Statistics

MATH 324

McGill University, Montréal, Canada

Fall 2018

In this section we will discuss:

(1) Principles of interval estimations or confidence intervals.

(2) Small-sample confidence intervals

(3) Large-sample confidence intervals

Sections 8.5-8.9

Introduction

- So far, we discussed methods of constructing point estimators for parameters of interest in parametric families (method of moments, MLE, ...)
- We also discussed how to investigate statistical properties of point estimators (bias, MSE, consistency, ...)
- Our next move would be to provide so-called an interval estimator for a parameter θ such that the interval would contain the true value of the parameter with certain confidence level.

Introduction

- So far, we discussed methods of constructing point estimators for parameters of interest in parametric families (method of moments, MLE, ...)
- We also discussed how to investigate statistical properties of point estimators (bias, MSE, consistency, ...)
- Our next move would be to provide so-called an interval estimator for a parameter θ such that the interval would contain the true value of the parameter with certain confidence level.

Introduction

- So far, we discussed methods of constructing point estimators for parameters of interest in parametric families (method of moments, MLE, ...)
- We also discussed how to investigate statistical properties of point estimators (bias, MSE, consistency, ...)
- Our next move would be to provide so-called an interval estimator for a parameter θ such that the interval would contain the true value of the parameter with certain confidence level.

Two-sided confidence intervals

- **Definition**:

  Given $\alpha \in (0, 1)$ and $\theta \in \Theta \subseteq \mathbb{R}$, the statistics $\hat{\theta}_{\mathsf{L}}$ and $\hat{\theta}_{\mathsf{U}}$ are called the lower and upper confidence limits, respectively, for a parameter θ if
  $$P(\hat{\theta}_{\mathsf{L}} \leq \theta \leq \hat{\theta}_{\mathsf{U}}) = 1 - \alpha.$$

- The random interval $[\hat{\theta}_{\mathsf{L}}, \hat{\theta}_{\mathsf{U}}]$ is called a two-sided confidence interval for θ, and $1 - \alpha$ is called the *confidence coefficient* or the *coverage probability*.

Two-sided confidence intervals

- Definition:

  Given $\alpha \in (0, 1)$ and $\theta \in \Theta \subseteq \mathbb{R}$, the statistics $\hat{\theta}_\mathrm{L}$ and $\hat{\theta}_\mathrm{U}$ are called the lower and upper confidence limits, respectively, for a parameter θ if
  $$P(\hat{\theta}_\mathrm{L} \leq \theta \leq \hat{\theta}_\mathrm{U}) = 1 - \alpha.$$

- The random interval $[\hat{\theta}_\mathrm{L}, \hat{\theta}_\mathrm{U}]$ is called a two-sided confidence interval for θ, and $1 - \alpha$ is called the *confidence coefficient* or the *coverage probability*.

One-sided confidence intervals

- Given $\alpha \in (0, 1)$ and $\theta \in \Theta \subseteq \mathbb{R}$, we could form one-sided confidence intervals using appropriate statistics $\hat{\theta}_{\mathsf{L}}$ and $\hat{\theta}_{\mathsf{U}}$:

  The interval $[\hat{\theta}_{\mathsf{L}}, \infty)$ is a lower one-sided confidence interval for θ if

  $$P(\hat{\theta}_{\mathsf{L}} \leq \theta) = 1 - \alpha$$

  The interval $(-\infty, \hat{\theta}_{\mathsf{U}}]$ is an upper one-sided confidence interval for θ if

  $$P(\theta \leq \hat{\theta}_{\mathsf{U}}) = 1 - \alpha.$$

McGill

One-sided confidence intervals

- Given $\alpha \in (0, 1)$ and $\theta \in \Theta \subseteq \mathbb{R}$, we could form one-sided confidence intervals using appropriate statistics $\hat{\theta}_{\mathrm{L}}$ and $\hat{\theta}_{\mathrm{U}}$:

  The interval $[\hat{\theta}_{\mathrm{L}}, \infty)$ is a lower one-sided confidence interval for θ if

$$P(\hat{\theta}_{\mathrm{L}} \leq \theta) = 1 - \alpha$$

  The interval $(-\infty, \hat{\theta}_{\mathrm{U}}]$ is an upper one-sided confidence interval for θ if

$$P(\theta \leq \hat{\theta}_{\mathrm{U}}) = 1 - \alpha.$$

Example 1

- Suppose that $X_1, X_2, \ldots, X_n$ are iid from $N(\mu, 1)$. We wish to construct a 95% confidence interval for the unknown parameter μ:

  1) Note that we have: $\sqrt{n}\,(\overline{X}_n - \mu) \sim N(0,1)$.

  2) Also, $P(-1.96 \leq Z \leq 1.96) = 0.95$, where $Z \sim N(0,1)$.

  3) Therefore, $P(-1.96 \leq \sqrt{n}\,(\overline{X}_n - \mu) \leq 1.96) = 0.95$.

- So, a 95% two-sided confidence interval (C.I.) for μ is given by

$$\left[\overline{X}_n - \frac{1.96}{\sqrt{n}}\ ,\ \overline{X}_n + \frac{1.96}{\sqrt{n}}\right].$$

Example 1

- Suppose that $X_1, X_2, \ldots, X_n$ are iid from $N(\mu, 1)$. We wish to construct a 95% confidence interval for the unknown parameter μ:

  1) Note that we have: $\sqrt{n}\,(\overline{X}_n - \mu) \sim N(0, 1)$.

  2) Also, $P(-1.96 \leq Z \leq 1.96) = 0.95$, where $Z \sim N(0, 1)$.

  3) Therefore, $P(-1.96 \leq \sqrt{n}\,(\overline{X}_n - \mu) \leq 1.96) = 0.95$.

- So, a 95% two-sided confidence interval (C.I.) for μ is given by

$$\left[\overline{X}_n - \frac{1.96}{\sqrt{n}}\ ,\ \overline{X}_n + \frac{1.96}{\sqrt{n}}\right].$$

Example 1

- Suppose that $X_1, X_2, \ldots, X_n$ are iid from $N(\mu, 1)$. We wish to construct a 95% confidence interval for the unknown parameter μ:

  1) Note that we have: $\sqrt{n}\,(\overline{X}_n - \mu) \sim N(0, 1)$.

  2) Also, $P(-1.96 \leq Z \leq 1.96) = 0.95$, where $Z \sim N(0, 1)$.

  3) Therefore, $P(-1.96 \leq \sqrt{n}\,(\overline{X}_n - \mu) \leq 1.96) = 0.95$.

- So, a 95% two-sided confidence interval (C.I.) for μ is given by

$$\left[\overline{X}_n - \frac{1.96}{\sqrt{n}}\ ,\ \overline{X}_n + \frac{1.96}{\sqrt{n}}\right].$$

Example 1: (one-sided confidence intervals for μ)

- Recall: if $Z \sim N(0,1)$, then $P(Z \geq 1.64) = P(Z \leq -1.64) = 0.95$.
- Therefore, we set

$$\begin{aligned} P(\sqrt{n}\,(\overline{X}_n - \mu) \leq 1.64) &= 0.95 \\ P(\sqrt{n}\,(\overline{X}_n - \mu) \geq -1.64) &= 0.95 \end{aligned}$$

- This results in the lower and upper 95% confidence intervals for μ:

$$\left[\overline{X}_n - \frac{1.64}{\sqrt{n}}\ ,\ +\infty\right)\ ,\ \left(-\infty\ ,\ \overline{X}_n + \frac{1.64}{\sqrt{n}}\right].$$

Example 1: (one-sided confidence intervals for μ)

- Recall: if $Z \sim N(0,1)$, then $P(Z \geq 1.64) = P(Z \leq -1.64) = 0.95$.
- Therefore, we set

$$\begin{aligned} P(\sqrt{n}\,(\overline{X}_n - \mu) \leq 1.64) &= 0.95 \\ P(\sqrt{n}\,(\overline{X}_n - \mu) \geq -1.64) &= 0.95 \end{aligned}$$

- This results in the lower and upper 95% confidence intervals for μ:

$$\left[\overline{X}_n - \frac{1.64}{\sqrt{n}}\ ,\ +\infty\right)\ ,\ \left(-\infty\ ,\ \overline{X}_n + \frac{1.64}{\sqrt{n}}\right].$$

The pivotal method for constructing confidence intervals

- Defintion:

  A *pivotal quantity* is a function of the data $X_1, X_2, \ldots, X_n$ and the parameter of interest θ, say,

  $$Q = Q(X_1, \ldots, X_n, \theta)$$

  whose distribution does not depend on θ.

- If we manage to construct such a function, then we find the quantiles of Q such that, for example,

  $$P(q_{1-\alpha/2} \leq Q \leq q_{\alpha/2}) = 1 - \alpha$$

  for a two-sided CI, and similarly for one-sided CI.

The pivotal method for constructing confidence intervals

- Defintion:

  A *pivotal quantity* is a function of the data $X_1, X_2, \ldots, X_n$ and the parameter of interest θ, say,

  $$Q = Q(X_1, \ldots, X_n, \theta)$$

  whose distribution does not depend on θ.

- If we manage to construct such a function, then we find the quantiles of Q such that, for example,

  $$P(q_{1-\alpha/2} \leq Q \leq q_{\alpha/2}) = 1 - \alpha$$

  for a two-sided CI, and similarly for one-sided CI.

...

- Then, solve the following inequalities to find a two-sided CI for θ:

$$q_{1-\alpha/2} \leq Q(X_1, \ldots, X_n, \theta) \leq q_{\alpha/2}.$$

And similarly for one-sided CI.

- In Example 1, a pivotal quantity is

$$Q(X_1, X_2, \ldots, X_n, \mu) = \sqrt{n}\,(\overline{X}_n - \mu).$$

- The fact that the distribution of Q does not depend on any unknown parameter is essential. It allows us to obtain appropriate quantiles $q_{1-\alpha/2}$ and $q_{\alpha/2}$.

...

- Then, solve the following inequalities to find a two-sided CI for θ:

$$q_{1-\alpha/2} \leq Q(X_1, \ldots, X_n, \theta) \leq q_{\alpha/2}.$$

  And similarly for one-sided CI.

- In Example 1, a pivotal quantity is

$$Q(X_1, X_2, \ldots, X_n, \mu) = \sqrt{n}\,(\overline{X}_n - \mu).$$

- The fact that the distribution of Q does not depend on any unknown parameter is essential. It allows us to obtain appropriate quantiles $q_{1-\alpha/2}$ and $q_{\alpha/2}$.

Example 2

- Let $X_1, X_2, \ldots, X_n$ be iid from $U(0, \theta)$, with the unknown parameter θ. Find a $100(1-\alpha)\%$ two-sided confidence interval for θ.

- Let $X_{(n)} = \max(X_1, X_2, \ldots, X_n)$. We know that,

$$P(X_{(n)} \leq y) = \left(\frac{y}{\theta}\right)^n, \quad 0 \leq y \leq \theta.$$

- Therefore, a pivotal quantity is:

$$Q(X_1, X_2, \ldots, X_n, \theta) = \left[\frac{X_{(n)}}{\theta}\right]^n.$$

- This is because:

Example 2

- Let $X_1, X_2, \ldots, X_n$ be iid from $U(0, \theta)$, with the unknown parameter θ. Find a $100(1-\alpha)\%$ two-sided confidence interval for θ.

- Let $X_{(n)} = \max(X_1, X_2, \ldots, X_n)$. We know that,

$$P(X_{(n)} \leq y) = \left(\frac{y}{\theta}\right)^n, \quad 0 \leq y \leq \theta.$$

- Therefore, a pivotal quantity is:

$$Q(X_1, X_2, \ldots, X_n, \theta) = \left[\frac{X_{(n)}}{\theta}\right]^n.$$

- This is because:

Example 2

- Let $X_1, X_2, \ldots, X_n$ be iid from $U(0, \theta)$, with the unknown parameter θ. Find a $100(1-\alpha)\%$ two-sided confidence interval for θ.

- Let $X_{(n)} = \max(X_1, X_2, \ldots, X_n)$. We know that,

$$P(X_{(n)} \leq y) = \left(\frac{y}{\theta}\right)^n \,, \;\; 0 \leq y \leq \theta.$$

- Therefore, a pivotal quantity is:

$$Q(X_1, X_2, \ldots, X_n, \theta) = \left[\frac{X_{(n)}}{\theta}\right]^n.$$

- This is because:

for any $x \in [0, 1]$,

$$P\left(\left[\frac{X_{(n)}}{\theta}\right]^n \leq x\right) = P\left(X_{(n)} \leq \theta x^{1/n}\right) = \left(\frac{\theta x^{1/n}}{\theta}\right)^n = x.$$

This implies that $Q(X_1, X_2, \dots, X_n, \theta) = \left[\frac{X_{(n)}}{\theta}\right]^n \sim U(0, 1)$.

- Next, we use Q to construct a CI for θ. Note that,

$$P\left(\frac{\alpha}{2} \leq Q(X_1, X_2, \dots, X_n, \theta) \leq 1 - \frac{\alpha}{2}\right) = 1 - \frac{\alpha}{2} - \frac{\alpha}{2} = 1 - \alpha.$$

- Solve the inequalities with respect to θ, we get a $100(1-\alpha)\%$ CI:

$$\left[\frac{X_{(n)}}{(1-\alpha/2)^{1/n}}, \frac{X_{(n)}}{(\alpha/2)^{1/n}}\right].$$

for any $x \in [0, 1]$,

$$P\left(\left[\frac{X_{(n)}}{\theta}\right]^n \leq x\right) = P\left(X_{(n)} \leq \theta x^{1/n}\right) = \left(\frac{\theta x^{1/n}}{\theta}\right)^n = x.$$

This implies that $Q(X_1, X_2, \ldots, X_n, \theta) = \left[\frac{X_{(n)}}{\theta}\right]^n \sim U(0, 1)$.

- Next, we use Q to construct a CI for θ. Note that,

$$P\left(\frac{\alpha}{2} \leq Q(X_1, X_2, \ldots, X_n, \theta) \leq 1 - \frac{\alpha}{2}\right) = 1 - \frac{\alpha}{2} - \frac{\alpha}{2} = 1 - \alpha.$$

- Solve the inequalities with respect to θ, we get a $100(1-\alpha)\%$ CI:

$$\left[\frac{X_{(n)}}{(1-\alpha/2)^{1/n}}, \frac{X_{(n)}}{(\alpha/2)^{1/n}}\right].$$

for any $x \in [0,1]$,

$$P\left(\left[\frac{X_{(n)}}{\theta}\right]^n \leq x\right) = P\left(X_{(n)} \leq \theta x^{1/n}\right) = \left(\frac{\theta x^{1/n}}{\theta}\right)^n = x.$$

This implies that $Q(X_1, X_2, \ldots, X_n, \theta) = \left[\frac{X_{(n)}}{\theta}\right]^n \sim U(0,1)$.

- Next, we use Q to construct a CI for θ. Note that,

$$P(\frac{\alpha}{2} \leq Q(X_1, X_2, \ldots, X_n, \theta) \leq 1 - \frac{\alpha}{2}) = 1 - \frac{\alpha}{2} - \frac{\alpha}{2} = 1 - \alpha.$$

- Solve the inequalities with respect to θ, we get a $100(1-\alpha)\%$ CI:

$$\left[\frac{X_{(n)}}{(1-\alpha/2)^{1/n}}, \frac{X_{(n)}}{(\alpha/2)^{1/n}}\right].$$

for any $x \in [0,1]$,

$$P\left(\left[\frac{X_{(n)}}{\theta}\right]^n \leq x\right) = P\left(X_{(n)} \leq \theta x^{1/n}\right) = \left(\frac{\theta x^{1/n}}{\theta}\right)^n = x.$$

This implies that $Q(X_1, X_2, \ldots, X_n, \theta) = \left[\frac{X_{(n)}}{\theta}\right]^n \sim U(0,1)$.

- Next, we use Q to construct a CI for θ. Note that,

$$P\left(\frac{\alpha}{2} \leq Q(X_1, X_2, \ldots, X_n, \theta) \leq 1 - \frac{\alpha}{2}\right) = 1 - \frac{\alpha}{2} - \frac{\alpha}{2} = 1 - \alpha.$$

- Solve the inequalities with respect to θ, we get a $100(1-\alpha)\%$ CI:

$$\left[\frac{X_{(n)}}{(1-\alpha/2)^{1/n}}, \frac{X_{(n)}}{(\alpha/2)^{1/n}}\right].$$

McGill

Illustration in R

- In class, we will illustrate the above interval via simulations in R.
- When we compute a CI based on the observed data, the resulting interval will not be random anymore, and hence it does not make sense to provide any probability statement about this interval. Be very careful about the interpretation of this interval.
  (will explain in class using R).

Confidence intervals for Normal samples

- Throughout this section, we assume that $X_1, X_2, \ldots, X_n$ is an iid sample from $N(\mu, \sigma^2)$ with unknown parameters (μ, σ^2).
- We will discuss:

(1) confidence intervals for μ,

(2) confidence intervals for the difference of means based on two random samples,

(3) confidence intervals for σ^2.

Confidence intervals for Normal samples

- Throughout this section, we assume that $X_1, X_2, \ldots, X_n$ is an iid sample from $N(\mu, \sigma^2)$ with unknown parameters (μ, σ^2).
- We will discuss:

(1) confidence intervals for μ,

(2) confidence intervals for the difference of means based on two random samples,

(3) confidence intervals for σ^2.

Confidence intervals for Normal samples

- Throughout this section, we assume that $X_1, X_2, \dots, X_n$ is an iid sample from $N(\mu, \sigma^2)$ with unknown parameters (μ, σ^2).
- We will discuss:

(1) confidence intervals for μ,
(2) confidence intervals for the difference of means based on two random samples,
(3) confidence intervals for σ^2.

(1). Confidence intervals for μ:

- Recall:

$$T_n = \frac{\sqrt{n}(\overline{X}_n - \mu)}{S_n} \sim t_{(n-1)}.$$

- Therefore, T_n is a *pivotal quantity* and can be used to construct confidence intervals for μ.
- Consider the quantiles of a Student t distribution with $(n-1)$ degrees of freedom such that:

$$P(-t_{(n-1;\frac{\alpha}{2})} \leq T_n \leq t_{(n-1;\frac{\alpha}{2})}) = 1 - \alpha.$$

(1). Confidence intervals for μ:

- Recall:

$$T_n = \frac{\sqrt{n}(\overline{X}_n - \mu)}{S_n} \sim t_{(n-1)}.$$

- Therefore, T_n is a *pivotal quantity* and can be used to construct confidence intervals for μ.

- Consider the quantiles of a Student t distribution with $(n-1)$ degrees of freedom such that:

$$P(-t_{(n-1;\frac{\alpha}{2})} \leq T_n \leq t_{(n-1;\frac{\alpha}{2})}) = 1 - \alpha.$$

(1). Confidence intervals for μ:

- Recall:

$$T_n = \frac{\sqrt{n}(\overline{X}_n - \mu)}{S_n} \sim t_{(n-1)}.$$

- Therefore, T_n is a *pivotal quantity* and can be used to construct confidence intervals for μ.

- Consider the quantiles of a Student t distribution with $(n-1)$ degrees of freedom such that:

$$P(-t_{(n-1;\frac{\alpha}{2})} \leq T_n \leq t_{(n-1;\frac{\alpha}{2})}) = 1 - \alpha.$$

Theorem:

- A $100(1-\alpha)\%$ two-sided confidence interval for μ:

$$\left[\overline{X}_n - t_{(n-1;\frac{\alpha}{2})} \times \frac{S_n}{\sqrt{n}} \;,\; \overline{X}_n + t_{(n-1;\frac{\alpha}{2})} \times \frac{S_n}{\sqrt{n}}\right]$$

- A $100(1-\alpha)\%$ lower one-sided confidence interval for μ:

$$\left[\overline{X}_n - t_{(n-1;\alpha)} \times \frac{S_n}{\sqrt{n}} \;,\; +\infty\right)$$

- A $100(1-\alpha)\%$ upper one-sided confidence interval for μ:

$$\left(-\infty \;,\; \overline{X}_n + t_{(n-1;\alpha)} \times \frac{S_n}{\sqrt{n}}\right]$$

Theorem:

- A $100(1-\alpha)\%$ two-sided confidence interval for μ:

$$\left[\overline{X}_n - t_{(n-1;\frac{\alpha}{2})} \times \frac{S_n}{\sqrt{n}} \;,\; \overline{X}_n + t_{(n-1;\frac{\alpha}{2})} \times \frac{S_n}{\sqrt{n}}\right]$$

- A $100(1-\alpha)\%$ lower one-sided confidence interval for μ:

$$\left[\overline{X}_n - t_{(n-1;\alpha)} \times \frac{S_n}{\sqrt{n}} \;,\; +\infty\right)$$

- A $100(1-\alpha)\%$ upper one-sided confidence interval for μ:

$$\left(-\infty \;,\; \overline{X}_n + t_{(n-1;\alpha)} \times \frac{S_n}{\sqrt{n}}\right]$$

Theorem:

- A $100(1-\alpha)\%$ two-sided confidence interval for μ:

$$\left[\overline{X}_n - t_{(n-1;\frac{\alpha}{2})} \times \frac{S_n}{\sqrt{n}} \;,\; \overline{X}_n + t_{(n-1;\frac{\alpha}{2})} \times \frac{S_n}{\sqrt{n}}\right]$$

- A $100(1-\alpha)\%$ lower one-sided confidence interval for μ:

$$\left[\overline{X}_n - t_{(n-1;\alpha)} \times \frac{S_n}{\sqrt{n}} \;,\; +\infty\right)$$

- A $100(1-\alpha)\%$ upper one-sided confidence interval for μ:

$$\left(-\infty \;,\; \overline{X}_n + t_{(n-1;\alpha)} \times \frac{S_n}{\sqrt{n}}\right]$$

(2). Confidence intervals for the difference of two Normal means

- Often, we are interested in comparing the means μ_1 and μ_2 of two populations. This is done based on two random samples:

$$X_1, X_2, \ldots, X_n \quad ; \quad Y_1, Y_2, \ldots, Y_m.$$

- Examples: comparing the effectiveness of two medications, two training methods, or two production methods, etc.

(2). Confidence intervals for the difference of two Normal means

- Often, we are interested in comparing the means μ_1 and μ_2 of two populations. This is done based on two random samples:

$$X_1, X_2, \ldots, X_n \quad ; \quad Y_1, Y_2, \ldots, Y_m.$$

- Examples: comparing the effectiveness of two medications, two training methods, or two production methods, etc.

Assumptions:

- It is easy to construct confidence intervals for the difference $\mu_1 - \mu_2$ under these conditions:

(i) $X_1, X_2, \ldots, X_n \stackrel{\text{iid}}{\sim} N(\mu_1, \sigma_1^2)$,

(ii) $Y_1, Y_2, \ldots, Y_m \stackrel{\text{iid}}{\sim} N(\mu_2, \sigma_2^2)$,

(iii) The two samples are mutually independent,

(iv) $\sigma_1^2 = \sigma_2^2 = \sigma^2$.

Notation:

- Denote the two sample means, variances, and a pooled variance:

$$\begin{aligned}
\overline{X}_n &= \frac{1}{n}\sum_{i=1}^{n} X_i \;\;, \;\; S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2, \\
\overline{Y}_n &= \frac{1}{m}\sum_{i=1}^{m} Y_i \;\;, \;\; S_m^2 = \frac{1}{m-1}\sum_{i=1}^{m}(Y_i - \overline{Y}_m)^2, \\
S^2_{\text{pooled}} &= \frac{1}{n+m-2}\left\{\sum_{i=1}^{n}(X_i - \overline{X}_n)^2 + \sum_{i=1}^{m}(Y_i - \overline{Y}_m)^2.\right\}
\end{aligned}$$

- Fact: $\frac{(n+m-2)S^2_{\text{pooled}}}{\sigma^2} \sim \chi^2_{(n+m-2)}$.

A pivotal quantity

- Under Conditions (i)-(iv), we have that:

$$\frac{(\overline{X}_n - \overline{Y}_m) - (\mu_1 - \mu_2)}{S_{\text{pooled}}\sqrt{\frac{1}{n} + \frac{1}{m}}} \sim t_{(n+m-2)}.$$

- Why? (Details will be discussed in class).
- The above quantity can then be used to construct confidence intervals for $\mu_1 - \mu_2$.

A pivotal quantity

- Under Conditions (i)-(iv), we have that:

$$\frac{(\overline{X}_n - \overline{Y}_m) - (\mu_1 - \mu_2)}{S_{\text{pooled}}\sqrt{\frac{1}{n} + \frac{1}{m}}} \sim t_{(n+m-2)}.$$

- Why? (Details will be discussed in class).
- The above quantity can then be used to construct confidence intervals for $\mu_1 - \mu_2$.

A pivotal quantity

- Under Conditions (i)-(iv), we have that:

$$\frac{(\overline{X}_n - \overline{Y}_m) - (\mu_1 - \mu_2)}{S_{\text{pooled}}\sqrt{\frac{1}{n} + \frac{1}{m}}} \sim t_{(n+m-2)}.$$

- Why? (Details will be discussed in class).
- The above quantity can then be used to construct confidence intervals for $\mu_1 - \mu_2$.

Theorem:

- A $100(1-\alpha)\%$ two-sided confidence interval for $\mu_1 - \mu_2$:

$$\left[(\overline{X}_n - \overline{Y}_m) - t_{(n+m-2;\frac{\alpha}{2})} \times S_{\text{pooled}}\sqrt{\frac{1}{n}+\frac{1}{m}}\ ,\right.$$

$$\left.(\overline{X}_n - \overline{Y}_m) + t_{(n+m-2;\frac{\alpha}{2})} \times S_{\text{pooled}}\sqrt{\frac{1}{n}+\frac{1}{m}}\ \right]$$

- $100(1-\alpha)\%$ upper and lower one-sided CI for $\mu_1 - \mu_2$:

$$\left[(\overline{X}_n - \overline{Y}_m) - t_{(n+m-2;\alpha)} \times S_{\text{pooled}}\sqrt{\frac{1}{n}+\frac{1}{m}}\ ,\ +\infty\right)$$

$$\left(-\infty\ ,\ (\overline{X}_n - \overline{Y}_m) + t_{(n+m-2;\alpha)} \times S_{\text{pooled}}\sqrt{\frac{1}{n}+\frac{1}{m}}\ \right]$$

(3) Confidence intervals for σ^2

- Recall: If $X_1, X_2, \ldots, X_n$ is iid from $N(\mu, \sigma^2)$, then

$$\frac{(n-1)S_n^2}{\sigma^2} = \frac{\sum_{i=1}^{n}(X_i - \overline{X}_n)^2}{\sigma^2} \sim \chi^2_{(n-1)}.$$

- Hence, we have a pivotal quantity that can be use to construct confidence intervals for σ^2.

- Note that unlike the Normal and Student t distributions, chi-squared distribution is not symmetric. We use equal-tail probability quantiles of the χ^2 distribution:

$$P\left(\chi^2_{n-1;1-\frac{\alpha}{2}} \leq \frac{(n-1)S_n^2}{\sigma^2} \leq \chi^2_{n-1;\frac{\alpha}{2}}\right) = 1-\alpha$$

McGill

(3) Confidence intervals for σ^2

- Recall: If $X_1, X_2, \ldots, X_n$ is iid from $N(\mu, \sigma^2)$, then

$$\frac{(n-1)S_n^2}{\sigma^2} = \frac{\sum_{i=1}^n (X_i - \overline{X}_n)^2}{\sigma^2} \sim \chi^2_{(n-1)}.$$

- Hence, we have a pivotal quantity that can be use to construct confidence intervals for σ^2.

- Note that unlike the Normal and Student t distributions, chi-squared distribution is not symmetric. We use equal-tail probability quantiles of the χ^2 distribution:

$$P\left(\chi^2_{n-1;1-\frac{\alpha}{2}} \le \frac{(n-1)S_n^2}{\sigma^2} \le \chi^2_{n-1;\frac{\alpha}{2}}\right) = 1 - \alpha$$

(3) Confidence intervals for σ^2

- Recall: If $X_1, X_2, \ldots, X_n$ is iid from $N(\mu, \sigma^2)$, then

$$\frac{(n-1)S_n^2}{\sigma^2} = \frac{\sum_{i=1}^{n}(X_i - \overline{X}_n)^2}{\sigma^2} \sim \chi^2_{(n-1)}.$$

- Hence, we have a pivotal quantity that can be use to construct confidence intervals for σ^2.

- Note that unlike the Normal and Student t distributions, chi-squared distribution is not symmetric. We use equal-tail probability quantiles of the χ^2 distribution:

$$P\left(\chi^2_{n-1;1-\frac{\alpha}{2}} \leq \frac{(n-1)S_n^2}{\sigma^2} \leq \chi^2_{n-1;\frac{\alpha}{2}}\right) = 1 - \alpha$$

(3) Confidence intervals for σ^2

- Recall: If $X_1, X_2, \ldots, X_n$ is iid from $N(\mu, \sigma^2)$, then

$$\frac{(n-1)S_n^2}{\sigma^2} = \frac{\sum_{i=1}^n (X_i - \overline{X}_n)^2}{\sigma^2} \sim \chi^2_{(n-1)}.$$

- Hence, we have a pivotal quantity that can be use to construct confidence intervals for σ^2.

- Note that unlike the Normal and Student t distributions, chi-squared distribution is not symmetric. We use equal-tail probability quantiles of the χ^2 distribution:

$$P\left(\chi^2_{n-1;1-\frac{\alpha}{2}} \leq \frac{(n-1)S_n^2}{\sigma^2} \leq \chi^2_{n-1;\frac{\alpha}{2}}\right) = 1 - \alpha$$

Theorem:

- A $100(1-\alpha)\%$ two-sided confidence interval for σ^2:

$$\left[\frac{(n-1)S_n^2}{\chi^2_{n-1;\frac{\alpha}{2}}}\ ,\ \frac{(n-1)S_n^2}{\chi^2_{n-1;1-\frac{\alpha}{2}}}\right]$$

Remarks

So far we have been working with iid sample(s) from Normal population(s):

- Confidence intervals for the mean or the difference of means are often meaningful (i.e., the coverage probability is about $1 - \alpha$) even for non-Normal samples, as long as the underlying distribution is roughly mound-shaped.
- However, Normality assumption is critical for the confidence interval for the variance. If the underlying distribution is not Normal, the confidence interval derived previously may perform very poorly (i.e., the coverage probability could be much lower than $1 - \alpha$).

Remarks

So far we have been working with iid sample(s) from Normal population(s):

- Confidence intervals for the mean or the difference of means are often meaningful (i.e., the coverage probability is about $1 - \alpha$) even for non-Normal samples, as long as the underlying distribution is roughly mound-shaped.
- However, Normality assumption is critical for the confidence interval for the variance. If the underlying distribution is not Normal, the confidence interval derived previously may perform very poorly (i.e., the coverage probability could be much lower than $1 - \alpha$).

Remarks

So far we have been working with iid sample(s) from Normal population(s):

- Confidence intervals for the mean or the difference of means are often meaningful (i.e., the coverage probability is about $1 - \alpha$) even for non-Normal samples, as long as the underlying distribution is roughly mound-shaped.
- However, Normality assumption is critical for the confidence interval for the variance. If the underlying distribution is not Normal, the confidence interval derived previously may perform very poorly (i.e., the coverage probability could be much lower than $1 - \alpha$).

- There are situations where the data are not Normal, and also exact pivotal quantities are difficult (if not impossible) to find.

- Solution:

  Approximate pivotal quantities, i.e., pivots whose distributions do not depend on θ when $n \to \infty$.

- There are situations where the data are not Normal, and also exact pivotal quantities are difficult (if not impossible) to find.

- Solution:

  Approximate pivotal quantities, i.e., pivots whose distributions do not depend on θ when $n \to \infty$.

Toward approximate pivots

- Let $\hat{\theta}_n$ be an estimator of θ, and denote $\sigma^2_{\hat{\theta}_n} = Var(\hat{\theta}_n)$. We can find many examples where, $n \to \infty$,

$$\frac{\hat{\theta}_n - \theta}{\sigma_{\hat{\theta}_n}} \xrightarrow{d} N(0, 1)$$

- Examples: (based on iid samples)

  Sample mean, difference in two sample means, sample proportion, difference in two sample proportions, the maximum likelihood estimators, many moment estimators.

McGill

Toward approximate pivots

- Let $\hat{\theta}_n$ be an estimator of θ, and denote $\sigma^2_{\hat{\theta}_n} = Var(\hat{\theta}_n)$. We can find many examples where, $n \to \infty$,

$$\frac{\hat{\theta}_n - \theta}{\sigma_{\hat{\theta}_n}} \xrightarrow{d} N(0, 1)$$

- Examples: (based on iid samples)

  Sample mean, difference in two sample means, sample proportion, difference in two sample proportions, the maximum likelihood estimators, many moment estimators.

Approximate pivotal quantities

- If $\hat{\sigma}_{\hat{\theta}_n}$, i.e. the standard error of $\hat{\theta}_n$, is a consistent estimator of $\sigma_{\hat{\theta}_n}$, then we have that, as $n \to \infty$,

$$\frac{\hat{\theta}_n - \theta}{\hat{\sigma}_{\hat{\theta}_n}} \xrightarrow{d} N(0, 1)$$

- Hence, an approximate pivotal quantity:

$$\frac{\hat{\theta}_n - \theta}{\hat{\sigma}_{\hat{\theta}_n}}.$$

Approximate pivotal quantities

- If $\hat{\sigma}_{\hat{\theta}_n}$, i.e. the standard error of $\hat{\theta}_n$, is a consistent estimator of $\sigma_{\hat{\theta}_n}$, then we have that, as $n \to \infty$,

$$\frac{\hat{\theta}_n - \theta}{\hat{\sigma}_{\hat{\theta}_n}} \xrightarrow{d} N(0, 1)$$

- Hence, an approximate pivotal quantity:

$$\frac{\hat{\theta}_n - \theta}{\hat{\sigma}_{\hat{\theta}_n}}.$$

Approximate confidence intervals for θ:

For large n,

- a $100(1-\alpha)\%$ approximate two-sided confidence interval:

$$\left[\hat{\theta}_n - z_{\frac{\alpha}{2}} \times \hat{\sigma}_{\hat{\theta}_n} \quad , \quad \hat{\theta}_n + z_{\frac{\alpha}{2}} \times \hat{\sigma}_{\hat{\theta}_n}\right]$$

- a $100(1-\alpha)\%$ approximate lower one-sided confidence interval:

$$\left[\hat{\theta}_n - z_{\alpha} \times \hat{\sigma}_{\hat{\theta}_n} \quad , \quad +\infty\right)$$

- a $100(1-\alpha)\%$ approximate upper one-sided confidence interval:

$$\left(-\infty \quad , \quad \hat{\theta}_n + z_{\alpha} \times \hat{\sigma}_{\hat{\theta}_n}\right]$$

(1) Mean

- Let $X_1, X_2, \ldots, X_n$ be an iid sample from a population with unknown mean μ and variance σ^2, such that $E(X^4) < \infty$.

- As $n \to \infty$,

$$\frac{\sqrt{n}(\overline{X}_n - \mu)}{S_n} \xrightarrow{d} N(0,1).$$

- A $100(1-\alpha)\%$ approximate two-sided confidence interval for μ:

$$\left[\overline{X}_n - z_{\frac{\alpha}{2}} \times \frac{S_n}{\sqrt{n}} \;,\; \overline{X}_n + z_{\frac{\alpha}{2}} \times \frac{S_n}{\sqrt{n}}\right]$$

(1) Mean

- Let $X_1, X_2, \ldots, X_n$ be an iid sample from a population with unknown mean μ and variance σ^2, such that $E(X^4) < \infty$.

- As $n \to \infty$,

$$\frac{\sqrt{n}(\overline{X}_n - \mu)}{S_n} \xrightarrow{d} N(0,1).$$

- A $100(1-\alpha)\%$ approximate two-sided confidence interval for μ:

$$\left[\overline{X}_n - z_{\frac{\alpha}{2}} \times \frac{S_n}{\sqrt{n}} \;,\; \overline{X}_n + z_{\frac{\alpha}{2}} \times \frac{S_n}{\sqrt{n}}\right]$$

(1) Mean

- Let $X_1, X_2, \ldots, X_n$ be an iid sample from a population with unknown mean μ and variance σ^2, such that $E(X^4) < \infty$.

- As $n \to \infty$,

$$\frac{\sqrt{n}(\overline{X}_n - \mu)}{S_n} \xrightarrow{d} N(0, 1).$$

- A $100(1 - \alpha)\%$ approximate two-sided confidence interval for μ:

$$\left[\overline{X}_n - z_{\frac{\alpha}{2}} \times \frac{S_n}{\sqrt{n}} \;,\; \overline{X}_n + z_{\frac{\alpha}{2}} \times \frac{S_n}{\sqrt{n}}\right]$$

(2) Difference in two population means

- Suppose that $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_m$ are two *independent* iid samples with means μ_1 and μ_2 and finite variances σ_1^2 and σ_2^2, respectively.

- By the CLT, as $n, m \to \infty$ such that $n/m \to \rho$, we have

$$\frac{(\overline{X}_n - \overline{Y}_m) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}} \xrightarrow{d} N(0, 1).$$

- If $E(X^4), E(Y^4) < \infty$, then it as $n, m \to \infty$ so that $n/m \to \rho$, we have $\sqrt{\frac{S_n^2}{n} + \frac{S_m^2}{m}} / \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}$ converges to one, in probability.

(2) Difference in two population means

- Suppose that $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_m$ are two *independent* iid samples with means μ_1 and μ_2 and finite variances σ_1^2 and σ_2^2, respectively.
- By the CLT, as $n, m \to \infty$ such that $n/m \to \rho$, we have

$$\frac{(\overline{X}_n - \overline{Y}_m) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}} \xrightarrow{d} N(0,1).$$

- If $E(X^4), E(Y^4) < \infty$, then it as $n, m \to \infty$ so that $n/m \to \rho$, we have $\sqrt{\frac{S_n^2}{n} + \frac{S_m^2}{m}} / \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}$ converges to one, in probability.

(2) Difference in two population means

- Suppose that $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_m$ are two *independent* iid samples with means μ_1 and μ_2 and finite variances σ_1^2 and σ_2^2, respectively.
- By the CLT, as $n, m \to \infty$ such that $n/m \to \rho$, we have

$$\frac{(\overline{X}_n - \overline{Y}_m) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}} \xrightarrow{d} N(0,1).$$

- If $E(X^4), E(Y^4) < \infty$, then it as $n, m \to \infty$ so that $n/m \to \rho$, we have $\sqrt{\frac{S_n^2}{n} + \frac{S_m^2}{m}} / \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}$ converges to one, in probability.

- Therefore, a $100(1-\alpha)\%$ approximate two-sided confidence interval for $\mu_1-\mu_2$, when n, m are large:

$$\left[(\overline{X}_n-\overline{Y}_m)-z_{\frac{\alpha}{2}}\times\sqrt{\frac{S_n^2}{n}+\frac{S_m^2}{m}}\;,\;(\overline{X}_n-\overline{Y}_m)+z_{\frac{\alpha}{2}}\times\sqrt{\frac{S_n^2}{n}+\frac{S_m^2}{m}}\right].$$

(3) Population proportion

- Suppose that $X_1, X_2, \ldots, X_n$ is an iid sample from a Bernoulli distribution with probability of success p (unknown).
- Consider the point estimator $\hat{p}_n = \overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$.
- By the CLT, as $n \to \infty$,

$$\frac{\hat{p}_n - p}{\sqrt{p(1-p)/n}} \xrightarrow{d} N(0,1)$$

and also $\hat{p}_n \xrightarrow{p} p$.

(3) Population proportion

- Suppose that $X_1, X_2, \ldots, X_n$ is an iid sample from a Bernoulli distribution with probability of success p (unknown).
- Consider the point estimator $\hat{p}_n = \overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$.
- By the CLT, as $n \to \infty$,

$$\frac{\hat{p}_n - p}{\sqrt{p(1-p)/n}} \xrightarrow{d} N(0, 1)$$

and also $\hat{p}_n \xrightarrow{p} p$.

(3) Population proportion

- Suppose that $X_1, X_2, \ldots, X_n$ is an iid sample from a Bernoulli distribution with probability of success p (unknown).
- Consider the point estimator $\hat{p}_n = \overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$.
- By the CLT, as $n \to \infty$,

$$\frac{\hat{p}_n - p}{\sqrt{p(1-p)/n}} \xrightarrow{d} N(0,1)$$

and also $\hat{p}_n \xrightarrow{p} p$.

- Therefore, a $100(1-\alpha)\%$ approximate two-sided confidence interval for p, when n is large:

$$\left[\hat{p}_n - z_{\frac{\alpha}{2}} \times \sqrt{\frac{\hat{p}_n(1-\hat{p}_n)}{n}} \;,\; \hat{p}_n + z_{\frac{\alpha}{2}} \times \sqrt{\frac{\hat{p}_n(1-\hat{p}_n)}{n}}\right].$$

(4) Difference in two population proportions

- Suppose that $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_m$ are two *independent* iid samples from two Bernoulli distributions with parameters p_1 and p_2, respectively.

- Consider the point estimators of p_1 and p_2:

$$\hat{p}_1 = \overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \quad , \quad \hat{p}_2 = \overline{Y}_m = \frac{1}{m}\sum_{i=1}^{m} Y_i.$$

- It is straightforward to show that, as $n, m \to \infty$ so that $n/m \to \rho$,

$$\frac{(\hat{p}_1 - \hat{p}_2) - (\hat{p}_2 - p_2)}{\sqrt{\hat{p}_1(1-\hat{p}_1)/n + \hat{p}_2(1-\hat{p}_2)/m}} \xrightarrow{d} N(0,1).$$

(4) Difference in two population proportions

- Suppose that $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_m$ are two *independent* iid samples from two Bernoulli distributions with parameters p_1 and p_2, respectively.

- Consider the point estimators of p_1 and p_2:

$$\hat{p}_1 = \overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \quad , \quad \hat{p}_2 = \overline{Y}_m = \frac{1}{m}\sum_{i=1}^{m} Y_i.$$

- It is straightforward to show that, as $n, m \to \infty$ so that $n/m \to \rho$,

$$\frac{(\hat{p}_1 - \hat{p}_2) - (\hat{p}_2 - p_2)}{\sqrt{\hat{p}_1(1-\hat{p}_1)/n + \hat{p}_2(1-\hat{p}_2)/m}} \xrightarrow{d} N(0,1).$$

(4) Difference in two population proportions

- Suppose that $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_m$ are two *independent* iid samples from two Bernoulli distributions with parameters p_1 and p_2, respectively.

- Consider the point estimators of p_1 and p_2:

$$\hat{p}_1 = \overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \quad , \quad \hat{p}_2 = \overline{Y}_m = \frac{1}{m}\sum_{i=1}^{m} Y_i.$$

- It is straightforward to show that, as $n, m \to \infty$ so that $n/m \to \rho$,

$$\frac{(\hat{p}_1 - \hat{p}_2) - (\hat{p}_2 - p_2)}{\sqrt{\hat{p}_1(1-\hat{p}_1)/n + \hat{p}_2(1-\hat{p}_2)/m}} \xrightarrow{d} N(0,1).$$

- Therefore, a $100(1-\alpha)\%$ approximate two-sided confidence interval for $p_1 - p_2$, when n, m are large:

$$\left[(\hat{p}_1 - \hat{p}_2) - z_{\frac{\alpha}{2}} \times \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n} + \frac{\hat{p}_2(1-\hat{p}_2)}{m}}\ ,\right.$$

$$\left.(\hat{p}_1 - \hat{p}_2) + z_{\frac{\alpha}{2}} \times \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n} + \frac{\hat{p}_2(1-\hat{p}_2)}{m}}\right].$$

(5) Approximate confidence interval for θ using the MLE theory

- Suppose $X_1, X_2, \ldots, X_n$ are i.i.d from a parametric model $f(\cdot;\theta)$, and θ is a one-dimensional unknown parameter. Let $\hat{\theta}_n = \hat{\theta}(X_1, X_2, \ldots, X_n)$ be the maximum likelihood estimator of θ. Under the Regularity Conditions, as $n \to \infty$,

$$\frac{\hat{\theta}_n - \theta}{\sqrt{[nI(\hat{\theta}_n)]^{-1}}} \xrightarrow{d} N(0,1)$$

where $I(\theta)$ is the Fisher information given by

$$I(\theta) = E\left\{ -\frac{\partial^2}{\partial\theta^2} \log f(X;\theta) \right\}$$

(5) Approximate confidence interval for θ using the MLE theory

- Suppose $X_1, X_2, \ldots, X_n$ are i.i.d from a parametric model $f(\cdot;\theta)$, and θ is a one-dimensional unknown parameter. Let $\hat{\theta}_n = \hat{\theta}(X_1, X_2, \ldots, X_n)$ be the maximum likelihood estimator of θ. Under the Regularity Conditions, as $n \to \infty$,

$$\frac{\hat{\theta}_n - \theta}{\sqrt{[nI(\hat{\theta}_n)]^{-1}}} \xrightarrow{d} N(0,1)$$

where $I(\theta)$ is the Fisher information given by

$$I(\theta) = E\left\{ -\frac{\partial^2}{\partial\theta^2} \log f(X;\theta) \right\}$$

...

- Note: In practice, $nI(\theta)$ is estimated either by $nI(\hat{\theta}_n)$ or by so called the *observed* Fisher information

$$-\frac{\partial^2 l_n(\theta)}{\partial \theta^2}\bigg|_{\theta=\hat{\theta}_n} = -\sum_{i=1}^{n} \frac{\partial^2}{\partial \theta^2} \log f(x_i;\theta)\bigg|_{\theta=\hat{\theta}_n}.$$

where $l_n(\theta)$ is the log-likelihood function of θ.

- In either case, a $100(1-\alpha)\%$ approximate two-sided confidence interval for θ, when n is large, is given by:

$$\left[\hat{\theta} - z_{\frac{\alpha}{2}} \times \sqrt{\frac{1}{n} \times [\widehat{I(\theta)}]^{-1}}\ ,\ \hat{\theta} + z_{\frac{\alpha}{2}} \times \sqrt{\frac{1}{n} \times [\widehat{I(\theta)}]^{-1}}\right]$$

...

- Note: In practice, $nI(\theta)$ is estimated either by $nI(\hat{\theta}_n)$ or by so called the *observed* Fisher information

$$-\frac{\partial^2 l_n(\theta)}{\partial\theta^2}\bigg|_{\theta=\hat{\theta}_n} = -\sum_{i=1}^{n} \frac{\partial^2}{\partial\theta^2} \log f(x_i;\theta)\bigg|_{\theta=\hat{\theta}_n}.$$

where $l_n(\theta)$ is the log-likelihood function of θ.

- In either case, a $100(1-\alpha)\%$ approximate two-sided confidence interval for θ, when n is large, is given by:

$$\left[\hat{\theta} - z_{\frac{\alpha}{2}} \times \sqrt{\frac{1}{n} \times [\widehat{I(\theta)}]^{-1}}\ ,\ \hat{\theta} + z_{\frac{\alpha}{2}} \times \sqrt{\frac{1}{n} \times [\widehat{I(\theta)}]^{-1}}\right]$$

McGill

Examples

- Now, we will apply the above confidence intervals to different data examples.

Example 3

- A manufacturer developed a new gunpowder and tested it in eight shells. The resulting muzzle velocities, in feet per second, were:

$$3005, 2925, 2935, 2965, 2995, 3005, 2937, 2905.$$

- Assume that the velocities are iid sample from a $N(\mu, \sigma^2)$. Compute 95% two-sided and one-sided confidence intervals for μ. Provide interpretation for the intervals.

Example 4:

- In a packing plant, a machine packs cartons with jars. It is supposed that a new machine will pack faster on the average than the machine currently used. To test that hypothesis, the times it takes each machine to pack ten cartons are recorded. The results in seconds are:

$$\text{old} : 42.7, 43.8, 42.5, 43.1, 44.0, 43.6, 43.3, 43.5, 41.7, 44.1$$

$$\text{new} : 42.1, 41.3, 42.4, 43.2, 41.8, 41.0, 41.8, 42.8, 42.3, 42.7$$

- Construct a two-sided 95% confidence for the difference in the respective means. Provide interpretation for the interval. (Assume that the timings for the old and new machines are independent iid samples from $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$, respectively, and $\sigma_1 = \sigma_2$.)

Example 5:

- Assume that the number of days needed to hatch an egg of a certain type of a rare lizard is distributed Normally. Using incubator, 13 eggs from different nests separately hatched. The sample mean is 18.97 weeks and the sample standard deviation is $\sqrt{10.7}$ weeks. Find a 90% confidence interval for the population variance. Provide interpretation for the interval.

Example 6

- Shopping times of 64 randomly selected customers in a supermarket averaged 33 minutes with a standard deviation of 16 minutes. Construct an approximate 90% confidence interval for the true mean shopping time per customer. Provide interpretation for the interval.

Example 7:

- We wish to compare the daily intake of selenium in two regions. In each region, 30 adults were tested and the results (in mg/day) were:

$$\overline{x}_n = 167.1 \;,\; s_n = 24.3 \;\;,\;\; \overline{y}_m = 140.9 \;,\; s_m = 17.6$$

- Find a 95% two-sided approximate confidence interval for the difference in mean daily intake of selenium in the two regions. Provide interpretation for the interval.

Example 8:

- A sample of $n = 1000$ voters, randomly selected from a city, showed 560 in favour of candidate Jones. Find an approximate 99% confidence interval for the population proportion in favour of candidate Jones. Provide interpretation for the interval.

Example 9:

- A medical researcher conjectures that smoking can result in wrinkled skin around the eyes. The researcher recruited 150 smokers and 250 nonsmokers to take part in an observational study and found that 95 of the smokers and 105 of the nonsmokers were seen to have prominent wrinkles around the eyes (based on a standardized wrinkle score administered by a person who did not know if the subject smoked or not). Find an approximate 95% confidence interval for the difference in the proportions of people who have wrinkled skin around their eyes in the two populations. Provide interpretation for the interval.

Example 10:

- Suppose $X_1, X_2, \ldots, X_n$ are i.i.d from a Poisson distribution $Pos(\lambda)$, and λ is the unknown parameter. Using the MLE theory, construct a $100(1-\alpha)\%$ approximate two-sided confidence interval for λ.

Sample size calculations

- How much data do we need in a study? (sample size n)
- Once we decide how accurate we wish the estimation to be for the population mean or proportion, large sample confidence intervals could be used to calculate the required sample size n.
- Note that sample sizes are typically calculated before the data have been collected.

Sample size calculations

- How much data do we need in a study? (sample size n)
- Once we decide how accurate we wish the estimation to be for the population mean or proportion, large sample confidence intervals could be used to calculate the required sample size n.
- Note that sample sizes are typically calculated before the data have been collected.

Sample size calculations

- How much data do we need in a study? (sample size n)
- Once we decide how accurate we wish the estimation to be for the population mean or proportion, large sample confidence intervals could be used to calculate the required sample size n.
- Note that sample sizes are typically calculated before the data have been collected.

Inference about the mean μ

- Recall that if the population variance σ^2 is known, then $100(1-\alpha)\%$ confidence interval for μ is:

$$\left[\overline{X}_n - z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}} \ , \ \overline{X}_n + z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}}\right]$$

- And we re-write the interval as

$$\left[\overline{X}_n - \varepsilon \, , \, \overline{X}_n + \varepsilon\right]$$

where ε is called precision error.

Inference about the mean μ

- Recall that if the population variance σ^2 is known, then $100(1-\alpha)\%$ confidence interval for μ is:

$$\left[\overline{X}_n - z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}} \;,\; \overline{X}_n + z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}}\right]$$

- And we re-write the interval as

$$\left[\overline{X}_n - \varepsilon \,,\, \overline{X}_n + \varepsilon\right]$$

where ε is called precision error.

...

- Then the required sample size n in order to achieve the precision error ε is

$$\varepsilon \leq z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}} \Longleftrightarrow n \geq \sigma^2 \times \left(\frac{z_{\frac{\alpha}{2}}}{\varepsilon}\right)^2$$

- In the above formula, $z_{\frac{\alpha}{2}}$ and ε are given. The variance σ^2 must be either known (rare), or estimated from previous data or from similar experiments.

...

- Then the required sample size n in order to achieve the precision error ε is

$$\varepsilon \leq z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}} \Longleftrightarrow n \geq \sigma^2 \times \left(\frac{z_{\frac{\alpha}{2}}}{\varepsilon}\right)^2$$

- In the above formula, $z_{\frac{\alpha}{2}}$ and ε are given. The variance σ^2 must be either known (rare), or estimated from previous data or from similar experiments.

Example 11

- Suppose we wish to estimate the average daily yield μ of a chemical and we wish the error of estimation to be less than 5 tons with probability 0.95. From a previous study $\sigma = 21$.

Example 12: (Inference about the mean p)

- In a study, the goal is to estimate the probability p that a person will react in manner A rather than B to a given stimulus in a psychological experiment. How many people must be included in the study for the error of estimation to be less than 0.04 with probability 0.90?

- Here $z_{(.1/2)} = z_{.05} = 1.64$. Also, the variance of a Bernoulli distribution is $p(1-p)$. Similar to Example 11, the required sample size satisfies

$$n \geq p(1-p) \times \left(\frac{z_{\frac{\alpha}{2}}}{\varepsilon}\right)^2$$

Example 12: (Inference about the mean p)

- In a study, the goal is to estimate the probability p that a person will react in manner A rather than B to a given stimulus in a psychological experiment. How many people must be included in the study for the error of estimation to be less than 0.04 with probability 0.90?

- Here $z_{(.1/2)} = z_{.05} = 1.64$. Also, the variance of a Bernoulli distribution is $p(1-p)$. Similar to Example 11, the required sample size satisfies

$$n \geq p(1-p) \times \left(\frac{z_{\frac{\alpha}{2}}}{\varepsilon}\right)^2$$

...

- Since for any $p \in (0,1)$, we have $p(1-p) \leq \frac{1}{4}$, we can make a conservative sample size choice as:

$$n \geq \frac{1}{4}\left(\frac{z_{\frac{\alpha}{2}}}{\varepsilon}\right)^2.$$

- We will the rest of the calculation in class.

Inference about the mean difference $\mu_1 - \mu_2$

- If the variances of the two populations are known, a $100(1-\alpha)\%$ two-sided confidence interval for $\mu_1 - \mu_2$ is given by

$$\left[(\overline{X}_n - \overline{Y}_m) - z_{\frac{\alpha}{2}} \times \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}\ ,\ (\overline{X}_n - \overline{Y}_m) + z_{\frac{\alpha}{2}} \times \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}\ \right].$$

- And we re-write the interval as

$$\left[(\overline{X}_n - \overline{Y}_m) - \varepsilon\ ,\ (\overline{X}_n - \overline{Y}_m) + \varepsilon\right]$$

where ε is called precision error.

...

- Then the required sample size, assuming $n = m$, in order to achieve the precision error ε is

$$\varepsilon \le z_{\frac{\alpha}{2}} \times \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}} \Longleftrightarrow n \ge (\sigma_1^2 + \sigma_2^2) \times \left(\frac{z_{\frac{\alpha}{2}}}{\varepsilon}\right)^2$$

- In the above formula, $z_{\frac{\alpha}{2}}$ and ε are given. The variances σ_1^2, σ_2^2 must be either known (rare), or estimated from previous data or from similar experiments.

...

- Then the required sample size, assuming $n = m$, in order to achieve the precision error ε is

$$\varepsilon \leq z_{\frac{\alpha}{2}} \times \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}} \Longleftrightarrow n \geq (\sigma_1^2 + \sigma_2^2) \times \left(\frac{z_{\frac{\alpha}{2}}}{\varepsilon}\right)^2$$

- In the above formula, $z_{\frac{\alpha}{2}}$ and ε are given. The variances σ_1^2, σ_2^2 must be either known (rare), or estimated from previous data or from similar experiments.

Example 13

- To compare two methods of training employees for performing an assembly operation, the selected employees will be divided into two groups of equal size and trained by Methods 1 and 2, respectively. After training, the length of assembly time will be recorded for each trainee. The experimenter wishes the difference in mean assembly times to be correct to within 1 minute with probability 0.95. How many workers must be included in each training group? Assume that $\sigma_1 \approx \sigma_2 = 2$.

- This will be discussed in class.