No. ✓

Date

$E(s'^2) = \frac{n-1}{n}\sigma^2 \neq \sigma^2 \rightarrow$ biase.

$E(s^2) = \frac{n-1}{n-1}\sigma^2 = \sigma^2 \rightarrow$ unbiase.

N is large enough if $n > 30$ ($n_+ > 15$ & $n_- > 15$)

The error of estimation: $\varepsilon = |\hat{\theta} - \theta|$. $\hat{\theta}$ RV

Prob that $\mu - 2\sigma < Y < \mu + 2\sigma$.

Distribution	Probability.
Normal	0.9544
Uniform	1.0
Exponential	0.9502.

Given $b = 2\sigma_{\hat{p}} = 2\sqrt{\frac{pq}{n}}$ ($\boxed{\hat{p}}$ unbiase $\rightarrow \sigma_p = \sqrt{\frac{pq}{n}}$).

the prob that $b < 2\sqrt{\frac{pq}{n}} \Rightarrow 0.95$.

interval estimator: A rule specifying the method to calculate two numbers that form the endpoints of the interval.

Confidence coefficient: The probability that a (random) ~~conf~~ confidence interval will enclose θ is called the confidence coefficient.

No.

Date

confidence coefficient identifies the fraction of the time, in repeated sampling, that the intervals constructed will contain the target parameter θ.

if $P(\hat{\theta}_L \le \theta \le \hat{\theta}_u) = \boxed{1-\alpha} \rightarrow$ confidence coefficient.

Two side confidence interval: $P(\theta_L \le \theta \le \theta_u)$.

One side confidence interval: $P(\theta_L \le \theta)$ or $P(\theta \le \theta_u)$.

check pivotal method at this page. P411 - P414.

We use $\hat{\theta} \pm Z_{\frac{\alpha}{2}} \cdot \sigma_{\hat{\theta}}$ for large sample est.

e.g. $\hat{P}_1 - \hat{P}_2 \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{P_1 q_1}{n_1} + \frac{P_2 q_2}{n_2}}$. ($q_2 = 1 - P_2$). maybe.

$\bar{Y} \pm Z_{\frac{\alpha}{2}} \left(\frac{\sigma}{\sqrt{n}}\right) \approx \bar{Y} \pm Z_{\frac{\alpha}{2}} \left(\frac{S}{\sqrt{n}}\right)$.

For small interval: $P(-t_{\frac{\alpha}{2}} \le T \le t_{\frac{\alpha}{2}})$

two side: $\bar{Y} \pm t_{\frac{\alpha}{2}} \left(\frac{S}{\sqrt{n}}\right) \leftarrow (\sigma_n)$

one side: $P(T \le t_\alpha) = 1-\alpha$.

$\frac{Y-\mu}{\frac{S}{\sqrt{n}}} \le t_\alpha$ $\quad Y-\mu \le \frac{S\, t_\alpha}{\sqrt{n}}$ $\quad Y - \frac{S\, t_\alpha}{\sqrt{n}} \le \mu$.

No.

Date

So $P(\bar{Y} - t_\alpha(\frac{S}{\sqrt{n}}) \le \mu) = 1-\alpha$.

large interval: $Z = \dfrac{(\bar{Y}_1 - \bar{Y}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$

Construct common variance S^2:

Given $Y_{11}, Y_{12}, Y_{13}, \cdots, Y_{1n}, Y_{21}, Y_{22}, Y_{23}, \cdots, Y_{2n}$

$\bar{Y}_1 = \frac{1}{n_1} \sum_{i=1}^{n} Y_{1i} \qquad \bar{Y}_2 = \frac{1}{n} \sum_{i=1}^{n} Y_{2i}$

Pooled square estimator:

$$S_p^2 = \frac{\sum_{i=1}^{n} (Y_{1i} - \bar{Y}_1)^2 + \sum_{i=1}^{n_2} (Y_{2i} - \bar{Y}_2)^2}{n_1 + n_2 - 2} = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}$$

$$W = \frac{(n_1 + n_2 - 2) S_p^2}{\sigma^2} = \frac{\sum_{i=1}^{n} (Y_{1i} - \bar{Y}_1)^2}{\sigma^2} + \frac{\sum_{i=1}^{n_2} (Y_{2i} - \bar{Y}_2)^2}{\sigma^2}$$

is the sum of 2 independent χ^2-dist RV with (n_1-1), (n_2-1) df.

df of $W = n_1 + n_2 - 2$ so

$$T = \frac{Z}{\sqrt{\frac{W}{V}}} = \left[\frac{(\bar{Y}_1 - \bar{Y}_2) - (\mu_1 - \mu_2)}{\sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \right] \Big/ \sqrt{\frac{(n_1+n_2-2) S_p^2}{\sigma^2 (n_1+n_2-2)}}$$

$$= \frac{(\bar{Y}_1 - \bar{Y}_2) - (\mu_1 - \mu_2)}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$ with $(n_2 + n_1 - 2)$ df

No.

Date

So for $\mu_1 - \mu_2$:

$$(\bar{Y}_1 - \bar{Y}_2) \pm t_{\frac{\alpha}{2}} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \quad \left(\begin{matrix} \text{independent } \checkmark \\ \sigma_1^2 = \sigma_2^2 \end{matrix} \right).$$

Estimate confidential interval for σ^2

Given $\quad S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (Y_i - \bar{Y})^2$

then $\quad \chi^2 \in \frac{(n-1)S^2}{\sigma^2}$

$$P\left[\chi_L^2 \leq \frac{(n-1)S^2}{\sigma^2} \leq \chi_U^2 \right] = 1-\alpha.$$

$$P\left[\chi^2_{1-\frac{\alpha}{2}} \leq \frac{(n-1)S^2}{\sigma^2} \leq \chi^2_{\frac{\alpha}{2}} \right] = 1-\alpha.$$

$$P\left[\frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}} \right] = 1-\alpha.$$

So $100(1-\alpha)\%$ CI for σ^2: $\sigma^2 \in \left[\frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}}, \frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}} \right]$

Sufficient statistic: an unbiased estimator with small variance is or can be made to be a function of a sufficient statistic.

Efficiency: Given two unbiased estimator $\hat{\theta}_1, \hat{\theta}_2$ of θ, note Var of $\hat{\theta}_1, \hat{\theta}_2$ as $V(\hat{\theta}_1), V(\hat{\theta}_2)$, then efficiency of $\hat{\theta}_1$ relative to $\hat{\theta}_2$, $\text{eff}(\hat{\theta}_1, \hat{\theta}_2)$

$$\text{eff}(\hat{\theta}_1, \hat{\theta}_2) = \frac{V(\hat{\theta}_2)}{V(\hat{\theta}_1)}$$

No.

Date

if $eff(\hat{\theta_1}, \hat{\theta_2}) < 1 \rightarrow$ Pick $\hat{\theta_2}$.

else $> 1 \rightarrow$ Pick $\hat{\theta_1}$.

Consistency: $P(|\frac{Y}{N} - p| \leq \varepsilon) = 1$ When $n \to +\infty$.

if true, $\frac{Y}{N}$ is a consistent estimator of P. /

$\frac{Y}{N}$ converges in probability to p.

經过大量实验结果 → 某个P值

estimator $\hat{\theta_n}$ is $\boxed{\text{consistent}}$ estimator if.

$\lim_{n\to\infty} P(|\hat{\theta_n} - \theta| \leq \varepsilon) = 1$

or $\lim_{n\to\infty} P(|\hat{\theta_n} - \theta| > \varepsilon) = 0$.

$\lim_{n\to\infty} V(\hat{\theta_n}) = 0 \leftarrow$ unbiased $\hat{\theta_n}$ consistent.

Proof: Let Y be any independent var. with

$E(Y) = \mu$ & $V(Y) = \sigma^2 < \infty$

By Tchebysheff's theorem:

$P(|Y - \mu| > k\sigma) \leq \frac{1}{k^2}$

if θ unbiase $\rightarrow E(\hat{\theta_n}) = \theta$.

So $P(|\hat{\theta} - \theta| > k\sigma_{\hat{\theta_n}}) \leq \frac{1}{k^2}$

let $k = \frac{\varepsilon}{\sigma_{\theta_n}}$

$P(|\hat{\theta_n} - \theta| > \varepsilon) = P\left(|\hat{\theta_n} - \theta| > \left[\frac{\varepsilon}{\sigma_{\hat{\theta_n}}}\right]\sigma_{\hat{\theta_n}}\right) \leq \frac{1}{\left(\frac{\varepsilon}{\sigma_{\hat{\theta_n}}}\right)^2} = \frac{V(\hat{\theta_n})}{\varepsilon^2}$