Math 505 Notes

Chapter 5

Jesse Crawford

Department of Mathematics
Tarleton State University

Fall 2009

Outline

1. Sections 5.1 and 5.4: Confidence Intervals
2. Sections 5.5 and 5.6: Hypothesis Tests

Definition

- Consider a statistical model with unknown parameter $\theta \in \mathbb{R}$.
- Let $X_1, \ldots, X_n$ be a sample from the model.
- Suppose (a, b) is a random interval based on this sample such that

$$P(a < \theta < b) = 1 - \alpha, \text{ where } \alpha \in [0, 1].$$

- Then (a, b) is a *confidence interval for* θ with *confidence level* or *confidence coefficient* $1 - \alpha$.

Example

- Consider a statistical model with finite positive variance σ^2 and mean μ.
- Let $X_1, \ldots, X_n$ be a large sample ($n \geq 30$).
- Let $\bar{X}_n$ and S_n be the sample mean and sample standard deviation, respectively.
- An *approximate* $1 - \alpha$ confidence interval for μ is

$$\left(\bar{X}_n - z_{\alpha/2}\frac{S_n}{\sqrt{n}}, \bar{X}_n + z_{\alpha/2}\frac{S_n}{\sqrt{n}}\right),$$

- where $z_{\alpha/2}$ is the $1 - \alpha/2$ quantile of a standard normal distribution, i.e.,

$$P(Z \leq z_{\alpha/2}) = 1 - \alpha/2.$$

Example

- Suppose the population is normally distributed, i.e., the model is

$$\{N(\mu, \sigma^2) \mid \mu \in \mathbb{R}, \sigma^2 > 0\}.$$

- Then,

$$\frac{\bar{X}_n - \mu}{S_n/\sqrt{n}}$$

  has a t-distribution with $n-1$ degrees of freedom (see Section 3.6).

- An *exact* $1-\alpha$ confidence interval for μ is

$$\left(\bar{X}_n - t_{\alpha/2,n-1}\frac{S_n}{\sqrt{n}}, \bar{X}_n + t_{\alpha/2,n-1}\frac{S_n}{\sqrt{n}}\right),$$

- where $t_{\alpha/2,n-1}$ is the $1-\alpha/2$ quantile of a t-distribution with $n-1$ degrees of freedom.
- This confidence interval is valid for any sample size $n \geq 2$.

Definition

Suppose Z and V are independent random variables such that $Z \sim N(0,1)$ and $V \sim \chi^2(n-1)$. Then the random variable

$$T = \frac{Z}{\sqrt{V/(n-1)}}$$

has a t-distribution with $n-1$ degrees of freedom.

Theorem (Student's Theorem)

Let $X_1, \ldots, X_n$ be a random sample from the distribution $N(\mu, \sigma^2)$, and let $\bar{X}$ and S^2 be the sample mean and variance.

- $\bar{X} \sim N(\mu, \sigma^2/n)$ *and* $(n-1)S^2/\sigma^2 \sim \chi^2(n-1)$
- $\bar{X}$ *and* S^2 *are independent*
- $$T = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$ *has a t-distribution with $n-1$ degrees of freedom.*

Example

- Consider two normally distributed populations with means μ_1 and μ_2 and the same variance σ^2.
- Let $X_1, \ldots, X_{n_1}$ and $Y_1, \ldots, Y_{n_2}$ be independent samples from these populations, and let $n = n_1 + n_2$.
- Define the *pooled sample variance* to be

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n - 2}.$$

- Then a $1 - \alpha$ confidence interval for $\mu_1 - \mu_2$ is

$$\bar{X} - \bar{Y} \pm t_{\alpha/2, n-2} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}.$$

Example

- Let $Y \sim b(n, p)$, where $n \in \mathbb{N}$ is known and $p \in (0, 1)$ is unknown.
- This is the setting of estimating an unknown population proportion p based on a sample of size n.
- For large sample sizes, an approximate $1 - \alpha$ confidence interval for p is

$$\hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}},$$

- where $\hat{p} = Y/n$ is the proportion of successes in the sample.
- (The approximation is conventionally considered valid when $n\hat{p} \geq 5$ and $n(1 - \hat{p}) \geq 5$, although some authors replace 5 with a more conservative value, such as 15.)
- Confidence interval for the difference of two proportions based on two independent samples:

$$\hat{p}_1 - \hat{p}_2 \pm z_{\alpha/2} \sqrt{\frac{\hat{p}_1(1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2(1 - \hat{p}_2)}{n_2}}.$$

Outline

1. Sections 5.1 and 5.4: Confidence Intervals

2. Sections 5.5 and 5.6: Hypothesis Tests

Definition

- Consider a statistical model $(P_\theta \mid \theta \in \Omega)$.
- Suppose that $\Omega = \omega_0 \cup \omega_1$, where $\omega_0 \cap \omega_1 = \emptyset$.
- We have two competing *statistical hypotheses*

$$\mathrm{H}_0 : \theta \in \omega_0 \text{ vs. } \mathrm{H}_1 : \theta \in \omega_1 .$$

- Exactly one of these hypotheses is true, and determining which one is true on the basis of sample data is a *statistical testing problem*, or *hypothesis testing problem*.
- The hypothesis $\mathrm{H}_0 : \theta \in \omega_0$ is called the *null hypothesis*, and the hypothesis $\mathrm{H}_1 : \theta \in \omega_1$ is called the *alternative hypothesis*.
- A procedure that selects one of these hypotheses on the basis of sample data is a *hypothesis test*.
- Given a sample $X_1, \ldots, X_n$ from the model, a test is determined by a critical region $C \subset \mathbb{R}^n$ and the following procedure:
  - If $(X_1, \ldots, X_n) \in C$, we reject H_0.
  - If $(X_1, \ldots, X_n) \notin C$, we do not reject H_0.

- Two types of errors:
  - Type I error: Rejecting H_0 when it is true.
  - Type II error: Not rejecting H_0 when it is false.
- The probability of making a type I error is at most

$$\alpha = \max_{\theta \in \omega_0} P_\theta[(X_1, \dots, X_n) \in C].$$

- α is called the *size* or *significance level* of the test. Common value of α are 0.05 and 0.01.
- The probability of rejecting the null hypothesis when the parameter is θ is

$$\gamma(\theta) = P_\theta[(X_1, \dots, X_n) \in C]$$

- The function $\gamma : \Omega \to [0, 1]$ is called the *power function*.
  - If $\theta \in \omega_0$, then $\gamma(\theta)$ is the probability of making a type I error.
  - If $\theta \in \omega_1$, then $\gamma(\theta)$ is the probability of *NOT* making a type II error.

- If $\theta \in \omega_0$, then $\gamma(\theta)$ is the probability of making a type I error.
- If $\theta \in \omega_1$, then $\gamma(\theta)$ is the probability of *NOT* making a type II error.
- We want $\gamma(\theta)$ to be small for $\theta \in \omega_0$.
- Since

$$\max_{\theta \in \omega_0} \gamma(\theta) = \alpha,$$

where α is small, this condition is met.

- We also want $\gamma(\theta)$ to be large for $\theta \in \omega_1$.

Definition

- Consider two tests C_1 and C_2 with significance level α and power functions γ_1 and γ_2.
- If $\gamma_1(\theta) \geq \gamma_2(\theta)$, for every $\theta \in \omega_1$, then C_1 is *uniformly more powerful* than C_2, and is clearly a better test than C_2.
- If C_1 is uniformly more powerful than C_2 for *any* test C_2 with significance level α, then C_1 is a *uniformly most powerful test*, and is the "optimal" test at significance level α.

Example

- Suppose $X_1, \ldots, X_n$ is a random sample from a $N(\mu, \sigma^2)$ distribution, and let $\mu_0 \in \mathbb{R}$.
- For testing

$$\mathrm{H}_0 : \mu = \mu_0 \text{ vs. } \mathrm{H}_1 : \mu \neq \mu_0,$$

- the following test has significance level α

$$\text{Reject } \mathrm{H}_0 \text{ if } |t| = \left| \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \right| \geq t_{\alpha/2, n-1}.$$

Definition

- Consider a hypothesis test with test statistic t.
- Let t_0 be an observed value of t.
- The *p*-value, or *observed significance level*, corresponding to t_0 is the probability of observing a value of t “more extreme” than t_0 under the null hypothesis.
- The null hypothesis is rejected if the *p*-value is less than or equal to α.