Chapter 8: Sampling distributions of estimators

Sections

- 8.1 Sampling distribution of a statistic
- 8.2 The Chi-square distributions
- 8.3 Joint Distribution of the sample mean and sample variance
    - Skip: p. 476 - 478
- 8.4 The *t* distributions
    - Skip: derivation of the pdf, p. 483 - 484
- 8.5 Confidence intervals
- 8.6 Bayesian Analysis of Samples from a Normal Distribution
- 8.7 Unbiased Estimators
- 8.8 Fisher Information

Review from Sections 8.1 - 8.4

Let $X_1, \dots, X_n$ be a random sample from $N(\mu, \sigma^2)$

- Chi-square distribution:
  χ^2_m is the same as $\text{Gamma}(\alpha = m/2, \beta = 1/2)$

$$\frac{n\,\widehat{\sigma_0^2}}{\sigma^2} \sim \chi^2_n \quad \text{where} \quad \widehat{\sigma_0^2} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2$$

- Also,

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

- The t_m distribution: If $Y \sim \chi^2_m$ and $Z \sim N(0,1)$ are independent then $\frac{Z}{\sqrt{Y/m}} \sim t_m$. Also,

$$\frac{\sqrt{n}(\overline{X}_n - \mu)}{\sigma'} \sim t_{n-1} \quad \text{where} \quad \sigma' = \left[\frac{\sum_{i=1}^{n}(X_i - \overline{X}_n)^2}{n-1}\right]^{1/2}$$

Confidence Interval – A frequentist tool

- Say we want to estimate θ, or in general $g(\theta)$
- We also want to know "how good" that estimate is.

Def: Confidence Interval (CI)

Let $X_1, \ldots, X_n$ be a random sample from $f(x|\theta)$, where θ is unknown (but not random). Let $g(\theta)$ be a real-valued function and let A and B be statistics where

$$P\left(A < g(\theta) < B\right) \geq \gamma \quad \forall\theta \ .$$

The random interval (A, B) is called a *$100\gamma\%$ confidence interval for $g(\theta)$*. If "=", the CI is *exact*.

- After the random variables $X_1, \ldots, X_n$ have been observed and the values of $A = a$ and $B = b$ have been computed, the interval (a, b) is called the *observed confidence interval*.

Confidence Interval - Mean of a Normal Distribution

Last time we saw the following example

- Let $X_1, \dots, X_n$ be a random sample from $N(\mu, \sigma^2)$
- Let

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \quad \text{and} \quad \sigma' = \left(\frac{\sum_{i=1}^{n}(X_i - \overline{X}_n)^2}{n-1}\right)^{1/2}$$

- Then we know that

$$U = \frac{\sqrt{n}(\overline{X}_n - \mu)}{\sigma'}$$

  has the t_{n-1} distribution.
- We can therefore calculate $\gamma = P(-c < U < c)$. Turning this around, we get

$$\gamma = P\left(\overline{X}_n - c\frac{\sigma'}{\sqrt{n}} < \mu < \overline{X}_n + c\frac{\sigma'}{\sqrt{n}}\right)$$

Confidence Interval - Mean of a Normal Distribution

- Let $T_m(x)$ denote the cdf of the t_m distribution.
- Given γ we can find c so that $P(-c < U < c) = \gamma$:

$$\gamma = P(-c < U < c) = 2T_{n-1}(c) - 1$$

since the t distribution is symmetric around 0. Solving for c we get

$$c = T_{n-1}^{-1}\left(\frac{\gamma + 1}{2}\right)$$

where T_{n-1}^{-1} is the quantile function for the t_{n-1} distribution.

- So a $100\gamma\%$ confidence interval for μ is

$$\left(\overline{X}_n - T_{n-1}^{-1}\left(\frac{\gamma + 1}{2}\right)\frac{\sigma'}{\sqrt{n}},\ \overline{X}_n + T_{n-1}^{-1}\left(\frac{\gamma + 1}{2}\right)\frac{\sigma'}{\sqrt{n}}\right)$$

Example – Hotdogs

Exercise 8.5.7 in the book

Data on calorie content in 20 different beef hot dogs from *Consumer Reports* (June 1986 issue):

$$186, 181, 176, 149, 184, 190, 158, 139, 175, 148,$$
$$152, 111, 141, 153, 190, 157, 131, 149, 135, 132$$

Assume that these numbers are observed values from a random sample of twenty independent $N(\mu, \sigma^2)$ random variables, where μ and σ^2 are unknown.

- Observed sample mean and σ' are

$$\overline{X}_n = 156.85 \quad \text{and} \quad \sigma' = 22.64201$$

- Find a 95% confidence interval for μ

Interpretation of a confidence interval

Confidence intervals are a Frequentist tool

We know that

$$P\left(\overline{X}_n - T_{n-1}^{-1}\left(\frac{\gamma+1}{2}\right)\frac{\sigma'}{\sqrt{n}} < \mu < \overline{X}_n + T_{n-1}^{-1}\left(\frac{\gamma+1}{2}\right)\frac{\sigma'}{\sqrt{n}}\right) = \gamma$$

After observing the data we observe the random interval

- For example: $(146.25, 167.45)$ is an observed 95% confidence interval for μ
- That does NOT mean that $P(146.25 < \mu < 167.45) = 0.95$.
  For this statement to make sense we need Bayesian thinking and Bayesian methods.

Interpretation of a confidence interval

Confidence intervals are a Frequentist tool

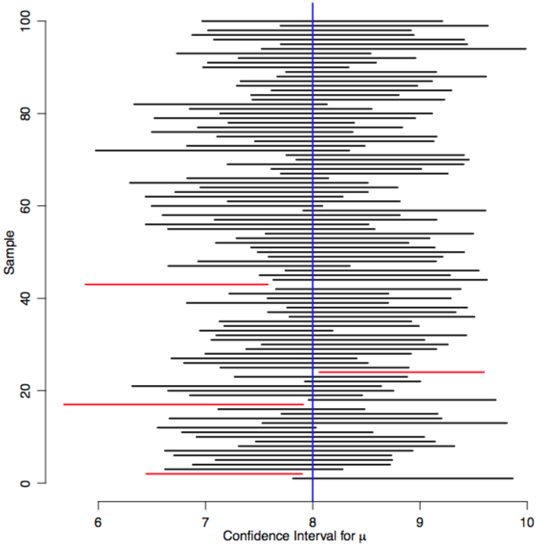
One way of thinking of this: Repeated samples.

- Take a random sample of size n from $N(\mu, \sigma^2)$ and calculate the 95% confidence interval
- Take another random sample (of the same size n) and do the same calculations.
- Repeat. Many times.
- Since there is a 95% chance that the random intervals cover the value of μ we expect 95% of the intervals to cover the actual value of μ

Problem: We never take more than one sample!

Properties of a confidence interval - Simulation Study

- I simulated n=20 r.v. from $N(8, 2^2)$ and calculated the 95% CI
- I repeated that 100 times
- 4 of the 100 intervals do not cover $\mu = 8$ (red intervals)



Non-symmetric confidence intervals

Mean of the normal distribution

More generally we want to find

$$P(c_1 < U < c_2) = \gamma$$

- *Symmetric confidence interval*: Equal probability on either side:

$$P(U \le c_1) = P(U \ge c_2) = \frac{1-\gamma}{2}$$

- Since the distribution of U is symmetric around 0, the shortest possible for μ is the symmetric confidence interval.
- *One-sided confidence interval*: All the extra probability is on one side.
  That is, either $c_1 = \infty$ or $c_2 = \infty$

One-sided Confidence Interval

Def: Lower bound

Let A be a statistic so that

$$P(A < g(\theta)) \geq \gamma \quad \forall \theta$$

- The random interval (A, ∞) is a *one-sided* $100\gamma\%$ *confidence interval for* $g(\theta)$
- A is a $100\gamma\%$ ***lower*** *confidence limit for* $g(\theta)$

One-sided Confidence Interval

Def: Upper bound

Let B be a statistic so that

$$P(g(\theta) < B) \geq \gamma \quad \forall\theta$$

- The random interval $(-\infty, B)$ is a *one-sided* $100\gamma\%$ *confidence interval for* $g(\theta)$.
- B is a $100\gamma\%$ ***upper*** *confidence limit for* $g(\theta)$

One-sided Confidence Interval - Mean of a normal

- Let $X_1, \ldots, X_n$ be a random sample from $N(\mu, \sigma^2)$, both μ and σ^2 unknown.
- Find the one-sided $100\gamma\%$ confidence intervals for μ
- Find the observed 95% upper confidence limit for μ for the hotdog example.

Confidence intervals for other distributions

Def: Pivotal

Let $\mathbf{X} = (X_1, \ldots, X_n)$ be a random sample from a distribution that depends on parameter θ. Let $V(\mathbf{X}, \theta)$ be a random variable whose distribution is the same for all θ. Then V is called a *pivotal quantity*.

To use this we need to be able to invert the pivotal relationship: find a function $r(v, \mathbf{x})$ so that

$$r(V(\mathbf{X}, \theta), \mathbf{X}) = g(\theta)$$

If the r function is increasing in v for every $\mathbf{x}$, V has a continuous distribution with cdf $F(v)$ and $\gamma_2 - \gamma_1 = \gamma$, then

$$A = r\left(F^{-1}(\gamma_1), \mathbf{X}\right) \quad \text{and} \quad B = r\left(F^{-1}(\gamma_2), \mathbf{X}\right)$$

are the endpoints of an exact $100\gamma\%$ confidence interval (Theorem 8.5.3).

Confidence interval using Pivotal quantities

Example: The rate parameter θ of the exponential distribution

- $X_1, \ldots, X_n$ i.i.d. $\mathrm{Expo}(\theta)$
- Find the $\gamma\%$ upper confidence limit for θ
- Find a symmetric $\gamma\%$ confidence interval for θ

Example: Variance of the normal distribution

- $X_1, \ldots, X_n$ i.i.d. $N(\mu, \sigma^2)$, both unknown.
- Find a symmetric $\gamma\%$ confidence interval for σ^2
- Find the observed symmetric $\gamma\%$ confidence interval for σ^2 for the hotdog example

Problems with interpretation of a confidence interval

- Example 8.5.11 is an interesting example.
- Say X_1, X_2 are i.i.d. $\text{Uniform}(\theta - 0.5, \theta + 0.5)$
- Let $Y_1 = \min(X_1, X_2)$ and $Y_2 = \max(X_1, X_2)$.
  Then (Y_1, Y_2) is a 50% confidence interval for θ
- However: If we observe Y_1 and Y_2 that are more than 0.5 apart, that is $y_2 - y_1 > 0.5$ then we know for a certainty that (y_1, y_2) contains θ! Yet we only assign 50% "confidence" to that interval, which ignores information we have.