MAT2377

Ali Karimnezhad

Version December 15, 2016

Comments

- These slides cover material from Chapter 6.

- In class, I may use a blackboard. I recommend reading these slides before you come to the class.

- I am planning to spend 3 lectures on this chapter.

- I am not re-writing the textbook. The reference book contains many interesting and practical examples.

- There may be some typos. The final version of the slides will be posted *after* the chapter is finished.

The Null and Alternative Hypotheses

- A statistical hypothesis is an assertion or conjecture concerning one or more populations.

- Null hypothesis, refers to any hypothesis we wish to test and is denoted by H_0

- The rejection of H_0 leads to the acceptance of an alternative hypothesis, denoted by H_1.

- The analyst arrives at one of the two following conclusions:

  - Reject H_0 in favor of H_1 because of sufficient evidence in the data; or

  - Fail to reject H_0 because of insufficient evidence in the data.

Examples:

Consider the null hypothesis that the average weight of male students in a certain college is 68 kilograms against the alternative hypothesis that it is unequal to 68. Assume the standard deviation of the population of weights to be $\sigma = 3.6$. We wish to test

$$H_0 : \qquad \mu = 68$$

$$H_1 : \qquad \mu \neq 68 \qquad (i.e., \mu < 68 \text{ or } \mu > 68).$$

- A sample mean (test statistic) that falls close to the hypothesized value of 68 would be considered evidence in favor of H_0.

- A sample mean that is considerably less than or more than 68 would be evidence inconsistent with H_0 and therefore favoring H_1.

- A critical region for the test statistic might arbitrarily be chosen to be the two intervals $\bar{x} < 67$ and $\bar{x} > 69$. The nonrejection region will then be the interval $67 \leq \bar{x} \leq 69$.

Type I and II errors

$$\alpha = P(\text{Type I Error}) = P(\text{reject } H_0 | H_0 \text{ is true}),$$

$$\beta = P(\text{Type II Error}) = P(\text{do not reject } H_0 | H_0 \text{ is not true}),$$

$$\beta^* = 1 - \beta = P(\text{reject } H_0 | H_0 \text{ is not true}).$$

Example (Continued):

$$\begin{aligned} \alpha \ &= \ P(\bar{X} < 67 \text{ or } \bar{X} > 69 | H_0 \text{ is true}) \\ &= \ P(\bar{X} < 67 \text{ or } \bar{X} > 69 | \mu = 68) = \ldots = 0.0950. \end{aligned}$$

The Approach to Hypothesis Testing with α

1. State the null and alternative hypotheses.

2. Choose a fixed significance level α.

3. Choose an appropriate test statistic and establish the critical region based on α.

4. Reject H_0 if the computed test statistic is in the critical region. Otherwise, do not reject.

5. Draw scientific or engineering conclusions.

The Approach to Hypothesis Testing with $pvalue$

A $pvalue$ is the lowest level (of significance) at which the observed value of the test statistic is significant.

1. State the null and alternative hypotheses.

2. Choose an appropriate test statistic.

3. Compute the $pvalue$ based on the computed value of the test statistic.

4. Use judgment based on the $pvalue$ and knowledge of the scientific system in deciding whether or not to reject H_0.

One- and Two-Sided Tests

One-Sided Tests

- $H_0 : \mu = \mu_0 \quad vs \quad H_1 : \mu = \mu_1 \quad (\mu_1 > \mu_0)$
  $H_0 : \mu = \mu_0 \quad vs \quad H_1 : \mu > \mu_0$

- $H_0 : \mu = \mu_0 \quad vs \quad H_1 : \mu = \mu_1 \quad (\mu_1 < \mu_0)$
  $H_0 : \mu = \mu_0 \quad vs \quad H_1 : \mu < \mu_0$

Two-Sided Test

- $H_0 : \mu = \mu_0 \quad vs \quad H_1 : \mu \neq \mu_0$

Examples:

- We are interested in the mean burning rate of a solid propellant used to power aircrew escape systems. We want to determine whether or not the mean burning rate is 50 cms/second. The sample of 10 specimens is tested and we observe $\bar{x} = 48.5$. Assume normality and $\sigma = 2.5$. State the null and alternative hypotheses to be used in testing this claim and determine where the critical region is located.

- A group of 100 adult American Catholics have been asked: 'Do you favor allowing women to be priests?'. 60 of them answered 'YES'. Is this the strong evidence that more than half American Catholics favor allowing women to be priests? State the null and alternative hypotheses to be used in this test and determine the location of the critical region.

Tests Concerning a Single Mean

Two-sided case with known variance:

Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from the Normal population with unknown mean μ and known variance σ^2.

We wish to test

$$H_0 : \mu = \mu_0 \quad vs \quad H_1 : \mu \neq \mu_0.$$

Remember that, a $100(1 - \alpha)\%$ confidence interval for μ is given by

$$\bar{x} - z_{\frac{\alpha}{2}} \sigma / \sqrt{n} < \mu < \bar{x} + z_{\frac{\alpha}{2}} \sigma / \sqrt{n} \quad \text{or} \quad -z_{\frac{\alpha}{2}} < \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} < z_{\frac{\alpha}{2}},$$

where $z_{\frac{\alpha}{2}}$ is the z-value leaving an area of $\frac{\alpha}{2}$ to the right.

Define $z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$. If $z > z_{\frac{\alpha}{2}}$ or $z < -z_{\frac{\alpha}{2}}$, reject H_0.

Define $pvalue = 2P(Z > |z|)$. If $\alpha > pvalue$, reject H_0.

Two-sided case with unknown variance:

Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from the Normal population with unknown mean μ and unknown variance σ^2.
We wish to test

$$H_0 : \mu = \mu_0 \quad vs \quad H_1 : \mu \neq \mu_0.$$

Remember that, a $100(1 - \alpha)\%$ confidence interval for μ is given by

$$\bar{x} - t_{\frac{\alpha}{2}}(n-1)\frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\frac{\alpha}{2}}(n-1)\frac{s}{\sqrt{n}} \quad \text{or} \quad -t_{\frac{\alpha}{2}}(n-1) < \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} < t_{\frac{\alpha}{2}}(n-1),$$

where $t_{\frac{\alpha}{2}}(n-1)$ is the t-value leaving an area of $\frac{\alpha}{2}$ to the right with $n-1$ degrees of freedom.

Define $t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$. If $t > t_{\frac{\alpha}{2}}(n-1)$ or $t < -t_{\frac{\alpha}{2}}(n-1)$, reject H_0.
Define $pvalue = 2P(T > |t|)$. If $\alpha > pvalue$, reject H_0.

$\boxed{\text{Examples:}}$

- Test the hypothesis that the average content of containers of a particular lubricant is 10 liters if the contents of a random sample of 10 containers are 10.2, 9.7, 10.1, 10.3, 10.1, 9.8, 9.9, 10.4, 10.3, and 9.8 liters.
  Use a 0.01 level of significance and assume that the distribution of contents is normal.

One-sided case with known variance:

Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from the Normal population with unknown mean μ and known variance σ^2.

- To test
$$H_0 : \mu = \mu_0 \quad vs \quad H_1 : \mu > \mu_0,$$
Define $z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$. If $z > z_\alpha$, reject H_0.
Define $pvalue = P(Z > z)$. If $\alpha > pvalue$, reject H_0.

- To test
$$H_0 : \mu = \mu_0 \quad vs \quad H_1 : \mu < \mu_0,$$
Define $z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$. If $z < -z_\alpha$, reject H_0.
Define $pvalue = P(Z < z)$. If $\alpha > pvalue$, reject H_0.

One-sided case with unknown variance:

Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from the Normal population with unknown mean μ and unknown variance σ^2.

- To test
$$H_0 : \mu = \mu_0 \quad vs \quad H_1 : \mu > \mu_0,$$
Define $t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$. If $t > t_\alpha(n-1)$, reject H_0.
Define $pvalue = P(T > t)$. If $\alpha > pvalue$, reject H_0.

- To test
$$H_0 : \mu = \mu_0 \quad vs \quad H_1 : \mu < \mu_0,$$
Define $t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$. If $t < -t_\alpha(n-1)$, reject H_0.
Define $pvalue = P(T < t)$. If $\alpha > pvalue$, reject H_0.

Example:

- Components are manufactured to have strength normally distributed with mean $\mu = 40$ units and standard deviation $\sigma = 1.2$ units. A modification have been tried, for which an increase in mean strength is claimed (the standard deviation remains the same). A random sample of $n = 12$ components produced using the modified process had strength

  42.5, 39.8, 40.3, 43.1, 39.6, 41.0, 39.9, 42.1, 40.7, 41.6, 42.1, 40.8,

  Do the data provide strong evidence that the mean strength exceeds 40 units? ($Use\ \alpha = 0.05$)

$$H_0 : \mu = 40 \quad vs \quad H_1 : \mu > 40$$

Observed value of the sample mean is $\bar{x} = 41.125$. Hence

$$
\begin{aligned}
pvalue \quad &= \quad P(\bar{X} \geq 41.125) = P(\bar{X} \geq 41.125) \\
&= \quad P\left(\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \geq \frac{41.125 - \mu_0}{\sigma/\sqrt{n}} \right) = P(Z \geq 3.25) \approx 0.006.
\end{aligned}
$$

So, $pvalue$ is smaller than α - reject H_0.

If the model '$\mu = 40$' is true, the event $\{\bar{X} \geq 41.125\}$ is very unlikely.

Example:

- Components are manufactured to have strength normally distributed with mean $\mu = 40$ units and standard deviation $\sigma = 1.2$ units. A modification have been tried, for which an increase in mean strength is claimed (the standard deviation remains the same). A random sample of $n = 12$ components produced using the modified process had strength

  ```
  42.5, 39.8, 40.3, 43.1, 39.6, 41.0, 39.9, 42.1, 40.7,
  41.6, 42.1, 40.8,
  ```

  Do the data provide strong evidence that the mean strength is not equal to 40 units? $(Use\ \alpha = 0.05)$

$$H_0 : \mu = 40 \quad vs \quad H_1 : \mu \neq 40$$

Observe that value of the sample mean is $\bar{x} = 41.125$, and $z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = 3.25$. Hence

$$pvalue \quad = \quad 2P(Z > |z|) = 2P(Z \geq 3.25) \approx 0.012.$$

So, $pvalue$ is smaller than α - reject H_0.

Example:

- The Edison Electric Institute has published figures on the number of kilowatt hours used annually by various home appliances. It is claimed that a vacuum cleaner uses an average of 46 kilowatt hours per year. If a random sample of 12 homes included n a planned study indicates that vacuum cleaners use an average of 42 kilowatt hours per year with a sample standard deviation of 11.9 kilowatt hours, does this suggest at the 0.05 level of significance that vacuum cleaners use, on average, less than 46 kilowatt hours annually? Assume the population of kilowatt hours to be normal.

Example:

- Suppose we have the following data

| 18.0 | 17.4 | 15.5 | 16.8 | 19.0 | 17.8 | 17.4 | 15.8 |
| 17.9 | 16.3 | 16.9 | 18.6 | 17.7 | 16.4 | 18.2 | 18.7 |

from $N(\mu, \sigma^2)$ with completely unknown μ and σ^2. Test

$$H_0 : \mu = 16.6 \quad vs \quad H_1 : \mu > 16.6$$

Use $\alpha = 0.05$.

Observe that value of the sample mean and standard deviation are 17.4 and 1.0386, respectively. Then, $t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = 3.081$. Hence

$$pvalue \quad = \quad P(T > t) = P(T > 3.081),$$

where $T \sim t_{15}$. From the t-tables we see that

$$P\left(T \geq 2.947\right) \approx 0.005 \quad \text{and} \quad P\left(T \geq 3.286\right) \approx 0.0025\,.$$

Hence the p-value is somewhere between these, i.e. in the interval $(0.0025, 0.005)$, in particular it is ≤ 0.05, strong evidence against $H_0 \colon \mu = 16.6$.

Test on proportion

Example:

A group of 100 adult American Catholics have been asked: 'Do you favor allowing women to be priests?'. 60 of them answered 'YES'. Is this the strong evidence that more than half American Catholics favor allowing women to be priests?

- Define X as the number of people who answered YES.
  Thus, $X \sim B(100, p)$.

- $H_0 : p = 0.5$, $H_1 : p > 0.5$,

- Under H_0, $X \sim B(100, 0.5)$

- $pvalue$:

$$
\begin{aligned}
P(X \geq 60) &= 1 - P(X < 60) = 1 - P(X \leq 59) \\
&= 1 - P\left(\frac{X - np}{\sqrt{np(1-p)}} \leq \frac{59 - np}{\sqrt{np(1-p)}} \right) \\
&= 1 - P(Z \leq 1.9) = 0.0287.
\end{aligned}
$$

- Reject H_0 for $\alpha = 0.05$, don't reject for $\alpha = 0.01$.

Paired test

Assumptions:

- $X_{11}, \ldots, X_{1n}$ is a random sample from population 1,

- $X_{21}, \ldots, X_{2n}$ is a random sample from population 2,

- Two populations are normal with means μ_1 and μ_2, respectively,

- The populations are not independent ,

- Variances are unknown.

We wish to to test

$$H_0 : \mu_1 = \mu_2 \;\; vs \;\; H_1 : \mu_1 \neq \mu_2.$$

Solution: Compute differences $D_i = X_{1i} - X_{2i}$ and consider t-test as follows:

The test statistics is

$$T = \frac{\bar{D}}{S_D/\sqrt{n}} \sim t_{n-1},$$

where

$$\bar{D} = \frac{1}{n}\sum_{i=1}^{n} D_i,$$

$$S_D^2 = \frac{1}{n-1}\sum_{i=1}^{n_1}(D_i - \bar{D})^2$$

Note: You can also consider one-sided alternatives for both two-sample and paired test.

Example:

- Ten engineers' knowledge of basic statistical concepts was measured on a scale of 100 before and after a short course in statistical quality control. The result are as follows:

Engineer	1	2	3	4	5	6	7	8	9	10
Before X_{1i}	43	82	77	39	51	66	55	61	79	43
After X_{2i}	51	84	74	48	53	61	59	75	82	48

Let μ_1 and μ_2 be the mean mean score before and after the course. Test $H_0 : \mu_1 = \mu_2$ against $H_1 : \mu_1 < \mu_2$.

Solution: The differences $D_i = X_{1i} - X_{2i}$ are:

Engineer	1	2	3	4	5	6	7	8	9	10
Before X_{1i}	43	82	77	39	51	66	55	61	79	43
After X_{2i}	51	84	74	48	53	61	59	75	82	48
Difference D_i	-8	-2	3	-9	-2	5	-4	-14	-3	-5

So, based on D_i values, the observed sample mean is $\bar{D} = -2.9$, the observed sample variance is $S_D^2 = 31.21$. The test statistic is

$$T = \frac{\bar{D}}{S_D/\sqrt{n}} \sim t_{n-1}.$$

We compute

$$pvalue = P(\bar{D} \leq -2.9) = P\left(\frac{\bar{D} - 0}{S_D/\sqrt{n}} \leq \frac{-2.9}{\sqrt{31.21/10}}\right)$$
$$= P(T < -1.64) = P(T > 1.64) \in (0.05, 0.1)$$

Thus, do not reject H_0 when $\alpha = 0.05$ or $\alpha = 0.01$.

Two Sample Test

Assumptions:

- $X_{11}, \ldots, X_{1n_1}$ is a random sample from population 1,

- $X_{21}, \ldots, X_{2n_2}$ is a random sample from population 2,

- Two populations are independent.

We want to test
$$H_0 : \mu_1 = \mu_2, \qquad H_1 : \mu_1 \neq \mu_2.$$

Let
$$\bar{X}_1 = \frac{1}{n_1} \sum_{i=1}^{n_1} X_{1i}, \qquad \bar{X}_2 = \frac{1}{n_2} \sum_{i=1}^{n_2} X_{2i}.$$

Case 1: Known variances σ_1^2 and σ_2^2

Example: A researcher is interested to assess whether an income in Alberta is higher than in Ontario. A sample of 100 people from Alberta has been interviewed yielding the sample mean $\bar{x}_1 = 33000$. A sample of 80 people from Ontario yields the sample mean $\bar{x}_2 = 32000$. Do we have enough evidence that people in Alberta get more on average than in Ontario? From the previous studies it is known that the population standard deviations are respectively, $\sigma_1 = 5000$ and $\sigma_2 = 2000$.

Solution: $H_0 : \mu_1 = \mu_2$; $H_1 : \mu_1 > \mu_2$. The observed difference is 1000. To compute

$$pvalue = P\left(\bar{X}_1 - \bar{X}_2 > 1000\right) = P\left(\frac{\bar{X}_1 - \bar{X}_2 - 0}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} > \frac{1000 - 0}{\sqrt{5000^2/100 + 2000^2/80}}\right)$$

$$= P(Z > 1.82) = 0.035$$

Reject H_0 when $\alpha = 0.05$, do not reject when $\alpha = 0.01$.

Case 2: Unknown variances $\sigma_1^2 = \sigma_2^2$; small sample size

Example: A researcher wants to test that on average a new fertilizer yields taller plants. Plants were divided into two groups: a control group treated with an old fertilizer and a study group treated with the new fertilizer. The following data are obtained:

Sample size	Sample Mean	Sample Variance
$n_1 = 8$	$\bar{x}_1 = 43.14$	$s_1^2 = 71.65$
$n_2 = 8$	$\bar{x}_2 = 47.79$	$s_2^2 = 52.66$

Test $H_0 : \mu_1 = \mu_2$ vs. $H_1 : \mu_1 < \mu_2$.

Solution: $H_0 : \mu_1 = \mu_2$; $H_1 : \mu_1 < \mu_2$. The observed difference is -4.65. To compute

$$
pvalue = P\left(\bar{X}_1 - \bar{X}_2 < -4.65\right) = P\left(\frac{\bar{X}_1 - \bar{X}_2 - 0}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} < \frac{-4.65 - 0}{7.88\sqrt{1/8 + 1/8}}\right)
$$

$$
= P(t_{n_1+n_2-2} < -1.18)
$$

$$
= P(t_{14} < -1.18) = P(t_{14} > 1.18) \in (0.1, 0.15).
$$

Here, s_p^2 is the pooled variance which is computed as follows:

$$
s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} = 62.155.
$$

Do not reject H_0 when $\alpha = 0.05$, do not reject when $\alpha = 0.01$.

Case 3: Unknown variances σ_1^2, σ_2^2; large sample size

Example: A researcher wants to test that on average a new fertilizer yields taller plants. Plants were divided into two groups: a control group treated with an old fertilizer and a study group treated with the new fertilizer. The following data are obtained:

Sample size	Sample Mean	Sample Variance
$n_1 = 100$	$\bar{x}_1 = 43.14$	$s_1^2 = 71.65$
$n_2 = 100$	$\bar{x}_2 = 47.79$	$s_2^2 = 52.66$

Test $H_0 : \mu_1 = \mu_2$ vs. $H_1 : \mu_1 < \mu_2$.

Solution: $H_0 : \mu_1 = \mu_2; \ H_1 : \mu_1 < \mu_2.$ The observed difference is $-4.65.$
To compute

$$pvalue = P\left(\bar{X}_1 - \bar{X}_2 < -4.65\right) = P\left(\frac{\bar{X}_1 - \bar{X}_2 - 0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} < \frac{-4.65 - 0}{\sqrt{\frac{71.65}{100} + \frac{52.66}{100}}}\right)$$

$$= P(Z < -4.18) = 0.$$

Reject H_0 when $\alpha = 0.05$, reject when $\alpha = 0.01.$

Recommendation: Make a Brief Review on Tests

Add *pvalue* and *results on proportion*

R-**commands**

- `t.test(x,mu=5)` is the command for $H_0 : \mu = 5$ against $H_1 : \mu \neq 5$ when σ is unknown (t-test)

- `t.test(x,mu=5,alternative="greater")` is the command for $H_0 :$ $\mu = 5$ against $H_1 : \mu > 5$ when σ is unknown (t-test)

- `t.test(x,mu=5,alternative="less")` is the command for $H_0 : \mu = 5$ against $H_1 : \mu < 5$ when σ is unknown (t-test)

- `t.test(x,y,paired=TRUE)` is the command for $H_0 : \mu_1 = \mu_2$ against $H_1 : \mu_1 \neq \mu_2$ in case of paired samples, when variances are unknown.

- `t.test(x,y,paired=TRUE,alternative="greater")` is the command for $H_0 : \mu_1 = \mu_2$ against $H_1 : \mu_1 > \mu_2$ in case of paired samples, when variances are unknown.

- `t.test(x,y,paired=TRUE,alternative="less")` is the command for $H_0 : \mu_1 = \mu_2$ against $H_1 : \mu_1 < \mu_2$ in case of paired samples, when variances are unknown.

- `t.test(x,y)` is the command for $H_0 : \mu_1 = \mu_2$ against $H_1 : \mu_1 \neq \mu_2$ in case of two independent samples, when variances are unknown.

- `t.test(x,y,alternative="greater")` is the command for $H_0 : \mu_1 = \mu_2$ against $H_1 : \mu_1 > \mu_2$ in case of two independent samples, when variances are unknown.

- `t.test(x,y,alternative="less")` is the command for $H_0 : \mu_1 = \mu_2$ against $H_1 : \mu_1 < \mu_2$ in case of two independent samples, when variances are unknown.