\exp\{-\lambda_2\}\,\lambda_2^{x_2}}{x_2!} - \frac{\exp\{-\lambda_2\}\,\lambda_2^{x_1}}{x_1!}\frac{\exp\{-\lambda_1\}\,\lambda_1^{x_2}}{x_2!} \\ = \; & \frac{\exp\{-\lambda_1\}\,\exp\{-\lambda_2\}\,\lambda_1^{x_1}\,\lambda_2^{x_1}}{x_1!\,x_2!}\left(\lambda_2^{x_2-x_1} - \lambda_1^{x_2-x_1}\right) \end{aligned}$$

which will always be positive for $x_1<x_2$, $\lambda_1<\lambda_2$ on the domain of the function. The density functions of all exponential family distributions are sign consistent.

A Composition Theorem

Karlin (at chapter 1 §2) shows the following to be true (we simplify the presentation to the extent that it is applicable to our purposes).

If $F(x,z)$ and $G(y,z)$ and are both SC_2, and

$$H(x,y) \;=\; \int_z F(x,z)\,G(y,z)\,dz$$

then $H(x,y)$ is SC_2 with $\varepsilon_H=\varepsilon_F\varepsilon_G$. Where z consists of a discrete set, the integral is interpreted as a sum.

4.10.2 Applicability to our Theorem 1

Put:

$$H(c,\alpha;\delta,n) \equiv \frac{1}{\sqrt{2}\pi\Gamma(n-1)}\sum_{j=0}^{\infty}\sum_{k=0}^{2j+1}\frac{\Gamma(n-1+\frac{k}{2})\,\Gamma(j+\alpha)}{\Gamma(2j-k+2)\,\Gamma(k+1)}\left(\delta\right)\left(-2\delta^2\right)^j\left(\frac{-c}{\delta\sqrt{n-1}}\right)^k ,$$

then when $\rho=0$, $\beta_{MU}-\frac{1}{2}=H(c,\frac{1}{2};\delta,n)$ and $\frac{d\beta_{MU}}{d\rho}=H(c,\frac{3}{2};\delta,n)$.

If $H(c,\alpha;\delta,n)$ is SC_2 (for $\alpha\in\{\frac{1}{2},\frac{3}{2}\}$), then (by the definition of SC_2, for $\epsilon>0$)

$$\begin{vmatrix} H(c,\frac{1}{2};\delta,n) & H(c+\epsilon,\frac{1}{2};\delta,n) \\ H(c,\frac{3}{2};\delta,n) & H(c+\epsilon,\frac{3}{2};\delta,n) \end{vmatrix} > 0$$

$$H(c,\tfrac{1}{2};\delta,n)\,H(c+\epsilon,\tfrac{3}{2};\delta,n) > H(c,\tfrac{3}{2};\delta,n)\,H(c+\epsilon,\tfrac{1}{2};\delta,n) .$$

Now $H(c,\frac{1}{2};\delta,n)=0 \Rightarrow H(c+\epsilon,\frac{1}{2};\delta,n)<0$, since power decreases with increasing c. Hence, when $\rho=0$ and $\beta_{MU}=\frac{1}{2}$

$$H(c,\tfrac{3}{2};\delta,n)\equiv\frac{d\beta_{MU}}{d\rho}>0 .$$

We therefore need to show that $H(c,\alpha;\delta,n)$ is SC_2. Put

$$\begin{aligned} F(j,\alpha) &= \Gamma(j+\alpha) \\ G(c,j;\delta,n) &= \sum_{k=0}^{2j+1}\frac{\Gamma(n-1+\frac{k}{2})}{\Gamma(2j-k+2)\,\Gamma(k+1)}\left(\delta\right)\left(-2\delta^2\right)^j\left(\frac{-c}{\delta\sqrt{n-1}}\right)^k \end{aligned}$$

By the composition theorem, $H(c,\alpha;\delta,n)$ is SC_2 if $F(j,\alpha)$ is SC_2 and $G(c,j;\delta,n)$ is SC_2.

$F(j,\alpha)$ is SC_2 since it is easily shown that

$$\begin{vmatrix} \Gamma(j+\alpha) & \Gamma(j+1+\alpha) \\ \Gamma(j+\alpha+\epsilon) & \Gamma(j+1+\alpha+\epsilon) \end{vmatrix} = \epsilon\Gamma(j+\alpha)\Gamma(j+\alpha+\epsilon) > 0$$

for $\alpha>0$, $\epsilon>0$, $j=0,1,2,\ldots$.

For $G(c,j;\delta,n)$, straightforward algebra shows that

$$\begin{vmatrix} G(c,j;\delta,n) & G(c+\epsilon,j;\delta,n) \\ G(c,j+1;\delta,n) & G(c+\epsilon,j+1;\delta,n) \end{vmatrix}$$

$$= \frac{(2\delta)^2}{\Gamma^2(n-1)}\sum_{k=0}^{2j+1}\sum_{l=0}^{2j+3}\frac{\Gamma(n-1+\frac{k}{2})\,\Gamma(n-1+\frac{l}{2})}{\Gamma(2j-k+2)\,\Gamma(2j-l+4)\,\Gamma(k+1)\,\Gamma(l+1)}$$

$$\times \left(\left(\frac{-(c+\epsilon)}{\delta\sqrt{n-1}}\right)^k \left(\frac{-c}{\delta\sqrt{n-1}}\right)^l - \left(\frac{-c}{\delta\sqrt{n-1}}\right)^k \left(\frac{-(c+\epsilon)}{\delta\sqrt{n-1}}\right)^l \right)$$

but there is no readily apparent way to determine if the expression has a consistent sign.

The composition theorem may not be used again to show that $G(c, j; \delta, n)$ is SC_2 since the function cannot be factorized into an expression in c and j, as the index of summation depends on j.

If were possible to show that $G(c, j; \delta, n)$, and hence $H(c, \alpha; \delta, n)$, are SC_2, it would follow that, at $\rho=0$, $\frac{d\beta_{MU}}{d\rho} > 0$ when $\beta_{MU} = \frac{1}{2}$.

4.10.3 Total Positivity and the general case, $\rho \neq 0$

Numerical analysis suggests that the function $\beta_{MU} - \frac{1}{2}$ is TP_2 in c and ρ when $\beta_{MU} > \frac{1}{2}$ and $\rho > 0$ (provided $c > 0$; the function would not be a 'power function' in the normal sense when $c < 0$). If it were possible to prove this algebraically, the theorem could be proved by the following lemma.

Lemma

Given

$$c_0 < c_1$$

$$\rho_0 < \rho_1$$

$$\beta(c_0, \rho_0)\,\beta(c_1, \rho_1) > \beta(c_0, \rho_1)\,\beta(c_1, \rho_0) \tag{4.8}$$

$$\beta(c_1, \rho) < \beta(c_0, \rho) \;\; \forall \rho \tag{4.9}$$

then

$$\beta(c, \rho_1) > \beta(c, \rho_0)$$

whenever

$$\beta(c, \rho_0) \geq 0$$

where

i) $\beta(c, \rho) = \beta(c, \rho; \delta, n) = \beta_{MU} - \frac{1}{2} \in (-\frac{1}{2}, \frac{1}{2})$, the power function considered as a function of c and ρ, with δ and n arbitrary and fixed;

ii) $\rho \in (-1, 1)$;

iii) $c \in (-\infty, \infty)$ is the quantile (critical value).

Proof.

Choose c_1, ρ_0 so that $\beta(c_1, \rho_0) = 0$. From (4.9) it follows that $\beta(c_0, \rho_0) > 0$ (for any $c_0 < c_1$), and hence by (4.8) that

$$\beta(c_1, \rho_1) \;\; > \;\; 0 = \beta(c_1, \rho_0). \tag{4.10}$$

Now choose c_0 so that $\beta(c_0, \rho_0) = \beta(c_1, \rho_1)$. By (4.9) again we have

$$\beta(c_0, \rho_0) = \beta(c_1, \rho_1) \;\; < \;\; \beta(c_0, \rho_1) \; . \tag{4.11}$$

Taking (4.10) and (4.11) together, we see that

$$\beta(c, \rho_1) > \beta(c, \rho_0)$$

whenever

$$\beta(c, \rho_0) \geq 0 \; . \qquad \square$$

(4.8) would follow from the total positivity of $\beta_{MU}(c, \rho; \delta, n) - \frac{1}{2}$; (4.9) is a consequence of power decreasing with increasing c.

4.11 *When $\rho = 0$: Remove alternating signs*

One of the primary difficulties in dealing with the expressions for the power function (4.6) and the derivative function (4.7) is that the signs of the terms in the power series alternate positive and negative. The negative signs may be removed by applying the transformation

$$H(\alpha, \beta, \gamma; x) \;\; \equiv \;\; (1-x)^{-\alpha} H(\alpha, \gamma-\beta, \gamma; \frac{x}{x-1})$$

(Erdélyi [12], p.109), and Kummer's transformation (see [1])

$$M(\alpha, \beta; x) \;\; \equiv \;\; e^{x} M(\beta-\alpha, \beta; -x) \; .$$

However, the resulting expressions have then lost the neat symmetry of (4.6) and (4.7).

$\beta - \frac{1}{2}$

$$= \frac{e^{-\frac{\delta_\rho^2}{2}}}{2} \sum_{j,k=0}^{\infty} \frac{\Gamma(n-1+j)}{\Gamma(n-1)\,\Gamma(j+k+\frac{3}{2})\,\Gamma(j+1)} \left(\tfrac{2(n-1)}{c^2+2(n-1)}\right)^{n-1} \left(\tfrac{c^2}{c^2+2(n-1)}\right)^{j} \left(\tfrac{\delta_\rho^2}{2}\right)^{j+k+\frac{1}{2}}$$

$$- \frac{e^{-\frac{\delta_\rho^2}{2}}}{2} \sum_{j,k=0}^{\infty} \frac{\Gamma(n-1+j+k+\frac{1}{2})}{\Gamma(n-1)\,\Gamma(j+k+\frac{3}{2})\,\Gamma(k+1)} \left(\tfrac{2(n-1)}{c^2+2(n-1)}\right)^{n-1} \left(\tfrac{c^2}{c^2+2(n-1)}\right)^{j+k+\frac{1}{2}} \left(\tfrac{\delta_\rho^2}{2}\right)^{k}$$

$$\frac{d\beta}{d\rho}$$

$$= \frac{e^{-\frac{\delta_\rho^2}{2}}}{2} \sum_{j=0}^{\infty} \frac{\Gamma(n-1+j)}{\Gamma(n-1)\,\Gamma(j+\frac{1}{2})\,\Gamma(j+1)} \left(\tfrac{2(n-1)}{c^2+2(n-1)}\right)^{n-1} \left(\tfrac{c^2}{c^2+2(n-1)}\right)^{j} \left(\tfrac{\delta_\rho^2}{2}\right)^{j+\frac{1}{2}}$$

$$+ \frac{e^{-\frac{\delta_\rho^2}{2}}}{2} \sum_{j=0}^{\infty} \frac{\Gamma(n-1+j+\frac{1}{2})}{\Gamma(n-1)\,\Gamma(j+\frac{3}{2})\,\Gamma(j+1)} \left(\tfrac{2(n-1)}{c^2+2(n-1)}\right)^{n-1} \left(\tfrac{c^2}{c^2+2(n-1)}\right)^{j+\frac{1}{2}} \left(\tfrac{\delta_\rho^2}{2}\right)^{j+1}$$

$$- \frac{e^{-\frac{\delta_\rho^2}{2}}}{2} \sum_{j=0}^{\infty} \frac{\Gamma(n-1+j+\frac{1}{2})}{\Gamma(n-1)\,\Gamma(j+\frac{1}{2})\,\Gamma(j+1)} \left(\tfrac{2(n-1)}{c^2+2(n-1)}\right)^{n-1} \left(\tfrac{c^2}{c^2+2(n-1)}\right)^{j+\frac{1}{2}} \left(\tfrac{\delta_\rho^2}{2}\right)^{j}$$

$$- \frac{e^{-\frac{\delta_\rho^2}{2}}}{2} \sum_{j=0}^{\infty} \frac{\Gamma(n-1+j+1)}{\Gamma(n-1)\,\Gamma(j+\frac{3}{2})\,\Gamma(j+1)} \left(\tfrac{2(n-1)}{c^2+2(n-1)}\right)^{n-1} \left(\tfrac{c^2}{c^2+2(n-1)}\right)^{j+1} \left(\tfrac{\delta_\rho^2}{2}\right)^{j+\frac{1}{2}}$$

The expressions do not appear to lead to a proof.

4.12 When $\rho=0$: Contiguous Horn functions

Another approach is to express the power and derivative functions (at $\rho = 0$) as hypergeometric functions in two variables (a.k.a. Horn functions) and apply known results for contiguous Horn functions. From the distribution functions at (A.2) and initially setting $\beta=\frac{1}{2}$, at $\rho=0$, this yields

$$\text{Given}: \quad \beta-\tfrac{1}{2} \equiv \frac{1}{\sqrt{\pi}} \Bigg[\tfrac{\delta_\rho}{\sqrt{2}}\, \Psi_1\!\left(\tfrac{1}{2}, n-1; \tfrac{1}{2}, \tfrac{3}{2}; \tfrac{-c^2}{2(n-1)}, \tfrac{-\delta_\rho^2}{2}\right) - \tfrac{c}{\sqrt{2(n-1)}} \frac{\Gamma(n-\frac{1}{2})}{\Gamma(n-1)}\, \Psi_1\!\left(\tfrac{1}{2}, n-\tfrac{1}{2}; \tfrac{3}{2}, \tfrac{1}{2}; \tfrac{-c^2}{2(n-1)}, \tfrac{-\delta_\rho^2}{2}\right) \Bigg] = 0$$

$$\text{show that } \frac{d\beta}{d\rho} \equiv \frac{1}{2\sqrt{\pi}} \Bigg[\tfrac{\delta_\rho}{\sqrt{2}}\, \Psi_1\!\left(\tfrac{3}{2}, n-1; \tfrac{1}{2}, \tfrac{3}{2}; \tfrac{-c^2}{2(n-1)}, \tfrac{-\delta_\rho^2}{2}\right) - \tfrac{c}{\sqrt{2(n-1)}} \frac{\Gamma(n-\frac{1}{2})}{\Gamma(n-1)}\, \Psi_1\!\left(\tfrac{3}{2}, n-\tfrac{1}{2}; \tfrac{3}{2}, \tfrac{1}{2}; \tfrac{-c^2}{2(n-1)}, \tfrac{-\delta_\rho^2}{2}\right) \Bigg] > 0$$

where $\Psi_1(\alpha,\beta,\gamma,\gamma';x,y)) \equiv \sum_{i,j=0}^{\infty} \frac{(\alpha)_{i+j}\,(\beta)_i}{(\gamma)_i\,(\gamma')_j} \frac{x^i}{i!} \frac{y^j}{j!}$ is the Horn function given by Erdélyi [12], p. 225.

The Horn functions in the condition and conclusion are contiguous in that the first parameters differ by 1.

Dr. R. G. Buschman of Langlois, Oregon, kindly provided (by personal communication) the relation

$$\begin{aligned}(\alpha-\beta-\gamma'+1)\,\Psi_1(\alpha,\beta,\gamma,\gamma';x,y) \;&-\; \alpha\,\Psi_1(\alpha+1,\beta,\gamma,\gamma';x,y) \\ &+ \beta\,\Psi_1(\alpha,\beta+1,\gamma,\gamma';x,y) \;+\; (\gamma'-1)\,\Psi_1(\alpha,\beta,\gamma,\gamma'-1;x,y) = 0\end{aligned}$$

for contiguous Horn functions. However, each application (substituting for the conclusion terms and subtracting the condition) always left at least one pair of terms to be shown to be > 0. Without some other relation, this method could not work.

Since the power must be increasing in δ and decreasing in c, the derivatives of the power function with respect to δ and c must always be positive and negative respectively. This provided two other relations involving Ψ_1 functions: however, we could not find the transformations that would enable them to be incorporated.

This approach also encounters the 'no k small enough' problem described above in 4.3.

4.13 When $\rho=0$: Parallel to the asymptotic proof

From the asymptotic proof of the theorem (at 4.4) and setting $\rho=0$ we have

$$\beta_{MU} \overset{n\to\infty}{\longrightarrow} \frac{1}{2} + \frac{\delta-c}{\sqrt{2\pi}} \exp\left\{-\tfrac{(\delta-c)^2}{2}\right\} M\!\left(1,\tfrac{3}{2};\tfrac{(\delta-c)^2}{2}\right)$$

and

$$\frac{d\beta_{MU}}{d\rho} \overset{n\to\infty}{\longrightarrow} \frac{\delta-c}{2\sqrt{2\pi}} \exp\left\{-\tfrac{(\delta-c)^2}{2}\right\} .$$

The two expressions differ in the confluent hypergeometric function $M\!\left(1,\tfrac{3}{2};\tfrac{(\delta-c)^2}{2}\right)$, and the asymptotic proof relies on the positiveness of this function, it being an infinite sum, each term of which is a power of $(\delta-c)^2$.

The mechanism of convergence of the power function for finite n (at (4.6)) is known

$$\frac{\Gamma(n-1+\frac{k}{2})}{\Gamma(n-1)}\left(\frac{-c}{\delta\sqrt{n-1}}\right)^{k} \overset{n\to\infty}{\longrightarrow} \left(\frac{-c}{\delta}\right)^{k}$$

and $((\delta-c)^2)^l$ may be re-expressed

$$((\delta-c)^2)^l \equiv \sum_{k=0}^{2l}\binom{2l}{k}\left(\delta\right)^{2l-k}\left(-c\right)^{k},$$

thus one has for the convergence of the power function

$$\left(\delta\right)^{2l}\sum_{k=0}^{2l}\binom{2l}{k}\frac{\Gamma(n-1+\frac{k}{2})}{\Gamma(n-1)}\left(\frac{-c}{\delta\sqrt{n-1}}\right)^{k} \overset{n\to\infty}{\longrightarrow} \left(\delta-c\right)^{2l}.$$

Given the righthand side, i. e. the asymptotic expression, is always positive, one may ask whether the lefthand side, the expression for finite n, is also always positive?

The question my be re-phased as follows. The expression

$$\frac{1}{\Gamma(2j+2)}\sum_{k=0}^{2j+1}\binom{2j+1}{k}\frac{\Gamma(n-1+\frac{k}{2})}{\Gamma(n-1)}\left(\frac{-c}{\delta\sqrt{n-1}}\right)^{k}$$

appears in both the power function (4.6) and the derivative function (4.7) (in the form $\sum_{k=0}^{2j+1}\frac{\Gamma(n-1+\frac{k}{2})}{\Gamma(n-1)\,\Gamma(2j-k+2)\,\Gamma(k+1)}\left(\frac{-c}{\delta\sqrt{n-1}}\right)^{k}$). When n is large, the asymptotic expression leads to a straightforward proof. What can be said about the expression for finite n?

Simplifying the question slightly, can we show that

$$f_n(x,j) \equiv \sum_{k=0}^{2j}\binom{2j}{k}\frac{\Gamma(n-1+k)}{\Gamma(n-1)}\left(\frac{-x}{n-1}\right)^{k}$$

is always positive? If it is, and it appears to be, the methods by which it were proved would lead to a proof of the theorem as a whole.

Numerical analysis suggests that it is. Figure 4.4 (plotted on logarithmic scales) shows that for the values of the parameters used there, the function was positive over a wide range of n. Figure 4.4 also suggests that

$$f_{n-1}(x,j) > f_n(x,j)$$

which if it could be shown algebraicly would show that $f_n(x,j) > 0$, since we know that $f_\infty(x,j) > 0$.

4.13.1 Difference of contiguous terms

Some algebra shows that

$$\begin{aligned} f_{n-1}(x,j) - f_n(x,j) &\equiv \sum_{k=0}^{2j} \binom{2j}{k} \frac{\Gamma(n-2+k)}{\Gamma(n-2)} \left(\frac{-x}{n-2}\right)^k - \sum_{k=0}^{2j} \binom{2j}{k} \frac{\Gamma(n-1+k)}{\Gamma(n-1)} \left(\frac{-x}{n-1}\right)^k \\ &= \sum_{k=0}^{2j} \binom{2j}{k} \left(-x\right)^k \left(\frac{\prod_{i=1}^{k}(n-3+i)}{(n-2)^k} - \frac{\prod_{i=1}^{k}(n-2+i)}{(n-1)^k} \right) \end{aligned}$$

and although further simplification is possible no conclusion could be reached.

4.13.2 Term by term expansion of the $\Gamma(\cdot)$ function.

The $\Gamma(\cdot)$ functions do not depend on j and hence the term by term expansion of $\frac{\Gamma(n-1+k)}{\Gamma(n-1)}\left(\frac{-x}{n-1}\right)^k$ may proceed without j being specified

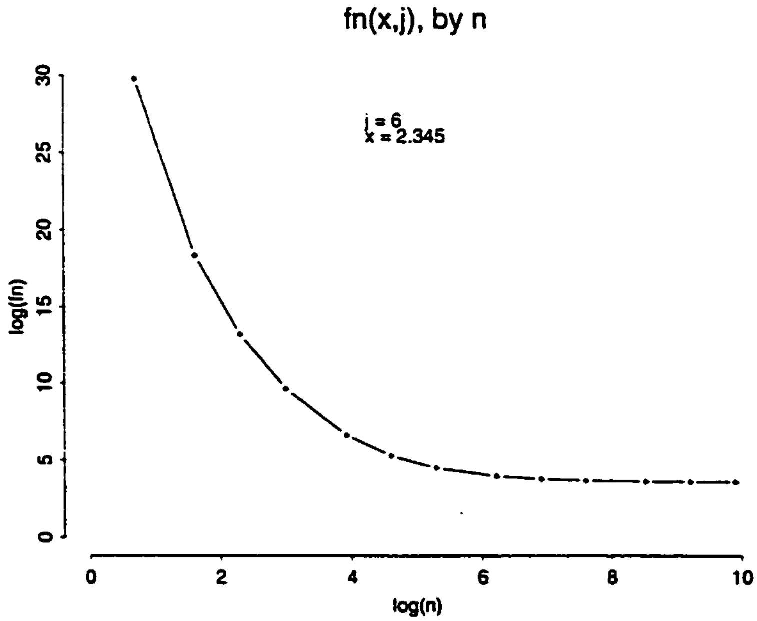


Figure 4.4: The value of the function discussed in 4.13, by n.

k	$\binom{2j}{k}\frac{\Gamma(n-1+k)}{\Gamma(n-1)}\left(\frac{-x}{n-1}\right)^k$
0	$\binom{2j}{k}\,1$
1	$\binom{2j}{k}\,(-x)$
2	$\binom{2j}{k}\,(-x)^2\left[1+\frac{1}{n-1}\right]$
3	$\binom{2j}{k}\,(-x)^3\left[1+\frac{3}{n-1}+\frac{2}{(n-1)^2}\right]$
4	$\binom{2j}{k}\,(-x)^4\left[1+\frac{6}{n-1}+\frac{11}{(n-1)^2}+\frac{6}{(n-1)^2}\right]$
5	$\binom{2j}{k}\,(-x)^5\left[1+\frac{10}{n-1}+\frac{35}{(n-1)^2}+\frac{50}{(n-1)^3}+\frac{24}{(n-1)^4}\right]$
6	$\binom{2j}{k}\,(-x)^6\left[1+\frac{15}{n-1}+\frac{85}{(n-1)^2}+\frac{225}{(n-1)^3}+\frac{274}{(n-1)^4}+\frac{120}{(n-1)^5}\right]$
$\vdots$	$\vdots$
generally	$\binom{2j}{k}\,(-x)^k\prod_{i=1}^{k}\left(1+\frac{i-1}{n-1}\right)$

It can be readily seen that

$$\sum_{k=0}^{2j}\binom{2j}{k}\frac{\Gamma(n-1+k)}{\Gamma(n-1)}\left(\frac{-x}{n-1}\right)^k \overset{n\to\infty}{\longrightarrow} \sum_{k=0}^{2j}\binom{2j}{k}\left(x\right)^k$$

but no pattern could be discerned in the coefficients of $\frac{1}{n-1}$, $\frac{1}{(n-1)^2}$, etc. However, the expansion does suggest a method for demonstrating the relation for fixed j.

4.13.3 Illustration for $j=0$, $j=1$, $j=2$, $j=3$

The relation is clearly true for $j=0$.

For $j=1$ we have

$$\begin{aligned} f_n(x,1) &\equiv \sum_{k=0}^{2}\binom{2}{k}\frac{\Gamma(n-1+k)}{\Gamma(n-1)}\left(\frac{-x}{n-1}\right)^k \\ &= \binom{2}{0}1+\binom{2}{1}(-x)+\binom{2}{2}(-x)^2\left[1+\frac{1}{n-1}\right] \\ &= (1-x)^2+\frac{x^2}{n-1} \quad > \quad 0\,. \end{aligned}$$

When $j=2$ it can be shown, by subtracting the binomial expansion defined by the first two terms of the power series, that

$$f_n(x,2) \equiv \sum_{k=0}^{4}\binom{4}{k}\frac{\Gamma(n-1+k)}{\Gamma(n-1)}\left(\frac{-x}{n-1}\right)^k$$

$$= (1-x)^4 + \tfrac{6x^2}{n-1}\left(1 - x\tfrac{3n-1}{3n-3}\right)^2 + \tfrac{x^4(9n+1)}{3(n-1)^3} \quad > \quad 0$$

When $j = 3$ it can similarly be shown that

$$\begin{aligned} f_n(x,3) &\equiv \sum_{k=0}^{6} \binom{6}{k} \frac{\Gamma(n-1+k)}{\Gamma(n-1)} \left(\tfrac{-x}{n-1}\right)^k \\ &= (1-x)^6 + \tfrac{15x^2}{n-1}\left(1 - x\tfrac{(3n-1)}{(3n-3)}\right)^4 + \tfrac{x^4(5)(9n+1)}{(n-1)^3}\left(1 - x\tfrac{(405n^2+180n-17)}{45(9n+1)(n-1)}\right)^2 \\ &\qquad + \tfrac{x^6(5^2 3^7 n^3 + 5^3 3^5 13n^2 + 5(3)^3 131n + 41)}{5(3)^4(9n+1)(n-1)^5} \quad > \quad 0 \end{aligned}$$

It can be seen that each expansion yields sum of perfect squares plus a remainder term. For any j, the number of terms in the expansion will always be odd: thus successive subtraction of the binomial expansion of the perfect square defined by the first two terms will always lead to a single remainder term. However, it could not be proved that the remainder term is always positive.

4.13.4 Expressing f_n as a ${}_2F_0$ hypergeometric function

The function $f_n(x,j)$ may be expressed as a ${}_2F_0$ hypergeometric function. A ${}_2F_0$ hypergeometric function is defined

$${}_2F_0(\alpha,\beta;z) \equiv \sum_m (\alpha)_m(\beta)_m \tfrac{z^m}{m!}$$

where $(\alpha)_m$ is Pochhammer's symbol (see 2.4), and the sum is taken from $m = 0$ to the term where the series terminates. The series must terminate since it does not converge. See Erdélyi [12], chapter IV for details. Hence

$$\begin{aligned} f_n(x,j) &\equiv \sum_{k=0}^{2j} \binom{2j}{k} \frac{\Gamma(n-1+k)}{\Gamma(n-1)} \left(\tfrac{-x}{n-1}\right)^k \\ &= {}_2F_0\left(-2j, n-1; \tfrac{x}{n-1}\right) . \end{aligned}$$

Erdélyi gives a number of relations for the ${}_2F_0$ function. The most promising (a combination of two relations) expresses the ${}_2F_0$ function as a sum of confluent hypergeometric functions.

$${}_2F_0(\alpha,\beta;\tfrac{-1}{x}) = x^\alpha \frac{\Gamma(\beta-\alpha)}{\Gamma(\beta)} M(\alpha,\alpha-\beta+1;x) + x^\beta \frac{\Gamma(\alpha-\beta)}{\Gamma(\alpha)} M(\beta,\beta-\alpha+1;x)$$

Thus

$$\begin{aligned} f_n(x,j) &\equiv {}_2F_0\left(-2j, n-1; \tfrac{x}{n-1}\right) \\ &= \left(\tfrac{-(n-1)}{x}\right)^{n-1} (-2j)_{1-n}\, M\left(n-1, n+2j; \tfrac{-(n-1)}{x}\right) \\ &\quad + \left(\tfrac{-(n-1)}{x}\right)^{-2j} (n-1)_{2j}\, M\left(-2j, 2-n-2j; \tfrac{-(n-1)}{x}\right) \end{aligned}$$

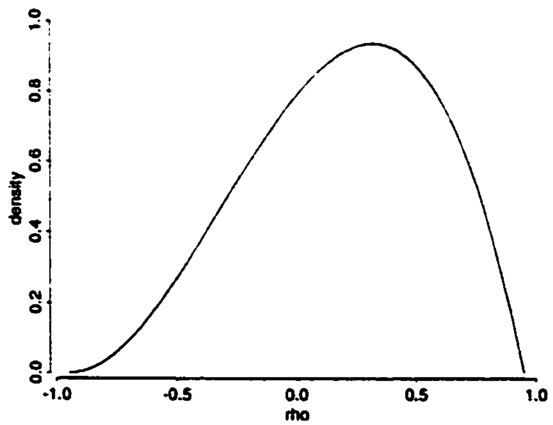
However, the first term can be negative when n is even, even when $\frac{-(n-1)}{x} > 0$. This is so because $(-2l)_{1-n}$ can be negative for some even ns. Thus the approach did not appear to lead to showing that $f_n > 0$.

4.14 Demonstration of the Theorem for random parameter values

2000 sets of random parameter values were generated as follows

n a random integer from 3 to 17;

ρ a Beta(3,2) random variable rescaled to the interval [-.95,.95]; this gave the following distribution



δ_ρ the absolute value of a N(2,1) random variable;

c the absolute value of ($\frac{\delta_\rho}{3}$ plus a N(1,1) random variable);

The parameter values were chosen so that in approximately half the instances, the conditions of the theorem were met, namely that $\rho \geq 0$ and $\beta \geq \frac{1}{2}$, and extreme values of the parameters were avoided.

For each of the 2000 cases, the power β and the derivative $\frac{d\beta}{d\rho}$ were computed using the distribution and derivative functions at (A.4) and (A.8). In eight instances, the calculation failed because the values of the parameters were too extreme. The remaining 1992 were distributed as follows

		$\beta \geq 0$ AND $\rho \geq 0$		
		True	False	Total
$\frac{d\beta}{d\rho} > 0$	True	796	453	1249
	False	0	743	743
	Total	796	1196	1992

It can be seen that in every case where the conditions were met, the conclusion was also true, thus demonstrating the theorem to be true for these 1992 cases.

Chapter 5

A TEST BASED ON $\bar{\vec{Y}}|\mathrm{V}$

As discussed in section 1.4.1, there are generally no sufficient statistics for each of the parameters of the bivariate normal distribution considered separately. However, the estimated variance-covariance matrix $\mathbf{v} \equiv \sum_{i=1}^{n} \vec{y}_i \vec{y}_i'$ is sufficient for the whole variance-covariance matrix $\Sigma = \sigma^2 \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$ under the null $\mu_1 = \mu_2 = 0$. Here, $\vec{y}_i \equiv \begin{pmatrix} y_{i1} \\ y_{i2} \end{pmatrix}$. This motivates the consideration of a test based on the distance between $\bar{y}_{\cdot 1}$ and $\bar{y}_{\cdot 2}$, conditioned on $\mathbf{v}$.

The derivation of the distribution is based on the derivation of $\bar{\vec{Y}}|\mathrm{V}$ for the general multivariate normal distribution given by Wang and McDermott in the Journal of the American Statistical Association, 1998 [42].

The data consist of n pairs from a bivariate normal distribution, and hence under the null the mean vector $\bar{\vec{Y}}$ distributes $\bar{\vec{Y}} \sim N_2(\vec{0}, \Sigma/n)$ with density function

$$f_{\bar{\vec{Y}}}(\bar{\vec{y}}) = \frac{n}{2\pi\sqrt{|\Sigma|}} \exp\left\{-\frac{n}{2}\left(\bar{\vec{y}}'\Sigma^{-1}\bar{\vec{y}}\right)\right\}.$$

Put $\mathrm{S} \equiv \mathrm{V} - n\bar{\vec{Y}}\bar{\vec{Y}}' = \sum_{i=1}^{n} \vec{Y}_i\vec{Y}_i' - n\bar{\vec{Y}}\bar{\vec{Y}}'$. Under the null, S has a Wishart$(\Sigma, n-1)$ distribution: Johnson & Kotz [22] give the density function

$$f_{\mathrm{S}}(\mathrm{s}) = \frac{|\mathrm{s}|^{(n-4)/2}}{4\pi|\Sigma|^{(n-1)/2}\,\Gamma(n-2)} \exp\left\{-\tfrac{1}{2}\mathrm{tr}(\Sigma^{-1}\mathrm{s})\right\}, \qquad \mathrm{s}\ \text{positive definite.}$$

Under the null $\bar{\vec{Y}}$ and S are independent, by the independence of sample means and variances in bivariate normal distributions; hence their joint density is the product of $f_{\bar{\vec{Y}}}(\bar{\vec{y}})$ and $f_{\mathrm{S}}(\mathrm{s})$ which simplifies to

$$f_{\bar{\vec{Y}},\mathrm{S}}(\bar{\vec{y}}, \mathrm{s}) = \frac{n\,|\mathrm{v} - n\bar{\vec{y}}\bar{\vec{y}}'|^{(n-4)/2}}{8\pi^2|\Sigma|^{n/2}\,\Gamma(n-2)} \exp\left\{-\frac{1}{2}\mathrm{tr}(\Sigma^{-1}\mathrm{v})\right\}$$

and also equals $f_{\bar{\vec{Y}},\mathrm{V}}(\bar{\vec{y}}, \mathrm{v})$ the joint density of $\bar{\vec{Y}}$ and V since the Jacobian of the transformation from $(\bar{\vec{Y}}, \mathrm{S})$ to $(\bar{\vec{Y}}, \mathrm{V})$ is 1.

The joint distribution of $\bar{\vec{Y}}$ and $\mathbf{V}$ is the product of the conditional distribution of $\bar{\vec{Y}}$ given $\mathbf{V}$ and the marginal distribution of $\mathbf{V}$. Since $\mathbf{V}$ is sufficient for $\mathbf{\Sigma}$ under the null, the conditional distribution of the data given $\mathbf{V}$ will be independent of $\mathbf{\Sigma}$; thus the conditional distribution of $\bar{\vec{Y}}$ will also be independent of $\mathbf{\Sigma}$ since $\bar{\vec{Y}}$ is a function of the data. The marginal distribution of $\mathbf{V}$ will be a function of $\mathbf{v}$ and $\mathbf{\Sigma}$. Thus the factorization of the joint density into the conditional and marginal can only be

$$\begin{aligned} f_{\bar{\vec{Y}},\mathbf{V}}(\bar{\vec{y}},\mathbf{v}) &= f_{\bar{\vec{Y}}|\mathbf{V}}(\bar{\vec{y}}|\mathbf{v}) \times f_{\mathbf{V}}(\mathbf{v}) \\ &= \frac{n}{8\pi^2\Gamma(n-2)}\left(|\mathbf{v}-n\bar{\vec{y}}\bar{\vec{y}}'|^{(n-4)/2} \times \frac{\exp\{-\frac{1}{2}\mathrm{tr}(\mathbf{\Sigma}^{-1}\mathbf{v})\}}{|\mathbf{\Sigma}|^{n/2}}\right) . \end{aligned}$$

Thus under the null

$$f_{\bar{\vec{Y}}|\mathbf{V}}(\bar{\vec{y}}|\mathbf{v}) = k_{\mathbf{v}}\,|\mathbf{v}-n\bar{\vec{y}}\bar{\vec{y}}'|^{(n-4)/2}$$

where $k_{\mathbf{v}}$ is the normalizing constant, the density being defined on the region where $\mathbf{s}$ is positive definite.

Large values of $|\bar{y}_{\cdot 1}-\bar{y}_{\cdot 2}|$ will tend to reject the hypothesis of equal means. The distribution is symmetric with respect to $\bar{y}_{\cdot 1}$ and $\bar{y}_{\cdot 2}$ and hence the p value is twice

$$\frac{\int_{\vec{u}:u_1-u_2\geq\bar{y}_{\cdot 1}-\bar{y}_{\cdot 2}} |\mathbf{v}-n\vec{u}\vec{u}'|^{(n-4)/2}\,d\vec{u}}{\int_{\vec{u}} |\mathbf{v}-n\vec{u}\vec{u}'|^{(n-4)/2}\,d\vec{u}} .$$

Two dimensional numerical integration was performed in SPLUS using the dcuhre algorithm developed by Genz et. al. [5]. The integrals extend to the boundary of the region where the density is defined, that is, the region where the density is greater than zero, (being the region where $\mathbf{s}$ is positive definite). The integration was facilitated by rotating the function and regions 45°. An illustration of the density in the region of integration for $n=7$ pairs, $\rho=.2345$ and randomly generated data. is shown in figure 5.1.

p values computed as above were compared with the p values from a paired t test for 100 data sets generated from a bivariate normal distribution with random parameter values as follows

n a random integer from $\{5,7,10,15,20\}$;

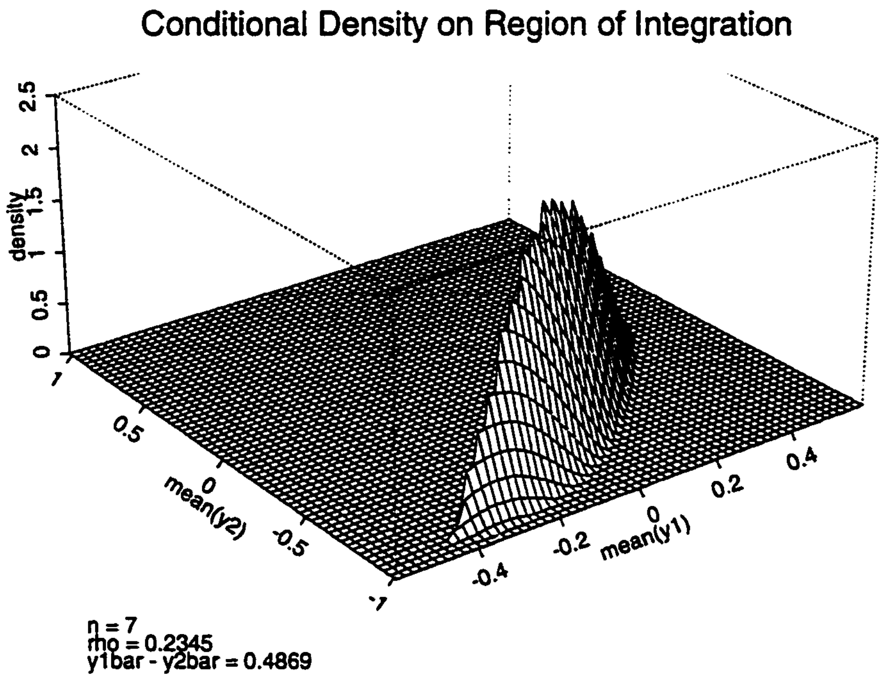


Figure 5.1: Illustration of the joint conditional density of $\bar{y}_{\cdot 1}$ and $\bar{y}_{\cdot 2}$ given $\mathbf{v}$ on the region of integration.

ρ a Uniform random variable on [-.9,.9];

$\mu_1 = 0$, μ_2 a $N(1,1)$ random variable, divided by $\sqrt{n}$.

None of the p values differed by more than 0.0007 between the two tests, indicating that the test based on $\bar{\vec{Y}}|\mathbf{V}$ was equivalent to a paired t test.

Chapter 6

TESTS BASED ON $T_2|\rho$

6.1 ρ known

If ρ were known, then the distribution of the two sample t statistic T_2 under the null would be fully specified, and unbiased tests may be constructed. The distribution function derived by Proschan [34] and in Appendix A may be used to compute critical values c_ρ under the null and power under the alternative.

The power achievable by such a test would exceed the power of both a two-sample test and a paired test when ρ was positive.

However, ρ cannot be known. Thus a value must be specified.

Specifying the value $\rho = 0$ reduces the test to an MU procedure. The distribution reduces to the distribution of a two-sample t statistic with bivariate normal data. The bias of the MU procedure is discussed in chapter 3 and the power illustrated in figure 1.2.

The MU procedure may be considered as a special case of a more general procedure, namely, testing the means of a bivariate normal distribution with a two sample t statistic and specifying the unknown parameter ρ at a value other than zero.

6.2 Specifying $\rho \neq 0$

Although the matching correlation ρ cannot be known, an investigator may believe under certain circumstances that it can be estimated with a high degree of precision. We therefore investigate the consequences of doing so. If the investigator specifies a value equal to the 'true' ρ, that is, if his guess is correct, the test will be unbiased. The question of interest then is, what is the consequence of mis-specifying ρ.

The bias, that is the type 1 error rate compared to the nominal size of the test, was computed using the distribution function at (A.4). First, for a given n and test size, a

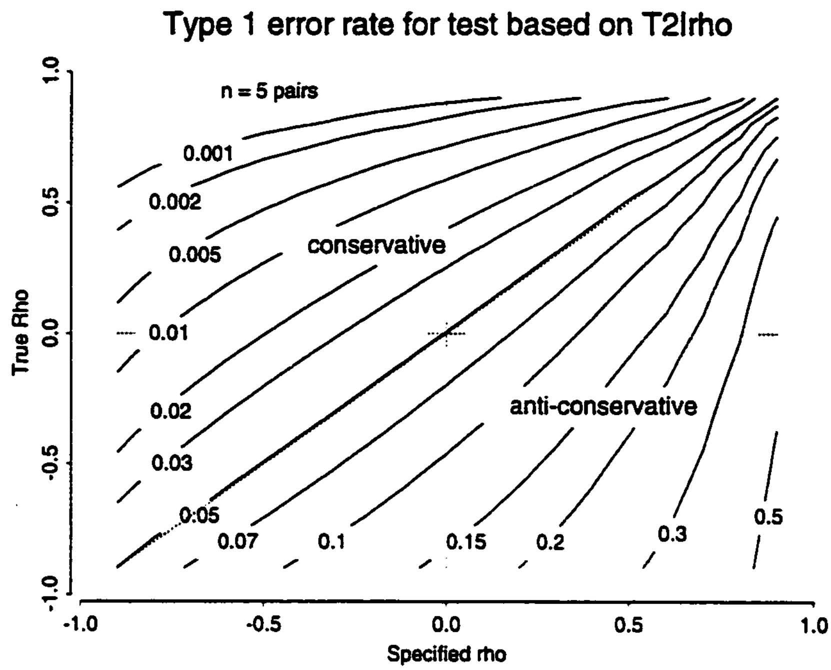


Figure 6.1: Bias of a test based on $T_2|\rho$, $n=5$ pairs, nominal size .05.

critical value is computed using the specified value of ρ. Then the probability of rejecting under the null is computed using the 'true' ρ. Figure 6.1 shows the result for a nominal size 0.05 test and $n=5$ pairs.

If the specified ρ is less than the 'true' ρ the test is biased conservative. Conversely, if the specified ρ is greater than the true ρ the bias is anti-conservative. This result may be compared with the bias of the MU procedure discussed in chapter 3. In effect, the MU procedure specifies ρ to be 0; if the true ρ was greater than 0, the test biased conservative; if less than 0, anti-conservative.

6.3 Specifying r for ρ

A brief evaluation was conducted of a test based on the distribution of $T_2|\rho$, using r as a 'plug-in' estimator of ρ for calculating critical values. For each of three values of the 'true' ρ, $\rho=-.5$, $\rho=0$, and $\rho=.5$, 5,000 datasets each comprising 5 pairs were generated by

Table 6.1: Bias of a test based on the distribution of $T_2|\rho$ using r as a 'plug-in' estimator of ρ, for various 'true' ρ, $n=5$ pairs, nominal size .05.

Type 1 error rate			
True $\rho =$	-0.5	0	0.5
Type 1 error rate	0.0774	0.0886	0.1012

simulation from a null bivariate normal distribution. For each dataset, Pearson's r was calculated, and the critical value found on the distribution of $T_2|\rho$ using r for ρ. Since the distribution function of $T_2|\rho$ at (A.4) converges only slowly for $|\rho|>.95$ the critical values in those ranges were calculated by extrapolation. If the two sample t statistic of the dataset was greater than the critical value, the test was taken to have rejected. The number of datasets rejecting as a proportion of the total was taken as an estimate of the type 1 error rate.

Table 6.1 shows the results. The type 1 error rate is greater than the nominal test size for all values of ρ, indicating the procedure would be biased anti-conservative.

Chapter 7

TESTS BASED ON $T_2|r$

In this chapter as in common practice r represents the Pearson product-moment sample correlation coefficient; in other chapters where it is necessary to distinguish between r and other estimators of ρ, the symbol r_1 is used. Recall also that T_2 represents the two sample t statistic; full details of notation are given in chapter 2.

As discussed in section 1.4.1, there are no sufficient statistics for the parameter ρ of the bivariate normal distribution, and since then r is not sufficient for ρ it cannot be expected that a test based on $T_2|r$ will be independent of ρ. Nevertheless, the intuitive appeal of a test based on $T_2|r$ - the anti-conservative bias and greater power of the two sample test for small positive values of ρ, the use of information about the effectiveness of matching - suggests such a test should be examined.

The distribution of $T_2|r$ under the null is derived in Appendix B. A simple form of the distribution function is

$$\begin{aligned} F_{T_2|r}(t|r) &= \frac{1}{2} + \frac{\Gamma\left(n-\frac{1}{2}\right)(1-\rho r)^{n-\frac{3}{2}}}{2\sqrt{2\pi}\,\Gamma^2(n-1)\,H\left(\frac{1}{2},\frac{1}{2},n-\frac{1}{2},\frac{1}{2}(1+\rho r)\right)} \\ &\quad \times \sum_{i,j=0}^{\infty} \frac{\Gamma\left(\frac{n+i-1}{2}\right)\Gamma\left(n+i+j-\frac{1}{2}\right)\Gamma\left(j+\frac{1}{2}\right)}{\Gamma\left(\frac{n+i}{2}\right)\Gamma\left(j+\frac{3}{2}\right)\Gamma(i+1)\Gamma(j+1)} \left(\rho r\right)^i \left(\frac{t\sqrt{1+\rho}}{\sqrt{2n-2}}\right) \left(\frac{-t^2(1+\rho)}{2n-2}\right)^j \end{aligned}$$

It can be seen that the distribution is not independent of ρ. In order for a critical value to be computed for a test, a value for ρ must be specified.

An intuitive choice for ρ is zero. Just as the MU procedure uses $\rho=0$ to compute a critical value for a test based on the distribution of $T_2|\rho$, so $\rho=0$ may be a starting point for a test based on $T_2|r$, or $T_2|r,\rho$ as it might more properly be described.

However, as can be seen from examination of the distribution function, when $\rho=0$ the distribution is independent of r, and hence is the same as the unconditional distribution, that is the two sample t distribution. The critical value would then be chosen with reference to that distribution, and the test would be the same as an MU procedure, that is, a two sample t test of matched data, with bias as shown in table 3.1 and power as illustrated in figure 1.2.

7.1 Generalizing to $\rho \neq 0$.

As with the MU procedure in the previous chapter, a more general test may be derived by specifying some other value for ρ, based on an estimated or assumed value, on which the computation of the critical value is based. Again, the question arises as to the consequences of mis-specifying ρ. The procedure is similar to that specified in 6.2. except that r-specific critical values are computed in each case.

Since r is a random function of the data. the procedure was evaluated using simulated datasets. For each of four values of 'true' ρ, $\rho=0$, $\rho=.25$, $\rho=.5$. and $\rho=.75$, 1000 datasets were generated from the null bivariate normal distribution. each consisting of 5 pairs. The two sample t statistic and the r statistic were computed for each dataset. Then, for each of four values of the specified ρ, the same four values, and for the r for each dataset, a critical value on the distribution of $T_2|r,\rho$ was computed and compared to the t statistic for the dataset.

The results are shown in figure 7.1. When $\rho=0$ was specified. the bias was, as expected, close to the bias of the procedure based on $T_2|\rho$ (at 6.2) and the MU procedure.

When other values of ρ were specified, the bias was in the same direction, but less than, the bias of the test based on $T_2|\rho$ at 6.2. If the ρ specified is less than the 'true' ρ the test is biased conservative; if the specified ρ is greater than the true ρ the bias is anti-conservative. The effect of conditioning on r was to reduce the bias of the procedure. In particular, the anti-conservative bias was reduced when ρ was over-specified. The effect may be noted by comparing the type 1 error rates on figures 7.1 and 6.1 for given values of true and specified ρ.

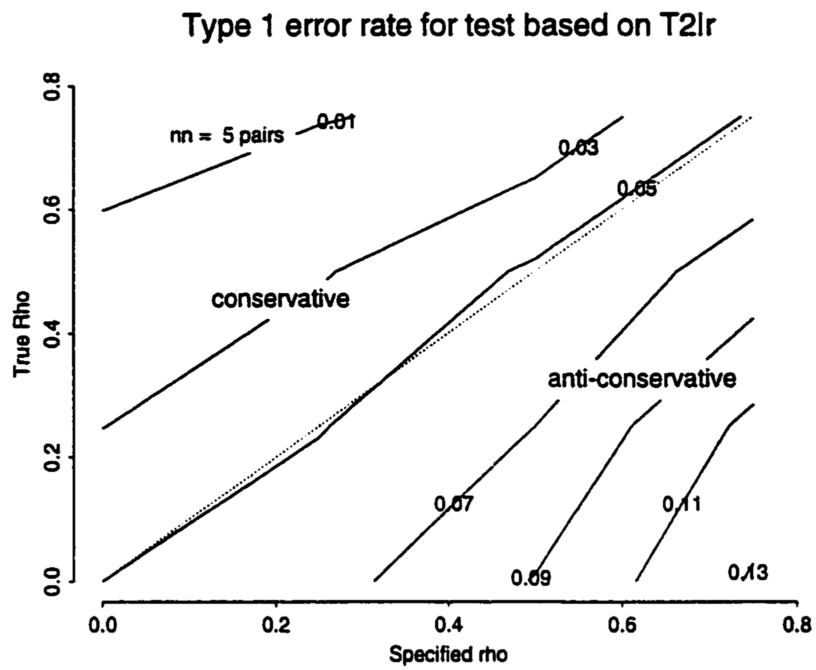


Figure 7.1: Bias of a test based on $T_2|r$, $n=5$ pairs, nominal size .05.

7.2 Specifying r for ρ

A brief evaluation was conducted of a test based on the distribution of $T_2|r, \rho$, using r as a 'plug-in' estimator of ρ. For each of three values of the 'true' ρ, $\rho = -.5$, $\rho = 0$, and $\rho = .5$, 5,000 datasets were generated by simulation, each consisting of 5 pairs from a null bivariate normal distribution. For each dataset, Pearson's r and the two-sample t statistic were calculated, and the t statistic was compared with the 0.05 size critical value on the distribution of $T|r, \rho$, using r for both r and ρ. Because the distribution function for $T|r, \rho$ at B.2 converges only slowly for $|r| > .95$, critical values in those ranges were calculated by extrapolation.

Table 7.1 shows the rejection rate under the null for a nominal size 0.05 test. For each value of ρ, the rate of rejection exceeded the nominal test size.

The results may be compared with the bias of a similar unconditional 'plug-in' test based on the distribution of $T_2|\rho$ at table 6.1.

Table 7.1: Bias of a test based on the distribution of $T_2|r, \rho$ using r as a 'plug-in' estimator of ρ, for various 'true' ρ, $n=5$ pairs, nominal size .05.

Type 1 error rate			
True $\rho =$	-0.5	0	0.5
Type 1 error rate	0.0796	0.0876	0.1064

Chapter 8

OTHER CONDITIONAL TESTS

8.1 Test based on $T_2|r_2$

There are three maximum likelihood estimators (MLEs) of ρ in the bivariate normal distribution, depending on the assumptions made. Pearson's r (sometimes denoted r_1) is the MLE of ρ when no assumption is made about the equality of means and variances. If the variances are assumed equal, but not the means, the maximum likelihood estimator of ρ is

$$r_2 = \frac{\sum_{i=1}^{n}(y_{i1}-\bar{y}_{\cdot 1})(y_{2i}-\bar{y}_{\cdot 2})}{\frac{1}{2}\left(\sum_{i=1}^{n}(y_{i1}-\bar{y}_{\cdot 1})^2+\sum_{i=1}^{n}(y_{i2}-\bar{y}_{\cdot 2})^2\right)} .$$

r_2 differs from r_1 in that the denominator is the arithmetic mean of the sums of squares; for r_1 the denominator is the geometric mean.

In order to assess whether investigating a test based on $T_2|r_2$ would be worthwhile, the distribution of $T_2|r_2$ was first estimated by simulation. For each of three values of ρ, $\rho=0$, $\rho=.2$ and $\rho=.5$, 20,000 datasets each consisting of 5 pairs were generated from a null bivariate normal distribution. For each dataset, r_2 was computed. The values of r_2 were grouped into intervals of width .1, and the .975 quantiles of the T_2 statistics for the corresponding datasets in each interval were computed, to estimate the .975 quantiles of the distribution of $T_2|r_2$.

Figure 8.1 shows the results. For comparison purposes, the .975 quantiles of $T_2|r_1$ for the same values of ρ are also shown.

The distribution of $T_2|r_2$ clearly depends on ρ. Also, the estimated quantiles are close to the quantiles of $T_2|r_1$, and the similarity is greatest in the intervals with the highest number of datasets. In particular, the estimated quantiles for $\rho=0$ appears to be substantially flat in r_2. It was therefore concluded that a test based on the distribution of $T_2|r_2$ would have similar limitations to a test based on $T_2|r_1$, namely that it would not be independent of ρ;

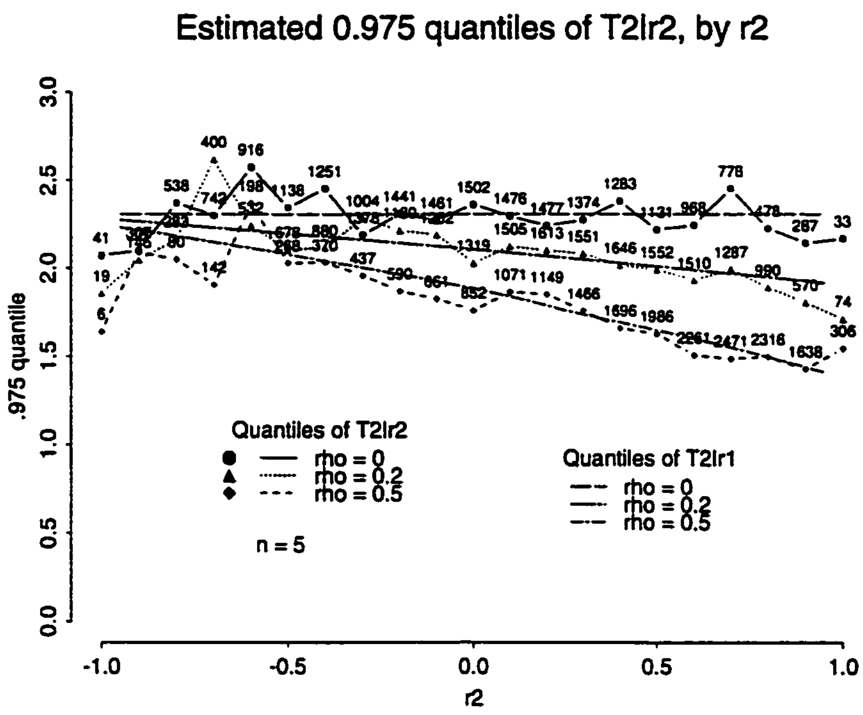


Figure 8.1: Estimated .975 quantiles of the distribution of $T_2|r_2$, $n=5$ pairs. Numbers above each point indicated the number of datasets used to estimate each quantile. Quantiles of $T_2|r_1$ are also shown.

thus an assumed or estimated value of ρ would have to be specified in order to calculate critical values. There are strong reasons for specifying $\rho=0$ under such circumstances, but it appears the distribution of $T_2|r_2$ will be independent of r_2 when $\rho=0$. Thus, the test would reduce to an unconditional two sample t test.

8.2 Test based on $T_2|r_3$

The maximum likelihood estimator of ρ in a bivariate normal distribution when the means and variances are both assumed equal is

$$r_3 = \frac{\sum_{i=1}^{n}(y_{i1}-\bar{y}_{..})(y_{i2}-\bar{y}_{..})}{\frac{1}{2}\left(\sum_{i=1}^{n}(y_{i1}-\bar{y}_{..})^2+\sum_{i=1}^{n}(y_{i2}-\bar{y}_{..})^2\right)}.$$

r_3 differs from r_2 in that the overall mean $\bar{y}_{..}$ replaces $\bar{y}_{.1}$ and $\bar{y}_{.2}$ each place they occur.

It was found that a test based on $T_2|r_3$ is equivalent to a paired t test; when one rejects, the other rejects. That this is so may be seen from the following sequence of lemmas.

Lemma 1

$$r_3 = \frac{2\,s_{12} - \frac{n\Delta^2}{2}}{s_1^2 + s_2^2 + \frac{n\Delta^2}{2}}$$

where s_1^2, s_2^2 and s_{12} are as defined at section 2.2 and $\Delta = \bar{y}_{\cdot 2} - \bar{y}_{\cdot 1}$.

Proof

$$\begin{aligned}
r_3 &= \frac{\sum_{i=1}^n (y_{i1} - \bar{y}_{\cdot\cdot})(y_{i2} - \bar{y}_{\cdot\cdot})}{\frac{1}{2}\left(\sum_{i=1}^n (y_{i1} - \bar{y}_{\cdot\cdot})^2 + \sum_{i=1}^n (y_{i2} - \bar{y}_{\cdot\cdot})^2\right)} \\
&= \frac{\sum_{i=1}^n (y_{i1} - \frac{\bar{y}_{\cdot 1}+\bar{y}_{\cdot 2}}{2})(y_{i2} - \frac{\bar{y}_{\cdot 1}+\bar{y}_{\cdot 2}}{2})}{\frac{1}{2}\left(\sum_{i=1}^n (y_{i1} - \frac{\bar{y}_{\cdot 1}+\bar{y}_{\cdot 2}}{2})^2 + \sum_{i=1}^n (y_{i2} - \frac{\bar{y}_{\cdot 1}+\bar{y}_{\cdot 2}}{2})^2\right)} \\
&= \frac{\sum_{i=1}^n (y_{i1} - \bar{y}_{\cdot 1} - \frac{\bar{y}_{\cdot 2}-\bar{y}_{\cdot 1}}{2})(y_{2i} - \bar{y}_{\cdot 2} + \frac{\bar{y}_{\cdot 2}-\bar{y}_{\cdot 1}}{2})}{\frac{1}{2}\left(\sum_{i=1}^n (y_{i1} - \bar{y}_{\cdot 1} - \frac{\bar{y}_{\cdot 2}-\bar{y}_{\cdot 1}}{2})^2 + \sum_{i=1}^n (y_{i2} - \bar{y}_{\cdot 2} + \frac{\bar{y}_{\cdot 2}-\bar{y}_{\cdot 1}}{2})^2\right)} \\
&= \frac{s_{12} - \frac{n\Delta^2}{4}}{\frac{1}{2}\left(s_1^2 + \frac{n\Delta^2}{4} + s_2^2 + \frac{n\Delta^2}{4}\right)} \\
&= \frac{2\,s_{12} - \frac{n\Delta^2}{2}}{s_1^2 + s_2^2 + \frac{n\Delta^2}{2}} \qquad \square
\end{aligned}$$

Lemma 2

$$t_p = \frac{t_2}{\sqrt{1 - r_3 - \frac{t_2^2(r_3+1)}{2(n-1)}}}$$

Proof

$$\begin{aligned}
t_p &= \frac{\sqrt{n}\Delta}{\sqrt{\frac{1}{n-1}\left(s_1^2 + s_2^2 - 2\,s_{12}\right)}} \\
&= \frac{t_2}{\sqrt{\frac{s_1^2+s_2^2-2\,s_{12}}{s_1^2+s_2^2}}} \qquad \left(\text{since } t_2 = \frac{\sqrt{n}\Delta}{\sqrt{\frac{1}{n-1}\left(s_1^2 + s_2^2\right)}}\right) \\
&= \frac{t_2}{\sqrt{1 - \frac{2\,s_{12}}{s_1^2+s_2^2}}} \\
&= \frac{t_2}{\sqrt{1 - \frac{1}{s_1^2+s_2^2}\left(r_3\left(s_1^2 + s_2^2 + \frac{n\Delta^2}{2}\right) + \frac{n\Delta^2}{2}\right)}} \quad \text{by Lemma 1,}
\end{aligned}$$

$$= \frac{t_2}{\sqrt{1 - r_3 - \frac{n\Delta^2(r_3+1)}{2(s_1^2+s_2^2)}}}$$

$$= \frac{t_2}{\sqrt{1 - r_3 - \frac{t_2^2(r_3+1)}{2(n-1)}}}$$ □

Lemma 3

For a fixed r_3, the mapping given by Lemma 2 is 1:1.

Proof

Using Lemma 2, write

$$t_p = \begin{cases} 0 & \text{when } t_2 = 0 \\ \frac{\text{sign}(t_2)}{\sqrt{\frac{1-r_3}{t_2^2} - \frac{r_3+1}{2(n-1)}}} & \text{else.} \end{cases}$$

It can be seen that the expression is monotonic in t_2. □

Lemma 4

$$P(T_2 \leq t_2 \,|\, r_3) = P(T_p \leq t_p)$$

Proof

The equality of Lemma 2 may be restated for the random variables

$$T_p = \frac{T_2}{\sqrt{1 - R_3 - \frac{T_2^2(R_3+1)}{2(n-1)}}} .$$

and so for a given value of R_3, say r_3,

$$T_p = \frac{T_2}{\sqrt{1 - r_3 - \frac{T_2^2(r_3+1)}{2(n-1)}}} .$$

Since the mapping is 1:1, whenever T_2 takes a certain value then T_p has a certain value. Suppose T_2 takes the value t_2; then T_p must take the value $\frac{t_2}{\sqrt{1-r_3-\frac{t_2^2(r_3+1)}{2(n-1)}}}$. But by Lemma 2, $\frac{t_2}{\sqrt{1-r_3-\frac{t_2^2(r_3+1)}{2(n-1)}}} = t_p$; hence

whenever, given r_3, $\quad T_2 = t_2 \quad$ then $\quad T_p = t_p$,

whenever, given r_3, $\quad T_2 \leq t_2 \quad$ then $\quad T_p \leq t_p$,

$$\text{so} \qquad P(T_2 \leq t_2 \,|\, r_3) \quad = \quad P(T_p \leq t_p) \, . \qquad \square$$

Thus a test based on the distribution of $T_2|r_3$ is equivalent to a paired t test.

Note that for $\frac{t_2}{\sqrt{1-r_3-\frac{t_2^2(r_3+1)}{2(n-1)}}}$ to be defined, $\left(1 - r_3 - \frac{t_2^2(r_3+1)}{2(n-1)}\right)$ must be positive

$$\begin{aligned} 1 - r_3 - \frac{t_2^2(r_3+1)}{2(n-1)} &> 0 \\ r_3\left(1 + \frac{t_2^2}{2(n-1)}\right) &< 1 - \frac{t_2^2}{2(n-1)} \\ r_3 &< \frac{2(n-1) - t_2^2}{2(n-1) + t_2^2} \end{aligned}$$

yielding an upper bound on r_3.

Chapter 9

KENWARD AND ROGER'S METHOD

In *Biometrics* in 1997 [24], Michael Kenward and James Roger propounded a method for estimating parameters of Gaussian linear models with small sample sizes. The method has gained wide acceptance and has been implemented as an option for the computation of denominator degrees of freedom in the SAS Proc Mixed procedure. We evaluated the method as it applies to our problem.

It was hoped the method would provide a smooth transition between the two sample test and the paired test as ρ moved away from 0. As discussed in chapter 1, when $\rho=0$ the variables of a bivariate normal distribution are independent. and the optimal test is then the two sample test with $2n-2$ degrees of freedom; when ρ is slightly different from zero, the paired test with $n-1$ degrees of freedom is optimal in the sense of being UMP unbiased and UMP invariant, but is less powerful than the two sample test for finite n because of loss of degrees of freedom.

9.1 *Background*

The Gaussian linear model for observations $\vec{Y}$ takes the form

$$\vec{Y} \sim N(\mathbf{X}\vec{\beta}, \mathbf{\Sigma}),$$

where $\mathbf{X}$ is a matrix of known covariates, $\vec{\beta}$ is a vector of unknown effects, and $\mathbf{\Sigma}$ is an unknown but usually structured covariance matrix. (We use boldface to denote matrices and $\vec{\cdot}$ to denote vectors).

Restricted Maximum Likelihood REML is well established as a method for estimating the parameters of a Gaussian linear model with structured covariance matrix. In particular, REML has desirable properties for the estimation of the parameters of the covariance matrix $\mathbf{\Sigma}$.

Usually in a Gaussian linear model, the objective is to make inference about the fixed effects $\vec{\beta}$, and the parameters of $\mathbf{\Sigma}$ are nuisance parameters. In making inference about $\vec{\beta}$, it may readily be shown that the variance of the estimator $\hat{\vec{\beta}}$ is $\Phi = \{\mathbf{X}^T\mathbf{\Sigma}^{-1}\mathbf{X}\}^{-1}$. However, since $\mathbf{\Sigma}$ is unknown, the REML estimate of $\mathbf{\Sigma}$, denoted $\hat{\mathbf{\Sigma}}$ is conventionally used in the computation. That is, for inference concerning $\vec{\beta}$, it is assumed that

$$\text{Var}\hat{\vec{\beta}} = \left\{\mathbf{X}^T\hat{\mathbf{\Sigma}}^{-1}\mathbf{X}\right\}^{-1}$$

This approach takes no account of the variability of the estimate $\hat{\mathbf{\Sigma}}$.

Kenward and Roger's method brings together the results of a number of papers by Harville [19] and others which address the problem of estimating the small sample variance of $\hat{\vec{\beta}}$. The problem is, firstly, that the estimate $\left\{\mathbf{X}^T\hat{\mathbf{\Sigma}}^{-1}\mathbf{X}\right\}^{-1}$ takes no account of the variability of $\hat{\mathbf{\Sigma}}$, which is of particular importance in small samples. Secondly, it is biased as an estimator of the true variance. By Taylor series expansion about the true value of $\mathbf{\Sigma}$ an adjusted estimator $\hat{\Phi}_A$ of the variance of $\hat{\vec{\beta}}$ in small samples is derived.

The adjusted estimator $\hat{\Phi}_A$ is then used in the test statistic. Suppose inference is to be made about a linear combination $\vec{L}'\vec{\beta}$ of the elements of $\vec{\beta}$. Then an adjusted Wald statistic

$$F(\vec{y}) = (\hat{\beta}-\beta)'L\left(L'\hat{\Phi}_A L\right)^{-1}L'(\hat{\beta}-\beta)$$

is formed using the adjusted estimate of the variance of $\hat{\beta}$, namely $\hat{\Phi}_A$. The distribution of this test statistic may approximated by a scaled $\mathcal{F}$ distribution, (compared with an exact, unscaled $\mathcal{F}$ distribution when n is large). The numerator degrees of freedom is 1; the denominator degrees of freedom and the scaling factor are to be derived.

A general form for the computation of the scaling factor and the denominator degrees of freedom takes a first order Taylor series expansion of $\left(L'\hat{\Phi}_A L\right)^{-1}$ and equates the expectation and variance of the thus expanded adjusted Wald statistic to the mean and variance of an $\mathcal{F}$ distribution.

In order for the method to yield the correct scaling factor and denominator degrees of freedom in cases where the distribution of the test statistic is exact and known, further terms of the Taylor series expansion, and thus modified estimates of the expectation and variance, are used. The formulæ are given in the paper.

The method using the first order Taylor series expansion of $\left(L'\hat{\Phi}_A L\right)^{-1}$ is referred to as the 'standard method'; the method using further terms of the Taylor series expansion is referred to as the 'modified method'.

9.2 Application to our problem

In our case, for n pairs, we have

$$\begin{aligned}
\mathbf{X} &\equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ \vdots & \vdots \\ 1 & 0 \\ 0 & 1 \end{pmatrix}_{2n\times 2} \\
\vec{\beta} &\equiv \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \\
\mathbf{\Sigma} &\equiv \mathbb{I}_n \otimes \sigma^2 \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}
\end{aligned}$$

where

$\mathbb{I}_K$ represents a $K \times K$ identity matrix;

$\otimes$ represents the Kronecker product of two matrices.

The estimators of the fixed effects and the REML estimators of the parameters of the covariance matrix are

$$\begin{aligned}
\begin{pmatrix} \hat{\mu}_1 \\ \hat{\mu}_2 \end{pmatrix} &= \begin{pmatrix} \bar{y}_{\cdot 1} \\ \bar{y}_{\cdot 2} \end{pmatrix} \\
\widehat{\sigma^2} &= \frac{1}{2(n-1)} \sum_{i=1}^{n} \left[(y_{i1}-\bar{y}_{\cdot 1})^2 + (y_{i2}-\bar{y}_{\cdot 2})^2\right] \\
\hat{\rho} &= \frac{\sum_i (y_{i1}-\bar{y}_{\cdot 1})(y_{i2}-\bar{y}_{\cdot 2})}{\frac{1}{2}\sum_i \left[(y_{i1}-\bar{y}_{\cdot 1})^2 + (y_{i2}-\bar{y}_{\cdot 2})^2\right]} .
\end{aligned}$$

(Recall this estimator of ρ is the MLE when the variances are assumed equal, a statistic we have referred to elsewhere as r_2).

Using their method, the adjusted estimator of the variance of the fixed effects is

$$\hat{\Phi}_A = \frac{\widehat{\sigma^2}}{n} \begin{bmatrix} 1 & \hat{\rho}^* \\ \hat{\rho}^* & 1 \end{bmatrix}$$

where

$$\hat{\rho}^* = \hat{\rho}\left(\frac{n-2+\hat{\rho}^2}{n-1}\right).$$

Note that when n or $|\hat{\rho}|$ is large, $\hat{\rho}^* \approx \hat{\rho}$; as $|\hat{\rho}|$ gets smaller, $\hat{\rho}^*$ approaches $\hat{\rho}\left(\frac{n-2}{n-1}\right)$.

In our case

$$L = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \text{ and } L'\beta = \mu_1 - \mu_2$$

and the (unscaled) test statistic is

$$F(\vec{y}) = \frac{n\left((\bar{y}_{\cdot 1} - \bar{y}_{\cdot 2}) - (\mu_1 - \mu_2)\right)^2}{2\widehat{\sigma^2}(1-\hat{\rho}^*)}$$

whose distribution can be approximated by an $\mathcal{F}$ distribution with 1 numerator degrees of freedom and denominator degrees of freedom to be determined.

Under the null hypothesis of $\mu_1 = \mu_2$, the square root of the test statistic is

$$\begin{aligned} t(\vec{y}) &= \frac{\sqrt{n}\,(\bar{y}_{\cdot 1} - \bar{y}_{\cdot 2})}{\sqrt{2\widehat{\sigma^2}(1-\hat{\rho}')}} \\ &= \frac{\sqrt{n}\,(\bar{y}_{\cdot 1} - \bar{y}_{\cdot 2})}{\sqrt{\frac{1}{n-1}\left(s_1 + s_2 - 2s_{12}\left(1 - \frac{1-\hat{\rho}^2}{n-1}\right)\right)}} \end{aligned} \tag{9.1}$$

(with s_1, s_2, s_{12} having their usual meaning, see chapter 2).

It may be seen that when $\hat{\rho} = \pm 1$ or when n is large, this is the paired t statistic; when $\hat{\rho} = 0$ and $n = 2$ it is the two-sample t statistic.

9.2.1 *Application of the modified method*

It would appear that, at least for certain values of the parameters, our problem would fall into the class of cases where 'the distribution of the test statistic is exact and known'. We therefore first evaluated the modified method.

The modified method uses higher order terms of the Taylor series expansion to derive the scaling factor and the denominator degrees of freedom. Applying the method gives denominator degrees of freedom for our problem of $n-1$ and a scaling factor of 1. Thus the statistic $t(\vec{y})$ at (9.1) is to be tested against a t distribution with $n-1$ degrees of freedom,

Table 9.1: Bias of a test based on Kenward & Roger's modified method, for various ρ, $n=5$ pairs, nominal size .05.

Bias of Kenward & Roger's modified method						
ρ	-0.5	-0.25	0	0.25	0.5	0.75
Type 1 error rate	0.0445	0.0405	0.033	0.041	0.0285	0.029

without scaling (the t_ν distribution being the distribution of the square root of a random variable having an $\mathcal{F}_{1,\nu}$ distribution).

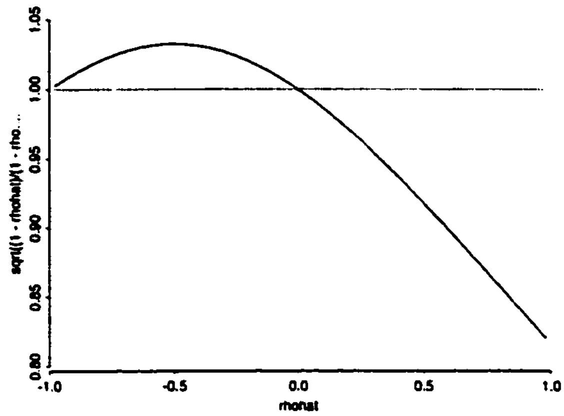
When $s_{12}>0$, $t(\bar{y})$ will be slightly smaller than the paired t statistic; when $s_{12}<0$, it will be slightly larger. It is therefore to be expected that the test will have properties similar to a paired t test, with some anti-conservative bias when $\rho<0$, and conservative bias and decreased power when $\rho>0$.

The test was evaluated by simulation. First for the bias under the null, for each of six values of ρ, 2000 datasets were generated each with 5 pairs, and the test applied to each. The results are shown in table 9.1

The test is biased conservative for all values of ρ. Since the test is similar to a paired test, significant bias was not expected. The explanation appears to be that the proportionate reduction in the test statistic compared to the paired t statistic when $\rho>0$ is greater than the proportionate increase when $\rho<0$. It may be shown that

$$t(\bar{y}) = t_p\sqrt{\frac{1-\hat{\rho}}{1-\hat{\rho}\left(1-\frac{1-\hat{\rho}^2}{n-1}\right)}}.$$

and plotting $\sqrt{\frac{1-\hat{\rho}}{1-\hat{\rho}\left(1-\frac{1-\hat{\rho}^2}{n-1}\right)}}$ against $\hat{\rho}$ illustrates the reason.



The power was evaluated again by simulation and was close to but less than the paired test, as would be expected given the conservative bias. The power is shown in figure 9.1.

In general, the modified method performed as expected, giving results close to those for

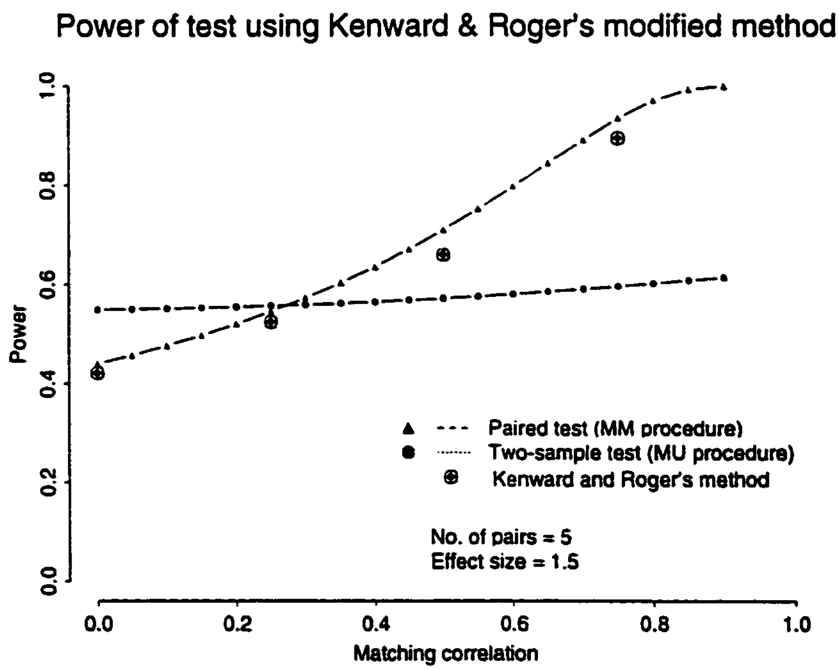


Figure 9.1: Power of a test based on Kenward & Roger's modified method, compared to the power of a two sample test and a paired test, for various ρ, $n = 5$ pairs, nominal size .05, effect size 1.5.

the paired test.

9.2.2 Application of the standard method

The standard method takes a first order Taylor series expansion of $\left(L'\hat{\Phi}_A L\right)^{-1}$ and equates the expectation and variance of the thus expanded adjusted Wald statistic to the mean and variance of an $\mathcal{F}$ distribution to derive an appropriate scaling factor and denominator degrees of freedom for the statistic.

Applying the method in our case gives, for the scaling factor (for a t statistic)

$$\lambda = \sqrt{\frac{(n-1)(3n^2+18n-25)}{(n+1)(3n^2+12n-11)}}$$

and for the denominator degrees of freedom

$$df = \frac{3n^2+18n-25}{3n-7} .$$

Thus the quantity $\lambda\, t(\bar{y})$ is to be compared to a t distribution with df degrees of freedom, (where the statistic $t(\bar{y})$ is as given at (9.1)).

For various values of n, the scaling factor λ and the degrees of freedom are

n	λ	df
2	0.5538	-23
3	0.7338	28
4	0.8189	19
5	0.8676	17.5
6	0.8985	17.36
7	0.9195	17.71
10	0.9540	19.78
20	0.9858	28.96
50	0.9974	58.57

For large n, the scaling factor approaches 1 and the degrees of freedom approaches $n+8\frac{1}{3}$. Since the test statistic $t(\bar{y})$ approaches the paired t statistic for large n, the method will approximate a paired test asymptotically.

We evaluated the method for $n=5$ pairs by simulation. The bias turned out to be small, and so the number of simulations was increased to 10,000 for each value of ρ, and a 95%

Table 9.2: Bias of a test based on Kenward & Roger's standard method, for various ρ, $n=5$ pairs, nominal size .05, with 95% C.I. for the type 1 error rate.

Bias of Kenward & Roger's standard method						
ρ	-0.5	-0.25	0	0.25	0.5	0.75
Type 1 error rate	0.0684	0.0708	0.0609	0.0562	0.0509	0.0474
95% C.I. - from	0.0635	0.0658	0.0562	0.0517	0.0466	0.0432
- to	0.0733	0.0758	0.0656	0.0607	0.0552	0.0516

confidence interval for the type 1 error rate calculated. Power was assessed based on 2,000 simulations, as for the modified method. The bias is shown in table 9.2. Power is shown in figure 9.2.

9.2.3 Discussion

The standard method has several desirable properties. The power is good; when $\rho=0$, the power is almost as high as the two sample test, and more powerful than the paired test for all positive values of ρ.

The drawback is the anti-conservative bias for the values of ρ most likely to be found in group randomized trials. The bias for negative ρ may be excused since many investigators may be willing to accept the argument that matching correlations are in general positive. The anti-conservative bias for $\rho \in [0, .25]$ is more of a limitation.

Reducing the degrees of freedom would correct the anti-conservative bias, and in this connection it is noted that the degrees of freedom in the table above do not decline monotonically with n, but (with the exception of the value for $n = 2$) are values of a convex function, increasing for values of n below 6.

The method is intended to be applicable to a wide variety of problems, and in that respect its performance when applied to this particular problem could not be criticized. It is noted that the two elements of the method, the scaling factor and the denominator degrees of freedom, go together. Using for example the computations for the degrees of

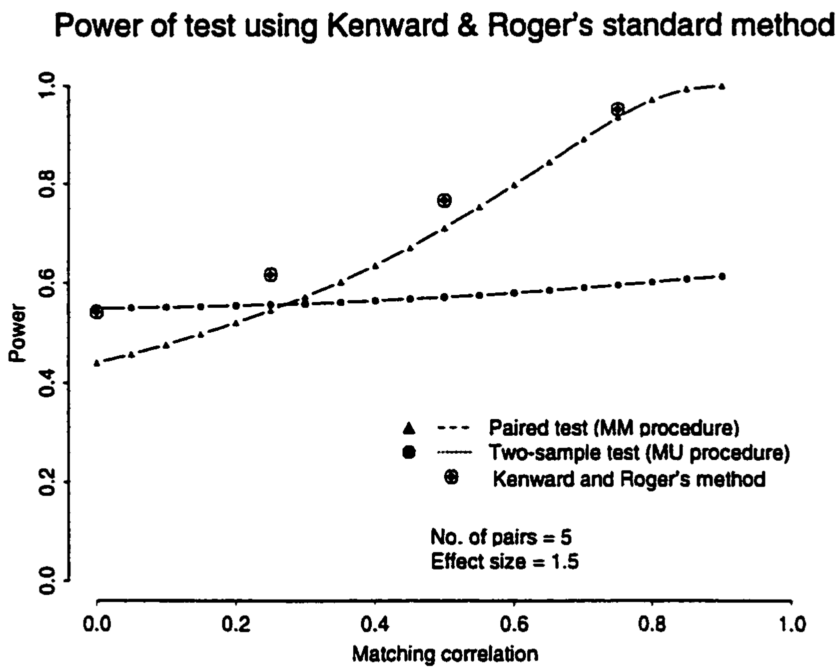


Figure 9.2: Power of a test based on Kenward & Roger's standard method, compared to the power of a two sample test and a paired test, for various ρ, $n = 5$ pairs, nominal size .05, effect size 1.5.

freedom without also applying the scaling factor to the statistic would appear to lead to biased results, in this case anti-conservative.

Chapter 10

FLEXIBLE CONDITIONAL TESTS

Notation: In this chapter the Pearson product-moment sample correlation coefficient is represented by r; elsewhere, where it is necessary to distinguish between the Pearson estimator and other estimators of ρ, r_1 is used.

10.1 Introduction

We noted in the introduction (section 1.4.1) that conditioning on the sufficient statistic for a nuisance parameter is a well recognized way of deriving tests. Often, the sufficient statistic is an estimator of the nuisance parameter. In our case, there is no sufficient statistic for ρ, and hence it cannot be expected that a conditional test based on an estimator of ρ will be independent of ρ; nevertheless, conditional tests were investigated in chapters 7 and 8 to determine the extent to which, after conditioning on the estimator of ρ, the conditional test remained dependent on ρ.

The MU procedure is unbiased when $\rho = 0$, (and the paired test is always unbiased). Hence, a reasonable criteria for evaluating other tests. such as conditional tests, is that they be unbiased when $\rho = 0$. Apart from any other arguments, constructing tests which were unbiased for values of ρ other than zero would involve an element of arbitrariness which would be hard to justify.

We considered three conditional tests based on the two sample statistic T_2. They conditioned on the three estimators r_1, r_2 and r_3. None of them yielded a useful test. The test using $T_2|r_3$ turned out to be equivalent to the paired test. The null distribution of $T_2|r_1$ was derived, but it turned out to be independent of r_1 when $\rho=0$. Thus the critical values which would make the test unbiased at $\rho=0$ were all the same, they did not vary with r_1 as they would in a conditional test. Hence, the test was the same as the unconditional test.

Similarly, it appeared from simulations that the distribution of $T_2|r_2$ was also independent of r_2 when $\rho=0$; thus again the test would reduce to an unconditional test.

The critical values of a conditional test are usually set by finding the quantiles of the conditional distributions that, for each value of the conditioning (given) statistic, correspond to a probability of $1-\frac{\alpha}{2}$. Thus in chapter 7, Tests based on $T_2|r$, the critical values $c_r \equiv c(r, \alpha)$ are the quantiles of the conditional distributions such that under the null

$$P(T_2 > c_r \,|\, r, \rho) \;=\; \frac{\alpha}{2} \quad \forall\, r\,. \tag{10.1}$$

This approach, though straightforward, is more stringent than necessary. In order to protect the test size, all that is necessary is that *on average*, taken over the distribution of the conditioning statistic, the type 1 error rate is no greater than the test size. That is, all that is required is that

$$\int_{r=-1}^{1} P(T_2 > c_r \,|\, r, \rho)\, f(r|\rho)\, dr \;\leq\; \frac{\alpha}{2} \tag{10.2}$$

This holds out the possibility of a conditional test conditioning on r that does not reduce to an unconditional test at $\rho = 0$. Instead of setting the critical values in the usual way according to the conditional distribution, as at (10.1), they may be set according to some other scheme provided that on average over the distribution of r the test is unbiased at $\rho=0$, that is, provided that (10.2) is satisfied.

The distinction will not make much difference if n is large, since then the distribution of $r|\rho$ will be closely gathered around ρ. Then, only those c_r for r close to ρ will have much impact on the characteristics of the test. But if n is small, the distribution of r will be more dispersed, and the c_r over a wider range of r will be material to the performance of the test.

Then, the critical values may be decreased for some values of r, thus increasing power, and increased elsewhere, thus decreasing power, so long as overall the size is protected; that is, so long as (10.2) is satisfied.

We showed in chapter 7 that conditioning on r reduced the bias of a test using T_2, compared to an unconditional test, when $\rho \neq 0$. It is hoped that bias will similarly be reduced when the conditional critical values c_r are set according to some other scheme.

Sets of critical values are therefore sought, other than those set according to the r-specific conditional distributions, which yield a test unbiased at $\rho=0$ and with improved bias and

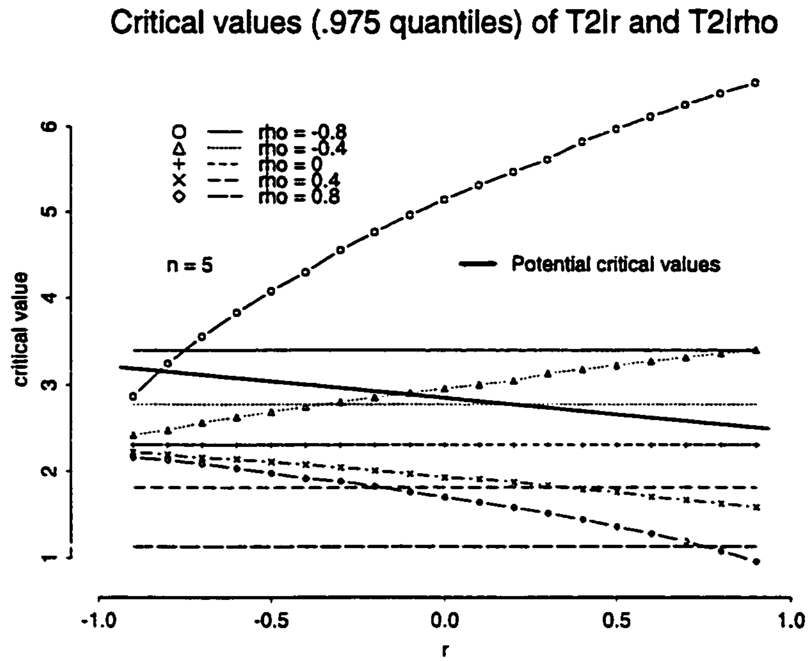


Figure 10.1: Critical values of tests based on $T_2|r$ and $T_2|\rho$, $n = 5$ pairs, nominal size .05. The potential critical values illustrate a set of c_r to be evaluated.

power characteristics for other values of ρ. In particular, sets of critical values which yield tests with characteristics appropriate for the analysis of group randomized trial as discussed in 1.4.1 are desired.

Any set of critical values c_r so determined will be n-specific, since the distribution of r will depend on n; however, that need not be a serious limitation since our interest is in small samples.

The approach may be illustrated by considering some sets of critical values in the $c-r$ space. Figure 10.1 shows sets for $n = 5$ pairs, nominal size .05. It shows for various ρ the sets of critical values for the test based on $T_2|\rho$ (chapter 6) and the test based on $T_2|r$ (chapter 7); and an illustrative possible set of linear critical values that might be evaluated as to whether or when they satisfy (10.2).

10.2 Choice of c_r

Section 1.4.1 discusses some of the characteristics that might be desirable in a test for group randomized trials, and some conditions that might be relaxed. In particular, complete unbiasedness for negative ρ and large positive ρ might not be considered essential. The question is, what set of c_r will achieve those objectives. What function (of r) should c_r be to give a test with those properties?

10.2.1 A theoretical approach

Suppose one has as a starting point a set of c_r satisfying (10.2). Consider two distinct points of the set; call them c_1 and c_2.

If the critical values are to be changed by moving c_1 and c_2 small distances, the change in probability at, say, c_1 will be the amount of change in c_1 times the conditional probability density at that point. Let Δ_1 and Δ_2 represent the change in c_1 and c_2 and $f_N(\cdot)$ represent the conditional probability density at each point under the null; then unbiasedness will be maintained if

$$\Delta_1 f_N(c_1) + \Delta_2 f_N(c_2) = 0$$

One of $\Delta_1 . \Delta_2$ will be positive and the other negative.
To achieve an increase in power under the alternative requires that

$$\Delta_1 f_A(c_1) + \Delta_2 f_A(c_2) > 0$$

where $f_A(\cdot)$ is the density under the alternative. Combining these two constraints gives (for Δ_1 positive)

$$\frac{f_A(c_1)}{f_N(c_1)} > \frac{f_A(c_2)}{f_N(c_2)} ,$$

suggesting that to find the critical values which should be increased or decreased to maintain unbiasedness while increasing power, one should seek the points c_r where $\frac{f_A(c_r)}{f_N(c_r)}$ is least and greatest.

An immediate practical difficulty is the derivation of $f_A(\cdot)$, that is $f_{T_2}(t_2|r;\rho,\delta)$ the density of T_2 given r under the alternate. The density under the null $f_{T_2}(t_2|r;\rho,\delta=0)$ is derived at Appendix B. Initial approaches to the distribution under the alternative suggest the derivation may not be straightforward, and the resulting function may not be computable at a reasonable speed.

The density under the alternative will depend on δ; hence $\frac{f_A}{f_N}$ may vary with δ. Thus there may not be a uniformly more powerful test; the best c_r may depend on δ, which would be a limitation.

The choice of a suitable function for $c_r = c(r, \alpha)$ is therefore approached from a more pragmatic direction. We will evaluate different functions for c_r, i) a linear function of r, ii) second and third order quadratic functions of r, iii) a smoothed step function, and iv) a flat/increasing function of r.

10.3 Linear function of r

Let c_r be a linear function of r

$$c_r = c_0(1 + \lambda r) .$$

Bringing together the requirement that the test be unbiased when $\rho = 0$, the constraint at (10.2) expressed in terms of the distribution function of $T_2|r$ (from Appendix B) and the density of r (see (B.3)) and the linear structure of c_r above, and integrating out r gives the following condition defining c_0 and λ

$$\frac{1}{2} - \frac{c_0}{2^{n-1}\Gamma\left(\frac{n}{2}\right)\sqrt{2(n-1)}} \sum_{j,k=0}^{\infty} \frac{\Gamma\left(n-\frac{1}{2}+j+k\right)\Gamma\left(j+k+\frac{1}{2}\right)}{\Gamma\left(\frac{n-1}{2}+k\right)\Gamma\left(j+\frac{3}{2}\right)\Gamma(j+1)\Gamma(k+1)} \left(\frac{-c_0^2}{2(n-1)}\right)^j \left(\frac{-c_0^2\lambda^2}{2(n-1)}\right)^k \leq \frac{\alpha}{2},$$

or in a computable form which will converge for all values of c_0 (provided $\lambda < 1$)

$$\begin{aligned}
&\frac{1}{2} - \frac{c_0}{2\,\Gamma(n-1)\sqrt{2(n-1)}} \left(\frac{2(n-1)}{2(n-1)+c_0^2}\right)^{n-\frac{1}{2}} \sum_{j=0}^{\infty} \frac{\Gamma\left(n-\frac{1}{2}+j\right)}{\Gamma\left(j+\frac{3}{2}\right)} \left(\frac{c_0^2}{c_0^2+2(n-1)}\right)^j \\
&- \frac{c_0}{2^{n-1}\Gamma\left(\frac{n}{2}\right)\sqrt{2(n-1)}} \left(\frac{2(n-1)}{2(n-1)+c_0^2}\right)^{n-\frac{1}{2}} \sum_{j,k=0}^{\infty} \\
&\frac{\Gamma\left(n+\frac{1}{2}+2j+k\right)\Gamma\left(k+j+\frac{3}{2}\right)\Gamma(k+j+1)}{\Gamma\left(\frac{n+1}{2}+j+k\right)\Gamma\left(j+\frac{3}{2}\right)\Gamma(j+1)\Gamma(k+j+2)\Gamma(k+1)} \left(\frac{c_0^2\lambda}{2(n-1)+c_0^2}\right)^{2j} \left(\frac{-c_0^2\lambda^2}{2(n-1)+c_0^2}\right)^{k+1} \\
&\leq \frac{\alpha}{2} . \qquad (10.3)
\end{aligned}$$

Solving (10.3) iteratively for c_0 with strict equality, $\alpha=.05$, $n=5$, and $\lambda=0, -.1, -.2, \cdots, -.7$, gave the following values for c_0

Values of λ and c_0 for linear c_r								
$\lambda =$	0	-0.1	-0.2	-0.3	-0.4	-0.5	-0.6	-0.7
$c_0 =$	2.306	2.3164	2.3484	2.4044	2.4889	2.6098	2.7803	3.0245
Referred to as	$c_{r(0)}$	$c_{r(1)}$	$c_{r(2)}$	$c_{r(3)}$	$c_{r(4)}$	$c_{r(5)}$	$c_{r(6)}$	$c_{r(7)}$

Notationally, the different sets of c_r are identified by $c_{r(\cdot)}$. Thus these parameters define eight sets of critical values, each set being a straight line with intercept c_0 at $r = 0$ and slope λ. Figure 10.2 illustrates them superimposed over the critical values of a conventional test based on $T_2|r$ and of a test based on $T_2|\rho$. The set $c_{r(0)}$ has zero slope and is thus an unconditional test, that is, a two sample test.

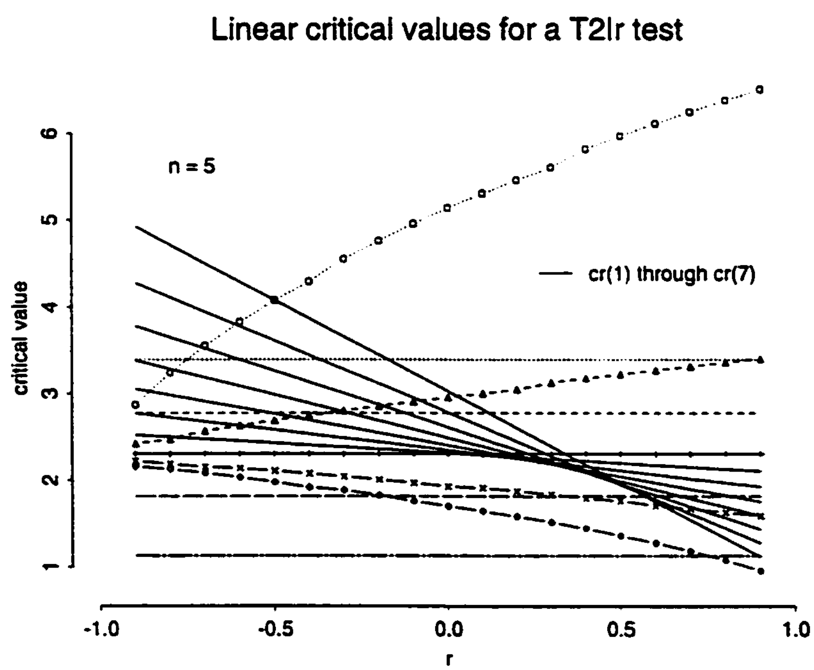


Figure 10.2: Eight sets of linear critical values for a test based on $T_2|r$, $n=5$ pairs, nominal size .05, unbiased at $\rho=0$. Critical values for a conventional test based on $T_2|r$ and a test based on $T_2|\rho$ are also shown.

10.3.1 Bias and Power of Tests with Linear Critical Values

The bias of five tests using $c_{r(0)}, c_{r(1)}, c_{r(3)}, c_{r(5)}$ and $c_{r(7)}$ was calculated for various ρ. The results are shown in table 10.1.

It can be seen that the bias decreases as the slope λ becomes steeper. The $c_{r(5)}$ critical values give a test with good bias. At $\rho=0$, the test is unbiased. For $\rho<0$ the bias is small anti-conservative. For $\rho>0$ the bias is conservative.

However, it was found that the power of a test using the $c_{r(5)}$ critical values was almost identical to a paired t test. As the bias is reduced, the power approaches that of the uniformly most powerful unbiased test. Power was evaluated by simulation of 2,000 datasets.

Also, as the slope increased over different $c_{r(\cdot)}$, the critical values at $r=0$ increased.

Therefore, a test using some other critical values. such as the $c_{r(3)}$ set, might be preferred. The power for $c_{r(3)}$ is illustrated in figure 10.3. The power is close to that of a two sample test when $\rho=0$ and better than a two sample test for ρ greater than approximately 0.15. Also, the test is unbiased when $\rho=0$. conservative when ρ is positive, and only slightly biased anti-conservative at $\rho=-.1$ (interpolated type 1 error rate = 0.054).

This $c_{r(3)}$ linear set of critical values will only be valid for $\alpha=.05$ and $n=5$; different values would be required for other test and sample sizes.

The conservativeness of the bias when ρ is positive suggests that more power may be

Table 10.1: Bias of tests based on $T_2|r$ and linear critical values, $n=5$ pairs, nominal size .05.

Type 1 error rate for five tests with linear critical values						
	$\rho=-.5$	$\rho=-.25$	$\rho=0$	$\rho=.25$	$\rho=.5$	$\rho=.75$
$c_{r(0)}$	0.1035	0.0746	0.05	0.0297	0.014	0.0038
$c_{r(1)}$	0.093	0.0706	0.05	0.0315	0.0158	0.0045
$c_{r(3)}$	0.0731	0.0625	0.05	0.0359	0.021	0.0072
$c_{r(5)}$	0.0551	0.0542	0.05	0.042	0.0296	0.0129
$c_{r(7)}$	0.0393	0.0458	0.05	0.0506	0.0451	0.028

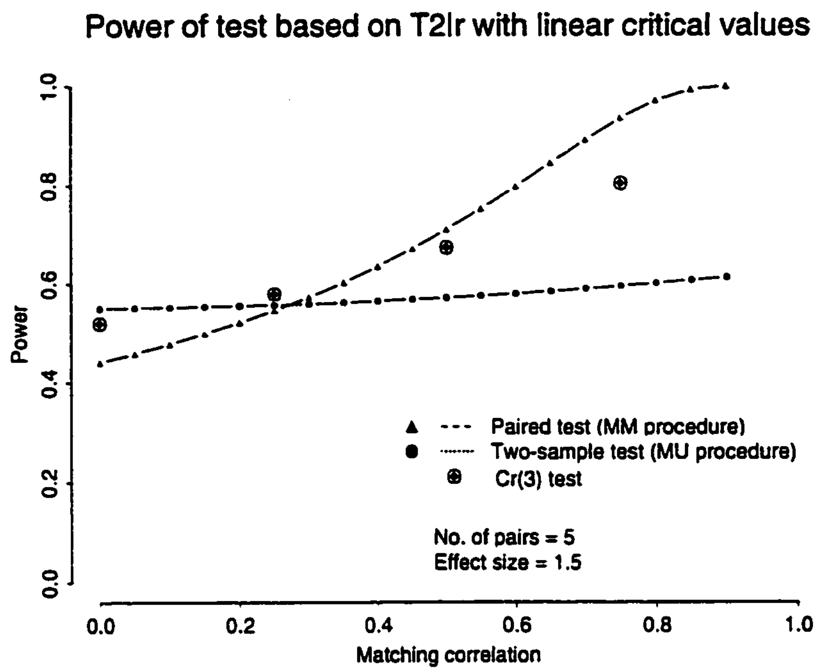


Figure 10.3: Power of a test using the $c_{r(3)}$ linear critical values, compared to the power of a two sample test and a paired test, for various ρ, $n=5$ pairs, nominal size .05, effect size 1.5.

obtained in this region, and there would seem to be no reason in principle why unbiasedness at $\rho = -.1$ as well as $\rho = 0$ should not be attainable with an appropriate function for the critical values c_r.

10.4 c_r as a second and third order power series in r

c_r as a second order power series in r is

$$c_r = a_0 + a_1 r + a_2 r^2.$$

The $a.$ were chosen by fitting the curve to three specified points in the $c-r$ space. It was found however that a second order curve provided insufficient flexibility for the curve of critical values. When the curve was convex, the critical values were larger than desired towards the extreme values of r; when the curve was concave, they were too small. The

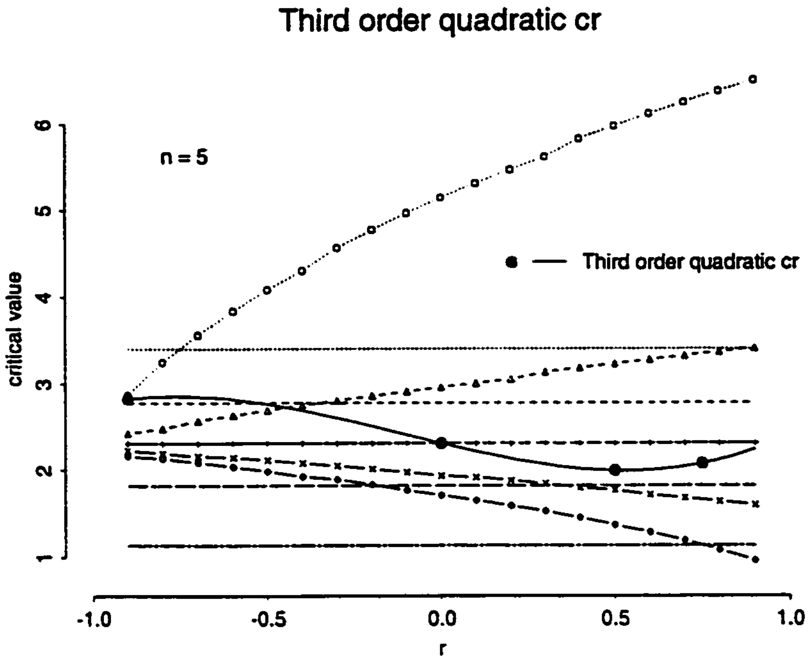


Figure 10.4: Critical values for a test based on $T_2|r$ using a third order quadratic function, $n=5$ pairs, nominal size .05. Critical values for a conventional test based on $T_2|r$ and a test based on $T_2|\rho$ are also shown.

second order curve did not conveniently allow the adjustments necessary to achieve the test properties sought.

10.4.1 c_r as a third order power series in r

c_r as a third order power series in r is

$$c_r = a_0 + a_1 r + a_2 r^2 + a_3 r^3.$$

Figure 10.4 illustrates one such curve, again superimposed over the critical values of a test based on $T_2|r$ and constraint (10.1), and the critical values for a test based on $T_2|\rho$. The curve was obtained by fitting to the points shown.

One of the desirable properties of a test for group randomized trials discussed in section 1.4.1 is that it not be anti-conservative in some interval below $\rho = 0$. The immediate objective then, since the tests with linear critical values in 10.3 are unbiased at $\rho=0$ but anti-conservative at $\rho<0$, is to find critical values which will yield an unbiased test at some point $\rho<0$; $\rho=-.1$ was selected as an arbitrary starting point.

A further consideration was how the test would compare to an unconditional test, that is, a two sample test. If the critical values were substantially greater than those of the two sample test, the test would not be appealing. The value of c_r at $r=0$ was therefore fixed to be the same as the critical value of an unconditional two sample test.

To remove anti-conservative bias in the region $\rho<0$, critical values in the region $r<0$ must be increased. To increase power, critical values must be reduced; and since the unconditional test is biased conservative in $\rho>0$, the region $r>0$ is the most promising place to do that.

A third order quadratic c_r was again found to be insufficiently flexible to achieve these properties. Because at the .975 quantile the density is decreasing rapidly, the anti-conservativeness engendered by a given decrease in c_r at, say, $r=.5$ requires a correspondingly much larger increase in c_r at $r=-.5$. When increasing critical values, one soon arrives at the point where the density is so low that further increase has little effect on the test.

10.5 Two flexible functions for c_r

10.5.1 Smoothed step function

The obvious way to define a function with a given value at $r=0$, larger in the region $r<0$ and smaller in $r>0$ is a step function with one step at $r=0$. Rather than a sharp step, a smoothed step is to be preferred. Such a function is shown in figure 10.5.

To optimize the smoothed step function in terms of the test properties discussed above, the critical values in $r<0$ are to be increased to try and eliminate the anti-conservative bias at $\rho<-.1$. Yet large increases render the test unattractive. To achieve critical values of 3 or 4 because the data gave an r of -.8 would render the test unappealing, particularly as r is a random quantity which may be markedly different from ρ. For the purposes of evaluating the test, r was set at 2.95 on the left side of the function (for an $n=5$ pairs, 0.05 size test). The right side of the step function was then set so that the test was unbiased at $\rho=-.1$.

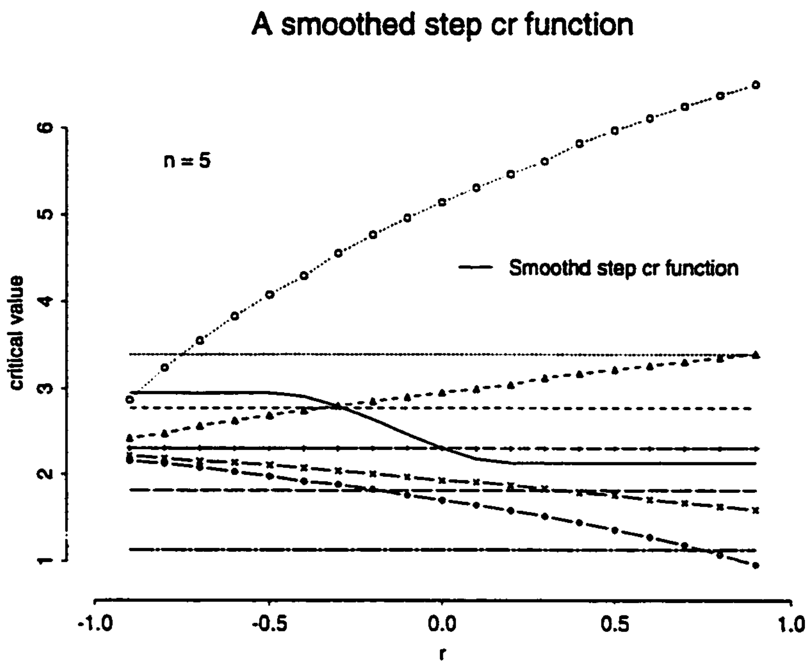


Figure 10.5: Critical values for a test based on $T_2|r$ using a smoothed step function; $n = 5$ pairs, nominal size .05. Critical values for a conventional test based on $T_2|r$ and a test based on $T_2|\rho$ are also shown.

The bias for other values of ρ is shown in table 10.2. and the power in figure 10.6.

Table 10.2: Bias of a test with smoothed step critical values, for various ρ, $n = 5$ pairs, nominal size .05.

	Type 1 error rate for a test based on $T_2\|r$ with smoothed step critical values					
$\rho=$	-.5	-.25	0	.25	.5	.75
	0.0694	0.058	0.0443	0.0295	0.0155	0.0045

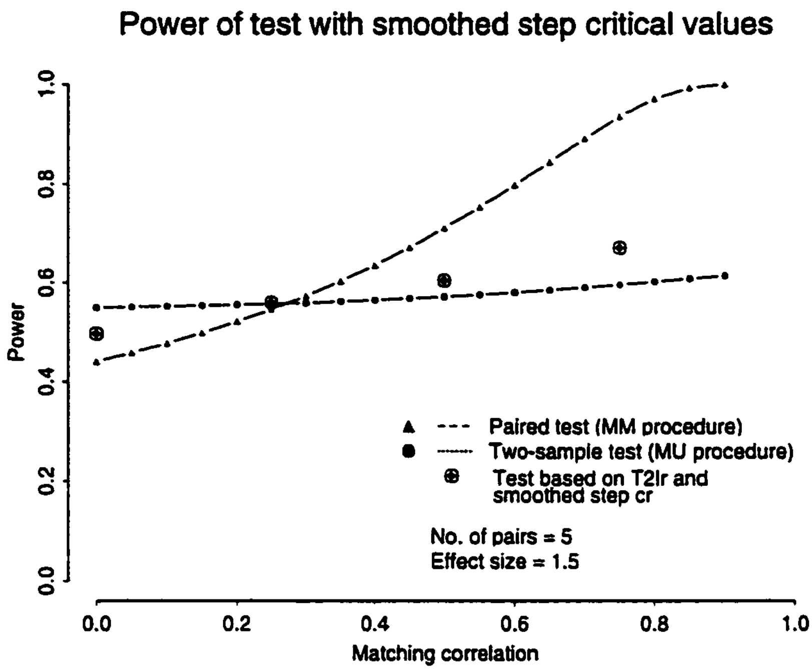


Figure 10.6: Power of a test with smoothed step critical values, compared to the power of a two sample test and a paired test, for various ρ, $n=5$ pairs, nominal size .05, effect size 1.5.

10.5.2 A flexible function

In setting the parameters of the smoothed step function in the previous section, the critical values on the left side of the step were arbitrarily set at 2.95. An alternative is to set the critical values for each r in $r<0$ equal to the the corresponding critical values for each ρ in $\rho<0$ for the unconditional test. Then, the critical value for each r would be no more onerous to achieve than would be the case if it represented the actual ρ of the matching correlation.

This approach gives an increasing set of critical values in the region $r<0$. Figure 10.7 illustrates the function. As with the smoothed step function, the constant value in $r>0$ may be set so as to give an unbiased (or conservatively biased) test at $\rho>0$ or $\rho>-.1$ and conservative for higher values of ρ. Evaluation of the properties of the tests show that the bias and power are similar to the test using the smoothed step function shown in table 10.2

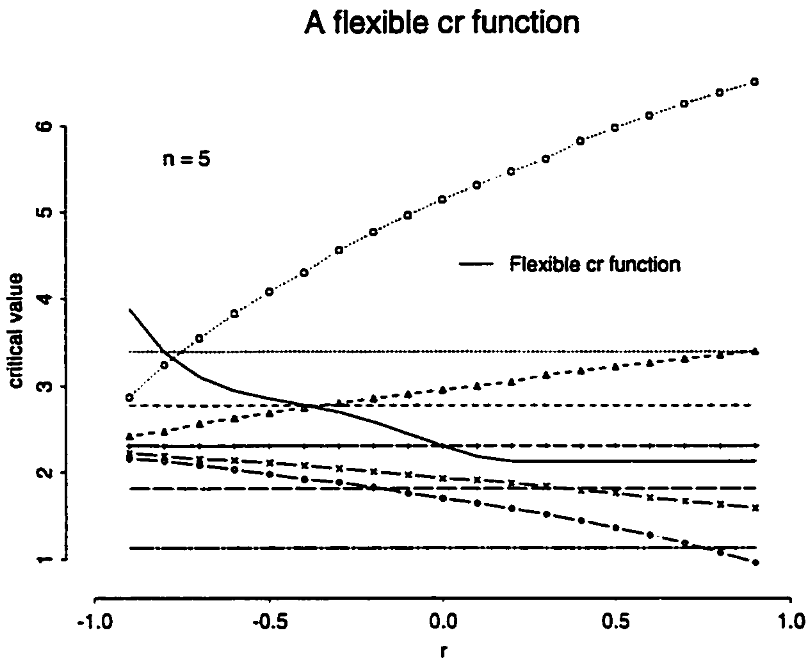


Figure 10.7: Critical values for a test based on $T_2|r$ using a flexible function, $n = 5$ pairs, nominal size .05. Critical values for a conventional test based on $T_2|r$ and a test based on $T_2|\rho$ are also shown.

and figure 10.6.

Chapter 11

A TWO-STAGE PROCEDURE

11.1 Background

In *Breaking the Matches* [9], Diehr considered a procedure where a paired t test is performed if the sample correlation r is significantly different from zero. otherwise a two-sample t test is performed. It was found that the procedure was anti-conservative and was not pursued further in that form.

Procedures involving repeated significance tests tend to have anti-conservative bias. The multiple testing does not of itself make a procedure anti-conservative, but choosing a more powerful test based on an examination of the data tends to make the overall procedure anti-conservative.

Conducting a paired test if r is significant or an unpaired test if it is not does not necessarily choose the more powerful test. For a fixed sample size n, the sample correlation r is significant if it exceeds some fixed value, that is if $|r| > r_\alpha$; thus the procedure amounts to conducting a paired test if $|r| > r_\alpha$, otherwise conducting a two-sample test. If r is significant but negative, a paired test would be conducted, which is known to be less powerful than a two sample test when $\rho < 0$.

The reason why the procedure might be anti-conservative may be surmised from knowledge of the distribution of the two sample t statistic when the data come from a bivariate normal distribution. The paired test is unbiased for all values of ρ. but we showed in table 3.1 that the two-sample test is conservative when $\rho > 0$ and anti-conservative when $\rho < 0$. It will be noted that the anti-conservative bias when $\rho = -.5$ is greater than the conservative bias when $\rho = +.5$ ($0.1035 - 0.05 = 0.0535$ for $\rho = -.5$, compared to $0.05 - 0.01402 = 0.03598$ for $\rho = +.5$, $n = 5$ pairs), and in general that the anti-conservative bias for a given negative ρ is greater than the conservative bias for the corresponding positive ρ. The distribution of

r is symmetric about ρ when $\rho=0$. Thus it might be expected that a procedure which uses the two sample test in a symmetrical interval would tend to be anti-conservative.

The question remains as to whether there is some other procedure which performs a paired test or alternatively a two sample test according to some condition, and which can be made unbiased.

For such a two stage procedure to be of any value, it should have better power than at least one of the constituent tests taken alone, and less bias than at least one of them. Since the power of the paired test is greater than that of the two sample test when $\rho > .27$ approx. for $n=5$ pairs, one might consider performing a paired test when say $r > .25$ and a two sample test when $r < .25$. However, since the two sample test is anti-conservative when $\rho < 0$, one would like to do a paired test when $r < 0$. Thus we consider a procedure where a paired or two sample test is performed according to whether r falls within or outside an asymmetrical interval.

11.2 Derivation

Stated formally, the procedure is

$$\begin{aligned} &\text{Perform a paired test} \quad \text{if} \quad r < r_{\text{lower}} \\ &\text{Perform a two sample test} \quad \text{if} \quad r_{\text{lower}} < r < r_{\text{upper}} \\ &\text{Perform a paired test} \quad \text{if} \quad r_{\text{upper}} > r \ . \end{aligned}$$

We will refer to r_{lower} and r_{upper} as the lower and upper cutpoints for r.

For unbiasedness we require that

$$\int_{r=-1}^{r_{\text{lower}}} P(T_p > c_{p,\alpha}|\rho,r)\, f(r|\rho)\, dr + \int_{r=r_{\text{lower}}}^{r_{\text{upper}}} P(T_2 > c_{2,\alpha}|\rho,r)\, f(r|\rho)\, dr + \int_{r=r_{\text{upper}}}^{1} P(T_p > c_{p,\alpha}|\rho,r)\, f(r|\rho)\, dr \ \leq \ \frac{\alpha}{2} \,. \qquad (11.1)$$

Since a test using T_2 will not be unbiased for all ρ. the test properties discussed at 1.4.1 suggest that the test be made unbiased when $\rho=0$ and the bias at $\rho \neq 0$ evaluated.

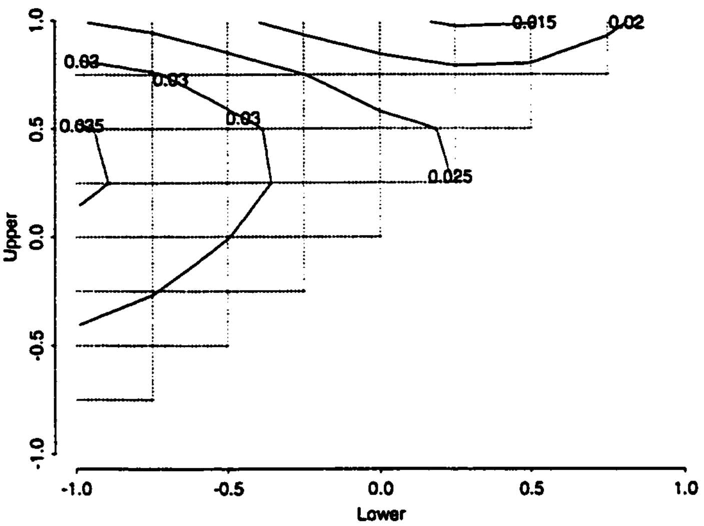
To find pairs of cutpoints satisfying (11.1), the distribution functions of $T_2|\rho$ (since T_2 is independent of r at $\rho = 0$) and of $T_p|r$ (Appendix C), and the density of $r|\rho$, were used initially to compute the probability in eight equal, contiguous bands of r between -1 and 1, by numerical integration. The probabilities are shown in the following table. n is held at 5 pairs and ρ at 0, the test size at 0.05 and non-r-specific critical values are used; that is, it is assumed that unconditional paired and two sample tests are to be performed.

Probability by band of r

Band from $r=$	to $r=$	$\int_r P(T_2 > c_{2,\alpha}\vert\rho, r)\, f(r\vert\rho)\, dr$	$\int_r P(T_p > c_{p,\alpha}\vert\rho, r)\, f(r\vert\rho)\, dr$
-1	-0.75	0.0018037	0.0002325
-0.75	-0.5	0.0030839	0.0005446
-0.5	-0.25	0.0036754	0.0009387
-0.25	0	0.0039370	0.0015242
0	0.25	0.0039370	0.0024506
0.25	0.5	0.0036754	0.0039820
0.5	0.75	0.0030839	0.0065272
0.75	1	0.0018037	0.0087347

The probability shown is one-sided: thus the probability sums to 0.025 for $T_2|r$ and $T_p|r$. $P(T_2 > c_{2,\alpha}|\rho, r)$ is independent of r at $\rho=0$ and hence $\int_r P(T_2 > c_{2,\alpha}|\rho, r)\, f(r|\rho)\, dr$ is simply the probability that R (Pearson's r considered as a random variable) falls in the given band of r for $n=5$ and $\rho=0$.

By interpolation within the bands, the probability in (11.1) for $r_{\text{lower}} \in \{-.99(.02).97\}$, $r_{\text{upper}} \in \{r_{\text{lower}}(.02).99\}$ was computed and mapped, as shown in the following figure.



Approximate pairs of cutpoints were found on the 0.025 probability contour, their values refined by interpolation within the bands, and their precision checked by numerical integration. Four pairs were selected for further evaluation.

Cutpoints		
r_{lower}	r_{upper}	Probability by integration
-0.1	0.68	0.024940
-0.25	0.76	0.024857
-0.4	0.82	0.024975
-0.5	0.86	0.024918

It appeared that if the interval for the two sample test was smaller than $(-0.1, 0.68)$ the procedure behaved closely similar to a paired test; if the interval for the two sample test is very large the procedure will be closely similar to a two sample test.

11.3 Evaluation

The bias of the procedure was computed by numerical integration of the conditional probability distribution functions for $T_2|r$ (Appendix B) and $T_p|r$ and the density of $r|\rho$ according to (11.1) for various ρ. The results are shown in table 11.1.

Table 11.1: Bias of a two stage procedure based on r, for various ρ, $n = 5$ pairs, nominal size .C5.

Type 1 error rate of the two stage procedure							
r_{lower}	r_{upper}	$\rho=-.5$	$\rho=-.25$	$\rho=0$	$\rho=.25$	$\rho=.5$	$\rho=.75$
-0.1	0.68	0.05364	0.05237	0.04988	0.048	0.0476	0.04847
-0.25	0.76	0.05776	0.05478	0.04971	0.04582	0.04488	0.04698
-0.4	0.82	0.06400	0.05833	0.04995	0.0434	0.04135	0.04466
-0.5	0.86	0.06931	0.06097	0.04984	0.04103	0.03774	0.04192

The power was computed for each of the pairs of cutpoints by simulation on 2000 datasets. It was found that the power approached that of the paired test for the cutpoints

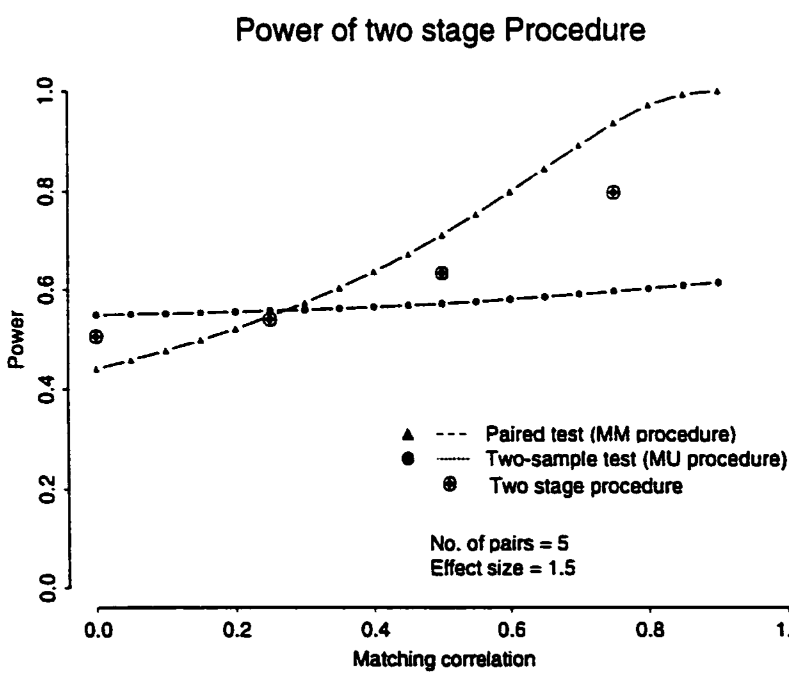


Figure 11.1: Power of a two stage procedure, compared to the power of a two sample test and a paired test, for various ρ, $n=5$ pairs, nominal size .05.

with the smaller ranges. The power for the (-.5,.86) cutpoints is shown in figure 11.1.

As might be expected of a procedure which sometimes conducts a paired test and sometimes a two sample test, the characteristics of the procedure are intermediate between those of its two constituent tests. Like the two sample test. it is biased anti-conservative when $\rho < 0$ and conservative when $\rho > 0$, but less so. The power falls between that of the two sample test and that of the paired test.

The procedure may be made more similar to either of its constituent tests by adjusting the cutpoints. As the cutpoints for r are moved closer to -1 and +1, a two sample test is performed more often, and the bias and power characteristics of the procedure move closer to those of that test. As the cutpoints are moved closer to 0, a paired test is more often conducted, the bias diminishes, and the power approaches that of the paired test.

In this respect the procedure is similar to the tests with linear critical values in chapter 10. When the linear critical values denoted by $c_{r(0)}$ were used, the test was the same as the two sample test, with the same bias and power; when the $c_{r(5)}$ critical values were used, the bias was substantially improved, but the power was closer to that of the paired test.

Chapter 12

DISCUSSION

12.1 The MU procedure

It was a disappointment that Theorem 1 was not proved in more generality. Extensive attempts were made. Proofs were accomplished for the special cases of n large, when the critical value $c=0$, for small effect size δ and power $\beta=\frac{1}{2}$, and when $\rho=0$ and $\delta=0$. While these limited proofs have little applicability in statistical terms, in mathematical terms they prove the relationship between the power and derivative functions in these limited cases. In several thousand computations with random parameter values, no set of values was found for which the theorem was not true.

A large proportion of the relevant literature was reviewed. *The Bateman Manuscript Project* [12] and Abramowitz and Stegun's *Handbook of Mathematical Functions* [1] provide comprehensive collections of the common relations for hypergeometric functions. Appell and Kampé de Fériet's *Fonctions Hypergéométriques et Hypersphériques* (1926) [3] is an extensive classical work on the topic. Exton's *Multiple Hypergeometric Functions and Applications* (1976) [13], and Srivastava's *Multiple Gaussian Hypergeometric Series* (1985) [38] provide more recent compilations. Two articles by Burchnall and Chaundy (1940 and 1941) [6, 7] are frequently cited. While it was not possible to examine in detail the applicability of every theorem and relation contained in these works, no explicit relation which would immediately lead to proof was found.

An aspect of the theorem, and more specifically a feature of the distribution and derivative functions which is particularly problematic is the binomial expansion extended to the case of finite n, addressed in section 4.13. The binomial expansion does not fit neatly into the mainstream of hypergeometric functions. Any future work might seek ways around this problem.

Considerable work remains to be done to complete the proof. If the asymptotic proof or the proof at $c=0$ is taken as a starting point, and assuming the problem of the binomial expansion is resolved, then the Kummer's transformation which is the essence of these two proofs would have to be extended to cover confluent hypergeometric functions in two and possibly three variables, from the present one; this would yield a proof for a 'quasi-derivative' function (a function intermediate between the power and derivative functions). The relation between the derivative and the quasi-derivative has been established, but the forms of the two functions are different from those used elsewhere, and work would be needed to link the results. It may be that some of the mathematics needed to accomplish a proof is not yet available.

The result is considered to be one that may however be relied upon.

As discussed at the beginning of chapter 3, the bias of the MU procedure may be surmised from examination of the statistics, and Diehr demonstrated the bias by simulation. Our results confirm and quantify those findings. that the MU procedure is biased conservative when the correlation induced by matching is positive, and anti-conservative when the correlation is negative. Quantitatively, the anti-conservative bias is such that the procedure would not be acceptable for large negative values of the correlation.

12.2 Other Tests for Matched Designs

12.2.1 Test if ρ known

If circumstances were to arise when the value of the matching correlation ρ could be known with substantial certainty, a test using the two sample t statistic and its distribution as derived at Appendix A would be unbiased and more powerful than either a paired test or a two sample test.

12.2.2 Tests reducing to the Paired Test

As discussed previously, conditioning on the sufficient statistics for the nuisance parameters is a natural and well recognized way to derive tests, and the test based on $\bar{\bar{Y}}|\mathbf{V}$ (chapter 5) was is a way to apply this principle to this problem. That this test reduces to the paired test reinforces the evidence that the paired test is the 'natural' test for the two sample bivariate normal problem; it is also the uniformly most powerful unbiased test and the uniformly most powerful invariant test.

The result that the conditional test based on $T_2|r_3$ was equivalent to the paired test was unexpected, and we were not able to find any theoretical explanation as to why this should be so, except for the algebraic relation given by the lemma in section 8.2. If the circumstances can be established when a matched design guarantees the matching correlation to be positive, the equivalence of the distribution of $T_2|r_3$ to that of the paired statistic may be a starting point for a more powerful test based on restriction of the parameter space.

12.2.3 Tests unbiased at $\rho=0$

The MU procedure may be viewed as a test based on the distribution of $T_2|\rho$ and the assumption that $\rho=0$; thus it is unbiased at $\rho=0$. Of the tests evaluated, those with this property of unbiasedness at $\rho=0$ are first considered.

All of the tests evaluated were biased anti-conservative when $\rho<0$ and biased conservative when $\rho>0$. When they were unbiased or biased conservative, they were more powerful than the paired test in an interval above $\rho=0$. When ρ was large, the paired test was more powerful.

The MU procedure (chapter 3) was biased anti-conservative when $\rho<0$ and conservative when $\rho>0$. The MU procedure may be regarded as 'ignored matching', and thus this shows that matching may be ignored when the correlation induced by matching is guaranteed to be positive. The power of the MU procedure was greater than the paired test for $\rho\in(0,.25)$ approximately for $n=5$; when $\rho>0.25$ the paired test was more powerful.

It will be recalled (chapter 7) that the test based on $T_2|r_1$ is equivalent to the MU procedure when the assumption that $\rho=0$ was used when setting the critical value, and has the same bias and power as the MU procedure.

The tests with linear critical values (section 10.3.1) constitute a progression from the MU procedure (the $c_{r(0)}$ critical values) to a test closely similar to the paired test. The test with the $c_{r(5)}$ critical values was almost unbiased; its power was close to that of the paired

Table 12.1: Comparison of four tests, each unbiased at $\rho=0$; $n=5$ pairs, nominal size .05, effect size 1.5.

	Type 1 error rate when $\rho=-.5$	Power when $\rho=0$	Power when $\rho=.5$
MU procedure (or test based on $T_2 \vert r_1$)	0.1035	0.5494	0.5717
Test with flexible, $c_{r(3)}$ conditional values	0.0731	0.5185	0.674
Two stage procedure with (-0.5,0.86) cutpoints	0.0693	0.5055	0.633
Paired t test	0.05	0.4401	0.7107

test.

Similarly the two stage procedure (section 11.3) may be viewed as a progression from the MU procedure to a test close to the paired test. If the interval of r in which a two sample test is conducted is small, the procedure is close to a paired test; if it is large, the procedure approaches an MU procedure.

Table 12.1 compares the bias and power of the flexible test with linear, $c_{r(3)}$ critical values and the two-stage procedure with the $(-0.5, 0.86)$ cutpoints with the two-sample test and the paired test. For purposes of comparison. the bias at $\rho=-.5$ and the power at $\rho=0$ and $\rho=.5$ are shown.

The table shows that there are in fact tests intermediate between the paired test and the two sample test. The test with linear critical values and the two-stage procedure both have less bias than the MU procedure, and both are more powerful than than MU when $\rho=.5$. The loss in power when $\rho=0$ is modest, and both are more powerful than the paired test for small values of ρ. In circumstances where a positive matching correlation is guaranteed, they could be considered preferable to either the paired test or the MU procedure.

These two are not the only tests that provide results intermediate between the two sample test and the paired test. The parameters of the test with the smoothed step critical values discussed at section 10.5.1 were set with the objective of reducing the bias at $\rho=-0.1$;

they could have been set so that it was unbiased at $\rho = 0$, and would have then provided results close to the test with linear critical values or the two-stage procedure. The test with linear critical values might by some be considered unappealing in that the critical value when $r = 0$ is 2.4044, greater than the 2.3060 for the two sample test. The smoothed step critical values may be set so that both the critical value when $r = 0$ is 2.3060 and the test is unbiased at $\rho = 0$.

These three tests may all be regarded as ways of getting round the difficulty encountered with the conventional test based on $T_2|r_1$. It will be recalled that attempting to make this test unbiased at $\rho = 0$ resulted in an unconditional test, as the distribution was independent of r_1 when $\rho = 0$ (section 8.1) and hence the critical values did not vary with r_1. Two of these tests set critical values that vary with r_1 by methods other than according to the r_1-specific conditional distributions; the other selects the test to be performed based on the value of r_1. The method of setting the critical values and the cutpoints includes a certain arbitrary element: however the arbitrariness is necessitated by the independence of T_2 and r_1 when $\rho = 0$.

12.2.4 Attempts to improve Bias Characteristics

If an investigator were to take the view that, while negative matching correlations could not be ruled out, they were unlikely to be found frequently, particularly under certain known matching schemes, and in any event were unlikely to be large, then a test which provided some protection against bias in some interval below $\rho = 0$ might be considered desirable.

The test based on the smoothed step critical values in section 10.5.1 with the steps set as described there, is unbiased (or conservative) when $\rho > -0.1$, thus providing some measure of protection against small negative ρ. The power is less than for the test with linear critical values and the two-stage procedure; compare for example figure 10.6 with figures 10.3 and 11.1. It may be seen that the power is lower in the test with the greater protection against anti-conservative bias.

As with the tests discussed in the previous section, attempts to remove some of the anti-conservative bias below $\rho = 0$ suffer from a similar arbitrariness. Should the test be unbiased for $\rho > -0.1$? or for $\rho > -0.2$? and so on. It will be recalled that the paired test is the uniformly most powerful test that is unbiased for all values of ρ, so eventually any attempt to remove anti-conservative bias can be no more powerful than a paired test.

It would seem from the results of section 10.5.1 that attempting to extend the range of ρ for which a test is unbiased or conservative reduces power at a rate that soon makes its power comparable with the paired test. In view of the other optimal properties of the paired test, there would seem to be little to be gained from trying to extend the range of unbiasedness below $\rho = 0$ in practical applications; in view of the arbitrariness of the extended range, the paired test might just as well be used.

12.2.5 Kenward and Roger's Method

It was hoped that Kenward and Roger's method would provide a smooth transition between the two sample test and the unpaired test, and to some extent it achieves this. The modified method is intended to yield the same results as in known cases, which would be expected to include the paired test, and in that respect it performed as expected, with results closely comparable.

The standard method is more promising. It differs from the other approaches discussed here in that it uses a different test statistic, one that adapts as the estimated strength of the matching correlation increases and decreases (expression (9.1)), thus suggesting that a method based on an adaptive statistic may offer more potential for handling cases of both effective matching and ineffective matching with the same test. In the instance examined, the test had slight anti-conservative bias over the range of ρ most likely to be found, $0 \geq \rho \geq 0.5$, but as discussed in the test, it may be possible to remedy that by use of a different computed degrees of freedom which did not increase as n got small. The power also increased substantially as ρ became large, a feature not of any great advantage in most trials. The statistic used may be viewed as a modified paired statistic, which suggests that a modified two sample statistic may provide more useful test characteristics, since the power increases less for the two sample test as ρ gets large.

Again, anti-conservativeness tended to increase as ρ became more negative.

12.2.6 *Two Stage Procedure*

The two stage procedure, performing a paired test when r was large (positive or negative) and a two sample test when r was small, possesses some intuitive appeal, and the distributions of Appendices B and C enable a procedure to be constructed which was unbiased when $\rho=0$. As with other methods, it is biased anti-conservative when $\rho<0$. The power is intermediate between the paired test and the two sample test. The asymmetric cutpoints reduce the intuitive appeal somewhat.

12.2.7 *Other tests*

Some other approaches to testing were investigated but not reported in detail as the results were unremarkable. Evaluation of a Bayesian approach with an assumed prior distribution for ρ showed it to have properties closely similar to those for a single fixed assumed value of ρ equal to the expectation of ρ under the prior. Parametric bootstrap methods yielded results similar to the plug-in results of sections 6.3 and 7.2. A test based on the statistic

$$t_{S12} \equiv \frac{\sqrt{n}\,(\bar{y}_{\cdot 1}-\bar{y}_{\cdot 2})}{\sqrt{\frac{1}{n-1}\left(s_1^2+s_2^2-s_{12}\right)}}$$

(note the denominator includes the term s_{12}; the paired statistic has $2s_{12}$ and the two sample has $0s_{12}$ here) had properties intermediate between the paired test and the two sample test, with again anti-conservative bias when $\rho<0$.

12.3 *Matching and Group Randomized Trials*

In section 1.4, we posed the question

> If a group randomized trial is matched, should a matched or an unmatched analysis be used? What are the consequences of conducting an analysis appropriate for an unmatched experiment if the experiment is matched?

The paired test of a matched experiment (the MM procedure) is an unbiased test of the treatment effect, and the power of this procedure for a typical case is shown in figure 1.2; the power increases with the effectiveness of the matching. A two sample test of a matched experiment (MU procedure) is biased conservative when the correlation induced by matching is positive, and anti-conservative when the correlation is negative. When the correlation is positive, figure 1.2 shows that the power is greater than for the paired test for smaller values of the correlation, and less for larger. Theorem 1 shows that the power will increase with the correlation when the power is at least 50%.

When therefore the correlation induced by the matching is positive, both methods of analysis have attractive properties. Neither method is anti-conservative, and the power of both increases with the correlation (provided the power is at least 50% in the case of the MU procedure).

As between the two methods, the choice comes down to which is more powerful (since neither method is anti-conservative). Their relative power depends on the correlation, which, since it is a parameter of the model, can never be known. The prudent course would therefore be to choose the two sample test. It is more powerful than the paired test for the most likely values of the correlation, and power would not be lost if the correlation was higher than expected. The paired test, on the other hand, loses power if the correlation is less than expected. It would be imprudent to design an experiment and compute sample sizes assuming a value for the correlation which would make the paired test more powerful. Then, if the correlation was less than expected, the experiment would be underpowered.

When, however, the correlation induced by the matching is negative, neither method of analysis is particularly attractive. As the correlation becomes more negative, the two sample test becomes increasingly anti-conservative, and the power of the paired test continues to decline.

In section 1.4, we also posed the question

> Should a group randomized trial be matched? Under what circumstances is a matched design preferable to an unmatched design?

If the above argument is correct, that is, if for a matched experiment a two sample test is to be preferred over a paired test, then a two sample test will be conducted whether the design is matched or unmatched.

If the correlation induced by the matching is zero, then the properties of the two designs are the same; the test is unbiased in both cases and the power is the same.

If the experiment is matched and the correlation is positive, the test will be conservative, and if it is negative, anti-conservative. Power will not change substantially with the correlation.

Thus again the answer to the question depends on whether the correlation induced by the matching is positive or negative. If it is positive, then from a statistical point of view a matched design is to be preferred.

If the correlation induced by a matched design is negative, then the question is more problematic, since from a statistical point of view an unmatched design would then be indicated. Set against this must be the convincing arguments for matching the trial in order to maintain its face validity, discussed at the start of section 1.2.

12.4 Further work

A general proof of Theorem 1 remains to be accomplished. The theorem would appear to have implications for small sample experiments in general, not only for group randomized trials, and a general proof would therefore appear to be worth pursuing.

12.4.1 Correlation Induced by Matching

As can be seen, a factor of crucial importance in the analysis of matched experiments is whether the correlation induced by the matching is positive or negative. All the tests and procedures considered which were unbiased when the correlation was zero were conservative when the correlation was positive and anti-conservative when it was negative (except for those that were unbiased for all values of the correlation). Power usually increased as the correlation became more positive, and, in the cases where generalizations could be made, decreased as the correlation became more negative.

It is clearly commonly believed that matching induces a positive correlation; many authors have commented in such terms. Some models, such as Martin's and Lehmann's, only allow for positive correlation.

Simulations show that rank matching tends to lead to a positive correlation. By rank matching is meant a procedure where the experimental units are arranged into pairs according to their ranks on some matching variable; the first and second ranked units form the first pair, the third and fourth the second, and so on. The experimental units within each pair are randomized to treatment or control. Simulations show that such a procedure can be modeled by a bivariate normal distribution with positive correlation. No reference to an exhaustive evaluation of rank matching was found.

On the other hand, it is difficult to dispute Smith & Murray's argument that the weights of the pairs of twin calves could be modeled as negatively correlated. Further, their experiment could be said to be a situation where the matching was close to 'perfect' - the pairs had a common mother - and where there was no treatment effect.

As we have shown here, a matching correlation can be negative when the treatment interaction effects are greater than, and in the opposite direction to, the main matching effects. Wacholder and Weinberg give a similar example of when a negative matching correlation is possible.

Uncertainty as to the circumstances when matching correlations can be relied on to be positive therefore complicates the evaluation of methods for matched designs. Further research clarifying the issue would be a great advantage.

BIBLIOGRAPHY

[1] Milton Abramowitz and Irene A. Stegun. *Handbook of Mathematical Functions.* Dover Publications, 1970.

[2] David B. Abrams et al. Cancer control at the workplace: the Working Well Trial. *Preventive Medicine*, 23:15–27, 1994.

[3] P. Appell et J. Kampé de Fériet. *Fonctions Hypergéométriques et Hypersphériques.* Gauthier-Villars et cie, Paris, 1926.

[4] S. L. Bellamy, R. Gibberd, L. Hancock, P. Howley, B. Kennedy, N. Klar, S. Lipsitz, and L. Ryan. Analysis of dichotomous outcome data for community intervention studies. *Statistical Methods in Medical Research*, 9(2):135–59, 2000.

[5] Jarle Berntsen, Terje O. Espelid, and Alan Genz. Algorithm 698: DCUHRE: An adaptive multidimensional integration routine for a vector of integrals. *Transactions of Mathematical Software*, 17(4):452–456, 1991.

[6] J. L. Burchnall and T. W. Chaundy. Expansions of Appell's double hypergeometric functions. *Quart. J. Math., Oxford Ser.*, 11:249–270, 1940.

[7] J. L. Burchnall and T. W. Chaundy. Expansions of Appell's double hypergeometric functions. II. *Quart. J. Math., Oxford Ser.*, 12:112–128, 1941.

[8] Jerome Cornfield. Randomization by group: A formal analysis. *American Journal of Epidemiology*, 108:100–102, 1978.

[9] Paula Diehr et al. Breaking the matches in a paired t-test for community interventions when the number of pairs is small. *Statistics in Medicine*, 14:1491–1504, 1995.

[10] Allan Donner, K. Stephen Brown, and Penny Brasher. A methodological review of non-therapeutic intervention trials employing cluster randomization, 1979-1989. *International Journal of Epidemiology*, 19(4):795–800, 1990.

[11] Andrew J. Dunning and Paula K. Diehr. Selection of matching variables in community health intervention trials. Technical Report 161, Department of Biostatistics, University of Washington, 1998.

[12] Arthur Erdélyi, editor. *Higher Transcendental Functions: The Bateman Manuscript Project.* McGraw-Hill, 1953.

[13] Harold Exton. *Multiple Hypergeometric Functions and Applications.* John Wiley & Son Ltd., 1976.

[14] Ziding Feng, Paula Diehr, Yutaka Yasui, Brent Evans, Shirley Beresford, and Thomas D. Koepsell. Explaining community-level variance in group randomized trials. *Statistics in Medicine*, 18:539–556, 1999.

[15] R. A. Fisher. *The Design of Experiments.* Oliver and Boyd, 1935.

[16] L. S. Freedman, S. B. Green, and D. P. Byar. Assessing the gain in efficiency due to matching in a community intervention study. *Statistics in Medicine*, 9:943–952, 1990.

[17] Michael H. Gail et al. Aspects of statistical design for the Community Intervention Trial for Smoking Cessation (COMMIT). *Controlled Clinical Trials*, 13:6–21, 1992.

[18] Mitchell H. Gail, Steven D. Mark, Raymond J. Carroll, Sylvan B. Green, and David Pee. On design considerations and randomization-based inference for community intervention trials. *Statistics in Medicine*, 15:1069–1092, 1996.

[19] David A. Harville and Alicia L. Carriquiry. Classical and Bayesian prediction as applied to an unbalanced mixed linear model. *Biometrics*, 48:987–1003, 1992.

[20] R. R. Hocking. *The Analysis of Linear Models.* Brooks/Cole, 1985.

[21] David R. Jacobs, Jr. et al. Community-wide prevention strategies: Evaluation design of the Minnesota Heart Health Program. *Journal of Chronic Disease*, 13(10):775–788, 1986.

[22] Norman L. Johnson and Samuel Kotz. *Continuous Multivariate Distributions.* Wiley, 1972.

[23] Samuel Karlin. *Total Positivity.* Stanford University Press, 1968.

[24] Michael G. Kenward and James H. Roger. Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics*, 53:983–997, 1997.

[25] Neil Klar and Allan Donner. The merits of matching in community intervention trials: A cautionary tale. *Statistics in Medicine*, 16:1753–1764. 1997.

[26] Neil Klar and Allan Donner. *Design and Analysis of Cluster Randomization Trials in Health Research.* Arnold, 2000.

[27] Thomas D. Koepsell, Donald C. Martin, Paula H. Diehr, Bruce M. Psaty, Edward H. Wagner, Edward B. Perrin, and Allen Cheadle. Data analysis and sample size issues in evaluations of community-based health promotion and disease prevention programs: A mixed-model analysis of variance approach. *Journal of Clinical Epidemiology*, 44:34–43, 1994.

[28] Marakatha Krishnan. Series representations of the doubly noncentral t distribution (corr: V64 p1702). *Journal of the American Statistical Association*, 63:1004, 1968.

[29] E. L. Lehmann. *Testing Statistical Hypotheses (Second Edition).* Wiley, 1986.

[30] S. L. Mann, A. V. Peterson, Jr., P. M. Marek, and K. A. Kealey. The Hutchinson Smoking Prevention Project trial: design and baseline characteristics. *Preventive Medicine*, 30:485–495, 2000.

[31] Donald C. Martin et al. The effect of matching on the power of randomized community intervention trials. *Statistics in Medicine*, 12:329–338, 1993.

[32] D. M. Murray, H. A. Feldman, and P. G. McGovern. Components of variance in a group-randomized trial analysed via a random-coefficients model: the rapid early action for coronary treatment (react) trial. *Statistical Methods in Medical Research*, 9(2):117–33, 2000.

[33] David M. Murray. *Design and Analysis of Group-randomized Trials*. Oxford University Press, 1998.

[34] Michael A. Proschan. On the distribution of the unpaired t-statistic with paired data. *Statistics in Medicine*, 15:1059–1063, 1996.

[35] Friedrich Pukelsheim and George P. H. Styan. Nonnegative definiteness of the estimated dispersion matrix in a multivariate linear model. *Bulletin de l'Académie Polonaise des Sciences, Série des Sciences Mathématiques*. 27:327–330. 1979.

[36] Shayle R. Searle, George Casella, and Charles E. McCulloch. *Variance Components*. Wiley, 1992.

[37] David W. Smith and Leigh W. Murray. An alternative to Eisenhart's Model II and mixed model in the case of negative variance estimates. *Journal of the American Statistical Association*, 79:145, 1984.

[38] H. M. Srivastava and Per W. Karlsson. *Multiple Gaussian Hypergeometric Series*. John Wiley & Son Ltd., 1985.

[39] Alan Stuart and J. Keith Ord. *Kendall's Advanced Theory of Statistics. Volume 1. Distribution Theory (Sixth Edition)*. Edward Arnold, 1994.

[40] Sholom Wacholder and Clarice Ring Weinberg. Paired versus two-sample design for a clinical trial of treatments with dichotomous outcome: Power considerations. *Biometrics*, 38:801, 1982.

[41] Edward H. Wagner et al. The evaluation of the Henry J. Kaiser Family Foundation's Community Health Promotion Grant Program: Design. *Journal of Clinical Epidemiology*, 44(7):685–699, 1991.

[42] Yining Wang and Michael P. McDermott. Conditional likelihood ratio test for a nonnegative normal mean vector. *Journal of the American Statistical Association*, 93:380–386, 1998.

Appendix A

THE DISTRIBUTION OF $T_2|\rho$

Recall that by $T_2|\rho$ we mean the two-sample t statistic when the observations are realizations of a correlated bivariate normal distribution (see *Notation* at 2).

In this Appendix, we

a) Express the distribution function of $T_2|\rho$ as a Power Series (section A.1);

b) Express the distribution function of $T_2|\rho$ as a guaranteed convergent Power Series suitable for computations (section A.2);

c) Derive the density of $T_2|\rho$ (section A.3);

d) Derive the asymptotic distribution of $T_2|\rho$ (section A.4);

e) Derive the derivative of the distribution function of $T_2|\rho$ with respect to ρ (section A.5);

f) Confirm by simulation the distribution of $T_2|\rho$ for negative ρ (section A.6);

g) Provide Splus code for the distribution function of $T_2|\rho$ (section A.7);

h) Provide Splus code for the density function of $T_2|\rho$ (section A.8).

A.1 The Distribution Function of $T_2|\rho$ as a Power Series

Proschan [34] showed (at equation (9)) that

$$F_{T_2}(t) \;=\; \int_{w=0}^{\infty} \left[G_{2n-2}\left(t; \delta_\rho, \frac{2\rho w}{1-\rho}\right)\right] f_{n-1}(w)\,dw$$

where

i) $T_2 \equiv \dfrac{\sqrt{n}(\bar{Y}_{2\cdot} - \bar{Y}_{1\cdot})}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(Y_{1i}-\bar{Y}_{1\cdot})^2 + (Y_{2i}-\bar{Y}_{2\cdot})^2}}$, the two-sample t-statistic;

ii) $\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} \sim N_2\left(\begin{matrix}\mu_1 \\ \mu_2\end{matrix}, \sigma^2\begin{pmatrix} 1 & \rho \\ \rho & 1\end{pmatrix}\right) \quad \rho > 0;$

iii) $G_\nu(\cdot\ ;\delta,\lambda)$ is the doubly non-central t distribution function with ν degrees of freedom and non-centrality parameters δ and λ;

iv) $\delta_\rho \equiv \sqrt{\frac{n}{2(1-\rho)}}\left(\frac{\mu_2-\mu_1}{\sigma}\right)$; and

v) $f_\nu(\cdot)$ is the $\mathcal{X}^2$ density function with ν degrees of freedom.

For the doubly non-central t distribution function, Krishnan [28] (at equation (11)) gives

$$\begin{aligned} G_\nu(t;\delta,\lambda) &= G(0;\delta) + \sum_{i,j=0}^{\infty} \frac{E(\lambda,i)\,E(\delta^2,j)\,a^{j+1/2}}{2\,\Gamma(\nu/2+i)\,\Gamma(j+3/2)} \\ &\quad \times \Bigg[\Gamma(i+j+\nu/2+1/2)\,H(j+1/2,1-i-\nu/2,j+3/2;a) \\ &\qquad + \frac{\delta}{j+1}\sqrt{\frac{a}{2}}\,\Gamma(i+j+\nu/2+1)\,H(j+1,1-i-\nu/2,j+2;a)\Bigg] \end{aligned}$$

where

i) $E(\alpha,k) \equiv \dfrac{e^{-\alpha/2}\left(\frac{\alpha}{2}\right)^k}{\Gamma(k+1)}$

ii) H is the Gauss hypergeometric function; see 2.4:

iii) $a \equiv \dfrac{t^2}{2n-2+t^2}$.

The $\mathcal{X}^2$ density function is well known

$$f_\nu(u) = \frac{u^{\nu/2-1}\exp\left(-\frac{u}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)2^{\nu/2}}.$$

Hence, and expanding $G_{2n-2}(0;\delta_\rho) = \frac{1}{2} - \frac{\delta_\rho}{\sqrt{2\pi}}M(\frac{1}{2},\frac{3}{2};-\delta_\rho^2/2)$.

$$\begin{aligned} F_{T_2}(t) \equiv P(T_2 \le t) &= \int_{w=0}^{\infty}\Bigg[\frac{1}{2} - \frac{\delta_\rho}{\sqrt{2\pi}}M(\tfrac{1}{2},\tfrac{3}{2};-\delta_\rho^2/2) \\ &\quad + \sum_{i,j=0}^{\infty} E\left(\frac{2\rho w}{1-\rho},i\right)\frac{E(\delta_\rho^2,j)\,a^{j+1/2}}{2\,\Gamma(n+i-1)\,\Gamma(j+\frac{3}{2})} \\ &\quad \times \Bigg(\Gamma(n+i+j-\tfrac{1}{2})\,H(j+\tfrac{1}{2},2-n-i,j+\tfrac{3}{2};a) \\ &\qquad + \frac{\delta_\rho}{j+1}\sqrt{\frac{a}{2}}\,\Gamma(n+i+j)\,H(j+1,2-n-i,j+2;a)\Bigg)\Bigg] \\ &\quad \times \frac{w^{(n-3)/2}\exp\left(-\frac{w}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)2^{(n-1)/2}}\,dw \end{aligned}$$

where M is the confluent hypergeometric function; see 2.4.

Integrating term by term

$$
\begin{aligned}
F_{T_2}(t) \;=\; & \frac{1}{2} - \frac{\delta_\rho}{\sqrt{2\pi}} M(\tfrac{1}{2}, \tfrac{3}{2}; -\delta_\rho^2/2) \\
& + \sum_{i,j=0}^{\infty} \frac{E(\delta_\rho^2, j)\ a^{j+1/2}}{2\,\Gamma(n+i-1)\ \Gamma(j+\frac{3}{2})} \\
& \times \left(\Gamma(n+i+j-\tfrac{1}{2})\ H(j+\tfrac{1}{2}, 2-n-i, j+\tfrac{3}{2}; a) \right. \\
& \qquad \left. + \frac{\delta_\rho}{j+1}\sqrt{\frac{a}{2}}\ \Gamma(n+i+j)\ H(j+1, 2-n-i, j+2; a) \right) \\
& \times \int_{w=0}^{\infty} E\left(\frac{2\rho w}{1-\rho}, i\right) \frac{w^{(n-3)/2} \exp\left(-\frac{w}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right) 2^{(n-1)/2}}\, dw\ .
\end{aligned}
$$

Evaluating the integral

$$
\begin{aligned}
& \int_{w=0}^{\infty} E\left(\frac{2\rho w}{1-\rho}, i\right) \frac{w^{(n-3)/2} \exp\left(-\frac{w}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right) 2^{(n-1)/2}}\, dw \\
= & \int_{w=0}^{\infty} \frac{e^{-\frac{\rho w}{1-\rho}} \left(\frac{\rho w}{1-\rho}\right)^i}{\Gamma(i+1)} \frac{w^{(n-3)/2} \exp\left(-\frac{w}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right) 2^{(n-1)/2}}\, dw \\
= & \frac{1}{\Gamma(i+1)\ \Gamma\left(\frac{n-1}{2}\right) 2^{(n-1)/2}} \\
& \times \int_{w=0}^{\infty} e^{-\frac{w(1+\rho)}{2(1-\rho)}} \left(\frac{\rho}{1-\rho}\right)^i w^{\frac{n+2i-3}{2}} \left(\frac{1+\rho}{2(1-\rho)}\right)^{\frac{n+2i-3}{2}} \left(\frac{2(1-\rho)}{1+\rho}\right)^{\frac{n+2i-3}{2}} dw \\
= & \frac{\left(\frac{\rho}{1-\rho}\right)^i \left(\frac{2(1-\rho)}{1+\rho}\right)^{\frac{n+2i-3}{2}}}{\Gamma(i+1)\ \Gamma\left(\frac{n-1}{2}\right) 2^{(n-1)/2}} \int_{w=0}^{\infty} e^{-\frac{w(1+\rho)}{2(1-\rho)}} \left(\frac{w(1+\rho)}{2(1-\rho)}\right)^{\frac{n+2i-3}{2}} dw, \\
& \text{put } u = \frac{w(1+\rho)}{2(1-\rho)};\ \ dw = \frac{2(1-\rho)}{1+\rho} du\ \ldots \\
= & \frac{\left(\frac{\rho}{1-\rho}\right)^i \left(\frac{2(1-\rho)}{1+\rho}\right)^{\frac{n+2i-1}{2}}}{\Gamma(i+1)\ \Gamma\left(\frac{n-1}{2}\right) 2^{(n-1)/2}} \int_{u=0}^{\infty} e^{-u} u^{\frac{n+2i-3}{2}}\, du. \\
= & \frac{\left(\frac{2\rho}{1+\rho}\right)^i \left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{\Gamma(i+1)\ \Gamma\left(\frac{n-1}{2}\right)} \Gamma\left(\frac{n+2i-1}{2}\right)\ .
\end{aligned}
$$

Then

$$
F_{T_2}(t) \;=\; \frac{1}{2} - \frac{\delta_\rho}{\sqrt{2\pi}} M(\tfrac{1}{2}, \tfrac{3}{2}; -\delta_\rho^2/2)
$$

$$
\begin{aligned}
&+\sum_{i,j=0}^{\infty}\frac{E(\delta_\rho^2,j)\ a^{j+1/2}}{2\,\Gamma(n+i-1)\ \Gamma(j+\frac{3}{2})}\\
&\times\left(\Gamma(n+i+j-\tfrac{1}{2})\ H(j+\tfrac{1}{2},2-n-i,j+\tfrac{3}{2};a)\right.\\
&\qquad\left.+\frac{\delta_\rho}{j+1}\sqrt{\frac{a}{2}}\,\Gamma(n+i+j)\ H(j+1,2-n-i,j+2;a)\right)\\
&\times\frac{\left(\frac{2\rho}{1+\rho}\right)^{i}\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{\Gamma(i+1)\ \Gamma\left(\frac{n-1}{2}\right)}\,\Gamma\left(\frac{n+2i-1}{2}\right)\\
=\ &\frac{1}{2}-\frac{\delta_\rho}{\sqrt{2\pi}}M(\tfrac{1}{2},\tfrac{3}{2};-\delta_\rho^2/2)\\
&+\frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{2\,\Gamma(\frac{n-1}{2})}\sum_{i,j=0}^{\infty}\frac{\Gamma(\frac{n+2i-1}{2})\ \left(\frac{2\rho}{1+\rho}\right)^{i}}{\Gamma(n+i-1)\ \Gamma(i+1)}\ \frac{e^{-\frac{\delta_\rho^2}{2}}\left(\frac{\delta_\rho^2}{2}\right)^{j}a^{j+1/2}}{\Gamma(j+1)\ \Gamma(j+\frac{3}{2})}\\
&\times\left(\Gamma(n+i+j-\tfrac{1}{2})\ H(j+\tfrac{1}{2},2-n-i,j+\tfrac{3}{2};a)\right.\\
&\qquad\left.+\frac{\delta_\rho}{j+1}\sqrt{\frac{a}{2}}\,\Gamma(n+i+j)\ H(j+1,2-n-i,j+2;a)\right)\\
=\ &\frac{1}{2}-\frac{\delta_\rho}{\sqrt{2\pi}}M(\tfrac{1}{2},\tfrac{3}{2};-\delta_\rho^2/2)\\
&+\frac{e^{-\frac{\delta_\rho^2}{2}}\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{2\,\Gamma(\frac{n-1}{2})}\sum_{i,j=0}^{\infty}\frac{\Gamma(\frac{n+2i-1}{2})\ \left(\frac{2\rho}{1+\rho}\right)^{i}}{\Gamma(n+i-1)\ \Gamma(i+1)}\\
&\times\left(\frac{\Gamma(n+i+j-\frac{1}{2})\ \left(\frac{\delta_\rho^2}{2}\right)^{j}a^{j+1/2}}{\Gamma(j+1)\ \Gamma(j+\frac{3}{2})}\ H(j+\tfrac{1}{2},2-n-i,j+\tfrac{3}{2};a)\right.\\
&\qquad\left.+\frac{\Gamma(n+i+j)\left(\frac{\delta_\rho^2}{2}\right)^{j+\frac{1}{2}}a^{j+1}}{\Gamma(j+\frac{3}{2})\ \Gamma(j+2)}\ H(j+1,2-n-i,j+2;a)\right).
\end{aligned}
$$

Putting $a \equiv \frac{t^2}{2n-2+t^2}$ and using $H(\alpha,\beta,\gamma;x) \equiv (1-x)^{-\alpha}H(\alpha,\gamma-\beta,\gamma;\frac{x}{x-1})$ (Erdélyi [12], p.109) gives

$$
F_{T_2}(t)\ =\ \frac{1}{2}-\frac{\delta_\rho}{\sqrt{2\pi}}M\left(\tfrac{1}{2},\tfrac{3}{2};-\tfrac{\delta_\rho^2}{2}\right)\ +\ \frac{e^{-\frac{\delta_\rho^2}{2}}\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{2\,\Gamma(\frac{n-1}{2})}\sum_{i,j=0}^{\infty}\frac{\Gamma(\frac{n-1}{2}+i)\ \left(\frac{2\rho}{1+\rho}\right)^{i}}{\Gamma(n+i-1)\ \Gamma(i+1)}\ \times
$$

$$\left[\frac{\Gamma(n+i+j-\frac{1}{2})\left(\frac{\delta_\rho^2}{2}\right)^j\left(\frac{t^2}{2n-2+t^2}\right)^{j+\frac{1}{2}}}{\Gamma(j+1)\,\Gamma(j+\frac{3}{2})}\left(\frac{2n-2}{2n-2+t^2}\right)^{-(j+\frac{1}{2})}H\!\left(j+\tfrac{1}{2},n+i+j-\tfrac{1}{2},j+\tfrac{3}{2};\tfrac{-t^2}{2n-2}\right)\right.$$
$$\left.+\frac{\Gamma(n+i+j)\left(\frac{\delta_\rho^2}{2}\right)^{j+\frac{1}{2}}\left(\frac{t^2}{2n-2+t^2}\right)^{j+1}}{\Gamma(j+\frac{3}{2})\,\Gamma(j+2)}\left(\frac{2n-2}{2n-2+t^2}\right)^{-(j+1)}H\!\left(j+1,n+i+j,j+2;\tfrac{-t^2}{2n-2}\right)\right].$$

Expanding the Gauss hypergeometric function (see 2.4) gives

$$F_{T_2}(t) = \frac{1}{2}-\frac{\delta_\rho}{\sqrt{2\pi}}M\!\left(\tfrac{1}{2},\tfrac{3}{2};-\tfrac{\delta_\rho^2}{2}\right)+\frac{e^{-\frac{\delta_\rho^2}{2}}\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{2\,\Gamma(\frac{n-1}{2})}\sum_{i,j,k=0}^{\infty}\frac{\Gamma(\frac{n-1}{2}+i)\left(\frac{2\rho}{1+\rho}\right)^i}{\Gamma(n+i-1)\,\Gamma(i+1)}\times$$
$$\left[\frac{\Gamma(n+i+j-\frac{1}{2})\left(\frac{\delta_\rho^2}{2}\right)^j\left(\frac{t^2}{2n-2}\right)^{j+\frac{1}{2}}}{\Gamma(j+1)\,\Gamma(j+\frac{3}{2})}\frac{\Gamma(j+k+\frac{1}{2})\,\Gamma(n+i+j+k-\frac{1}{2})\,\Gamma(j+\frac{3}{2})}{\Gamma(j+\frac{1}{2})\,\Gamma(n+i+j-\frac{1}{2})\,\Gamma(j+k+\frac{3}{2})\,\Gamma(k+1)}\left(\frac{-t^2}{2n-2}\right)^k\right.$$
$$\left.+\frac{\Gamma(n+i+j)\left(\frac{\delta_\rho^2}{2}\right)^{j+\frac{1}{2}}\left(\frac{t^2}{2n-2}\right)^{j+1}}{\Gamma(j+\frac{3}{2})\,\Gamma(j+2)}\frac{\Gamma(j+k+1)\Gamma(n+i+j+k)\Gamma(j+2)}{\Gamma(j+1)\Gamma(n+i+j)\Gamma(j+k+2)\Gamma(k+1)}\left(\frac{-t^2}{2n-2}\right)^k\right]$$
$$= \frac{1}{2}-\frac{\delta_\rho}{\sqrt{2\pi}}M\!\left(\tfrac{1}{2},\tfrac{3}{2};-\tfrac{\delta_\rho^2}{2}\right)+\frac{e^{-\frac{\delta_\rho^2}{2}}\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{2\,\Gamma(\frac{n-1}{2})}\sum_{i,j,k=0}^{\infty}\frac{\Gamma(\frac{n-1}{2}+i)\left(\frac{2\rho}{1+\rho}\right)^i}{\Gamma(n+i-1)\,\Gamma(i+1)}$$
$$\times\left[\frac{\Gamma(j+k+\frac{1}{2})\,\Gamma(n+i+j+k-\frac{1}{2})\,(-1)^k}{\Gamma(j+k+\frac{3}{2})\,\Gamma(j+\frac{1}{2})\,\Gamma(j+1)\,\Gamma(k+1)}\left(\frac{\delta_\rho^2}{2}\right)^j\left(\frac{t^2}{2n-2}\right)^{j+k+\frac{1}{2}}\right.$$
$$\left.+\frac{\Gamma(j+k+1)\,\Gamma(n+i+j+k)\,(-1)^k}{\Gamma(j+k+2)\,\Gamma(j+\frac{3}{2})\,\Gamma(j+1)\,\Gamma(k+1)}\left(\frac{\delta_\rho^2}{2}\right)^{j+\frac{1}{2}}\left(\frac{t^2}{2n-2}\right)^{j+k+1}\right]. \quad \text{(A.1)}$$

Diagonalizing the j,k infinite sums by $\sum_{i,j=0}^{\infty} c(i,j)\,x^i\,y^j = \sum_{i=0}^{\infty}\sum_{j=0}^{i} c(i-j,j)\,x^{i-j}\,y^j$ gives

$$F_{T_2}(t) = \frac{1}{2}-\frac{\delta_\rho}{\sqrt{2\pi}}M\!\left(\tfrac{1}{2},\tfrac{3}{2};-\tfrac{\delta_\rho^2}{2}\right)+\frac{e^{-\frac{\delta_\rho^2}{2}}\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{2\,\Gamma(\frac{n-1}{2})}\sum_{i,k=0}^{\infty}\sum_{j=0}^{k}\frac{\Gamma(\frac{n-1}{2}+i)\left(\frac{2\rho}{1+\rho}\right)^i}{\Gamma(n+i-1)\,\Gamma(i+1)}$$
$$\times\left[\frac{\Gamma(k+\frac{1}{2})\,\Gamma(n+i+k-\frac{1}{2})\,(-1)^{k-j}}{\Gamma(k+\frac{3}{2})\,\Gamma(j+\frac{1}{2})\,\Gamma(j+1)\,\Gamma(k-j+1)}\left(\frac{\delta_\rho^2}{2}\right)^j\left(\frac{t^2}{2n-2}\right)^{k+\frac{1}{2}}\right.$$
$$\left.+\frac{\Gamma(k+1)\,\Gamma(n+i+k)\,(-1)^{k-j}}{\Gamma(k+2)\,\Gamma(j+\frac{3}{2})\,\Gamma(j+1)\,\Gamma(k-j+1)}\left(\frac{\delta_\rho^2}{2}\right)^{j+\frac{1}{2}}\left(\frac{t^2}{2n-2}\right)^{k+1}\right]$$

$$= \frac{1}{2} - \frac{\delta_\rho}{\sqrt{2\pi}} M\left(\tfrac{1}{2}, \tfrac{3}{2}; -\tfrac{\delta_\rho^2}{2}\right) + \frac{e^{-\frac{\delta_\rho^2}{2}} \left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{2\,\Gamma(\frac{n-1}{2})} \sum_{i,k=0}^{\infty} \sum_{j=0}^{k} \frac{\Gamma(\frac{n-1}{2}+i)\left(\frac{2\rho}{1+\rho}\right)^i}{\Gamma(n+i-1)\,\Gamma(i+1)}$$
$$\times \left[\frac{\Gamma(k+\frac{1}{2})\,\Gamma(n+i+k-\frac{1}{2})}{\Gamma(k+\frac{3}{2})\,\Gamma(j+\frac{1}{2})\,\Gamma(j+1)\,\Gamma(k-j+1)} \left(-\tfrac{\delta_\rho^2}{2}\right)^j \left(\tfrac{t}{\sqrt{2n-2}}\right) \left(-\tfrac{t^2}{2n-2}\right)^k \right.$$
$$\left. - \frac{\Gamma(k+1)\,\Gamma(n+i+k)}{\Gamma(k+2)\,\Gamma(j+\frac{3}{2})\,\Gamma(j+1)\,\Gamma(k-j+1)} \left(\tfrac{\delta_\rho}{\sqrt{2}}\right) \left(-\tfrac{\delta_\rho^2}{2}\right)^j \left(-\tfrac{t^2}{2n-2}\right)^{k+1}\right].$$

Expressing $\Gamma(k-j+1)$ in terms of Pochhammer's symbol (see 2.5) gives

$$F_{T_2}(t) = \frac{1}{2} - \frac{\delta_\rho}{\sqrt{2\pi}} M\left(\tfrac{1}{2}, \tfrac{3}{2}; -\tfrac{\delta_\rho^2}{2}\right) + \frac{e^{-\frac{\delta_\rho^2}{2}} \left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{2\,\Gamma(\frac{n-1}{2})} \sum_{i,k=0}^{\infty} \sum_{j=0}^{k} \frac{\Gamma(\frac{n-1}{2}+i)\left(\frac{2\rho}{1+\rho}\right)^i}{\Gamma(n+i-1)\,\Gamma(i+1)}$$
$$\times \left[\frac{\Gamma(k+\frac{1}{2})\,\Gamma(n+i+k-\frac{1}{2})}{\Gamma(k+\frac{3}{2})\,\Gamma(j+\frac{1}{2})\,\Gamma(j+1)\,\Gamma(k+1)\,(k+1)_{-j}} \left(-\tfrac{\delta_\rho^2}{2}\right)^j \left(\tfrac{t}{\sqrt{2n-2}}\right) \left(-\tfrac{t^2}{2n-2}\right)^k \right.$$
$$\left. - \frac{\Gamma(k+1)\,\Gamma(n+i+k)}{\Gamma(k+2)\,\Gamma(j+\frac{3}{2})\,\Gamma(j+1)\,\Gamma(k+1)\,(k+1)_{-j}} \left(\tfrac{\delta_\rho}{\sqrt{2}}\right) \left(-\tfrac{\delta_\rho^2}{2}\right)^j \left(-\tfrac{t^2}{2n-2}\right)^{k+1}\right]$$
$$= \frac{1}{2} - \frac{\delta_\rho}{\sqrt{2\pi}} M\left(\tfrac{1}{2}, \tfrac{3}{2}; -\tfrac{\delta_\rho^2}{2}\right) + \frac{e^{-\frac{\delta_\rho^2}{2}} \left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{2\,\Gamma(\frac{n-1}{2})} \sum_{i,k=0}^{\infty} \sum_{j=0}^{k} \frac{\Gamma(\frac{n-1}{2}+i)\left(\frac{2\rho}{1+\rho}\right)^i}{\Gamma(n+i-1)\,\Gamma(i+1)}$$
$$\times \left[\frac{\Gamma(k+\frac{1}{2})\,\Gamma(n+i+k-\frac{1}{2})\,(-k)_j}{\Gamma(k+\frac{3}{2})\,\Gamma(j+\frac{1}{2})\,\Gamma(j+1)\,\Gamma(k+1)\,(-1)^j} \left(-\tfrac{\delta_\rho^2}{2}\right)^j \left(\tfrac{t}{\sqrt{2n-2}}\right) \left(-\tfrac{t^2}{2n-2}\right)^k \right.$$
$$\left. - \frac{\Gamma(n+i+k)\,(-k)_j}{\Gamma(k+2)\,\Gamma(j+\frac{3}{2})\,\Gamma(j+1)\,(-1)^j} \left(\tfrac{\delta_\rho}{\sqrt{2}}\right) \left(-\tfrac{\delta_\rho^2}{2}\right)^j \left(-\tfrac{t^2}{2n-2}\right)^{k+1}\right].$$

Expressing the j power series as a confluent hypergeometric function (see 2.4) gives

$$F_{T_2}(t) = \frac{1}{2} - \frac{\delta_\rho}{\sqrt{2\pi}} M\left(\tfrac{1}{2}, \tfrac{3}{2}; -\tfrac{\delta_\rho^2}{2}\right) + \frac{e^{-\frac{\delta_\rho^2}{2}} \left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{2\,\Gamma(\frac{n-1}{2})} \sum_{i,k=0}^{\infty} \frac{\Gamma(\frac{n-1}{2}+i)\left(\frac{2\rho}{1+\rho}\right)^i}{\Gamma(n+i-1)\,\Gamma(i+1)}$$
$$\times \left[\frac{\Gamma(k+\frac{1}{2})\,\Gamma(n+i+k-\frac{1}{2})}{\Gamma(k+\frac{3}{2})\,\Gamma(k+1)} \left(\tfrac{t}{\sqrt{2n-2}}\right) \left(-\tfrac{t^2}{2n-2}\right)^k \frac{1}{\Gamma(\frac{1}{2})} M\left(-k, \tfrac{1}{2}; \tfrac{\delta_\rho^2}{2}\right) \right.$$
$$\left. - \frac{\Gamma(n+i+k)}{\Gamma(k+2)} \left(\tfrac{\delta_\rho}{\sqrt{2}}\right) \left(-\tfrac{t^2}{2n-2}\right)^{k+1} \frac{1}{\Gamma(\frac{3}{2})} M\left(-k, \tfrac{3}{2}; \tfrac{\delta_\rho^2}{2}\right)\right].$$

Applying Kummer's relation $M(\alpha,\beta;x)=e^x\,M(\beta-\alpha,\beta;-x)$ (see *Handbook of Mathematical Functions* [1]) gives

$$\begin{aligned}
F_{T_2}(t) &= \frac{1}{2}-\frac{\delta_\rho}{\sqrt{2\pi}}M\left(\tfrac{1}{2},\tfrac{3}{2};-\tfrac{\delta_\rho^2}{2}\right)+\frac{e^{-\frac{\delta_\rho^2}{2}}\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{2\,\Gamma\left(\frac{n-1}{2}\right)}\sum_{i,k=0}^{\infty}\frac{\Gamma\left(\frac{n-1}{2}+i\right)\left(\frac{2\rho}{1+\rho}\right)^i}{\Gamma(n+i-1)\,\Gamma(i+1)}\\
&\times\left[\frac{\Gamma\left(k+\frac{1}{2}\right)\Gamma\left(n+i+k-\frac{1}{2}\right)}{\Gamma\left(k+\frac{3}{2}\right)\Gamma(k+1)}\left(\frac{t}{\sqrt{2n-2}}\right)\left(-\frac{t^2}{2n-2}\right)^k\frac{1}{\Gamma\left(\frac{1}{2}\right)}e^{\frac{\delta_\rho^2}{2}}M\left(k+\tfrac{1}{2},\tfrac{1}{2};-\tfrac{\delta_\rho^2}{2}\right)\right.\\
&\quad\left.-\frac{\Gamma(n+i+k)}{\Gamma(k+2)}\left(\frac{\delta_\rho}{\sqrt{2}}\right)\left(-\frac{t^2}{2n-2}\right)^{k+1}\frac{1}{\Gamma\left(\frac{3}{2}\right)}e^{\frac{\delta_\rho^2}{2}}M\left(k+\tfrac{3}{2},\tfrac{3}{2};-\tfrac{\delta_\rho^2}{2}\right)\right].
\end{aligned}$$

Note that $\dfrac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{\Gamma\left(\frac{n-1}{2}\right)}\displaystyle\sum_{i=0}^{\infty}\frac{\Gamma\left(\frac{n-1}{2}+i\right)\left(\frac{2\rho}{1+\rho}\right)^i}{\Gamma(i+1)}=1$: expand the confluent hypergeometric functions as power series (see 2.4) gives

$$\begin{aligned}
F_{T_2}(t) &= \frac{1}{2}-\frac{\delta_\rho}{\sqrt{2\pi}}\frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{\Gamma\left(\frac{n-1}{2}\right)}\sum_{i=0}^{\infty}\frac{\Gamma\left(\frac{n-1}{2}+i\right)\left(\frac{2\rho}{1+\rho}\right)^i}{\Gamma(i+1)}\frac{\Gamma\left(\frac{3}{2}\right)}{\Gamma\left(\frac{1}{2}\right)}\sum_{j=0}^{\infty}\frac{\Gamma\left(\frac{1}{2}+j\right)}{\Gamma\left(\frac{3}{2}+j\right)\Gamma(j+1)}\left(-\frac{\delta_\rho^2}{2}\right)^j\\
&+\frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{2\,\Gamma\left(\frac{n-1}{2}\right)}\sum_{i,j,k=0}^{\infty}\frac{\Gamma\left(\frac{n-1}{2}+i\right)\left(\frac{2\rho}{1+\rho}\right)^i}{\Gamma(n+i-1)\,\Gamma(i+1)}\\
&\times\left[\frac{\Gamma\left(k+\frac{1}{2}\right)\Gamma\left(n+i+k-\frac{1}{2}\right)}{\Gamma\left(k+\frac{3}{2}\right)\Gamma(k+1)}\left(\frac{t}{\sqrt{2n-2}}\right)\left(-\frac{t^2}{2n-2}\right)^k\frac{\Gamma\left(j+k+\frac{1}{2}\right)}{\Gamma\left(k+\frac{1}{2}\right)\Gamma\left(j+\frac{1}{2}\right)\Gamma(j+1)}\left(-\frac{\delta_\rho^2}{2}\right)^j\right.\\
&\quad\left.-\frac{\Gamma(n+i+k)}{\Gamma(k+2)}\left(\frac{\delta_\rho}{\sqrt{2}}\right)\left(-\frac{t^2}{2n-2}\right)^{k+1}\frac{\Gamma\left(j+k+\frac{3}{2}\right)}{\Gamma\left(k+\frac{3}{2}\right)\Gamma\left(j+\frac{3}{2}\right)\Gamma(j+1)}\left(-\frac{\delta_\rho^2}{2}\right)^j\right]
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2}-\frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{2\,\Gamma\left(\frac{n-1}{2}\right)}\sum_{i=0}^{\infty}\frac{\Gamma\left(\frac{n-1}{2}+i\right)\left(\frac{2\rho}{1+\rho}\right)^i}{\Gamma(i+1)}\times\\
&\left[\sum_{j=0}^{\infty}\frac{\Gamma\left(j+\frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(j+\frac{3}{2}\right)\Gamma(j+1)}\left(\frac{\delta_\rho}{\sqrt{2}}\right)\left(-\frac{\delta_\rho^2}{2}\right)^j\right. \qquad \text{(A)}\\
&-\sum_{j,k=0}^{\infty}\frac{\Gamma\left(n+i+k-\frac{1}{2}\right)\Gamma\left(j+k+\frac{1}{2}\right)}{\Gamma(n+i-1)\,\Gamma\left(j+\frac{1}{2}\right)\Gamma\left(k+\frac{3}{2}\right)\Gamma(j+1)\,\Gamma(k+1)}\left(\frac{t}{\sqrt{2n-2}}\right)\left(\frac{-\delta_\rho^2}{2}\right)^j\left(\frac{-t^2}{2n-2}\right)^k\\
&\left.+\sum_{j,k=0}^{\infty}\frac{\Gamma(n+i+k)\,\Gamma\left(j+k+\frac{3}{2}\right)}{\Gamma(n+i-1)\,\Gamma\left(j+\frac{3}{2}\right)\Gamma\left(k+\frac{3}{2}\right)\Gamma(k+2)\,\Gamma(j+1)}\left(\frac{\delta_\rho}{\sqrt{2}}\right)\left(\frac{-\delta_\rho^2}{2}\right)^j\left(\frac{-t^2}{2n-2}\right)^{k+1}\right].
\end{aligned}$$

It can be seen that term (A) is equal to the last term evaluated at $k=-1$; hence the series may be combined

$$F_{T_2}(t) = \frac{1}{2} - \frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{2\,\Gamma(\frac{n-1}{2})} \sum_{i=0}^{\infty} \frac{\Gamma(\frac{n-1}{2}+i)\,\left(\frac{2\rho}{1+\rho}\right)^i}{\Gamma(i+1)} \times$$
$$\left[\sum_{\substack{j=0\\k=-1}}^{\infty} \frac{\Gamma(n+i+k)\,\Gamma(j+k+\frac{3}{2})}{\Gamma(n+i-1)\,\Gamma(j+\frac{3}{2})\,\Gamma(k+\frac{3}{2})\,\Gamma(k+2)\,\Gamma(j+1)} \left(\frac{\delta_\rho}{\sqrt{2}}\right)\left(\frac{-\delta_\rho^2}{2}\right)^j\left(\frac{-t^2}{2n-2}\right)^{k+1}\right.$$
$$\left. - \sum_{j,k=0}^{\infty} \frac{\Gamma(n+i+k-\frac{1}{2})\,\Gamma(j+k+\frac{1}{2})}{\Gamma(n+i-1)\,\Gamma(j+\frac{1}{2})\,\Gamma(k+\frac{3}{2})\,\Gamma(j+1)\,\Gamma(k+1)} \left(\frac{t}{\sqrt{2n-2}}\right)\left(\frac{-\delta_\rho^2}{2}\right)^j\left(\frac{-t^2}{2n-2}\right)^{k}\right]$$
$$= \frac{1}{2} + \frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{2\,\Gamma(\frac{n-1}{2})} \sum_{i=0}^{\infty} \frac{\Gamma(\frac{n-1}{2}+i)\,\left(\frac{2\rho}{1+\rho}\right)^i}{\Gamma(i+1)} \sum_{j,k=0}^{\infty}$$
$$\left[\frac{\Gamma(n+i+k-\frac{1}{2})\,\Gamma(j+k+\frac{1}{2})}{\Gamma(n+i-1)\,\Gamma(j+\frac{1}{2})\,\Gamma(k+\frac{3}{2})\,\Gamma(j+1)\,\Gamma(k+1)} \left(\frac{t}{\sqrt{2n-2}}\right)\left(\frac{-\delta_\rho^2}{2}\right)^j\left(\frac{-t^2}{2n-2}\right)^{k}\right.$$
$$\left. - \frac{\Gamma(n+i+k-1)\,\Gamma(j+k+\frac{1}{2})}{\Gamma(n+i-1)\,\Gamma(j+\frac{3}{2})\,\Gamma(k+\frac{1}{2})\,\Gamma(j+1)\,\Gamma(k+1)} \left(\frac{\delta_\rho}{\sqrt{2}}\right)\left(\frac{-\delta_\rho^2}{2}\right)^j\left(\frac{-t^2}{2n-2}\right)^{k}\right]. \qquad \text{(A.2)}$$

Diagonalizing the j,k infinite sums by $\sum_{i,j=0}^{\infty} c(i,j)\,x^i\,y^j = \sum_{i=0}^{\infty}\sum_{j=0}^{i} c(i-j,j)\,x^{i-j}\,y^j$ gives

$$F_{T_2}(t) = \frac{1}{2} + \frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{2\,\Gamma(\frac{n-1}{2})} \sum_{i=0}^{\infty} \frac{\Gamma(\frac{n-1}{2}+i)\,\left(\frac{2\rho}{1+\rho}\right)^i}{\Gamma(i+1)} \sum_{k=0}^{\infty}\sum_{j=0}^{k}$$
$$\left[\frac{\Gamma(n+i+k-j-\frac{1}{2})\,\Gamma(k+\frac{1}{2})}{\Gamma(n+i-1)\,\Gamma(j+\frac{1}{2})\,\Gamma(k-j+\frac{3}{2})\,\Gamma(j+1)\,\Gamma(k-j+1)} \left(\frac{t}{\sqrt{2n-2}}\right)\left(\frac{-\delta_\rho^2}{2}\right)^j\left(\frac{-t^2}{2n-2}\right)^{k-j}\right.$$
$$\left. - \frac{\Gamma(n+i+k-j-1)\,\Gamma(k+\frac{1}{2})}{\Gamma(n+i-1)\,\Gamma(j+\frac{3}{2})\,\Gamma(k-j+\frac{1}{2})\,\Gamma(j+1)\,\Gamma(k-j+1)} \left(\frac{\delta_\rho}{\sqrt{2}}\right)\left(\frac{-\delta_\rho^2}{2}\right)^j\left(\frac{-t^2}{2n-2}\right)^{k-j}\right]$$
$$= \frac{1}{2} + \frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{2\,\Gamma(\frac{n-1}{2})} \sum_{i=0}^{\infty} \frac{\Gamma(\frac{n-1}{2}+i)\,\left(\frac{2\rho}{1+\rho}\right)^i}{\Gamma(i+1)} \sum_{k=0}^{\infty}\sum_{j=0}^{k}$$
$$\left[\frac{\Gamma(n+i+k-j-\frac{1}{2})\,\Gamma(k+\frac{1}{2})\,(-1)^k}{\Gamma(n+i-1)\,\Gamma(j+\frac{1}{2})\,\Gamma(k-j+\frac{3}{2})\,\Gamma(j+1)\,\Gamma(k-j+1)} \left(\frac{\delta_\rho^2}{2}\right)^j\left(\frac{t^2}{2n-2}\right)^{k-j+\frac{1}{2}}\right.$$
$$\left. - \frac{\Gamma(n+i+k-j-1)\,\Gamma(k+\frac{1}{2})\,(-1)^k}{\Gamma(n+i-1)\,\Gamma(j+1)\,\Gamma(k-j+1)\,\Gamma(j+\frac{3}{2})\,\Gamma(k-j+\frac{1}{2})} \left(\frac{\delta_\rho^2}{2}\right)^{j+\frac{1}{2}}\left(\frac{t^2}{2n-2}\right)^{k-j}\right].$$

Note that wherever j occurs in the first term, $j+\frac{1}{2}$ occurs in the second term; hence the j series may be combined

$$F_{T_2}(t) = \frac{1}{2} + \frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{2\,\Gamma\left(\frac{n-1}{2}\right)} \sum_{i=0}^{\infty} \frac{\Gamma\left(\frac{n-1}{2}+i\right)\left(\frac{2\rho}{1+\rho}\right)^i}{\Gamma(i+1)} \sum_{k=0}^{\infty}\sum_{j=0}^{2k+1}$$
$$\frac{\Gamma\left(n+i+k-\frac{j}{2}-\frac{1}{2}\right)\Gamma\left(k+\frac{1}{2}\right)(-1)^k(-1)^j}{\Gamma(n+i-1)\,\Gamma\left(\frac{j}{2}+\frac{1}{2}\right)\Gamma\left(k-\frac{j}{2}+\frac{3}{2}\right)\Gamma\left(\frac{j}{2}+1\right)\Gamma\left(k-\frac{j}{2}+1\right)}\left(\frac{\delta_\rho^2}{2}\right)^{\frac{j}{2}}\left(\frac{t^2}{2n-2}\right)^{k-\frac{j}{2}+\frac{1}{2}}.$$

Applying the identity $\Gamma(x)\,\Gamma\left(x+\frac{1}{2}\right) = \frac{\Gamma(2x)\,\Gamma\left(\frac{1}{2}\right)}{2^{2x-1}}$, reversing the indexing of the j series, and for clarity switching the j and k indices gives

$$F_{T_2}(t) = \frac{1}{2} + \frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{\sqrt{2}\,\pi\,\Gamma\left(\frac{n-1}{2}\right)} \times \tag{A.3}$$
$$\sum_{i,j=0}^{\infty}\sum_{k=0}^{2j+1} \frac{\Gamma\left(\frac{n-1}{2}+i\right)\Gamma\left(n+i-1+\frac{k}{2}\right)\Gamma\left(j+\frac{1}{2}\right)(-2)^j}{\Gamma(n+i-1)\,\Gamma(2j-k+2)\,\Gamma(i+1)\,\Gamma(k+1)}\left(\frac{2\rho}{1+\rho}\right)^i\left(-\delta_\rho\right)^{2j-k+1}\left(\frac{t}{\sqrt{n-1}}\right)^k.$$

The above expression will not converge for $\rho \le -\frac{1}{3}$. An expression for the distribution function which converges for all values of the parameters is derived in the following section.

A.2 The distribution function of $T_2|\rho$ as a guaranteed convergent power series suitable for computing

From A.1 we have

$$F_{T_2}(t) = \frac{1}{2} - \frac{\delta_\rho}{\sqrt{2\pi}}M\left(\frac{1}{2},\frac{3}{2};-\frac{\delta_\rho^2}{2}\right) + \frac{e^{-\frac{\delta_\rho^2}{2}}\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{2\,\Gamma\left(\frac{n-1}{2}\right)} \sum_{i,j,k=0}^{\infty} \frac{\Gamma\left(\frac{n-1}{2}+i\right)\left(\frac{2\rho}{1+\rho}\right)^i}{\Gamma(n+i-1)\,\Gamma(i+1)} \times$$
$$\left[\frac{\Gamma\left(j+k+\frac{1}{2}\right)\Gamma\left(n+i+j+k-\frac{1}{2}\right)(-1)^k}{\Gamma\left(j+k+\frac{3}{2}\right)\Gamma\left(j+\frac{1}{2}\right)\Gamma(j+1)\,\Gamma(k+1)}\left(\frac{\delta_\rho^2}{2}\right)^j\left(\frac{t^2}{2n-2}\right)^{j+k+\frac{1}{2}}\right.$$
$$\left.+\frac{\Gamma(j+k+1)\,\Gamma(n+i+j+k)(-1)^k}{\Gamma(j+k+2)\,\Gamma\left(j+\frac{3}{2}\right)\Gamma(j+1)\,\Gamma(k+1)}\left(\frac{\delta_\rho^2}{2}\right)^{j+\frac{1}{2}}\left(\frac{t^2}{2n-2}\right)^{j+k+1}\right].$$

Note that wherever j occurs in the first term, $j+\frac{1}{2}$ occurs in the second term; hence the j series may be combined

$$F_{T_2}(t) = \frac{1}{2} - \frac{\delta_\rho}{\sqrt{2\pi}}M\left(\frac{1}{2},\frac{3}{2};-\frac{\delta_\rho^2}{2}\right) + \frac{e^{-\frac{\delta_\rho^2}{2}}\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{2\,\Gamma\left(\frac{n-1}{2}\right)} \sum_{i,j,k=0}^{\infty} \frac{\Gamma\left(\frac{n-1}{2}+i\right)\left(\frac{2\rho}{1+\rho}\right)^i}{\Gamma(n+i-1)\,\Gamma(i+1)} \times$$

$$\frac{\Gamma\left(\frac{j}{2}+k+\frac{1}{2}\right)\Gamma\left(n+i+\frac{j}{2}+k-\frac{1}{2}\right)(-1)^k}{\Gamma\left(\frac{j}{2}+k+\frac{3}{2}\right)\Gamma\left(\frac{j}{2}+\frac{1}{2}\right)\Gamma(\frac{j}{2}+1)\Gamma(k+1)}\left(\frac{\delta_\rho^2}{2}\right)^{\frac{j}{2}}\left(\frac{t^2}{2n-2}\right)^{\frac{j}{2}+k+\frac{1}{2}}.$$

Since $\Gamma(x)\,\Gamma(x+\frac{1}{2}) = \dfrac{\Gamma(2x)\,\Gamma(\frac{1}{2})}{2^{2x-1}}$

$$F_{T_2}(t) = \frac{1}{2} - \frac{\delta_\rho}{\sqrt{2\pi}}M\left(\tfrac{1}{2},\tfrac{3}{2};-\tfrac{\delta_\rho^2}{2}\right) + \frac{e^{-\frac{\delta_\rho^2}{2}}\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{2\sqrt{\pi}\,\Gamma\left(\frac{n-1}{2}\right)}\sum_{i,j,k=0}^{\infty}\frac{\Gamma\left(\frac{n-1}{2}+i\right)}{\Gamma(n+i-1)\,\Gamma(i+1)}\left(\frac{2\rho}{1+\rho}\right)^i \times$$

$$\frac{\Gamma\left(\frac{j}{2}+k+\frac{1}{2}\right)\Gamma\left(n+i+\frac{j-1}{2}+k\right)}{\Gamma\left(\frac{j}{2}+k+\frac{3}{2}\right)\Gamma(j+1)\Gamma(k+1)}\left(\frac{t\,\delta_\rho}{\sqrt{n-1}}\right)^j\left(\frac{-t^2}{2n-2}\right)^k\left(\frac{t}{\sqrt{2n-2}}\right).$$

We transform the i series. Express the i series as a Gauss hypergeometric function (see 2.4)

$$\sum_{i=0}^{\infty}\frac{\Gamma\left(\frac{n-1}{2}+i\right)\Gamma\left(n+i+\frac{j-1}{2}+k\right)}{\Gamma(n+i-1)\,\Gamma(i+1)}\left(\frac{2\rho}{1+\rho}\right)^i$$
$$= \frac{\Gamma\left(\frac{n-1}{2}\right)\Gamma\left(n+\frac{j-1}{2}+k\right)}{\Gamma(n-1)}H\left(\tfrac{n-1}{2},n+\tfrac{j-1}{2}+k,n-1;\tfrac{2\rho}{1+\rho}\right).$$

Apply Erdélyi's [12] transformation(4) p.111; express the resulting Gauss hypergeometric function as a power series

$$= \frac{\Gamma\left(\frac{n-1}{2}\right)\Gamma\left(n+\frac{j-1}{2}+k\right)}{\Gamma(n-1)}\left(1-\frac{\rho}{1+\rho}\right)^{-(n+\frac{j-1}{2}+k)}H\left(\frac{n+\frac{j-1}{2}+k}{2},\frac{n+\frac{j-1}{2}+k}{2}+\frac{1}{2},\frac{n}{2};\rho^2\right)$$

$$= \frac{\Gamma\left(\frac{n-1}{2}\right)\Gamma\left(n+\frac{j-1}{2}+k\right)}{\Gamma(n-1)}(1+\rho)^{n+\frac{j-1}{2}+k}\frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n+\frac{j-1}{2}+k}{2}\right)\Gamma\left(\frac{n+\frac{j-1}{2}+k}{2}+\frac{1}{2}\right)}$$

$$\times\sum_{i=0}^{\infty}\frac{\Gamma\left(\frac{n+\frac{j-1}{2}+k}{2}+i\right)\Gamma\left(\frac{n+\frac{j-1}{2}+k}{2}+\frac{1}{2}+i\right)}{\Gamma\left(\frac{n}{2}+i\right)\Gamma(i+1)}\left(\rho^2\right)^i.$$

Since $\Gamma(x)\,\Gamma(x+\frac{1}{2}) = \dfrac{\Gamma(2x)\,\Gamma(\frac{1}{2})}{2^{2x-1}}$

$$= \frac{\Gamma(n-1)\,\Gamma(\frac{1}{2})\,\Gamma\left(n+\frac{j-1}{2}+k\right)}{2^{n-1-1}\Gamma(n-1)}(1+\rho)^{n+\frac{j-1}{2}+k}\frac{2^{n+\frac{j-3}{2}+k}}{\sqrt{\pi}\,\Gamma\left(n+\frac{j-1}{2}+k\right)}$$

$$\times\sum_{i=0}^{\infty}\frac{\sqrt{\pi}\,\Gamma\left(n+\frac{j-1}{2}+k+2i\right)}{2^{n+\frac{j-3}{2}+k+2i}\,\Gamma\left(\frac{n}{2}+i\right)\Gamma(i+1)}\left(\rho^2\right)^i$$

$$= 2\sqrt{\pi}\left(\tfrac{1+\rho}{2}\right)^{n-1}(1+\rho)^{\frac{j+1}{2}+k}\sum_{i=0}^{\infty}\frac{\Gamma\left(n+\frac{j-1}{2}+k+2i\right)}{\Gamma(\frac{n}{2}+i)\,\Gamma(i+1)}\left(\tfrac{\rho^2}{4}\right)^{i}.$$

Thus

$$F_{T_2}(t) = \frac{1}{2}-\frac{\delta_\rho}{\sqrt{2\pi}}M\left(\tfrac{1}{2},\tfrac{3}{2};-\tfrac{\delta_\rho^2}{2}\right)+\frac{e^{-\frac{\delta_\rho^2}{2}}\left(\frac{1-\rho^2}{4}\right)^{\frac{n-1}{2}}}{\Gamma(\frac{n-1}{2})}\times$$

$$\sum_{i,j,k=0}^{\infty}\frac{\Gamma\left(\frac{j}{2}+k+\frac{1}{2}\right)\Gamma\left(n+\frac{j-1}{2}+k+2i\right)}{\Gamma\left(\frac{j}{2}+k+\frac{3}{2}\right)\Gamma(\frac{n}{2}+i)\,\Gamma(i+1)\Gamma(j+1)\Gamma(k+1)}\left(\tfrac{\rho^2}{4}\right)^{i}\left(\tfrac{t\delta_\rho\sqrt{1+\rho}}{\sqrt{n-1}}\right)^{j}\left(\tfrac{-t^2(1+\rho)}{2n-2}\right)^{k}\left(\tfrac{t\sqrt{1+\rho}}{\sqrt{2n-2}}\right).$$

We transform the k series

$$\sum_{k=0}^{\infty}\frac{\Gamma\left(\frac{j}{2}+k+\frac{1}{2}\right)\Gamma\left(n+\frac{j-1}{2}+k+2i\right)}{\Gamma\left(\frac{j}{2}+k+\frac{3}{2}\right)\Gamma(k+1)}\left(\tfrac{-t^2(1+\rho)}{2n-2}\right)^{k}$$

$$= \sum_{k=0}^{\infty}\frac{\Gamma\left(\frac{j+1}{2}+k\right)\,\Gamma\left(n+\frac{j-1}{2}+2i+k\right)}{\Gamma\left(\frac{j+3}{2}+k\right)\,\Gamma(k+1)}\left(\tfrac{-t^2(1+\rho)}{2(n-1)}\right)^{k}$$

$$= \frac{\Gamma\left(\frac{j+1}{2}\right)\,\Gamma\left(n+\frac{j-1}{2}+2i\right)}{\Gamma\left(\frac{j+3}{2}\right)}H\left(\tfrac{j+1}{2},n+\tfrac{j-1}{2}+2i,\tfrac{j+3}{2};\tfrac{-t^2(1+\rho)}{2(n-1)}\right)$$

$$= \frac{\Gamma\left(\frac{j+1}{2}\right)\,\Gamma\left(n+\frac{j-1}{2}+2i\right)}{\Gamma\left(\frac{j+3}{2}\right)}\left(1-\tfrac{-t^2(1+\rho)}{2(n-1)}\right)^{-(n+\frac{j-1}{2}+2i)}H\left(1,n+\tfrac{j-1}{2}+2i,\tfrac{j+3}{2};\frac{\frac{-t^2(1+\rho)}{2(n-1)}}{\frac{-t^2(1+\rho)}{2(n-1)}-1}\right)$$

(see Erdélyi [12], transformation(4) p.105);

$$= \frac{\Gamma\left(\frac{j+1}{2}\right)\,\Gamma\left(n+\frac{j-1}{2}+2i\right)}{\Gamma\left(\frac{j+3}{2}\right)}\left(\tfrac{2(n-1)}{2(n-1)+t^2(1+\rho)}\right)^{n+\frac{j-1}{2}+2i}H\left(1,n+\tfrac{j-1}{2}+2i,\tfrac{j+3}{2};\tfrac{t^2(1+\rho)}{t^2(1+\rho)+2(n-1)}\right)$$

$$= \Gamma\left(\tfrac{j+1}{2}\right)\left(\tfrac{2(n-1)}{2(n-1)+t^2(1+\rho)}\right)^{n+\frac{j-1}{2}+2i}\sum_{k=0}^{\infty}\frac{\Gamma\left(n+\frac{j-1}{2}+2i+k\right)}{\Gamma\left(\frac{j+3}{2}+k\right)}\left(\tfrac{t^2(1+\rho)}{t^2(1+\rho)+2(n-1)}\right)^{k}.$$

Thus

$$F_{T_2}(t) = \frac{1}{2}-\frac{\delta_\rho}{\sqrt{2\pi}}M\left(\tfrac{1}{2},\tfrac{3}{2};-\tfrac{\delta_\rho^2}{2}\right)+\frac{e^{-\frac{\delta_\rho^2}{2}}}{\Gamma(\frac{n-1}{2})}\left(\tfrac{(n-1)\sqrt{1-\rho^2}}{2(n-1)+t^2(1+\rho)}\right)^{n-1}$$

$$\times\sum_{i,j,k=0}^{\infty}\frac{\Gamma\left(n+\frac{j-1}{2}+2i+k\right)\Gamma\left(\frac{j+1}{2}\right)}{\Gamma(\frac{n}{2}+i)\,\Gamma\left(\frac{j+3}{2}+k\right)\Gamma(i+1)\Gamma(j+1)}\left(\tfrac{\rho(n-1)}{2(n-1)+t^2(1+\rho)}\right)^{2i}$$

$$\times\left(\tfrac{t\delta_\rho\sqrt{2(1+\rho)}}{\sqrt{2(n-1)+t^2(1+\rho)}}\right)^{j}\left(\tfrac{t\sqrt{1+\rho}}{\sqrt{2(n-1)+t^2(1+\rho)}}\right)\left(\tfrac{t^2(1+\rho)}{2(n-1)+t^2(1+\rho)}\right)^{k}. \quad \text{(A.4)}$$

The above distribution function will always converge and hence is suitable for computer use.

A.3 The Density of $T_2|\rho$

Any of the distribution functions from A.1 may be differentiated to give a density function. A convenient form may be derived as follows. At (A.1) we have

$$F_{T_2}(t) = \frac{1}{2} - \frac{\delta_\rho}{\sqrt{2\pi}} M\left(\tfrac{1}{2}, \tfrac{3}{2}; -\tfrac{\delta_\rho^2}{2}\right) + \frac{e^{-\frac{\delta_\rho^2}{2}}\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{2\,\Gamma\left(\frac{n-1}{2}\right)} \sum_{i,j,k=0}^{\infty} \frac{\Gamma\left(\frac{n-1}{2}+i\right)\left(\frac{2\rho}{1+\rho}\right)^i}{\Gamma(n+i-1)\,\Gamma(i+1)} \times$$
$$\left[\frac{\Gamma\left(j+k+\frac{1}{2}\right)\Gamma\left(n+i+j+k-\frac{1}{2}\right)(-1)^k}{\Gamma\left(j+k+\frac{3}{2}\right)\Gamma\left(j+\frac{1}{2}\right)\Gamma(j+1)\Gamma(k+1)}\left(\frac{\delta_\rho^2}{2}\right)^j\left(\frac{t^2}{2n-2}\right)^{j+k+\frac{1}{2}}\right.$$
$$\left.+\frac{\Gamma(j+k+1)\Gamma(n+i+j+k)(-1)^k}{\Gamma(j+k+2)\Gamma\left(j+\frac{3}{2}\right)\Gamma(j+1)\Gamma(k+1)}\left(\frac{\delta_\rho^2}{2}\right)^{j+\frac{1}{2}}\left(\frac{t^2}{2n-2}\right)^{j+k+1}\right].$$

Differentiating with respect to t gives

$$f_{T_2}(t) = \frac{e^{-\frac{\delta_\rho^2}{2}}\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{2\,\Gamma\left(\frac{n-1}{2}\right)} \sum_{i,j,k=0}^{\infty} \frac{\Gamma\left(\frac{n-1}{2}+i\right)\left(\frac{2\rho}{1+\rho}\right)^i}{\Gamma(n+i-1)\,\Gamma(i+1)} \times$$
$$\left[\frac{\left(j+k+\frac{1}{2}\right)\Gamma\left(j+k+\frac{1}{2}\right)\Gamma\left(n+i+j+k-\frac{1}{2}\right)(-1)^k}{\Gamma\left(j+k+\frac{3}{2}\right)\Gamma\left(j+\frac{1}{2}\right)\Gamma(j+1)\Gamma(k+1)}\left(\frac{\delta_\rho^2}{2}\right)^j\left(\frac{t^2}{2n-2}\right)^{j+k-\frac{1}{2}}\left(\frac{2t}{2n-2}\right)\right.$$
$$\left.+\frac{(j+k+1)\Gamma(j+k+1)\Gamma(n+i+j+k)(-1)^k}{\Gamma(j+k+2)\Gamma\left(j+\frac{3}{2}\right)\Gamma(j+1)\Gamma(k+1)}\left(\frac{\delta_\rho^2}{2}\right)^{j+\frac{1}{2}}\left(\frac{t^2}{2n-2}\right)^{j+k}\left(\frac{2t}{2n-2}\right)\right]$$
$$= \frac{e^{-\frac{\delta_\rho^2}{2}}\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{\sqrt{2n-2}\,\Gamma\left(\frac{n-1}{2}\right)} \sum_{i,j,k=0}^{\infty} \frac{\Gamma\left(\frac{n-1}{2}+i\right)\left(\frac{2\rho}{1+\rho}\right)^i}{\Gamma(n+i-1)\,\Gamma(i+1)} \times$$
$$\left[\frac{\Gamma\left(n+i+j+k-\frac{1}{2}\right)(-1)^k}{\Gamma\left(j+\frac{1}{2}\right)\Gamma(j+1)\Gamma(k+1)}\left(\frac{\delta_\rho^2}{2}\right)^j\left(\frac{t^2}{2n-2}\right)^{j+k}\right.$$
$$\left.+\frac{\Gamma(n+i+j+k)(-1)^k}{\Gamma\left(j+\frac{3}{2}\right)\Gamma(j+1)\Gamma(k+1)}\left(\frac{\delta_\rho^2}{2}\right)^{j+\frac{1}{2}}\left(\frac{t^2}{2n-2}\right)^{j+k+\frac{1}{2}}\left(\frac{t^2}{2n-2}\right)^{j+k+\frac{1}{2}}\right].$$

Note that wherever j occurs in the first term, $j+\frac{1}{2}$ occurs in the second term; hence the j series may be combined

$$f_{T_2}(t) = \frac{e^{-\frac{\delta_\rho^2}{2}}\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{\sqrt{2n-2}\,\Gamma\left(\frac{n-1}{2}\right)} \sum_{i,j,k=0}^{\infty} \frac{\Gamma\left(\frac{n-1}{2}+i\right)\left(\frac{2\rho}{1+\rho}\right)^i}{\Gamma(n+i-1)\,\Gamma(i+1)} \times$$

$$\frac{\Gamma\left(n+i+\frac{j}{2}+k-\frac{1}{2}\right)(-1)^k}{\Gamma\left(\frac{j}{2}+\frac{1}{2}\right)\Gamma(\frac{j}{2}+1)\,\Gamma(k+1)}\left(\frac{\delta_\rho^2}{2}\right)^{\frac{j}{2}}\left(\frac{t^2}{2n-2}\right)^{\frac{j}{2}+k} .$$

Since $\Gamma(x)\,\Gamma(x+\frac{1}{2}) = \dfrac{\Gamma(2x)\,\Gamma(\frac{1}{2})}{2^{2x-1}}$

$$\begin{aligned} f_{T_2}(t) &= \frac{e^{-\frac{\delta_\rho^2}{2}}\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{\sqrt{2n-2}\,\Gamma(\frac{n-1}{2})}\sum_{i,j,k=0}^{\infty}\frac{\Gamma(\frac{n-1}{2}+i)\,\left(\frac{2\rho}{1+\rho}\right)^i}{\Gamma(n+i-1)\,\Gamma(i+1)} \times \\ &\quad \frac{\Gamma\left(n+i+\frac{j}{2}+k-\frac{1}{2}\right)(-1)^k\,2^j}{\Gamma(j+1)\,\Gamma(\frac{1}{2})\,\Gamma(k+1)}\left(\frac{\delta_\rho^2}{2}\right)^{\frac{j}{2}}\left(\frac{t^2}{2n-2}\right)^{\frac{j}{2}+k} \\ &= \frac{e^{-\frac{\delta_\rho^2}{2}}\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{\sqrt{2\pi(n-1)}\,\Gamma(\frac{n-1}{2})}\sum_{i,j,k=0}^{\infty}\frac{\Gamma(\frac{n-1}{2}+i)\,\Gamma\left(n+i+\frac{j-1}{2}+k\right)\left(\frac{2\rho}{1+\rho}\right)^i}{\Gamma(n+i-1)\,\Gamma(i+1)\,\Gamma(j+1)\,\Gamma(k+1)}\left(\frac{t\delta_\rho}{\sqrt{n-1}}\right)^j\left(\frac{-t^2}{2n-2}\right)^k . \end{aligned}$$

Note that $\sum_{i=0}^{\infty}\dfrac{\Gamma(\alpha+i)}{\Gamma(i+1)}x^i \;=\; \Gamma(\alpha)\,(1-x)^{-\alpha}$; hence

$$\begin{aligned} f_{T_2}(t) &= \frac{e^{-\frac{\delta_\rho^2}{2}}\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{\sqrt{2\pi(n-1)}\,\Gamma(\frac{n-1}{2})}\sum_{i,j=0}^{\infty}\frac{\Gamma(\frac{n-1}{2}+i)\,\Gamma\left(n+i+\frac{j-1}{2}\right)\left(\frac{2\rho}{1+\rho}\right)^i}{\Gamma(n+i-1)\,\Gamma(i+1)\,\Gamma(j+1)}\left(\frac{t\delta_\rho}{\sqrt{n-1}}\right)^j \\ &\qquad \times\left(1-\frac{-t^2}{2n-2}\right)^{-(n+i+\frac{j-1}{2})} \\ &= \frac{e^{-\frac{\delta_\rho^2}{2}}\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}\left(\frac{2n-2}{2n-2+t^2}\right)^{n-\frac{1}{2}}}{\sqrt{2\pi(n-1)}\,\Gamma(\frac{n-1}{2})} \\ &\quad \times\sum_{i,j=0}^{\infty}\frac{\Gamma(\frac{n-1}{2}+i)\,\Gamma\left(n+i+\frac{j-1}{2}\right)}{\Gamma(n+i-1)\,\Gamma(i+1)\,\Gamma(j+1)}\left(\frac{4\rho(n-1)}{(1+\rho)(2n-2+t^2)}\right)^i\left(\frac{t\delta_\rho\sqrt{2}}{\sqrt{2n-2+t^2}}\right)^j . \end{aligned}$$

The j series is confluent and will always converge. When t is small, convergence of the i series will be problematic for large $|\rho|$. Convergence may be assured by transforming the i series.

Express the i series as a Gauss hypergeometric function (see 2.4)

$$\begin{aligned} &\sum_{i=0}^{\infty}\frac{\Gamma(\frac{n-1}{2}+i)\,\Gamma\left(n+i+\frac{j-1}{2}\right)}{\Gamma(n+i-1)\,\Gamma(i+1)}\left(\frac{4\rho(n-1)}{(1+\rho)(2n-2+t^2)}\right)^i \\ &\qquad = \frac{\Gamma(\frac{n-1}{2})\,\Gamma\left(n+\frac{j-1}{2}\right)}{\Gamma(n-1)}\,H\left(\frac{n-1}{2},n+\frac{j-1}{2},n-1;\frac{4\rho(n-1)}{(1+\rho)(2n-2+t^2)}\right) . \end{aligned}$$

Apply Erdélyi's [12] transformation(4) p.111; express the resulting Gauss hypergeometric function as a power series

$$= \frac{\Gamma\left(\frac{n-1}{2}\right)\Gamma\left(n+\frac{j-1}{2}\right)}{\Gamma(n-1)} \left(1-\frac{2\rho(n-1)}{(1+\rho)(2n-2+t^2)}\right)^{-(n+\frac{j-1}{2})} H\left(\frac{n+\frac{j-1}{2}}{2}, \frac{n+\frac{j-1}{2}}{2}+\frac{1}{2}, \frac{n}{2}; \left(\frac{\rho(2n-2)}{2n-2+t^2(1+\rho)}\right)^2\right)$$

$$= \frac{\Gamma\left(\frac{n-1}{2}\right)\Gamma\left(n+\frac{j-1}{2}\right)}{\Gamma(n-1)} \left(\frac{(1+\rho)(2n-2+t^2)}{2n-2+t^2(1+\rho)}\right)^{n+\frac{j-1}{2}} \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n+\frac{j-1}{2}}{2}\right)\Gamma\left(\frac{n+\frac{j-1}{2}}{2}+\frac{1}{2}\right)}$$

$$\times \sum_{i=0}^{\infty} \frac{\Gamma\left(\frac{n+\frac{j-1}{2}}{2}+i\right)\Gamma\left(\frac{n+\frac{j-1}{2}}{2}+\frac{1}{2}+i\right)}{\Gamma\left(\frac{n}{2}+i\right)\Gamma(i+1)} \left(\frac{\rho(2n-2)}{2n-2+t^2(1+\rho)}\right)^{2i}.$$

Since $\Gamma(x)\,\Gamma(x+\frac{1}{2}) = \dfrac{\Gamma(2x)\,\Gamma(\frac{1}{2})}{2^{2x-1}}$

$$= \frac{\Gamma(n-1)\,\Gamma\left(\frac{1}{2}\right)\Gamma\left(n+\frac{j-1}{2}\right)}{2^{n-1-1}\Gamma(n-1)} \left(\frac{(1+\rho)(2n-2+t^2)}{2n-2+t^2(1+\rho)}\right)^{n+\frac{j-1}{2}} \frac{2^{n+\frac{j-3}{2}}}{\sqrt{\pi}\,\Gamma\left(n+\frac{j-1}{2}\right)}$$

$$\times \sum_{i=0}^{\infty} \frac{\sqrt{\pi}\,\Gamma\left(n+\frac{j-1}{2}+2i\right)}{2^{n+\frac{j-3}{2}+2i}\,\Gamma\left(\frac{n}{2}+i\right)\Gamma(i+1)} \left(\frac{\rho(2n-2)}{2n-2+t^2(1+\rho)}\right)^{2i}$$

$$= 2\sqrt{\pi} \left(\frac{(1+\rho)(2n-2+t^2)}{2(2n-2+t^2(1+\rho))}\right)^{n-1} \left(\frac{(1+\rho)(2n-2+t^2)}{2n-2+t^2(1+\rho)}\right)^{\frac{j+1}{2}} \sum_{i=0}^{\infty} \frac{\Gamma\left(n+\frac{j-1}{2}+2i\right)}{\Gamma\left(\frac{n}{2}+i\right)\Gamma(i+1)} \left(\frac{\rho(2n-2)}{2(2n-2+t^2(1+\rho))}\right)^{2i}.$$

Thus

$$f_{T_2}(t) = \frac{2e^{-\frac{\delta_\rho^2}{2}}\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}\left(\frac{(1+\rho)(n-1)}{2n-2+t^2(1+\rho)}\right)^{n-\frac{1}{2}}}{\sqrt{n-1}\,\Gamma\left(\frac{n-1}{2}\right)}$$

$$\times \sum_{i,j=0}^{\infty} \frac{\Gamma\left(n+\frac{j-1}{2}+2i\right)}{\Gamma\left(\frac{n}{2}+i\right)\Gamma(i+1)\Gamma(j+1)} \left(\frac{\rho(n-1)}{2n-2+t^2(1+\rho)}\right)^{2i} \left(\frac{t\delta_\rho\sqrt{2(1+\rho)}}{\sqrt{2n-2+t^2(1+\rho)}}\right)^{j}. \quad \text{(A.5)}$$

A.4 The Asymptotic Distribution of $T_2|\rho$

From (A.2) we have

$$F_{T_2}(t) = \frac{1}{2} + \frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{2\Gamma\left(\frac{n-1}{2}\right)} \sum_{i=0}^{\infty} \frac{\Gamma\left(\frac{n-1}{2}+i\right)\left(\frac{2\rho}{1+\rho}\right)^{i}}{\Gamma(i+1)} \sum_{j,k=0}^{\infty}$$

$$\left[\frac{\Gamma(n+i+k-\frac{1}{2})\,\Gamma(j+k+\frac{1}{2})}{\Gamma(n+i-1)\,\Gamma(j+\frac{1}{2})\,\Gamma(k+\frac{3}{2})\,\Gamma(j+1)\,\Gamma(k+1)}\left(\frac{t}{\sqrt{2n-2}}\right)\left(\frac{-\delta_\rho^2}{2}\right)^j\left(\frac{-t^2}{2n-2}\right)^k\right.$$
$$\left.-\frac{\Gamma(n+i+k-1)\,\Gamma(j+k+\frac{1}{2})}{\Gamma(n+i-1)\,\Gamma(j+\frac{3}{2})\,\Gamma(k+\frac{1}{2})\,\Gamma(j+1)\,\Gamma(k+1)}\left(\frac{\delta_\rho}{\sqrt{2}}\right)\left(\frac{-\delta_\rho^2}{2}\right)^j\left(\frac{-t^2}{2n-2}\right)^k\right].$$

For large n

$$\frac{\Gamma(n+i-1+k+\frac{1}{2})}{\Gamma(n+i-1)} = (n+i-1)^{k+\frac{1}{2}}\left(1+O\left(\tfrac{1}{n+i-1}\right)\right)$$

and

$$\frac{\Gamma(n+i-1+k)}{\Gamma(n+i-1)} = (n+i-1)^{k}\left(1+O\left(\tfrac{1}{n+i-1}\right)\right).$$

Hence

$$F_{T_2}(t) = \frac{1}{2}+\frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{2\,\Gamma(\frac{n-1}{2})}\sum_{i=0}^{\infty}\frac{\Gamma(\frac{n-1}{2}+i)}{\Gamma(i+1)}\left(\frac{2\rho}{1+\rho}\right)^i\left(1+O\left(\tfrac{1}{n+i-1}\right)\right)\sum_{j,k=0}^{\infty}$$
$$\left[\frac{\Gamma(j+k+\frac{1}{2})\,(-1)^k}{\Gamma(j+\frac{1}{2})\,\Gamma(k+\frac{3}{2})\,\Gamma(j+1)\,\Gamma(k+1)}\left(\frac{-\delta_\rho^2}{2}\right)^j\left(\frac{t^2}{2}\left(1+\frac{i}{n-1}\right)\right)^{k+\frac{1}{2}}\right.$$
$$\left.-\frac{\Gamma(j+k+\frac{1}{2})\,(-1)^k}{\Gamma(j+\frac{3}{2})\,\Gamma(k+\frac{1}{2})\,\Gamma(j+1)\,\Gamma(k+1)}\left(\frac{\delta_\rho}{\sqrt{2}}\right)\left(\frac{-\delta_\rho^2}{2}\right)^j\left(\frac{t^2}{2}\left(1+\frac{i}{n-1}\right)\right)^k\right].$$

Consider the i series

$$\frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{\Gamma(\frac{n-1}{2})}\sum_{i=0}^{\infty}\frac{\Gamma(\frac{n-1}{2}+i)}{\Gamma(i+1)}\left(\frac{2\rho}{1+\rho}\right)^i\left(1+\frac{i}{n-1}\right)^k\left(1+O\left(\tfrac{1}{n+i-1}\right)\right).$$

When $k=0$ and n is large, the expression simply equals 1, since $\sum_{i=0}^{\infty}\frac{\Gamma(\frac{n-1}{2}+i)}{\Gamma(i+1)}\left(\frac{2\rho}{1+\rho}\right)^i = \Gamma(\frac{n-1}{2})\left(\frac{1+\rho}{1-\rho}\right)^{\frac{n-1}{2}}$ and $O\left(\frac{1}{n+i-1}\right)\to 0$. For $k>0$ write

$$\frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{\Gamma(\frac{n-1}{2})}\sum_{i=0}^{\infty}\frac{\Gamma(\frac{n-1}{2}+i)}{\Gamma(i+1)}\left(\frac{2\rho}{1+\rho}\right)^i\sum_{l=0}^{k}\binom{k}{l}\left(\frac{i}{n-1}\right)^l\left(1+O\left(\tfrac{1}{n+i-1}\right)\right).$$

For $l=0$ (and n large), the expression equals 1.

For $l>0$, when $i=0$, the expression equals 0; thus $\sum_{i=0}^{\infty}\cdots=\sum_{i=1}^{\infty}\cdots$. Re-indexing yields

$$\frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{\Gamma(\frac{n-1}{2})}\sum_{i=0}^{\infty}\frac{\Gamma(\frac{n-1}{2}+i+1)}{\Gamma(i+1)}\left(\frac{2\rho}{1+\rho}\right)^i\left(\frac{2\rho}{1+\rho}\right)\sum_{l=1}^{k}\binom{k}{l}\left(\frac{i+1}{n-1}\right)^{l-1}\left(\frac{1}{n-1}\right)\left(1+O\left(\tfrac{1}{n+i}\right)\right)$$

Expanding $(i+1)^{l-1}$ as a power series gives

$$\frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{\Gamma(\frac{n-1}{2})}\sum_{i=0}^{\infty}\frac{\Gamma(\frac{n-1}{2}+i+1)}{\Gamma(i+1)}\left(\frac{2\rho}{1+\rho}\right)^{i}\left(\frac{2\rho}{1+\rho}\right)\sum_{l=1}^{k}c_{k,l}\,i^{l-1}\left(\frac{1}{n-1}\right)^{l}\left(1+O\left(\frac{1}{n+i}\right)\right)$$

Note that $c_{k,1}=\binom{k}{l}$.

For $l=1$ (and n large), the expression equals

$$\frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{\Gamma(\frac{n-1}{2})}\Gamma(\tfrac{n+1}{2})\left(\frac{1+\rho}{1-\rho}\right)^{\frac{n+1}{2}}\left(\frac{2\rho}{1+\rho}\right)c_{k,1}\left(\frac{1}{n-1}\right) = \binom{k}{1}\left(\frac{\rho}{1-\rho}\right)$$

For $l>1$, when $i=0$, the expression equals 0; thus $\sum_{i=0}^{\infty}\cdots=\sum_{i=1}^{\infty}\cdots$. Re-indexing yields

$$\frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{\Gamma(\frac{n-1}{2})}\sum_{i=0}^{\infty}\frac{\Gamma(\frac{n-1}{2}+i+2)}{\Gamma(i+1)}\left(\frac{2\rho}{1+\rho}\right)^{i}\left(\frac{2\rho}{1+\rho}\right)^{2}\sum_{l=2}^{k}c_{k,l}\,(i+1)^{l-2}\left(\frac{1}{n-1}\right)^{l}\left(1+O\left(\frac{1}{n+i+1}\right)\right)$$

Expanding $(i+1)^{l-2}$ as a power series gives

$$\frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{\Gamma(\frac{n-1}{2})}\sum_{i=0}^{\infty}\frac{\Gamma(\frac{n-1}{2}+i+2)}{\Gamma(i+1)}\left(\frac{2\rho}{1+\rho}\right)^{i}\left(\frac{2\rho}{1+\rho}\right)^{2}\sum_{l=2}^{k}c'_{k,l}\,i^{l-2}\left(\frac{1}{n-1}\right)^{l}\left(1+O\left(\frac{1}{n+i+1}\right)\right)$$

It can be shown that $c'_{k,2}=\binom{k}{2}$.

For $l=2$ (and n large), the expression equals

$$\begin{aligned}\frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{\Gamma(\frac{n-1}{2})}\Gamma(\tfrac{n+3}{2})\left(\frac{1+\rho}{1-\rho}\right)^{\frac{n+3}{2}}\left(\frac{2\rho}{1+\rho}\right)^{2}c'_{k,2}\left(\frac{1}{n-1}\right)^{2} &= \left(1+\frac{2}{n-1}\right)\left(\frac{\rho}{1-\rho}\right)^{2}\binom{k}{2}\\ &= \binom{k}{2}\left(\frac{\rho}{1-\rho}\right)^{2}\end{aligned}$$

Continuing through $l=k$, one arrives at, for $k>0$

$$\begin{aligned}&\frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{\Gamma(\frac{n-1}{2})}\sum_{i=0}^{\infty}\frac{\Gamma(\frac{n-1}{2}+i)}{\Gamma(i+1)}\left(\frac{2\rho}{1+\rho}\right)^{i}\left(1+\frac{i}{n-1}\right)^{k}\left(1+O\left(\frac{1}{n+i-1}\right)\right)\\ &\qquad\to\ 1+\sum_{l=1}^{k}\binom{k}{l}\left(\frac{\rho}{1-\rho}\right)^{l} = \left(1+\frac{\rho}{1-\rho}\right)^{k} = \left(\frac{1}{1-\rho}\right)^{k}\end{aligned}$$

and since the expression equals 1 for $k=0$, the relation is true for all $k=0,1,2,\ldots$.

Similarly

$$\frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{\Gamma\left(\frac{n-1}{2}\right)} \sum_{i=0}^{\infty} \frac{\Gamma\left(\frac{n-1}{2}+i\right)}{\Gamma(i+1)} \left(\frac{2\rho}{1+\rho}\right)^{i} \left(1+\frac{i}{n-1}\right)^{k+\frac{1}{2}} \left(1+O\left(\frac{1}{n+i-1}\right)\right) \;\rightarrow\; \left(\frac{1}{1-\rho}\right)^{k+\frac{1}{2}}$$

Thus

$$F_{T_2}(t) \;\rightarrow\; \frac{1}{2}+\frac{1}{2}\sum_{j,k=0}^{\infty}\left[\frac{\Gamma(j+k+\frac{1}{2})\,(-1)^k}{\Gamma(j+\frac{1}{2})\,\Gamma(k+\frac{3}{2})\,\Gamma(j+1)\,\Gamma(k+1)}\left(\frac{-\delta_\rho^2}{2}\right)^{j}\left(\frac{t^2}{2(1-\rho)}\right)^{k+\frac{1}{2}} - \frac{\Gamma(j+k+\frac{1}{2})\,(-1)^k}{\Gamma(j+\frac{3}{2})\,\Gamma(k+\frac{1}{2})\,\Gamma(j+1)\,\Gamma(k+1)}\left(\frac{\delta_\rho}{\sqrt{2}}\right)\left(\frac{-\delta_\rho^2}{2}\right)^{j}\left(\frac{t^2}{2(1-\rho)}\right)^{k}\right]$$

Diagonalizing the j,k infinite sums by $\sum_{i,j=0}^{\infty} c(i,j)\, x^i\, y^j \;=\; \sum_{i=0}^{\infty}\sum_{j=0}^{i} c(i-j,j)\, x^{i-j}\, y^j$ gives

$$F_{T_2}(t) \;\rightarrow$$

$$\frac{1}{2}+\frac{1}{2}\sum_{k=0}^{\infty}\sum_{j=0}^{k}\left[\frac{\Gamma(k+\frac{1}{2})}{\Gamma(j+\frac{1}{2})\,\Gamma(k-j+\frac{3}{2})\,\Gamma(j+1)\,\Gamma(k-j+1)}\left(\frac{t}{\sqrt{2(1-\rho)}}\right)\left(\frac{-\delta_\rho^2}{2}\right)^{j}\left(\frac{-t^2}{2(1-\rho)}\right)^{k-j} - \frac{\Gamma(k+\frac{1}{2})}{\Gamma(j+\frac{3}{2})\,\Gamma(k-j+\frac{1}{2})\,\Gamma(j+1)\,\Gamma(k-j+1)}\left(\frac{\delta_\rho}{\sqrt{2}}\right)\left(\frac{-\delta_\rho^2}{2}\right)^{j}\left(\frac{-t^2}{2(1-\rho)}\right)^{k-j}\right]$$

$$= \frac{1}{2}+\frac{1}{2}\sum_{k=0}^{\infty}\sum_{j=0}^{k}\left[\frac{\Gamma(k+\frac{1}{2})\,(-1)^k}{\Gamma(j+\frac{1}{2})\,\Gamma(k-j+\frac{3}{2})\,\Gamma(j+1)\,\Gamma(k-j+1)}\left(\frac{\delta_\rho^2}{2}\right)^{j}\left(\frac{t^2}{2(1-\rho)}\right)^{k-j+\frac{1}{2}} - \frac{\Gamma(k+\frac{1}{2})\,(-1)^k}{\Gamma(j+\frac{3}{2})\,\Gamma(k-j+\frac{1}{2})\,\Gamma(j+1)\,\Gamma(k-j+1)}\left(\frac{\delta_\rho^2}{2}\right)^{j+\frac{1}{2}}\left(\frac{t^2}{2(1-\rho)}\right)^{k-j}\right]$$

Note that wherever j occurs in the first term, $j+\frac{1}{2}$ occurs in the second term; hence the j series may be combined

$$F_{T_2}(t) \;\rightarrow\; \frac{1}{2}+ \frac{1}{2}\sum_{k=0}^{\infty}\sum_{j=0}^{2k+1}\frac{\Gamma(k+\frac{1}{2})\;(-1)^k\,(-1)^j}{\Gamma\left(\frac{j}{2}+\frac{1}{2}\right)\Gamma\left(k-\frac{j}{2}+\frac{3}{2}\right)\Gamma(\frac{j}{2}+1)\,\Gamma(k-\frac{j}{2}+1)}\left(\frac{\delta_\rho^2}{2}\right)^{\frac{j}{2}}\left(\frac{t^2}{2(1-\rho)}\right)^{k-\frac{j}{2}+\frac{1}{2}}$$

Since $\Gamma(x)\,\Gamma(x+\frac{1}{2}) = \dfrac{\Gamma(2x)\,\Gamma(\frac{1}{2})}{2^{2x-1}}$

$$F_{T_2}(t) \;\rightarrow\; \frac{1}{2}+\frac{1}{2}\sum_{k=0}^{\infty}\sum_{j=0}^{2k+1}\frac{\Gamma(k+\frac{1}{2})\;(-1)^k\,(-1)^j\,2^j\,2^{2k-j+1}}{\Gamma(j+1)\,\Gamma(\frac{1}{2})\,\Gamma(2k-j+2)\,\Gamma(\frac{1}{2})}\left(\frac{\delta_\rho^2}{2}\right)^{\frac{j}{2}}\left(\frac{t^2}{2(1-\rho)}\right)^{k-\frac{j}{2}+\frac{1}{2}}$$

$$
\begin{aligned}
&= \frac{1}{2} + \frac{1}{\pi\sqrt{2}} \sum_{k=0}^{\infty} \sum_{j=0}^{2k+1} \frac{\Gamma(k+\frac{1}{2})\,(-2)^k}{\Gamma(j+1)\,\Gamma(2k-j+2)} (-\delta_\rho)^j \left(\tfrac{t}{\sqrt{1-\rho}}\right)^{2k+1-j} \\
&= \frac{1}{2} + \frac{1}{\pi\sqrt{2}} \sum_{k=0}^{\infty} \frac{\Gamma(k+\frac{1}{2})\,(-2)^k}{\Gamma(2k+2)} \left(\tfrac{t}{\sqrt{1-\rho}} - \delta_\rho\right)^{2k+1} \\
&= \frac{1}{2} + \frac{1}{\pi\sqrt{2}} \sum_{k=0}^{\infty} \frac{\Gamma(k+\frac{1}{2})\,(-2)^k\,\Gamma(\frac{1}{2})}{\Gamma(k+1)\,\Gamma(k+\frac{3}{2})\,2^{2k+1}} \left(\tfrac{t}{\sqrt{1-\rho}} - \delta_\rho\right)^{2k+1} \qquad \text{(A.6)} \\
&= \frac{1}{2} + \frac{\left(\frac{t}{\sqrt{1-\rho}} - \delta_\rho\right)}{\sqrt{2\pi}} \frac{\Gamma(\frac{3}{2})}{\Gamma(\frac{1}{2})} \sum_{k=0}^{\infty} \frac{\Gamma(k+\frac{1}{2})}{\Gamma(k+\frac{3}{2})\,\Gamma(k+1)} \left(- \frac{\left(\frac{t}{\sqrt{1-\rho}} - \delta_\rho\right)^2}{2} \right)^k \\
&= \frac{1}{2} + \frac{\left(\frac{t}{\sqrt{1-\rho}} - \delta_\rho\right)}{\sqrt{2\pi}} M\left(\tfrac{1}{2}, \tfrac{3}{2}; - \frac{\left(\frac{t}{\sqrt{1-\rho}} - \delta_\rho\right)^2}{2} \right)
\end{aligned}
$$

where $M(\ )$ is the confluent hypergeometric function: see 2.4. Abramowitz and Stegun [1] at **26.2.10** give

$$
\begin{aligned}
\Phi(x) &= \tfrac{1}{2} + \frac{1}{\sqrt{2\pi}} \sum_{i=0}^{\infty} \frac{(-1)^i\, x^{2i+1}}{i!\, 2^i\, (2i+1)} \\
&= \tfrac{1}{2} + \frac{x}{\sqrt{2\pi}} M\left(\tfrac{1}{2}, \tfrac{3}{2}; -\tfrac{x^2}{2}\right)
\end{aligned}
$$

where $\Phi(\cdot)$ is the standard normal distribution function. Hence asymptotically

$$
T_2(t|\rho) \sim N\left(\tfrac{t}{\sqrt{1-\rho}} - \delta_\rho \, . \, 1\right) \quad .
$$

The result may also be obtained by considering the asymptotic distribution of $Y_2 - Y_1$ and $\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(Y_{1i}-\bar{Y}_{1\cdot})^2 + (Y_{2i}-\bar{Y}_{2\cdot})^2}$.

A.5 The Derivative with respect to ρ of the distribution function of $T_2|\rho$

We state two forms of the derivative with respect to ρ of the distribution function of $T_2|\rho$. The first is algebraically simple; the second is more convergent andd suitable for computing.

$$
\frac{dF_{T_2|\rho}}{d\rho} = \frac{e^{-\frac{\delta_\rho^2}{2}} \left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{2\sqrt{2\pi}\,(1-\rho)\,\Gamma(\frac{n-1}{2})} \times \qquad \text{(A.7)}
$$

$$\left[\frac{2}{1+\rho}\sum_{i,j,k=0}^{\infty}\frac{\Gamma(\frac{n+1}{2}+i)\,\Gamma\left(n+i+\frac{j-1}{2}+k\right)\,(-\frac{1}{2})^k}{\Gamma(n+i)\,\Gamma(i+1)\,\Gamma(j+1)\,\Gamma(k+1)}\left(\frac{2\rho}{1+\rho}\right)^i\left(\delta_\rho\right)^j\left(\frac{t}{\sqrt{n-1}}\right)^{j+2k+1}\right.$$
$$\left.-\sum_{i,j,k=0}^{\infty}\frac{\Gamma(\frac{n-1}{2}+i)\,\Gamma\left(n+i+\frac{j-2}{2}+k\right)\,(-\frac{1}{2})^k}{\Gamma(n+i-1)\,\Gamma(i+1)\,\Gamma(j+1)\,\Gamma(k+1)}\left(\frac{2\rho}{1+\rho}\right)^i\left(\delta_\rho\right)^{j+1}\left(\frac{t}{\sqrt{n-1}}\right)^{j+2k}\right]$$

Note that $\sum_{i=0}^{\infty}\frac{\Gamma(\alpha+i)}{\Gamma(i+1)}x^i = \Gamma(\alpha)\,(1-x)^{-\alpha}$; hence

$$\frac{dF_{T_2|\rho}}{d\rho} = \frac{e^{-\frac{\delta_\rho^2}{2}}\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{2\sqrt{2\pi}\,(1-\rho)\,\Gamma(\frac{n-1}{2})}\left(\frac{2(n-1)}{2(n-1)+t^2}\right)^{n-1}\left[\frac{2}{1+\rho}\times\right. \tag{A.8}$$
$$\sum_{i,j=0}^{\infty}\frac{\Gamma(\frac{n-1}{2}+i+1)\,\Gamma\left(n-1+i+\frac{j+1}{2}\right)}{\Gamma(n-1+i+1)\,\Gamma(i+1)\,\Gamma(j+1)}\left(\frac{2\rho}{1+\rho}\right)^i\left(\delta_\rho\right)^j\left(\frac{t}{\sqrt{n-1}}\right)^{j+1}\left(\frac{2(n-1)}{2(n-1)+t^2}\right)^{i+\frac{j+1}{2}}$$
$$\left.-\sum_{i,j=0}^{\infty}\frac{\Gamma(\frac{n-1}{2}+i)\,\Gamma\left(n-1+i+\frac{j}{2}\right)}{\Gamma(n-1+i)\,\Gamma(i+1)\,\Gamma(j+1)}\left(\frac{2\rho}{1+\rho}\right)^i\left(\delta_\rho\right)^{j+1}\left(\frac{t}{\sqrt{n-1}}\right)^j\left(\frac{2(n-1)}{2(n-1)+t^2}\right)^{i+\frac{j}{2}}\right]$$

A.6 Confirmation of the Distribution of $T_2|\rho$ for negative ρ by simulation

Proschan's derivation of the distribution of $T_2|\rho$ was restricted to $\rho > 0$. The distribution function $F_{T_2}(t)$ as we have derived it at (A.3) and (A.4) is well-defined for $\rho \in (-1, 1)$. We confirmed that our form at (A.4)

$$F_{T_2}(t) = \frac{1}{2} - \frac{\delta_\rho}{\sqrt{2\pi}}M\left(\frac{1}{2},\frac{3}{2};-\frac{\delta_\rho^2}{2}\right) + \frac{t\,e^{-\frac{\delta_\rho^2}{2}}\sqrt{1+\rho}}{\Gamma(\frac{n-1}{2})\,\sqrt{2(n-1)+t^2(1+\rho)}}\left(\frac{(n-1)\sqrt{1-\rho^2}}{2(n-1)+t^2(1+\rho)}\right)^{n-1}$$
$$\times\sum_{i,j,k=0}^{\infty}\frac{\Gamma\left(n+\frac{j-1}{2}+2i+k\right)\Gamma\left(\frac{j+1}{2}\right)}{\Gamma(\frac{n}{2}+i)\,\Gamma\left(\frac{j+3}{2}+k\right)\Gamma(i+1)\Gamma(j+1)}\left(\frac{\rho(n-1)}{2(n-1)+t^2(1+\rho)}\right)^{2i}$$
$$\times\left(\frac{t\,\delta_\rho\sqrt{2(1+\rho)}}{\sqrt{2(n-1)+t^2(1+\rho)}}\right)^j\left(\frac{t^2(1+\rho)}{t^2(1+\rho)+2(n-1)}\right)^k$$

is correct for $\rho < 0$ by simulation. Referring to table A.1, for each set of parameter values, we generated 5,000 bivariate normal datasets each comprising 5 pairs, and counted the proportion for which the two sample t-statistic exceeded the value shown. The proportion is shown in the column P_{sim}; the probability as calculated by the distribution function is shown in the column P_{calc}.

Table A.1: $P(T_2|\rho<t)$ for negative ρ by simulation and by computation.

$\frac{\mu_2-\mu_1}{\sigma}$	t	ρ	P_{sim}	P_{calc}
1	2.306	-0.5	0.678	0.6776
1	2.306	-0.8	0.667	0.655
1.5	2.306	-0.25	0.454	0.4537
1.5	2.306	-0.85	0.456	0.4497
2	2.306	-0.2	0.2314	0.2256
2	2.306	-0.45	0.2522	0.2424
2	2.306	-0.75	0.2592	0.258

A.7 Splus *code for the Distribution Function of* $T_2|\rho$

The following Splus code implements the expression for the distribution function of $T_2|\rho$ at (A.4).

```
pmut_function(q,Delta,nn,rho,rel.tol=.Machine$double.eps^.25,verbose=F){
if(any(!is.numeric(q),!is.numeric(Delta),!is.numeric(nn),!is.numeric(rho),
length(q)!=1,length(Delta)!=1,length(nn)!=1,length(rho)!=1,nn<3,nn>100))
stop("quantile, Delta, nn and rho must be numeric and of length 1 and
3<=nn<=100")
if(rho==1)stop("Function undefined for rho=1")
# this returns P(Tmu<q | Delta,nn,rho) by the triple sum method.
# Delta=(mu2-mu1)/sigma.
delta.rho_Delta*sqrt(nn/2/(1-rho))
summ_0
if(q!=0&rho^2!=1){
xi_(rho*(nn-1)/(2*(nn-1)+q^2*(1+rho)))^2
xj_delta.rho*q*sqrt(2*(1+rho)/(2*(nn-1)+q^2*(1+rho)))
xk_q^2*(1+rho)/(2*(nn-1)+q^2*(1+rho))
```

```
func.ij_function(i,j,k,xi,xj,n)exp(lgamma(n+2*i+(j-1)/2+k)+lgamma((j+1)/2)
-lgamma(n/2+i)-lgamma((j+3)/2+k)-lgamma(i+1)-lgamma(j+1))*xi^i*xj^j
func.jk_function(j,k,i,xj,xk,n)exp(lgamma(n+2*i+(j-1)/2+k)+lgamma((j+1)/2)
-lgamma(n/2+i)-lgamma((j+3)/2+k)-lgamma(i+1)-lgamma(j+1))*xj^j*xk^k
func.ik_function(i,k,j,xi,xk,n)exp(lgamma(n+2*i+(j-1)/2+k)+lgamma((j+1)/2)
-lgamma(n/2+i)-lgamma((j+3)/2+k)-lgamma(i+1)-lgamma(j+1))*xi^i*xk^k
if(verbose==T)print(paste("xi = ",signif(xi,7),"xj = ",signif(xj,7),"xk =
",signif(xk,7),sep=""))
rks_4*round(rank(abs(c(xi,xj,xk)))+c(.1,-.1,0),0)-5
 # this must return two positive integers (the size of the first "outer")
 # and a -1.
i_rks[1]
j_rks[2]
k_rks[3]
# Note the first sum is a big sum based on the ranks of xi, xj, xk
sumij_10^60*(i+1)*(j+1)
sumjk_10^60*(j+1)*(k+1)
sumik_10^60*(i+1)*(k+1)
# Note: wch==1, do ij sum
#       wch==2, do jk sum
#       wch==3, do ik sum
while((abs(sumij)+abs(sumik)+abs(sumjk))/abs(summ)>rel.tol){
wch_match(max(abs(c(sumij,sumjk,sumik))),abs(c(sumij,sumjk,sumik)))
i_i+(wch==2)
j_j+(wch==3)
k_k+(wch==1)
is_0:i
js_0:j
ks_0:k
terms_switch(wch,xk^k*outer(is,js,FUN=func.ij,k=k,xi=xi,xj=xj,n=nn),
```

```
xi^i*outer(js,ks,FUN=func.jk,i=i,xj=xj,xk=xk,n=nn),
xj^j*outer(is,ks,FUN=func.ik,j=j,xi=xi,xk=xk,n=nn))
summ_summ+sum(terms)
sumij_(wch==1)*sum(terms) + (wch!=1)*sumij +
(wch==2)*sum(terms[,dim(terms)[2]]) + (wch==3)*sum(terms[,dim(terms)[2]])
sumjk_(wch==2)*sum(terms) + (wch!=2)*sumjk +
(wch==1)*sum(terms[dim(terms)[1],]) + (wch==3)*sum(terms[dim(terms)[1],])
sumik_(wch==3)*sum(terms) + (wch!=3)*sumik +
(wch==1)*sum(terms[,dim(terms)[2]]) + (wch==2)*sum(terms[dim(terms)[1],])
if(verbose==T)print(paste("wch = ",wch,", i = ",i,", j = ",j,", k = ",k,",
sum(terms) = ",signif(sum(terms),4),", summ = ",signif(summ,4),sep=""))
   } }
1/2 - delta.rho * confl.hypergeom(.5,1.5,-delta.rho^2/2) / sqrt(2*pi) +
q * exp(-delta.rho^2/2) * gamma(nn/2) * (1-rho^2)^((nn-1)/2) / 2 /
sqrt(2*pi) / gamma(nn-1) * (2*(nn-1)/(2*(nn-1)+q^2*(1+rho)))^(nn-1) *
sqrt(2*(1+rho)/(2*(nn-1)+q^2*(1+rho))) * summ  }

confl.hypergeom_function(alpha, gamma, x, tol = .Machine$double.eps^.5,
maxterms = 1000,verbose=F){
if(any(!is.numeric(alpha),!is.numeric(gamma),!is.numeric(x),length(alpha)!=1,
length(gamma)!=1,gamma%%1==0&gamma<1))stop("alpha,gamma and x must be
numeric; alpha and gamma must be of length 1; gamma cannot be 0 or a
negative integer")
numterm <- 0
result <- term <- abs.sum <- 1
if(verbose==T)print(c("term","sum"))
while(max(abs(term/result)) > tol) {
numterm <- numterm + 1
if(numterm > maxterms)stop("maxterms (1000?) added without reaching tolerance")
term <- term * (alpha + numterm - 1) * x /numterm /(gamma + numterm - 1)
```

```
result <- result + term
abs.sum_abs.sum+abs(term)
if(verbose==T)print(c(term,result))
}
if(max(abs.sum/result)>10^10)warning("Value returned by confluent
hypergeometric {confl.hypergeom()} function may be wrong")
result}
```

A.8 Splus *code for the Density Function of* $T_2|\rho$

The following Splus code implements the expression for the density function of $T_2|\rho$ at (A.5).

```
dmut_function(q,Delta,nn,rho,rel.tol=.Machine$double.eps^.25,verbose=F){
if(any(!is.numeric(q),!is.numeric(Delta),!is.numeric(nn),!is.numeric(rho),
sum(length(q)!=1,length(Delta)!=1,length(rho)!=1)>1,length(nn)!=1,nn<3,nn>100))
stop("quantile, Delta and rho must be numeric and only one of them can be
vector valued; nn must be numeric, length(nn)=1 and 3<=nn<=100")
if(any(abs(rho)==1))stop("Function undefined for rho=+/-1")
# this returns the density of (Tmu<q | Delta,nn,rho) by the double sum method.
# One vector-valued arguement is allowed (not nn).  Delta=(mu2-mu1)/sigma.
delta.rho_Delta*sqrt(nn/2/(1-rho))
summ_rep(0,length(Delta*rho*q))
xi_(rho*(nn-1)/(2*(nn-1)+q^2*(1+rho)))^2*Delta^0 # for correct length
xj_delta.rho*q*sqrt(2*(1+rho)/(2*(nn-1)+q^2*(1+rho)))
func.i_function(i,xi,j,n)exp(lgamma(n+2*i+(j-1)/2)-lgamma(n/2+i)-lgamma(i+1)
-lgamma(j+1))*xi^i
func.j_function(j,xj,i,n)exp(lgamma(n+2*i+(j-1)/2)-lgamma(n/2+i)-lgamma(i+1)
-lgamma(j+1))*xj^j
if(verbose==T)print(paste("xi = ",signif(xi,7),"xj = ",signif(xj,7)))
i_5*(max(abs(xi))>max(abs(xj)))-1
j_5*(1-(max(abs(xi))>max(abs(xj))))-1
```

```
sumi_rep(10^60*(i+1),length(Delta*rho*q))
sumj_rep(10^60*(j+1),length(Delta*rho*q))
# Note: wch==0, do i sum
#       wch==1, do j sum
while(max(abs(sumi)+abs(sumj))/max(abs(summ))>rel.tol){
wch_(max(abs(sumi))<max(abs(sumj)))*1
i_i+wch
j_j+(1-wch)
is_0:i
js_0:j
terms_switch(wch+1,
(is+1)^0%*%t(as.matrix(xj^j))*outer(is,xi,FUN=func.i,j=j,n=nn),
(js+1)^0%*%t(as.matrix(xi^i))*outer(js,xj,FUN=func.j,i=i,n=nn))
smm_apply(terms,2,sum)
summ_summ+smm
sumi_(1-wch)*smm + wch*(sumi + terms[dim(terms)[1],])
sumj_wch*smm + (1-wch)*(sumj + terms[dim(terms)[1],])
if(verbose==T)print(paste("wch=",wch,", i=",i,", j=",j,",
max(abs(terms))=",signif(max(abs(terms)),4),", max(abs(summ))=",
signif(max(abs(summ)),4), sep=""))
    }
exp(-delta.rho^2/2) * gamma(nn/2) * ((1-rho)/(1+rho))^((nn-1)/2) /
sqrt(2*pi*(nn-1)) / gamma(nn-1) *
(2*(1+rho)*(nn-1)/(2*(nn-1)+q^2*(1+rho)))^((2*nn-1)/2) * summ
    }
```

Appendix B

THE DISTRIBUTION OF $T_2|r$

We derive the distribution of $T_2|r$, that is, the distribution of the two sample t statistic T_2 conditioned on the Pearson product-moment sample correlation coefficient $r \equiv r_1$, under the null $\mu_2 = \mu_1$. In section B.1 we derive the distribution function; in section B.2 a form of the distribution function that is convergent for all values of the arguments and that would be suitable for computation purposes is provided; in section B.3 a density function is given.

Notation

In this Appendix,

$$s_1^2 \equiv \frac{1}{n}\sum_{i=1}^{n}(y_{i1}-\bar{y}_{\cdot 1})^2;$$

$$s_2^2 \equiv \frac{1}{n}\sum_{i=1}^{n}(y_{i2}-\bar{y}_{\cdot 2})^2;$$

$$s_{12} \equiv \frac{1}{n}\sum_{i=1}^{n}(y_{i1}-\bar{y}_{\cdot 1})(y_{i2}-\bar{y}_{\cdot 2})^2;$$

These meanings are consistent with the authors cited here, but differ from the meanings given in section 2.

B.1 The Distribution Function of $T_2|r$

Recall that $\begin{pmatrix} y_{11} \\ y_{12} \end{pmatrix} \cdots \begin{pmatrix} y_{n1} \\ y_{n2} \end{pmatrix}$ represents n pairs of observations from

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} \sim N_2\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \sigma^2\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right)$$

and that the two sample t-statistic $T_2 \equiv \dfrac{(\bar{Y}_{\cdot 1}-\bar{Y}_{\cdot 2})}{\sqrt{\frac{1}{n-1}\left(S_1^2+S_2^2\right)}}$

The steps of the derivation are

i) State the distribution of $\bar{D} \equiv \bar{Y}_{\cdot 2} - \bar{Y}_{\cdot 1}$;

ii) State the joint density of S_1, S_2 and R;

iii) Derive the joint density of $\bar{D}, S_1, S_2$ and R from i) and ii);

iv) Derive the joint density of $\bar{D}, S_1$ and S_2 conditioned on r by dividing iii) by the density of R. It will be found that this conditional joint density may be factorized into a function of $\bar{D}$ and a function of S_1, S_2 and r.

v) Derive the distribution function and hence the density of $W^2 \equiv S_1^2 + S_2^2 \,|\, r$ from iv);

vi) Recombine the density of $\bar{D}$ with that of $W|r$ to give the joint density of $\bar{D}$ and W conditioned on r;

vii) Noting that $T_2 = \frac{\bar{D}\sqrt{n-1}}{W}$, integrate the density in vi) over the region of $\bar{d}$-w space for which $T_2 < t$.

For convenience, we set σ^2 to 1.

Under the null, the distribution of $\bar{D} \equiv \bar{Y}_{\cdot 2} - \bar{Y}_{\cdot 1}$ is Normal with mean 0 and variance $\frac{2}{n}(1-\rho)$; thus its density is given by

$$dP = \frac{\sqrt{n}}{2\sqrt{\pi(1-\rho)}} \exp\left\{ -\frac{n}{4(1-\rho)} \bar{d}^2 \right\} d\bar{d} \tag{B.1}$$

For the joint density of S_1, S_2 and R, *Kendall's Advanced Theory of Statistics* [39] gives at (16.59)

$$\begin{aligned} dP &= \frac{n^{n-1}}{\pi \sigma_1^{n-1} \sigma_2^{n-1} (1-\rho^2)^{(n-1)/2} \Gamma(n-2)} \\ &\quad \times \exp\left\{ -\frac{n}{2(1-\rho^2)} \left[\frac{s_1^2}{\sigma_1^2} - \frac{2\rho r s_1 s_2}{\sigma_1 \sigma_2} + \frac{s_2^2}{\sigma_2^2} \right] \right\} s_1^{n-2} s_2^{n-2} (1-r^2)^{(n-4)/2} \, ds_1 \, ds_2 \, dr \end{aligned} \tag{B.2}$$

Expressions (B.1) and (B.2) may be combined to give the joint density of $\bar{D}, S_1^2, S_2^2$ and R, since the distributions of the sample means and the sample variances and covariances are independent in bivariate normal distributions. The joint density of $\bar{D},, S_1^2$ and S_2^2 conditioned on $R{=}r$ may then be obtained by dividing by the density of R. For the density of R we use the form given in *Kendall* at (16.66)

$$dP = \frac{(n-2)\,\Gamma(n)}{(n-1)\sqrt{2}\,\Gamma(\frac{1}{2})\,\Gamma(n-\frac{1}{2})} (1-\rho^2)^{(n-1)/2} (1-r^2)^{(n-4)/2} (1-\rho r)^{3/2-n}$$

$$\times \; H(\tfrac{1}{2},\tfrac{1}{2},n-\tfrac{1}{2},\tfrac{1}{2}(1+\rho r))\,dr \tag{B.3}$$

where H is the Gauss hypergeometric function (see 2.4). The conditional density may again be factorized; the density of $\bar{D}$ is still given by equation (B.1); after setting σ_1 and σ_2 to 1, the joint density of the sample variances given r is

$$\begin{aligned} dP &= \frac{\sqrt{2}\,\Gamma(n-\frac{1}{2})\,n^{n-1}}{\sqrt{\pi}(1-\rho^2)^{n-1}\Gamma^2(n-1)} \exp\left\{-\frac{n}{2(1-\rho^2)}\,(s_1^2-2\rho r s_1 s_2+s_2^2)\right\} \\ &\quad \times s_1^{n-2}\,s_2^{n-2}\,(1-\rho r)^{n-3/2}\,H^{-1}\left(\tfrac{1}{2},\tfrac{1}{2},n-\tfrac{1}{2},\tfrac{1}{2}(1+\rho r)\right)\,ds_1\,ds_2 \end{aligned} \tag{B.4}$$

where H^{-1} is the reciprocal of the hypergeometric function and $\Gamma^2(\cdot)$ is $(\Gamma(\cdot))^2$.

Put $W^2 \equiv S_1^2+S_2^2$; $W>0$; then

$$\begin{aligned} P(W\le w\,|\,r) &= P(W^2\le w^2\,|\,r) \\ &= P(S_1^2+S_2^2\le w^2\,|\,r) \\ &= \int_{s_1,s_2:s_1^2+s_2^2\le w^2} f_{S_1,S_2}(s_1,s_2\,|\,r)\,ds_1\,ds_2 \\ &\propto \int_{s_1,s_2:s_1^2+s_2^2\le w^2} \exp\left\{-\frac{n}{2(1-\rho^2)}\,(s_1^2-2\rho r s_1 s_2+s_2^2)\right\} s_1^{n-2}\,s_2^{n-2}\,ds_1\,ds_2 \end{aligned}$$

from (B.4). Make the transformation

$$s_1=\sqrt{\frac{2(1-\rho^2)}{n}u}\,\sin\theta, \qquad s_2=\sqrt{\frac{2(1-\rho^2)}{n}u}\,\cos\theta$$

whose Jacobean is

$$\begin{aligned} \frac{\partial(s_1,s_2)}{\partial(u,\theta)} &= \begin{vmatrix} \sqrt{\frac{(1-\rho^2)}{2nu}}\sin\theta & \sqrt{\frac{(1-\rho^2)}{2nu}}\cos\theta \\ \sqrt{\frac{2(1-\rho^2)}{n}u}\cos\theta & -\sqrt{\frac{2(1-\rho^2)}{n}u}\sin\theta \end{vmatrix} \\ &= -\frac{1-\rho^2}{n}(\sin^2\theta+\cos^2\theta) \;=\; -\frac{1-\rho^2}{n} \end{aligned}$$

For the limits of integration, note that $s_1^2+s_2^2\le w^2 \Rightarrow u\le\frac{nw^2}{2(1-\rho^2)}$, that s_1,s_2 real $\Rightarrow u\ge 0$ and that $s_1,s_2\ge 0\Rightarrow \theta\in[0,\frac{\pi}{2}]$. Thus

$$\begin{aligned} P(W\le w\,|\,r) &\propto \int_{u=0}^{\frac{nw^2}{2(1-\rho^2)}}\int_{\theta=0}^{\frac{\pi}{2}} \exp\{-u(1-2\rho r\sin\theta\cos\theta)\} \\ &\quad \times \left(\tfrac{2(1-\rho^2)}{n}\right)^{n-2} u^{n-2}(\sin\theta\cos\theta)^{n-2}\left(\tfrac{1-\rho^2}{n}\right)\,du\,d\theta \end{aligned}$$

$$= \left(\tfrac{1-\rho^2}{n}\right)^{n-1} \int_{u=0}^{\frac{nw^2}{2(1-\rho^2)}} \int_{\theta=0}^{\frac{\pi}{2}} e^{-u} \exp\{u\rho r \sin 2\theta\} u^{n-2} (\sin 2\theta)^{n-2}\, du\, d\theta$$

$$= \left(\tfrac{1-\rho^2}{n}\right)^{n-1} \int_{u=0}^{\frac{nw^2}{2(1-\rho^2)}} e^{-u} u^{n-2} \int_{\theta=0}^{\frac{\pi}{2}} \sum_{i=0}^{\infty} \frac{(u\rho r)^i}{i!} (\sin 2\theta)^{n+i-2}\, du\, d\theta$$

$$= \left(\tfrac{1-\rho^2}{n}\right)^{n-1} \int_{u=0}^{\frac{nw^2}{2(1-\rho^2)}} e^{-u} u^{n-2} \sum_{i=0}^{\infty} \frac{(u\rho r)^i}{i!} \frac{\sqrt{\pi}\,\Gamma\left(\frac{n+i-2+1}{2}\right)}{2\,\Gamma\left(\frac{n+i-2+2}{2}\right)}\, du$$

$$= \left(\tfrac{1-\rho^2}{n}\right)^{n-1} \sum_{i=0}^{\infty} \frac{(\rho r)^i}{i!} \frac{\sqrt{\pi}\,\Gamma\left(\frac{n+i-1}{2}\right)}{2\,\Gamma\left(\frac{n+i}{2}\right)} \int_{u=0}^{\frac{nw^2}{2(1-\rho^2)}} e^{-u} u^{n+i-2}\, du$$

$$= \left(\tfrac{1-\rho^2}{n}\right)^{n-1} \sum_{i=0}^{\infty} \frac{(\rho r)^i}{i!} \frac{\sqrt{\pi}\,\Gamma\left(\frac{n+i-1}{2}\right)}{2\,\Gamma\left(\frac{n+i}{2}\right)} \int_{u=0}^{\frac{nw^2}{2(1-\rho^2)}} \sum_{j=0}^{\infty} \frac{(-u)^j}{j!} u^{n+i-2}\, du$$

$$= \left(\tfrac{1-\rho^2}{n}\right)^{n-1} \sum_{i=0}^{\infty} \frac{(\rho r)^i}{i!} \frac{\sqrt{\pi}\,\Gamma\left(\frac{n+i-1}{2}\right)}{2\,\Gamma\left(\frac{n+i}{2}\right)} \sum_{j=0}^{\infty} \frac{(-1)^j}{j!} \int_{u=0}^{\frac{nw^2}{2(1-\rho^2)}} u^{n+i+j-2}\, du$$

$$= \left(\tfrac{1-\rho^2}{n}\right)^{n-1} \sum_{i=0}^{\infty} \frac{(\rho r)^i}{i!} \frac{\sqrt{\pi}\,\Gamma\left(\frac{n+i-1}{2}\right)}{2\,\Gamma\left(\frac{n+i}{2}\right)} \sum_{j=0}^{\infty} \frac{(-1)^j}{j!} \frac{\left(\frac{nw^2}{2(1-\rho^2)}\right)^{n+i+j-1}}{n+i+j-1}$$

The derivative with respect to w may be taken to give the density

$$dP \propto \left(\tfrac{1-\rho^2}{n}\right)^{n-1} \sum_{i=0}^{\infty} \frac{(\rho r)^i}{i!} \frac{\sqrt{\pi}\,\Gamma\left(\frac{n+i-1}{2}\right)}{2\,\Gamma\left(\frac{n+i}{2}\right)} \sum_{j=0}^{\infty} \frac{(-1)^j}{j!} \left(\frac{n}{2(1-\rho^2)}\right)^{n+i+j-1} 2w^{2n+2i+2j-3} dw$$

$$= \frac{\sqrt{\pi}\, w^{2n-3}}{2^{n-1}} \sum_{i=0}^{\infty} \frac{\Gamma\left(\frac{n+i-1}{2}\right)}{\Gamma\left(\frac{n+i}{2}\right) i!} \left(\frac{n\rho r w^2}{2(1-\rho^2)}\right)^i \sum_{j=0}^{\infty} \frac{1}{j!} \left(\frac{-nw^2}{2(1-\rho^2)}\right)^j dw$$

$$= \frac{\sqrt{\pi}\, w^{2n-3}}{2^{n-1}} \exp\left\{\frac{-nw^2}{2(1-\rho^2)}\right\} \sum_{i=0}^{\infty} \frac{\Gamma\left(\frac{n+i-1}{2}\right)}{\Gamma\left(\frac{n+i}{2}\right) i!} \left(\frac{n\rho r w^2}{2(1-\rho^2)}\right)^i dw$$

With the constants we have

$$dP = \frac{\sqrt{2}\,\Gamma(n-\frac{1}{2})\, n^{n-1}}{\sqrt{\pi}(1-\rho^2)^{n-1}\Gamma^2(n-1)} (1-\rho r)^{n-3/2} H^{-1}\left(\tfrac{1}{2}, \tfrac{1}{2}, n-\tfrac{1}{2}, \tfrac{1}{2}(1+\rho r)\right)$$

$$\times \frac{\sqrt{\pi}\, w^{2n-3}}{2^{n-1}} \exp\left\{\frac{-nw^2}{2(1-\rho^2)}\right\} \sum_{i=0}^{\infty} \frac{\Gamma\left(\frac{n+i-1}{2}\right)}{\Gamma\left(\frac{n+i}{2}\right) i!} \left(\frac{n\rho r w^2}{2(1-\rho^2)}\right)^i dw$$

$$= \frac{\sqrt{2}\,\Gamma(n-\frac{1}{2})}{\Gamma^2(n-1)} \left(\frac{n}{2(1-\rho^2)}\right)^{n-1} (1-\rho r)^{n-3/2} H^{-1}\left(\tfrac{1}{2}, \tfrac{1}{2}, n-\tfrac{1}{2}, \tfrac{1}{2}(1+\rho r)\right)$$

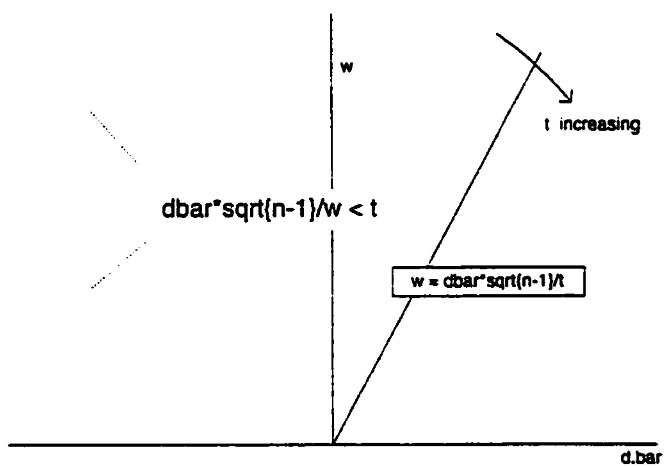
$$\times\; w^{2n-3} \exp\left\{\frac{-nw^2}{2(1-\rho^2)}\right\} \sum_{i=0}^{\infty} \frac{\Gamma\left(\frac{n+i-1}{2}\right)}{\Gamma\left(\frac{n+i}{2}\right) i!} \left(\frac{n\rho r w^2}{2(1-\rho^2)}\right)^i dw$$

Combining this with (B.1), the joint density of $\bar{D}$ and W conditioned on r is given by

$$\begin{aligned} dP &= \frac{\Gamma(n-\frac{1}{2})}{\Gamma^2(n-1)}\sqrt{\frac{n}{2\pi(1-\rho)}}\left(\frac{n}{2(1-\rho^2)}\right)^{n-1}(1-\rho r)^{n-3/2}H^{-1}\left(\tfrac{1}{2},\tfrac{1}{2},n-\tfrac{1}{2},\tfrac{1}{2}(1+\rho r)\right) \\ &\times \exp\left\{\frac{-n\bar{d}^2}{4(1-\rho)}\right\} w^{2n-3}\exp\left\{\frac{-nw^2}{2(1-\rho^2)}\right\}\sum_{i=0}^{\infty}\frac{\Gamma\left(\frac{n+i-1}{2}\right)}{\Gamma\left(\frac{n+i}{2}\right)i!}\left(\frac{n\rho r w^2}{2(1-\rho^2)}\right)^i dw\, d\bar{d} \end{aligned}$$

Now $T_2 \equiv \frac{\bar{D}\sqrt{n-1}}{W}$; thus

$$\begin{aligned} P(T_2 \le t\,|\,r) &= P\left(\frac{\bar{D}\sqrt{n-1}}{W} \le t\,|\,r\right) \\ &= \int_{\bar{d},w:\frac{\bar{d}\sqrt{n-1}}{w}\le t} f_{\bar{D},W}(\bar{d},w\,|\,r)\,d\bar{d}\,dw \end{aligned}$$



$$\begin{aligned} &= \int_{w=0}^{\infty}\int_{\bar{d}=-\infty}^{\frac{tw}{\sqrt{n-1}}} f_{\bar{D},W}(\bar{d},w\,|\,r)\,d\bar{d}\,dw \\ &= \int_{w=0}^{\infty}\int_{\bar{d}=-\infty}^{0} f_{\bar{D},W}(\bar{d},w\,|\,r)\,d\bar{d}\,dw + \int_{w=0}^{\infty}\int_{\bar{d}=0}^{\frac{tw}{\sqrt{n-1}}} f_{\bar{D},W}(\bar{d},w\,|\,r)\,d\bar{d}\,dw \\ &= \frac{1}{2} + \int_{w=0}^{\infty}\int_{\bar{d}=0}^{\frac{tw}{\sqrt{n-1}}} f_{\bar{D},W}(\bar{d},w\,|\,r)\,d\bar{d}\,dw \end{aligned}$$

by the symmetry of the distribution, (since the density must be the same when $\bar{d} \equiv \bar{y}_1 - \bar{y}_2$ as when $\bar{d} \equiv \bar{y}_2 - \bar{y}_1$). The integral of the density with respect to $\bar{d}$ may be accomplished by expressing the exponential function in $\bar{d}$ as an infinite sum and integrating term by term

$$= \frac{1}{2} + \int_{w=0}^{\infty}\int_{\bar{d}=0}^{\frac{tw}{\sqrt{n-1}}} \frac{\Gamma(n-\frac{1}{2})}{\Gamma^2(n-1)}\sqrt{\frac{n}{2\pi(1-\rho)}}\left(\frac{n}{2(1-\rho^2)}\right)^{n-1}(1-\rho r)^{n-3/2}H^{-1}\left(\tfrac{1}{2},\tfrac{1}{2},n-\tfrac{1}{2},\tfrac{1}{2}(1+\rho r)\right)$$

$$\times \ \exp\left\{\frac{-n\bar{d}^2}{4(1-\rho)}\right\} w^{2n-3} \exp\left\{\frac{-nw^2}{2(1-\rho^2)}\right\} \sum_{i=0}^{\infty} \frac{\Gamma\left(\frac{n+i-1}{2}\right)}{\Gamma\left(\frac{n+i}{2}\right) i!} \left(\frac{n\rho r w^2}{2(1-\rho^2)}\right)^i d\bar{d}\, dw$$

$$= \frac{1}{2} + \int_{w=0}^{\infty} \frac{\Gamma(n-\frac{1}{2})}{\Gamma^2(n-1)} \sqrt{\frac{n}{2\pi(1-\rho)}} \left(\frac{n}{2(1-\rho^2)}\right)^{n-1} (1-\rho r)^{n-3/2} H^{-1}(\tfrac{1}{2}, \tfrac{1}{2}, n-\tfrac{1}{2}, \tfrac{1}{2}(1+\rho r))$$

$$\times \sum_{j=0}^{\infty} \int_{\bar{d}=0}^{\frac{tw}{\sqrt{n-1}}} \left(\frac{-n\bar{d}^2}{4(1-\rho)}\right)^j \frac{1}{j!} d\bar{d}\, w^{2n-3} \exp\left\{\frac{-nw^2}{2(1-\rho^2)}\right\} \sum_{i=0}^{\infty} \frac{\Gamma(\frac{n+i-1}{2})}{\Gamma(\frac{n+i}{2})\, i!} \left(\frac{n\rho r w^2}{2(1-\rho^2)}\right)^i dw$$

$$= \frac{1}{2} + \int_{w=0}^{\infty} \frac{\Gamma(n-\frac{1}{2})}{\Gamma^2(n-1)} \sqrt{\frac{n}{2\pi(1-\rho)}} \left(\frac{n}{2(1-\rho^2)}\right)^{n-1} (1-\rho r)^{n-3/2} H^{-1}(\tfrac{1}{2}, \tfrac{1}{2}, n-\tfrac{1}{2}, \tfrac{1}{2}(1+\rho r))$$

$$\times \sum_{j=0}^{\infty} \left(\frac{-n}{4(1-\rho)}\right)^j \frac{1}{j!} \frac{\bar{d}^{2j+1}}{2j+1}\Bigg|_{\bar{d}=0}^{\frac{tw}{\sqrt{n-1}}} w^{2n-3} \exp\left\{\frac{-nw^2}{2(1-\rho^2)}\right\} \sum_{i=0}^{\infty} \frac{\Gamma(\frac{n+i-1}{2})}{\Gamma(\frac{n+i}{2})\, i!} \left(\frac{n\rho r w^2}{2(1-\rho^2)}\right)^i dw$$

$$= \frac{1}{2} + \int_{w=0}^{\infty} \frac{\Gamma(n-\frac{1}{2})}{\Gamma^2(n-1)} \sqrt{\frac{n}{2\pi(1-\rho)}} \left(\frac{n}{2(1-\rho^2)}\right)^{n-1} (1-\rho r)^{n-3/2} H^{-1}(\tfrac{1}{2}, \tfrac{1}{2}, n-\tfrac{1}{2}, \tfrac{1}{2}(1+\rho r))$$

$$\times \sum_{j=0}^{\infty} \left(\tfrac{-n}{4(1-\rho)}\right)^j \frac{1}{j!\,(2j+1)} \left(\tfrac{t}{\sqrt{n-1}}\right)^{2j+1} w^{2n+2j-2} \exp\left\{\tfrac{-nw^2}{2(1-\rho^2)}\right\} \sum_{i=0}^{\infty} \frac{\Gamma(\frac{n+i-1}{2})}{\Gamma(\frac{n+i}{2})\, i!} \left(\tfrac{n\rho r w^2}{2(1-\rho^2)}\right)^i dw$$

The second integral may be evaluated by making the change of variable $u = \frac{nw^2}{2(1-\rho^2)}$, so $dw = \frac{1}{\sqrt{2}}\left(\frac{1-\rho^2}{n}\right)^{\frac{1}{2}} u^{-\frac{1}{2}} du$

$$= \frac{1}{2} + \frac{\Gamma(n-\frac{1}{2})}{\Gamma^2(n-1)} \sqrt{\frac{n}{2\pi(1-\rho)}} \left(\frac{n}{2(1-\rho^2)}\right)^{n-1} (1-\rho r)^{n-3/2} H^{-1}(\tfrac{1}{2}, \tfrac{1}{2}, n-\tfrac{1}{2}, \tfrac{1}{2}(1+\rho r))$$

$$\times \sum_{j=0}^{\infty} \left(\frac{-n}{4(1-\rho)}\right)^j \frac{1}{j!\,(2j+1)} \left(\frac{t}{\sqrt{n-1}}\right)^{2j+1} \sum_{i=0}^{\infty} \frac{\Gamma(\frac{n+i-1}{2})}{\Gamma(\frac{n+i}{2})\, i!} (\rho r)^i$$

$$\times \int_{w=0}^{\infty} \left(\frac{2u(1-\rho^2)}{n}\right)^{n+j-1} e^{-u} u^i \frac{1}{\sqrt{2}} \left(\frac{1-\rho^2}{n}\right)^{\frac{1}{2}} u^{-\frac{1}{2}} du$$

$$= \frac{1}{2} + \frac{\Gamma(n-\frac{1}{2})}{\Gamma^2(n-1)} \sqrt{\frac{n}{2\pi(1-\rho)}} \left(\frac{n}{2(1-\rho^2)}\right)^{n-1} (1-\rho r)^{n-3/2} H^{-1}(\tfrac{1}{2}, \tfrac{1}{2}, n-\tfrac{1}{2}, \tfrac{1}{2}(1+\rho r))$$

$$\times \sum_{j=0}^{\infty} \left(\frac{-n}{4(1-\rho)}\right)^j \frac{1}{2\,j!\,(2j+1)} \left(\frac{t}{\sqrt{n-1}}\right)^{2j+1} \sum_{i=0}^{\infty} \frac{\Gamma(\frac{n+i-1}{2})}{\Gamma(\frac{n+i}{2})\, i!} (\rho r)^i \left(\frac{2(1-\rho^2)}{n}\right)^{n+j-\frac{1}{2}}$$

$$\times \int_{u=0}^{\infty} e^{-u} u^{(n+i+j-\frac{1}{2})-1} du$$

$$= \frac{1}{2} + \frac{\Gamma(n-\frac{1}{2})}{\Gamma^2(n-1)} \sqrt{\frac{n}{2\pi(1-\rho)}} \left(\frac{n}{2(1-\rho^2)}\right)^{n-1} (1-\rho r)^{n-3/2} H^{-1}(\tfrac{1}{2}, \tfrac{1}{2}, n-\tfrac{1}{2}, \tfrac{1}{2}(1+\rho r))$$

$$\times \sum_{j=0}^{\infty} \left(\frac{-n}{4(1-\rho)}\right)^j \frac{1}{2\,j!\,(2j+1)} \left(\frac{t}{\sqrt{n-1}}\right)^{2j+1} \sum_{i=0}^{\infty} \frac{\Gamma(\frac{n+i-1}{2})}{\Gamma(\frac{n+i}{2})\, i!} (\rho r)^i \left(\frac{2(1-\rho^2)}{n}\right)^{n+j-\frac{1}{2}}$$

$$\times\ \Gamma\left(n+i+j-\tfrac{1}{2}\right)$$

The expression may be simplified, yielding

$$\begin{aligned} P(T_2 \le t\,|\,r) &= \frac{1}{2} + \frac{\Gamma\left(n-\frac{1}{2}\right)(1-\rho r)^{n-\frac{3}{2}}}{2\sqrt{2\pi}\,\Gamma^2(n-1)\,H\left(\frac{1}{2},\frac{1}{2},n-\frac{1}{2},\frac{1}{2}(1+\rho r)\right)} \\ &\times \sum_{i,j=0}^{\infty} \frac{\Gamma\left(\frac{n+i-1}{2}\right)\Gamma\left(n+i+j-\frac{1}{2}\right)\Gamma\left(j+\frac{1}{2}\right)}{\Gamma\left(\frac{n+i}{2}\right)\Gamma\left(j+\frac{3}{2}\right)\Gamma(i+1)\,\Gamma(j+1)} \left(\rho r\right)^i \left(\frac{t\sqrt{1+\rho}}{\sqrt{2n-2}}\right) \left(\frac{-t^2(1+\rho)}{2n-2}\right)^j \end{aligned}$$

B.2 A convergent form of the distribution function

A form of the distribution function which will converge for all values of the arguments may be obtained by expressing the i series as a Gauss hypergeometric function and applying the transformation $H(\alpha,\beta,\gamma;x) \equiv (1-x)^{-\alpha}H(\alpha,\gamma-\beta,\gamma;\frac{x}{x-1})$ (see Erdélyi [12], p.105), followed by re-expressing the resulting Gauss hypergeometric function as a power series, in a manner similar to that used for the distribution of $T_2|\rho$ (see A.2) . This yields

$$\begin{aligned} P(T_2 \le t\,|\,r) &= \frac{1}{2} + \frac{\Gamma\left(n-\frac{1}{2}\right)(1-\rho r)^{n-3/2}\left(\frac{2(n-1)}{2(n-1)+t^2(1+\rho)}\right)^{n-1}}{2\sqrt{2}\,\Gamma^2(n-1)\,H\left(\frac{1}{2},\frac{1}{2},n-\frac{1}{2},\frac{1}{2}(1+\rho r)\right)} \left(\frac{t\sqrt{1+\rho}}{\sqrt{2(n-1)+t^2(1+\rho)}}\right) \\ &\times \sum_{i,j=0}^{\infty} \frac{\Gamma\left(\frac{n+i-1}{2}\right)\Gamma\left(n+i+j-\frac{1}{2}\right)}{\Gamma\left(\frac{n+i}{2}\right)\Gamma(i+1)\,\Gamma\left(j+\frac{3}{2}\right)} \left(\frac{2\rho r(n-1)}{2(n-1)+t^2(1+\rho)}\right)^i \left(\frac{t^2(1+\rho)}{t^2(1+\rho)+2(n-1)}\right)^j \end{aligned}$$

B.3 The density function of $T_2|r$

The derivative of the distribution function with respect to t yields the density function

$$\begin{aligned} f_{T_2}(t\,|\,r) &= \frac{\Gamma\left(n-\frac{1}{2}\right)(1-\rho r)^{n-3/2}}{\sqrt{2\pi}\,\Gamma^2(n-1)\,H\left(\frac{1}{2},\frac{1}{2},n-\frac{1}{2},\frac{1}{2}(1+\rho r)\right)} \left(\frac{2n-2}{2n-2+t^2(1+\rho)}\right)^{n-1} \sqrt{\frac{1+\rho}{2n-2+t^2(1+\rho)}} \\ &\times \sum_{i=0}^{\infty} \frac{\Gamma\left(\frac{n+i-1}{2}\right)\Gamma\left(n+i-\frac{1}{2}\right)}{\Gamma\left(\frac{n+i}{2}\right)\Gamma(i+1)} \left(\frac{\rho r(2n-2)}{2n-2+t^2(1+\rho)}\right)^i \end{aligned}$$

Appendix C

THE DISTRIBUTION OF $T_P|r$

We derive the distribution of $T_p|r$, that is, the distribution of the paired t statistic T_2 conditioned on the Pearson product-moment sample correlation coefficient $r \equiv r_1$, under the null $\mu_2 = \mu_1$. In section C.1 we derive the distribution function; in section C.2 a form of the distribution function that is convergent for all values of the arguments, suitable for computation purposes, is provided.

Notation

In this Appendix,

$s_1^2 \equiv \frac{1}{n}\sum_{i=1}^{n}(y_{i1}-\bar{y}_{\cdot 1})^2$;

$s_2^2 \equiv \frac{1}{n}\sum_{i=1}^{n}(y_{i2}-\bar{y}_{\cdot 2})^2$;

$s_{12} \equiv \frac{1}{n}\sum_{i=1}^{n}(y_{i1}-\bar{y}_{\cdot 1})(y_{i2}-\bar{y}_{\cdot 2})^2$;

These meanings are consistent with the authors cited here, but differ from the meanings given in section 2.

C.1 The Distribution Function of $T_p|r$

Recall that $\begin{pmatrix} y_{11} \\ y_{12} \end{pmatrix} \cdots \begin{pmatrix} y_{n1} \\ y_{n2} \end{pmatrix}$ represents n pairs of observations from

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} \sim N_2\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \sigma^2\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right)$$

and that the paired t-statistic $T_p \equiv \dfrac{(\bar{Y}_{\cdot 1}-\bar{Y}_{\cdot 2})}{\sqrt{\frac{1}{n-1}\left(S_1^2+S_2^2-2rS_1S_2\right)}}$ (see 2).

The derivation is similar to the derivation of the distribution of $T_2|r$ in Appendix B. The steps of the derivation are as follows; the first four are identical to those for $T_2|r$.

i) State the distribution of $\bar{D} \equiv \bar{Y}_{\cdot 2} - \bar{Y}_{\cdot 1}$;

ii) State the joint density of S_1, S_2 and R;

iii) Derive the joint density of $\bar{D}, S_1, S_2$ and R from i) and ii);

iv) Derive the joint density of $\bar{D}$, S_1 and S_2 conditioned on r by dividing iii) by the density of R. It will be found that this conditional joint density may be factorized into a function of $\bar{D}$ and a function of S_1, S_2 and r.

v) Derive the distribution function and hence the density of $W^2 \equiv S_1^2 + S_2^2 - 2rS_1S_2 \,|\, r$ from iv);

vi) Recombine the density of $\bar{D}$ with that of $W|r$ to give the joint density of $\bar{D}$ and W conditioned on r;

vii) Noting that $T_p = \frac{\bar{D}\sqrt{n-1}}{W}$, integrate the density in vi) over the region of $\bar{d}$-w space for which $T_p < t$.

For convenience, we set σ^2 to 1.

The first four steps of the derivation are the same as for the distribution of $T_2|r$ (see section B.1) and we do not reproduce them here. They yield for the joint density of S_1 and S_2 conditioned on r (B.4)

$$\begin{aligned} dP &= \frac{\sqrt{2}\,\Gamma(n-\frac{1}{2})\,n^{n-1}}{\sqrt{\pi}(1-\rho^2)^{n-1}\Gamma^2(n-1)} \exp\left\{-\frac{n}{2(1-\rho^2)}\left(s_1^2 - 2\rho r s_1 s_2 + s_2^2\right)\right\} \\ &\quad \times s_1^{n-2}\, s_2^{n-2}\,(1-\rho r)^{n-3/2}\, H^{-1}\left(\tfrac{1}{2}, \tfrac{1}{2}, n-\tfrac{1}{2}, \tfrac{1}{2}(1+\rho r)\right) ds_1\, ds_2 \qquad \text{(C.1)} \end{aligned}$$

Put $W^2 \equiv s_1^2 + s_2^2 - 2rs_1s_2$; $W > 0$ ($W > 0$ is allowable by Cauchy-Schwartz); then

$$\begin{aligned} P(W \le w \,|\, r) &= P(W^2 \le w^2 \,|\, r) \\ &= P(s_1^2 + s_2^2 - 2rs_1s_2 \le w^2 \,|\, r) \\ &= \int_{s_1, s_2 : s_1^2 + s_2^2 - 2rs_1s_2 \le w^2} f_{s_1, s_2}(s_1, s_2 \,|\, r)\, ds_1\, ds_2 \\ &\propto \int_{s_1, s_2 : s_1^2 + s_2^2 - 2rs_1s_2 \le w^2} \exp\left\{-\frac{n}{2(1-\rho^2)}\left(s_1^2 - 2\rho r s_1 s_2 + s_2^2\right)\right\} s_1^{n-2}\, s_2^{n-2}\, ds_1\, ds_2 \end{aligned}$$

from (C.1). Make the transformation

$$s_1 = \sqrt{\frac{2(1-\rho^2)}{n}u}\sin\theta, \qquad s_2 = \sqrt{\frac{2(1-\rho^2)}{n}u}\cos\theta$$

whose Jacobean is

$$\begin{aligned}\frac{\partial(s_1,s_2)}{\partial(u,\theta)} &= \begin{vmatrix} \sqrt{\frac{(1-\rho^2)}{2nu}}\sin\theta & \sqrt{\frac{(1-\rho^2)}{2nu}}\cos\theta \\ \sqrt{\frac{2(1-\rho^2)}{n}u}\cos\theta & -\sqrt{\frac{2(1-\rho^2)}{n}u}\sin\theta \end{vmatrix} \\ &= -\frac{1-\rho^2}{n}(\sin^2\theta+\cos^2\theta) = -\frac{1-\rho^2}{n}\end{aligned}$$

For the limits of integration, note that $s_1^2+s_2^2-2rs_1s_2 \le w^2 \Rightarrow u \le \frac{nw^2}{2(1-\rho^2)(1-r\sin 2\theta)}$, that s_1, s_2 real $\Rightarrow u \ge 0$ and that $s_1, s_2 \ge 0 \Rightarrow \theta \in [0, \frac{\pi}{2}]$. Thus

$$\begin{aligned}P(W\le w\,|\,r) &\propto \int_{u=0}^{\frac{nw^2}{2(1-\rho^2)(1-r\sin 2\theta)}}\int_{\theta=0}^{\frac{\pi}{2}} \exp\{-u(1-2\rho r\sin\theta\cos\theta)\} \\ &\qquad \times \left(\tfrac{2(1-\rho^2)}{n}\right)^{n-2} u^{n-2}(\sin\theta\cos\theta)^{n-2}\left(\tfrac{1-\rho^2}{n}\right)du\,d\theta \\ &= \left(\tfrac{1-\rho^2}{n}\right)^{n-1}\int_{\theta=0}^{\frac{\pi}{2}}(\sin 2\theta)^{n-2}\int_{u=0}^{\frac{nw^2}{2(1-\rho^2)(1-r\sin 2\theta)}}\sum_{i=0}^{\infty}u^{n+i-2}\frac{(\rho r\sin 2\theta-1)^i}{i!}\,du\,d\theta \\ &= \left(\tfrac{1-\rho^2}{n}\right)^{n-1}\int_{\theta=0}^{\frac{\pi}{2}}(\sin 2\theta)^{n-2}\sum_{i=0}^{\infty}u^{n+i-1}\frac{(\rho r\sin 2\theta-1)^i}{(n+i-1)\,i!}\,d\theta\Bigg|_{u=0}^{\frac{nw^2}{2(1-\rho^2)(1-r\sin 2\theta)}} \\ &= \left(\tfrac{1-\rho^2}{n}\right)^{n-1}\int_{\theta=0}^{\frac{\pi}{2}}(\sin 2\theta)^{n-2}\sum_{i=0}^{\infty}\frac{n^{n+i-1}w^{2(n+i-1)}}{2^{n+i-1}(1-\rho^2)^{n+i-1}(1-r\sin 2\theta)^{n+i-1}}\frac{(\rho r\sin 2\theta-1)^i}{(n+i-1)\,i!}\,d\theta \\ &= \sum_{i=0}^{\infty}\frac{n^i w^{2(n+i-1)}}{2^{n+i-1}(1-\rho^2)^i\,(n+i-1)\,i!}\int_{\theta=0}^{\frac{\pi}{2}}\frac{(\sin 2\theta)^{n-2}\,(\rho r\sin 2\theta-1)^i}{(1-r\sin 2\theta)^{n+i-1}}\,d\theta \\ &= \sum_{i=0}^{\infty}\frac{(-1)^i\,n^i\,w^{2(n+i-1)}}{2^{n+i-1}(1-\rho^2)^i\,(n+i-1)\,i!}\int_{\theta=0}^{\frac{\pi}{2}}\frac{(\sin\theta)^{n-2}\,(1-\rho r\sin\theta)^i}{(1-r\sin\theta)^{n+i-1}}\,d\theta \\ &= \sum_{i=0}^{\infty}\frac{(-1)^i\,n^i\,w^{2(n+i-1)}}{2^{n+i-1}(1-\rho^2)^i\,(n+i-1)\,i!} \\ &\qquad \times\int_{\theta=0}^{\frac{\pi}{2}}(\sin\theta)^{n-2}\sum_{j=0}^{i}\binom{i}{j}(-\rho r\sin\theta)^j\sum_{k=0}^{\infty}\binom{n+i+k-2}{k}(r\sin\theta)^k\,d\theta \\ &= \sum_{i=0}^{\infty}\frac{(-1)^i\,n^i\,w^{2(n+i-1)}}{2^{n+i-1}(1-\rho^2)^i\,(n+i-1)\,i!}\end{aligned}$$

$$\times \sum_{j=0}^{i} \binom{i}{j} (-\rho r)^j \sum_{k=0}^{\infty} \binom{n+i+k-2}{k} r^k \int_{\theta=0}^{\frac{\pi}{2}} (\sin\theta)^{n+j+k-2}\, d\theta$$

$$= \sum_{i=0}^{\infty} \frac{(-1)^i\, n^i\, w^{2(n+i-1)}}{2^{n+i-1}(1-\rho^2)^i\,(n+i-1)\,i!}$$

$$\times \sum_{j=0}^{i} \binom{i}{j} (-\rho r)^j \sum_{k=0}^{\infty} \binom{n+i+k-2}{k} r^k \frac{\sqrt{\pi}\,\Gamma\left(\frac{n+j+k-1}{2}\right)}{2\,\Gamma\left(\frac{n+j+k}{2}\right)}$$

The derivative with respect to w may be taken to give the density

$$dP \;\propto\; \sum_{i=0}^{\infty}\sum_{j=0}^{i}\sum_{k=0}^{\infty} \frac{(-1)^i\, n^i\, w^{2n+2i-3}\, i!\,(-\rho r)^j\,\Gamma(n+i+k-1)\, r^k\, \sqrt{\pi}\,\Gamma\left(\frac{n+j+k-1}{2}\right)}{2^{n+i-2}(1-\rho^2)^i\, i!\,\Gamma(j+1)\,\Gamma(i-j+1)\,\Gamma(k+1)\,\Gamma(n+i-1)\,2\,\Gamma\left(\frac{n+j+k}{2}\right)}\, dw$$

$$= \sum_{i=0}^{\infty}\sum_{j=0}^{i}\sum_{k=0}^{\infty} \frac{(-1)^i\, n^i\, w^{2n+2i-3}\,(-\rho r)^j\,\Gamma(n+i+k-1)\, r^k\, \sqrt{\pi}\,\Gamma\left(\frac{n+j+k-1}{2}\right)}{2^{n+i-1}(1-\rho^2)^i\,\Gamma(j+1)\,\Gamma(i-j+1)\,\Gamma(k+1)\,\Gamma(n+i-1)\,\Gamma\left(\frac{n+j+k}{2}\right)}\, dw$$

Restoring the constants gives ...

$$dP = \frac{\sqrt{2}\,\Gamma(n-\frac{1}{2})\, n^{n-1}}{\sqrt{\pi}(1-\rho^2)^{n-1}\Gamma^2(n-1)} (1-\rho r)^{n-3/2}\, H^{-1}\left(\tfrac{1}{2},\tfrac{1}{2},n-\tfrac{1}{2},\tfrac{1}{2}(1+\rho r)\right)$$

$$\times \sum_{i=0}^{\infty}\sum_{j=0}^{i}\sum_{k=0}^{\infty} \frac{(-1)^i\, n^i\, w^{2n+2i-3}\,(-\rho r)^j\,\Gamma(n+i+k-1)\, r^k\, \sqrt{\pi}\,\Gamma\left(\frac{n+j+k-1}{2}\right)}{2^{n+i-1}(1-\rho^2)^i\,\Gamma(j+1)\,\Gamma(i-j+1)\,\Gamma(k+1)\,\Gamma(n+i-1)\,\Gamma\left(\frac{n+j+k}{2}\right)}\, dw$$

$$= \frac{\sqrt{2}\,\Gamma(n-\frac{1}{2})}{w\,\Gamma^2(n-1)} \left(\frac{nw^2}{2(1-\rho^2)}\right)^{n-1} (1-\rho r)^{n-3/2}\, H^{-1}\left(\tfrac{1}{2},\tfrac{1}{2},n-\tfrac{1}{2},\tfrac{1}{2}(1+\rho r)\right)$$

$$\times \sum_{i=0}^{\infty}\sum_{j=0}^{i}\sum_{k=0}^{\infty} \frac{\Gamma(n+i+k-1)\,\Gamma\left(\frac{n+j+k-1}{2}\right)}{\Gamma(n+i-1)\,\Gamma\left(\frac{n+j+k}{2}\right)\,\Gamma(i-j+1)\,\Gamma(j+1)\,\Gamma(k+1)}$$

$$\times \left(\frac{-nw^2}{2(1-\rho^2)}\right)^i \left(-\rho r\right)^j \left(r\right)^k dw \tag{C.2}$$

Thus, combining (B.1) and (C.2), the joint density of $\bar{D}$ and W conditioned on r is

$$dP = \frac{\sqrt{n}}{2\sqrt{\pi(1-\rho)}} \exp\left\{\frac{-n}{4(1-\rho)}\,\bar{d}^2\right\} \frac{\sqrt{2}\,\Gamma(n-\frac{1}{2})}{w\,\Gamma^2(n-1)} \left(\frac{nw^2}{2(1-\rho^2)}\right)^{n-1} \frac{(1-\rho r)^{n-3/2}}{H\left(\frac{1}{2},\frac{1}{2},n-\frac{1}{2},\frac{1}{2}(1+\rho r)\right)}$$

$$\times \sum_{i=0}^{\infty}\sum_{j=0}^{i}\sum_{k=0}^{\infty} \frac{\Gamma(n+i+k-1)\,\Gamma\left(\frac{n+j+k-1}{2}\right)}{\Gamma(n+i-1)\,\Gamma\left(\frac{n+j+k}{2}\right)\,\Gamma(i-j+1)\,\Gamma(j+1)\,\Gamma(k+1)}$$

$$\times \left(\frac{-nw^2}{2(1-\rho^2)}\right)^i \left(-\rho r\right)^j \left(r\right)^k d\bar{d}\,dw$$

$$= \sqrt{\frac{n}{2\pi(1-\rho)}}\,\frac{\Gamma(n-\frac{1}{2})}{w\,\Gamma^2(n-1)}\left(\frac{nw^2}{2(1-\rho^2)}\right)^{n-1}\frac{(1-\rho r)^{n-3/2}\exp\left\{\frac{-n}{4(1-\rho)}\bar{d}^2\right\}}{H\left(\frac{1}{2},\frac{1}{2},n-\frac{1}{2},\frac{1}{2}(1+\rho r)\right)}$$

$$\times \sum_{i=0}^{\infty}\sum_{j=0}^{i}\sum_{k=0}^{\infty}\frac{\Gamma(n+i+k-1)\,\Gamma\left(\frac{n+j+k-1}{2}\right)}{\Gamma(n+i-1)\,\Gamma\left(\frac{n+j+k}{2}\right)\,\Gamma(i-j+1)\,\Gamma(j+1)\,\Gamma(k+1)}$$

$$\times \left(\frac{-nw^2}{2(1-\rho^2)}\right)^i\left(-\rho r\right)^j\left(r\right)^k\,d\bar{d}\,dw$$

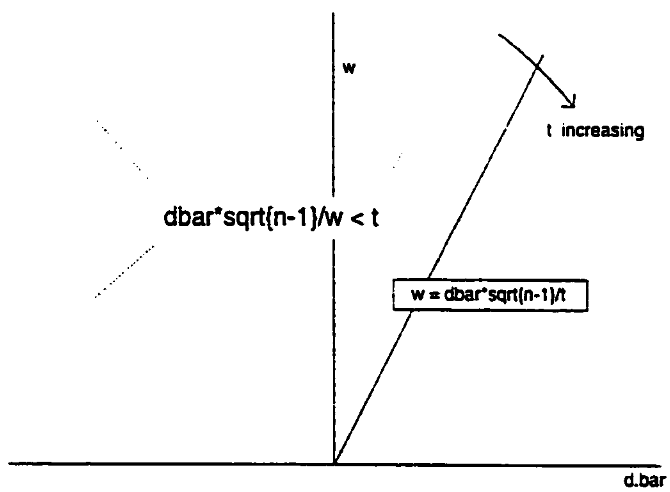
Now...

$$T_p \equiv \frac{\bar{D}\sqrt{n-1}}{W}$$

Thus...

$$P(T_p \le t\,|\,r) = P\left(\frac{\bar{D}\sqrt{n-1}}{W} \le t\,|\,r\right)$$

$$= \int_{\bar{d},w:\frac{\bar{d}\sqrt{n-1}}{w}\le t} f_{\bar{D},W}(\bar{d},w\,|\,r)\,d\bar{d}\,dw$$



$$= \int_{w=0}^{\infty}\int_{\bar{d}=-\infty}^{\frac{tw}{\sqrt{n-1}}} f_{\bar{D},W}(\bar{d},w\,|\,r)\,d\bar{d}\,dw$$

$$= \int_{w=0}^{\infty}\int_{\bar{d}=-\infty}^{0} f_{\bar{D},W}(\bar{d},w\,|\,r)\,d\bar{d}\,dw + \int_{w=0}^{\infty}\int_{\bar{d}=0}^{\frac{tw}{\sqrt{n-1}}} f_{\bar{D},W}(\bar{d},w\,|\,r)\,d\bar{d}\,dw$$

$$= \frac{1}{2} + \int_{w=0}^{\infty}\int_{\bar{d}=0}^{\frac{tw}{\sqrt{n-1}}} f_{\bar{D},W}(\bar{d},w\,|\,r)\,d\bar{d}\,dw$$

by the symmetry of the distribution, (since the density must be the same when $\bar{d} \equiv \bar{y}_1 - \bar{y}_2$ as when $\bar{d} \equiv \bar{y}_2 - \bar{y}_1$);

$$= \frac{1}{2} + \int_{w=0}^{\infty} \int_{\bar{d}=0}^{\frac{tw}{\sqrt{n-1}}} \sqrt{\frac{n}{2\pi(1-\rho)}} \frac{\Gamma(n-\frac{1}{2})}{w\,\Gamma^2(n-1)} \left(\tfrac{nw^2}{2(1-\rho^2)}\right)^{n-1} \frac{(1-\rho r)^{n-3/2} \exp\left\{\frac{-n}{4(1-\rho)}\bar{d}^2\right\}}{H\left(\frac{1}{2},\frac{1}{2},n-\frac{1}{2},\frac{1}{2}(1+\rho r)\right)}$$
$$\times \sum_{i=0}^{\infty}\sum_{j=0}^{i}\sum_{k=0}^{\infty} \frac{\Gamma(n+i+k-1)\,\Gamma\left(\frac{n+j+k-1}{2}\right)}{\Gamma(n+i-1)\,\Gamma\left(\frac{n+j+k}{2}\right)\,\Gamma(i-j+1)\,\Gamma(j+1)\,\Gamma(k+1)}$$
$$\times \left(\tfrac{-nw^2}{2(1-\rho^2)}\right)^i \left(-\rho r\right)^j \left(r\right)^k d\bar{d}\,dw$$

The integral with respect to $\bar{d}$ may be accomplished by expressing the exponential function in $\bar{d}$ as an infinite sum and integrating term by term

$$\begin{aligned}
\int_{\bar{d}=0}^{\frac{tw}{\sqrt{n-1}}} \exp\left\{\tfrac{-n}{4(1-\rho)}\bar{d}^2\right\} d\bar{d} &= \sum_{l=0}^{\infty} \int_{\bar{d}=0}^{\frac{tw}{\sqrt{n-1}}} \left(\frac{-n\bar{d}^2}{4(1-\rho)}\right)^l \frac{1}{l!}\, d\bar{d} \\
&= \sum_{l=0}^{\infty} \left(\frac{-n}{4(1-\rho)}\right)^l \frac{1}{l!} \frac{\bar{d}^{2l+1}}{2l+1}\Bigg|_{\bar{d}=0}^{\frac{tw}{\sqrt{n-1}}} \\
&= \sum_{l=0}^{\infty} \left(\frac{-n}{4(1-\rho)}\right)^l \frac{1}{l!} \frac{1}{2l+1} \left(\frac{tw}{\sqrt{n-1}}\right)^{2l+1} \\
&= \sum_{l=0}^{\infty} \frac{\Gamma(2l+1)}{\Gamma(2l+2)\,\Gamma(l+1)} \left(\tfrac{tw}{\sqrt{n-1}}\right) \left(\tfrac{-n\,t^2 w^2}{4(1-\rho)(n-1)}\right)^l
\end{aligned}$$

Since we must next integrate w over the interval $[0,\infty)$, we express the power series as a confluent hypergeometric function, apply Kummer's transformation (see [1]), and re-express the resulting confluent hypergeometric function as a power series in order to give a convergent integral. Thus

$$\begin{aligned}
\int_{\bar{d}=0}^{\frac{tw}{\sqrt{n-1}}} \exp\left\{\tfrac{-n}{4(1-\rho)}\bar{d}^2\right\} d\bar{d} &= \left(\tfrac{tw}{\sqrt{n-1}}\right) \sum_{l=0}^{\infty} \frac{\Gamma(l+\frac{1}{2})\,\Gamma(l+1)\,2^{2l}/\sqrt{\pi}}{\Gamma(l+1)\,\Gamma(l+\frac{3}{2})\,2^{2l+1}/\sqrt{\pi}\,\Gamma(l+1)} \left(\tfrac{-n\,t^2 w^2}{4(1-\rho)(n-1)}\right)^l \\
&= \left(\tfrac{tw}{2\sqrt{n-1}}\right) \sum_{l=0}^{\infty} \frac{\Gamma(l+\frac{1}{2})}{\Gamma(l+\frac{3}{2})\,\Gamma(l+1)} \left(\tfrac{-n\,t^2 w^2}{4(1-\rho)(n-1)}\right)^l \\
&= \left(\tfrac{tw}{2\sqrt{n-1}}\right) \frac{\Gamma(\frac{1}{2})}{\Gamma(\frac{3}{2})} M\left(\tfrac{1}{2}, \tfrac{3}{2}; \tfrac{-n\,t^2 w^2}{4(1-\rho)(n-1)}\right) \\
&= \left(\tfrac{tw}{2\sqrt{n-1}}\right) \frac{\Gamma(\frac{1}{2})}{\Gamma(\frac{3}{2})} e^{\frac{-n\,t^2 w^2}{4(1-\rho)(n-1)}} M\left(1, \tfrac{3}{2}; \tfrac{n\,t^2 w^2}{4(1-\rho)(n-1)}\right) \\
&= \left(\tfrac{tw}{2\sqrt{n-1}}\right) \frac{\Gamma(\frac{1}{2})}{\Gamma(\frac{3}{2})} e^{\frac{-n\,t^2 w^2}{4(1-\rho)(n-1)}} \frac{\Gamma(\frac{3}{2})}{\Gamma(1)} \sum_{l=0}^{\infty} \frac{\Gamma(l+1)}{\Gamma(l+\frac{3}{2})\,\Gamma(l+1)} \left(\tfrac{n\,t^2 w^2}{4(1-\rho)(n-1)}\right)^l
\end{aligned}$$

$$= \left(\frac{tw\sqrt{\pi}}{2\sqrt{n-1}}\right) e^{\frac{-nt^2w^2}{4(1-\rho)(n-1)}} \sum_{l=0}^{\infty} \frac{1}{\Gamma\left(l+\frac{3}{2}\right)} \left(\frac{nt^2w^2}{4(1-\rho)(n-1)}\right)^l$$

Recombining the terms gives

$$\begin{aligned}
P(T_p \le t\,|\,r) &= \frac{1}{2} + \int_{w=0}^{\infty} \sqrt{\frac{n}{2\pi(1-\rho)}} \frac{\Gamma(n-\frac{1}{2})}{w\,\Gamma^2(n-1)} \left(\frac{nw^2}{2(1-\rho^2)}\right)^{n-1} \frac{(1-\rho r)^{n-3/2}}{H\left(\frac{1}{2},\frac{1}{2},n-\frac{1}{2},\frac{1}{2}(1+\rho r)\right)} \\
&\times \sum_{i=0}^{\infty}\sum_{j=0}^{i}\sum_{k=0}^{\infty} \frac{\Gamma(n+i+k-1)\,\Gamma\left(\frac{n+j+k-1}{2}\right)}{\Gamma(n+i-1)\,\Gamma\left(\frac{n+j+k}{2}\right)\Gamma(i-j+1)\,\Gamma(j+1)\,\Gamma(k+1)} \\
&\times \left(\frac{-nw^2}{2(1-\rho^2)}\right)^i \left(-\rho r\right)^j \left(r\right)^k \left(\frac{tw\sqrt{\pi}}{2\sqrt{n-1}}\right) e^{\frac{-nt^2w^2}{4(1-\rho)(n-1)}} \sum_{l=0}^{\infty} \frac{1}{\Gamma\left(l+\frac{3}{2}\right)} \left(\frac{nt^2w^2}{4(1-\rho)(n-1)}\right)^l dw \\
&= \frac{1}{2} + \int_{w=0}^{\infty} \frac{t}{2} \sqrt{\frac{n}{2(1-\rho)(n-1)}} \frac{\Gamma(n-\frac{1}{2})}{\Gamma^2(n-1)} \frac{(1-\rho r)^{n-3/2}}{H\left(\frac{1}{2},\frac{1}{2},n-\frac{1}{2},\frac{1}{2}(1+\rho r)\right)} \\
&\times \sum_{i,k,l=0}^{\infty}\sum_{j=0}^{i} \frac{\Gamma(n+i+k-1)\,\Gamma\left(\frac{n+j+k-1}{2}\right)}{\Gamma(n+i-1)\,\Gamma\left(\frac{n+j+k}{2}\right)\Gamma(i-j+1)\,\Gamma\left(l+\frac{3}{2}\right)\Gamma(j+1)\,\Gamma(k+1)} \\
&\times (-1)^i \left(\frac{nw^2}{2(1-\rho^2)}\right)^{n+i-1} \left(-\rho r\right)^j \left(r\right)^k e^{\frac{-nt^2w^2}{4(1-\rho)(n-1)}} \left(\frac{nt^2w^2}{4(1-\rho)(n-1)}\right)^l dw
\end{aligned}$$

Put $u = u(w) = \frac{nt^2w^2}{4(1-\rho)(n-1)}$. $w^2 = \frac{4u(1-\rho)(n-1)}{nt^2}$. $w = \frac{1}{t}\sqrt{\frac{4u(1-\rho)(n-1)}{n}}$. $\frac{du}{dw} = \frac{2nt^2w}{4(1-\rho)(n-1)}$. $dw = \frac{4(1-\rho)(n-1)}{2nt^2w}du = \frac{4(1-\rho)(n-1)}{2nt^2}t\sqrt{\frac{n}{4u(1-\rho)(n-1)}}du = \frac{1}{t}\sqrt{\frac{(1-\rho)(n-1)}{un}}du$. However, the transformation removes the distinction between positive and negative t. We therefore add $\frac{1}{2}$ for all the probability in $t<0$ and confine our result to $t \ge 0$. The integral may then be evaluated

$$\begin{aligned}
P(T_p \le t\,|\,r) &= 1 + \int_{u=0}^{\infty} \frac{t}{2} \sqrt{\frac{n}{2(1-\rho)(n-1)}} \frac{\Gamma(n-\frac{1}{2})}{\Gamma^2(n-1)} \frac{(1-\rho r)^{n-3/2}}{H\left(\frac{1}{2},\frac{1}{2},n-\frac{1}{2},\frac{1}{2}(1+\rho r)\right)} \\
&\times \sum_{i,k,l=0}^{\infty}\sum_{j=0}^{i} \frac{\Gamma(n+i+k-1)\,\Gamma\left(\frac{n+j+k-1}{2}\right)}{\Gamma(n+i-1)\,\Gamma\left(\frac{n+j+k}{2}\right)\Gamma(i-j+1)\,\Gamma\left(l+\frac{3}{2}\right)\Gamma(j+1)\,\Gamma(k+1)} \\
&\times (-1)^i \left(\frac{n}{2(1-\rho^2)}\frac{4u(1-\rho)(n-1)}{nt^2}\right)^{n+i-1} \left(-\rho r\right)^j \left(r\right)^k e^{-u} \left(u\right)^l \frac{1}{t}\sqrt{\frac{(1-\rho)(n-1)}{un}}\,du \\
&= 1 + \frac{1}{2}\sqrt{\frac{1}{2}} \frac{\Gamma(n-\frac{1}{2})}{\Gamma^2(n-1)} \frac{(1-\rho r)^{n-3/2}}{H\left(\frac{1}{2},\frac{1}{2},n-\frac{1}{2},\frac{1}{2}(1+\rho r)\right)} \\
&\times \sum_{i,k,l=0}^{\infty}\sum_{j=0}^{i} \frac{\Gamma(n+i+k-1)\,\Gamma\left(\frac{n+j+k-1}{2}\right)}{\Gamma(n+i-1)\,\Gamma\left(\frac{n+j+k}{2}\right)\Gamma(i-j+1)\,\Gamma\left(l+\frac{3}{2}\right)\Gamma(j+1)\,\Gamma(k+1)} \\
&\times (-1)^i \left(\frac{2(n-1)}{t^2(1+\rho)}\right)^{n+i-1} \left(-\rho r\right)^j \left(r\right)^k \int_{u=0}^{\infty} e^{-u} \left(u\right)^{(n+i+l-\frac{1}{2})-1} du
\end{aligned}$$

$$= 1+\frac{1}{2}\sqrt{\frac{1}{2}}\,\frac{\Gamma(n-\frac{1}{2})}{\Gamma^2(n-1)}\,\frac{(1-\rho r)^{n-3/2}}{H\left(\frac{1}{2},\frac{1}{2},n-\frac{1}{2},\frac{1}{2}(1+\rho r)\right)}$$

$$\times \sum_{i,k,l=0}^{\infty}\sum_{j=0}^{i}\frac{\Gamma(n+i+k-1)\,\Gamma\left(\frac{n+j+k-1}{2}\right)}{\Gamma(n+i-1)\,\Gamma\left(\frac{n+j+k}{2}\right)\,\Gamma(i-j+1)\,\Gamma(l+\frac{3}{2})\,\Gamma(j+1)\,\Gamma(k+1)}$$

$$\times (-1)^i\left(\frac{2(n-1)}{t^2(1+\rho)}\right)^{n+i-1}\left(-\rho r\right)^j\left(r\right)^k\,\Gamma(n+i+l-\tfrac{1}{2})$$

$$= 1+\frac{1}{2}\sqrt{\frac{1}{2}}\,\frac{\Gamma(n-\frac{1}{2})}{\Gamma^2(n-1)}\left(\frac{2(n-1)}{t^2(1+\rho)}\right)^{n-1}\frac{(1-\rho r)^{n-3/2}}{H\left(\frac{1}{2},\frac{1}{2},n-\frac{1}{2},\frac{1}{2}(1+\rho r)\right)}$$

$$\times \sum_{i,k,l=0}^{\infty}\sum_{j=0}^{i}\frac{\Gamma(n+i+k-1)\,\Gamma\left(\frac{n+j+k-1}{2}\right)\Gamma(n+i+l-\frac{1}{2})}{\Gamma(n+i-1)\,\Gamma\left(\frac{n+j+k}{2}\right)\,\Gamma(i-j+1)\,\Gamma(l+\frac{3}{2})\,\Gamma(j+1)\,\Gamma(k+1)}$$

$$\times \left(\frac{-2(n-1)}{t^2(1+\rho)}\right)^i\left(-\rho r\right)^j\left(r\right)^k$$

The l series may be expressed in closed form (see Erdélyi [12], p. 111 (46))

$$\begin{aligned}
\sum_{l=0}^{\infty}\frac{\Gamma(n+i+l-\frac{1}{2})}{\Gamma(l+\frac{3}{2})} &= \sum_{l=0}^{\infty}\frac{\Gamma(n+i-\frac{1}{2}+l)\,\Gamma(1+l)}{\Gamma(\frac{3}{2}+l)\,\Gamma(l+1)}\,1^l \\
&= \sum_{l=0}^{\infty}\frac{\Gamma(n+i-\frac{1}{2}+l)\,\Gamma(1+l)}{\Gamma(\frac{3}{2}+l)\,\Gamma(l+1)}\,1^l \\
&= \frac{\Gamma(n+i-\frac{1}{2})}{\Gamma(\frac{3}{2})}\,H(n+i-\tfrac{1}{2},1,\tfrac{3}{2};1) \\
&= \frac{\Gamma(n+i-\frac{1}{2})}{\Gamma(\frac{3}{2})}\,\frac{\Gamma(\frac{3}{2})\,\Gamma(\frac{3}{2}-n-i+\frac{1}{2},-1)}{\Gamma(\frac{3}{2}-n-i+\frac{1}{2},)\,\Gamma(\frac{3}{2}-1)} \\
&= \frac{\Gamma(n+i-\frac{1}{2})}{\Gamma(\frac{1}{2})}\,\frac{\Gamma(\frac{3}{2}-n-i+\frac{1}{2}-1)}{\Gamma(\frac{3}{2}-n-i+\frac{1}{2}-1)\,(\frac{3}{2}-n-i+\frac{1}{2}-1)} \\
&= \frac{\Gamma(n+i-\frac{1}{2})}{\Gamma(\frac{1}{2})}\,\frac{-1}{(n+i-1)}
\end{aligned}$$

Recombining the terms and un-diagonalizing the i,j infinite sums by $\sum_{i=0}^{\infty}\sum_{j=0}^{i} c(i-j,j)\,x^{i-j}\,y^j = \sum_{i,j=0}^{\infty} c(i,j)\,x^i\,y^j$ therefore gives, for $t\geq 0$

$$P(T_p\leq t\,|\,r) = 1+\frac{1}{2}\sqrt{\frac{1}{2}}\,\frac{\Gamma(n-\frac{1}{2})}{\Gamma^2(n-1)}\left(\frac{2(n-1)}{t^2(1+\rho)}\right)^{n-1}\frac{(1-\rho r)^{n-3/2}}{H\left(\frac{1}{2},\frac{1}{2},n-\frac{1}{2},\frac{1}{2}(1+\rho r)\right)}$$

$$\times \sum_{i,k=0}^{\infty}\sum_{j=0}^{i}\frac{\Gamma(n+i+k-1)\,\Gamma\left(\frac{n+j+k-1}{2}\right)}{\Gamma(n+i-1)\,\Gamma\left(\frac{n+j+k}{2}\right)\,\Gamma(i-j+1)\,\Gamma(j+1)\,\Gamma(k+1)}$$

$$\times \left(\tfrac{-2(n-1)}{t^2(1+\rho)}\right)^i \left(-\rho r\right)^j \left(r\right)^k \frac{\Gamma(n+i-\frac{1}{2})}{\Gamma(\frac{1}{2})} \frac{-1}{(n+i-1)}$$

$$= 1 - \tfrac{1}{2}\sqrt{\tfrac{1}{2\pi}}\,\frac{\Gamma(n-\frac{1}{2})}{\Gamma^2(n-1)} \left(\tfrac{2(n-1)}{t^2(1+\rho)}\right)^{n-1} \frac{(1-\rho r)^{n-3/2}}{H\left(\frac{1}{2},\frac{1}{2},n-\frac{1}{2},\frac{1}{2}(1+\rho r)\right)}$$

$$\times \sum_{i,k=0}^{\infty}\sum_{j=0}^{i} \frac{\Gamma(n+i-\frac{1}{2})\,\Gamma(n+i+k-1)\,\Gamma\left(\frac{n+j+k-1}{2}\right)}{\Gamma(n+i)\,\Gamma\left(\frac{n+j+k}{2}\right)\,\Gamma(i-j+1)\,\Gamma(j+1)\,\Gamma(k+1)}$$

$$\times \left(\tfrac{-2(n-1)}{t^2(1+\rho)}\right)^i \left(-\rho r\right)^j \left(r\right)^k$$

$$= 1 - \frac{\Gamma(n-\frac{1}{2})\,(1-\rho r)^{n-3/2}}{2\sqrt{2\pi}\,\Gamma^2(n-1)\,H\left(\frac{1}{2},\frac{1}{2},n-\frac{1}{2},\frac{1}{2}(1+\rho r)\right)} \left(\tfrac{2(n-1)}{t^2(1+\rho)}\right)^{n-1}$$

$$\times \sum_{i,j,k=0}^{\infty} \frac{\Gamma(n+i+j-\frac{1}{2})\,\Gamma(n+i+j+k-1)\,\Gamma\left(\frac{n+j+k-1}{2}\right)}{\Gamma(n+i+j)\,\Gamma\left(\frac{n+j+k}{2}\right)\,\Gamma(i+1)\,\Gamma(j+1)\,\Gamma(k+1)}$$

$$\times \left(\tfrac{-2(n-1)}{t^2(1+\rho)}\right)^i \left(\tfrac{2\rho r(n-1)}{t^2(1+\rho)}\right)^j \left(r\right)^k$$

C.2 A convergent form of the distribution function

A form of the distribution function which will converge for all values of the arguments may be obtained by expressing the i series as a Gauss hypergeometric function and applying the transformation $H(\alpha,\beta,\gamma;x) \equiv (1-x)^{-\alpha}H(\alpha,\gamma-\beta,\gamma;\frac{x}{x-1})$ (see Erdélyi [12], p.105), followed by re-expressing the resulting Gauss hypergeometric function as a power series, in a manner similar to that used for the distribution of $T_2|\rho$ (see A.2) . This yields, for $t \geq 0$

$$P(T_p \leq t\,|\,r) = 1 - \frac{\Gamma(n-\frac{1}{2})\,(1-\rho r)^{n-3/2}}{2\pi\sqrt{2}\,\Gamma^2(n-1)\,H\left(\frac{1}{2},\frac{1}{2},n-\frac{1}{2},\frac{1}{2}(1+\rho r)\right)} \left(\tfrac{2(n-1)}{t^2(1+\rho)+2(n-1)}\right)^{n-1}$$

$$\times \sum_{i,j,k=0}^{\infty} \frac{\Gamma(i+\frac{1}{2})\,\Gamma(n+j-\frac{1}{2})\,\Gamma(n+i+j+k-1)\,\Gamma\left(\frac{n+j+k-1}{2}\right)}{\Gamma(n+i+j)\,\Gamma\left(\frac{n+j+k}{2}\right)\,\Gamma(i+1)\,\Gamma(j+1)\,\Gamma(k+1)}$$

$$\times \left(\tfrac{2(n-1)}{t^2(1+\rho)+2(n-1)}\right)^i \left(\tfrac{2\rho r(n-1)}{t^2(1+\rho)+2(n-1)}\right)^j \left(\tfrac{r\,t^2(1+\rho)}{t^2(1+\rho)+2(n-1)}\right)^k$$

Appendix D

NEGATIVE CORRELATION FROM TREATMENT×MATCHING INTERACTION

In section 1.3, 'Models and Parameterizations for Matched Designs', at page 13, conditions under which an additive random effects model allows a negative correlation between the observations in the treatment and control groups are discussed. The complete derivation of the correlation is provided here.

Let X be the matching variable, Y be the response. Suppose

$$Y = \beta_0 + \beta_1 X + \beta_2 I_{(\text{treatment})} + \beta_3 X I_{(\text{treatment})} + \epsilon$$

where $X \sim (0, \sigma_x^2)$, $\epsilon \sim (0, \sigma^2)$, $X \perp \epsilon$, and

a) β_0 represents the overall mean;
b) β_1 represents the matching effect;
c) β_2 represents the treatment effect;
d) β_3 represents the treatment×matching interaction.

Put $Y_1 \equiv Y | I_{(\text{treatment})} = 0$, $Y_2 \equiv Y | I_{(\text{treatment})} = 1$.

$$\begin{aligned}
Y_1 &= \beta_0 + \beta_1 X + \epsilon \\
Y_2 &= \beta_0 + \beta_1 X + \beta_2 + \beta_3 X + \epsilon \\
E[Y_1] &= E[E[Y_1|X]] = E[\beta_0 + \beta_1 X] = \beta_0 \\
E[Y_2] &= E[E[Y_2|X]] = E[\beta_0 + \beta_1 X + \beta_2 + \beta_3 X] = \beta_0 + \beta_2 \\
Var(Y_1) &= \beta_1^2 \sigma_x^2 + \sigma^2 \\
Var(Y_2) &= (\beta_1 + \beta_3)^2 \sigma_x^2 + \sigma^2
\end{aligned}$$

$E[(Y_1 - E[Y_1])(Y_2 - E[Y_2])]$

$$
\begin{aligned}
&= E[(Y_1 - \beta_0)(Y_2 - \beta_0 - \beta_2)] \\
&= E[Y_1Y_2] - E[Y_1](\beta_0 + \beta_2) - \beta_0 E[Y_2] + \beta_0(\beta_0 + \beta_2)] \\
&= E[E[Y_1Y_2|X]] - \beta_0(\beta_0 + \beta_2) - \beta_0(\beta_0 + \beta_2) + \beta_0(\beta_0 + \beta_2)] \\
&= E[E[Y_1|X].E[Y_2|X]] - \beta_0(\beta_0 + \beta_2) \\
&\qquad \text{since conditional on } X,\ Y_1 \perp Y_2 \\
&= E[(\beta_0 + \beta_1 X)(\beta_0 + \beta_1 X + \beta_2 + \beta_3 X)] - \beta_0(\beta_0 + \beta_2) \\
&= E[\beta_0(\beta_0 + \beta_2) + \beta_0(\beta_1 + \beta_3)X + \beta_1(\beta_0 + \beta_2)X + \beta_1(\beta_1 + \beta_3)X^2] - \beta_0(\beta_0 + \beta_2) \\
&= \beta_1(\beta_1 + \beta_3)E[X^2] \\
&= \beta_1(\beta_1 + \beta_3)\sigma_x^2
\end{aligned}
$$

$$
Corr(Y_1, Y_2) = \frac{\beta_1(\beta_1 + \beta_3)\sigma_x^2}{\sqrt{(\beta_1^2\sigma_x^2 + \sigma^2)((\beta_1 + \beta_3)^2\sigma_x^2 + \sigma^2)}}
$$

$$
\begin{aligned}
Corr(Y_1, Y_2) < 0 &\Leftrightarrow \beta_1(\beta_1 + \beta_3)\sigma_x^2 < 0 \\
&\Leftrightarrow \beta_1(\beta_1 + \beta_3) < 0 \\
&\Leftrightarrow |\beta_3| > |\beta_1| \text{ AND } \beta_1\beta_3 < 0
\end{aligned}
$$

What this means is that the correlation between two outcome variables can be negative if the treatment×matching interaction effect is greater than, and in the opposite direction to, the main treatment effect.

Appendix E

EQUATING COEFFICIENTS

In section 4.5, we outline a proof of theorem 1 at $\beta_{MU}=0$ for small values of δ. The proof relies on expressing the critical value c as a power series in δ. Details of the derivation are as follows.

From (A.3) we have

$$F_{T_2}(t) = \frac{1}{2} + \frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{\sqrt{2}\,\pi\,\Gamma\left(\frac{n-1}{2}\right)} \times$$
$$\sum_{i,j=0}^{\infty} \sum_{k=0}^{2j+1} \frac{\Gamma\left(\frac{n-1}{2}+i\right)\Gamma\left(n+i-1+\frac{k}{2}\right)\Gamma\left(j+\frac{1}{2}\right)(-2)^j}{\Gamma(n+i-1)\,\Gamma(2j-k+2)\,\Gamma(i+1)\,\Gamma(k+1)} \left(\frac{2\rho}{1+\rho}\right)^i \left(-\delta_\rho\right)^{2j-k+1} \left(\frac{t}{\sqrt{n-1}}\right)^k$$

Thus when $\beta_{MU}=\frac{1}{2}$

$$0 = \frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{\sqrt{2}\,\pi\,\Gamma\left(\frac{n-1}{2}\right)} \times$$
$$\sum_{i,j=0}^{\infty} \sum_{k=0}^{2j+1} \frac{\Gamma\left(\frac{n-1}{2}+i\right)\Gamma\left(n+i-1+\frac{k}{2}\right)\Gamma\left(j+\frac{1}{2}\right)(-2)^j}{\Gamma(n+i-1)\,\Gamma(2j-k+2)\,\Gamma(i+1)\,\Gamma(k+1)} \left(\frac{2\rho}{1+\rho}\right)^i \left(-\delta_\rho\right)^{2j-k+1} \left(\frac{c}{\sqrt{n-1}}\right)^k$$

Put

$$c = \sum_{l=0}^{\infty} c_l \delta^l = c_0 + c_1\delta + c_2\delta^2 + c_3\delta^3 + \cdots$$

Thus

$$\begin{aligned} c^k &= \left(c_0 + c_1\delta + c_2\delta^2 + \cdots\right)^k \\ &= c_0^k + k\,c_0^{k-1}\left(c_1\delta + c_2\delta^2\right) + \frac{k(k-1)}{2} c_0^{k-2}\left(c_1\delta\right)^2 + O\left(\delta^3\right) \end{aligned}$$

and

$$0 = \frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{\sqrt{2}\,\pi\,\Gamma\left(\frac{n-1}{2}\right)} \times$$

$$\sum_{i,j=0}^{\infty}\sum_{k=0}^{2j+1}\frac{\Gamma(\frac{n-1}{2}+i)\,\Gamma(n+i-1+\frac{k}{2})\,\Gamma(j+\frac{1}{2})\,(-2)^j}{\Gamma(n+i-1)\,\Gamma(2j-k+2)\,\Gamma(i+1)\,\Gamma(k+1)}\left(\tfrac{2\rho}{1+\rho}\right)^i\left(-\delta_\rho\right)^{2j-k+1}\left(\tfrac{1}{\sqrt{n-1}}\right)^k\times$$
$$\left(c_0^k+k\,c_0^{k-1}\,(c_1\delta+c_2\delta^2)+\frac{k(k-1)}{2}c_0^{k-2}\,(c_1\delta)^2+O(\delta^3)\right)$$

Equating coefficients of δ^0, which only occurs when $k=2j+1$ gives

$$0 = \frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{\sqrt{2}\,\pi\,\Gamma(\frac{n-1}{2})}\times$$
$$\sum_{i,j=0}^{\infty}\frac{\Gamma(\frac{n-1}{2}+i)\,\Gamma\left(n+i-1+\frac{2j+1}{2}\right)\Gamma(j+\frac{1}{2})\,(-2)^j}{\Gamma(n+i-1)\,\Gamma(i+1)\,\Gamma(2j+2)}\left(\tfrac{2\rho}{1+\rho}\right)^i\left(\tfrac{1}{\sqrt{n-1}}\right)^{2j+1}c_0^{2j+1}$$

Clearly $c_0=0$. Then

$$\begin{aligned} c^k &= \left(c_1\delta+c_2\delta^2+\cdots\right)^k \\ &= (c_1\delta)^k+k\,(c_1\delta)^{k-1}(c_2\delta^2)+O(\delta^3) \end{aligned}$$

and

$$0 = \frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{\sqrt{2}\,\pi\,\Gamma(\frac{n-1}{2})}\times$$
$$\sum_{i,j=0}^{\infty}\sum_{k=0}^{2j+1}\frac{\Gamma(\frac{n-1}{2}+i)\,\Gamma(n+i-1+\frac{k}{2})\,\Gamma(j+\frac{1}{2})\,(-2)^j}{\Gamma(n+i-1)\,\Gamma(2j-k+2)\,\Gamma(i+1)\,\Gamma(k+1)}\left(\tfrac{2\rho}{1+\rho}\right)^i\left(-\delta_\rho\right)^{2j-k+1}\left(\tfrac{1}{\sqrt{n-1}}\right)^k\times$$
$$\left((c_1\delta)^k+k\,(c_1\delta)^{k-1}(c_2\delta^2)+O(\delta^3)\right)\qquad\text{(E.1)}$$

Equating coefficients of δ, which only occurs when $j=0$ gives

$$0 = \frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{\sqrt{2}\,\pi\,\Gamma(\frac{n-1}{2})}\times$$
$$\sum_{i=0}^{\infty}\sum_{k=0}^{1}\frac{\Gamma(\frac{n-1}{2}+i)\,\Gamma(n+i-1+\frac{k}{2})\,\Gamma(\frac{1}{2})}{\Gamma(n+i-1)\,\Gamma(2-k)\,\Gamma(i+1)\,\Gamma(k+1)}\left(\tfrac{2\rho}{1+\rho}\right)^i\left(\tfrac{-1}{\sqrt{1-\rho}}\right)^{-k+1}\left(\tfrac{c_1}{\sqrt{n-1}}\right)^k$$

$$\frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{\Gamma(\frac{n-1}{2})}\sum_{i=0}^{\infty}\frac{\Gamma(\frac{n-1}{2}+i)}{\Gamma(i+1)}\left(\tfrac{2\rho}{1+\rho}\right)^i\left(\tfrac{1}{\sqrt{1-\rho}}\right)$$
$$= \frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{\Gamma(\frac{n-1}{2})}\sum_{i=0}^{\infty}\frac{\Gamma(\frac{n-1}{2}+i)\,\Gamma(n+i-1+\frac{1}{2})}{\Gamma(n+i-1)\,\Gamma(i+1)}\left(\tfrac{2\rho}{1+\rho}\right)^i\left(\tfrac{c_1}{\sqrt{n-1}}\right)$$

On the lefthand side, note that $\frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{\Gamma(\frac{n-1}{2})}\sum_{i=0}^{\infty}\frac{\Gamma(\frac{n-1}{2}+i)}{\Gamma(i+1)}\left(\frac{2\rho}{1+\rho}\right)^i = 1$. On the righthand side, the sum may be expressed as a Gauss hypergeometric function and the transformations $H(\alpha,\beta,\gamma;x)\equiv(1-x)^{\gamma-\alpha-\beta}H((\gamma-\alpha,\gamma-\beta,\gamma;x)$ (Erdélyi [12], p.105) and $H(\alpha,\beta,2\beta;x)\equiv(1-\frac{x}{2})^{-\alpha}H\left(\frac{\alpha}{2},\frac{\alpha+1}{2},\beta+\frac{1}{2};\left(\frac{x}{2-x}\right)^2\right)$ (Erdélyi [12], p.111) applied in turn.

$$\begin{aligned}
\left(\frac{1}{\sqrt{1-\rho}}\right) &= \left(\frac{c_1}{\sqrt{n-1}}\right)\frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{\Gamma(\frac{n-1}{2})}\frac{\Gamma(\frac{n-1}{2})\,\Gamma(n-\frac{1}{2})}{\Gamma(n-1)}H\left(\frac{n-1}{2},n-\frac{1}{2},n-1,\frac{2\rho}{1+\rho}\right)\\
&= \left(c_1\right)\frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}\Gamma(n-\frac{1}{2})}{\Gamma(n-1)\sqrt{n-1}}\left(\frac{1-\rho}{1+\rho}\right)^{-\frac{n}{2}}H\left(\frac{n-1}{2},-\frac{1}{2},n-1,\frac{2\rho}{1+\rho}\right)\\
&= \left(c_1\right)\frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{-1}{2}}\Gamma(n-\frac{1}{2})}{\Gamma(n-1)\sqrt{n-1}}\left(\frac{1}{1+\rho}\right)^{\frac{1}{2}}H(\tfrac{1}{4},-\tfrac{1}{4},\tfrac{n}{2};\rho^2)\\
c_1 &= \frac{\sqrt{n-1}\,\Gamma(n-1)}{\Gamma(n-\frac{1}{2})\,H(\frac{1}{4},-\frac{1}{4},\frac{n}{2};\rho^2)}
\end{aligned}$$

Referring to (E.1), equating coefficients of δ^2, which only occurs when $j=0$ gives

$$\begin{aligned}
0 &= \frac{\left(\frac{1-\rho}{1+\rho}\right)^{\frac{n-1}{2}}}{\sqrt{2}\,\pi\,\Gamma(\frac{n-1}{2})}\times\\
&\sum_{i=0}^{\infty}\sum_{k=0}^{1}\frac{\Gamma(\frac{n-1}{2}+i)\,\Gamma(n+i-1+\frac{k}{2})\,\Gamma(\frac{1}{2})}{\Gamma(n+i-1)\,\Gamma(2-k)\,\Gamma(i+1)\,\Gamma(k+1)}\left(\frac{2\rho}{1+\rho}\right)^i\left(\frac{-1}{\sqrt{1-\rho}}\right)^{-k+1}\left(\frac{1}{\sqrt{n-1}}\right)^k k\,(c_1)^{k-1}(c_2)
\end{aligned}$$

and it may be seen that $c_2=0$. Thus

$$c = \delta\frac{\sqrt{n-1}\,\Gamma(n-1)}{\Gamma(n-\frac{1}{2})\,H(\frac{1}{4},-\frac{1}{4},\frac{n}{2};\rho^2)}+O(\delta^3)$$

Put

$$\left(H\left(\tfrac{1}{4},-\tfrac{1}{4},\tfrac{n}{2};\rho^2\right)\right)^{-1} = h_0+h_1\rho^2+h_2\rho^4+h_3\rho^6+\cdots$$

Then, expanding the Gauss hypergeometric function

$$H\left(\tfrac{1}{4},-\tfrac{1}{4},\tfrac{n}{2};\rho^2\right)\left[h_0+h_1\rho^2+h_2\rho^4+h_3\rho^6+\cdots\right] = 1$$

$$1 = \sum_{i=0}^{\infty}\frac{\Gamma(\frac{1}{4}+i)\,\Gamma(-\frac{1}{4}+i)\,\Gamma(\frac{n}{2})}{\Gamma(\frac{1}{4})\,\Gamma(-\frac{1}{4})\,\Gamma(\frac{n}{2}+i)}\frac{\rho^{2i}}{i!}\left[h_0+h_1\rho^2+h_2\rho^4+h_3\rho^6+\cdots\right]$$

Equate coefficients of ρ^0 ...

$$h_0 = 1$$

Equating coefficients of ρ^2

$$\begin{aligned} 0 &= h_1 + \frac{\left(-\frac{1}{4}\right)\left(\frac{1}{4}\right)}{\left(\frac{n}{2}\right)} h_0 \\ h_1 &= \frac{1}{8n} \end{aligned}$$

Equating coefficients of ρ^4

$$\begin{aligned} 0 &= h_2 + \frac{\left(-\frac{1}{4}\right)\left(\frac{1}{4}\right)}{\left(\frac{n}{2}\right)} h_1 + \frac{\left(-\frac{1}{4}\right)\left(\frac{3}{4}\right)\left(\frac{1}{4}\right)\left(\frac{5}{4}\right)}{\left(\frac{n}{2}\right)\left(\frac{n}{2}+1\right)} \frac{1}{2} h_0 \\ &= h_2 + \frac{\left(-\frac{1}{4}\right)\left(\frac{1}{4}\right)}{\left(\frac{n}{2}\right)} h_1 + \frac{\left(\frac{3}{4}\right)\left(\frac{5}{4}\right)}{\left(\frac{n}{2}+1\right)} \frac{1}{2} (-h_1) \\ h_2 &= \frac{1}{8n} \left(\frac{\left(\frac{1}{4}\right)\left(\frac{1}{4}\right)}{\left(\frac{n}{2}\right)} + \frac{\left(\frac{3}{4}\right)\left(\frac{5}{4}\right)}{\left(\frac{n}{2}+1\right)} \frac{1}{2} \right) \\ &= \frac{1}{16(8n)} \left(\frac{2}{n} + \frac{15}{(n+2)} \right) \\ &= \frac{1}{16(8n)} \left(\frac{2n+4+15n}{n\,(n+2)} \right) \\ &= \frac{17n+4}{128n^2\,(n+2)} \end{aligned}$$

Thus

$$c = \delta\sqrt{n-1}\,\frac{\Gamma(n-1)}{\Gamma(n-\frac{1}{2})} \left(1 + \frac{1}{8n}\rho^2 + \frac{17n+4}{128n^2\,(n+2)}\rho^4 + O(\rho^6) \right) + O(\delta^3)$$

VITA

Andrew Dunning is a graduate student at the University of Washington. He returned to school in his late forties to take up Biostatistics, one of the best life decisions he has made.