T-estimation

7.2 point estimation

① point estimator

(Def 7.1) point estimator used by sample statistics to estimate ppl parameter. Such as $\bar{x}, S, \hat{p}$

① Estimator of σ^2

$(\hat\sigma^2 =)\ S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$ property $\begin{cases} E(S^2)=\sigma^2 & \text{unbiased for } \sigma^2. \\ Var(S^2) \end{cases}$

② Estimator of $\hat{p}$

$\hat{p} = \frac{x}{n}$ property $\begin{cases} E(\hat{p}) = p \\ Var(\hat{p}) = \frac{p(1-p)}{n}. \end{cases}$

7.3 Confidence Intervals (CI ppl)

① for μ:

$P[-z_{\frac{\alpha}{2}} \le \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \le z_{\frac{\alpha}{2}}] = 1-\alpha$ $100(1-\alpha)\%$ confidence interval for μ: $(\bar{x} - z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}})$

property: ① Validity of this confidence interval: $\begin{cases} \text{exact normal distribution} \\ \text{normal approximation} \end{cases}$

② one-sided C.I: $P[\frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \le z_\alpha] = 1-\alpha \to (\bar{x} - z_\alpha\frac{\sigma}{\sqrt{n}}, \infty)$

$P[-z_\alpha \le \frac{\bar{x}-\mu}{\sigma/\sqrt{n}}] = 1-\alpha \longrightarrow \mu \le \bar{x} + z_\alpha\frac{\sigma}{\sqrt{n}}$

③ $100(1-\alpha)\%$ CI for unequal tails: $\bar{x} - z_{\alpha_1}\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x} + z_{\alpha_2}\frac{\sigma}{\sqrt{n}}$, $\alpha_1+\alpha_2=\alpha$

④ width $= 2z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$ margin of error: $\frac{1}{2}$width $= z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$,

⑤ $100(1-\alpha)\%$ prediction/probability interval for $\bar{x}$: $\mu - z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \le \bar{x} \le \mu + z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$

② for σ^2 & σ:

$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1) \to P[\chi^2_{1-\frac{\alpha}{2}}(n-1) \le \frac{(n-1)S^2}{\sigma^2} \le \chi^2_{\frac{\alpha}{2}}(n-1)] = 1-\alpha$

CLT doesn't apply to S^2

$100(1-\alpha)\%$ CI for σ^2 or σ:

$\left(\frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}(n-1)} \le \sigma^2 \le \frac{(n-1)S^2}{\chi^2_{(1-\frac{\alpha}{2})}(n-1)}\right)$ $\left(\sqrt{\frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}} \le \sigma \le \sqrt{\frac{(n-1)S^2}{\chi^2_{(1-\frac{\alpha}{2})}(n-1)}}\right)$

③ for μ (σ unknown)

$\frac{\bar{x}-\mu}{S/\sqrt{n}} = \frac{\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{(n-1)S^2}{\sigma^2}/n-1}} \overset{\text{thm 6.4}}{=} t(n-1)$

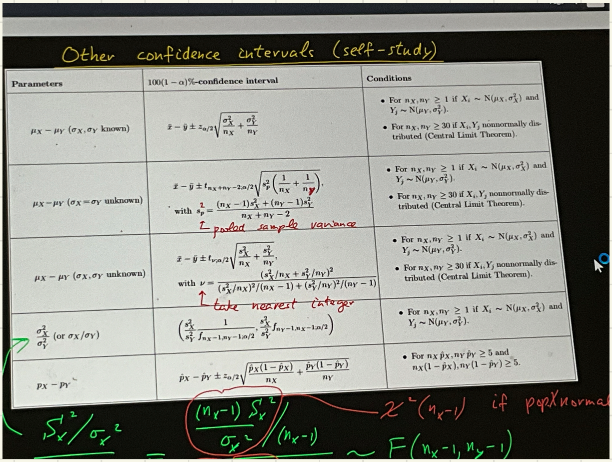
$\Rightarrow 100(1-\alpha)\%$ CI for μ: $(\bar{x} - t_{n-1,\frac{\alpha}{2}}\frac{S}{\sqrt{n}}, \bar{x} + t_{n-1,\frac{\alpha}{2}}\frac{S}{\sqrt{n}})$

④ for $\hat{p}$:

$P[-z_{\frac{\alpha}{2}} \le \frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}} \le z_{\frac{\alpha}{2}}] \approx 1-\alpha$ $100(1-\alpha)\%$ CI for p: $\hat{p} - z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \le p \le \hat{p} + z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

$2z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \le 2B$, $n \ge \frac{\hat{p}(1-\hat{p})}{B^2}(z_{\frac{\alpha}{2}})^2$ $n \ge (\frac{z_{\frac{\alpha}{2}}}{2B})^2$

7.4 Confidence Intervals (2 ppl)



Other confidence intervals (self-study)

Parameters	$100(1-\alpha)\%$-confidence interval	Conditions
$\mu_X - \mu_Y$ (σ_X, σ_Y known)	$\bar{x} - \bar{y} \pm z_{\alpha/2}\sqrt{\frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}}$	• For $n_X, n_Y \ge 1$ if $X_i \sim N(\mu_X, \sigma_X^2)$ and $Y_j \sim N(\mu_Y, \sigma_Y^2)$. • For $n_X, n_Y \ge 30$ if X_i, Y_j nonnormally distributed (Central Limit Theorem).
$\mu_X - \mu_Y$ ($\sigma_X = \sigma_Y$ unknown)	$\bar{x} - \bar{y} \pm t_{n_X+n_Y-2;\alpha/2}\sqrt{s_p^2\left(\frac{1}{n_X} + \frac{1}{n_Y}\right)}$, with $s_p^2 = \frac{(n_X-1)s_X^2 + (n_Y-1)s_Y^2}{n_X+n_Y-2}$ ↑ pooled sample variance	• For $n_X, n_Y \ge 1$ if $X_i \sim N(\mu_X, \sigma_X^2)$ and $Y_j \sim N(\mu_Y, \sigma_Y^2)$. • For $n_X, n_Y \ge 30$ if X_i, Y_j nonnormally distributed (Central Limit Theorem).
$\mu_X - \mu_Y$ (σ_X, σ_Y unknown)	$\bar{x} - \bar{y} \pm t_{\nu;\alpha/2}\sqrt{\frac{s_X^2}{n_X} + \frac{s_Y^2}{n_Y}}$, with $\nu = \frac{(s_X^2/n_X + s_Y^2/n_Y)^2}{(s_X^2/n_X)^2/(n_X-1) + (s_Y^2/n_Y)^2/(n_Y-1)}$ ↑ take nearest integer	• For $n_X, n_Y \ge 1$ if $X_i \sim N(\mu_X, \sigma_X^2)$ and $Y_j \sim N(\mu_Y, \sigma_Y^2)$. • For $n_X, n_Y \ge 30$ if X_i, Y_j nonnormally distributed (Central Limit Theorem).
$\frac{\sigma_X^2}{\sigma_Y^2}$ (or σ_X/σ_Y)	$\left(\frac{s_X^2}{s_Y^2}\frac{1}{f_{n_X-1,n_Y-1;\alpha/2}}, \frac{s_X^2}{s_Y^2}f_{n_Y-1,n_X-1;\alpha/2}\right)$	• For $n_X, n_Y \ge 1$ if $X_i \sim N(\mu_X, \sigma_X^2)$ and $Y_j \sim N(\mu_Y, \sigma_Y^2)$.
$p_X - p_Y$	$\hat{p}_X - \hat{p}_Y \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_X(1-\hat{p}_X)}{n_X} + \frac{\hat{p}_Y(1-\hat{p}_Y)}{n_Y}}$	• For $n_X\hat{p}_X, n_Y\hat{p}_Y \ge 5$ and $n_X(1-\hat{p}_X), n_Y(1-\hat{p}_Y) \ge 5$.

$S_X^2/\sigma_X^2 = \frac{(n_X-1)S_X^2}{\sigma_X^2}\Big/(n_X-1) \sim F(n_X-1, n_Y-1)$

$\frac{(n_X-1)S_X^2}{\sigma_X^2} \sim \chi^2(n_X-1)$ if pop X normal

paired sample: (dependent)

① $W_i = x_i - y_i$ $(i = 1, \dots, n)$

② $W_1, W_2, W_3, \dots, W_n \longrightarrow$ i.i.d.

③ $\bar{W} - t_{n-1,\frac{\alpha}{2}}\frac{S_W}{\sqrt{n}} \le \mu_W \le \bar{W} + t_{n-1,\frac{\alpha}{2}}\frac{S_W}{\sqrt{n}}$