This is a work of the United States Government. In accordance with 17 U.S.C. 105, no copyright protection is available for such works under U.S. Law.All Rights Reserved. Access to this work was provided by the University of Maryland, Baltimore County (UMBC) ScholarWorks@UMBC digital repository on the Maryland Shared Open Access (MD-SOAR) platform.

Please provide feedback

Please support the ScholarWorks@UMBC repository by emailing scholarworks-group@umbc.edu and telling us what having access to this work means to you and why it's important to you. Thank you.

RESEARCH REPORT SERIES
(Statistics #2020-02)

Comparison of Local Powers of Some Exact Tests for a Common Normal Mean with Unequal Variances

Yehenew G. Kifle[1],
Alain M. Moluh[1],
Bimal K. Sinha[1, 2]

[1]Department of Mathematics and Statistics, University of Maryland Baltimore County;
[2]Center for Statistical Research and Methodology, U.S. Census Bureau

Center for Statistical Research & Methodology
Research and Methodology Directorate
U.S. Census Bureau
Washington, D.C. 20233

Report Issued: May 20, 2020

Disclaimer: This report is released to inform interested parties of research and to encourage discussion. The views expressed are those of the authors and not those of the U.S. Census Bureau.

Comparison of local powers of some exact tests for a common normal mean with unequal variances

Yehenew G. Kifle[1], Alain M. Moluh[1], Bimal K. Sinha[1,2*]

Abstract

The inferential problem of drawing inference about a common mean μ of several independent normal populations with unequal variances has drawn universal attention, and there are many exact tests for testing a null hypothesis $H_0 : \mu = \mu_0$ against two-sided alternatives. In this paper we provide a review of their local powers and a comparison. It turns out that, in the case of equal sample size, a uniform comparison and ordering of the exact tests based on their local power can be carried out even when the variances are unknown.

Keywords: *Common Mean*; *Exact Test*; *Local Power*; *Meta-Analysis*; *P-value*

1 Introduction

The inferential problem of drawing inference about a common mean μ of several independent normal populations with unequal variances has drawn universal attention, and there are many exact tests for testing a null hypothesis $H_0 : \mu = \mu_0$ against two-sided alternatives $H_1 : \mu \neq \mu_0$. In this paper we provide a review of their local powers and a comparison.

A well-known context of this problem occurred when Meier [1953] was approached to draw inference about the mean of albumin in plasma protein in human subjects based on results from four experiments, reproduced below (Table 1).

Another scenario happened when Eberhardt et al. [1989] had results from four experiments about nonfat milk powder and the problem was to draw inference about the mean Selenium in nonfat milk powder by combining the results from four methods (Table 2).

*Corresponding author: `sinha@umbc.edu`
[1]Department of Mathematics and Statistics, University of Maryland Baltimore County, USA.
[2]Center for Statistical Research and Methodology, U.S. Census Bureau, USA.
Disclaimer: This article is released to inform interested parties of ongoing research and to encourage discussion. The views expressed are those of the authors and not those of the U.S. Census Bureau.

Table 1: **Percentage of albumin in plasma protein of four different experiments**

Experiment	n_i	Mean	Variance
A	12	62.3	12.986
B	15	60.3	7.84
C	7	59.5	33.433
D	16	61.5	18.51

Table 2: **Selenium content in nonfat milk powder using four methods**

Methods	n_i	Mean	Variance
Atomic absorption Spectrometry	8	105	85.711
Neutron activation: Instrumental	12	109.75	20.748
Neutron activation: Radiochemical	14	109.5	2.729
Isotope dilution mass spectrometry	8	113.25	33.64

A similar situation arises in the context of environmental data analysis when upon identifying a hot-spot in a contaminated area, samples are drawn and sent to several labs simultaneously and then the resulting data are combined for eventual analysis. This parallel data analysis is especially important for subsequent adoption of remedial actions in case the mean contamination level at the site is found to exceed a certain threshold.

A possible application scenario at the Census Bureau may arise if there is a need to draw inference about average wage of college graduates in a specified age group in a certain state. County level information can be collected from each county via simple random sampling. Then under a model-based approach with a common overall mean wage across the state and heterogeneous county level variances, the results developed in this paper can be useful. Of course, in case of complex surveys involving survey weights, our current formulation of the inference problem will not be applicable.

A general formulation of the problem can be stated as follows. There are k normal populations with a common mean μ and different variances $\sigma_1^2, \cdots, \sigma_k^2$. Based on a sample of size n_i from the i^{th} population, we want to test $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$. Obviously, there exist k independent t-tests based on $t_i = \frac{\sqrt{n_i}(\bar{X}_i - \mu)}{S_i}$ that follows central t distribution with ν_i degrees of freedom. Note that the assumption of normality is crucial in our subsequent discussion. Most meta-analysis applications are based on this assumption [Hartung et al., 2008].

The natural meta-analysis question now is: how to combine the results from the k independent t-tests? As one can expect, there are many ways of accomplishing this task based on some exact and some asymptotic procedures. Let us first briefly review the asymptotic procedures for testing hypothesis about common mean μ.

In the trivial case when the k population variances are completely known, the common mean μ can easily be estimated using the maximum likelihood estimator $\hat{\mu} = [\sum_{i=1}^k \frac{n_i}{\sigma_i^2} \bar{X}_i][\sum_{j=1}^k \frac{n_j}{\sigma_j^2}]^{-1}$ with $Var(\hat{\mu}) = \left[\sum_{i=1}^k \frac{n_i}{\sigma_i^2}\right]^{-1}$. This estimator $\hat{\mu}$ is the minimum variance unbiased estimator under normality as well as the best linear unbiased estimator without normality for estimating μ. A simple test based on standard normal z is obvious in this case.

However, in most cases, the population variances are unknown and a familiar estimate, known as the Graybill-Deal estimate can be used [Graybill and Deal, 1959]. This unbiased estimator $\hat{\mu}_{GD}$ together with its variance are given as

$$\hat{\mu}_{GD} = \frac{\sum_{i=1}^{k} \frac{n_i}{S_i^2} \bar{X}_i}{\sum_{j=1}^{k} \frac{n_j}{S_j^2}} \quad \text{with} \quad Var(\hat{\mu}_{GD}) = E\left[\left(\sum_{i=1}^{k} \frac{n_i \sigma_i^2}{S_i^4}\right) \Big/ \left(\sum_{i=1}^{k} \frac{n_i}{S_i^2}\right)^2\right].$$

Khatri and Shah [1974] proposed exact variance expression for $\hat{\mu}_{GD}$, which is complicated and cannot be easily implemented. To address this inferential problem, Meier [1953] derived a first order approximation of the variance of $\hat{\mu}_{GD}$ as

$$Var(\hat{\mu}_{GD}) = \left[\sum_{i=1}^{k} \frac{n_i}{\sigma_i^2}\right]^{-1} \left[1+2\sum_{i=1}^{k} \frac{1}{n_i-1} c_i(1-c_i) + O\left(\sum_{i=1}^{k} \frac{1}{(n_i-1)^2}\right)\right]; \;\; c_i = \frac{n_i/\sigma_i^2}{\sum_{j=1}^{k} n_j/\sigma_j^2}.$$

Sinha [1985] in the same spirit derived an unbiased estimator of the variance of $\hat{\mu}_{GD}$ that is a convergent series. A first order approximation of this estimator is

$$\widehat{Var}_{(1)}(\hat{\mu}_{GD}) = \frac{1}{\sum_{i=1}^{k} \frac{n_i}{S_i^2}} \left[1 + \sum_{i=1}^{k} \frac{4}{n_i+1}\left(\frac{n_i/S_i^2}{\sum_{j=1}^{k} n_j/S_j^2} - \frac{n_i^2/S_i^4}{(\sum_{j=1}^{k}(n_j/S_j^2))^2}\right)\right].$$

The above estimator is comparable to Meier's (1953) approximate estimator

$$\widehat{Var}_{(2)}(\hat{\mu}_{GD}) = \frac{1}{\sum_{i=1}^{k} \frac{n_i}{S_i^2}} \left[1 + \sum_{i=1}^{k} \frac{4}{n_i-1}\left(\frac{n_i/S_i^2}{\sum_{j=1}^{k} n_j/S_j^2} - \frac{n_i^2/S_i^4}{(\sum_{j=1}^{k}(n_j/S_j^2))^2}\right)\right].$$

The "classical" meta-analysis variance estimator, $\widehat{Var}_{(3)}(\hat{\mu}_{GD})$, and approximate variance estimator proposed by Hartung [1999] $\widehat{Var}_{(4)}(\hat{\mu}_{GD})$ are the two other variance estimators of $\hat{\mu}_{GD}$ which are given by

$$\widehat{Var}_{(3)}(\hat{\mu}_{GD}) = \frac{1}{\sum_{i=1}^{k} \frac{n_i}{S_i^2}} \quad \& \quad \widehat{Var}_{(4)}(\hat{\mu}_{GD}) = \frac{1}{k-1} \sum_{i=1}^{k} \left(\frac{n_i/S_i^2}{\sum_{j=1}^{k} n_j/S_j^2}\right)(\bar{X}_i - \hat{\mu}_{GD})^2.$$

We should mention that a parametric bootstrap approach based on Graybill-Deal estimate was suggested in Malekzadeh and Kharrati-Kopaei [2018] to draw inference about μ which works quite well in large samples. Likewise, inference based on the MLE of μ suggested in Chang and Pal [2008] is also asymptotic in nature. As mentioned earlier, the central focus of this paper is to critically examine some exact tests for the common mean. A power comparison of these available exact tests is then a natural desire. In this paper this is precisely what we accomplish by comparing six exact tests based on their local powers.

The organization of the paper is as follows. In section 2 we provide a brief description of the six exact tests with their references. The pdf of non-central t which naturally plays a pivotal role for studying power of t tests is given along with its local expansion (in terms of its non-centrality parameter). Section 3, a core section of the paper, provides expressions of local powers of all the proposed tests. Appendix I at the end contains proofs of all technical results. Section 4 contains some numerical (power) comparisons in the case of equal sample sizes. We conclude this paper with some remarks in Section 5.

2 Review of six exact tests for H_0 versus H_1

Consider k independent normal populations where the i^{th} population follows a normal distribution with mean $\mu \in \mathbb{R}$ and variance $\sigma_i^2 > 0$. Let $\bar{X}_i$ denote the sample mean, S_i^2 the (unbiased) sample variance, and n_i the sample size of the i^{th} population. Then, we have $\bar{X}_i \sim \mathcal{N}(\mu, \frac{\sigma_i^2}{n_i})$ and $\frac{(n_i-1)S_i^2}{\sigma_i^2} \sim \chi^2_{\nu_i}$, where $\nu_i = (n_i - 1)$ and $i = 1, \cdots, k$. Note that the statistics $\{\bar{X}_i, S_i^2, i = 1, \cdots, k\}$ are all mutually independent.

A generic notation for a t statistic based on a sample of size n is $t_{obs} = \sqrt{n}(\bar{x} - \mu_0)/s$. We can refer to this t computed from a given data set as the observed value of our test statistic, and reject H_0 when $|t_{obs}| > t_{\nu;\alpha/2}$, where ν is the degrees of freedom and α is Type I error level. A test for H_0 based on a P-value on the other hand is based on $P_{obs} = P[|t_\nu| > |t_{obs}|]$, and we reject H_0 at level α if $P_{obs} < \alpha$. Here t_ν stands for the central t variable with ν degrees of freedom and $t_{\nu;\alpha/2}$ stands for the upper $\alpha/2$ critical value of t_ν. It is easy to check that the two approaches are obviously equivalent.

A random P-value which has a $Uniform(0,1)$ distribution under H_0 is defined as $P_{ran} = P[|t_\nu| > |t_{ran}|]$, where $t_{ran} = \sqrt{n}(\bar{X} - \mu_0)/S$. All suggested tests for H_0 are based on P_{obs} and t_{obs} values and their properties, including size and power, are studied under P_{ran} and t_{ran}. To simplify notations, we will denote P_{obs} by small p and P_{ran} by large P. Six exact tests based on t_{obs} and p values from k independent studies as available in the literature are listed below.

2.1 *P*-value based exact tests

2.1.1 Tippett's test

This minimum P-value test was proposed by Tippett et al. [1931], who noted that, if $P_1, \cdots, P_k$ are independent p-values from continuous test statistics, then each has a $uniform$ distribution under H_0. Suppose that $P_{(1)}, \cdots, P_{(k)}$ are ordered p-values. According to this method, the common mean null hypothesis $H_0 : \mu = \mu_0$ is rejected at α level of significance if $P_{(1)} < \left[1 - (1-\alpha)^{\frac{1}{k}}\right]$. Incidentally, this test is equivalent to the test based on $M_t = max_{1 \le i \le k}|t_i|$ suggested by Cohen and Sackrowitz [1984].

2.1.2 Wilkinson's test

This test statistic proposed by Wilkinson [1951] is a generalization of Tippett's test that uses not just the smallest but the r^{th} smallest p-value ($P_{(r)}$) as a test statistic. The common mean null hypothesis $H_0 : \mu = \mu_0$ will be rejected if $P_{(r)} < d_{r,\alpha}$, where $P_{(r)}$ follows a $Beta$ distribution with parameters r and $(k - r + 1)$ under H_0 and $d_{r,\alpha}$ satisfies $Pr\{P_{(r)} < d_{r,\alpha} \,|\, H_0\} = \alpha$.

2.1.3 Inverse normal test

This exact test procedure which involves transforming each p-value to the corresponding normal score was proposed independently by Stouffer et al. [1949] and Lipták [1958]. Using this inverse normal method, hypothesis about the common μ will be rejected at α level of significance if $\left[\sum_{i=1}^{k}\Phi^{-1}(P_i)\right]\left[\sqrt{k}\right]^{-1} < -z_\alpha$, where Φ^{-1} denotes the inverse of the cdf of a standard normal distribution and z_α stands for the upper α level cutoff point of a standard normal distribution.

2.1.4 Fisher's inverse χ^2-test

This inverse χ^2-test is one of the most widely used exact test procedures for combining k independent p-values [Fisher, 1932]. This procedure uses the $\prod_{i=1}^{k} P_i$ to combine the k independent p-values. Then, using the connection between $uniform$ and χ^2 distributions, the hypothesis about the common μ will be rejected if $-2\sum_{i=1}^{k}\ln(P_i) > \chi^2_{2k,\alpha}$, where $\chi^2_{2k,\alpha}$ denotes the upper α critical value of a χ^2-distribution with $2k$ degrees of freedom.

2.2 Exact test based on a modified t

Fairweather [1972] consider a test based on a weighted linear combination of the t_i's. In this paper, we consider a variation of this test based on a weighted linear combination of $|t_i|$ as we are testing a non-directional alternative. Our test statistic T_1 is given as $\sum_{i=1}^{k} w_{1i}|t_i|$, where $w_{1i} \propto [Var(|t_i|)]^{-1}$ with $Var(|t_i|) = \left[[\nu_i(\nu_i-2)^{-1}] - \left([\Gamma(\frac{\nu_i-1}{2})\sqrt{\nu_i}][\Gamma(\frac{\nu_i}{2})\sqrt{\pi}]^{-1}\right)^2\right]$. The null hypothesis $H_0 : \mu = \mu_0$ will be rejected if $T_1 > d_{1\alpha}$, where $Pr\{T_1 > d_{1\alpha} \mid H_0\} = \alpha$. In applications $d_{1\alpha}$ is computed by simulation.

2.3 Exact test based on a modified F

Jordan and Krishnamoorthy [1996] considered a weighted linear combination of the F-test statistics F_i, namely T_2, which is given as $\sum_{i=1}^{k} w_{2i}F_i$, $F_i = t_i^2 \sim F(1,\nu_i)$, and $w_{2i} \propto [Var(F_i)]^{-1}$ with $Var(F_i) = [2\nu_i^2(\nu_i-1)][(\nu_i-2)^2(\nu_i-4)]^{-1}$ for $\nu_i > 4$. The null hypothesis $H_0 : \mu = \mu_0$ will be rejected if $T_2 > d_{2\alpha}$, where $Pr\{T_2 > d_{2\alpha} \mid H_0\} = \alpha$. In applications $d_{2\alpha}$ is computed by simulation.

We mention in passing that Philip et al. [1999] studied some properties of the confidence interval for the common mean μ based on Fisher's test and inverse normal test.

The pdfs of t statistic under the null and alternative hypothesis which will be required in the sequel are given below. $\delta = \sqrt{n}(\mu_1 - \mu_0)/\sigma$ below stands for the non-centrality parameter when μ_1 is chosen as an alternative value. Later, we will denote $(\mu_1 - \mu_0)$ by Δ.

$$
\begin{aligned}
f_\nu(t) &= \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\,\Gamma(\frac{\nu}{2})}\left(1+\frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}} \\
f_{\nu;\delta}(t) &= \frac{\nu^{\frac{\nu}{2}}\exp\left(\frac{-\nu\delta^2}{2(t^2+\nu)}\right)}{\sqrt{\pi}\,\Gamma(\frac{\nu}{2})2^{\frac{\nu-1}{2}}\left(t^2+\nu\right)^{\frac{\nu+1}{2}}}\int_0^\infty y^\nu \exp\left[-\frac{1}{2}\big(y-\frac{\delta t}{\sqrt{t^2+\nu}}\big)^2\right]dy
\end{aligned}
$$

First and second derivatives of $f_{\nu;\delta}(t)$ evaluated at $\delta = 0$ (equivalently, $\Delta = 0$) which will play a pivotal role in the study of local powers of the proposed tests appear below.

$$\frac{\partial f_{\nu;\delta}(t)}{\partial \delta}\Big|_{\delta=0} = \frac{t}{\sqrt{2\pi}\left(\frac{t^2}{\nu}+1\right)^{\frac{\nu+2}{2}}}$$

$$\frac{\partial^2 f_{\nu;\delta}(t)}{\partial \delta^2}\Big|_{\delta=0} = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sqrt{\nu\pi}}\left[\frac{t^2-1}{(\frac{t^2}{\nu}+1)^{\frac{\nu+3}{2}}}\right]$$

3 Expressions of local powers of the six proposed tests

In this section we provide the expressions of local powers of the suggested exact tests. A common premise is that we derive an expression of the power of a test under $\Delta \neq 0$, and carry out its Taylor expansion around $\Delta = 0$. It turns out that due to two-sided nature of our tests, the first term vanishes, and we retain terms of order $O(\Delta^2)$.

The final expressions of the local powers of the proposed tests are given below in the general case and also in the special case when $n_1 = \cdots = n_k = n$, and $\nu_1 = \cdots = \nu_k = \nu = n-1$. All throughout, we write $\Psi = \sum_{i=1}^{k} \frac{1}{\sigma_i^2}$ which is relevant in the special case. For detailed proofs of all technical results below we refer to the Appendix section of this paper.

3.1 Local power of Tippett test $[LP(T)]$

$$\begin{aligned} LP(T) &\approx \alpha + (1-\alpha)^{\frac{k-1}{k}}\frac{\Delta^2}{2}\left(\sum_{i=1}^{k}\frac{n_i}{\sigma_i^2}|\xi_{\nu_i T}(a_\alpha)|\right) \qquad (1)\\ &= \alpha + \left[\frac{n\Delta^2}{2}\Psi\right]\left[(1-\alpha)^{\frac{k-1}{k}}\right]|\xi_{\nu T}(a_\alpha)| \qquad \text{[special case]} \end{aligned}$$

where $\xi_{\nu T}(a_\alpha) = \int_{-t_\nu(\frac{a_\alpha}{2})}^{t_\nu(\frac{a_\alpha}{2})} \frac{\partial^2 f_{\nu;\delta}(t)}{\partial\delta^2}\Big|_{\delta=0} dt$; $a_\alpha = [1-(1-\alpha)^{\frac{1}{k}}]$. It turns out that $\xi_{\nu T}(a_\alpha) < 0$.

3.2 Local power of Wilkinson test $[LP(W_r)]$

$$\begin{aligned} LP(W_r) &\approx \alpha + \binom{k-1}{r-1} d_{r;\alpha}^{r-1}(1-d_{r;\alpha})^{k-r}\frac{\Delta^2}{2}\left[\sum_{i=1}^{k}\frac{n_i}{\sigma_i^2}|\xi_{iW}(d_{r,\alpha})|\right] \qquad (2)\\ &= \alpha + \left[\frac{n\Delta^2}{2}\Psi\right]\binom{k-1}{r-1}|\xi_{\nu W}(d_{r;\alpha})| d_{r;\alpha}^{r-1}(1-d_{r;\alpha})^{k-r} \qquad \text{[special case]} \end{aligned}$$

where $\xi_{\nu W}(d_{r;\alpha})$ is equivalent to $\xi_{\nu T}(a_\alpha)$ with $a_\alpha = d_{r;\alpha}$. It turns out that $\xi_{\nu W}(d_{r;\alpha}) < 0$.

Remark: For the special case $r = 1$, $LP(W_r) = LP(T)$, as expected,
because $d_{1;\alpha} = [1-(1-\alpha)^{\frac{1}{k}}]$, implying $(1-d_{1;\alpha})^{k-1} = (1-\alpha)^{\frac{k-1}{k}}$.

3.3 Local power of Inverse Normal test $[LP(INN)]$

$$LP(INN) \approx \alpha + \frac{\Delta^2}{\sqrt{k}}\phi(z_\alpha)\sum_{i=1}^{k}\frac{n_i\nu_i}{\sigma_i^2}\left[\frac{z_\alpha[B_{\nu_i} - C_{\nu_i}]}{2\sqrt{k}} - A_{\nu_i}\right] \tag{3}$$

$$= \alpha + \left[\frac{n\Delta^2}{2}\Psi\right]\frac{2\nu}{\sqrt{k}}\phi(z_\alpha)\left[\frac{z_\alpha[B_\nu - C_\nu]}{2\sqrt{k}} - A_\nu\right] \qquad \text{[special case]}$$

where $A_\nu = \int_{-\infty}^{\infty} u\phi(u)Q_\nu(u)du$; $B_\nu = \int_{-\infty}^{\infty} u^2\phi(u)Q_\nu(u)du$; $C_\nu = \int_{-\infty}^{\infty} \phi(u)Q_\nu(u)du$; $Q_\nu(u) = \left[\frac{x^2-1}{x^2+\nu}\right]_{x=t_\nu(\frac{c}{2}),\ c=\Phi(u)}$; $\phi(u)$ is standard normal pdf and $\Phi(u)$ is standard normal cdf.

3.4 Local power of Fisher test $[LP(F)]$

$$LP(F) \approx \alpha + \frac{\Delta^2}{2}\left[\sum_{i=1}^{k}\frac{n_i\nu_i}{\sigma_i^2}D_{\nu_i}\right]\left[E\Big\{\{\ln(T/2)\}I_{\{T\geq\chi^2_{2k;\alpha}\}}\Big\}_{T\sim\chi^2_{2k}} - \alpha D_0\right] \tag{4}$$

$$= \alpha + \left[\frac{n\Delta^2}{2}\Psi\right]\nu D_\nu\left[E\Big\{\{\ln(T/2)\}I_{\{T\geq\chi^2_{2k;\alpha}\}}\Big\}_{T\sim\chi^2_{2k}} - \alpha D_0\right] \qquad \text{[special case]}$$

where $D_0 = E\big[\log(q)\big]$; $D_\nu = E\big[U\psi_\nu(U)\big]$; $U \sim$exp[2]; $q \sim$gamma$[1,k]$; $T \sim$gamma$[2,k]$, $\psi_\nu(u) = \left[\frac{x^2-1}{x^2+\nu}\right]_{x=t_\nu(\frac{c}{2}),\ c=\exp(-\frac{u}{2})}$.

3.5 Local power of a modified t test $[LP(T_1)]$

$$LP(T_1) \approx \alpha + \frac{\Delta^2}{2}\left(\sum_{j=1}^{k}\frac{n_j\nu_j}{\sigma_j^2}\right)E_{H_0}\left[\left\{\frac{(t_j^2-1)\nu_j}{t_j^2+\nu_j}\right\}I_{\{\sum_{i=1}^k w_{1i}|t_i|>d_1\alpha\}}\right] \tag{5}$$

$$= \alpha + \left[\frac{n\Delta^2}{2}\Psi\right]\nu E_{H_0}\left[\left\{\frac{(t_1^2-1)\nu}{t_1^2+\nu}\right\}I_{\{\sum_{i=1}^k|t_i|>d_{1\alpha}\}}\right] \qquad \text{[special case]}$$

3.6 Local power of a modified F test $[LP(T_2)]$

$$LP(T_2) \approx \alpha + \frac{\Delta^2}{2}\left(\sum_{j=1}^{k}\frac{n_j}{\sigma_j^2}\right)E_{H_0}\left[\left\{\frac{[F_j-1]\nu_j}{F_j+\nu_j}\right\}I_{\{\sum_{i=1}^k w_{2i}F_i>d_{2\alpha}\}}\right] \tag{6}$$

$$= \alpha + \left[\frac{n\Delta^2}{2}\Psi\right]E_{H_0}\left[\left\{\frac{[F_1-1]\nu}{F_1+\nu}\right\}I_{\{\sum_{i=1}^k F_i>d_{2\alpha}\}}\right] \qquad \text{[special case]}$$

4 Comparison of local powers

It is interesting to observe from the above expressions that in the special case of equal sample size, local powers can be readily compared, irrespective of the values of the unknown variances (involved through Ψ, which is a common factor in all the expressions of local power).

Table 3 represents values of the 2^{nd} term of local power given above in equations 1 to 6, apart from the common term $\left[\frac{n\Delta^2}{2}\Psi\right]$ for different values of k, n and choices of r ($\leq k$) with maximum local power. Similarly, the 2^{nd} term of local power of $Wilkinson$ test for different values of n, k and r ($\leq k$) are presented in Table 4. All throughout we have used $\alpha = 5\%$.

Here are some interesting observations: comparing *Tippett* and $Wilkinson$ tests, we note that $Wilkinson$ test for some $r > 1$ always outperforms *Tippett* test, and the optimal choice of r seems to increase with k (Table 4 and Figure 2). Among the other tests, Figure 1 with $\Psi = 1$ reveals that $Fisher's$ test fares the best uniformly in the design parameters n and k. Both $modified$ t and $modified$ F tests perform reasonably well for all values of k and n (Table 3).

Some limited local power computations in case of unequal sample sizes are reported in Table 5. It again follows that $Fisher's$ test has an edge over all other tests. Our recommendation based on the local power comparison of the available exact tests is to advocate the use of $Fisher's$ test in all scenarios.

Table 3: Comparison of the 2^{nd} term of local powers [without $n\Delta^2\Psi/2$] of six exact tests for different values of k and n (equal sample size)

Exact Test	**k=5**			**k=10**			**k=15**		
	n=15	**n=25**	**n=40**	**n=15**	**n=25**	**n=40**	**n=15**	**n=25**	**n=40**
Tippett	0.0575	0.0633	0.0667	0.0322	0.0361	0.0383	0.0227	0.0257	0.0275
Wilkinson	0.0633	0.0664	0.0681	0.0412	0.0430	0.0441	0.0324	0.0338	0.0346
Inv Normal	0.0730	0.0731	0.0730	0.0576	0.0584	0.0588	0.0491	0.0501	0.0504
Fisher	0.1050	0.1179	0.1191	0.0844	0.0861	0.0877	0.0641	0.0705	0.0716
Modified t	0.0724	0.0738	0.0754	0.0471	0.0508	0.0529	0.0378	0.0417	0.0426
Modified F	0.0722	0.0768	0.0798	0.0479	0.0523	0.0556	0.0389	0.0410	0.0427

5 Conclusion

Based on our computations of local powers of the available exact tests, we have noted that a uniform comparison of them, irrespective of the values of the unknown variances, can be readily made in case of equal sample size, and it turns out that $Fisher's$ exact test performs the best. Some limited computations of local powers in case of unequal sample sizes also reveal the superiority of $Fisher's$ test compared to the other exact tests.

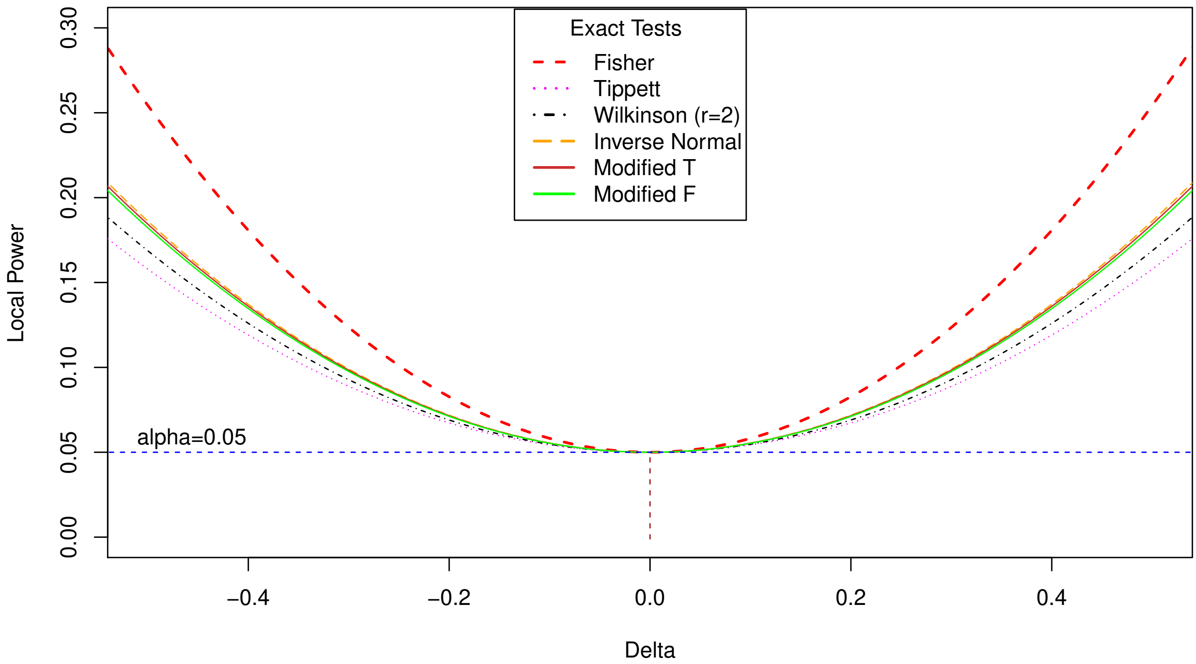

Figure 1: Comparison of local powers of six exact tests for n=15, k=5 and $\Psi = 1$.

Table 4: Comparison of the 2^{nd} term of local powers [without $n\Delta^2\Psi/2$] of $Wilkinson's$ test for different values of k, n (equal sample size) and r ($\leq k$)

r	k=5			k=10			k=15		
	n=15	**n=25**	**n=40**	**n=15**	**n=25**	**n=40**	**n=15**	**n=25**	**n=40**
1	0.0575	0.0633	0.0667	0.0322	0.0361	0.0383	0.0227	0.0257	0.0275
2	0.0633	0.0664	0.0681	0.0395	0.0422	0.0437	0.0292	0.0315	0.0328
3	0.0587	0.0603	0.0611	0.0412	0.0430	0.0441	0.0317	0.0335	0.0345
4	0.0494	0.0501	0.0504	0.0404	0.0417	0.0425	0.0324	0.0338	0.0346
5	0.0359	0.0361	0.0362	0.0384	0.0393	0.0398	0.0322	0.0333	0.0339
6				0.0355	0.0361	0.0364	0.0314	0.0323	0.0327
7				0.0320	0.0324	0.0326	0.0302	0.0309	0.0312
8				0.0279	0.0281	0.0282	0.0287	0.0292	0.0295
9				0.0230	0.0231	0.0231	0.0270	0.0273	0.0275
10				0.0168	0.0168	0.0168	0.0250	0.0253	0.0254
11							0.0229	0.0230	0.0231
12							0.0205	0.0206	0.0206
13							0.0178	0.0179	0.0179
14							0.0147	0.0148	0.0148
15							0.0108	0.0109	0.0109

Table 5: Coefficients of $\Delta^2/2$ in the 2^{nd} term of local powers of six exact tests for different values of k, n (unequal sample sizes) and σ^2

Exact Test	**k=2**	**k=3**	**k=4**
	$n_1 = 10,\ \sigma_1^2 = 1$ $n_2 = 20,\ \sigma_2^2 = 2$	$n_1 = 10,\ \sigma_1^2 = 1$ $n_2 = 20,\ \sigma_2^2 = 2$ $n_3 = 30,\ \sigma_3^2 = 3$	$n_1 = 10,\ \sigma_1^2 = 1$ $n_2 = 20,\ \sigma_2^2 = 2$ $n_3 = 30,\ \sigma_3^2 = 3$ $n_4 = 40,\ \sigma_4^2 = 4$
Tippett	2.3332	2.6606	2.9108
Inv Normal	1.7719	2.4772	3.0894
Fisher	3.4033	4.2150	5.0382
Modified t	2.5295	2.9658	3.4026
Modified F	2.4663	3.0839	3.4484
Wilkinson $(r = 1)$	2.3332	2.6606	2.9108
Wilkinson $(r = 2)$	1.9464	2.5472	2.9611
Wilkinson $(r = 3)$		1.8972	2.5389
Wilkinson $(r = 4)$			1.8447

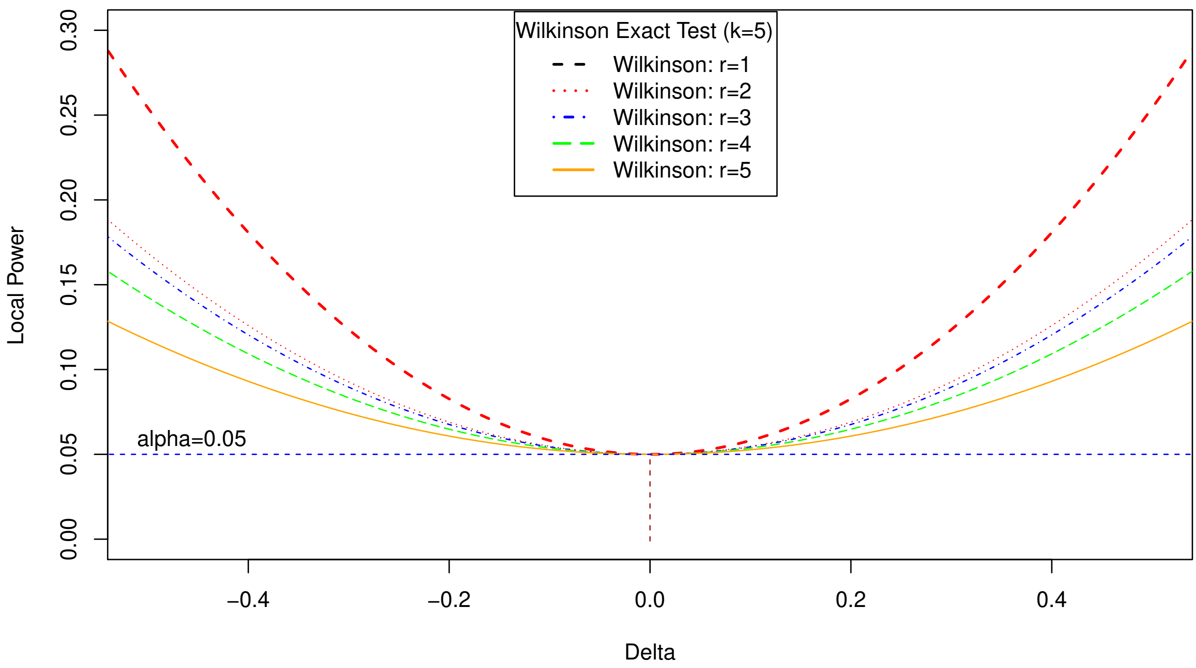

Figure 2: Comparison of local powers of $Wilkinson$ exact test for n=15, k=5 and $\Psi = 1$.

Acknowledgments

Bimal K. Sinha is thankful to Dr. Tommy Wright at the US Census Bureau for helpful comments and encouragement. We are also thankful to Professor Thomas Mathew and Professor Gaurisankar Datta for some helpful comments.

Appendix I: Proofs of local powers of six exact tests

We begin by stating a result related to the distribution of a P-value under the alternative hypothesis $H_0 : \mu = \mu_1$, which will be crucial for providing the main results on local power of all tests based on the P-values. We denote $F_\nu(\cdot)$ to represent the cdf of a central t-distribution with ν degrees of freedom.

Lemma 1

$$Pr\{P > c|H_1\} \approx (1-c) + \frac{\Delta^2}{2}\left[\frac{n}{\sigma^2}\xi_\nu(c)\right]. \tag{7}$$

Proof:

$$\begin{aligned}
Pr\{P > c|H_1\} &= Pr\left\{Pr\left[|t_\nu| > \left|\frac{\sqrt{n}(\bar{X}-\mu_0)}{S}\right|\right] > c|H_1\right\}\\
&= Pr\left\{1 - \left[F_\nu\left(\left|\frac{\sqrt{n}(\bar{X}-\mu_0)}{S}\right|\right) - F_\nu\left(-\left|\frac{\sqrt{n}(\bar{X}-\mu_0}{S}\right|\right)\right] > c|H_1\right\}\\
&= Pr\left\{\left[F_\nu\left(\left|\frac{\sqrt{n}(\bar{X}-\mu_0)}{S}\right|\right) - F_\nu\left(-\left|\frac{\sqrt{n}(\bar{X}-\mu_0}{S}\right|\right)\right] < 1-c|H_1\right\}\\
&= Pr\left\{\left|\frac{\sqrt{n}(\bar{X}-\mu_0)}{S}\right| < t_\nu\big(\frac{c}{2}\big)|H_1\right\}\\
&= Pr\left\{-t_\nu\big(\frac{c}{2}\big) < \frac{\sqrt{n}(\bar{X}-\mu_0)}{S} < t_\nu\big(\frac{c}{2}\big)|H_1\right\}\\
&= Pr\left\{-t_\nu\big(\frac{c}{2}\big) < t_\nu(\delta) < t_\nu\big(\frac{c}{2}\big)|H_1\right\}\\
&= \int_{-t_\nu(\frac{c}{2})}^{t_\nu(\frac{c}{2})} f(x|\nu,\delta)\,dx \qquad \left[f(x|\nu,\delta) \sim \text{non-central } t_\nu\left(\delta = \frac{\sqrt{n}}{\sigma}\Delta\right)\right]\\
&\approx \int_{-t_\nu(\frac{c}{2})}^{t_\nu(\frac{c}{2})} f(x|\nu,0)dx + \delta_i\left(\frac{\partial f}{\partial\delta}\right)\Big|_{\delta=0}dx + \frac{\delta^2}{2}\left(\frac{\partial^2 f}{\partial\delta^2}\right)\Big|_{\delta=0}dx\\
&\approx (1-c) + \frac{n}{2\sigma^2}\Delta^2\int_{-t_\nu(\frac{c}{2})}^{t_\nu(\frac{c}{2})}\left\{\frac{\partial^2 f(x|\nu,\delta)}{\partial\delta^2}\Big|_{\delta=0}\right\}dx\\
&\approx (1-c) + \frac{\Delta^2}{2}\left[\frac{n}{\sigma^2}\xi_\nu(c)\right]
\end{aligned}$$

where $\xi_\nu(c) = \int_{-t_\nu(\frac{c}{2})}^{t_\nu(\frac{c}{2})}\left\{\frac{\partial^2 f(x|\nu,\delta)}{\partial\delta^2}\Big|_{\delta=0}\right\}dx = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sqrt{\nu\pi}}\int_{-t_\nu(\frac{c}{2})}^{t_\nu(\frac{c}{2})}\left(\frac{x^2-1}{[\frac{x^2}{\nu}+1]^{\frac{\nu+3}{2}}}\right)dx$. It turns out that $\xi_\nu(c) < 0$.

I. Local power of Tippett test $[LP(T)]$

Recall that *Tippett* test rejects the null hypothesis if $P_{(1)} < \left[1-(1-\alpha)^{\frac{1}{k}}\right] = a_\alpha$. This leads to

$$\text{Power} = 1 - \prod_{i=1}^{k} Pr\{P_i > a_\alpha | H_1\}.$$

Applying Lemma 1, the local power of *Tippett* test is calculated as follows:

$$\begin{aligned}
\text{Local power} &\approx 1 - \prod_{i=1}^{k}\left[(1-a_\alpha) + \frac{\Delta^2}{2}\left(\frac{n_i}{\sigma_i^2}\xi_{\nu_i T}(a_\alpha)\right)\right] \\
&\approx 1 - \prod_{i=1}^{k}\left[(1-\alpha)^{\frac{1}{k}} + \frac{\Delta^2}{2}\left(\frac{n_i}{\sigma_i^2}\xi_{\nu_i T}(a_\alpha)\right)\right] \\
&\approx 1 - \left[(1-\alpha) + (1-\alpha)^{\frac{k-1}{k}}\frac{\Delta^2}{2}\left(\sum_{i=1}^{k}\frac{n_i}{\sigma_i^2}\xi_{\nu_i T}(a_\alpha)\right)\right] \\
&\approx \alpha + (1-\alpha)^{\frac{k-1}{k}}\frac{\Delta^2}{2}\left(\sum_{i=1}^{k}\frac{n_i}{\sigma_i^2}|\xi_{\nu_i T}(a_\alpha)|\right).
\end{aligned}$$

For the special case $n_1 = \cdots = n_k = n$; $\nu_1 = \cdots = \nu_k = \nu = n-1$ and $\xi_{\nu_1 T}(a_\alpha) = \cdots = \xi_{\nu_k T}(a_\alpha) = \xi_{\nu T}(a_\alpha)$, the local power of *Tippett* test reduces to:

$$\begin{aligned}
\text{LP(T)} &\approx \alpha + (1-\alpha)^{\frac{k-1}{k}}\frac{n\Delta^2}{2}|\xi_{\nu T}(a_\alpha)|\left(\sum_{i=1}^{k}\frac{1}{\sigma_i^2}\right) \\
&= \alpha + \left[\frac{n\Delta^2}{2}\Psi\right]\left[(1-\alpha)^{\frac{k-1}{k}}\right]|\xi_{\nu T}(a_\alpha)| \quad \text{where } \Psi = \sum_{i=1}^{k}\frac{1}{\sigma_i^2}.
\end{aligned}$$

II. Local power of Wilkinson test $[LP(W_r)]$

Using r^{th} smallest p-value $P_{(r)}$ as a test statistic, the null hypothesis will be rejected if $P_{(r)} < d_{r,\alpha}$, where $P_{(r)} \sim Beta[r, k-r+1]$ under H_0 and $d_{r,\alpha}$ satisfies $\alpha = Pr\{P_{(r)} < d_{r,\alpha} \mid H_0\} = \int_0^{d_{r,\alpha}} \frac{u^{r-1}(1-u)^{k-r}}{B[r,k-r+1]}du$. This leads to

$$\begin{aligned}
\text{Power} &= Pr[P_{(r)} < d_{r,\alpha}|H_1] \\
&= \sum_{l=r}^{k} Pr\{P_{i_1}, \ldots, P_{i_l} < d_{r,\alpha} < P_{i_{l+1}}, \ldots, P_{i_k}|H_1\}
\end{aligned}$$

where $(i_1, \cdots, i_l, i_{l+1}, \cdots, i_k)$ is a permutation of $(1, \cdots, k)$. Applying Lemma 1, we get

$$
\begin{aligned}
Pr\{P_{i_1}, \quad \ldots \quad &, P_{i_l} < d_{r,\alpha} < P_{i_{l+1}}, \ldots, P_{i_k} | H_1\} \\
&\approx \left\{\prod_{j=1}^{l} \left(d_{r,\alpha} - \frac{n_{i_j}\Delta^2}{2\sigma_{i_j}^2}\xi_{i_jW}(d_{r,\alpha})\right)\right\}\left\{\prod_{j=l+1}^{k} \left(1 - d_{r,\alpha} + \frac{n_{i_j}\Delta^2}{2\sigma_{i_j}^2}\xi_{i_jW}(d_{r,\alpha})\right)\right\} \\
&\approx \left\{d_{r,\alpha}^{l} - d_{r,\alpha}^{l-1}\frac{\Delta^2}{2}\left(\sum_{j=1}^{l}\frac{n_{i_j}}{\sigma_{i_j}^2}\xi_{i_jW}(d_{r,\alpha})\right)\right\} \times \\
&\quad \left\{(1-d_{r,\alpha})^{k-l} + (1-d_{r,\alpha})^{k-l-1}\frac{\Delta^2}{2}\left(\sum_{j=l+1}^{k}\frac{n_{i_j}}{\sigma_{i_j}^2}\xi_{i_jW}(d_{r,\alpha})\right)\right\} \\
&\approx d_{r,\alpha}^{l}(1-d_{r,\alpha})^{k-l} + \frac{\Delta^2}{2}\left\{d_{r,\alpha}^{l}(1-d_{r,\alpha})^{k-l-1}\left(\sum_{j=l+1}^{k}\frac{n_{i_j}}{\sigma_{i_j}^2}\xi_{i_jW}(d_{r,\alpha})\right)\right. \\
&\quad \left. -d_{r,\alpha}^{l-1}(1-d_{r,\alpha})^{k-l}\left(\sum_{j=1}^{l}\frac{n_{i_j}}{\sigma_{i_j}^2}\xi_{i_jW}(a_{d_{r,\alpha}})\right)\right\}.
\end{aligned}
$$

Permuting $(i_1, \ldots, i_k)$ over $(1, \ldots, k)$, we get for any fixed l $(r \le l \le k)$,

$$
\begin{aligned}
\text{1st term} &= \binom{k}{l} d_{r,\alpha}^{l}(1-d_{r,\alpha})^{k-l} \\
\text{2nd term} &= \frac{\Delta^2}{2} d_{r,\alpha}^{l}(1-d_{r,\alpha})^{k-l-1}\left\{\binom{k-1}{k-l-1}\left(\sum_{i=1}^{k}\frac{n_i}{\sigma_i^2}\xi_{iW}(d_{r,\alpha})\right)\right\} \\
\text{3rd term} &= -\frac{\Delta^2}{2} d_{r,\alpha}^{l-1}(1-d_{r,\alpha})^{k-l}\left\{\binom{k-1}{l-1}\left(\sum_{i=1}^{k}\frac{n_i}{\sigma_i^2}\xi_{iW}(d_{r,\alpha})\right)\right\}.
\end{aligned}
$$

The 2nd term above follows upon noting that when $\left[\sum_{j=l+1}^{k}\frac{n_{i_j}}{\sigma_{i_j}^2}\xi_{i_jW}(d_{r,\alpha})\right]$ is permuted over $(i_{l+1} < \cdots < i_k) \subset (1, \ldots, k)$, each term $\frac{n_i}{\sigma_i^2}\xi_{iW}(d_{r,\alpha})$ appears exactly $\binom{k-1}{k-l-1}$ times, for each $i = 1, \cdots, k$. The 3rd term, likewise, follows upon noting that when $\left[\sum_{j=1}^{l}\frac{n_{i_j}}{\sigma_{i_j}^2}\xi_{i_jW}(d_{r,\alpha})\right]$ is permuted over $(i_1 < \cdots < i_l) \subset (1, \ldots, k)$, each term $\frac{n_i}{\sigma_i^2}\xi_{iW}(d_{r,\alpha})$ appears exactly $\binom{k-1}{l-1}$ times, for each $i = 1, \cdots, k$.

Adding the above three terms and simplifying, we get

$$
LP(W_r) \approx \alpha + \binom{k-1}{r-1} d_{r;\alpha}^{r-1}(1-d_{r;\alpha})^{k-r}\frac{\Delta^2}{2}\left[\sum_{i=1}^{k}\frac{n_i}{\sigma_i^2}|\xi_{iW}(d_{r,\alpha})|\right].
$$

For the special case $n_1 = \cdots = n_k = n$; $\nu_1 = \cdots = \nu_k = \nu = n-1$ and $\xi_{\nu_1W}(d_{r;\alpha}) = \cdots = \xi_{\nu_kW}(d_{r;\alpha}) = \xi_{\nu W}(d_{r;\alpha})$, the local power of $Wilkinson$ test reduces to:

$$
\begin{aligned}
LP(W_r) &\approx \alpha + \binom{k-1}{r-1} d_{r;\alpha}^{r-1}(1-d_{r;\alpha})^{k-r}\frac{n\Delta^2}{2}|\xi_{iW}(d_{r,\alpha})|\left(\sum_{i=1}^{k}\frac{1}{\sigma_i^2}\right) \\
&= \alpha + \left[\frac{n\Delta^2}{2}\Psi\right]\binom{k-1}{r-1}|\xi_{\nu W}(d_{r;\alpha})| d_{r;\alpha}^{r-1}(1-d_{r;\alpha})^{k-r} \text{ where } \Psi = \sum_{i=1}^{k}\frac{1}{\sigma_i^2}.
\end{aligned}
$$

III. Local power of Inverse Normal test $[LP(INN)]$

Under this test, the null hypothesis will be rejected if $\frac{1}{\sqrt{k}}\sum_{i=1}^{k} U_i < -z_\alpha$, where $U_i = \Phi^{-1}(P_i)$, Φ^{-1} is the inverse cdf and z_α is the upper α level critical value of a standard normal distribution. This leads to

$$\text{Power} = Pr\left\{\frac{1}{\sqrt{k}}\sum_{i=1}^{k} U_i < -z_\alpha | H_1\right\}.$$

First, let us determine the pdf of U under H_1, $f_{H_1}(u)$, via its cdf $F_{H_1}(u) = Pr\{U \leq u|H_1\}$.

$$\begin{aligned}
Pr\{U \leq u|H_1\} &= Pr\{\Phi(U) \leq \Phi(u)|H_1\} \\
&= Pr\{P \leq \Phi(u)|H_1\} \qquad \big[U = \Phi^{-1}(P) \implies P = \Phi(U)\big] \\
&= 1 - Pr\{P > \Phi(u)|H_1\} \\
&\approx 1 - \left[[1-\Phi(u)] + \frac{n\Delta^2}{2\sigma^2}\big[\xi_\nu(c)\big]_{c=\Phi(u)}\right] \qquad \big[\text{upon applying Lemma 1}\big] \\
&\approx \Phi(u) - \frac{n\Delta^2}{2\sigma^2}\big[\xi_\nu(c)\big]_{c=\Phi(u)}.
\end{aligned}$$

This implies

$$\begin{aligned}
f_{H_1}(u) &\approx \frac{d}{du}\left[\Phi(u) - \frac{n\Delta^2}{2\sigma^2}\big[\xi_\nu(c)\big]_{c=\Phi(u)}\right] \\
&\approx \phi(u)\left[1 - \frac{n\Delta^2}{2\sigma^2}\left(\frac{d}{du}\big[\xi_\nu(c)\big]_{c=\Phi(u)}\right)\right] \\
&\approx \frac{\phi(u)\big[1 + \frac{n\nu\Delta^2}{\sigma^2}Q_\nu(u)\big]}{1 + \frac{n\nu\Delta^2}{\sigma^2}\int_{-\infty}^{\infty}\phi(u)Q_\nu(u)du}, \quad Q_\nu(u) = \left[\frac{x^2-1}{x^2+\nu}\right]_{x=t_\nu(\frac{c}{2}),\ c=\Phi(u)}.
\end{aligned}$$

Here we have used the fact that $\frac{d}{du}[\xi_\nu(c)] = \frac{d}{dc}[\xi_\nu(c)]\frac{dc}{du}$, $\frac{d}{dc}[\xi_\nu(c)] = Q_\nu(\cdot)$ given above, upon simplification, and $\frac{dc}{du} = \phi(u)$. The denominator in the last expression is a normalizing constant.

Let us define A_ν, B_ν and C_ν as $A_\nu = \int_{-\infty}^{\infty} u\phi(u)Q_\nu(u)du$, $B_\nu = \int_{-\infty}^{\infty} u^2\phi(u)Q_\nu(u)du$ and $C_\nu = \int_{-\infty}^{\infty}\phi(u)Q_\nu(u)du$. Using these three quantities, we now approximate the distribution of U as:

$$\begin{aligned}
U \quad &\sim \quad N[E(U), Var(U)] \quad \text{where} \quad E(u) = \int_{-\infty}^{\infty} uf_{H_1}(u)du \approx \frac{n\nu\Delta^2}{\sigma^2}A_\nu \quad \text{and} \\
&Var(U) = \int_{-\infty}^{\infty} u^2 f_{H_1}(u)du \approx 1 + \frac{n\nu\Delta^2}{\sigma^2}[B_\nu - C_\nu].
\end{aligned}$$

This leads to:

$$\begin{aligned}
\frac{1}{\sqrt{k}}\sum_{i=1}^{k} U_i &\sim N\left[\frac{1}{\sqrt{k}}\sum_{i=1}^{k} E(U_i), \frac{1}{k}\sum_{i=1}^{k} Var(U_i)\right] \\
&\sim N\left[\frac{\Delta^2}{\sqrt{k}}\delta_1, 1 + \frac{\Delta^2}{k}\delta_2\right]
\end{aligned}$$

$$\text{where} \quad \delta_1 = \sum_{i=1}^{k}\frac{n_i\nu_i}{\sigma_i^2}A_{\nu_i} \ \text{ and } \ \delta_2 = \sum_{i=1}^{k}\frac{n_i\nu_i}{\sigma_i^2}[B_{\nu_i} - C_{\nu_i}].$$

Using the above result, the local power of inverse normal test is obtained by approximating its $Power = Pr\left\{\frac{1}{\sqrt{k}}\sum_{i=1}^{k} U_i < -z_\alpha | H_1\right\}$ as

$$\begin{aligned}
\text{Local power (INN)} &\approx \Phi\left[\frac{-z_\alpha - \frac{\Delta^2}{\sqrt{k}}\delta_1}{\sqrt{1+\frac{\Delta^2}{k}\delta_2}}\right] \\
&\approx \Phi\left[-z_\alpha - \frac{\Delta^2}{\sqrt{k}}\delta_1 + \frac{z_\alpha}{2}\frac{\Delta^2}{k}\delta_2\right] \\
&\approx \Phi\left[-z_\alpha + \frac{\Delta^2}{\sqrt{k}}\left(\frac{z_\alpha}{2\sqrt{k}}\delta_2 - \delta_1\right)\right] \\
&\approx \Phi(-z_\alpha) + \frac{\Delta^2}{\sqrt{k}}\phi(z_\alpha)\left[\frac{z_\alpha}{2\sqrt{k}}\delta_2 - \delta_1\right] \\
&\approx \alpha + \frac{\Delta^2}{\sqrt{k}}\phi(z_\alpha)\left[\frac{z_\alpha}{2\sqrt{k}}\delta_2 - \delta_1\right].
\end{aligned}$$

Substituting back the expressions for δ_1 and δ_2 results in:

$$LP(INN) \approx \alpha + \frac{\Delta^2}{\sqrt{k}}\phi(z_\alpha)\sum_{i=1}^{k}\frac{n_i\nu_i}{\sigma_i^2}\left[\frac{z_\alpha[B_{\nu_i} - C_{\nu_i}]}{2\sqrt{k}} - A_{\nu_i}\right].$$

For the special case $n_1 = \cdots = n_k = n$ and $\nu_1 = \cdots = \nu_k = \nu = n-1$, the local power of *Inverse Normal test* reduces to:

$$\begin{aligned}
LP(INN) &\approx \alpha + \frac{n\nu\Delta^2}{\sqrt{k}}\phi(z_\alpha)\left(\sum_{i=1}^{k}\frac{1}{\sigma_i^2}\right)\left[\frac{z_\alpha[B_\nu - C_\nu]}{2\sqrt{k}} - A_\nu\right] \\
&= \alpha + \left[\frac{n\Delta^2}{2}\Psi\right]\frac{2\nu}{\sqrt{k}}\phi(z_\alpha)\left[\frac{z_\alpha[B_\nu - C_\nu]}{2\sqrt{k}} - A_\nu\right] \quad \text{where } \Psi = \sum_{i=1}^{k}\frac{1}{\sigma_i^2}.
\end{aligned}$$

IV. Local power of Fisher test $[LP(F)]$

According to Fisher's exact test, the null hypothesis will be rejected if $\sum_{i=1}^{k} U_i > \chi^2_{2k;\alpha}$, where $U_i = -2\ln(P_i)$, and $\chi^2_{2k;\alpha}$ is the upper α level critical value of a χ^2-distribution with $2k$ degrees of freedom. This leads to

$$\text{Power} = Pr\left\{\sum_{i=1}^{k} U_i > \chi^2_{2k;\alpha} | H_1\right\}.$$

In a similar way to the inverse normal test in Appendix III, first let us determine the pdf of U under H_1, $g_{H_1}(u)$, via its cdf $G_{H_1}(u) = Pr\{U \le u | H_1\}$.

$$
\begin{aligned}
Pr\{U \le u|H_1\} &= Pr\{-2\ln(P) \le u|H_1\} \\
&= Pr\{\ln(P) > -u/2|H_1\} \\
&= Pr\{P > \exp(-u/2)|H_1\} \\
&\approx [1-\exp(-u/2)] + \frac{n\Delta^2}{2\sigma^2}\big[\xi_\nu(c)\big]_{c=\exp(-u/2)} \quad \big[\text{upon applying Lemma 1}\big].
\end{aligned}
$$

This implies

$$
\begin{aligned}
g_{H_1}(u) &\approx \frac{d}{du}\left[1-\exp(-u/2) + \frac{n\Delta^2}{2\sigma^2}\big[\xi_\nu(c)\big]_{c=\exp(-u/2)}\right] \\
&\approx \frac{\frac{1}{2}\exp(-u/2)\big[1+\frac{n\nu\Delta^2}{\sigma^2}\Psi_\nu(u)\big]}{1+\frac{n\nu\Delta^2}{\sigma^2}\big[\int_0^\infty \frac{1}{2}\exp(-u/2)\Psi_\nu(u)du\big]}, \quad \Psi_\nu(u) = \left[\frac{x^2-1}{x^2+\nu}\right]_{x=t_\nu(\frac{c}{2}),\ c=\exp(-u/2)}.
\end{aligned}
$$

The denominator again stands for a normalizing constant.

Define $D_0 = \int_0^\infty \frac{1}{\Gamma(k)}\exp(-u)u^{k-1}\ln(u)du$ and $D_\nu = \int_0^\infty \frac{1}{2}\exp(-u/2)(u-2)\Psi_\nu(u)du$. Using these quantities, we can now approximate the distribution of U as:

$$
U \sim Gamma[\beta = 2, \gamma_\nu] \quad \text{where } \gamma_\nu = \big[1+\frac{n\nu\Delta^2}{2\sigma^2}D_\nu\big].
$$

Here Gamma$[\beta,\gamma_\nu]$ stands for a Gamma random variable with scale parameter β and shape parameter γ_ν with the pdf $f(x) = [e^{-x/\beta}x^{\gamma_\nu - 1}]/[\beta^{\gamma_\nu}\Gamma(\gamma_\nu)]$. By the additive property of independent $Gamma[\beta = 2, \gamma_{\nu_1}], \cdots, Gamma[\beta = 2, \gamma_{\nu_k}]$ corresponding to $U_1, \cdots, U_k$, we readily get the approximate distribution of $(U_1 + \cdots + U_k)$ as:

$$
\sum_{i=1}^k U_i \sim Gamma\big[\beta = 2, k+\Delta^2 A\big] \quad \text{where } A = \frac{1}{2}\sum_{i=1}^k \frac{n_i\nu_i}{\sigma_i^2}D_{\nu_i}.
$$

The local power of Fisher test under H_1 is then obtained as follows:

$$
\begin{aligned}
\text{Local power (F)} &\approx \int_{\chi^2_{2k;\alpha}}^\infty \frac{\exp(-t/2)t^{k+A\Delta^2-1}}{2^{k+A\Delta^2}\Gamma(k+A\Delta^2)}dt \qquad \left[\text{since } \sum_{i=1}^k U_i \sim Gamma\big[\beta = 2, k+\Delta^2 A\big]\right] \\
&= Q(\Delta^2).
\end{aligned}
$$

We now expand $Q(\Delta^2)$ around $\Delta^2 = 0$ to get

$$
\begin{aligned}
\text{Local power (F)} &\approx \alpha + \Delta^2 \int_{\chi^2_{2k;\alpha}}^\infty \frac{\exp(-t/2)t^{k-1}}{2^k}\left[\frac{\partial}{\partial\Delta^2}\left(\frac{(t/2)^{A\Delta^2}}{\Gamma(k+A\Delta^2)}\right)_{\Delta^2=0}\right]dt \\
&\approx \alpha + \Delta^2 \int_{\chi^2_{2k;\alpha}}^\infty \frac{\exp(-t/2)t^{k-1}}{2^k}\left[\frac{A\ln(t/2)}{\Gamma(k)} - \frac{A\int_0^\infty \exp(-u)u^{k-1}\ln(u)du}{\Gamma^2(k)}\right]dt \\
&\approx \alpha + \Delta^2 A \int_{\chi^2_{2k;\alpha}}^\infty \frac{\exp(-t/2)t^{k-1}}{2^k\Gamma(k)}\left[\ln(t/2) - \frac{\int_0^\infty \exp(-u)u^{k-1}\ln(u)du}{\Gamma(k)}\right]dt \\
&\approx \alpha + \Delta^2 A\left[E\left\{\{\ln(T/2)\}I_{\{T\ge\chi^2_{2k;\alpha}\}}\right\}_{T\sim\chi^2_{2k}} - \alpha D_0\right].
\end{aligned}
$$

Substituting back the expressions for A results in:

$$LP(F) \approx \alpha + \frac{\Delta^2}{2}\Big[\sum_{i=1}^{k}\frac{n_i\nu_i}{\sigma_i^2}D_{\nu_i}\Big]\Big[E\Big\{\{\ln(T/2)\}I_{\{T\geq\chi^2_{2k;\alpha}\}}\Big\}_{T\sim\chi^2_{2k}} - \alpha D_0\Big].$$

For the special case $n_1 = \cdots = n_k = n$ and $\nu_1 = \cdots = \nu_k = \nu = n-1$, the local power of *Fisher* test reduces to:

$$\begin{aligned} LP(F) &\approx \alpha + \frac{n\Delta^2}{2}\nu D_\nu\Big[\sum_{i=1}^{k}\frac{1}{\sigma_i^2}\Big]\Big[E\Big\{\{\ln(T/2)\}I_{\{T\geq\chi^2_{2k;\alpha}\}}\Big\}_{T\sim\chi^2_{2k}} - \alpha D_0\Big] \\ &= \alpha + \Big[\frac{n\Delta^2}{2}\Psi\Big]\nu D_\nu\Big[E\Big\{\{\ln(T/2)\}I_{\{T\geq\chi^2_{2k;\alpha}\}}\Big\}_{T\sim\chi^2_{2k}} - \alpha D_0\Big] \quad \text{where } \Psi = \sum_{i=1}^{k}\frac{1}{\sigma_i^2}. \end{aligned}$$

V. Local power of a modified t test $[LP(T_1)]$

Using this exact test based on a modified t, the null hypothesis $H_0 : \mu = \mu_0$ will be rejected if $T_1 > d_{1\alpha}$, where $T_1 = \sum_{i=1}^{k} w_{1i}|t_i|$, $w_{1i} \propto [Var(|t_i|)]^{-1}, Var(|t_i|) = [\nu_i(\nu_i-2)^{-1}] - \big([\Gamma(\frac{\nu_i-1}{2})\sqrt{\nu_i}][\Gamma(\frac{\nu_i}{2})\sqrt{\pi}]^{-1}\big)^2$, and $Pr\{T_1 > d_{1\alpha}\,|\,H_0\} = \alpha$. In applications $d_{1\alpha}$ is computed by simulation. This leads to

$$\begin{aligned} \text{Power of } T_1 &= Pr\Big\{\sum_{i=1}^{k} w_{1i}|t_i| > d_{1\alpha}|H_1\Big\} \\ &= \int\cdots\int_{\sum_{i=1}^{k} w_{1i}|t_i|>d_{1\alpha}} \prod_{i=1}^{k}\big[f_{\nu_i,\delta_i}(t_i)\big]\mathrm{d}t_i \qquad \Big[\delta_i = \frac{\sqrt{n_i}\Delta}{\sigma_i}\Big] \\ &\approx \int\cdots\int_{\sum_{i=1}^{k} w_{1i}|t_i|>d_{1\alpha}} \prod_{i=1}^{k}\Big[f_{\nu_i}(t_i) + \delta_i\frac{\partial f_{\nu_i,\delta_i}(t_i)}{\partial\delta_i}\Big|_{\delta_i=0} + \frac{\delta_i^2}{2}\frac{\partial^2 f_{\nu_i,\delta_i}(t_i)}{\partial\delta_i^2}\Big|_{\delta_i=0}\Big]\mathrm{d}t_i \\ &\approx \alpha + \sum_{j=1}^{k}\frac{\delta_j^2\nu_j}{2}\Big[\int\cdots\int_{\sum_{i=1}^{k} w_{1i}|t_i|>d_{1\alpha}}\Big\{\prod_{i=1}^{k} f_{\nu_i}(t_i)\Big\}\Big\{\frac{\frac{\partial^2 f_{\nu_j,\delta_j}(t_j)}{\partial\delta^2}\big|_{\delta=0}}{f_{\nu_j}(t_j)}\Big\}\Big]\prod_{i=1}^{k}\mathrm{d}t_i \\ &\approx \alpha + \sum_{j=1}^{k}\frac{\delta_j^2\nu_j}{2}\Big[E_{H_0}\Big[\Big\{\frac{\frac{\partial^2 f_{\nu_j,\delta_j}(t_j)}{\partial\delta_j^2}\big|_{\delta_j=0}}{f_{\nu_j}(t_j)}\Big\}I_{\{\sum_{i=1}^{k} w_{1i}|t_i|>d_{1\alpha}\}}\Big]\Big] \\ &\approx \alpha + \sum_{j=1}^{k}\frac{\delta_j^2\nu_j}{2}\Big[E_{H_0}\Big[\Big\{\frac{(t_j^2-1)\nu_j}{t_j^2+\nu_j}\Big\}I_{\{\sum_{i=1}^{k} w_{1i}|t_i|>d_{1\alpha}\}}|H_0\Big]\Big] \\ &\approx \alpha + \frac{\Delta^2}{2}\Big(\sum_{j=1}^{k}\frac{n_j\nu_j}{\sigma_j^2}E_{H_0}\Big[\Big\{\frac{(t_j^2-1)\nu_j}{t_j^2+\nu_j}\Big\}I_{\{\sum_{i=1}^{k} w_{1i}|t_i|>d_{1\alpha}\}}\Big]\Big) \text{ using } \Big[\delta_j = \frac{\sqrt{n_j}\Delta}{\sigma_j}\Big]. \end{aligned}$$

$E_{H_0}[\cdot]$ above is computed by simulation. It is easy to verify from Section 3 that the product terms $\Big\{\frac{\partial f_{\nu_i,\delta_i}(t_i)}{\partial\delta_i}\Big|_{\delta_i=0}\Big\} \times \Big\{\frac{\partial f_{\nu_j,\delta_j}(t_j)}{\partial\delta_j}\Big|_{\delta_j=0}\Big\}$ involve (t_it_j), apart from t_i^2 and t_j^2, whose integral over $\{\sum_{i=1}^{k} w_{1i}|t_i| > d_{1\alpha}\}$ under H_0 is zero.

For the special case $n_1 = \cdots = n_k = n$ and $\nu_1 = \cdots = \nu_k = \nu = n-1$ which implies $w_{11} = \cdots = w_{1k} = 1$, the local power of this exact test based on modified t reduces to:

$$
\begin{aligned}
LP(T_1) &\approx \alpha + \frac{n\Delta^2}{2}\Big(\sum_{j=1}^{k}\frac{1}{\sigma_j^2}\Big)\nu E_{H_0}\Big[\Big\{\frac{(t_1^2-1)\nu}{t_1^2+\nu}\Big\}I_{\{\sum_{i=1}^{k}|t_i|>d_{1\alpha}\}}\Big] \\
&= \alpha + \Big[\frac{n\Delta^2}{2}\Psi\Big]\nu E_{H_0}\Big[\Big\{\frac{(t_1^2-1)\nu}{t_1^2+\nu}\Big\}I_{\{\sum_{i=1}^{k}|t_i|>d_{1\alpha}\}}\Big] \quad \text{where } \Psi = \sum_{j=1}^{k}\frac{1}{\sigma_j^2}.
\end{aligned}
$$

VI. Local power of a modified F test $[LP(T_2)]$

According to this exact test based on a modified F, the null hypothesis $H_0 : \mu = \mu_0$ will be rejected if $T_2 > d_{2\alpha}$, where $T_2 = \sum_{i=1}^{k} w_{2i}F_i$, $F_i \sim F(1,\nu_i)$, $w_{2i} \propto [Var(F_i)]^{-1} = [2\nu_i^2(\nu_i-1)]^{-1}[(\nu_i-2)^2(\nu_i-4)]$, and $Pr\{T_2 > d_{2\alpha}\,|\,H_0\} = \alpha$. In applications $d_{2\alpha}$ is computed by simulation. This leads to

$$
\begin{aligned}
\text{Power of } T_2 &= Pr\Big\{\sum_{i=1}^{k} w_{2i}F_i > d_{2\alpha}|H_1\Big\} \\
&= \int\cdots\int_{\sum_{i=1}^{k} w_{2i}F_i>d_{2\alpha}} \prod_{i=1}^{k}\big[f_{\nu_i,\lambda_i}(F_i)\big]\mathrm{d}F_i \qquad \Big[f_{\nu,\lambda}(F) \sim \text{non-central } F_{1,\nu}\Big(\lambda = \frac{n\Delta^2}{\sigma^2}\Big)\Big].
\end{aligned}
$$

Note that $f_{\nu,\lambda}(F)$ and its local expansion around $\lambda = 0$ are give by

$$
\begin{aligned}
f_{\nu,\lambda}(F) &= \exp(-\frac{\lambda}{2})\sum_{j=0}^{\infty}\frac{(\frac{\lambda}{2})^j}{j!}\Bigg[\frac{(\frac{\nu_1}{\nu_2})^{\frac{\nu_1+2j}{2}}\Gamma(\frac{\nu_1+\nu_2+2j}{2})}{\Gamma(\frac{\nu_1+2j}{2})\Gamma(\frac{\nu_2}{2})}\Bigg]\Bigg[\frac{F^{\frac{\nu_1+2j}{2}-1}}{\big(1+F\frac{\nu_1}{nu_2}\big)^{\frac{\nu_1+\nu_2+2j}{2}}}\Bigg] \\
&\approx f_\nu(F)\big(1-\frac{\lambda}{2}\big) + \Bigg[\frac{(\frac{\lambda}{2})(\frac{\nu_1}{\nu_2})^{\frac{\nu_1+2}{2}}\Gamma(\frac{\nu_1+\nu_2+2}{2})}{\Gamma(\frac{\nu_1+2}{2})\Gamma(\frac{\nu_2}{2})}\Bigg]\Bigg[\frac{F^{\nu_1}}{\big(1+F\frac{\nu_1}{\nu_2}\big)^{\frac{\nu_1+\nu_2+2}{2}}}\Bigg] \\
&= f_\nu(F) + \frac{\lambda}{2}\big[f_\nu^*(F) - f_\nu(F)\big], \quad \text{where } f_\nu^*(F) = \Big(\frac{1}{\nu}\Big)^{\frac{3}{2}}\Bigg[\frac{F}{(1+\frac{F}{\nu})^{\frac{\nu+3}{2}}B[\frac{3}{2},\frac{\nu}{2}]}\Bigg].
\end{aligned}
$$

Using the above first order expansion of $f_{\nu,\lambda}(F)$ leads to the following local power of T_2.

$$
\begin{aligned}
LP(T_2) &\approx \int\cdots\int_{\sum_{i=1}^{k} w_{2i}F_i>d_{2\alpha}} \Bigg[\prod_{i=1}^{k} f_{\nu_i}(F_i) + \sum_{j=1}^{k}\frac{\lambda_j}{2}\Big(f_{\nu_j}^*(F_j) - f_{\nu_j}(F_j)\Big)\Big\{\prod_{i\neq j}\big[f_{\nu_i}(F_i)\big]\Big\}\Bigg]\prod_{i=1}^{k}\mathrm{d}F_i \\
&\approx \alpha + \Bigg(\sum_{j=1}^{k}\frac{\lambda_j}{2}E_{H_0}\Bigg[\Bigg\{\frac{f_{\nu_j}^*(F_j) - f_{\nu_j}(F_j)}{f_{\nu_j}(F_j)}\Bigg\}I_{\{\sum_{i=1}^{k} w_{2i}F_i>d_{2\alpha}\}}\Bigg]\Bigg)
\end{aligned}
$$

$E_{H_0}[\cdot]$ stands for expectation w.r.t $F_1,\ldots,F_k$ under $H_0[F_i \sim F(1,\nu_i)]$.

$$
\begin{aligned}
&\approx \alpha + \Bigg(\sum_{j=1}^{k}\frac{\lambda_j}{2}E_{H_0}\Bigg[\Bigg\{\frac{F_j-1}{\frac{F_j}{\nu_j}+1}\Bigg\}I_{\{\sum_{i=1}^{k} w_{2i}F_i>d_{2\alpha}\}}\Bigg]\Bigg) \\
&\approx \alpha + \frac{\Delta^2}{2}\Bigg(\sum_{j=1}^{k}\frac{n_j}{\sigma_j^2}E_{H_0}\Bigg[\Bigg\{\frac{[F_j-1]\nu_j}{F_j+\nu_j}\Bigg\}I_{\{\sum_{i=1}^{k} w_{2i}F_i>d_{2\alpha}\}}\Bigg]\Bigg) \quad \text{using } \Bigg[\lambda_j = \frac{n_j\Delta^2}{\sigma_j^2}\Bigg].
\end{aligned}
$$

$E_{H_0}[\cdot]$ is obtained by simulation.

For the special case $n_1 = \cdots = n_k = n$ and $\nu_1 = \cdots = \nu_k = \nu = n-1$ which implies $w_{21} = \cdots = w_{2k} = 1$, the local power of this exact test based on modified F reduces to:

$$\begin{aligned} LP(T_2) &\approx \alpha + \frac{n\Delta^2}{2}\left(\sum_{j=1}^{k}\frac{1}{\sigma_j^2}\right)E_{H_0}\left[\left\{\frac{[F_1-1]\nu}{F_1+\nu}\right\}I_{\{\sum_{i=1}^{k}F_i > d_{2\alpha}\}}\right] \\ &= \alpha + \left[\frac{n\Delta^2}{2}\Psi\right]E_{H_0}\left[\left\{\frac{[F_1-1]\nu}{F_1+\nu}\right\}I_{\{\sum_{i=1}^{k}F_i > d_{2\alpha}\}}\right] \quad \text{where } \Psi = \sum_{j=1}^{k}\frac{1}{\sigma_j^2}. \end{aligned}$$

References

Chang, C.-H. and Pal, N. (2008). Testing on the common mean of several normal distributions. *Computational Statistics and Data Analysis*, 53(2):321–333.

Cohen, A. and Sackrowitz, H. (1984). Testing hypotheses about the common mean of normal distributions. *Journal of Statistical Planning and Inference*, 9(2):207–227.

Eberhardt, K. R., Reeve, C. P., and Spiegelman, C. H. (1989). A minimax approach to combining means, with practical examples. *Chemometrics and Intelligent Laboratory Systems*, 5(2):129–148.

Fairweather, W. R. (1972). A method of obtaining an exact confidence interval for the common mean of several normal populations. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 21(3):229–233.

Fisher, R. (1932). Statistical methods for research workers. ed. 4, 307 pp., illus. *Edinburgh and London.*

Graybill, F. A. and Deal, R. (1959). Combining unbiased estimators. *Biometrics*, 15(4):543–550.

Hartung, J. (1999). An alternative method for meta-analysis. *Biometrical Journal: Journal of Mathematical Methods in Biosciences*, 41(8):901–916.

Hartung, J., Knapp, G., and Sinha, B. (2008). Statistical meta-analysis with application. *John Wiley and Sons, Inc., USA.*

Jordan, S. M. and Krishnamoorthy, K. (1996). Exact confidence intervals for the common mean of several normal populations. *Biometrics*, pages 77–86.

Khatri, C. and Shah, K. (1974). Estimation of location parameters from two linear models under normality. *Communications in Statistics-Theory and Methods*, 3(7):647–663.

Lipták, T. (1958). On the combination of independent tests. *Magyar Tud Akad Mat Kutato Int Kozl*, 3:171–197.

Malekzadeh, A. and Kharrati-Kopaei, M. (2018). Inferences on the common mean of several normal populations under heteroscedasticity. *Computational Statistics*, 33(3):1367–1384.

Meier, P. (1953). Variance of a weighted mean. *Biometrics*, 9(1):59–73.

Philip, L., Sun, Y., and Sinha, B. K. (1999). On exact confidence intervals for the common mean of several normal populations. *Journal of Statistical Planning and Inference*, 81(2):263–277.

Sinha, B. K. (1985). Unbiased estimation of the variance of the Graybill-Deal estimator of the common mean of several normal populations. *Canadian Journal of Statistics*, 13(3):243–247.

Stouffer, S. A., Suchman, E. A., DeVinney, L. C., Star, S. A., and Williams Jr, R. M. (1949). The American soldier: Adjustment during army life. *Princeton University Press.*

Tippett, L. H. C. et al. (1931). The methods of statistics. *The Methods of Statistics.*

Wilkinson, B. (1951). A statistical consideration in psychological research. *Psychological Bulletin*, 48(2):156.