Lec 1/30

Monday, January 30, 2017 15:02

ch 11 Interval estimation §11.1 & §11.2 estimation of means

Def: Confidence Interval:

if $\hat{\theta}_1, \hat{\theta}_2$ are random variables (usually statistics of a RS $X_1, \dots, X_n$)

such that $P(\hat{\theta}_1 < \theta < \hat{\theta}_2) = 1-\alpha$

then the interval $(\hat{\theta}_1, \hat{\theta}_2)$ is a $(1-\alpha)\times 100\,\%$ $\overbrace{\text{confidence interval}}^{\text{CI}}$ for θ.

Note: CIs based on sample data, give a range of possible values for θ.

Consider estimation of means

Goal: from a random sample, obtain an estimate of a range of values for the mean μ of a pop.

Ex: Dog food production line.

Boxes filled w/ dog biscuits, estimate mean weight of a box.

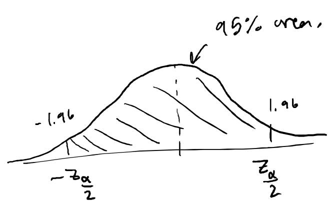
Take a RS of 100 boxes $X_1, \dots, X_{100}$, find $\bar{X} = \frac{\sum_{i=1}^{100} X_i}{100} = 1.15$ lbs

Since we expect variability from sample to sample, it's unlikely that $\mu = 1.15$.

we need to include an estimate of variability.

we know $Var(\bar{X}) = \frac{Var(X_i)}{100} = \frac{\sigma^2}{100}$. Spoze for now that σ^2 is known.

Central Limit Theorem: $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$ as $n \to \infty$



$$\int_{-1.96}^{1.96} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}\, dx = 0.95$$

$$\Rightarrow P\left(-1.96 < \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < 1.95\right) \approx 0.95$$

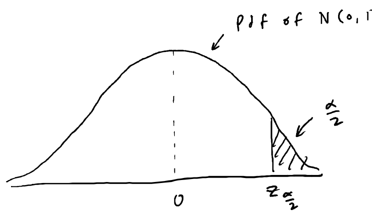
$$-1.96 < \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < 1.96$$

$$\Leftrightarrow \quad \underbrace{\overline{X} - 1.96\frac{\sigma}{\sqrt{n}}}_{\hat{\theta}_1} < \mu < \underbrace{\overline{X} + 1.96\frac{\sigma}{\sqrt{n}}}_{\hat{\theta}_2} \qquad \text{so CI is } (\hat{\theta}_1, \hat{\theta}_2)$$

$(\hat{\theta}_1, \hat{\theta}_2)$ has about a 95% chance of containing μ. (approx CI).

iff X_i s are normal, this is exact.

Thm 11.2 If $\overline{X}$ is mean of RS of size n from a normal population ($X_i \overset{iid}{\sim}$ Normal) with a known σ^2, then $\left(\overline{X} - z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}, \overline{X} + z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right)$ is a $1-\alpha \times 100\%$ confidence interval for the mean of the population. $\left(\overline{X} \pm z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right)$



90% CI $\Rightarrow \alpha = 0.1 \Rightarrow z_{\frac{\alpha}{2}} = 1.65$

99% CI $\Rightarrow \alpha = 0.01 \Rightarrow z_{\frac{\alpha}{2}} = 2.58$

95% CI $\Rightarrow \alpha = 0.05 \Rightarrow z_{\frac{\alpha}{2}} = 1.96$

68% CI $\Rightarrow \alpha = 0.32 \Rightarrow z_{\frac{\alpha}{2}} = 1$

Back to example: $\overline{X} = 1.15$ lbs, assume $\sigma = 0.17$ lbs.

95% CI $= \left(1.15 \pm 1.96 \cdot \frac{0.17}{10}\right) = \left(1.15 \pm 1.96 \cdot 0.017\right) = (1.12, 1.18)$

↳ should finish calculation w/ calculator on quiz/exam

__Remark:__ Once you use values of sample, CI is not a RV anymore. μ is either in CI or not. (not a 95% chance).

95% of intervals constructed in this way contain μ.

Before collecting data:

- Probability statement
- interval $\left(\bar{X} \pm z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right)$ is random

After collecting data:

- Confidence statement
- how often is the method successful?