فترات الثقة للفرق بين متوسطين

② Confidence Intervals for difference between two means :-

Given two different normal-populations with means and variances are $\theta_1, \theta_2, \sigma_1^2$ and σ_2^2 respectively.

نفرض لدينا مجتمعين طبيعيين الاول $N(\theta_1,\sigma_1^2)$ والثاني $N(\theta_2,\sigma_2^2)$
عندئذٍ لايجاد فترات الثقة للفرق بين المتوسطين هناك حالتين :

1) σ_1 and σ_2 are known :

عندما σ_1 و σ_2 معلومتين

Assume we tack two random samples of size n_1 and n_2 from two populations and we obtained $\bar{X}_1$ and $\bar{X}_2$, then

$(1-\alpha)100\%$ C.I. for $(\theta_1-\theta_2)$ may be obtained as:

$$\therefore \theta_{\bar{X}_1-\bar{X}_2} = E(\bar{X}_1-\bar{X}_2) = \theta_1-\theta_2$$

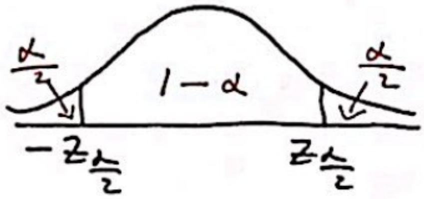
$$\sigma^2_{\bar{X}_1-\bar{X}_2} = Var(\bar{X}_1-\bar{X}_2) = Var(\bar{X}_1)+Var(\bar{X}_2)$$
$$= \frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}$$

$$\Rightarrow (\bar{X}_1-\bar{X}_2) \sim N\left(\theta_1-\theta_2 \,,\, \frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}\right)$$

$$\Rightarrow Z = \frac{(\bar{X}_1-\bar{X}_2)-(\theta_1-\theta_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

Now we know that من الرسم

$$Pr(-Z_{\frac{\alpha}{2}} < Z < Z_{\frac{\alpha}{2}}) = 1-\alpha$$



وبالتعويض عن قيمة Z بما يساويها

$$\Rightarrow P_r\left(-Z_{\frac{\alpha}{2}} < Z < Z_{\frac{\alpha}{2}}\right) = 1-\alpha$$

$$\Rightarrow P_r\left(-Z_{\frac{\alpha}{2}} < \frac{(\bar{X}_1-\bar{X}_2)-(\theta_1-\theta_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} < Z_{\frac{\alpha}{2}}\right) = 1-\alpha$$

وبضرب المتراجحة ب $\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$ نحصل على

$$Pr\left(-Z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}} < (\bar{X}_1-\bar{X}_2)-(\theta_1-\theta_2) < Z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right) = 1-\alpha$$

وبطرح $(\bar{X}_1-\bar{X}_2)$ من المتراجحة والضرب في (-1) نحصل على

$$Pr\left[(\bar{X}_1-\bar{X}_2)-Z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}} < \theta_1-\theta_2 < (\bar{X}_1-\bar{X}_2)+Z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right] = 1-\alpha$$

$\Rightarrow (1-\alpha)100\%$ C.I. for $(\theta_1-\theta_2)$ is :-

$$\boxed{(\bar{X}_1-\bar{X}_2)-Z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}} < \theta_1-\theta_2 < (\bar{X}_1-\bar{X}_2)+Z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$$

هذا يعني انه باحتمالية $(1-\alpha)$ القيمة الحقيقية للفرق بين المتوسطين واقعة بين الحد الادنى والحد الاعلى.

Ex:
Construct 99% C.I. for the actual differences between the average life times of two kinds of bulbs given that a r.s. of size (40) and (50) of two types are tested and found on average the first kind lasted (418) hours and the second kind lasted (402) hours on continuous. use the population standard deviations are $\sigma_1 = 26$ and $\sigma_2 = 22$

Solu:

$\alpha = 0.01 \Rightarrow Z_{\frac{\alpha}{2}} = Z_{0.005} = 2.575$

the 99% C.I. for $\theta_1 - \theta_2$ (actual difference)

$$(418-402) - 2.575\sqrt{\frac{26^2}{40}+\frac{22^2}{50}} < \theta_1 - \theta_2 < (418-402) + 2.575\sqrt{\frac{26^2}{40}+\frac{22^2}{50}}$$

$$12 - 2.575\sqrt{\frac{26^2}{40}+\frac{22^2}{50}} < \theta_1 - \theta_2 < 12 + 2.575\sqrt{\frac{26^2}{40}+\frac{22^2}{50}}$$

$$\Rightarrow 2.73 < \theta_1 - \theta_2 < 28.27$$

This means on average, the first kind is superior to the second kind because both limits are positive

هذا يعني أنه معدل العمر الزمني للنوع الاول اكبر منه معدل العمر الزمني للنوع الثاني لان حدي الثقة للفرق بين المتوسطين قيمهم موجبة .

ملاحظة :. لو كان احد الحدود سالب والاخر موجب عليه قد يكون الفرق بينهما صفر ولا نعرف اي المتوسطين اكبر .

2) σ_1 and σ_2 are unknown

عندما σ_1 و σ_2 غير معلومين هناك حالتين

1. n_1 and $n_2 \geq 30$ (عندما حجم العينة كبير)

$(1-\alpha)100\%$ C.I. for $\theta_1 - \theta_2$, is

$$(\bar{X}_1 - \bar{X}_2) - Z_{\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}} < \theta_1 - \theta_2 < (\bar{X}_1 - \bar{X}_2) + Z_{\frac{\alpha}{2}}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}$$

This means σ_1 and σ_2 are replaced in the samples standard deviations

يعني فقط في هذه الحالة نستبدل σ_1 و σ_2 في الحالة الاولى بـ S_1 و S_2 التي تمثل الانحراف المعياري للعينتين حيث أن

where $S_1^2 = \frac{1}{n_1-1}\sum(x_i - \bar{x}_1)^2$, $S_2^2 = \frac{1}{n_2-1}\sum(x_i - \bar{x}_2)^2$

نفرض تجانس التباين

2 n_1 and $n_2 < 30$ (عندما حجم العينة صغير)

In this case we only consider $(\sigma_1^2 = \sigma_2^2 = \sigma^2)$

then

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\theta_1 - \theta_2)}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}} \sim N(0,1)$$

$$\Rightarrow Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\theta_1 - \theta_2)}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim N(0,1)$$ باخراج σ^2 عامل مشترك

σ^2 may be estimated by that is called pooled (التجميعي) Variance (التباين) (S_p^2)

بما أنه في هذه الحالة σ_1^2 و σ_2^2 غير معلومة وافترضنا حالة تجانس التباين اي $\sigma_1^2 = \sigma_2^2 = \sigma^2$ ولكن σ^2 ايضاً غير معلومة لذلك سوف نقوم باخذ المقدر لها وهو ما يسمى بالتباين التجميعي (S_p^2) حيث أنه :

$$S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}$$

تابع تكملة ال اشتقاق حل

للاطلاع

H.W:

Show that S_p^2 is unbiased estimator for σ^2

Solu:

$$S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2} = \frac{\Sigma(x_i-\bar{x}_1)^2 + \Sigma(x_i-\bar{x}_2)^2}{n_1+n_2-2}$$

بقسمة البسط والمقام على $n_1 n_2$

$$= \frac{\frac{\Sigma(x_i-\bar{x}_1)^2}{n_1 n_2} + \frac{\Sigma(x_i-\bar{x}_2)^2}{n_1 n_2}}{\frac{n_1+n_2-2}{n_1 n_2}} = \frac{\frac{1}{n_2}\frac{(n_1-1)}{n_1}\sigma^2 + \frac{1}{n_1}\frac{(n_2-1)}{n_2}\sigma^2}{\frac{n_1+n_2-2}{n_1 n_2}}$$

$$= \frac{\frac{n_1\sigma^2 - \sigma^2 + n_2\sigma^2 - \sigma^2}{n_1 n_2}}{\frac{n_1+n_2-2}{n_1 n_2}} = \frac{\sigma^2(n_1+n_2-2)}{(n_1+n_2-2)} = \sigma^2$$

$\therefore E S_p^2 = E\sigma^2 = \sigma^2$ ($\therefore S_p^2$ is unbiased estimator for σ^2).

$$\because \frac{(n_1-1)S_1^2}{\sigma^2} \sim \chi^2_{(n_1-1)}$$

$$\frac{(n_2-1)S_2^2}{\sigma^2} \sim \chi^2_{(n_2-1)}$$

$$\Longrightarrow \frac{(n_1-1)S_1^2}{\sigma^2} + \frac{(n_2-1)S_2^2}{\sigma^2} \sim \chi^2_{(n_1+n_2-2)}$$

$$\because t = \frac{Z}{\sqrt{\frac{Y}{v}}} \quad \text{where } Z \sim N(0,1), \quad Y \sim \chi^2_v$$

$$\Longrightarrow t = \frac{\dfrac{(\bar{x}_1-\bar{x}_2)-(\theta_1-\theta_2)}{\sigma\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}}{\sqrt{\dfrac{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{\sigma^2}}{n_1+n_2-2}}}$$

$$\Longrightarrow t = \frac{(\bar{x}_1-\bar{x}_2)-(\theta_1-\theta_2)}{\sigma\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \cdot \frac{\sigma\sqrt{n_1+n_2-2}}{\sqrt{(n_1-1)S_1^2+(n_2-1)S_2^2}}$$

$$\Longrightarrow t = \frac{(\bar{x}_1-\bar{x}_2)-(\theta_1-\theta_2)}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim t(0,1,n_1+n_2-2)$$

$\Longrightarrow$ $(1-\alpha)100\%$ C.I. for $(\theta_1-\theta_2)$ when σ_1^2 and σ_2^2 unknown and $n_1, n_2 < 30$ is:-

$$\boxed{(\bar{x}_1-\bar{x}_2)-t_{\frac{\alpha}{2},n_1+n_2-2}\, S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} < \theta_1-\theta_2 < (\bar{x}_1-\bar{x}_2)+t_{\frac{\alpha}{2},n_1+n_2-2}\, S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$

Ex

If $x_1, x_2, \ldots, x_{10}$ and $y_1, y_2, \ldots, y_7$ be two independent samples from two Normal populations with means and variances are :-

$\theta_1, \theta_2, \sigma_1^2$ and σ_2^2, yield $\bar{X} = 4.2$, $\bar{y} = 3.4$

$S_1^2 = 49$, $S_2^2 = 32$

Construct 90 % C.I. for $(\theta_1 - \theta_2)$

Solu:

$$S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2} = \frac{9(49)+6(32)}{15} = 41.53$$

$$\alpha = 0.10 \Rightarrow t_{\frac{\alpha}{2}, n_1+n_2-2} = t_{0.05, 15} = 1.753$$

(from table of t-dist.)

Then 90% C.I. for $(\theta_1 - \theta_2)$ is

$$(4.2 - 3.4) - t_{0.05,15}\sqrt{41.53}\sqrt{\frac{1}{10}+\frac{1}{7}} < \theta_1 - \theta_2 < (4.2-3.4) + t_{0.05,15}\sqrt{41.53}\sqrt{\frac{1}{10}+\frac{1}{7}}$$

$$0.8 - 1.753\sqrt{4.53}\sqrt{\frac{1}{10}+\frac{1}{7}} < \theta_1 - \theta_2 < 0.8 + 1.753\sqrt{41.53}\sqrt{\frac{1}{10}+\frac{1}{7}}$$

$$\Rightarrow -5.16 < \theta_1 - \theta_2 < 6.76$$

هنا لا نستطيع معرفة اي مجتمع ينتج افضل لان الفرق بين المتوسطين بين قيمتين سالبة وموجبة ويمكن ان يكون مساوي للصفر.