Confidence Intervals for One Population

- **Estimating the population mean (μ) when σ is <u>known</u>.**

  1. Use $Z = \dfrac{\bar{X} - \mu}{\sigma/\sqrt{n}}$. Z follows the $N(0, 1)$.

  2. The $(1 - \alpha)100\%$ confidence interval for μ is $\left[\bar{X} - z_{\frac{\alpha}{2}} \dfrac{\sigma}{\sqrt{n}}, \bar{X} + z_{\frac{\alpha}{2}} \dfrac{\sigma}{\sqrt{n}}\right]$.

  a. The 90% C.I. for μ is $\left[\bar{x} - 1.645 \dfrac{\sigma}{\sqrt{n}}, \bar{x} + 1.645 \dfrac{\sigma}{\sqrt{n}}\right]$.

  b. The 95% C.I. for μ is $\left[\bar{x} - 1.96 \dfrac{\sigma}{\sqrt{n}}, \bar{x} + 1.96 \dfrac{\sigma}{\sqrt{n}}\right]$.

  c. The 99% C.I. for μ is $\left[\bar{x} - 2.576 \dfrac{\sigma}{\sqrt{n}}, \bar{x} + 2.576 \dfrac{\sigma}{\sqrt{n}}\right]$.

  3. The *Margin of Error*, $M = Z_{\frac{\alpha}{2}} \dfrac{\sigma}{\sqrt{n}}$.

  4. For a specified margin of error M, the required sample size is $n = \left(\dfrac{Z_{\frac{\alpha}{2}}\sigma}{M}\right)^2$.

- **One-Sided Confidence Intervals when σ is <u>known</u>.**

  1. A *large-sample upper confidence bound for* μ is

$$\mu < \bar{x} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

  2. A *large-sample lower confidence bound for* μ is

$$\mu > \bar{x} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

- **Estimating the population mean (μ) when σ is <u>unknown</u>.**

  1. Use $T = \dfrac{\bar{X} - \mu}{s/\sqrt{n}}$. T follows the student t-distribution with $n - 1$ degrees of freedom.

  2. The $(1 - \alpha)100\%$ confidence interval for μ is $\left[\bar{x} - (t_{\{\frac{\alpha}{2},(n-1)\}}) \dfrac{s}{\sqrt{n}}, \bar{x} + (t_{\{\frac{\alpha}{2},(n-1)\}}) \dfrac{s}{\sqrt{n}}\right]$.

  3. The *Margin of Error*, $M = (t_{\{\frac{\alpha}{2},(n-1)\}}) \dfrac{s}{\sqrt{n}}$.

- **Prediction Interval.** A *prediction interval* (PI) for a single observation to be selected from a normal population distribution is

$$\bar{x} - (t_{\{\frac{\alpha}{2},(n-1)\}}) s \sqrt{1 + \frac{1}{n}}$$

- **Tolerance Interval.** Let k be a number between 0 and 100. A *tolerance interval* for capturing at least $k\%$ of the values in a normal population distribution with a confidence level of 95% has the form

$$\bar{x} - (\textit{tolerance critical value}) \cdot s$$

- **Estimating the population proportion (p) when the sample size is large.**

  1. The unbiased estimator of p is the sample proportion $\hat{p} = \dfrac{Y}{n}$, where Y is the number of successes in the sample.
  Recall that if $Y \sim \text{bin}(n, p)$, then $E(Y) = np$ and $Var(Y) = np(1-p)$.
  Therefore, $E(\hat{p}) = p$ and $Var(\hat{p}) = \frac{p(1-p)}{n}$.

  2. The $(1-\alpha)100\%$ confidence interval for p is $\left[\hat{p} - (z_{\frac{\alpha}{2}})\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}, \hat{p} + (z_{\frac{\alpha}{2}})\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}\right]$.

  3. The *Margin of Error*, $M = (z_{\frac{\alpha}{2}})\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}} \quad \Rightarrow \quad n = \dfrac{z^2_{\alpha/2}\hat{p}(1-\hat{p})}{M^2} \leq \dfrac{z^2_{\alpha/2}(0.25)}{M^2}$.

- **Estimating the population proportion (p).**

  The $(1-\alpha)100\%$ confidence interval for p is

$$\left[\frac{\hat{p} + \frac{z^2_{\alpha/2}}{2n} - z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n} + \frac{z^2_{\alpha/2}}{4n^2}}}{1 + z^2_{\alpha/2}/n}, \frac{\hat{p} + \frac{z^2_{\alpha/2}}{2n} - z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n} + \frac{z^2_{\alpha/2}}{4n^2}}}{1 + z^2_{\alpha/2}/n}\right]$$

- **Estimating the population variance (σ^2) and population s.d. (σ).**

  1. Use the statistic, $X^2 = \dfrac{(n-1)S^2}{\sigma^2} = \dfrac{\sum(X_i - \bar{X})^2}{\sigma^2}$.
  X^2 follows the chi-squared (χ^2)) distribution with $(n-1)$ degrees of freedom.

  2. The $(1-\alpha)100\%$ confidence interval for σ^2 of a normal population is

$$\left[\frac{(n-1)s^2}{\chi^2_{\alpha/2, df=n-1}}, \frac{(n-1)s^2}{\chi^2_{1-\alpha/2, df=n-1}}\right]$$