

MATH 450: Mathematical statistics

November 19th, 2020

Lecture 22: Inferences based on two samples

Week 2	●	Chapter 6: Statistics and Sampling Distributions
Week 4	●	Chapter 7: Point Estimation
Week 7	●	Chapter 8: Confidence Intervals
Week 10	●	Chapter 9: Tests of Hypotheses
Week 12	●	Chapter 10: Two-sample testing
Week 14	●	Regression

Key steps in statistical inference

- Understand statistical models [Chapter 6]
- Come up with reasonable estimates of the parameters of interest [Chapter 7]
- Quantify the confidence with the estimates [Chapter 8]
- Testing with one sample [Chapter 9]
- Testing with two samples [Chapter 10]

Contexts

- The central mega-example: population mean μ
- Difference between two population means

Inferences based on two samples

10.1 Difference between two population means

- z-test
- confidence intervals

10.2 The two-sample t test and confidence interval

10.3 Analysis of paired data

Chapter 9: Hypothesis testing (with one sample)

Test about a population mean

- Null hypothesis

$$H_0 : \mu = \mu_0$$

- The alternative hypothesis will be either:

- $H_a : \mu > \mu_0$
- $H_a : \mu < \mu_0$
- $H_a : \mu \neq \mu_0$

- Three settings

- normal population with known σ
- large-sample tests
- a normal population with unknown σ

Hypothesis testing for one parameter

- 1 Identify the parameter of interest
- 2 Determine the null value and state the null hypothesis
- 3 State the appropriate alternative hypothesis
- 4 Give the formula for the test statistic
- 5 State the rejection region for the selected significance level α
- 6 Compute statistic value from data
- 7 Decide whether H_0 should be rejected and state this conclusion in the problem context

Sample solution

- Parameter of interest: $\mu =$ true average activation temperature
- Hypotheses

$$H_0 : \mu = 130$$

$$H_a : \mu \neq 130$$

- Test statistic:

$$z = \frac{\bar{x} - 130}{1.5/\sqrt{n}}$$

- Rejection region: either $z \leq -z_{0.005}$ or $z \geq z_{0.005} = 2.58$
- Substituting $\bar{x} = 131.08$, $n = 25 \rightarrow z = 2.16$.
- Note that $-2.58 < 2.16 < 2.58$. We fail to reject H_0 at significance level 0.01.
- The data does not give strong support to the claim that the true average differs from the design value.

DEFINITION

The ***P*-value** (or *observed significance level*) is the smallest level of significance at which H_0 would be rejected when a specified test procedure is used on a given data set. Once the *P*-value has been determined, the conclusion at any particular level α results from comparing the *P*-value to α :

1. $P\text{-value} \leq \alpha \Rightarrow$ reject H_0 at level α .
2. $P\text{-value} > \alpha \Rightarrow$ do not reject H_0 at level α .

Testing by P-value method

DECISION
RULE BASED
ON THE
P-VALUE

Select a significance level α (as before, the desired type I error probability).
Then reject H_0 if $P\text{-value} \leq \alpha$; do not reject H_0 if $P\text{-value} > \alpha$

Remark: the smaller the P-value, the more evidence there is in the sample data against the null hypothesis and for the alternative hypothesis.

Sample solution

1. Parameter of interest: $\mu =$ true average wafer thickness
2. Null hypothesis: $H_0: \mu = 245$
3. Alternative hypothesis: $H_a: \mu \neq 245$
4. Formula for test statistic value: $z = \frac{\bar{x} - 245}{s/\sqrt{n}}$
5. Calculation of test statistic value: $z = \frac{246.18 - 245}{3.60/\sqrt{50}} = 2.32$
6. Determination of P -value: Because the test is two-tailed,
$$P\text{-value} = 2[1 - \Phi(2.32)] = .0204$$
7. Conclusion: Using a significance level of .01, H_0 would not be rejected since $.0204 > .01$. At this significance level, there is insufficient evidence to conclude that true average thickness differs from the target value.

P-values for t -tests

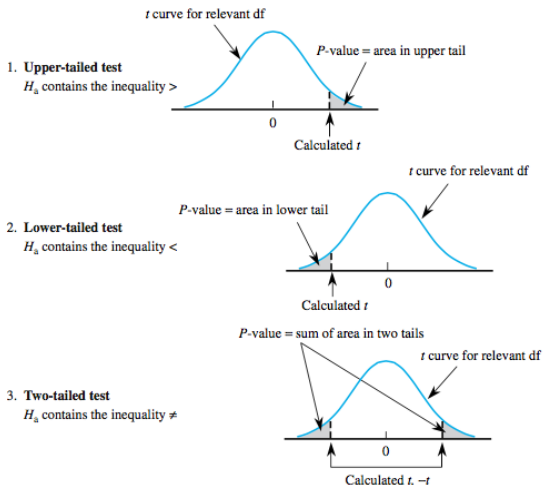


Figure 9.8 P -values for t tests

A P-value:

- is not the probability that H_0 is true
- is not the probability of rejecting H_0
- is the probability, calculated assuming that H_0 is true, of obtaining a test statistic value at least as contradictory to the null hypothesis as the value that actually resulted

Example 3

A sample of 12 radon detectors of a certain type was selected, and each was exposed to 100 pCi/L of radon. The resulting readings were as follows:

105.6 90.9 91.2 96.9 96.5 91.3

100.1 105.0 99.6 107.7 103.3 92.4

Does this data suggest that the population mean reading under these conditions differs from 100? What is the P-value of the test? What is the conclusion with level of significance $\alpha = 0.05$

Example 4

On the label, Pepperidge Farm bagels are said to weigh four ounces each (113 grams). A random sample of six bagels resulted in the following weights (in grams):

117.6 109.5 111.6 109.2 119.1 110.8

- a.** Based on this sample, is there any reason to doubt that the population mean is at least 113 grams?

Two-sample inference

Example

Let μ_1 and μ_2 denote true average decrease in cholesterol for two drugs. From two independent samples X_1, X_2, \dots, X_m and Y_1, Y_2, \dots, Y_n , we want to test:

$$H_0 : \mu_1 = \mu_2$$

$$H_a : \mu_1 \neq \mu_2$$

- This lecture: independent samples

Assumption

- 1 X_1, X_2, \dots, X_m is a random sample from a population with mean μ_1 and variance σ_1^2 .
 - 2 Y_1, Y_2, \dots, Y_n is a random sample from a population with mean μ_2 and variance σ_2^2 .
 - 3 The X and Y samples are independent of each other.
- Next lecture: paired-sample test

Problem

Assume that

- X_1, X_2, \dots, X_m is a random sample from a population with mean μ_1 and variance σ_1^2 .
- Y_1, Y_2, \dots, Y_n is a random sample from a population with mean μ_2 and variance σ_2^2 .
- The X and Y samples are independent of each other.

Compute (in terms of $\mu_1, \mu_2, \sigma_1, \sigma_2, m, n$)

- (a) $E[\bar{X} - \bar{Y}]$
- (b) $\text{Var}[\bar{X} - \bar{Y}]$ and $\sigma_{\bar{X} - \bar{Y}}$

Proposition

The expected value of $\bar{X} - \bar{Y}$ is $\mu_1 - \mu_2$, so $\bar{X} - \bar{Y}$ is an unbiased estimator of $\mu_1 - \mu_2$. The standard deviation of $\bar{X} - \bar{Y}$ is

$$\sigma_{\bar{X}-\bar{Y}} = \sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}$$

Normal distributions with known variances

Chapter 8: Confidence intervals

Assume further that the distributions of X and Y are normal and σ_1, σ_2 are known:

Problem

(a) *What is the distribution of*

$$\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}}$$

(b) *Compute*

$$P \left[-1.96 \leq \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}} \leq 1.96 \right]$$

(c) *Construct a 95% CI for $\mu_1 - \mu_2$ (in terms of $\bar{x}, \bar{y}, m, n, \sigma_1, \sigma_2$).*

Confidence intervals

When both population distributions are normal, standardizing $\bar{X} - \bar{Y}$ gives a random variable Z with a standard normal distribution. Since the area under the z curve between $-z_{\alpha/2}$ and $z_{\alpha/2}$ is $1 - \alpha$, it follows that

$$P\left(-z_{\alpha/2} < \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}} < z_{\alpha/2}\right) = 1 - \alpha$$

Manipulation of the inequalities inside the parentheses to isolate $\mu_1 - \mu_2$ yields the equivalent probability statement

$$P\left(\bar{X} - \bar{Y} - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}} < \mu_1 - \mu_2 < \bar{X} - \bar{Y} + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}\right) = 1 - \alpha$$

Testing the difference between two population means

- Setting: independent normal random samples X_1, X_2, \dots, X_m and Y_1, Y_2, \dots, Y_n with known values of σ_1 and σ_2 . Constant Δ_0 .
- Null hypothesis:

$$H_0 : \mu_1 - \mu_2 = \Delta_0$$

- Alternative hypothesis:

(a) $H_a : \mu_1 - \mu_2 > \Delta_0$

(b) $H_a : \mu_1 - \mu_2 < \Delta_0$

(c) $H_a : \mu_1 - \mu_2 \neq \Delta_0$

- When $\Delta = 0$, the test (c) becomes

$$H_0 : \mu_1 = \mu_2$$

$$H_a : \mu_1 \neq \mu_2$$

Testing the difference between two population means

Problem

Assume that we want to test the null hypothesis $H_0 : \mu_1 - \mu_2 = \Delta_0$ against each of the following alternative hypothesis

(a) $H_a : \mu_1 - \mu_2 > \Delta_0$

(b) $H_a : \mu_1 - \mu_2 < \Delta_0$

(c) $H_a : \mu_1 - \mu_2 \neq \Delta_0$

by using the test statistic:

$$z = \frac{(\bar{x} - \bar{y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}}.$$

What is the rejection region in each case (a), (b) and (c)?

Proposition

Null hypothesis: $H_0: \mu_1 - \mu_2 = \Delta_0$

Test statistic value: $z = \frac{\bar{x} - \bar{y} - \Delta_0}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}}$

Alternative Hypothesis

$$H_a: \mu_1 - \mu_2 > \Delta_0$$

$$H_a: \mu_1 - \mu_2 < \Delta_0$$

$$H_a: \mu_1 - \mu_2 \neq \Delta_0$$

Rejection Region for Level α Test

$$z \geq z_\alpha \text{ (upper-tailed test)}$$

$$z \leq -z_\alpha \text{ (lower-tailed test)}$$

$$\text{either } z \geq z_{\alpha/2} \text{ or } z \leq -z_{\alpha/2} \text{ (two-tailed test)}$$

Practice problem

Each student in a class of 21 responded to a questionnaire that requested their GPA and the number of hours each week that they studied. For those who studied less than 10 h/week the GPAs were

2.80, 3.40, 4.00, 3.60, 2.00, 3.00, 3.47, 2.80, 2.60, 2.00

and for those who studied at least 10 h/week the GPAs were

3.00, 3.00, 2.20, 2.40, 4.00, 2.96, 3.41, 3.27, 3.80, 3.10, 2.50

Assume that the distribution of GPA for each group is normal and both distributions have standard deviation $\sigma_1 = \sigma_2 = 0.6$. Treating the two samples as random, is there evidence that true average GPA differs for the two study times? Carry out a test of significance at level .05.

1. The parameter of interest is $\mu_1 - \mu_2$, the difference between true mean GPA for the < 10 (conceptual) population and true mean GPA for the ≥ 10 population.
2. The null hypothesis is $H_0: \mu_1 - \mu_2 = 0$.
3. The alternative hypothesis is $H_a: \mu_1 - \mu_2 \neq 0$; if H_a is true then μ_1 and μ_2 are different. Although it would seem unlikely that $\mu_1 - \mu_2 > 0$ (those with low study hours have higher mean GPA) we will allow it as a possibility and do a two-tailed test.
4. With $\Delta_0 = 0$, the test statistic value is

$$z = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}}$$

5. The inequality in H_a implies that the test is two-tailed. For $\alpha = .05$, $\alpha/2 = .025$ and $z_{\alpha/2} = z_{.025} = 1.96$. H_0 will be rejected if $z \geq 1.96$ or $z \leq -1.96$.

6. Substituting $m = 10$, $\bar{x} = 2.97$, $\sigma_1^2 = .36$, $n = 11$, $\bar{y} = 3.06$, and $\sigma_2^2 = .36$ into the formula for z yields

$$z = \frac{2.97 - 3.06}{\sqrt{\frac{.36}{10} + \frac{.36}{11}}} = \frac{-.09}{.262} = -.34$$

That is, the value of $\bar{x} - \bar{y}$ is only one-third of a standard deviation below what would be expected when H_0 is true.

7. Because the value of z is not even close to the rejection region, there is no reason to reject the null hypothesis. This test shows no evidence of any relationship between study hours and GPA. ■

Large-sample tests/confidence intervals

- Central Limit Theorem: \bar{X} and \bar{Y} are approximately normal when $n > 30 \rightarrow$ so is $\bar{X} - \bar{Y}$. Thus

$$\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}}$$

is approximately standard normal

- When n is sufficiently large $S_1 \approx \sigma_1$ and $S_2 \approx \sigma_2$
- Conclusion:

$$\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{m} + \frac{S_2^2}{n}}}$$

is approximately standard normal when n is sufficiently large

If $m, n > 40$, we can ignore the normal assumption and replace σ by S

Proposition

Use of the test statistic value

$$z = \frac{\bar{x} - \bar{y} - \Delta_0}{\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}}$$

along with the previously stated upper-, lower-, and two-tailed rejection regions based on z critical values gives large-sample tests whose significance levels are approximately α . These tests are usually appropriate if both $m > 40$ and $n > 40$. A P -value is computed exactly as it was for our earlier z tests.

Proposition

Provided that m and n are both large, a CI for $\mu_1 - \mu_2$ with a confidence level of approximately $100(1 - \alpha)\%$ is

$$\bar{x} - \bar{y} \pm z_{\alpha/2} \sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}$$

where $-$ gives the lower limit and $+$ the upper limit of the interval. An upper or lower confidence bound can also be calculated by retaining the appropriate sign and replacing $z_{\alpha/2}$ by z_{α} .

Example

Let μ_1 and μ_2 denote true average tread lives for two competing brands of size P205/65R15 radial tires.

(a) Test

$$H_0 : \mu_1 = \mu_2$$

$$H_a : \mu_1 \neq \mu_2$$

at level 0.05 using the following data: $m = 45$, $\bar{x} = 42,500$, $s_1 = 2200$, $n = 45$, $\bar{y} = 40,400$, and $s_2 = 1900$.

(b) Construct a 95% CI for $\mu_1 - \mu_2$.