A SUMMARY OF CONFIDENCE INTERVAL ESTIMATION OF STANDARD AND CERTAIN NON-CENTRALITY PARAMETERS

by

Homer Thornton Hayslett, Jr.

Thesis submitted to the Graduate Faculty of the

Virginia Polytechnic Institute

in candidacy for the degree of

MASTER OF SCIENCE

in

Statistics

May, 1961

Blacksburg, Virginia

TABLE OF CONTENTS

Chapter | | Page

I. INTRODUCTION

The area of confidence interval estimation is an important one in applied statistics. We frequently like to know just how confident we can be that a certain estimated interval does, in fact, cover the true parameter.

This paper is a review of confidence interval estimation on some of the familiar parameters of the normal distribution and a presentation of confidence interval estimation of some not so familiar parameters. The dimension-free parameters, such as $\frac{\mu^2}{\sigma^2}$, are of especial interest since they are independent of the unit of measurement used in the original data.

Confidence interval estimation on the correlation coefficient and on the non-centrality parameters in the χ^2, t^2, and F distributions may be obtained by interpolation in detailed tables of the percentage points of the exact non-central distributions. Although such tables are available, they are quite voluminous, and may not be easily accessible. For that reason, Fisher's z-transformation has been used rather widely, even though tables of the exact distribution have been available. In this thesis similar techniques will be described based upon improved variance stabilizing transformations of the non-central χ^2 and F distributions which represent approximations to the exact distributions as good as that based upon Hotelling's improvement of Fisher's z-transformation. They were obtained by methods analogous to Hotelling's [4], and studied by Bargmann [1] and Hofer [5].

II. REVIEW OF LITERATURE

A treatment of the parameters included in the first five sections of this thesis can be found in many standard textbooks or manuals.

Hotelling [4] has discussed improvements of the mean and variance of the well-known Fisher z-transformation used to put confidence bounds on the correlation coefficient.

Roy and Potthoff [6] present confidence bounds on the parameters μ_1/μ_2 and σ_x^2/σ_y^2 for an underlying bivariate normal distribution.

The variance-stabilizing transformation, used to put confidence limits on the non-centrality parameter in the non-central t^2, F, and χ^2 distribution, is discussed by Bargmann [1].

III. STATEMENT AND DEVELOPMENT OF FORMULAS

3.1 Parameter μ.

When we have a sample of size N from a normal distribution and wish to make a confidence statement about μ when σ^2 is known, we use the fact that

$$(3.1.1) \qquad \frac{(\bar{x} - \mu)\sqrt{N}}{\sigma} = N(0, 1).$$

The usual two-sided confidence statement with equal tail proportions is

$$(3.1.2) \qquad \Pr\{\bar{x} - \phi^{-1}(1 - \tfrac{\alpha}{2})\tfrac{\sigma}{\sqrt{N}} < \mu < \bar{x} + \phi^{-1}(1 - \tfrac{\alpha}{2})\tfrac{\sigma}{\sqrt{N}}\} = 1 - \alpha$$

where $\phi^{-1}(1 - \frac{\alpha}{2})$ is the upper $\frac{\alpha}{2}$ point of the standard normal distribution.

If σ^2 is unknown, the statistic is

$$(3.1.3) \qquad (\bar{x} - \mu)\sqrt{N}/s = t_{N-1}$$

with corresponding confidence statement

$$(3.1.4) \qquad \Pr\{\bar{x} - t\sqrt{S_{xx}/N(N-1)} < \mu < \bar{x} + t\sqrt{S_{xx}/N(N-1)}\} = 1 - \alpha,$$

where t is the upper $\alpha/2$ point of the t distribution with N - 1 degrees of freedom, and $S_{xx} = \Sigma x_i^2 - (\Sigma x_i)^2/N$.

3.2 Parameter σ^2.

For the parameter σ^2, when μ is known, the fact that

(3.2.1) $\Sigma(x_i - \mu)^2/\sigma^2 = \chi^2_N$

is used to give the confidence statement with equal tail proportions

(3.2.2) $\Pr\{\Sigma(x_i - \mu)^2/\chi^2_{1-\alpha/2} < \sigma^2 < \Sigma(x_i - \mu)^2/\chi^2_{\alpha/2}\} = 1 - \alpha$,

where $\chi^2_{\alpha/2}$ denotes the lower $\alpha/2$ point of the χ^2-distribution with N degrees of freedom.

If μ is unknown, the statistic

(3.2.3) $S_{xx}/\sigma^2 = \chi^2_{N-1}$

is used and the confidence statement is

(3.2.4) $\Pr\{S_{xx}/\chi^2_{1-\alpha/2} < \sigma^2 < S_{xx}/\chi^2_{\alpha/2}\} = 1 - \alpha$,

where S_{xx} is defined in (3.1.4) and χ^2 has N - 1 degrees of freedom.

3.3 Parameter $\mu_1 - \mu_2$.

If we have two samples from two independent normal populations with a known common σ^2 and wish to make a confidence statement about $\mu_1 - \mu_2$, the appropriate distribution is

$$(3.3.1) \qquad [(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)]/\sigma\sqrt{1/N_1 + 1/N_2} = N(0, 1).$$

The confidence statement is

$$(3.3.2) \qquad \Pr\{(\bar{x}_1 - \bar{x}_2) - \phi^{-1}(1 - \alpha/2)\sigma\sqrt{1/N_1 + 1/N_2} < \mu_1 - \mu_2$$
$$< (\bar{x}_1 - \bar{x}_2) + \phi^{-1}(1 - \alpha/2)\sigma\sqrt{1/N_1 + 1/N_2}\} = 1 - \alpha.$$

When σ^2 is unknown, the expression used to put confidence bounds on $\mu_1 - \mu_2$ is

$$(3.3.3) \qquad [(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)]/s_p\sqrt{1/N_1 + 1/N_2} = t_{N_1+N_2-2}$$

and the confidence statement is

$$\Pr\{(\bar{x}_1-\bar{x}_2)-ts_p\sqrt{1/N_1+1/N_2} < \mu_1-\mu_2 < (\bar{x}_1-\bar{x}_2)+ts_p\sqrt{1/N_1 \quad 1/N_2}\}$$
$$(3.3.4) \qquad = 1 - \alpha$$

where t is the upper $\alpha/2$ point of the t distribution with $N_1 + N_2 - 2$ degrees of freedom and s_p is the pooled estimate of the standard deviation, given by

$$(3.3.5) \qquad s_p = \sqrt{(S_{xx_1} + S_{xx_2})/(N_1 + N_2 - 2)},$$

where $S_{xx_1} = \Sigma x_{1i}^2 - (\Sigma x_{1i})^2/N_1$ and $S_{xx_2} = \Sigma x_{2i}^2 - (\Sigma x_{2i})^2/N_2$.

3.4 Parameter σ_1^2/σ_2^2.

Suppose we have a sample from each of two independent normal populations and wish to put confidence bounds on σ_1^2/σ_2^2. Then,

$$(3.4.1) \qquad s_1^2\sigma_2^2/s_2^2\sigma_1^2 = F ,$$

where $s_1^2 = \Sigma(x_{1i} - \mu_1)^2/N_1$, $s_2^2 = \Sigma(x_{2i} - \mu_2)^2/N_2$, and F has (N_1, N_2) degrees of freedom if μ_1 and μ_2 are known, or $s_1^2 = \Sigma(x_{1i} - \bar{x}_1)^2/(N_1 - 1)$, $s_2^2 = \Sigma(x_{2i} - \bar{x}_2)^2/(N_2 - 1)$, and F has (N_1-1, N_2-1) degrees of freedom if μ_1 and μ_2 are unknown, gives the confidence statement

$$(3.4.2) \qquad \Pr\{s_1^2/s_2^2F_{1-\alpha/2}) < \sigma_1^2/\sigma_2^2 < s_1^2/(s_2^2F_{\alpha/2})\} = 1 - \alpha,$$

where $F_{1-\alpha/2}$ is the upper $\alpha/2$ point of the F distribution and $F_{\alpha/2}$ is the lower $\alpha/2$ point.

3.5 Correlation Coefficient.

When we have a sample of size N from a bivariate normal population and wish to put confidence bounds on the population correlation coefficient ρ, the familiar Fisher z-transformation is used, i.e.,

$$(3.5.1) \qquad z = \tanh^{-1} r = \tfrac{1}{2} \ln[(1 + r)/(1 - r)],$$

where r is the sample correlation coefficient. Since

(3.5.2) $$E(z) \cong \tanh^{-1}\rho = \tfrac{1}{2} \ln [(1+\rho)/(1-\rho)]$$

and

(3.5.3) $$\operatorname{var}(z) \cong 1/(N-3) ,$$

(3.5.4) $$[z - E(z)]/\sqrt{\operatorname{var}(z)} = \sqrt{N-3}(\tanh^{-1}r - \tanh^{-1}\rho) \cong N(0,1)$$

and the confidence statement for ρ is

(3.5.5) $$\Pr\{z - \phi^{-1}(1-\alpha/2)/\sqrt{N-3} < \tanh^{-1}\rho < z + \phi^{-1}(1-\alpha/2)/\sqrt{N-3}\} \cong 1-\alpha.$$

If a more exact confidence statement is desired, the z-transformation can be improved. The correction for the bias (the bias of z is the excess of $E(z)$ over $\tanh^{-1}\rho$) of z is $\rho/(2N-5)$. For var(z), $1/(N - 8/3)$ should be used instead of $1/(N-3)$. These corrections usually make both the upper and lower bound in (3.5.5) smaller, but generally the improvement is negligible.

References:

Standard textbooks

Hotelling (1953) [4].

3.6 Parameter μ_1/μ_2.

Suppose we have two samples of size N_1 and N_2, respectively, from two independent normal populations, $x_1 = N(\mu_1, \sigma^2)$ and $x_2 = N(\mu_2, \sigma^2)$ and desire to put confidence bounds on $\gamma = \mu_1/\mu_2$. The common σ^2 may be either known or unknown, although usually σ^2 is not known. Introduce

(3.6.1) $z = x_1 - \gamma x_2$.

Since

(3.6.2) $E(\bar{z}) = E(\bar{x}_1 - \gamma\bar{x}_2) = 0$

and

(3.6.3) $\mathrm{var}(\bar{z}) = \mathrm{var}(\bar{x}_1 - \gamma\bar{x}_2) = \sigma^2/N_1 + \gamma^2\sigma^2/N_2$,

then

(3.6.4) $\bar{z}/\sigma\sqrt{1/N_1 + \gamma^2/N_2} = N(0, 1)$

and

(3.6.5) $\bar{z}^2/\sigma^2(1/N_1 + \gamma^2/N_2) = \chi_1^2$.

If σ^2 is unknown, then

(3.6.6) $\bar{z}/s_p\sqrt{1/N_1 + \gamma^2/N_2} = t_{N_1+N_2-2}$

and

(3.6.7) $\bar{z}^2/s_p^2(1/N_1 + \gamma^2/N_2) = t_{N_1+N_2-2}^2 = F_{(1,N_1+N_2-2)}$

where s_p^2 is the pooled mean square of the two samples, as defined in (3.3.5). We can say

(3.6.8) $\Pr\{(\bar{x}_1 - \gamma\bar{x}_2)^2/\sigma^2(1/N_1 + \gamma^2/N_2) < \chi_{1-\alpha}^2\} = 1 - \alpha$,

where $\chi_{1-\alpha}^2$ represents the upper tail value of χ^2 with one degree of freedom for level α. If σ^2 is unknown, $\chi_{1-\alpha}^2$ is replaced by the upper α point of the F distribution with $(1, N_1 + N_2 - 2)$ degrees of freedom. For simplicity of notation we will call these upper α points χ_u^2 and F_u, respectively. From (3.6.8) we get

(3.6.9) $\Pr\{(\bar{x}_1 - \gamma\bar{x}_2)^2 < \chi_u^2\sigma^2(1/N_1 + \gamma^2/N_2)\} = 1 - \alpha$

and if one sets

(3.6.10) $\chi_u^2\sigma^2/N_1 = b_1$ and

(3.6.11) $\chi_u^2\sigma^2/N_2 = b_2$,

(3.6.9) becomes

(3.6.12) $\Pr\{(\bar{x}_1 - \gamma\bar{x}_2)^2 < b_1 + \gamma^2 b_2\} = 1 - \alpha$.

If σ^2 is unknown,

(3.6.13) $b_1 = F_u s_p^2/N_1$ and

(3.6.14) $b_2 = F_u s_p^2/N_2$,

and replace the values given in (3.6.10) and (3.6.11).

We can state the inequality in brackets in the following form, by expanding the left-hand side and collecting powers of γ:

(3.6.15) $\gamma^2(\bar{x}_2^2 - b_2) - 2\gamma\bar{x}_1\bar{x}_2 < b_1 - \bar{x}_1^2$,

which becomes

$$(3.6.16)\quad [\gamma - \bar{x}_1\bar{x}_2/(\bar{x}_2^2 - b_2)]^2 < [b_2\bar{x}_1^2 + b_1\bar{x}_2^2 - b_1b_2]/(\bar{x}_2^2-b_2)^2$$

if $\bar{x}_2^2 - b_2 > 0$.

It will now be shown that the condition $\bar{x}_2^2 - b_2 > 0$ is equivalent to rejection of the hypothesis that $\mu_2 = 0$ in a two-tailed test with significance level α. To this end we will formulate the following

Theorem: A necessary and sufficient condition for the existence of a real-valued confidence interval on $\gamma = \mu_1/\mu_2$, of coefficient $1 - \alpha$, is the rejection of the null hypothesis H_0: $\mu_2 = 0$ vs. the alternative $\mu_2 \neq 0$ at significance level α.

Proof: If σ^2 is known, the critical region for the two-tailed test of H_0: $\mu_2 = 0$ can be written as

$$(3.6.17)\quad \frac{N_2\bar{x}_2^2}{\sigma^2} > \chi^2_{1-\alpha} ,$$

where χ^2 is based upon one degree of freedom and is equal to the χ^2_u used in (3.6.10) and (3.6.11). Hence, if H_0: $\mu_2 = 0$ is rejected, $\bar{x}_2^2 > \frac{\chi^2_u\sigma^2}{N_2} = b_2$ and it follows that $\bar{x}_2^2 - b_2 > 0$, which proves the necessity of the condition.

In order to show sufficiency, we must show that the right-hand side of (3.6.16) is positive. We can say, since $\bar{x}_2^2 > b_2$ and $b_2 > 0$ by (3.6.11), that

$$(3.6.18)\quad b_2\bar{x}_1^2 + b_1\bar{x}_2^2 - b_1b_2 > b_2\bar{x}_1^2 + b_1b_2 - b_1b_2 = b_2\bar{x}_1^2 > 0,$$

which shows that the right-hand side of (3.6.16) is positive and proves the theorem.

If σ^2 is unknown it will be estimated by s_p^2; (3.6.10) and (3.6.11) should be replaced by (3.6.13) and (3.6.14) and the proof is the same as above. Note that even though the test is made on the mean of the second sample only, a pooled estimate of σ^2 from both samples is used.

Subject to the conditions in the preceding theorem we can now find the confidence interval for γ. Let us set

$$(3.6.19) \quad (b_1\bar{x}_2^2 + b_2\bar{x}_1^2 - b_1b_2)^{\frac{1}{2}} = c.$$

Now $a^2 < k$ implies that $-\sqrt{k} < a < \sqrt{k}$; hence we get from (3.6.16) the interval

$$(3.6.20) \quad -c/(\bar{x}_2^2 - b_2) < \gamma - \bar{x}_1\bar{x}_2/(\bar{x}_2^2 - b_2) < c/(\bar{x}_2^2 - b_2)$$

and the final confidence statement is

$$(3.6.21) \quad \Pr\{(\bar{x}_1\bar{x}_2 - c)/(\bar{x}_2^2 - b_2) < \gamma < (\bar{x}_1\bar{x}_2 + c)/(\bar{x}_2^2 - b)\} = 1 - \alpha.$$

For a bivariate normal population

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \left\{ \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \right\},$$

S. N. Roy and R. F. Potthoff [6] give the confidence bounds

$$(3.6.22)\quad \frac{(\bar{x}_1\bar{x}_2-ks_1s_2r)-[(\bar{x}_1\bar{x}_2-ks_1s_2r)^2-(\bar{x}_1^2-ks_1^2)(\bar{x}_2^2-ks_2^2)]^{\frac{1}{2}}}{\bar{x}_2^2-ks_2^2} \leq \gamma \leq \frac{(\bar{x}_1\bar{x}_2-ks_1s_2r)+[(\bar{x}_1\bar{x}_2-ks_1s_2r)^2-(\bar{x}_1^2-ks_1^2)(\bar{x}_2^2-ks_2^2)]^{\frac{1}{2}}}{\bar{x}_2^2-ks_2^2}$$

where r is the sample correlation coefficient, s_1^2 and s_2^2 are the unbiased estimates of σ_1^2 and σ_2^2 respectively, N is sample size, and $k = t_{1-\alpha/2}^2/N$, where $t_{1-\alpha/2}$ is the upper $\alpha/2$ point of the t distribution with $N - 1$ degrees of freedom. The confidence bounds on γ in (3.6.22) are meaningful only if

$$(3.6.23)\quad \bar{x}_1^2/s_1^2 + \bar{x}_2^2/s_2^2 \geq 2\bar{x}_1\bar{x}_2 r/s_1s_2 + k\bar{x}_1^2\bar{x}_2^2(1-r^2)/s_1^2s_2^2.$$

When this condition is not satisfied bounds on μ_1/μ_2 should not be attempted.

References:

Bliss (1935) [2]

Fieller (1940) [3]

Roy and Potthoff (1958) [6]

3.7 Parameter σ_x^2/σ_y^2, Correlated Populations.

Suppose that we have a sample of paired observations, say

$$x_1, x_2, \ldots, x_N$$
$$y_1, y_2, \ldots, y_N$$

from two correlated normal populations (actually one bivariate population, but many authors treat the paired t-test situation and wish to put confidence bounds on σ_x^2/σ_y^2 as if it were based upon two populations). Set $u_i = x_i + (\sigma_x/\sigma_y)y_i$ and $v_i = x_i - (\sigma_x/\sigma_y)y_i$. Then

$$(3.7.1) \qquad \mathrm{cov}(u,v) = \mathrm{cov}(x + \frac{\sigma_x}{\sigma_y}\, y,\ x - \frac{\sigma_x}{\sigma_y}\, y)$$

$$= \sigma_x^2 - \frac{\sigma_x^2}{\sigma_y^2}\,\sigma_y^2 = 0$$

and it follows that u and v are uncorrelated. Hence,

$$(3.7.2) \qquad r_{u,v}\sqrt{N-2}/\sqrt{1-r_{u,v}^2} = t_{N-2} \quad \text{and}$$

$$(3.7.3) \qquad \Pr\{-t < r_{u,v}\sqrt{N-2}/\sqrt{1-r_{u,v}^2} < t\} = 1 - \alpha,$$

where t is the upper $\alpha/2$ point of the t-distribution, and

$$(3.7.4) \qquad r_{u,v}^2 = [\Sigma(u_i-\bar{u})(v_i-\bar{v})]^2/[\Sigma(u_i-\bar{u})^2][\Sigma(v_i-\bar{v})^2] \ .$$

Using the previously defined values of u_i and v_i and setting $\sigma_x/\sigma_y = \sqrt{\lambda}$, $x_i - \bar{x} = x_i'$ and $y_i - \bar{y} = y_i'$, (3.7.4) becomes

$$(3.7.5) \qquad r_{u,v}^2 = \frac{[\Sigma(x_i'^2-\lambda y_i'^2)]^2}{[\Sigma x_i'^2+2\sqrt{\lambda}\Sigma x_i'y_i'+\lambda\Sigma y_i'^2][\Sigma x_i'^2-2\sqrt{\lambda}\Sigma x_i'y_i'+\lambda\Sigma y_i'^2]}$$

Setting $\Sigma x_i'^2/(N-1) = s_x^2$, $\Sigma y_i'^2/(N-1) = s_y^2$ and $\Sigma x_i'y_i'/[(\Sigma x_i'^2)(\Sigma y_i'^2)]^{\frac{1}{2}} = r_{xy} = r$, we obtain

$$(3.7.6) \quad r^2_{u,v} = (s_x^2 - \lambda s_y^2)^2/[(s_x^2 + \lambda s_y^2)^2 - 4\lambda r^2 s_x^2 s_y^2].$$

Then (3.7.3) may be written

$$(3.7.7) \quad \Pr\{-t < (s_x^2 - \lambda s_y^2)\sqrt{N-2}/2\sqrt{\lambda}s_x s_y\sqrt{1-r^2} < t\} = 1 - \alpha,$$

which gives

$$(3.7.8) \quad \Pr\{-2\sqrt{\lambda}\ \sqrt{1-r^2}s_x s_y t < \sqrt{N-2}(s_x^2-\lambda s_y^2) < 2\sqrt{\lambda}\ \sqrt{1-r^2}s_x s_y t\} = 1 - \alpha.$$

Recall that $-a < \alpha < a$ implies that $\alpha^2 < a$. Hence,

$$(3.7.9) \quad \Pr\{(N-2)(s_x^2-\lambda s_y^2)^2 < 4\lambda(1-r^2)s_x^2 s_y^2 t^2\} = 1-\alpha \ ,$$

which becomes, upon setting $s_x^2/s_y^2 = k$,

$$(3.7.10) \quad \Pr\{(N-2)(k-\lambda)^2 < 4\lambda(1-r^2)kt^2\} = 1 - \alpha,$$ which gives

$$(3.7.11) \quad \Pr\{(k-\lambda)^2 < 4\lambda(1 - r^2)kt^2/(N - 2)\} = 1 - \alpha,$$

which may be stated as

$$(3.7.12) \quad \Pr\{\lambda^2 - 2\lambda[k + 2kt^2(1 - r^2)/(N - 2)] < -k^2\} = 1 - \alpha.$$

Completing the square we get

$$(3.7.13) \quad \Pr\ \{\lambda-[k+2kt^2(1-r^2)/(N-2)]\}^2 < 4k^2t^2(1-r^2)/(N-2) + 4k^2t^4(1-r^2)^2/(N-2)^2 = 1- \alpha.$$

Letting $a = t\sqrt{1 - r^2}$ we obtain

$$(3.7.14)$$
$$\Pr\{-2ak\sqrt{N-2+a^2}/(N-2) < \lambda-k-2a^2k/(N-2) < 2ak\sqrt{N-2+a^2}/(N-2)\} = 1-\alpha,$$

which gives the final confidence statement for $\lambda = \sigma_x^2/\sigma_y^2$,

$$(3.7.15) \quad \Pr\{\, k+2ak(a-\sqrt{N-2+a^2})/(N-2) < \lambda$$

$$< k+2ak(a+\sqrt{N-2+a^2})/(N-2)\} = 1-\alpha$$

where a, k, N, t are defined as before.

Reference:

Roy and Potthoff (1958) [6]

3.8 Non-centrality Parameter of t^2.

Suppose we have a sample of N observations, $x_1, x_2, \ldots, x_N$, from a normal population with mean μ and variance σ^2 and want to test the null hypothesis that $\mu = 0$. If H_0: $\mu = 0$ is true, the statistic $\bar{x}\sqrt{N}/s$ has a t-distribution with N - 1 degrees of freedom. If, however, $\mu \neq 0$ in the population, the statistic $N\bar{x}^2/s^2$ has the non-central t^2-distribution with non-centrality parameter $\gamma^2 = N\mu^2/\sigma^2$. If we desire a confidence statement on μ^2/σ^2, we may proceed as follows: Using the fact that

$$(3.8.1) \quad (\bar{x}\sqrt{N}/s)^2 = t^2 = F$$

from the variance-stabilizing transformation of non-central F,

$$(3.8.2) \quad z = \cosh^{-1}(w/a),$$

where in the case with a single degree of freedom in the numerator

$$(3.8.3) \quad a = \sqrt{(N-2)/(N-3)} \text{ and}$$

(3.8.4) $w = 1 + F/(N - 1)$,

then z is approximately normal with

(3.8.5) $E(z) = \zeta - (\coth\zeta)/(N - 5)$ and

(3.8.6) $\mathrm{var}(z) = 2/(N - 5)$, where

(3.8.7) $\zeta = \cosh^{-1}(\gamma^2/\sqrt{(N-2)(N-3)} + a)$.

Since $[z - E(z)]/\sqrt{\mathrm{var}(z)} \cong N(0, 1)$ we can say that

(3.8.8) $$\sqrt{(N-5)/2}\{ z - \zeta + (\gamma^2/\sqrt{N-2} + \sqrt{N-2}) \div [(N-5)\sqrt{\gamma^4/(N-2) + 1 + 2\gamma^2}]\} \cong N(0, 1),$$

and

(3.8.9) $$\Pr\{\phi^{-1}(\alpha/2) < \sqrt{(N-5)/2}\, [z - \zeta + (\gamma^2/\sqrt{N-2} + \sqrt{N-2}) \div (N-5)\sqrt{\gamma^4/(N-2) + 1 + 2\gamma^2}] < \phi^{-1}(1 - \alpha/2)\} \cong 1-\alpha,$$

where $\phi^{-1}(\alpha/2)$ denotes the lower (negative) value of the abscissa which has $\alpha/2\%$ of the area under the normal curve to the left of it. Setting $\phi^{-1}(\alpha/2) = -\phi^{-1}(1 - \alpha/2)$,

(3.8.10) $\gamma^2/\sqrt{N-2} = v$, and

(3.8.11) $(v + \sqrt{N-2})/\sqrt{v^2 + 1 + 2\gamma^2} = u$

(3.8.9) becomes

(3.8.12) $$\Pr\{-\phi^{-1}(1 - \alpha/2)\sqrt{2/(N-5)} < z - \zeta + u/(N-5) < \phi^{-1}(1 - \alpha/2)\sqrt{2/(N-5)}\} = 1 - \alpha,$$

which may be written as

(3.8.13) $\Pr\{z_L < \zeta - u/(N-5) < z_U\} = 1 - \alpha,$ where

(3.8.14) $z_L = z - \phi^{-1}(1 - \alpha/2)\sqrt{2/(N-5)}$ and

(3.8.15) $z_U = z + \phi^{-1}(1 - \alpha/2)\sqrt{2/(N-5)}$.

As a first approximation to the confidence bounds on we may say

(3.8.16) $\Pr\{z_L < \zeta < z_U\} \cong 1 - \alpha$.

Setting

(3.8.17) $\sqrt{(N - 2)(N - 3)} = k$

we have, from (3.8.7),

(3.8.18) $\cosh \zeta = \gamma^2/k + a$

and we may form the equations

(3.8.19) $\cosh z_L = \gamma_L^2/k + a,$ and

(3.8.20) $\cosh z_U = \gamma_U^2/k + a,$

where γ_L^2 and γ_U^2 are the lower and upper limits, respectively, of the approximate confidence interval for the non-centrality parameter γ^2. Equations (3.8.19) and (3.8.20) may be solved for γ_L^2 and γ_U^2, which give our first approximate values of the lower and upper limits of γ^2: $\gamma_{L_0}^2$ and $\gamma_{U_0}^2$, say.

A more exact statement than (3.8.16) is

$$(3.8.21)\quad \Pr\{z'_L < \zeta < z'_U\} = 1 - \alpha,$$

where

$$(3.8.22)\quad z'_L = z_L + u_L/(N - 5)$$

and

$$(3.8.23)\quad z'_U = z_U + u_U/(N - 5),$$

where

$$(3.8.24)\quad u_L = (v_L + \sqrt{N - 2})/\sqrt{v_L^2 + 1 + 2\gamma_L^2}$$

and

$$(3.8.25)\quad v_L = \gamma_L^2/\sqrt{N - 2}\ .$$

The quantities u_U and v_U are defined accordingly in terms of the upper limit of the confidence interval for γ^2.

Only the lower limit will be dealt with in the remainder of this section. The upper limit can be found in an exactly analogous manner. In (3.8.19) the equation $\cosh z_L = \gamma_L^2/k + a$ was given, from which we can find a first approximate value for γ_L^2, $\gamma_{L_0}^2$. But a more exact lower limit for ζ is $z'_L = z_L + u_L/(N - 5)$, and instead of (3.8.19) we can form

$$(3.8.26)\quad \cosh[z_L + u_L/(N - 5)] = \gamma_L^2/k + a,$$

which yields

$$(3.8.27)\quad f(\gamma_L^2) = \gamma_L^2/k + a - \cosh[z_L + u_L/(N - 5)] = 0.$$

The solution of this equantion yields the improved value for the lower limit of the confidence interval.

We have a first estimate of the root of equation (3.3.27). Using the Newton Method, better approximations to the root are given by the formula

$$(3.8.28)\quad \gamma^2_{L_{n+1}} = \gamma^2_{L_n} - f(\gamma^2_{L_n})/f'(\gamma^2_{L_n}) \qquad n = 0, 1, 2, \ldots .$$

Differentiating $f(\gamma^2_L)$ with respect to γ^2_L we have

$$\begin{aligned} \partial f/\partial\gamma^2_L &= 1/k - \partial/\partial\gamma^2_L \cosh[z_L + u_L/(N-5)] \\ &= 1/k - \sinh[z_L + u_L/(N-5)]\, \partial/\partial\gamma^2_L [z_L \\ &\qquad + u_L/(N-5)] \\ &= 1/k - \{\sinh[z_L + u_L/(N-5)]\, \partial u_L/\partial v_L \\ &\qquad \cdot \partial v_L/\partial\gamma^2_L\}/(N-5) \end{aligned}$$

where v_L is defined as in (3.3.25) and

$$\begin{aligned} (3.3.30)\quad u_L &= (v_L + \sqrt{N-2})/\sqrt{v^2_L+1+2\gamma^2_L} \\ &= (v_L + \sqrt{N-2})/\sqrt{v^2_L + 1 + 2v_L\sqrt{N-2}} \quad . \end{aligned}$$

We have

$$(3.8.31)\quad \partial u_L/\partial v_L = -(N-3)/(v^2_L + 1 + 2v_L\sqrt{N-2})^{3/2} \quad \text{and}$$

$$(3.8.32)\quad \partial v_L/\partial\gamma^2_L = 1/\sqrt{N-2}\ .$$

Then

(3.8.33) $\partial u_L/\partial v_L \quad \partial v_L/\partial \gamma_L^2 = -(N-3)/\sqrt{N-2}(v_L^2+1+2v_L\sqrt{N-2})^{3/2}$,

(3.8.34) $\partial f/\partial \gamma_L^2 = 1/k + \{(N-3) \sinh[z_L + u_L/(N-5)]\}$

$\div [(N-5)\sqrt{N-2}(v_L^2+1+2v_L\sqrt{N-2})^{3/2}]$,

and

(3.8.35) $\gamma_{L_{n+1}}^2 = \gamma_{L_n}^2 - \{\gamma_{L_n}^2/k + a - \cosh[z_L + u_L/(N-5)]\}$

$\div\{1/k+[(N-3)\sinh[z_L+u_L/(N-5)] /[(N-5)\sqrt{N-2}(v_L^2+1+2v_L\sqrt{N-2})^{3/2}]\}$,

where u_L and v_L are functions of $\gamma_{L_n}^2$. Equation (3.8.35) yields closer approximations to the lower limit of the confidence interval, i.e., the root of (3.8.27). As mentioned previously, an analogous expression gives the improved approximations to the upper limit of the confidence interval for γ^2.

Since $a < b < c$ implies that $\cosh a < \cosh b < \cosh c$ only if $a > 0$, $b > 0$, and $c > 0$, we must impose the condition that $z_L' > 0$ so that our confidence bounds will be valid. The quantity u_L is always positive. Hence $z_L' > 0$ is certainly satisfied if $z_L > 0$, i.e., if

(3.8.36) $z > \sqrt{2/(N-5)}\ \phi^{-1}(1 - \alpha/2)$.

If we used the improved transformation of the non-central F statistic with approximate normality in a procedure analogous to the use of the Fisher-z transformation for tests concerning correlation coefficients we would have

(3.8.37) $\sqrt{(N-5)/2}\,(z - \zeta) \to N(0, 1)$

and we would reject H_0: $\zeta = 0$ if either

(3.8.38) $z\sqrt{(N-5)/2} > \phi^{-1}(1 - \alpha/2)$ or

(3.8.39) $z\sqrt{(N-5)/2} < \phi^{-1}(\alpha/2)$

and accept the alternative hypothesis that $\zeta \neq 0$.

Statement (3.8.38) is identical to (3.8.36). Then we can expect the above procedure to yield a valid confidence statement of γ^2 if the null hypothesis (here $\mu/\sigma = 0$) has been rejected at level α. However, if the null hypothesis is accepted the value $\gamma^2 = 0$ would have to be included in the confidence interval. But since γ^2 can never be negative it is meaningless to construct a confidence interval for γ^2 which includes zero. This limitation would not hold for a non-central t-distribution with parameter μ/σ (positive or negative), but it has a rather complex form. The non-central t^2-distribution is more easily manageable and contains the positive non-centrality parameter μ^2/σ^2, on which meaningful confidence bounds can be stated only if the null hypothesis $\mu^2/\sigma^2 = 0$ is rejected.

After finding γ_L^2 and γ_U^2 to the desired accuracy we have the confidence statement

(3.8.40) $\Pr\{\gamma_L^2 < \gamma^2 < \gamma_U^2\} = 1 - \alpha.$

Confidence limits on μ^2/σ^2 are obtained by dividing each member of (3.8.38) by N, and the final confidence statement is

(3.8.41) $\Pr\{\gamma_L^2/N < \mu^2/\sigma^2 < \gamma_U^2/N\} = 1 - \alpha.$

Reference:

Bargmann (1958) [1].

3.9 Non-centrality Parameter of F.

Suppose we have a one-way classification with k classes and are testing the null hypothesis that there is no difference between class means. Various procedures have been recommended to describe departures from rejected null-hypotheses. One method assumes the treatment effects to be random and estimates the variance of this component. Confidence bounds for this variance component are exactly solvable only for the simplest case of equal numbers of observations in each group. For unequal numbers of observations in each group, or more complicated designs, exact estimation methods and distributions of estimates are quite complicated. In many applications, and in almost all textbooks, estimation procedures are recommended which are simple but may not be very satisfactory. Confidence bounds based on these approximate techniques are also proposed.

As an alternative, one may retain the Model I type analysis and express the degree of departure from the null hypothesis by a standardized, dimension-free measure of departure (the square of a distance function) known as the non-centrality parameter. By comparing the expected mean squares in Model I and in a variance component model we note certain analogies between the ratio of variance components and the non-centrality parameter.

The non-centrality parameter in a simple one-way classification is

$$(3.9.1) \qquad \gamma^2 = \sum_{i=1}^{k} n_i(\mu_i - \bar{\mu})^2/\sigma^2$$

where μ_i is the mean of the i-th population and $\bar{\mu} = (\Sigma n_i \mu_i)/(\Sigma n_i)$. Here $\sigma^2\gamma^2$ is analogous to $(n - \sum_i n_i^2/n)\sigma_t^2$ (where σ_t^2 denotes the variance component due to treatments) because the expectation of the sum-of-squares between groups in Model I is $(k-1)\sigma^2 + \gamma^2\sigma^2$ as defined above; and the expectation of the sum of squares between groups in Model II is $(k-1)\sigma_e^2 + (n - \Sigma n_i^2/n)\sigma_t^2$. Similar analogies can be established for more complicated designs.

If all the n_i are equal (to r, say) we may obtain confidence bound on γ^2/r. If the n_i are unequal, we may consider the proportion, p_i, of the total number of observations contained in each class. Then

$$(3.9.2) \qquad \gamma^2 = n \sum_{i=1}^{k} p_i(\mu_i - \bar{\mu})^2/\sigma^2$$

and we might desire to put confidence limits on

$$(3.9.3) \qquad \gamma^2/n = \sum_i p_i(\mu_i - \bar{\mu})^2/\sigma^2 \ .$$

If desirable, (3.9.2) may be divided by (k - 1) and (3.9.3) may be multiplied by k/(k - 1) in order to obtain parameters which are formally analogous to the ratio σ_t^2/σ^2 in Model II. Each of these parameters is dimension free, i.e., does not depend upon the unit of measurement used in describing the original data.

In order to put confidence limits on γ^2 we use the improved variance-stabilizing transformation

$$(3.9.4) \qquad z = \cosh^{-1}(w/a)$$

which is approximately normal with

$$(3.9.5) \qquad E(z) = \zeta - \coth\zeta/(n - 4) \quad \text{and}$$

$$(3.9.6) \qquad \mathrm{var}(z) = 2/(n - 4), \quad \text{where}$$

$$(3.9.7) \qquad a = \sqrt{(m + n - 2)/(n - 2)}$$

$$(3.9.8) \qquad w = 1 + mF/n.$$

The condition $w > a$ implies, for $n > 2$, that $F > E(F)$; i.e., rejection of H_0 is sufficient to insure real values of Z.

(3.9.9) $\zeta = \cosh^{-1}[\gamma^2/\sqrt{(m + n - 2)(n - 2)} + a],$

where m and n are the degrees of freedom for F, the statistic obtained in testing the null hypothesis. Now since

(3.9.10) $[z - E(z)]/\sqrt{var\ (z)} \cong N(0, 1)$

we may say that

(3.9.11) $\sqrt{(n-4)/2}[z - \zeta + (\gamma^2/\sqrt{m+n-2} + \sqrt{m+n-2})$

$$\div (n-4)\sqrt{\gamma^4/(m+n-2)+m+2\gamma^2}] = N(0, 1).$$

Setting

(3.9.12) $\gamma^2/\sqrt{m+n-2} = v$ and

(3.9.13) $(v + \sqrt{m+n-2})/\sqrt{v^2+m+2\gamma^2} = u,$

we have the confidence statement

(3.9.14) $Pr\{\phi^{-1}(\alpha/2) < \sqrt{(n-4)/2}[z-\zeta+u/(n-4)] < \phi^{-1}(1-\alpha/2)\}=1-\alpha$

which may be written

(3.9.15) $Pr\{z - \sqrt{2/(n-4)}\phi^{-1}(1 - \alpha/2) < \zeta - u/(n-4)$

$$< z + \sqrt{2/(n-4)}\phi^{-1}(1 - \alpha/2)\} = 1 - \alpha.$$

As an approximation to the confidence bounds on ζ we may use

(3.9.16) $Pr\{z_L < \zeta < z_U\} \cong 1 - \alpha$ where

(3.9.17) $z_L = z - \sqrt{2/(n-4)}\ \phi^{-1}(1 - \alpha/2)$ and

(3.9.18) $z_U = z + \sqrt{2/(n-4)}\ \phi^{-1}(1 - \alpha/2)$.

Setting $\sqrt{(m+n-2)(n-2)} = k$, from (3.9.9) we have

(3.9.19) $\cosh \zeta = \gamma^2/k + a$.

Then we may form the equations

(3.9.20) $\cosh z_L = \gamma_L^2/k + a$ and

(3.9.21) $\cosh z_U = \gamma_U^2/k + a$

where γ_L^2 and γ_U^2 are the approximate lower and upper limits, respectively, of the confidence bounds on γ^2. Equations (3.9.20) and (3.9.21) may be solved for γ_L^2 and γ_U^2, which yield first approximation values of the lower and upper limits for γ^2; $\gamma_{L_0}^2$ and $\gamma_{U_0}^2$, say.

Consider (3.9.15) again. The confidence statement

(3.9.22) $\Pr\{z_L + u/(n-4) < \zeta < z_U + u/(n-4)\} = 1 - \alpha$

may be written as

(3.9.23) $\Pr\{z_L' < \zeta < z_U'\} = 1 - \alpha$,

where z_L' is a function of the final γ_L^2 and z_U' is a function of the final γ_U^2. Given a first approximate value of γ_L^2 ($\gamma_{L_0}^2$ above) we may obtain a first guess for z_L' (z_{L_0}', say) and improve it by an iterative technique until we find a value of γ_L^2 which satisfies (3.9.22).

The derivation of the improved estimates for the lower limit, γ_L^2, will be given. An analogous procedure is used to derive the improved estimates for the upper limit. From (3.9.20) we had

(3.9.24) $\cosh z_L = \gamma_L^2/k + a.$

But now, instead of z_L we have $z_L + u/(n-4)$, which gives us the improved lower confidence bound. So instead of (3.9.24) above, we have

(3.9.25) $\cosh [z_L + u/(n-4)] = \gamma_L^2/k + a$

which yields

(3.9.26) $f(\gamma_L^2) = \gamma_L^2/k + a - \cosh[z_L + u/(n-4)] = 0.$

Using the Newton method, successively better approximations to the root of this equation are given by the formula

(3.9.27) $\gamma^2_{L_{n+1}} = \gamma^2_{L_n} - f({}^{2}_{L_n})/f'(\gamma^2_{L_n})$. $n = 0, 1, 2, \ldots$

Differentiating $f(\gamma_L^2)$ with respect to γ_L^2 we have

(3.9.28) $\partial f/\partial\gamma_L^2 = 1/k - \partial/\partial\gamma_L^2 \cosh[z_L + u/(n-4)]$

$$= 1/k - \sinh[z_L + u/(n-4)]\partial/\partial\gamma_L^2[z_L + u/(n-4)]$$

$$= 1/k - \{\sinh[z_L + u/(n-4)]\partial u/\partial v \cdot \partial v/\partial\gamma_L^2\}/(n-4)$$

where, for simplicity of notation, u and v are written without subscripts but are understood to be u and v as defined in (3.9.12) and (3.9.13) with γ^2 replaced by $\gamma^2_{L_n}$. Then, using the fact that $\gamma^2_L = v\sqrt{m+n-2}$, we have

$$(3.9.29)\quad \partial u/\partial v = \partial/\partial v[(v + \sqrt{m+n-2})/(v^2 + m + 2v\sqrt{m+n-2})^{\frac{1}{2}}]$$

$$(3.9.30)\quad \partial u/\partial v = -(n-2)/(v^2 + m + 2v\sqrt{m+n-2})^{3/2} .$$

Also

$$(3.9.31)\quad \partial v/\partial\gamma^2_L = 1/\sqrt{m+n-2} ,$$

$$(3.9.32)\quad \partial u/\partial v \cdot \partial v/\partial\gamma^2_L = -(n-2)/[(v^2+m+2v\sqrt{m+n-2})^{3/2}\sqrt{m+n-2}],$$

$$(3.9.33)\quad \partial f/\partial\gamma^2_L = 1/k + (n-2)\sinh[z_L + u/(n-4)]$$

$$\div [(n-4)\sqrt{m+n-2}\,(v^2 + m + 2v\sqrt{m+n-2})^{3/2}],$$

and, from (3.9.2) we have

$$(3.9.34)\quad \gamma^2_{L_{n+1}} = \gamma^2_{L_n} - \{\gamma^2_L/k + a - \cosh[z_L + u/(n-4)]\}$$

$$\div \{\tfrac{1}{k} + (n-2)\sinh[z_L+u/(n-4)] /[(n-4)\sqrt{m+n-2}(v^2+m+2v\sqrt{m+n-2})^{\frac{3}{2}}]\}.$$

Equation (3.9.34) yields closer approximations to the lower bound of the confidence interval for γ^2. As previously stated, the improved estimates of γ^2_U are found in an exactly analogous manner, and to find $\gamma^2_{U_{n+1}}$ simply replace $\gamma^2_{L_n}$ by $\gamma^2_{U_n}$ in (3.9.34). Note that u and v will now be defined in terms of $\gamma^2_{U_n}$.

Since $a < b < c$ implies that $\cosh a < \cosh b < \cosh c$ only if $a > 0$, $b > 0$, and $c > 0$, our confidence bounds are valid only if $z_L + u/(n-4) > 0$ and $z_U + u/(n-4) > 0$.

After finding the lower and upper limits of γ^2 to the desired accuracy, we have the confidence statement

$$(3.9.35) \quad \Pr\{\gamma_L^2 < \gamma^2 < \gamma_U^2\} = 1 - \alpha.$$

If the number in each class is the same, say r, the confidence statement for $\sum_i (\mu_i - \mu)^2/\sigma^2$ is

$$(3.9.36) \quad \Pr\{\gamma_L^2/r < \sum_i (\mu_i - \mu)^2/\sigma^2 < \gamma_U^2/r\} = 1 - \alpha.$$

If the number in each class is not the same, the confidence statement for $\sum_i p_i(\mu_i - \mu)/\sigma^2$ is

$$(3.9.37) \quad \Pr\{\gamma_L^2/\Sigma n_i < \sum_i p_i(\mu_i - \mu)^2/\sigma^2 < \gamma_U^2/\Sigma n_i\} = 1 - \alpha,$$

where p_i represents the proportion of the total sample contained in the i-th group.

Reference:

Bargmann (1958) [1].

3.10 <u>Non-centrality parameter of χ^2</u>.

Suppose we perform a χ^2 goodness-of-fit test and find that $\sum_i (O_i - E_i)^2/E_i = u$. If the null hypothesis is true, the statistic has approximately a central χ^2 distribution.

For any given alternative model we can construct another set of "expected values" which we will call "postulated values under the alternative." If such an alternative is true, the statistic u defined above will have approximately a non-central χ^2 distribution with non-centrality parameter

$$(3.10.1) \quad \gamma^2 = \sum_i (P_i - E_i)^2/E_i$$

where the P_i are the "postulated values."

We would like to put confidence limits on γ^2. If a statistic u has the non-central χ^2 distribution with ν degrees of freedom and non-centrality parameter γ^2, then

$$(3.10.2) \quad \sqrt{u - \nu/2} - \sqrt{\gamma^2 + \nu/2} + 1/2\sqrt{\gamma^2 + \nu/2} \cong N(0, 1).$$

If we set

$$(3.10.3) \quad \sqrt{u - \nu/2} = x \quad \text{and}$$

$$(3.10.4) \quad \sqrt{\gamma^2 + \nu/2} = \delta$$

the confidence statement on the quantity in (3.10.2) is

$$(3.10.5) \quad \Pr\{\phi^{-1}(\alpha/2) < x - \delta + 1/2\delta < \phi^{-1}(1 - \alpha/2)\} = 1 - \alpha.$$

To assure that x is real, u must be larger than $\nu/2$. Since $E(u - \nu/2) = \nu/2$ under H_0, and $E(u - \nu/2) > \nu/2$ under any alternative, rejection of the null hypothesis is sufficient to assure that x is real.

Replacing $\emptyset^{-1}(1 - \alpha/2)$ by $\emptyset$ and $\emptyset^{-1}(\alpha/2)$ by $-\emptyset$, our confidence interval becomes

(3.10.6) $[-\emptyset < x - \delta + 1/2\delta < \emptyset]$

which is equivalent to

(3.10.7) $[x - \emptyset < \delta - 1/2\delta < x + \emptyset]$.

Dealing with only the lower part of this inequality we have

(3.10.8) $\delta - 1/2\delta > x - \emptyset$,

and, after multiplying both sides by δ and completing the square on δ, we get

(3.10.9) $[\delta - (x-\emptyset)/2]^2 > 1/2 + (x-\emptyset)^2/4$.

Using the fact that $a^2 > k$ implies that either $a > \sqrt{k}$ or $a < -\sqrt{k}$, we get from (3.10.9) the two inequalities

(3.10.10) $\delta > (x-\emptyset)/2 + \frac{1}{2}\sqrt{2 + (x-\emptyset)^2}$ and

(3.10.11) $\delta < (x-\emptyset)/2 - \frac{1}{2}\sqrt{2 + (x-\emptyset)^2}$.

Now, dealing with the upper part of (3.10.7), after multiplying both sides by δ and completing the square in δ, we get

(3.10.12) $[\delta - (x+\emptyset)/2]^2 < \frac{1}{2} + (x+\emptyset)^2/4$.

Since $a^2 < k$ implies that $-\sqrt{k} < a < \sqrt{k}$, (3.10.12) implies that

(3.10.13) $(x+\emptyset)/2 - \frac{1}{2}\sqrt{2+ (x+\emptyset)^2} < \delta < (x+\emptyset)/2 + \frac{1}{2}\sqrt{2 + (x+\emptyset)^2}$.

The confidence interval for δ is thus the intersection of the interval (3.10.13) with the union of the intervals (3.10.10) and (3.10.11). Since $\delta < 0$ is meaningless, we must insure that the intersection of (3.10.11) and (3.10.13) is empty, for both the lower bound of (3.10.13) and the interval (3.10.11) lie below zero. Thus, we must insure that

(3.10.14) $$(x-\phi)/2 - \tfrac{1}{2}\sqrt{2 + (x-\phi)^2} \leq (x+\phi)/2 - \tfrac{1}{2}\sqrt{2 + (x+\phi)^2} .$$

For this inequality to hold it is sufficient that $x \geq \phi$ and $\phi \geq 0$. Since for $\gamma^2 = 0$, $E(x) \geq 0$, rejection of the hypothesis $\gamma^2 = 0$ at level α is sufficient to insure that the lower intersection is empty. Hence, the desired confidence interval is then the intersection of (3.10.10) and (3.10.13); i.e.,

(3.10.15) $$(x-\phi)/2 + \tfrac{1}{2}\sqrt{2 + (x-\phi)^2} < \delta < (x+\phi)/2 + \tfrac{1}{2}\sqrt{2 + (x+\phi)^2},$$

which covers δ with probability $1 - \alpha$. Since both bounding quantities are positive it follows that

(3.10.16) $$\Pr\{[(x-\phi)/2 + \tfrac{1}{2}\sqrt{2 + (x-\phi)^2}]^2 < \delta^2 < [(x+\phi)/2 + \tfrac{1}{2}\sqrt{2 + (x+\phi)^2}]^2\} = 1 - \alpha.$$

Replacing δ^2 by $\gamma^2 + \nu/2$, our final probability statement is

$$(3.10.17)\quad \Pr\{\tfrac{1}{4}[x - \phi + \sqrt{2 + (x-\phi)^2}]^2 - \nu/2 < \gamma^2 < \tfrac{1}{4}[x + \phi + \sqrt{2 + (x+\phi)^2}]^2 - \nu/2\} = 1 - \alpha$$

which gives the confidence bounds on the non-centrality parameter of the non-central χ^2 distribution. It may appear as if the sample size is of no consequence in the above confidence statement. This is due to the fact that the sample size is implicit in γ^2 as defined above. For example, in a one-way classification analysis of variance with r replications per treatment

$$(3.10.18)\quad \gamma^2 = r \sum_i (\mu_i - \mu)^2/\sigma^2 .$$

If we want confidence bounds on

$$(3.10.19)\quad \gamma^2/r = \sum_i(\mu_i - \mu)^2/\sigma^2$$

we need only divide the left and right-hand sides by r, the number of replications.

A similar situation holds in the goodness-of-fit tests. The non-centrality parameter is

$$(3.10.20)\quad \gamma^2 = \sum_i (P_i - E_i)^2/E_i ,$$

where P_i is the <u>number</u> of items in the i-th category under the alternative hypothesis and E_i is the expected <u>number</u> of items in the i-th category under H_0. If we prefer confidence bounds on a parameter consisting of <u>proportions</u>, we would have to put bounds on γ^2/N, where N is the total number of items, since P_i and E_i represent N x(proportions). That is, since $P_i = Np_i$ and $E_i = Ne_i$, then

$$(3.10.21) \quad \gamma^2/N = 1/N \sum_i (Np_i - Ne_i)^2/Ne_i$$

$$= \sum_i (p_i - e_i)^2/e_i \; .$$

Reference:

Hofer (1960) [5]

IV. NUMERICAL EXAMPLES

This chapter is composed of computational procedures with worked examples that illustrate the estimation of confidence intervals presented in Chapter Three. Due to the simplicity of the methods no detailed computational procedures are given for the parameters presented in the first four sections of Chapter Three. All data used in the examples are artificial.

4.1 Parameter μ; σ^2 Known.

Confidence Statement:

$$\Pr\{\bar{x} - \phi^{-1}\sigma/\sqrt{N} < \mu < \bar{x} + \phi^{-1}\sigma/\sqrt{N}\} = 1 - \alpha$$

Data: $x_i = 2, 5, 7, 4, 7$; $\sigma^2 = 4$; $\alpha = .05$

The 95% confidence interval is thus

$$[5 - (1.96)2/\sqrt{5} < \mu < 5 + (1.96)2/\sqrt{5}], \text{ or}$$

$$[3.25 < \mu < 6.75] .$$

Parameter μ, σ^2 Unknown.

Confidence Statement:

$$\Pr\{\bar{x} - t\sqrt{S_{xx}/N(N-1)} < \mu < \bar{x} + t\sqrt{S_{xx}/N(N-1)}\} = 1 - \alpha .$$

Data: $x_i = 11, 7, 13, 9$; $\alpha = .05$

The 95% confidence interval is thus

$$[10 - (3.182)\sqrt{1.67} < \mu < 10 + (3.182)\sqrt{1.67}];$$

$$[5.89 < \mu < 14.11] .$$

4.2 Parameter σ^2; μ Known.

Confidence Statement:

$$\Pr\{[\Sigma(x_i-\mu)^2]/\chi^2_{1-\alpha/2} < \sigma^2 < [\Sigma(x_i-\mu)^2]/\chi^2_{\alpha/2}\} = 1-\alpha.$$

Data: $x_i = 6, 8, 9, 5, 8, 6$; $\mu = 7$; $\alpha = .10$

The 90% confidence interval is thus

$$[12/12.59 < \sigma^2 < 12/1.64];$$

$$[.95 < \sigma^2 < 7.32].$$

Parameter σ^2, μ Unknown.

Confidence Statement: $\Pr\{S_{xx}/\chi^2_{1-\alpha/2} < \sigma^2 < S_{xx}/\chi^2_{\alpha/2}\} = 1 - \alpha.$

Data: $x_i = 11, 12, 13, 12, 10, 14$; $\alpha = .05$

The 90% confidence interval is thus

$$[10/12.83 < \sigma^2 < 10/.831];$$

$$[.78 < \sigma^2 < 12.03].$$

4.3 Parameter $\mu_1 - \mu_2$; Common σ^2 Known.

Confidence Statement:

$$\Pr\{(\bar{x}_1 - \bar{x}_2) - \emptyset^{-1}(1 - \alpha/2)\sigma\sqrt{1/N_1 + 1/N_2} < \mu_1 - \mu_2$$

$$< (\bar{x}_1 - \bar{x}_2) + \emptyset^{-1}(1 - \alpha/2)\sigma\sqrt{1/N_1 + 1/N_2}\} = 1 - \alpha$$

Data: $x_{1i} = 5,3,4,7,5,6$; $x_{2i} = 3,6,4,3,4$; $\sigma^2 = 1$, $\alpha = .05$.

The 95% confidence interval is thus

$$[1 - (1.96)\sqrt{1/6 + 1/5} < \mu_1-\mu_2 < 1 + (1.96)\sqrt{1/6 + 1/5}];$$

$$[-.19 < \mu_1 - \mu_2 < 2.19] .$$

Parameter $\mu_1 - \mu_2$; Common σ^2 Unknown.

Confidence Statement:

$$\Pr\{(\bar{x}_1-\bar{x}_2)-s_p t\sqrt{1/N_1+1/N_2} < \mu_1-\mu_2 < (\bar{x}_1-\bar{x}_2)+s_p t\sqrt{1/N_1+1/N_2}\} = 1-\alpha$$

Data: $x_{1i} = 5, 3, 4, 7, 5, 6$; $x_{2i} = 3, 6, 4, 3, 4$; $\alpha = .01$

The 99% confidence interval is thus

$$[1 - (3.25)(4/3)(.61) < \mu_1 - \mu_2 < (3.25)(4/3)(.61)];$$

$$[-1.64 < \mu_1 - \mu_2 < 3.64] .$$

4.4 Parameter σ_1^2/σ_2^2.

Confidence Statement:

$$\Pr\{(s_1^2/s_2^2)/F_{1-\alpha/2} < \sigma_1^2/\sigma_2^2 < (s_1^2/s_2^2)/F_{\alpha/2}\} = 1 - \alpha.$$

Data: $x_{1i} = 1.3, 8.6, 4.2, 3.8, 5.1, 6.0, 3.9, 4.5, 4.0, 7.1$;

$x_{2i} = 7.0, 5.4, 3.8, 4.2, 4.6, 3.9, 5.8, 2.6$;

$\alpha = .10$

The 90% confidence interval is thus

$$[(4.0428/1.8713)/3.68 < \frac{\sigma_1^2}{\sigma_2^2} < (4.0428/1.8713)/.304];$$

$$[.59 < \sigma_1^2/\sigma_2^2 < 7.11] .$$

4.5 Parameter ρ.

Data: x_i = 10.2, 10.1, 9.8, 10.1, 9.2, 8.6, 7.9, 8.7, 10.3, 10.4;

y_i = 4.0, 4.1, 4.0, 4.0, 4.2, 3.1, 1.9, 2.0, 3.0, 3.7;

$\alpha = .10$

Computational Procedure:

1. Calculate Σx, Σy, Σx^2, Σy^2, and Σxy from the data.
2. Calculate $r = [N\Sigma xy - \Sigma x\Sigma y]/\sqrt{[N\Sigma x^2 - (\Sigma x)^2][N\Sigma y^2 - (\Sigma y)^2]}$
3. Find $z = \tanh^{-1} r = \frac{1}{2} \log(1+r)/(1-r)$
4. Calculate $1/\sqrt{N - 3}$
5. Calculate $\phi^{-1}(1 - \alpha/2)/\sqrt{N - 3} = c$
6. Calculate $z - c$ and $z + c$
7. Find $\rho_L = \tanh(z - c)$
8. Find $\rho_U = \tanh(z + c)$
9. $\Pr\{\rho_L < \rho < \rho_U\} = 1 - \alpha$

Computation:

1. $\Sigma x = 95.3$, $\Sigma y = 34.0$, $\Sigma x^2 = 915.05$, $\Sigma y^2 = 122.36$,
   $\Sigma xy = 328.90$
2. $r = .718$
3. $z = .904$
4. $1/\sqrt{N - 3} = .378$
5. $c = .622$
6. $z - c = .282$, $z + c = 1.526$
7. $\rho_L = \tanh(.282) = .275$

8. $\rho_U = \tanh(1.526) = .910.$

The 90% confidence interval is thus

$$[.275 < \rho < .910] .$$

4.6 Parameter μ_1/μ_2.

Data: $x_{1i} = 5.0, 6.9, 7.0, 5.2, 8.2, 7.3, 5.1, 4.6, 6.3, 6.7;$

$x_{2i} = 6.1, 4.3, 4.8, 5.0, 3.7, 4.9, 5.1, 4.4, 4.8, 4.9,$
$5.4, 3.5;$

$\alpha = .05$

Computational Procedure:

1. Look up $F(1, N_1 + N_2 - 2)$, for the chosen α level.
2. Compute s_p^2 from the two samples.
3. Compute $\bar{x}_1$ and $\overline{x_1^2}$ from the data.
4. Compute $\bar{x}_2$ and $\overline{x_2^2}$ from the data.
5. Compute $\bar{x}_1\bar{x}_2$.
6. Compute $b_1 = Fs_p^2/N_1$.
7. Compute $b_2 = Fs_p^2/N_2$.
8. Compute $d = \overline{x_2^2} - b_2$ and note whether $d > 0$ or not. If $\overline{x_2^2} - b_2 < 0$, valid confidence bounds cannot be put on $\gamma = \mu_1/\mu_2$.
9. Compute b_1b_2.
10. Compute $b_2\overline{x_1^2}$ and $b_1\overline{x_2^2}$.

11. Calculate $c = \sqrt{b_2\bar{x}_1^2 + b_1\bar{x}_2^2 - b_1b_2}$

12. Compute $(\bar{x}_1\bar{x}_2 - c)/d$ and $(\bar{x}_1\bar{x}_2 + c)/d$.

13. $\Pr\{(\bar{x}_1\bar{x}_2 - c)/d < \gamma < (\bar{x}_1\bar{x}_2 + c)/d\} = 1 - \alpha$

If σ^2 is known or may be assumed known, Step 2 is omitted, s_p^2 is replaced by the known σ^2, and F is replaced by χ_1^2.

Computation:

1. $\alpha = .05$, $F_{.05}(1,20) = 4.35$
2. $s_p^2 = .9135$
3. $\bar{x}_1 = 6.230$, $\bar{x}_1^2 = 38.8129$
4. $\bar{x}_2 = 4.7417$, $\bar{x}_2^2 = 22.4837$
5. $\bar{x}_1\bar{x}_2 = 29.5408$
6. $b_1 = .3974$
7. $b_2 = .3311$
8. $d = 22.1526 > 0$
9. $b_1b_2 = .1316$
10. $b_2\bar{x}_1^2 = 12.8510$, $b_1\bar{x}_2^2 = 8.9350$
11. $c = 4.6534$
12. $(\bar{x}_1\bar{x}_2 - c)/d = 1.1235$, $(\bar{x}_1\bar{x}_2 + c)/d = 1.5436$
13. The 95% confidence interval is thus $[1.12 < \gamma < 1.54]$.

4.7. <u>Parameter</u> $\underline{\sigma_x^2/\sigma_y^2}$ (<u>x and y are correlated</u>).

Data: $x_i = 28, 18, 22, 27, 25, 30, 21, 20, 27, 21;$

$y_i = 19, 38, 42, 25, 15, 31, 22, 37, 30, 24;$ $\alpha = .10$

Computational Procedure:

1. Compute Σx, Σy, Σx^2, Σy^2, and Σxy from the data.

2. Compute $r^2 = [N\Sigma xy - \Sigma x \Sigma y]^2/[N\Sigma x^2 - (\Sigma x)^2][N\Sigma y^2 - (\Sigma y)^2]$

3. Compute $\sqrt{1 - r^2}$

4. Find t for the upper (positive) $\alpha/2$ level with N - 2 degrees of freedom and form $a = t\sqrt{1 - r^2}$

5. Compute $\sqrt{N - 2 + a^2}$

6. Compute $b_L = a - \sqrt{N - 2 + a^2}$ and $b_U = a + \sqrt{N - 2 + a^2}$

7. Compute $k = s_x^2/s_y^2$

8. Compute $2ak/(N - 2)$

9. Compute $C_L = 2akb_L/(N-2)$ and $C_U = 2akb_U/(N-2)$

10. Compute $k + C_L$ and $k + C_U$

11. $\Pr\{k + C_L < \sigma_x^2/\sigma_y^2 < k + C_U\} = 1 - \alpha$

Computation:

1. $\Sigma x = 239$, $\Sigma y = 283$, $\Sigma x^2 = 5857$, $\Sigma y^2 = 8709$, $\Sigma xy = 6636$

2. $r^2 = .1607$

3. $\sqrt{1-r^2} = .9161$

4. $t_{.95} = 1.86$, $a = 1.7039$

5. $\sqrt{N-2+a^2} = 3.3020$

6. $b_L = -1.5981$, $b_U = 5.0059$

7. $k = .2070$

8. $2ak/(N-2) = .0882$

9. $C_L = -.1409$, $C_U = .4414$

10. $k + C_L = .0661$, $k + C_U = .6484$

11. The 90% confidence interval is thus $[.07 < \sigma_x^2/\sigma_y^2 < .65]$.

4.8. Confidence Bounds on the Non-Centrality Parameter in the Non-Central t^2 Distribution.

Computational Procedure:

Steps 1 through 8 apply to both upper and lower bounds, but only the formulas appropriate for the lower bound are shown in steps 9 through 18.

1. Compute $\bar{x}$ and s^2 from the data.
2. Compute $t^2 = F = N\bar{x}^2/s^2$
3. Compute $w = 1 + F/(N-1)$
4. Compute $a = \sqrt{(N-2)/(N-3)}$
5. Compute w/a
6. Find $z = \cosh^{-1}(w/a)$
7. For the chosen α compute $c = \Phi^{-1}(1-\alpha/2)\sqrt{2/(N-5)}$ and form $z_L = z-c$ and $z_U = z+c$
8. Look up $\cosh(z_L)$ and $\cosh(z_U)$
9. Find $\sqrt{(N-2)(N-3)} = k$ and compute $\gamma^2_{L_0} = k[\cosh(z_L) - a]$
10. Compute $b = \gamma^2_{L_0}/k$
11. Compute $v = \gamma^2_{L_0}/\sqrt{N-2}$ (Record $\sqrt{N-2}$)
12. Compute $s = \sqrt{v^2 + 1 + 2\gamma^2_{L_0}}$ and s^3
13. Compute $u = (v + \sqrt{N-2})/s$
14. Compute $y = z_L + u/(N-5)$ and look up $\sinh(y)$ and $\cosh(y)$
15. Compute $d = a + b - \cosh(y)$
16. Compute $e = [(N-3)\sinh(y)]/[(N-5)\sqrt{N-2}\, s^3]$
17. Compute $f = d/(1/k + e)$

18. Compute $\gamma^2_{L_1} = \gamma^2_{L_0} - f$

Steps 10 through 18 are reiterated using each new value $\gamma^2_{L_{i+1}}$, $i = 0,1,2,\ldots$, in place of $\gamma^2_{L_0}$ until the desired accuracy for γ^2_L is obtained (usually until the quantity d in step 15 equals zero). When finding the upper bound of γ^2 use z_U instead of z_L in step 9 and $\gamma^2_{U_i}$ instead of $\gamma^2_{L_i}$ $(i=0,1,2,\ldots)$ in steps 10 through 18.

Data: 2.2, 3.1, 1.8, 1.0, 4.1, 3.5, 2.9, 2.2, 1.1, 3.2, 2.5

Computation:

Lower Limit. Iteration 1

1. $\bar{x} = 2.5091$, $s^2 = .9449$
2. $F = 73.2896$
3. $w = 8.3290$
4. $a = 1.0607$
5. $\frac{w}{a} = 7.8524$
6. $z = 2.7499$
7. $\alpha = .10$, $\Phi^{-1}(.95) = 1.6449$, $c = .9497$, $z_L = 1.8002$,
   $z_U = 3.6996$
8. $\cosh(z_L) = 3.1081$, $\cosh(z_U) = 20.2283$
9. $k = 8.4853$, $\gamma^2_{L_0} = 17.3728$
10. $b = 2.0474$
11. $v = 5.7909$ $(\sqrt{N-2} = 3)$
12. $s = 8.3235$, $s^3 = 576.65$
13. $u = 1.0562$

14. $y = 1.9762$, $\sinh(y) = 3.5384$, $\cosh(y) = 3.6770$

15. $d = -.5689$

16. $e = .0027$

17. $f = -4.7172$

18. $\gamma^2_{L_1} = 22.0900$

Iteration 2.

10. $b = 2.6033$

11. $v = 7.3633$

12. $s = 9.9699$, $s^3 = 990.99$

13. $u = 1.0395$

14. $y = 1.9735$, $\sinh(y) = 3.5285$, $\cosh(y) = 3.6674$

15. $d = -.0034$

16. $e = .0016$

17. $f = -.0285$

18. $\gamma^2_{L_2} = 22.1185$

Iteration 3.

10. $b = 2.6067$

11. $v = 7.3728$

12. $s = 9.9797$, $s^3 = 993.93$

13. $u = 1.0394$

14. $y = 1.9734$, $\sinh(y) = 3.5281$, $\cosh(y) = 3.6671$

15. $d = .0003$

16. $e = .0016$

17. $f = .0025$

18. $\gamma^2_{L_3} = 22.1160$

Iteration 4.

10. $b = 2.6064$

11. $v = 7.3720$

12. $s = 9.9789$, $s^3 = 993.68$

13. $u = 1.0394$

14. $y = 1.9734$, $\sinh(y) = 3.5281$, $\cosh(y) = 3.6671$

15. $d = 0$

Cease iterating. $\gamma^2_{L_4} = 22.1160$

Upper limit. Iteration 1.

9. $\gamma^2_{U_0} = 162.6428$

10. $b = 19.1676$

11. $v = 54.2143$

12. $s = 57.1443$, $s^3 = 186,603$

13. $u = 1.0012$

14. $y = 3.8665$, $\sinh(y) = 23.9033$, $\cosh(y) = 23.9242$

15. $d = -3.6959$

16. $e = .0001$

17. $f = -31.3212$

18. $\gamma^2_{U_1} = 193.9640$

Iteration 2.

10. $b = 22.8588$

11. $v = 64.6547$

12. $s = 67.5955$, $s^3 = 308,855$

13. $u = 1.0009$

14. $y = 3.8664$, $\sinh(y) = 23.9009$, $\cosh(y) = 23.9219$

15. $d = -.0024$

16. $e = .0000$

17. $f = -.0204$

18. $\gamma^2_{U_2} = 193.9844$

Iteration 3.

10. $b = 22.8612$

11. $v = 64.6615$

12. $s = 67.6024$, $s^3 = 308,948$

13. $u = 1.0009$

14. $y = 3.8664$, $\sinh(y) = 23.9009$, $\cosh(y) = 23.9219$

15. $d = 0$

Cease iterating. $\gamma^2_{U_3} = 193.9844$

The 90% confidence interval is thus $[22.12 < \gamma^2 < 193.98]$; or $[2.011 < \mu^2/\sigma^2 < 17.635]$.

4.9. Confidence Bounds on the Non-Centrality Parameter in the Non-Central F Distribution.

Steps 1 through 8 apply to both upper and lower bounds, but only the formulas appropriate for the lower bound are shown in steps 9 through 18.

Computational Procedure:

1. Compute F from the data.

2. Compute $w = mF/n + 1$

3. Compute $a = \sqrt{(m + n - 2)/(n - 2)}$

4. Compute w/a

5. Find $z = \cosh^{-1}(w/a)$

6. Find $\Phi^{-1}(1-\alpha/2)$ for the chosen α and compute
   $c = \Phi^{-1}(1-\alpha/2)\ \sqrt{2/(n-4)}$

7. Compute $z_L = z - c$ and $z_U = z + c$ and look up $\cosh(z_L)$ and $\cosh(z_U)$

8. Compute $k = \sqrt{(m+n-2)(n-2)}$ and $1/k$

9. Compute $\gamma^2_{L_0} = k[\cosh(z_L) - a]$

10. Compute $b = \gamma^2_{L_0}/k$

11. Compute $v = \gamma^2_{L_0} / \sqrt{m+n-2}$; (Record $\sqrt{m+n-2}$)

12. Compute $s = \sqrt{v^2 + m + 2\gamma^2_{L_0}}$ and s^3

13. Compute $u = (v + \sqrt{m + n - 2})/s$

14. Compute $y = z_L + u/(n-4)$ and look up $\sinh(y)$ and $\cosh(y)$

15. Compute $d = a + b - \cosh(y)$

16. Compute $e = [(n-2)\sinh(y)]/[(n-4)\sqrt{m + n-2}\ s^3]$

17. Compute $f = d/(1/k + e)$

18. Compute $\gamma^2_{L_1} = \gamma^2_{L_0} - f$

Steps 10 through 18 are reiterated using each new value $\gamma^2_{L_{i+1}}$ $(i=0,1,2,\ldots)$ in place of $\gamma^2_{L_0}$ until the desired accuracy for γ^2_L is obtained (usually until the quantity d in step 15 equals zero). When finding the upper bound of γ^2 use z_U instead of z_L in step 9 to find $\gamma^2_{U_0}$; and use $\gamma^2_{U_i}$ instead of $\gamma^2_{L_i}$ $(i=0,1,2,\ldots)$ in steps 10 through 18.

Data (5 groups, 11 observations per group):

	(1)	(2)	(3)	(4)	(5)	
	4	1	6	6	4	
	1	7	7	1	5	
	2	6	9	10	4	
	2	2	10	3	3	
	3	6	8	7	6	
	1	5	7	4	5	
	2	4	5	9	3	
	1	1	10	6	6	
	4	5	8	1	5	
	3	4	9	4	4	
	1	8	7	2	4	
Σ	24	49	86	53	49	$\Sigma\Sigma = 261$

Analysis of Variance.

Source	SS	d.f.	MS
Between	178.0728	4	44.5182
Within	198.3636	50	3.9673
Total	376.4364	54	

Computation:

Lower limit. Iteration 1.

1. $F = 11.2213$

2. $w = 1.8977$

3. $a = 1.0408$

4. $w/a = 1.8233$

5. $z = 1.2083$

6. $\alpha = .10$, $\phi^{-1}(.95) = 1.6449$, $c = .3430$

7. $z_L = .8653$, $z_U = 1.5513$

   $\cosh(z_L) = 1.3983$, $\cosh(z_U) = 2.4648$

8. $k = 49.96$, $1/k = .0200$

9. $\gamma^2_{L_0} = 17.8607$

10. $b = .3575$

11. $v = 2.4768$, $\sqrt{m + n - 2} = 7.2111$

12. $s = 6.7717$, $s^3 = 310.52$

13. $u = 1.4306$

14. $y = .8964$, $\sinh(y) = 1.0214$, $\cosh(y) = 1.4294$

15. $d = -.0311$

16. $e = .0005$

17. $f = -1.5171$

18. $\gamma^2_{L_1} = 19.3778$

Iteration 2.

10. $b = .3879$

11. $v = 2.6872$

12. $s = 7.0694$, $s^3 = 353.31$

13. $u = 1.4002$

14. $y = .8957$, $\sinh(y) = 1.0204$, $\cosh(y) = 1.4287$

15. $d = 0$

Cease iterating. $\gamma^2_{L_2} = 19.3778$

Upper limit. Iteration 1.

9. $\gamma^2_{U_0} = 71.1430$

10. $b = 1.4240$

11. $v = 9.8658$, $\sqrt{m + n - 2} = 7.2111$

12. $s = 15.6083$, $s^3 = 3802.5$

13. $u = 1.0941$

14. $y = 1.5751$, $\sinh(y) = 2.3121$, $\cosh(y) = 2.5191$

15. $d = -.0543$

16. $e = .0001$

17. $f = -2.7015$

18. $\gamma^2_{U_1} = 73.8445$

Iteration 2.

10. $b = 1.4781$

11. $v = 10.2404$

12. $s = 16.0173$, $s^3 = 4109.3$

13. $u = 1.0895$

14. $y = 1.5750$, $\sinh(y) = 2.3119$, $\cosh(y) = 2.5189$

15. $d = 0$

Cease iterating. $\gamma^2_{U_2} = 73.8445$

The 90% confidence interval is thus $[19.38 < \gamma^2 < 73.84]$.

4.10. Confidence Bounds on the Non-Centrality Parameter in the Non-Central χ^2 Distribution.

Computational Procedure:

1. Record the statistic $u = \Sigma(O_i - E_i)^2/E_i$ which has been calculated from the sample.
2. Look up $\phi^{-1}(1 - \alpha/2) = \phi$ for the desired α- level.
3. Calculate $x = \sqrt{u - \nu/2}$
4. Compute $x - \phi$ and $(x-\phi)^2$; check that $\phi \geq 0$ and $x \geq \phi$.
5. Compute $x + \phi$ and $(x + \phi)^2$
6. Compute $a = \sqrt{2 + (x - \phi)^2}$
7. Compute $b = 1/4\ (x - \phi + a)^2$
8. Compute $b - \nu/2$
9. Compute $c = \sqrt{2 + (x + \phi)^2}$
10. Compute $d = 1/4\ (x + \phi + c)^2$
11. Find $d - \nu/2$
12. $\Pr\{b - \nu/2 < \gamma^2 < d - \nu/2\} = 1-\alpha$

Data: $O_i = 18,10,29,11,7,23,8,13,6$

$E_i = 13,15,23,18,13,17,13,8,5$

Computation:

1. $u = 18.012$
2. $\alpha = .05$, $\phi = 1.960$
3. $x = 3.7433$
4. $x-\phi = 1.7833$ $(x-\phi)^2 = 3.1802$ $(\phi \geq 0,\ x \geq \phi)$.
5. $x + \phi = 5.7033$ $(x + \phi)^2 = 32.5276$

6. $a = 2.2760$

7. $b = 4.1195$

8. $b - \nu/2 = 0.1195$

9. $c = 5.8760$

10. $d = 33.5201$

11. $d - \nu/2 = 29.5201$

The 95% confidence interval is thus $[.12 < \Upsilon^2 < 29.52]$.

Now, $E(u) = \Upsilon^2 + \nu$.

Hence $\hat{\Upsilon}^2 \cong 18 - 8 = 10$; so we see that the confidence interval is skewed to the right about the point estimate of Υ^2.

Example 2.

In this example confidence bounds for various values of α will be computed, given the

Data: $O_i = 7, 20, 9, 12, 6, 22, 19$

$E_i = 12, 18, 14, 10, 7, 11, 23$

1. $u = 16.3298$

2. $\alpha = .20, \quad \phi = 1.282$

3. $x = 3.6510$

4. $x - \phi = 2.3690, \quad (x - \phi)^2 = 5.6122 \qquad (\phi > 0,\ x > \phi)$

5. $x + \phi = 4.9330, \quad (x + \phi)^2 = 24.3345$

6. $a = 2.7590$

7. $b = 6.5741$

8. $b - \nu/2 = 3.5741$

9. $c = 5.1317$

10. $d = 25.3245$

11. $d - \nu/2 = 22.3245$

The 80% confidence interval is thus $[3.57 < \Upsilon^2 < 22.32]$.

1. $u = 16.3298$

2. $\alpha = .10, \quad \phi = 1.645$

3. $x = 3.6510$

4. $x - \phi = 2.0060 \quad (x - \phi)^2 = 4.0240 \quad (\phi > 0 \ , \ x > \phi)$

5. $x + \phi = 5.2960 \quad (x + \phi)^2 = 28.0476$

6. $a = 2.4544$

7. $b = 4.9738$

8. $b - \nu/2 = 1.9738$

9. $c = 5.4816$

10. $d = 29.0392$

11. $d - \nu/2 = 26.0392$

The 90 % confidence interval is thus $[1.97 < \Upsilon^2 < 26.04]$.

1. $u = 16.3298$

2. $\alpha = .05, \quad \phi = 1.960$

3. $x = 3.6510$

4. $x - \phi = 1.6910, \quad (x - \phi)^2 = 2.8595 \quad (\phi > 0 \ , \ x > \phi)$

5. $x + \phi = 5.6110, \quad (x + \phi)^2 = 31.4833$

6. $a = 2.2044$

7. $b = 3.7935$

8. $b - \nu/2 = .7935$

9. $c = 5.7865$

10. $d = 32.4758$

11. $d - \nu/2 = 29.4758$

The 95% confidence interval is thus

$$[.79 < \gamma^2 < 29.48].$$

V. ACKNOWLEDGEMENTS

The author would like to express his appreciation to Dr. Rolf E. Bargmann for his assistance and counsel during the writing of this thesis.

Also, appreciation is expressed to Dr. Boyd Harshbarger for the patience and solicitude with which he viewed the progress of this thesis, and to Dr. H. A. David, who kindly consented to read it in manuscript form.

Finally, the author wishes to thank and for typing the final copy.

VI. BIBLIOGRAPHY

[1] Bargmann, R. E., "Some Interpretations in the Analysis of Transformed Data", Virginia Polytechnic Institute Technical Report (1958).

[2] Bliss, C. I., "The Comparison of Dosage-Mortality Data", Annals of Applied Biology, vol. 22 (1935), p. 307.

[3] Fieller, E. C., "The Biological Standardization of Insulin", Journal of the Royal Statistical Society Supplement, vol. 7 (1940), p. 1.

[4] Hotelling, H., "New Light on the Correlation Coefficient and its Transforms", Journal of the Royal Statistical Society, Series B, vol. 15 (1953), p. 193.

[5] Hofer, G., "A Variance Stabilizing Transformation of the Non-Central X^2 - distribution", Unpublished Masters thesis, Virginia Polytechnic Institute (1960).

[6] Roy, S. N., Potthoff, R. F., "Confidence Bounds on Vector Analogues of the 'Ratio of Means' and the 'Ratio of Variances' for Two Correlated Normal Variates and some Associated Tests", Annals of Mathematical Statistics, vol. 29 (1958), p. 829.

The vita has been removed from the scanned document

ABSTRACT

In this thesis, confidence bounds on simple and more complex parameters are stated along with detailed computational procedures for finding these confidence bounds from the given data.

Confidence bounds on the more familiar parameters, i.e., μ, σ^2, $\mu_1 - \mu_2$, and σ_1^2/σ_2^2, are briefly presented for the sake of completeness. The confidence statements for the less familiar parameters and combinations of parameters are treated in more detail.

In the cases of the non-centrality parameters of the non-central t^2, F and χ^2 distributions, a variance-stabilizing transformation is used, a normal approximation is utilized, and confidence bounds are put on the parameter. In the non-central t^2 and non-central F distributions iterative procedures are used to obtain confidence bounds on the non-centrality parameter, i.e., a first guess is made which is improved until the desired accuracy is obtained. This procedure is unnecessary in the non-central χ^2 distribution, since the expressions for the upper and lower limits can be reduced to closed form.

Computational procedures and completely worked examples are included.