Confidence intervals and tests of hypothesis, chapter 4.3

Grethe Hystad

October 16, 2012

Example

The math SAT score is approximately normally distributed with a mean of 500 and standard deviation of 100. A high school claim that the students from their high school have a higher average math SAT score than 500. That is, if μ is the mean, they claim that $\mu > 500$. This conjecture is called a statistical hypothesis. We want to test the parameter of the distribution, in this case the mean. Some people don't believe in this claim, and so they think that the mean math SAT score of the students from this high school will be the same or maybe even lower, that is, at most $\mu = 500$. We want to test this.

We want to test:

$$H_0 : \mu = 500 \quad \text{against} \quad H_1 : \mu > 500.$$

- H_0 is called the **null hypothesis**
- H_1 is called the **alternative hypothesis**

Assume the standard deviation is about the same at $\sigma = 100$. We take a random sample, $X_1, ..., X_{49}$, of $n = 49$ math SAT scores from students from this high school. Then $\overline{X}$ is approximately normally distributed with mean μ and standard deviation $\frac{\sigma}{\sqrt{n}} = \frac{100}{\sqrt{49}}$, i.e. $\overline{X}$ is $N(\mu, \frac{10000}{49})$. Suppose $\overline{x} = 530$.

Question: Do we accept or reject the null-hypothesis?

- Consider $P\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \leq z_\alpha\right) = 1 - \alpha$. which is equivalent to

- Consider $P\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \leq z_{\alpha}\right) = 1-\alpha.$ which is equivalent to
- $P\left(\overline{X} - z_{\alpha}\frac{\sigma}{\sqrt{n}} \leq \mu\right) = 1-\alpha.$ That is, we are $100(1-\alpha)\%$ confident that the new mean is above $\overline{X} - z_{\alpha}\frac{\sigma}{\sqrt{n}}$.

- Consider $P\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \leq z_\alpha\right) = 1-\alpha.$ which is equivalent to
- $P\left(\overline{X} - z_\alpha \frac{\sigma}{\sqrt{n}} \leq \mu\right) = 1-\alpha.$ That is, we are $100(1-\alpha)\%$ confident that the new mean is above $\overline{X} - z_\alpha \frac{\sigma}{\sqrt{n}}$.
- If $\alpha = 0.05$, the critical value is $z_{0.05} = 1.645$ $(qnorm(0.95) = 1.645)$

- Consider $P\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \leq z_\alpha\right) = 1-\alpha.$ which is equivalent to
- $P\left(\overline{X} - z_\alpha \frac{\sigma}{\sqrt{n}} \leq \mu\right) = 1-\alpha.$ That is, we are $100(1-\alpha)\%$ confident that the new mean is above $\overline{X} - z_\alpha \frac{\sigma}{\sqrt{n}}$.
- If $\alpha = 0.05$, the critical value is $z_{0.05} = 1.645$ $(qnorm(0.95) = 1.645)$
- If $n = 49$, $\sigma = 100$, $\overline{x} = 53$, we have $P\left(530 - 1.645 \cdot \frac{100}{7} \leq \mu\right) = P(506.5 \leq \mu) = 0.95$

- Consider $P\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \leq z_\alpha\right) = 1-\alpha.$ which is equivalent to
- $P\left(\overline{X} - z_\alpha \frac{\sigma}{\sqrt{n}} \leq \mu\right) = 1-\alpha.$ That is, we are $100(1-\alpha)\%$ confident that the new mean is above $\overline{X} - z_\alpha \frac{\sigma}{\sqrt{n}}$.
- If $\alpha = 0.05$, the critical value is $z_{0.05} = 1.645$ $(qnorm(0.95) = 1.645)$
- If $n = 49$, $\sigma = 100$, $\overline{x} = 53$, we have $P\left(530 - 1.645 \cdot \frac{100}{7} \leq \mu\right) = P(506.5 \leq \mu) = 0.95$
- Thus, we are 95% confident that the true mean of the math SAT score of the students from this high school is above 506.5.

- Consider $P\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \leq z_\alpha\right) = 1-\alpha.$ which is equivalent to
- $P\left(\overline{X} - z_\alpha \frac{\sigma}{\sqrt{n}} \leq \mu\right) = 1-\alpha.$ That is, we are $100(1-\alpha)\%$ confident that the new mean is above $\overline{X} - z_\alpha \frac{\sigma}{\sqrt{n}}$.
- If $\alpha = 0.05$, the critical value is $z_{0.05} = 1.645$ $(qnorm(0.95) = 1.645)$
- If $n = 49$, $\sigma = 100$, $\overline{x} = 53$, we have $P\left(530 - 1.645 \cdot \frac{100}{7} \leq \mu\right) = P(506.5 \leq \mu) = 0.95$
- Thus, we are 95% confident that the true mean of the math SAT score of the students from this high school is above 506.5.
- A confidence interval for μ is $(506.5, \infty)$ and we see that 500 is not in that interval. Thus, we are 95% confident that the high school is correct, that their mean is significantly higher than 500. Hence we accept the alternative hypothesis, $H_1 : \mu > 500$. (Reject the null-hypothesis).

Type I and II errors

In general in testing:

$$H_0 : \mu = \mu_0 \quad \textit{against} \quad H_1 : \mu > \mu_0$$

we can make two types of errors:

- Rejecting the null hypothesis when it is true is called a **type I error**.
- Accepting the null hypothesis when it is false is called a **type II error**.

The probability of type I error is called the **significance level,** α**,** of the test.

Critical region

There is a critical region, C, such that if the data $(x_1, ..., x_n) \in C$, then H_0 is rejected and if $(x_1, ..., x_n) \notin C$, then H_0 is accepted.

- The probability of type I error is

$$\alpha = P[(X_1, ..., X_n) \in C; H_0]$$

  which is the probability that the data $(X_1, ..., X_n)$ is in C when H_0 is true. (Rejecting the hypothesis H_0 when it is true.)

Critical region

There is a critical region, C, such that if the data $(x_1, ..., x_n) \in C$, then H_0 is rejected and if $(x_1, ..., x_n) \notin C$, then H_0 is accepted.

- The probability of type I error is

  $$\alpha = P[(X_1, ..., X_n) \in C; H_0]$$

  which is the probability that the data $(X_1, ..., X_n)$ is in C when H_0 is true. (Rejecting the hypothesis H_0 when it is true.)
- The probability of type II error is

  $$\beta = P[(X_1, ..., X_n) \notin C; H_1]$$

  which is the probability that the data $(X_1, ..., X_n)$ is not in C when H_1 is true. (Accepting the null hypothesis when it is false)

Critical region

- We specify the critical region by the value of the test statistics, $\overline{X}$, for which H_0 is rejected. We define

  $$C = \{\overline{x} : \overline{x} \geq c_\alpha\}$$

  for a nonnegative constant, c_α, where

  $$P\left(\overline{X} \geq c_\alpha; H_0\right) = \alpha.$$

  Thus, the probability that $\overline{X} \in C$ if H_0 is true is α.

Critical region

- We specify the critical region by the value of the test statistics, $\overline{X}$, for which H_0 is rejected. We define

$$C = \{\overline{x} : \overline{x} \geq c_\alpha\}$$

  for a nonnegative constant, c_α, where

$$P\left(\overline{X} \geq c_\alpha; H_0\right) = \alpha.$$

  Thus, the probability that $\overline{X} \in C$ if H_0 is true is α.
- When H_0 is true, $\overline{X}$ is $N(\mu_0, \sigma^2/n)$. Let $Z = \frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$ which is then $N(0,1)$.

Critical region

- We specify the critical region by the value of the test statistics, $\overline{X}$, for which H_0 is rejected. We define

$$C = \{\overline{x} : \overline{x} \geq c_\alpha\}$$

  for a nonnegative constant, c_α, where

$$P\left(\overline{X} \geq c_\alpha; H_0\right) = \alpha.$$

  Thus, the probability that $\overline{X} \in C$ if H_0 is true is α.
- When H_0 is true, $\overline{X}$ is $N(\mu_0, \sigma^2/n)$. Let $Z = \frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$ which is then $N(0,1)$.
- We want to determine the critical value c_α.

Let z_α so that $P(Z \geq z_\alpha) = \alpha$. Then

$$\alpha = P\left(\frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} \geq z_\alpha\right) = P\left(\overline{X} \geq \mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}}\right),$$

where

$$c_\alpha = \mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}}.$$

Test of hypothesis

One sided composite hypothesis test about mean with known variance:

$$H_0 : \mu = \mu_0 \quad \textit{against} \quad H_1 : \mu > \mu_0.$$

We reject H_0 at the $(\alpha \times 100)\%$ significance level if

$$z \geq z_\alpha$$

which is equivalent to

$$\bar{x} \geq \mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}}.$$

Test of hypothesis

- In the previous example with $n = 49$, $\sigma = 100$, $\mu_0 = 500$, $\overline{x} = 530$, $\alpha = 0.05$,
  we reject H_0 if $Z \geq z_{0.05} = 1.645$. We have

$$Z = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{530 - 500}{100/7} = 2.1 > 1.645 = z_{0.05}.$$

  Thus, we reject H_0 at the 5% significance level.

Test of hypothesis

- In the previous example with $n = 49$, $\sigma = 100$, $\mu_0 = 500$, $\overline{x} = 530$, $\alpha = 0.05$,
  we reject H_0 if $Z \geq z_{0.05} = 1.645$. We have
  $$Z = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{530 - 500}{100/7} = 2.1 > 1.645 = z_{0.05}.$$
  Thus, we reject H_0 at the 5% significance level.
- Alternatively we could compute
  $$c_{0.05} = \mu_0 + z_{0.05}\frac{\sigma}{\sqrt{n}} = 500 + 1.645\frac{100}{7} = 523.5.$$
  Since $\overline{X} = 530 > 523.5$, we reject the null hypothesis. Thus $P(\overline{X} \geq 523.5; H_0) = 0.05$. Thus, if H_0 is true, there is only 5% chance that the mean $\overline{X}$ is greater than 523.5.

$$H_0 : \mu = \mu_0 \quad \textit{against} \quad H_1 : \mu < \mu_0$$

Let z_α so that $P(Z \leq -z_\alpha) = \alpha$. Then

$$\alpha = P\left(\frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} \leq -z_\alpha\right) = P\left(\overline{X} \leq \mu_0 - z_\alpha \frac{\sigma}{\sqrt{n}}\right),$$

where

$$c_\alpha = \mu_0 - z_\alpha \frac{\sigma}{\sqrt{n}}.$$

Test of hypothesis

One sided composite hypothesis test about mean with known variance:

$$H_0 : \mu = \mu_0 \quad \textit{against} \quad H_1 : \mu < \mu_0.$$

We reject H_0 at the $(\alpha \times 100)\%$ significance level if

$$z \leq -z_\alpha$$

which is equivalent to

$$\overline{x} \leq \mu_0 - z_\alpha \frac{\sigma}{\sqrt{n}}.$$

$$H_0 : \mu = \mu_0 \quad \textit{against} \quad H_1 : \mu \neq \mu_0$$

Let $z_{\alpha/2}$ so that $P(|Z| \geq z_{\alpha/2}) = \alpha$. Then

$$\alpha = P\left(\left| \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} \right| \geq z_{\alpha/2} \right) = P\left(\mu_0 + z_{\alpha/2} \frac{\sigma/2}{\sqrt{n}} \leq \overline{X} \leq \mu_0 - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right).$$

This is equivalent to rejecting H_0 if μ_0 is not in the confidence interval $\left(\overline{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}},\ \overline{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$.

Test of hypothesis

Two sided composite hypothesis test about mean with known variance:

$$H_0 : \mu = \mu_0 \quad \textit{against} \quad H_1 : \mu \neq \mu_0.$$

We reject H_0 at the $(\alpha \times 100)\%$ significance level if

$$|z| \geq z_{\alpha/2}$$

which is equivalent to

$$|\overline{x} - \mu_0| \geq z_{\alpha/2} \frac{\sigma}{\sqrt{n}}.$$

Hypothesis

- A **Simple hypothesis** is a hypothesis that completely specifies the distribution of X. For example testing

$$H_0 : \mu = 3 \quad \text{against} \quad H_1 : \mu = 5$$

- A **composite hypothesis** is composed of at least two simple hypothesis. For example testing

$$H_0 : \mu = 3 \quad \text{against} \quad H_1 : \mu > 3.$$

Example

Let X be the resistance of a thermistor at temperature 25^0 C. The resistance of thermistors from manufacturer A is normally distributed with mean $10,000$ ohm and standard deviation 4000 ohm. It is hoped that a different manufacturer, called manufacturer B, produces thermistors that is normally distributed with mean $12,000$ ohm and standard deviation 4000 ohm. That is we want to construct a hypothesis test to test whether the increase in mean of the resistance was realized. We will construct the simple hypothesis:

$$H_0 : \mu_0 = 10,000 \quad \text{against} \quad H_1 : \mu_1 = 12,000.$$

We take a random sample of $n = 25$ thermistors. We have $\overline{X}$ is $N(10000, \frac{4000^2}{25})$ if H_0 is true and $\overline{X}$ is $N(12000, \frac{4000^2}{25})$ if H_1 is true. We want the critical region to be defined by $\overline{x} > 11,500$.

example continues

Example

(A) Find the significance level α for the test.
(B) Find the probability for type II error.

Solution

Answers:
(A) $\alpha = P(\overline{X} \geq 11,500; H_0) = 0.0304$
(B) $\beta = P(\overline{X} < 11,500; H_1) = 0.2660$
Details are given in class.

example continues

Example

In the previous example we want to keep $\alpha = 0.0304$ but we want to lower β to $\beta = 0.10$. This will require a larger sample size than $n = 25$ and a different critical region, $C = \{\overline{x} : \overline{x} \geq c_{0.034}\}$.
Find n and $c_{0.0304}$ so that the probability for type I and II errors are as given above.

Solution

Answer: $n = 40$ *and* $c_{0.0304} = 11,186$
Details are given in class.