6. Hypothesis testing and confidence intervals

We have the following goals:

- Understand how to construct confidence intervals for parameters in a linear model.
- Understand how to test statistical hypotheses about a linear model.
- In particular, to ask and answer the question: "Are the data consistent with a hypothesis that a covariate, or a collection of covariates, are unimportant?" (What is the fundamental scientific importance of the slightly contorted logical reasoning in this question?)
- Learn to use R to carry out these tasks.
- See how the linear model includes and extends basic tests for means of one and two samples.

First, we'll review hypothesis testing by working through some notes on "Topics in comparing means of one or two samples."

Confidence intervals

- An interval $[u, v]$ constructed using the data $\mathbf{y}$ is said to **cover** a parameter θ if $u \leq \theta \leq v$.
- $[u, v]$ is a 95% **confidence interval** (CI) for θ if the same construction, applied to a large number of draws from the model, would cover θ 95% of the time.
- A **parameter** is a name for any unknown constant in a model. In linear models,each component $\beta_1, \ldots, \beta_p$ of the **coefficient vector** $\boldsymbol{\beta}$ is a parameter. So is the variance σ^2 of the measurement error.
- A confidence interval is the usual way to represent the amount of uncertainty in an estimated parameter.
- The parameter is not random. According to the model, it has a fixed but unknown value.The observed interval $[u, v]$ is also not random. An interval $[U, V]$ constructed using a realization $\mathbf{Y}$ from the model is random.
- If the model is appropriate, then it is reasonable to treat the data $\mathbf{y}$ like a realization from the model.

A confidence interval for the coefficient of a linear model

- Consider estimating β_1 in the linear model $\mathbf{Y} = \mathbb{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$.
- Recall that $\mathrm{E}[\hat{\beta}_1] = \beta_1$ and $\mathrm{SD}(\hat{\beta}_1) = \sigma\sqrt{\left[\left(\mathbb{X}^{\mathrm{T}}\mathbb{X}\right)^{-1}\right]_{11}}$.

Question 6.1. Supposing we can make a normal approximation, show that $\mathrm{P}\big[\,\hat{\beta}_1 - 1.96\,\mathrm{SD}(b_1) \le \beta_1 \le \hat{\beta}_1 + 1.96\,\mathrm{SD}(b_1)\,\big] = 0.95$

- Therefore, an approximate 95% CI for β_1 is

$$\boxed{\big[\,b_1 - 1.96\,\mathrm{SE}(b_1)\,,\; b_1 + 1.96\,\mathrm{SE}(b_1)\,\big]}$$

where $\mathbf{y} = \mathbb{X}\mathbf{b} + \mathbf{e}$ with $\mathrm{SE}(b_1) = s\sqrt{\left[\left(\mathbb{X}^{\mathrm{T}}\mathbb{X}\right)^{-1}\right]_{11}}$.

A CI for association between unemployment and mortality

-