LECTURE 10
HYPOTHESIS TESTING: ONE-SAMPLE Z-TEST

PSY2002

Hye Won Suk

STATISTICAL INFERENCE

- Last time we learned estimation. From this lecture we will start learning hypothesis testing.

- **Statistical Inference**
  - Estimation
    - Point estimation
    - Interval estimation
  - Hypothesis Testing
    - Z-test
    - t-test
    - Analysis of Variance (ANOVA)
    - And so on…

ONE-SAMPLE TESTS

- A one-sample (or single-sample) test is used when a single sample is collected and the researcher wants to examine if there is strong evidence that a parameter (e.g., population mean) of the population, from which the sample is extracted, is different from a hypothesized value.

- Let's look at some examples of one-sample tests.

EXAMPLE 1

- A car manufacturer claims that the fuel efficient of a new model Q is normally distributed with mean of $\mu = 35$ miles per gallon on the highway, and a standard deviation of $\sigma = 5$. The American Automobile Association randomly selects $n = 16$ new Qs and measures their fuel efficiency. The average efficiency is $\bar{X} = 33$ miles per gallon. Based on this, can you say that the car manufacturer is truthful in advertising Q's fuel efficiency as being 35 miles per gallon?

Bowen, C-C. (2016). *Straightforward Statistics*. Sage: Thousand Oaks, CA.

EXAMPLE 2

- National data on student loans indicate that the student loan amount has a normal distribution with mean of $\mu = \$30{,}000$ and a standard deviation of $\sigma = \$8{,}000.$ A random sample of $n = 25$ students from a private university reported a mean student debt load of $\bar{X} = \$34{,}500.$ Is the average student loan from this private university higher than $\$30{,}000$?

Bowen, C-C. (2016). *Straightforward Statistics*. Sage: Thousand Oaks, CA.

A WORKING EXAMPLE: PRENATAL CARE STUDY

- Birthweight is one of the best indicators of the health of a baby. It is also an excellent predictor of some difficulties that infants may experience in their first weeks of life.

- In the United States, mothers who live in poverty generally have babies with lower birthweight than those who do not live in poverty.

- It is known that the birthweight for infants of women living in poverty is normally distributed with mean $\mu = 2800\ g$ and standard deviation $\sigma = 500\ g$.

http://ccnmtl.columbia.edu/projects/qmss/the_ttest/onesample_ttest.html

A WORKING EXAMPLE: PRENATAL CARE STUDY

- Recently, a local hospital introduced an innovative new prenatal care program to reduce the number of low birthweight babies born in the hospital. 25 mothers, all of whom live in poverty, participated in this program. The babies born to these women had an average birthweight of $\bar{X} = 3075\ g$.

- The question is whether this program has been effective at improving the birthweights of babies born to poor women.

http://ccnmtl.columbia.edu/projects/qmss/the_ttest/onesample_ttest.html

Population of babies born to poor mothers who did *not* participate in the prenatal care program

Sample of babies born to poor mothers who participated in the prenatal care program

$$\mu = 2800$$
$$\sigma = 500$$

$$\bar{X} = 3075$$
$$n = 25$$

- Based on this sample, can we conclude that the prenatal care program is effective?
- In other words, can we conclude that this sample is not from the population with $\mu = 2800$?

TWO TYPES OF ONE-SAMPLE TEST

- Two types of one-sample tests:
  - A <u>one sample *Z*-test</u> is used when the population standard deviation (σ) is known.
  - A <u>one sample *t*-test</u> is used when the population standard deviation (σ) is NOT known.

One-sample *Z*-test

THE LOGIC OF HYPOTHESIS TESTING

First, the researcher assumes that the prenatal care has NO effect on birthweight ($\mu_{prenatal} = 2800\ g$).

The research gathers evidence (the sample mean) to see if the prenatal care has an effect on birthweight.

If there is enough evidence (if the sample mean is noticeably different from $2800\ g$), the researcher concludes that there really is an effect.

If there is not enough evidence (if the sample mean is around $2800\ g$), the researcher concludes that there is not enough evidence to say that there is an effect.

AN ANALOG FOR HYPOTHESIS TESTING

In a jury trial, the trial begins with an assumption that the defendant did not commit a crime (innocent until proven guilty).

The prosecutor gathers evidence to show that the defendant really did commit a crime.

If there is enough evidence, the jury rejects the assumption and concludes that the defendant is guilty of a crime.

If there is not enough evidence, the jury fails to find the defendant guilty. Note that the jury does not conclude that the defendant is innocent, simply that there is not enough evidence for a guilty verdict.

BASIC STEPS OF A HYPOTHESIS TEST

- Whatever hypothesis test you use, you should always follow these five steps.

  - Step 1: State the hypotheses
  - Step 2: Set the criteria for a decision
  - Step 3: Collect data and compute test statistics
    - *Z*-test: calculate *Z*-statistic
  - Step 4: Make a decision
  - Step 5: State a conclusion

STEP 1: STATE THE HYPOTHESES

- There are two opposing hypotheses.

  - The null hypothesis (H_0)
  - The alternative hypothesis (H_1)

- The null hypothesis (H_0) states that in the general population there is <u>no change</u>, <u>no difference</u>, or <u>no relationship</u>.

- The alternative hypothesis (H_1) is the opposite to the null hypothesis. We accept this alternative hypothesis when the null hypothesis is rejected.

- In the working example,
  - The null hypothesis (H_0):
    - The prenatal care program has <u>no effect</u> on the birthweights of babies.
    - The population mean birthweight of babies born to poor mothers who participate in the program <u>is the same as</u> that of babies born to poor mothers who do not participate in the program, which is 2800 grams.
    - $\mu_{prenatal} = 2800\ g$

- The alternative hypothesis (H_1):
  - The prenatal care program has an effect on birthweights of babies.
  - The population mean birthweight of babies born to poor mothers who participate in the program is different from that of babies born to poor mothers who do not participate in the program ($2800\ g$).
  - $\mu_{prenatal} \neq 2800\ g$
- Note that the alternative hypothesis simply states that there is some change. It does not specify the amount of change. It also does not specify the direction of change (increase or decrease).

STEP 2: SET THE CRITERIA FOR A DECISION

- Step 2: Set the criteria for a decision.
  - $\alpha = 0.05$
  - The alpha level (or level of significance) indicates that the extreme 5% of the scores will be treated as the values that are too extreme to be observed when the null hypothesis is true.
  - By convention, we use $\alpha = 0.05$ unless otherwise specified.

STEP 3

- Step 3: Compute a test statistic (or *Z*-statistic).

  - If the null hypothesis is true (the prenatal care program has no effect), the followings will hold (recall the central limit theorem).

  - The sample mean birth weight of 25 babies will be normally distributed since the original distribution of birth weights is a normal distribution.

  - The expected value of mean $(\mu_{\bar{X}}) = \mu = 2800$
  - The standard error of mean $(\sigma_{\bar{X}}) = \frac{\sigma}{\sqrt{n}} = \frac{500}{\sqrt{25}} = \frac{500}{5} = 100$

Population of babies born to poor mothers who did *not* participate in the prenatal care program

Sample of babies born to poor mothers who participated in the prenatal care program

$$\mu = 2800$$
$$\sigma = 500$$

$$\bar{X} = 3075$$
$$n = 25$$

The sample means (of sample size 25) form a normal distribution with mean of 2800 and standard deviation of 100.

- Step 3: Compute the test statistic (or Z-statistic).

$$Z = \frac{\bar{X} - \mu_{\bar{X}}}{\sigma_{\bar{\mathrm{X}}}} = \frac{3075 - 2800}{100} = \frac{275}{100} = 2.75$$

  - When we assume that the null hypothesis is true (the prenatal care has no effect), the Z-score corresponding to the sample mean of $3075\ g$ is 2.75.

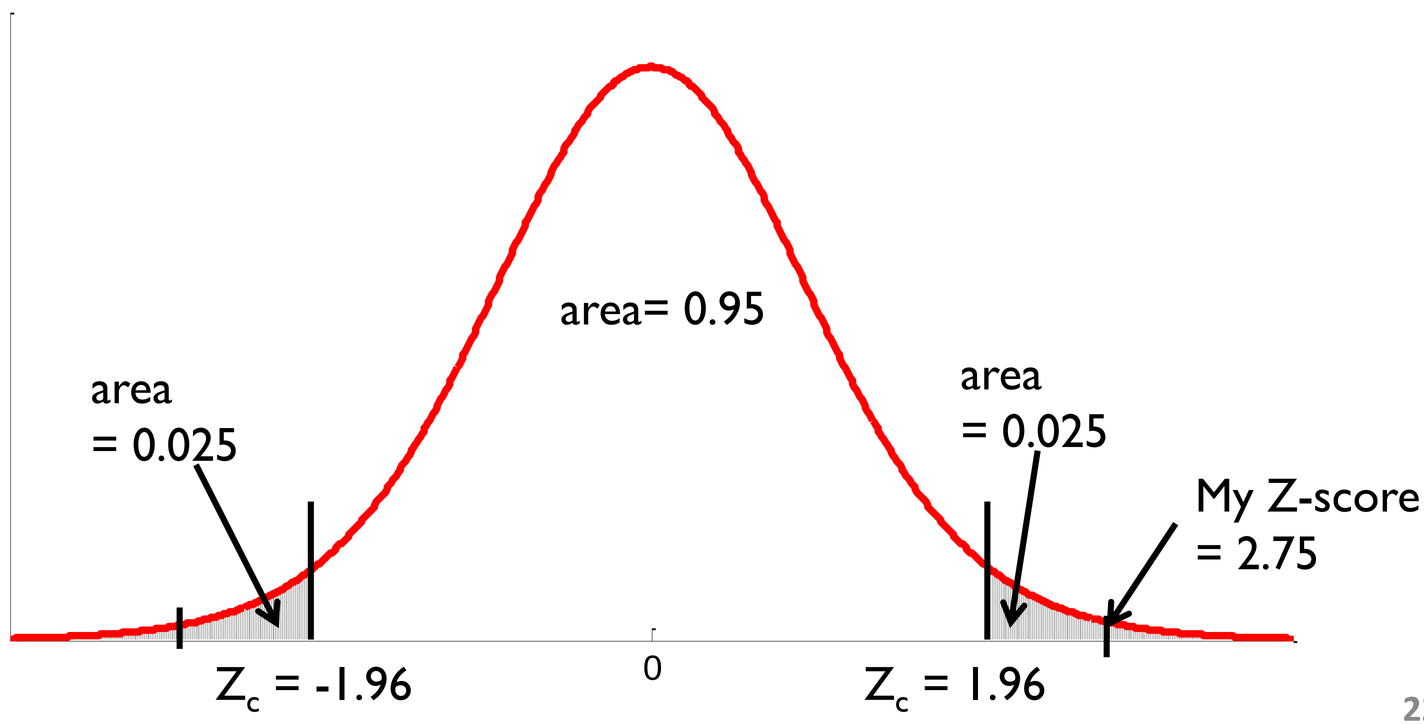
- The Z-score follows the standard normal distribution because…

    - X (original birthweight) forms a normal distribution
    - Therefore, $\bar{X}$ forms a normal distribution (by the central limit theorem).
    - Therefore, Z (the standardized score of $\bar{X}$) forms the standard normal distribution.

    - Therefore, we can use the standard normal distribution to obtain the probability for the Z-score.

STEP 4

- Step 4: Make a decision.
  - My *Z*-score is in the critical region (it belongs to the extreme 5%). That is, it is *very unlikely* to obtain $Z = 2.75$ when the null hypothesis is true. My *Z*-score provides strong evidence against the null hypothesis. ➔ Reject H_0.



- Another equivalent way to make a decision is to obtain the *p*-value and compare it to the level of significance ($\alpha = 0.05$)

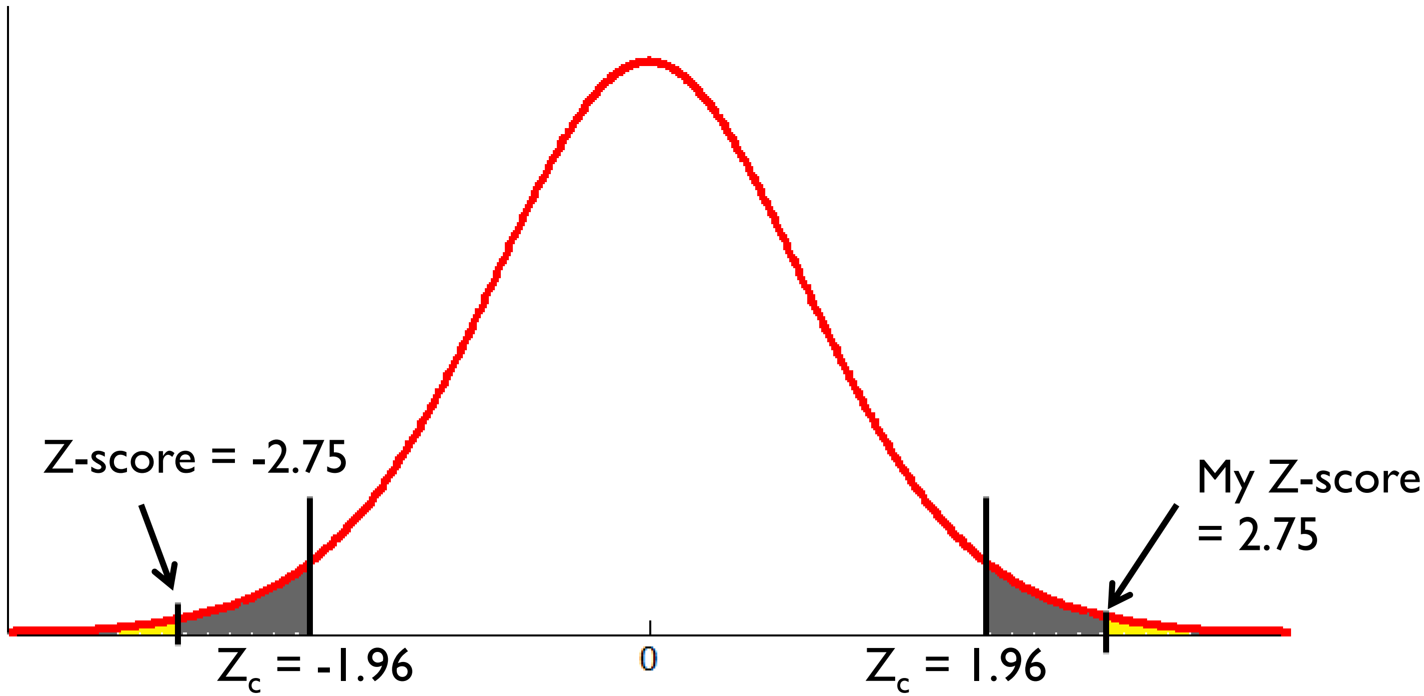
- $p = P(Z \leq -2.75 \text{ or } Z \geq 2.75) = 2(.0030) = .0060$

- My *p*-value (p=.0060) is smaller than the level of significance ($\alpha = 0.05$). It indicates that my *Z*-score falls within the extreme 5%. It means that it is *very unlikely* (less than 5% chance) to obtain my *Z*-score when H_0 is true.

- The data provide strong evidence against the null hypothesis.

- Therefore, we reject the null hypothesis.

- p = .006 (yellow area)



STEP 5

- Step 5. State a conclusion.
  - The prenatal care has a <u>significant effect</u> on the birthweights of babies born to poor women (Z = 2.75, p=.006).

- In the conclusion, you need to report a verbal interpretation of the result as well as *Z* and *p* values.

- *Z* = 2.75 indicates that a *Z*-score was used as the test statistic (i.e., Z-test was used) and its value was 2.75.

- *p*=.006 indicates that the probability of observing my *Z*-score (or more extreme values) under the H_0 is very small (less than .05), i.e., smaller than the alpha level. It is *very unlikely* to obtain my *Z*-score when the null hypothesis is true.

SUMMARY

- The basic five steps of hypothesis testing
  - Step 1: State the hypotheses
  - Step 2: Set the criteria for a decision
    - Set the significance level (α) and obtain critical values.
  - Step 3: Collect data and compute sample statistics
    - *Z*-test: calculate *Z*-statistic
  - Step 4: Make a decision
  - Step 5: State a conclusion
    - Report a verbal interpretation, *Z*-score, and *p*-value.