A HYPOTHESIS TESTING

by K. X. Karakostas
Department of Mathematics, University of Ioannina, Creece

1. - Introduction

Let $X_1, X_2, \cdots, X_n$ and $Y_1, Y_2, \cdots, Y_m$ be two independent and identically distributed (i.i.d) random samples (r.s) from normal populations with means μ_X and μ_Y and variances σ_X^2 and σ_Y^2, respectively. It is well known (see e.g Casella, Berger (1990)) that when the variances of the two populations are unknown, but equal, the test statistic for testing the hypothesis $H_o:\mu_X=\mu_Y$, is the following

$$t_{Old} = \frac{\overline{X}-\overline{Y}}{\sqrt{(n-1)S_X^2+(m-1)S_Y^2/(n+m-2)}\sqrt{(n+m)/(nm)}}, \tag{1}$$

where $\overline{X}$, $\overline{Y}$, S_X^2 and S_Y^2 are the sample means and variances of the two samples, respectively. Under the null hypothesis H_o the distribution of (1) is the central t with n+m-2 degrees of freedom. This test is both likelihood and a significance test. When the population variances are unknown and unequal the test statistic, for testing the same hypothesis, is,

$$t_{Welch} = \frac{\overline{X}-\overline{Y}}{\sqrt{\frac{S_x^2}{n}+\frac{S_y^2}{m}}} \tag{2}$$

The distribution of the above statistic, under the null hypothesis, is approximated by a central t distribution whose degrees of freedom are obtained approximately from various formulas. (See Cochran and Cox (1950), Satterthwaite (1946)).

Let's now think of a slightly different problem. More precisely we consider the previous hypothesis H_o, but this time we suppose that the variance σ_X^2 is known and σ_Y^2 is unknown. Problems of such a type can arise when we test the effectiveness of two methods and the precision of one of them is known, e.g. from previous experience.

2. -Testing the hypothesis H_o:μ_X=μ_Y with σ_X^2 known and σ_Y^2 unknown

In the sequel we will present three methods for testing the previously mentioned hypothesis. The first is based on the significance test; the second one is the well-known likelihood ratio technique, while the third is rather different. Without loss of generality, in what follows, we can assume that $\sigma_X^2=1$.

2.1. -Significance test approach

It is well known that $\bar{X} \sim N(\mu_X,\ 1/n)$ and $\bar{Y} \sim N(\mu_Y,\ \sigma_Y^2/m)$. Hence, under the null hypothesis H_o:μ_X=μ_Y $\bar{X}-\bar{Y} \sim N\left(0,\ \frac{1}{n}+\frac{\sigma_Y^2}{m}\right)$. Assuming that 1/n≈0 and using the fact that $\frac{1}{\sigma_Y^2}\sum_{j=1}^{N}(Y_j-\bar{Y})^2 \sim \aleph_{N-1}^2$ and $\bar{X}$, $\bar{Y}$ and S_Y^2 are independent we get that the statistic $t_{New}=\frac{\sqrt{m}(\bar{X}-\bar{Y})}{S_Y}$ follows the central t-distribution with m-1 degrees of freedom. That is

$$t_{New}=\frac{\sqrt{m}(\bar{X}-\bar{Y})}{S_Y} \sim t_{m-1} \tag{3}$$

The above test is an approximate test of significance, and as it will be come clear in the next section it cannot be derived using the likelihood approach.

2.2. - Likelihood ratio approach

Under H_o:μ_X=μ_Y(=μ) the likelihood function of the two samples is given by

$$L_o=\frac{1}{\left(\sqrt{\sigma_y^2}\right)^m\left(\sqrt{2\pi}\right)^{n+m}}\exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2\right\}\exp\left\{-\frac{1}{2\sigma_y^2}\sum_{j=1}^{m}(y_j-\mu)^2\right\}. \tag{4}$$

To find the maximum likelihood estimates (m.l.e), $\hat{\mu}$ and $\hat{\sigma}_{y,o}^2$, of μ and σ_Y^2, respectively, we take the natural logarithm of both sides and we get

$$\ln L_o=\ln\left(\frac{1}{\sqrt{2\pi}}\right)^{n+m}-\frac{m}{2}\ln\sigma_y^2-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)^2-\frac{1}{2\sigma_y^2}\sum_{j=1}^{m}(y_j-\mu)^2 .$$

Hence

$$\frac{\partial \ln L_o}{\partial \mu}=\sum_{i=1}^{n}(x_i-\mu)+\frac{1}{\sigma_y^2}\sum_{j=1}^{m}(y_j-\mu) \quad \text{and} \quad \frac{\partial \ln L_o}{\partial \sigma_y^2}=-\frac{m}{2\sigma_y^2}+\frac{1}{2\sigma_y^4}\sum_{j=1}^{m}(y_j-\mu)^2 .$$

By putting each of the above equations to zero we get the next system of equations

$$\left(m+n\sigma_y^2\right)\mu = m\overline{Y} + n\sigma_y^2\overline{X} \quad \kappa\alpha\iota \quad m\sigma_y^2 = \sum_{j=1}^{n}(y_j-\mu)^2 . \tag{5}$$

If we solve this system (For the solution see Appendix), with respect of μ and σ_Y^2, we get the m.l. e., $\hat{\mu}$ and $\hat{\sigma}_{Y,o}^2$ of μ and σ_Y^2, respectively..

Under H_α:$\mu_X \neq \mu_Y$ the likelihood function of the two samples is given by

$$L_\alpha = \frac{1}{\left(\sqrt{2\pi}\right)^{n+m}\left(\sqrt{\sigma_Y^2}\right)^m}\exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu_x)^2\right\}\exp\left\{-\frac{1}{2\sigma_y^2}\sum_{j=1}^{m}(y_j-\mu_y)^2\right\}. \tag{6}$$

Working as above we obtain that the maximum likelihood estimates of μ_X, μ_Y and σ_Y^2 are given by the formulae

$$\hat{\mu}_X = \overline{X},\ \hat{\mu}_Y = \overline{Y} \ \kappa\alpha\iota\ \hat{\sigma}_{Y,\alpha}^2 = \frac{1}{m}\sum_{j=1}^{m}\left(Y_j - \overline{Y}\right)^2 . \tag{7}$$

From (4), (6), (5) και (7) we have the maximum likelihood ratio test, λ, is given by

$$\lambda = \frac{\left(\sqrt{\hat{\sigma}_{y,o}^2}\right)^{-m}\exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\hat{\mu})^2\right\}\exp\left\{-\frac{1}{2\hat{\sigma}_{y,o}^2}\sum_{j=1}^{m}(y_j-\hat{\mu})^2\right\}}{\left(\sqrt{\hat{\sigma}_{y,\alpha}^2}\right)^{-m}\exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\overline{X})^2\right\}\exp\left\{-\frac{m}{2}\right\}} \leq k \ \text{ or}$$

$$\lambda = \left(\frac{\hat{\sigma}_{y,\alpha}^2}{\hat{\sigma}_{y,o}^2}\right)^{m/2}\exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\hat{\mu})^2\right\}\exp\left\{-\frac{1}{2\hat{\sigma}_{y,o}^2}\sum_{j=1}^{m}(y_j-\hat{\mu})^2\right\}\exp\left\{\frac{1}{2}\sum_{i=1}^{n}(x_i-\overline{X})^2\right\} \leq k^*$$

The term $-\frac{1}{2\hat{\sigma}_{y,o}^2}\sum_{j=1}^{m}(y_j-\hat{\mu})^2$ can be written as

$$-\frac{1}{2\hat{\sigma}_{y,o}^2}\sum_{j=1}^{m}(y_j-\hat{\mu})^2 + \frac{1}{2\hat{\sigma}_{y,o}^2}m\hat{\sigma}_{y,o}^2 - \frac{1}{2\hat{\sigma}_{y,o}^2}m\hat{\sigma}_{y,o}^2 = -\frac{1}{2\hat{\sigma}_{y,o}^2}\left\{\sum_{j=1}^{m}(y_j-\hat{\mu})^2 - m\hat{\sigma}_{y,o}^2\right\} - \frac{m}{2}.$$

From the second of the equations (5) we have that $\sum_{j=1}^{m}(y_j-\hat{\mu})^2 - m\hat{\sigma}_{y,o}^2 = 0$, and hence the maximum likelihood ratio takes the form

$$\lambda = \left(\frac{\hat{\sigma}_{y,\alpha}^2}{\hat{\sigma}_{y,o}^2}\right)^{m/2}\exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\left[(x_i-\hat{\mu})^2-(x_i-\overline{X})^2\right]\right\} \leq k^{**}.$$

But $\sum_{i=1}^{n}\left[(x_i-\hat{\mu})^2-(x_i-\overline{X})^2\right] = \sum_{i=1}^{n}\left(x_i-\hat{\mu}+x_i-\overline{X}\right)\left(x_i-\hat{\mu}-x_i+\overline{X}\right)$

$$= \sum_{i=1}^{n}\left(2x_i-\overline{X}-\hat{\mu}\right)\left(\overline{X}-\hat{\mu}\right)$$

$$=\left(\bar{X}-\hat{\mu}\right)\sum_{i=1}^{n}\left(2x_i-\bar{X}-\hat{\mu}\right)=\left(\bar{X}-\hat{\mu}\right)\left(2n\bar{X}-n\bar{X}-n\hat{\mu}\right)$$

$$=\left(\bar{X}-\hat{\mu}\right)\left(n\bar{X}-n\hat{\mu}\right)=n\left(\bar{X}-\hat{\mu}\right)^2.$$

Therefore $\lambda=\left(\dfrac{\hat{\sigma}^2_{y,\alpha}}{\hat{\sigma}^2_{y,o}}\right)^{m/2}\exp\left\{-\dfrac{n}{2}\left(\bar{X}-\hat{\mu}\right)^2\right\}\le k$.

Using the first of the equations (5) we obtain

$$\bar{X}-\hat{\mu}=\bar{X}-\frac{m\bar{Y}+n\bar{X}\hat{\sigma}^2_{y,o}}{m+n\hat{\sigma}^2_{y,o}}=\frac{m\bar{X}-m\bar{Y}}{m+n\hat{\sigma}^2_{y,o}}=\frac{m\left(\bar{X}-\bar{Y}\right)}{m+n\hat{\sigma}^2_{y,o}}$$

and finally

$$\lambda=\left(\frac{\hat{\sigma}^2_{y,\alpha}}{\hat{\sigma}^2_{y,o}}\right)^{m/2}\exp\left\{-\frac{n}{2}\frac{m^2\left(\bar{X}-\bar{Y}\right)^2}{\left(m+n\hat{\sigma}^2_{y,o}\right)^2}\right\}\le k^{***}\ . \qquad (8)$$

2.2.1- Case m/n≈0

The exact distribution of the quantity in (8) cannot be derived analytically. To simplify things we use the approximation $m/n\approx 0$. This can be the case e.g. when for the old method a wealth of observations is available whereas this is not the case for the new method, (e.g. because of cost).

If, in equations (5), we put $m/n\approx 0$, we get

$$\hat{\mu}=\bar{X}\ \ \text{και}\ \ \hat{\sigma}^2_{y,o}=\frac{1}{m}\sum_{j=1}^{m}(y_j-\bar{X})^2\,.$$

Also, in this case, (8) can be written as $\lambda=\left(\dfrac{\hat{\sigma}^2_{y,\alpha}}{\hat{\sigma}^2_{y,o}}\right)^{m/2}\le k^{***}$. Substituting $\hat{\sigma}^2_{y,\alpha}$ from (7) and $\hat{\sigma}^2_{y,o}$ from the previous relationship, we get

$$\lambda=\frac{\sum_{j=1}^{m}(y_j-\bar{Y})^2}{\sum_{j=1}^{m}(y_j-\bar{X})^2}\le k^{***}. \qquad (9)$$

But

$$\sum_{j=1}^{m}(y_j-\bar{Y})^2=\sum_{j=1}^{m}\left[(y_j-\bar{X})+(\bar{X}-\bar{Y})\right]^2=\sum_{j=1}^{m}\left[(y_j-\bar{X})+2(y_j-\bar{X})(\bar{X}-\bar{Y})+(\bar{X}-\bar{Y})^2\right]$$

$$=\sum_{j=1}^{m}(y_j-\bar{X})^2+2(\bar{X}-\bar{Y})\sum_{j=1}^{m}(y_j-\bar{X})+m(\bar{X}-\bar{Y})^2$$

$$= \sum_{j=1}^{m}(y_j - \bar{X})^2 - 2m(\bar{X} - \bar{Y})^2 + m(\bar{X} - \bar{Y})^2$$

$$= \sum_{j=1}^{m}(y_j - \bar{X})^2 - m(\bar{X} - \bar{Y})^2 .$$

Hence (9) can be written as

$$\lambda = \frac{\sum_{j=1}^{m}(y_j - \bar{Y})^2}{\sum_{j=1}^{m}(y_j - \bar{Y})^2 + m(\bar{X} - \bar{Y})^2} \le k^{***} \quad \text{or} \quad \lambda = \frac{1}{1 + \frac{m(\bar{X} - \bar{Y})^2}{\sum_{j=1}^{m}(y_j - \bar{Y})^2}} \le k^{***} \quad \text{or}$$

$$\lambda = \frac{1}{1 + \frac{t_{New}^2}{m-1}} \le k^{**},$$

where

$$t_{New} = \frac{\sqrt{m(m-1)}(\bar{X} - \bar{Y})}{\sqrt{\sum_{j=1}^{m}(y_j - \bar{Y})^2}} = \frac{\sqrt{m}(\bar{X} - \bar{Y})}{S_Y} . \qquad (10)$$

Under the null hypothesis $H_o{:}\mu_X{=}\mu_Y$, the distribution of the above statistic, (10), is the central t-distribution with m-1 degrees of freedom.

Remark The assumption $n/m \approx 0$ it is not so helpful. This is so because, in this case, the likelihood statistic in (8) can be written in a form, which follows the chi-square distribution with one (1) degree of freedom.

2.3 - A different approach

In this section we will present a different approach for testing the hypothesis $H_o{:}\mu_X{=}\mu_Y$. This approach it is based on the "synthesis" of two other hypotheses and resembles, in some way, the union-intersection method of Roy (1957).

More precisely suppose that the hypothesis for testing is the $H_o{:}\mu_X{=}\mu_Y$ while the alternative is the $H_\alpha{:}\mu_X{\neq}\mu_Y$. The null hypothesis H_o can be separated into two partial hypotheses

$$H_o^{(1)} : \mu_X = \mu_o \ \text{και} \ H_o^{(2)} : \mu_Y = \mu_o ,$$

where μ_o is a specified constant. Suppose now that for that value of μ_o, at least one of the hypothesis $H_o^{(1)} : \mu_X = \mu_o$ and $H_o^{(2)} : \mu_Y = \mu_o$ is being rejected. Then $H_o{:}\mu_X{=}\mu_Y$ has to be rejected too. Alternatively, if both of the hypothesis $H_o^{(1)}: \mu_X = \mu_o$ and $H_o^{(2)}: \mu_Y = \mu_o$ can not be rejected, then we can not reject H_o:μ_X=μ_Y too. The value of μ_o is the estimated common value of μ_X and μ_Y, under H_o:μ_X=μ_Y, as it is given by the solution of the system (5). Denote by γ_z and γ_t the powers of testing the hypothesis $H_o^{(1)}: \mu_X = \mu_o$ and $H_o^{(2)}: \mu_Y = \mu_o$, respectively. Then the power of the different approach procedure, $\gamma_{DiffAppr}$ is given by $\gamma_{DiffAppr}=\gamma_z(1-\gamma_t)+(1-\gamma_z)\gamma_t+\gamma_z\gamma_t$, or $\gamma_{DiffAppr}=\gamma_z+\gamma_t-\gamma_z\gamma_t$. Obviously $\gamma_{DiffAppr}\geq\gamma_z$ and $\gamma_{DiffAppr}\geq\gamma_t$.

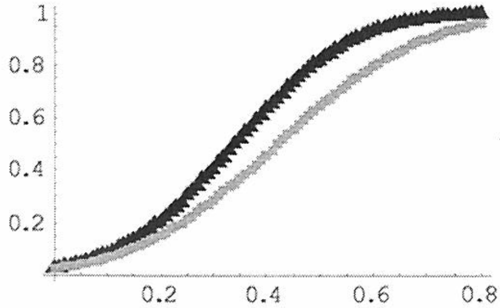
3. – Power comparisons

As it is well known, (see e.g. Cassella, Berger (1990)), the power, γ, of a test is defined as γ=P(reject H_o |H_α is true).More ever, if the test statistic follows, under the null hypothesis, a t distribution with v degrees of freedom then, under the alternative hypothesis H_α, the distribution of the test statistic follows, in general, a non-central t-distribution with v degrees of freedom and non-centrality parameter δ. The analytical form of such a distribution is

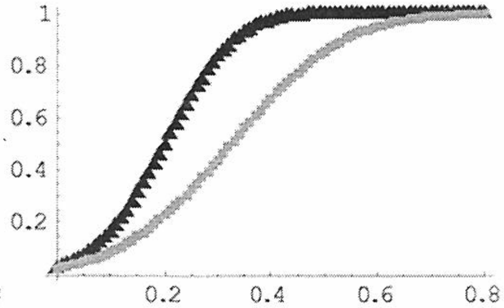
$$f(t;v,\delta) = \frac{1}{2^{(v+1)/2}\Gamma(\frac{v}{2})\sqrt{\pi v}} \int_0^\infty x^{(v-1)/2} \exp\left\{-\frac{1}{2}\left[x + \left(t\sqrt{\frac{x}{v}} - \delta\right)^2\right]\right\} dx\,,\ t\in R \qquad (11)$$

<u>*Significance test approach*</u>.

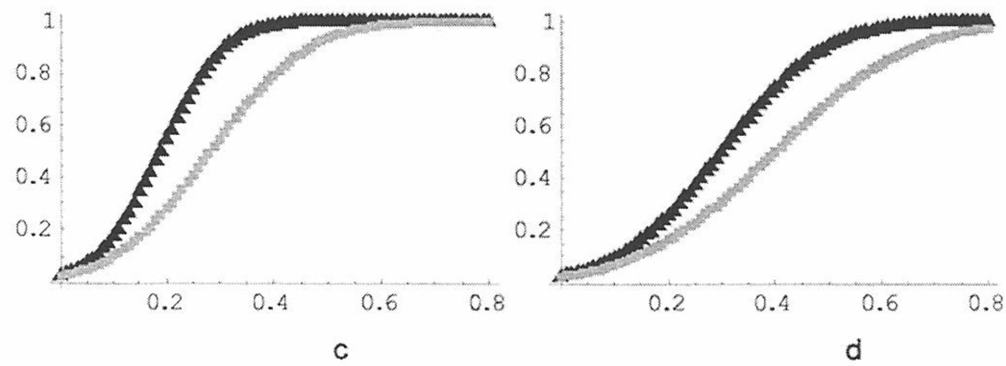
<u>A) Variances are equal</u>. Suppose that the variances σ_X^2 and σ_Y^2 are equal, say to σ^2. In this case we have to compare the powers of the test statistics in (1) and (3). The power, γ_{Old}, of this test statistic in (1), is γ_{Old}= P ($|t_{Old}|>t_{\alpha/2;n+m-2}|H_\alpha$ is true), while the power, γ_{New} , for the test statistic in (3), is γ_{New}= P($|t_{New}|>t_{\alpha/2;m-1}|H_\alpha$ is true). The non-centrality parameters are, $\delta_{Old}=|\mu_X-\mu_Y|\frac{1}{\sigma}\left(\frac{nm}{n+m}\right)^{1/2}$ and $\delta_{New}=|\mu_X-\mu_Y|(m/\sigma)^{1/2}$, respectively. No theoretical results can be obtained to see if γ_{Old} is greater or smaller to γ_{New}. However numerical computations have shown that γ_{New} is greater to γ_{Old}. (See graph below)



a



b



Graph1. Power comparisons for the test statistic t_{old} with t_{New}, when the variances are equal for α=0.05. a) m=35, n=60; b) m=97, n=60; c) m=110, n=87; d) m=45, n=55. (The black curve is the power of the new test while the gray is for the old one).

If the sample sizes are equal, i.e m=n (=N), then we can show, theoretically, that γ_{New} is greater to γ_{Old}. To see that note that in this case the power, γ_{Old}, is $\gamma_{Old}= P(|t_{Old}|>t_{\alpha/2;2(N-1)}|H_\alpha$ is true), while the power, γ_{New} is $\gamma_{New}= P(|t_{New}|>t_{\alpha/2;N-1}|H_\alpha$ is true). Also, the non centrality parameters are, $\delta_{Old}=|\mu_X-\mu_Y|(N/2\sigma)^{1/2}$ and $\delta_{New}=|\mu_X-\mu_Y|(N/\sigma)^{1/2}$, respectively.

Proposition In the case n=m (=N), if the variances σ_X^2 and σ_Y^2 are equal, say to σ^2, and 1/N can be taken, approximately, to be equal to zero, then

$$\gamma_{New}\geq\gamma_{Old}.$$

Proof. Obviously

$\gamma_{Old}= P(|t_{Old}|>t_{\alpha/2;2(N-1)}|H_\alpha$ is true)=P($t_{Old}> t_{\alpha/2;2(N-1)}|H_\alpha$ is true)+ P($t_{Old}<-t_{\alpha/2;2(N-1)}|H_\alpha$ is true).

Hence, using (11), we can write

$$\gamma_{Old}=\frac{1}{2^N\Gamma(N-1)\sqrt{2(N-1)\pi}}\left[\int_0^\infty x^{N-2}\left(\int_{t_{\alpha/2;2(N-1)}}^\infty \exp\left\{-\tfrac{1}{2}\left[x+\left(t\sqrt{\tfrac{x}{2(N-1)}}-\delta_{Old}\right)^2\right]\right\}dt\right)dx\right.$$

$$\left.+\int_0^\infty x^{N-2}\left(\int_{-\infty}^{-t_{\alpha/2;2(N-1)}} \exp\left\{-\tfrac{1}{2}\left[x+\left(t\sqrt{\tfrac{x}{2(N-1)}}-\delta_{Old}\right)^2\right]\right\}dt\right)dx\right]$$

and similarly

$$\gamma_{New}=\frac{1}{2^{N/2}\Gamma((N-1)/2)\sqrt{(N-1)\pi}}\left[\int_0^\infty x^{(N-2)/2}\left(\int_{t_{\alpha/2;N-1}}^\infty \exp\left\{-\tfrac{1}{2}\left[x+\left(t\sqrt{\tfrac{x}{N-1}}-\delta_{New}\right)^2\right]\right\}dt\right)dx\right.$$

$$\left.+\int_0^\infty x^{(N-2)/2}\left(\int_{-\infty}^{-t_{\alpha/2;N-1}} \exp\left\{-\tfrac{1}{2}\left[x+\left(t\sqrt{\tfrac{x}{N-1}}-\delta_{New}\right)^2\right]\right\}dt\right)dx\right]$$

Now, since $t_{\alpha/2;N-1}\approx t_{\alpha/2;2(N-1)}$, especially in the case where $1/N\approx0$, we get that $\gamma_{New}\geq\gamma_{Old}$ if

$$\frac{1}{2^{N/2}\Gamma((N-1)/2)\sqrt{(N-1)\pi}}\exp\left\{-\tfrac{1}{2}\left[x+\left(t\sqrt{\tfrac{x}{N-1}}-\delta_{New}\right)^2\right]\right\}\geq$$

$$\frac{1}{2^{N}\Gamma(N-1)\sqrt{2(N-1)\pi}}\exp\left\{-\tfrac{1}{2}\left[x+\left(t\sqrt{\tfrac{x}{2(N-1)}}-\delta_{Old}\right)^2\right]\right\}$$

or equivalently

$$-\frac{1}{2}\left[x+\left(t\sqrt{\frac{x}{N-1}}-\delta_{New}\right)^2\right]\geq-\frac{1}{2}\left[x+\left(t\sqrt{\frac{x}{2(N-1)}}-\delta_{Old}\right)^2\right]+c$$

where $c=\ln\left(\Gamma(\frac{N-1}{2})/2^{\frac{5}{2}}\Gamma(N-1)\right)$. Note that c is a negative number. By putting, in the last inequality, $\delta_{Old}=\delta_{New}\frac{\sqrt{2}}{2}$ and after some algebraic manipulation we arrive at

$$\delta_{New}^2-2t\sqrt{\frac{x}{N-1}}\delta_{New}+t^2\frac{x}{N-1}-4c\geq 0.$$

But this inequality is always true, and hence we have completed the proof of the proposition.

B) Variances are unequal. If the variances are unequal, that is $\sigma_X^2\neq\sigma_Y^2$, then the power of the test statistic in (3) must be compared with that of test statistic in (2). In this case, see Brentari, Carpita and Dancelli (2001), $\gamma_{Welch}=P(|t_{Welch}|>t_{\alpha/2;m-1}\,|\,H_\alpha\text{ is true})$ and the non-centrality parameter is $\delta_{Welch}=|\mu_X-\mu_Y|(m/\sigma_Y)^{1/2}$. Hence, in this case, the powers of the two test statistics coincide.

Likelihood ratio approach

Case $m/n\approx 0$. If the variances are equal, then $\gamma_{New}=P(|t_{New}|>t_{\alpha/2;m-1}|H_\alpha$ is true), where now t_{New} is the statistic in (10) and $\gamma_{Old}=P(|t_{Old}|>t_{\alpha/2};n|H_\alpha$ is true), where t_{Old} is the statistic in (1), written appropriately. Then, under the alternative hypothesis, each of these statistics follow the non-central t distribution with non-centrality parameters $\delta_{New}=|\mu_X-\mu_Y|/\sigma m^{1/2}$ and $\delta_{Old}=|\mu_X-\mu_Y|/\sigma m^{1/2}$, respectively. In this case $\delta_{New}=\delta_{Old}$ and hence $\gamma_{New}\leq\gamma_{Old}$. The last inequality comes from the Pearson-Hartley (1951) table for the power of the $F_{1,v}$ distribution.

If the variances are unequal, then $\gamma_{Welch}=P(|t_{Welch}|>t_{\alpha/2;m-1}\,|\,H_\alpha\text{ is true})$, where t_{Welch} is the statistic in (2). Following the work of Brentari, Carpita and Dancelli (2001) we can show that the two tests coincide. Hence $\gamma_{New}=\gamma_{Welch}$.

Different approach. In this case, if the variances are equal, ($\sigma_X=\sigma_Y=1$) then $\gamma_{DiffAppr}$ is slightly les than γ_{Old}. For example if n=46, m=37, α=0.05 $\mu_X=5$ and $\mu_Y=5.6$ we get that

$\gamma_{DiffAppr}$=0.72 and γ_{Old}=0.76. In the case of unequal variances $\gamma_{DiffAppr}$ is much better than γ_{Welch}. For example for n=46, m=37, α=0.05, μ_X=4, μ_Y=8, σ_X=1 and σ_Y=11 we obtain $\gamma_{DiffAppr}$=0.77 and γ_{Welch}=0.11. In both cases (variances equal or unequal) the value of μ_o is obtained from the formula (12)

4. – Numerical comparisons

In following three examples the mean of one population is double the mean of the other, so we expect the test to reject the hypothesis of equality of the two means.

Example 1. (*Case m=n*)

Using a computer we created two independent random samples from the normal distributions N(4, 49) and N(7, 100), each of size 45. The two samples gave $\overline{X} = 4{,}034$, $S_x = 6{,}75$, $\overline{Y} = 7{,}19$ and $S_Y = 8{,}93$. In order our first sample to be from a normal distribution with variance equal to one (1) we divide each observation, for both samples, by seven (7) to get N(0.57, 1) and N(1, 2.04). The Levene's test does not reject the hypothesis of equal variances, so we can make use of the statistic in (1). This statistic cannot reject the null hypothesis, (p=0,062). However, the t test given by relation (3) rejects the null hypothesis (p=0,01).

Example 2 (*Case m/n≈0*)

We created two independent random samples. The first, of size n=270, from the normal N(4, 1) and the second, of size m=20, from the normal distribution N(8, 121). The two samples gave $\overline{X} = 3{,}98$, $S_x = 1{,}03$, $\overline{Y} = 8{,}59$ και $S_Y = 9{,}56$. The test, for the equality of variances (Leven's test), rejects the hypothesis that the variances are equal, (p=<0,001). For that reason we use the Welch's (1938) test, (see equation (2)) for testing the hypothesis H_o:μ_X=μ_Y. The Welch's statistic gave p=0,0439. The corresponding value for the test statistic in (10) is p=0,0152.

Example 3 (*Case § 2.3*)

We created two independent random samples. The first, of size n=35, from the normal N(4, 1) and the second, of size m=42, from the normal distribution N(8, 225). The two samples gave $\overline{X} = 4{,}23$, $S_x = 0{,}97$, $\overline{Y} = 7{,}72$ και $S_Y = 14{,}26$. The solution of the system (5) gave $\hat{\mu}$=4,62. From this we conclude that for α=0,05 the null hypothesis is rejected since p_z=0.0214 and p_t=0.1661. The test statistic t_{Welch} in (2) cannot reject the null hypothesis, (p=0,1212).

5. Conclusions

From the previous analysis we can state the following:

a) The significance test t_{New},in (3) is more powerful than the t_{Old} in (1), in the case of equal variances, and coincides with the t_{Welch} in (2), in the case of unequal variances.

b) The likelihood ratio test in (10), in the case of equal variances, is less powerful than the test in (1), while in the case of unequal variances, coincides with the test in (2).

c) The power for the different approach, presented in § 2.3, is slightly less powerful than the test in (1), in the case of equal variances, but it is more powerful than the test in (2), if the variances are unequal.

Appendix - Solution of the system (5)

From the first of equations (5) we get

$$\hat{\mu}=\frac{m\bar{Y}+n\bar{X}\hat{\sigma}_{y,o}^2}{m+n\hat{\sigma}_{y,o}^2}. \qquad (12)$$

Making use of this relationship we can write

$$\sum_{j=1}^{m}(y_j-\hat{\mu})^2=\sum_{j=1}^{m}\left(y_j-\frac{m\bar{Y}+n\bar{X}\hat{\sigma}_{y,o}^2}{m+n\hat{\sigma}_{y,o}^2}\right)^2=\sum_{j=1}^{m}\left(\frac{my_j+n\hat{\sigma}_{y,o}^2y_j-m\bar{Y}-n\bar{X}\hat{\sigma}_{y,o}^2}{m+n\hat{\sigma}_{y,o}^2}\right)^2$$

$$=\frac{1}{\left(m+n\hat{\sigma}_{y,o}^2\right)^2}\sum_{j=1}^{m}\left\{m(y_j-\bar{Y})+n(y_j-\bar{X})\hat{\sigma}_{y,o}^2\right\}^2$$

$$=\frac{1}{\left(m+n\hat{\sigma}_{y,o}^2\right)^2}\sum_{j=1}^{m}\left\{m^2(y_j-\bar{Y})^2+n^2(y_j-\bar{X})^2\left(\hat{\sigma}_{y,o}^2\right)^2+2mn\hat{\sigma}_{y,o}^2(y_j-\bar{Y})(y_j-\bar{X})\right\}$$

$$=\frac{1}{\left(m+n\hat{\sigma}_{y,o}^2\right)^2}\left\{m^2\sum_{j=1}^{m}(y_j-\bar{Y})^2+n^2\left(\hat{\sigma}_{y,o}^2\right)^2\sum_{j=1}^{m}(y_j-\bar{X})^2+2mn\hat{\sigma}_{y,o}^2\sum_{j=1}^{m}(y_j-\bar{Y})(y_j-\bar{X})\right\}$$

$$=\frac{1}{\left(m+n\hat{\sigma}_{y,o}^2\right)^2}\left\{m^2\sum_{j=1}^{m}(y_j-\bar{Y})^2+n^2\left(\hat{\sigma}_{y,o}^2\right)^2\sum_{j=1}^{m}(y_j-\bar{X})^2+2mn\hat{\sigma}_{y,o}^2\sum_{j=1}^{m}(y_j-\bar{Y})y_j\right\}$$

$$=\frac{1}{\left(m+n\hat{\sigma}_{y,o}^2\right)^2}\left\{m^2\sum_{j=1}^{m}(y_j-\bar{Y})^2+n^2\left(\hat{\sigma}_{y,o}^2\right)^2\sum_{j=1}^{m}(y_j-\bar{X})^2+2mn\hat{\sigma}_{y,o}^2\sum_{j=1}^{m}(y_j-\bar{Y})^2\right\}$$

From the second of the equations (5) we get $m\hat{\sigma}_{y,o}^2=\sum_{j=1}^{m}(y_j-\hat{\mu})^2$ or, by replacing this term on the right hand side of the equal sign with the last result we get

$$m\hat{\sigma}_{y,o}^2\left(m+n\hat{\sigma}_{y,o}^2\right)^2=$$

$$=\left\{m^2\sum_{j=1}^{m}(y_j-\bar{Y})^2+n^2\left(\hat{\sigma}_{y,o}^2\right)^2\sum_{j=1}^{m}(y_j-\bar{X})^2+2mn\hat{\sigma}_{y,o}^2\sum_{j=1}^{m}(y_j-\bar{Y})^2\right\}$$

or $m\hat{\sigma}^2_{y,o}\left[m^2+n^2\left(\hat{\sigma}^2_{y,o}\right)^2+2mn\hat{\sigma}^2_{y,o}\right]=$

$$=\left\{m^2\sum_{j=1}^{m}(y_j-\overline{Y})^2+n^2\left(\hat{\sigma}^2_{y,o}\right)^2\sum_{j=1}^{m}(y_j-\overline{X})^2+2mn\hat{\sigma}^2_{y,o}\sum_{j=1}^{m}(y_j-\overline{Y})^2\right\}$$

or $m^3\hat{\sigma}^2_{y,o}+mn^2\left(\hat{\sigma}^2_{y,o}\right)^3+2m^2n\left(\hat{\sigma}^2_{y,o}\right)^2=$

$$=\left\{m^2\sum_{j=1}^{m}(y_j-\overline{Y})^2+n^2\left(\hat{\sigma}^2_{y,o}\right)^2\sum_{j=1}^{m}(y_j-\overline{X})^2+2mn\hat{\sigma}^2_{y,o}\sum_{j=1}^{m}(y_j-\overline{Y})^2\right\}$$

and finally

$$mn^2\left(\hat{\sigma}^2_{y,o}\right)^3+\left\{2m^2n-n^2\sum_{j=1}^{m}(y_j-\overline{X})^2\right\}\left(\hat{\sigma}^2_{y,o}\right)^2+\left\{m^3-2mn\sum_{j=1}^{m}(y_j-\overline{Y})^2\right\}\hat{\sigma}^2_{y,o}-$$

$$-m^2\sum_{j=1}^{m}(y_j-\overline{Y})^2=0.$$

Putting $A=mn^2$, $B=2m^2n-n^2\sum_{j=1}^{m}(y_j-\overline{X})^2$, $\Gamma=m^3-2mn\sum_{j=1}^{m}(y_j-\overline{Y})^2$ and

$\Delta=-m^2\sum_{j=1}^{m}(y_j-\overline{Y})^2$ we obtain

$$A(\hat{\sigma}^2_{y,o})^3+B(\hat{\sigma}^2_{y,o})^2+\Gamma\hat{\sigma}^2_{y,o}+\Delta=0 \qquad (13)$$

References

Brentari, E., Carpita, M. and Dancelli, L (2001) *An Improvement in the Welch Approximate t Test*, Metron LIX, 99-112.

Casella, G, Berger, R. L. (1990) *Statistical Inference*, Duxbury Press, Belmont, California

Cochran W. G, Cox, G. M (1950) *Experimental Designs*, New York, John Wiley & Sons, Inc.

Pearson, E. S., Hartley, H. O. (1951) *Charts of the Power Function of the Analysis of Variance Tests, Derived from the non-central F Distribution*, Biometrika 38, 112-130

Roy, S. N.(1957) *Some Aspects of Multivariate Analysis*, John Wiley

Satterthwaite, F. E. (1946) *An Approximate Distribution of Estimates of Variance Components*, Biometrics Bulletin (now called Biometrics) 12, 110-114

Welch, B. L (1938) *The Significance of the Difference Between Two Means when the Population Variances are Unequal*, Biometrika 29, 350-362.