02: Interval Estimation

* for any (p) paramtr, an intrval → confidence intrval within the (p) paramtr may be expected to line with a certain degree of confidence say 'α' (alpha).
* In othr words, gvn a r's of 'n' independnt values $x_1, x_2, \dots x_n$ of r.v 'x' hving the pdf $f(x, \theta)$, θ being a paramtr, we wish to find t_1 & t_2 of the () of $x_1, x_2 \dots x_n$, such that

$$\boxed{P(t_1 < \theta < t_2) = 1 - \alpha}$$

* The intrval (t_1, t_2) is random intrval.
* The intrval (t_1, t_2) → confidnce intrval / fiducial intrval & $1-\alpha$ → confidnce (coef) of intrval t_1, t_2.
* The limits t_1 & t_2 → confidnce limits
* A confidnce intrval for unknown paramtrs gives both & indication of numerical value of an unknown paramtr as well as a measure of hw confidnt we are of that numerical value.

($1-\alpha$ → confidence intrval)

* If we take $\alpha = 0.05$, the 95% posses the meaning that if 100 intervals are constructed based on 100 different samples of the same size from the (P), 95 of the will include the true value of the paramtr.

* By accepting 95% confidence interval for the paramtr, the freq of wrong estimate is approximatly = 5%

⇒ Confidence interval for mean of a normal (P) ~~[illegible]~~ $N(\mu, \sigma)$

<u>Case 1</u> :

To estimate when σ is known,

To estimate μ, let us draw a r.s $x_1, x_2, \ldots x_n$ of size n from normal distrn

let $\bar{x}$ be the mean of r.s of size n drawn from the normal (P) $N(\mu, \sigma)$.

then,

$\bar{x} \longrightarrow N(\mu, \sigma/\sqrt{n})$

$\therefore\ z = \dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}} \rightarrow N(0,1).$ $\left[\begin{array}{l} x \rightarrow N(\mu,\sigma) \\ z = \dfrac{x-\mu}{\sigma}. \end{array}\right]$

from the area property, of the standard normal distri~ we gt,

$$\boxed{P\left\{ |z| \leq z_{\alpha/2} \right\} = 1-\alpha.}$$

$$P\left\{ -z_{\alpha/2} \leq z \leq z_{\alpha/2} \right\} = 1-\alpha.$$

$$P\left\{ -z_{\alpha/2} \leq \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \leq z_{\alpha/2} \right\} = 1-\alpha.$$

we need only μ at here

$\times\ \sigma/\sqrt{n}$ (to cancel $\sigma/\sqrt{n}$), $P\left\{ -z_{\alpha/2}\,\sigma/\sqrt{n} \leq \bar{x}-\mu \leq z_{\alpha/2}\,\sigma/\sqrt{n} \right\} = 1-\alpha.$

$-\bar{x}$ (to cancel $\bar{x}$) $P\left\{ -z_{\alpha/2}\,\sigma/\sqrt{n} - \bar{x} \leq -\mu \leq z_{\alpha/2}\,\sigma/\sqrt{n} - \bar{x} \right\} = 1-\alpha.$

(multiply with $-$) $P\left\{ -z_{\alpha/2}\,\sigma/\sqrt{n} + \bar{x} \leq \mu \leq z_{\alpha/2}\,\sigma/\sqrt{n} + \bar{x} \right\} = 1-\alpha.$

to cancel $-$ at μ. $P\left\{ \bar{x} + z_{\alpha/2}\,\sigma/\sqrt{n} \geq \mu \geq \bar{x} - z_{\alpha/2}\,\sigma/\sqrt{n} \right\} = 1-\alpha.$

$$P\left\{ \bar{x} - z_{\alpha/2}\,\sigma/\sqrt{n} \leq \mu \leq \bar{x} + z_{\alpha/2}\,\sigma/\sqrt{n} \right\} = 1-\alpha.$$

Here the interval,

$$\left\{ \bar{x} - z_{\alpha/2}\,\sigma/\sqrt{n},\ \bar{x} + z_{\alpha/2}\,\sigma/\sqrt{n} \right\}.$$

is called 100 $(1-\alpha)$ %. confidnce interval. for the mean μ. of normal (P)

Note,

1) If $\alpha = 0.05$, $z_{\alpha/2} = 1.96$, so the 95% confidnce interval for μ is $\left\{ \bar{x} - 1.96\,\sigma/\sqrt{n},\ \bar{x} + 1.96\,\sigma/\sqrt{n} \right\}$.

2) If $\alpha = 0.02$, $z_{\alpha/2} = 2.326$, so 98% C. interval for μ is $\left\{ \bar{x} - 2.326\,\sigma/\sqrt{n},\ \bar{x} + 2.326\,\sigma/\sqrt{n} \right\}$

3) If $\alpha = 0.01$, $Z_{\alpha/2} = 2.58$, so 99% of c. interval for μ is $\{\bar{x} - 2.58\,\sigma/\sqrt{n}, \bar{x} + 2.58\,\sigma/\sqrt{n}\}$

4) If $\alpha = 0.10$, $Z_{\alpha/2} = 1.645$, so 90% of c. intrval for μ is $\{\bar{x} - 1.645\,\sigma/\sqrt{n}, \bar{x} + 1.645\,\sigma/\sqrt{n}\}$

case : 2 =

when σ is unknown, n is large $(n \geq 30)$ →
when the samples is drawn from a normal (P) or not by central limit theorem.

$$Z = \frac{\bar{x} - \mu}{s/\sqrt{n}} \longrightarrow N(0,1) \quad \text{as } n \to \infty.$$

from the area property of normal distri~

$$P\{|Z| \leq Z_{\alpha/2}\} = 1-\alpha.$$

$$P\{-Z_{\alpha/2} \leq Z \leq Z_{\alpha/2}\} = 1-\alpha.$$

$$P\left\{-Z_{\alpha/2} \leq \frac{\bar{x}-\mu}{s/\sqrt{n}} \leq Z_{\alpha/2}\right\} = 1-\alpha.$$

$$P\{-Z_{\alpha/2}\, s/\sqrt{n} \leq \bar{x}-\mu \leq Z_{\alpha/2}\, s/\sqrt{n}\} = 1-\alpha.$$

$$P\{-\bar{x}-Z_{\alpha/2}\, s/\sqrt{n} \leq -\mu \leq \bar{x}+Z_{\alpha/2}\, s/\sqrt{n}\} = 1-\alpha.$$

$$P\{\bar{x} + Z_{\alpha/2}\, s/\sqrt{n} \geq \mu \leq \bar{x} + Z_{\alpha/2}\, s/\sqrt{n}\} = 1-\alpha.$$

$$P\{\bar{x} - Z_{\alpha/2}\, s/\sqrt{n} \leq \mu \leq \bar{x} + Z_{\alpha/2}\, s/\sqrt{n}\} = 1-\alpha.$$

Thus the $100(1-\alpha)\%$ confidence interval for μ,

$$\left[\bar{x} - Z_{\alpha/2}\, s/\sqrt{n},\ \bar{x} + Z_{\alpha/2}\, s/\sqrt{n}\right]$$

case : 3 =

when σ is unknown, n is small $(n < 30)$ →

Let $x_1, x_2 \ldots x_n$ be a r.s drawn from $N(\mu, \sigma)$ where σ is unknown,

Let $\bar{x}$ be sample mean & s^2 be its sample varience.

$$\boxed{t = \frac{\bar{x}-\mu}{s/\sqrt{n-1}} \longrightarrow t_{(n-1)} \text{ d.f.}}$$

Hence $100(1-\alpha)\%$ CI for μ is,

$$P\{|t| \le t_{\alpha/2}\} = 1-\alpha$$

$$P\{-t_{\alpha/2} \le t \le t_{\alpha/2}\} = 1-\alpha.$$

$$P\left\{-t_{\alpha/2} \le \frac{\bar{x}-\mu}{s/\sqrt{n-1}} \le t_{\alpha/2}\right\} = 1-\alpha.$$

$$P\{-t_{\alpha/2}\, s/\sqrt{n-1} \le \bar{x}-\mu \le t_{\alpha/2}\, s/\sqrt{n-1}\} = 1-\alpha.$$

$$P\{-\bar{x}-t_{\alpha/2}\, s/\sqrt{n-1} \le -\mu \le -\bar{x}+t_{\alpha/2}\, s/\sqrt{n-1}\} = 1-\alpha.$$

$$P\{\bar{x}+t_{\alpha/2}\, s/\sqrt{n-1} \le \mu \le \bar{x}-t_{\alpha/2}\, s/\sqrt{n-1}\} = 1-\alpha$$

$$P\{\bar{x}-t_{\alpha/2}\, s/\sqrt{n-1} \le \mu \le \bar{x}+t_{\alpha/2}\, s/\sqrt{n-1}\} = 1-\alpha.$$

Thus $100(1-\alpha)\%$ of CI for μ is,

$$\left[\bar{x}-t_{\alpha/2}\, s/\sqrt{n-1},\ \bar{x}+t_{\alpha/2}\, s/\sqrt{n-1}\right].$$

where $t_{\alpha/2}$ is obtained by referring the student's t table for $(n-1)$ df & (prob) α.

⇒ Interval estimate of the difference of 2 (p) means =

Case 1 =

when σ_1, σ_2 known,

let $\bar{x}_1$ be mean of the sample of size n_1 taken from a (p) with mean μ_1 & SD σ_1.

then, $\bar{x}_1 \longrightarrow N\left(\mu_1, \frac{\sigma_1}{\sqrt{n_1}}\right)$

let $\bar{x}_2$ be the mean of the sample of size n_2 taken from a (p) with mean μ_2 & SD σ^2.

Then,

$$\bar{x}_2 \longrightarrow N\left(\mu_2, \frac{\sigma_2}{\sqrt{n_2}}\right)$$

By additive property,

$$\bar{x}_1-\bar{x}_2 \longrightarrow N\left(\mu_1-\mu_2, \sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right)$$

$$\therefore, z = \frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \longrightarrow N(0,1).$$

by area property of Normal distri~

$$P\{|z| \le z_{\alpha/2}\} = 1-\alpha.$$

$$P\{-z_{\alpha/2} \le z \le z_{\alpha/2}\} = 1-\alpha.$$

$$P\left\{-z_{\alpha/2} \le \frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \le z_{\alpha/2}\right\} = 1-\alpha.$$

$$P\left\{-z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}} \le (\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2) \le z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right\} = 1-\alpha$$

$$P\left\{-(\bar{x}_1-\bar{x}_2)-z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}} \le -\mu_1-\mu_2 \le -(\bar{x}_1-\bar{x}_2)+z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right\} = 1-\alpha.$$

$$P\left\{(\bar{x}_1-\bar{x}_2) - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}} \le \mu_1-\mu_2 \le (\bar{x}_1-\bar{x}_2) + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right\} = 1-\alpha$$

Thus $100(1-\alpha)\%$ CI for μ is,

$$\left[(\bar{x}_1-\bar{x}_2) - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{2}},\ (\bar{x}_1-\bar{x}_2) + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right].$$

where $z_{\alpha/2} \rightarrow$ determined from normal table.

* when $\alpha = 0.05$, $z_{\alpha/2} = 1.96$, so $100(1-\alpha)\% = 95\%$

CI for $\mu_1-\mu_2$ is,

$$\left[(\bar{x}_1-\bar{x}_2) - 1.96\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}},\ (\bar{x}_1-\bar{x}_2) + 1.96\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right].$$

when $\alpha = 0.01$ & $z_{\alpha/2} = 2.58$

So the 99% for $\mu_1-\mu_2$ is.

$$\left[(\bar{x}_1-\bar{x}_2) - 2.58\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}},\ (\bar{x}_1-\bar{x}_2) + 2.58\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right]$$

case: 2 =

when σ_1, σ_2 unknown, n_1, n_2 large.

In this case we replace σ_1 & σ_2 respectively by their estimate s_1 & s_2, so confidence interval for μ_1 & μ_2 is,

$$(\bar{x}_1-\bar{x}_2) - z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}},\ (\bar{x}_1-\bar{x}_2) + z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$$

* $\alpha = 0.5$, $z_{\alpha/2} = 1.96$

so 95% CI for $\mu_1-\mu_2$ is,

$$(\bar{x}_1-\bar{x}_2) - 1.96\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}},\ (\bar{x}_1-\bar{x}_2) + 1.96\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$$

* $\alpha = 0.02$, $z_{\alpha/2} = 2.326$.

* so 98% C I (confidence interval) for $\mu_1-\mu_2$,

$$(\bar{x}_1-\bar{x}_2) - 2.326\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}},\ (\bar{x}_1-\bar{x}_2) + 2.326\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$$

* $\alpha = 0.01$, $z_{\alpha/2} = 2.58$.

so 99% of CI of $\mu_1-\mu_2$ is,

$$(\bar{x}_1-\bar{x}_2) - 2.58\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}},\ (\bar{x}_1-\bar{x}_2) + 2.58\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$$

case: 3 =

when $\sigma_1 = \sigma_2 = \sigma$ unknown, n_1, n_2 small,

Here $t = \dfrac{(\bar{x}_1-\bar{x}_2) - (\mu_1-\mu_2)}{\sqrt{\left(\frac{\sigma^2}{n_1}\right)+\frac{\sigma_2^2}{n_2}}}$ $\rightarrow$ Student's t distri~ with $v=(n_1+n_2-2)$ df.

$$\sigma^2 = \frac{n_1 s_1^2 + n_2 s_2^2}{n_1+n_2-2}$$

refer the t curve for $v = (n_1 + n_2 - 2)$ df (degree of freedom)
& probability level $p = \alpha$,
the table value of t, is $t_{\alpha/2}$, then
we have,

$$P\{|t| > t_{\alpha/2}\} = \alpha$$

$$P\{|t| \le t_{\alpha/2}\} = 1 - \alpha$$

$$P\{-t_{\alpha/2} \le t \le t_{\alpha/2}\} = 1 - \alpha$$

$$P\left\{-t_{\alpha/2} \le \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{\sigma^2}{n_1}\right) + \frac{\sigma^2}{n_2}}} \le t_{\alpha/2}\right\} = 1 - \alpha$$

$$\cancel{(\bar{x}_1 - \bar{x}_2) - t_{\alpha/2}\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}}$$

$$\cancel{(\bar{x}_1 - \bar{x}_2) + t_{\alpha/2}\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}}$$

$$P\left\{-t_{\alpha/2}\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}} \le (\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2) \le t_{\alpha/2}\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}\right\} = 1 - \alpha$$

$$P\left\{-(\bar{x}_1 - \bar{x}_2) - t_{\alpha/2}\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}} \le -(\mu_1 - \mu_2) \le -(\bar{x}_1 - \bar{x}_2) + t_{\alpha/2}\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}\right\} = 1 - \alpha$$

$$P\left\{(\bar{x}_1 - \bar{x}_2) - t_{\alpha/2}\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}} \le \mu_1 - \mu_2 \le (\bar{x}_1 - \bar{x}_2) + t_{\alpha/2}\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}\right\} = 1 - \alpha$$

Thus $100(1-\alpha)\%$ CI for $\mu_1 - \mu_2$ is,

$$\left[(\bar{x}_1 - \bar{x}_2) - t_{\alpha/2}\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}},\ (\bar{x}_1 - \bar{x}_2) + t_{\alpha/2}\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}\right]$$

where $t_{\alpha/2}$ is obtained by refering the t table for $(n_1 + n_2 - 2)$ df.

⇒ CI for the varience of normal (P) $N(\mu, \sigma)$

let s^2 be the varience of sample of size n, $(n < 30)$ drawn from $N(\mu, \sigma)$ we know that the statistic $\chi^2 = \frac{ns^2}{\sigma^2}$ $\rightarrow \chi^2_{(n-1)}$ df.

Now by refering χ^2 table, we find $\chi^2_{1-\alpha/2}$ and $\chi^2_{\alpha/2}$ such that,

$$P\{\chi^2_{1-\alpha/2} \le \chi^2 \le \chi^2_{\alpha/2}\} = 1 - \alpha$$

where $\chi^2_{1-\alpha/2}$ & $\chi^2_{\alpha/2}$ are obtained by refering the table for $(n-1)$ df & (prob) $1 - \alpha/2$ & $\alpha/2$ respectivly

$$P\left\{\chi^2_{1-\alpha/2} \le \frac{ns^2}{\sigma^2} \le \chi^2_{\alpha/2}\right\} = 1 - \alpha$$

$$P\left\{\frac{1}{\chi^2_{1-\alpha/2}} \ge \frac{\sigma^2}{ns^2} \ge \frac{1}{\chi^2_{\alpha/2}}\right\} = 1 - \alpha$$

$$P\left\{\frac{ns^2}{\chi^2_{1-\alpha/2}} \ge \sigma^2 \ge \frac{ns^2}{\chi^2_{\alpha/2}}\right\} = 1 - \alpha$$

$$P\left\{\frac{ns^2}{\chi^2_{\alpha/2}} \le \sigma^2 \le \frac{ns^2}{\chi^2_{1-\alpha/2}}\right\} = 1 - \alpha$$

$100(1-\alpha)\%$ for σ^2 is $\left[\frac{ns^2}{\chi^2_{\alpha/2}}, \frac{ns^2}{\chi^2_{1-\alpha/2}}\right]$

⇒ CI for the proportional of success of a binomial (P) =

let $p' = \left(\frac{x}{n}\right)$ be the proportional of success of sample of size 'n' drawn from a binomial (P) with paramtrs n & p where p is unknown & n is assumed to be known.

Then we know that,

$$Z = \frac{p' - p}{\sqrt{\frac{pq}{n}}} \to N(0,1) \text{ for large } n.$$

from normal table we get,

$$P\{|Z| \le Z_{\alpha/2}\} = 1-\alpha.$$

$$P\{-Z_{\alpha/2} \le Z \le Z_{\alpha/2}\} = 1-\alpha.$$

$$P\left\{-Z_{\alpha/2} \le \frac{p'-p}{\sqrt{\frac{pq}{n}}} \le Z_{\alpha/2}\right\} = 1-\alpha.$$

$$P\left\{-Z_{\alpha/2}\sqrt{\frac{pq}{n}} \le p'-p \le Z_{\alpha/2}\sqrt{\frac{pq}{n}}\right\} = 1-\alpha.$$

$$P\left\{-p'-Z_{\alpha/2}\sqrt{\frac{pq}{n}} \le -p \le p'+Z_{\alpha/2}\sqrt{\frac{pq}{n}}\right\} = 1-\alpha.$$

$$P\left\{p'-Z_{\alpha/2}\sqrt{\frac{pq}{n}} \le p \le p'+Z_{\alpha/2}\sqrt{\frac{pq}{n}}\right\} = 1-\alpha.$$

$$\left[p' - Z_{\alpha/2}\sqrt{\frac{pq}{n}}\ ,\ p' + Z_{\alpha/2}\sqrt{\frac{pq}{n}}\right]$$

Since p is unknown we can replace p & q by their unbiased (E) p' & q'

Thus 100(1−α)% CI for P is,

$$\left[p' - Z_{\alpha/2}\sqrt{\frac{p'q'}{n}}\ ,\ p' + Z_{\alpha/2}\sqrt{\frac{p'q'}{n}}\right]$$

where $Z_{\alpha/2}$ can be determined from the normal table for a given α,

Note,

* when α = 0.05, $Z_{\alpha/2} = 1.96$, so 95% CI for P,

$$\left[p' - 1.96\sqrt{\frac{p'q'}{n}}\ ,\ p' + 1.96\sqrt{\frac{p'q'}{n}}\right]$$

* α = 0.02, $Z_{\alpha/2} = 2.326$, so 98% CI for P,

$$\left[p' - 2.326\sqrt{\frac{p'q'}{n}}\ ,\ p' + 2.326\sqrt{\frac{p'q'}{n}}\right]$$

* α = 0.01, $Z_{\alpha/2} = 2.58$, so 99% CI for P,

$$\left[p' - 2.58\sqrt{\frac{p'q'}{n}}\ ,\ p' + 2.58\sqrt{\frac{p'q'}{n}}\right]$$

⇒ Interval estimate of the difference of proportions of 2 binomial (P) =

$$p_1' - p_2' \to N\left(p_1 - p_2, \sqrt{\frac{p_1q_1}{n} + \frac{p_2q_2}{n}}\right)$$

where n_1 & n_2 are larger.

$$\therefore\ Z = \frac{(p_1' - p_2')(p_1 - p_2)}{\sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}}} \to N(0,1)$$

100(1−α)% CI for $(p_1 - p_2)$ is,

$$\left[(p_1' - p_2') - Z_{\alpha/2}\sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}}\ ,\ (p_1' - p_2') + Z_{\alpha/2}\sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}}\right]$$

p_1, q_1 & p_2, q_2 are unknown we are estimated as $p_1 = p_1'$, $q_1 = q_1'$, $p_2 = p_2'$ & $q_2 = q_2'$

Thus $100(1-\alpha)\%$ CI for (P_1-P_2) is,

$$\left(p_1'-p_2'\right) - z_{\alpha/2}\sqrt{\frac{p_1'q_1'}{n_1}+\frac{p_2'q_2'}{n_2}},\ \left(p_1'-p_2'\right)+z_{\alpha/2}\sqrt{\frac{p_1'q_1'}{n_1}+\frac{p_2'q_2'}{n_2}}$$

Note,

* The 95% CI for (P_1-P_2) is $\alpha=0.05$, $z_{\alpha/2}=1.96$

$$\left(p_1'-p_2'\right) - 1.96\sqrt{\frac{p_1'q_1'}{n_1}+\frac{p_2'q_2'}{n_2}},\ \left(p_1'-p_2'\right)+1.96\sqrt{\frac{p_1'q_1'}{n_1}+\frac{p_2'q_2'}{n_2}}$$

* 98%, $\alpha=0.02$, $z_{\alpha/2}=2.326$.
* 99%, $\alpha=0.01$, $z_{\alpha/2}=2.58$

⇒ CI for (ρ) correlation (coe) when (ρ) is normal =

From the theory of sampling distri~ we know that the statistic,

$$t=\frac{r-\rho}{\sqrt{\frac{1-r^2}{n-2}}} \longrightarrow t \text{ distri~ with } (n-2) \text{ df}$$

where the (ρ) is normal & sample size 'n' is small.

$$P\left\{-t_{\alpha/2} \le t \le t_{\alpha/2}\right\} = 1-\alpha$$

$$P\left\{-t_{\alpha/2} \le \frac{r-\rho}{\sqrt{\frac{1-r^2}{n-2}}} \le t_{\alpha/2}\right\} = 1-\alpha$$

$$P\left\{-t_{\alpha/2}\sqrt{\frac{1-r^2}{n-2}} \le r-\rho \le t_{\alpha/2}\sqrt{\frac{1-r^2}{n-2}}\right\} = 1-\alpha$$

$$P\left\{-r-t_{\alpha/2}\sqrt{\frac{1-r^2}{n-2}} \le -\rho \le -r+t_{\alpha/2}\sqrt{\frac{1-r^2}{n-2}}\right\} = 1-\alpha$$

$$P\left\{r-t_{\alpha/2}\sqrt{\frac{1-r^2}{n-2}} \le \rho \le r+t_{\alpha/2}\sqrt{\frac{1-r^2}{n-2}}\right\} = 1-\alpha$$

$$\left[r-t_{\alpha/2}\sqrt{\frac{1-r^2}{n-2}},\ r+t_{\alpha/2}\sqrt{\frac{1-r^2}{n-2}}\right] \longrightarrow 100(1-\alpha)\%.$$

$t_{\alpha/2}$ is obtained by referring t table for $(n-2)$ df & (prob) α.

Q

1) For a bernouli trials with the prob of success p, $\sum_{i=1}^{n} x_i$, the no. of success in n trials is a sufficient (e) for p. (written)

2) Examine sufficiency of $\sum_{i=1}^{n} x_i^2$ for σ^2 in $N(0,\sigma^2)$ distri~ (written)

3) P.T for a binomial distri~ with density (), $f(x,p) = nC_x p^x q^{n-x}$, $x=0,1,2 \ldots n$, the max likelyhood (e) for p is x/n.

A) Since $f(x,p)$ is the (prob) out of n trials, it itself is the L. ().

$q \to 1-p$

$$L = nC_x\, p^x q^{n-x}.$$

$$\log L = \log nC_x + \log p^x + \log q^{n-x}.$$

$$= \log nC_x + x\log p + (n-x)\log(1-p)$$

$\log x \to \frac{1}{x}$

$$\frac{\partial \log L}{\partial p} = \frac{x}{p} + (n-x)\times\frac{1}{1-p}\times -1$$

$$= \frac{x}{p} - \frac{(n-x)}{1-p}$$

$$\frac{\partial \log L}{\partial p} = 0 \Rightarrow \frac{x}{p} - \frac{(n-x)}{1-p} = 0$$

$$\hat{p} = \frac{x}{n}.$$

$$\frac{\partial^2 \log L}{\partial p^2} = -\frac{x}{p^2} - \frac{(n-x)}{(1-p)^2} < 0$$

2 fo -ue ഞ്ഞങ്ങൾ , Aus ~ -ve.

2) for a Bernoulli's trials with a (prob) of success p, $\sum_{i=1}^{n} x_i$ the no. of success in the n trials is a sufficient (e) of p.

A) $(x_1, x_2, \ldots, x_n \sim B(p))$ with parameter p.

ഒരു r.s -ലെ ഓരോന്നും [illegible] bernouli ആണ്

$$f(x_i, p) = p^{x_i} q^{1-x_i}$$
$$= p^{x_i}(1-p)^{1-x_i} \qquad x_i = 0, 1$$
$$i = 1, 2 \ldots n.$$

To prove $\Rightarrow \sum_{i=1}^{n} x_i \to p$.

$(\Leftarrow \to p)$ $L(x_1, x_2 \ldots x_n; p) = L_1(t, p) \; L_2(x_1, x_2 \ldots x_n)$

$$L(x_1, x_2 \ldots x_n; p) = f(x_1, x_2 \ldots x_n; p)$$
$$= \prod_{i=1}^{n} f(x_i, p).$$
$$= p^{x_1}(1-p)^{1-x_1} \times p^{x_2}(1-p)^{1-x_2} \ldots \times p^{x_n}(1-p)^{1-x_n}$$
$$= p^{x_1+x_2+\ldots x_n}(1-p)^{\underbrace{(1-x_1)+(1-x_2)\ldots(1-x_n)}}$$

1 -ന്നെ -ന്നും n times x വരുന്നു.

$$= p^{\sum x_i}(1-p)^{n-(x_1+x_2\ldots x_n)}$$
$$= p^{\sum x_i}(1-p)^{n-\sum x_i}$$
$$\cancel{= L\left(\sum_{i=1}^{n} x_i, p\right)}$$

(ഇവിടെ L_2 ഒന്നുമില്ലാത്തതിനാൽ ഉള്ളതാണ്,) x1 ആവശ്യം,

(i.e) $\cancel{= L\left(\sum_{i=1}^{n} x_i, p\right) \times 1}$

$$= p^{\sum x_i}(1-p)^{n-\sum x_i} \times 1$$
$$= L_1\left(\sum_{i=1}^{n} x_i, p\right) \; L_2(x_1, x_2 \ldots x_n)$$

$\therefore$ $\sum x_i$ is s. (e) for p.

3) S.T $\bar{x}$ is sufficient (e) of p when samples of size are taken from a binomial (p) with paramtrs N & p (N gvn).

x) $x_1, x_2 \dots x_n \sim B(N, p)$ $\bar{x} = \frac{1}{n} \Sigma x_i$

$x_i \sim B(N, p)$.

$N \to$ given (i.e) known

$p \to$ unknown.

$L(x_1, x_2 \dots x_n; \theta) = L_1(t, \theta)\ L_2(x_1, x_2 \dots x_n)$

(i.e) $L(x_1, x_2 \dots x_n; p) = L_1(\bar{x}, p)\ L_2(x_1, x_2 \dots x_n)$

$f(x, p) = N C_x\, p^{x_i} (1-p)^{N-x_i}$..

$L(x_1, x_2 \dots x_n; p) = f(x_1, x_2 \dots x_n; p)$

$= \prod_{i=1}^{n} f(x_i, p)$

$= N C_{x_1} p^{x_1} (1-p)^{N-x_1}$.

$N C_{x_2} p^{x_2} (1-p)^{N-x_2}$ ....

$N C_{x_n} p^{x_n} (1-p)^{N-x_n}$

$n\bar{x} = \Sigma x_i$. $= p^{x_1 + x_2 \dots + x_n} (1-p)^{\underbrace{N - x_1 + N - x_2 \dots N - x_n}_{n \text{ times} \to nN}}$

$N C_{x_1} N C_{x_2} \dots N C_{x_n}$

$= p^{n\bar{x}} (1-p)^{nN - (x_1 + x_2 + \dots x_n)}$.

$N C_{x_1} N C_{x_2} \dots N C_{x_n}$

$= \underbrace{p^{n\bar{x}} (1-p)^{nN - n\bar{x}}}\ \underbrace{N C_{x_1} N C_{x_2} \dots N C_{x_n}}$

$= L(\bar{x}, p) \cdot L_2(x_1, x_2 \dots x_n)$

$\therefore \bar{x}$ is S.(E) of p