(This lecture's notes are not based on the actual live lecture, they are a transcript of professor Siannis' notes and are added only for co pletion)

Example

Let $X_1, \dots, X_v$ r.s. of $N(\mu, \sigma^2)$, where σ^2 is unknown.

a) Find CI with confidence level $(1-a)$ for μ

b) Find a relation that must be satisfied by the magnitude of the sample v, so that the width of the CI is less than 2σ with probability $100(1-\beta)\%$

Solution

a) We proved in the last lecture that the CI is
$$\left(\bar{X} - t_{v-1,a/2} \cdot \frac{S}{\sqrt{v}},\ \bar{X} + t_{v-1,a/2} \cdot \frac{S}{\sqrt{v}}\right)$$

b) The length of the CI is $2t_{v-1,a/2} \cdot \frac{S}{\sqrt{v}}$, hence we have:
$$P\left(2t_{v-1,a/2} \cdot \frac{S}{\sqrt{v}} < 2\sigma\right) = 1-\beta \Leftrightarrow P\left(\frac{S}{\sigma} < \frac{\sqrt{v}}{t_{v-1,a/2}}\right) = 1-\beta \Leftrightarrow$$
$$\Leftrightarrow P\left(\frac{S}{\sigma} \geq \frac{\sqrt{v}}{t_{v-1,a/2}}\right) = \beta \Leftrightarrow P\left(\frac{S^2}{\sigma^2} \geq \frac{v}{t^2_{v-1,a/2}}\right) = \beta \Leftrightarrow$$
$$\Leftrightarrow P\left(\frac{(v-1)S}{\sigma} \geq \frac{v(v-1)}{t^2_{v-1,a/2}}\right) = \beta$$

So $\frac{v(v-1)}{t^2_{v-1,a/2}} = \chi^2_{v-1,\beta}$, where $\chi^2_{v-1,\beta}$ is the point for which $P(X > \chi^2_{v-1,\beta}) = \beta$ for $X \sim \chi^2_{v-1}$

So $v(v-1) = \chi^2_{v-1,\beta} \cdot t^2_{v-1,a/2}$, which is the sought relation.

Note

CI for the mean μ of normal distribution

i) If σ^2 is known and $v > 30$, then from CLT (Central Limit Theorem) $\frac{\bar{X}-\mu}{\sigma/\sqrt{v}} \sim N(0,1)$ towards infinity and the CI is $\bar{X} \pm Z_{a/2} \cdot \frac{\sigma}{\sqrt{v}}$

ii) If σ^2 is unknown and $v > 30$ the r.v. $\frac{\bar{X}-\mu}{s/\sqrt{v}} \sim t_{v-1}$ towards infinity and the CI is $\bar{X} \pm t_{v-1,a/2} \cdot s/\sqrt{v}$

CI for the difference of means of two normal populations

(If the populations are not normal, but the magnitudes of the samples are $v_1, v_2 > 30$, then the following still hold true because of the CLT)

Let $X_1, \dots, X_{v_1}$ r.s. of $N(\mu_1, \sigma_1^2)$ and $Y_1, \dots, Y_{v_2}$ r.s of $N(\mu_2, \sigma_2^2)$. We have $\bar{X} \sim N(\mu_1, \frac{\sigma_1^2}{v_1})$, $\bar{Y} \sim N(\mu_2, \frac{\sigma_2^2}{v_2})$ and $\bar{X}, \bar{Y}$ are independent so $\bar{X} - \bar{Y} \sim N(\mu_1 - \mu_2, \frac{\sigma_1^2}{v_1} + \frac{\sigma_2^2}{v_2})$

Moment-generating function: $M_{\bar{X}-\bar{Y}}(t) = E[e^{t(\bar{X}-\bar{Y})}] =$
$= E[e^{t\bar{X}}] \cdot E[e^{-t\bar{Y}}] = M_{\bar{X}}(t) \cdot M_{\bar{Y}}(-t)$

For $X \sim N(\mu, \sigma^2)$ we know that $M_X(t) = \exp\{\mu t + \sigma^2 \frac{t^2}{2}\}$,
so $M_{\bar{X}-\bar{Y}}(t) = M_{\bar{X}}(t) \cdot M_{\bar{Y}}(-t) = \exp\{\mu_1 t + \frac{\sigma_1^2}{v_1} \cdot \frac{t^2}{2} - \mu_2 t + \frac{\sigma_2^2}{v_2} \cdot \frac{t^2}{2}\} =$
$= \exp\{(\mu_1 - \mu_2) \cdot t + (\frac{\sigma_1^2}{v_1} + \frac{\sigma_2^2}{v_2}) \cdot \frac{t^2}{2}\}$
which is the moment-generating function of $N(\mu_1 - \mu_2, \frac{\sigma_1^2}{v_1} + \frac{\sigma_2^2}{v_2})$

Finally: $Z = \frac{(\bar{X}-\bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2/v_1 + \sigma_2^2/v_2}} \sim N(0,1)$

i) If σ_1^2, σ_2^2 are known, then the boundaries of the CI with confidence level $1-a$ for $\mu_1 - \mu_2$ are:
$$\bar{X} - \bar{Y} \pm Z_{a/2} \sqrt{\sigma_1^2/v_1 + \sigma_2^2/v_2}$$

ii) If σ_1^2, σ_2^2 are unknown then:

a) if $v_1, v_2 > 30$, then an ~~au~~ approximate CI with confidence level $1-a$ for $\mu_1 - \mu_2$ is:

$$\bar{X} - \bar{Y} \pm Z_{a/2} \sqrt{S_1^2/v_1 + S_2^2/v_2}$$

b) if $v_1, v_2 < 30$, then the case is complicated. We usually suppose that $\sigma_1^2 = \sigma_2^2 = \sigma^2$. Then:

$\frac{(v_1-1)S_1^2}{\sigma^2} \sim \chi^2_{v_1-1}$, $\frac{(v_2-1)S_2^2}{\sigma^2} \sim \chi^2_{v_2-1}$ and these are independent so $\frac{(v_1-1)S_1^2 + (v_2-1)S_2^2}{\sigma^2} \sim \chi^2_{v_1+v_2-2} \Rightarrow$

$$\Rightarrow Z = \frac{(\bar{X}-\bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\sigma^2/v_1 + \sigma^2/v_2}} \sim N(0,1)$$

Since X, Y are independent, X, S_1^2, Y, S_2^2 are also independent

$$\text{so } T = \frac{\dfrac{(\bar{X}-\bar{Y}) - (\mu_1-\mu_2)}{\sqrt{\sigma^2\left(\frac{1}{v_1} + \frac{1}{v_2}\right)}}}{\sqrt{\dfrac{(v_1-1)\cdot S_1^2 + (v_2-1)S_2^2}{\sigma^2} \Big/ v_1+v_2-2}} = \frac{(\bar{X}-\bar{Y}) - (\mu_1-\mu_2)}{\sqrt{\dfrac{(v_1-1)S_1^2 + (v_2-1)S_2^2}{v_1+v_2-2}\left(\frac{1}{v_1} + \frac{1}{v_2}\right)}} \sim t_{v_1+v_2-2}$$

and the boundaries for the CI with c.l. $1-a$ for $\mu_1 - \mu_2$ are

$$(\bar{X} - \bar{Y}) \pm t_{v_1+v_2-2, a/2} \sqrt{\frac{(v_1-1)S_1^2 + (v_2-1)S_2^2}{v_1+v_2-2}\left(\frac{1}{v_1} + \frac{1}{v_2}\right)}$$

Example

If d is a predetermined value for the error margin, find v so that the $100(1-a)\%$ CI for μ is not largen than $(\bar{X} - d, \bar{X} + d)$.

Solution

Essentially, we want $P(|\bar{X} - \mu| \leq d) = 1 - a \Leftrightarrow$

$\Leftrightarrow P(-d \leq \mu - \bar{X} \leq d) = 1 - a \Leftrightarrow P(\bar{X} - d \leq \mu \leq \bar{X} + d) = 1 - a$,

but $P(\bar{X} - Z_{a/2} \cdot \frac{\sigma}{\sqrt{v}} \leq \mu \leq \bar{X} - Z_{a/2} \frac{\sigma}{\sqrt{v}}) = 1 - a$, so $d = Z_{a/2} \cdot \frac{\sigma}{\sqrt{v}} \Rightarrow$

$\Rightarrow d^2 = Z_{a/2}^2 \frac{\sigma^2}{v} \Rightarrow v = Z_{a/2}^2 \frac{\sigma^2}{d^2}$

Example

Let $X_1, \ldots, X_{16}$ r.s. of $N(\mu, \sigma^2 = 1)$ with $\bar{X} = 2$

Find a 95% CI for μ. It is given that $Z_{0,025} = 1,96$

Solution

The CI is $\bar{X} \pm Z_{a/2} \cdot \frac{\sigma}{\sqrt{n}} = 2 \pm 1,96 \cdot \frac{1}{\sqrt{16}} = 2 \pm 0,49$

so CI: $[1.51, 2.49]$