Def

A $100(1-\alpha)\%$ confidence interval (CI) for a parameter θ is a random interval $[L,R]$ such that $P(L \leq \theta \leq R) = 1-\alpha$ regardless of the value of θ, where α is called confidence level.

Ex

A chocolate bar's length is denoted by $X \sim N(\mu, 0.5^2)$

Obervations: 13.07, 13.28, 12.36, 13.09, 14.12, 13.11, 12.11, 13.08 , $n=8$

Find a 95% CI for μ.

$$X_i \sim N(\mu, \sigma^2), \quad \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \sim N\left(\mu, \frac{\sigma^2}{n}\right), \quad \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

$$P[L \leq \mu \leq R] = 0.95$$
$$P[L-\overline{X} \leq \mu-\overline{X} \leq R-\overline{X}] = 0.95$$
$$P[\overline{X}-R \leq \overline{X}-\mu \leq \overline{X}-L] = 0.95$$
$$P\left[\frac{\overline{X}-R}{\sigma/\sqrt{n}} \leq \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \leq \frac{\overline{X}-L}{\sigma/\sqrt{n}}\right] = 0.95$$

Standard, normal, distribution

We now want to find L such that: 1. $P\left[Z \geq \frac{\overline{X}-L}{\sigma/\sqrt{n}}\right] = \frac{\alpha}{2} = 0.025$

2. $\frac{\overline{X}-L}{\sigma/\sqrt{n}} = -\frac{\overline{X}-R}{\sigma/\sqrt{n}}$

$Q(p) = \min\{x : F(x) \geq p\}$ $\quad p \in (0,1)$

$Z_{\frac{\alpha}{2}} = Z_{0.025} = \frac{\overline{X}-L}{\sigma/\sqrt{n}} \Rightarrow L = \overline{X} - Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$

$R = \overline{X} + Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$

$\overline{X} = 13.02$, $\sigma = 0.5$, $n=8$, $Z_{\frac{\alpha}{2}} = 1.96$

CI: $(12.6, 13.37)$

Proposition

$X_1, \ldots, X_n \sim N(\mu, \sigma^2)$, μ is unknown. Then $(1-\alpha)\cdot 100\%$ CI is given by $\left(\overline{X} \pm Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}\right)$.

Theorem (Central limit theorem)

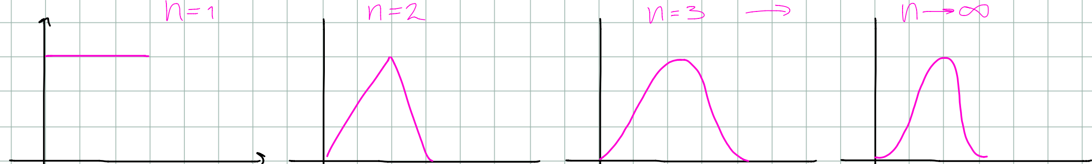
If $X_1 \ldots X_n$ are independent and identically distributed as X, where the distribution of X is "decent" ($E[X] < \infty$, $Var[X] < \infty$), then $\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$

Note: $\sim$ approximately

Ex

becomes more and more normally distributed

$n=1$ $\quad$ $n=2$ $\quad$ $n=3$ $\longrightarrow$ $\quad$ $n \to \infty$



Proposition

Let $X_1 \ldots X_n$ be a random sample with $\sigma^2 = Var[X_1]$. Then $(1-\alpha)100\%$ CI is given by: $\left(\overline{X} \pm Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}\right)$

Find σ^2, μ

$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$ is unbiased estimator of σ^2. $\sqrt{S^2}$ is a biased estimator for σ.

$X_i \sim N(\mu, \sigma^2)$, n is small

Proposition

If $X_1 \ldots X_n$ is a sample from a $N \sim (\mu, \sigma^2)$, $\frac{\overline{X}-\mu}{S/\sqrt{n}} \sim T_{n-1}$, called t or student, distribution with $n-1$ degrees of freedom.

Proposition

$X_1, \dots X_n \sim N(\mu, \sigma^2)$, μ, σ^2 - unknown, (n is small: <30). Then $(1-\alpha)100\%$ CI is given by: $\left(\overline{X} \pm t_{\frac{\alpha}{2}, n-1} \cdot \sqrt{\frac{S^2}{n}}\right)$

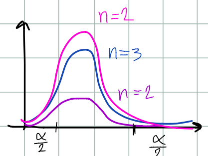
For VLE: Inverse T
n-1 df
correct tail
$\frac{\alpha}{2}$

Exact formula: $f(x) = \frac{\Gamma(n+1)}{2\sqrt{\pi n}\,\Gamma(\frac{n}{2})} \cdot \left(1 + \frac{x^2}{n}\right)^{-\frac{(n+1)}{2}}$

Find σ^2

Proposition

If $X_1, \dots, X_n \sim N(\mu, \sigma^2)$ and are independent $\frac{(n-1)S^2}{\sigma^2} = \sum_{i=1}^{n} \frac{(X_i - \overline{X})^2}{\sigma^2} \sim \chi_{n-1}$ - Chi-squared with n-1 degrees of freedom.



$P[L \leq \sigma^2 \leq R] = 1-\alpha \longrightarrow \frac{(n-1)S^2}{\sigma^2} \longrightarrow$ Table

Proposition

$X_1, \dots, X_n \sim N(\mu, \sigma^2)$, μ, σ^2 - unknown. Then $(1-\alpha)100\%$ CI is given by: $\left(\frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}, n-1}}, \frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}, n-1}}\right)$

Summation

$\left(\overline{X} \pm z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{n}\right)$. Normal, (0, 1)

$\left(\overline{X} \pm t_{\frac{\alpha}{2}, n-1} \cdot \sqrt{\frac{S^2}{n}}\right)$ t-distr; n-1 df

$\left(\frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}, n-1}}, \frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}, n-1}}\right)$ Chi-square; n-1 df