Day 8-1

Prakash Balachandran
Department of Mathematics & Statistics
Boston University

Monday, October 24, 2011

Section 5.1: Estimating $\mu$

- Example: (putting everything together) Below are the sale prices of 12 randomly selected 3-year-old Chevy Corvettes. Construct a 90% confidence interval for both the population mean and variance. Conclusions?

\[
\begin{align*}
41844, & 41500, 39995, 36995, 40990, 37995, 41995, 38900, 42995, 36995, 43995, 35950 \\
\end{align*}
\]

Solution:
So the data is approximately normal and has no outliers.
\[
\bar{X} = 40,012, \ s = 2,615.20, \ n = 12, \ t_{1-\frac{\alpha}{2}} = 1.796, \\
\]
\[
\bar{X} \pm t_{1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}} = 40,012 \pm 1.796 \frac{2,615.20}{\sqrt{12}} = 40,012 \pm 1,355.90 \\
\]
is the confidence interval for the mean.
For the confidence interval for the variance, $c_{1-\frac{\alpha}{2}} = 4.575, c_{\frac{\alpha}{2}} = 19.675,$
\[
\frac{(n-1) s^2}{c_{\frac{\alpha}{2}}} \leq \sigma^2 \leq \frac{(n-1) s^2}{c_{1-\frac{\alpha}{2}}} \\
\Rightarrow 11 \frac{2615.20^2}{19.675} \leq \sigma^2 \leq 11 \frac{2615.20^2}{4.575}. \\
\Rightarrow 3.8237 \times 10^6 \leq \sigma^2 \leq 1.6444 \times 10^7 \\
\Rightarrow 1,955 \leq \sigma \leq 4,055
\]
Section 5.2: Hypothesis Tests for $\mu$

- A *hypothesis* is a statement or claim regarding a characteristic of two or more populations (in this case, the population mean of a single population).
- *Hypothesis-testing* is a procedure, based on sample evidence and probability, used to test claims regarding the characteristic.
- Procedure of hypothesis testing:
  1. Determine the null and alternative hypotheses,
  2. Understand Type I and Type II errors,
  3. Analyze and state the conclusions to the hypothesis tests.

1: Determining the Null and Alternative Hypotheses

- The *null hypothesis* denoted $H_0$ is a statement to be tested. The null hypothesis is assumed true until evidence indicates otherwise. In the case of the population mean, it is a statement about the value of the population mean.
- The *alternative hypothesis* denoted $H_1$ is a claim to be tested. We are trying to find evidence for the alternative hypothesis. In the case of the population mean, it will be a claim regarding the value of the population parameter.
- Examples:
  1. The Blue Book value of a 3-year-old Chevy Corvette is $37,500. You wish to test the claim that the mean price of a used 3-year-old Corvette in the Boston area is different than $37,500.
  2. The packaging on a light bulb states that the bulb will last 500 hours under normal use. A consumer would like to know if the mean lifetime of a club is less than 500 hours.
  3. The average return for a certain class of mutual funds is $1,000. A mutual fund manage claims that his return is more than this amount.
- The previous examples illustrate the 3 categories that hypothesis tests can split up into. For two fixed numbers $\mu_0$ and $\mu_1$:
  1. (two-tailed test)

\[
H_0 : \mu = \mu_0 \\
H_1 : \mu \neq \mu_1
\]

  2. (left-tailed test)

\[
H_0 : \mu = \mu_0 \\
H_1 : \mu < \mu_1
\]

  3. (right-tailed test)

\[
H_0 : \mu = \mu_0
\]
$H_1 : \mu > \mu_1$

## 2: Understanding Type I and Type II Errors

- There are four outcomes in hypothesis testing:
  1. We reject $H_0$ when $H_1$ is true.
  2. We reject $H_1$ when $H_0$ is true.
  3. We reject $H_0$ when $H_0$ is true. This is called a *Type I Error*.
  4. We accept $H_0$ when $H_0$ is false ($H_1$ is true). This is called a *Type II Error*.

- We denote:

  $\alpha = \mathbb{P} [\text{Type I Error}] = \mathbb{P} [\text{Reject } H_0 | H_0 \text{ True}]$.

  $\beta = \mathbb{P} [\text{Type II Error}] = \mathbb{P} [\text{Accept } H_0 | H_1 \text{ True}]$

and refer to $\alpha$ as the *level of significance*, or probability of making a Type I error. As in Confidence Intervals, this will be the minimal error we want to make.

- A pharmaceutical company has developed a new antibiotic. Among competing antibiotics, an average of 200 out of 1000 children who take the antibiotics experience headaches as a side effect. The FDA tests,

  $H_0 : \mu = 200,$

  $H_1 : \mu > 200.$

What would it mean to make a Type I versus a Type II error?

- Someone asked in class today what $\mathbb{P} [H_0 \text{ false}, H_1 \text{ false}]$ would mean. This doesn’t quite fit into the four cases above, since we’re always *conditioning* on either $H_0$ or $H_1$ being actually true, so we’re not really interested in this probability. On the other hand, we are very interested in

  $\mathbb{P} [H_0 \text{ false}|H_1 \text{ false}] = \mathbb{P} [H_0 \text{ false}|H_0 \text{ true}]$

would be a Type I error, and

  $\mathbb{P} [H_1 \text{ false}|H_0 \text{ false}] = \mathbb{P} [H_1 \text{ false}|H_1 \text{ true}]$

would be a Type II error, which we’ve already accounted for.