HANU Probability and Statistics

Lecture Contents

- Hypothesis Testing II

Hypothesis Testing

Tests on Proportion: single sample

Two-tailed:

$$H_0\colon\ p = p_0,$$
$$H_1\colon\ p \neq p_0,$$

One-tailed:

$$H_0\colon\ p = p_0,$$
$$H_1\colon\ p < p_0,$$

Or

$$H_0\colon\ p = p_0,$$
$$H_1\colon\ p \neq p_0,$$

Hypothesis Testing

Tests on Proportion: single sample

1. H_0: $p = p_0$.
2. One of the alternatives H_1: $p < p_0$, $p > p_0$, or $p \neq p_0$.
3. Choose a level of significance equal to α.
4. Test statistic: Binomial variable X with $p = p_0$.
5. Computations: Find x, the number of successes, and compute the appropriate P-value.
6. Decision: Draw appropriate conclusions based on the P-value.

Hypothesis Testing

Tests on Proportion: single sample

Example: A builder claims that heat pumps are installed in 70% of all homes being constructed today in the city of Richmond, Virginia. Would you agree with this claim

if a random survey of new homes in this city showed that 8 out of 15 had heat pumps installed? Use a 0.10 level of significance.

Hypothesis Testing

Tests on Proportion: single sample

1. H_0: $p = 0.7$.
2. H_1: $p \neq 0.7$.
3. $\alpha = 0.10$.
4. Test statistic: Binomial variable X with $p = 0.7$ and $n = 15$.
5. Computations: $x = 8$ and $np_0 = (15)(0.7) = 10.5$. Therefore, from Table A.1, the computed P-value is

$$P = 2P(X \leq 8 \text{ when } p = 0.7) = 2\sum_{x=0}^{8} b(x; 15, 0.7) = 0.2622 > 0.10.$$

6. Decision: Do not reject H_0. Conclude that there is insufficient reason to doubt the builder's claim.

Hypothesis Testing

Tests on Proportion: single sample

We could also use normal distribution to approximate binomial distribution

$$z = \frac{x - np_0}{\sqrt{np_0 q_0}} = \frac{\hat{p} - p_0}{\sqrt{p_0 q_0 / n}},$$

Hypothesis Testing

Tests on Proportion: single sample

Example:

A commonly prescribed drug for relieving nervous tension is believed to be only 60% effective. Experimental results with a new drug administered to a random sample of 100 adults who were suffering from nervous tension show that 70 received relief. Is this sufficient evidence to conclude that the new drug is superior to the one commonly prescribed? Use a 0.05 level of significance.

Hypothesis Testing

Tests on Proportion: single sample

Example:

1. H_0: $p = 0.6$.
2. H_1: $p > 0.6$.
3. $\alpha = 0.05$.
4. Critical region: $z > 1.645$.

Hypothesis Testing

Tests on Proportion: single sample

Example:

5. Computations: $x = 70$, $n = 100$, $\hat{p} = 70/100 = 0.7$, and

$$z = \frac{0.7 - 0.6}{\sqrt{(0.6)(0.4)/100}} = 2.04, \quad P = P(Z > 2.04) < 0.0207.$$

6. Decision: Reject H_0 and conclude that the new drug is superior.

Hypothesis Testing

Tests on Proportion: two sample

H_0: $p_1 = p_2$

H_1: $p_1 - p_2 \neq 0$

$p_1 - p_2 > 0$

$p_1 - p_2 < 0$

We could use normal distribution to approximate binomial distribution.

Hypothesis Testing

Tests on Proportion: two sample

Compute mean difference:

$$\mu_{\widehat{P}_1 - \widehat{P}_2} = p_1 - p_2$$

Variance
$$\sigma^2_{\widehat{P}_1 - \widehat{P}_2} = \frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}.$$

Compute Z-statistics

$$Z = \frac{(\widehat{P}_1 - \widehat{P}_2) - (p_1 - p_2)}{\sqrt{p_1 q_1 / n_1 + p_2 q_2 / n_2}}.$$

Hypothesis Testing

Tests on Proportion: two sample

Example:

A vote is to be taken among the residents of a town and the surrounding county to determine whether a proposed chemical plant should be constructed. The construction site is within the town limits, and for this reason many voters in the county believe that the proposal will pass because of the large proportion of town voters who favor the construction. To determine if there is a significant difference in the proportions of town voters and county voters favoring the proposal, a poll is taken. If 120 of 200 town voters favor the proposal and 240 of 500 county residents favor it, would you agree that the proportion of town voters favoring the proposal is higher than the proportion of county voters? Use an $\alpha = 0.05$ level of significance.

Hypothesis Testing

Tests on Proportion: two sample

Example:

Let p_1 and p_2 be the true proportions of voters in the town and county, respectively, favoring the proposal.

1. H_0: $p_1 = p_2$.
2. H_1: $p_1 > p_2$.
3. $\alpha = 0.05$.
4. Critical region: $z > 1.645$.

Hypothesis Testing

Tests on Proportion: two sample

Example:

5. Computations:

$$\hat{p}_1 = \frac{x_1}{n_1} = \frac{120}{200} = 0.60, \quad \hat{p}_2 = \frac{x_2}{n_2} = \frac{240}{500} = 0.48, \quad \text{and}$$

$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{120 + 240}{200 + 500} = 0.51.$$

Therefore,

$$z = \frac{0.60 - 0.48}{\sqrt{(0.51)(0.49)(1/200 + 1/500)}} = 2.9,$$

$$P = P(Z > 2.9) = 0.0019.$$

6. Decision: Reject H_0 and agree that the proportion of town voters favoring the proposal is higher than the proportion of county voters.

Hypothesis Testing

Tests on Variance: single sample

H0: σ2 = σ_0^2

H1: $\sigma^2 < \sigma_0^2$ or $\sigma^2 > \sigma_0^2$ or $\sigma^2 \neq \sigma_0^2$

If the sample is from normal distribution then the statistics has chi-squared distribution (n-1 degree of freedom)

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2},$$

Hypothesis Testing

Tests on Variance: single sample

Critical region for one-tailed tests:

$$\chi^2 < \chi^2_{1-\alpha} \qquad \chi^2 > \chi^2_{\alpha}$$

Critical region for two-tailed tests:

$$\chi^2 < \chi^2_{1-\alpha/2} \text{ or } \chi^2 > \chi^2_{\alpha/2}$$

Hypothesis Testing

Tests on Variance: single sample

Example: A manufacturer of car batteries claims that the life of the company's batteries is approximately normally distributed with a standard deviation equal to 0.9 year. If a random sample of 10 of these batteries has a standard deviation of 1.2 years, do you think that $\sigma > 0.9$ year? Use a 0.05 level of significance.

Hypothesis Testing

Tests on Variance: single sample

Example:.

1. H_0: $\sigma^2 = 0.81$.
2. H_1: $\sigma^2 > 0.81$.
3. $\alpha = 0.05$.
4. Critical region: From Figure 10.19 we see that the null hypothesis is rejected when $\chi^2 > 16.919$, where $\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$, with $v = 9$ degrees of freedom.

Hypothesis Testing

Tests on Variance: single sample

Example:.

5. Computations: $s^2 = 1.44$, $n = 10$, and

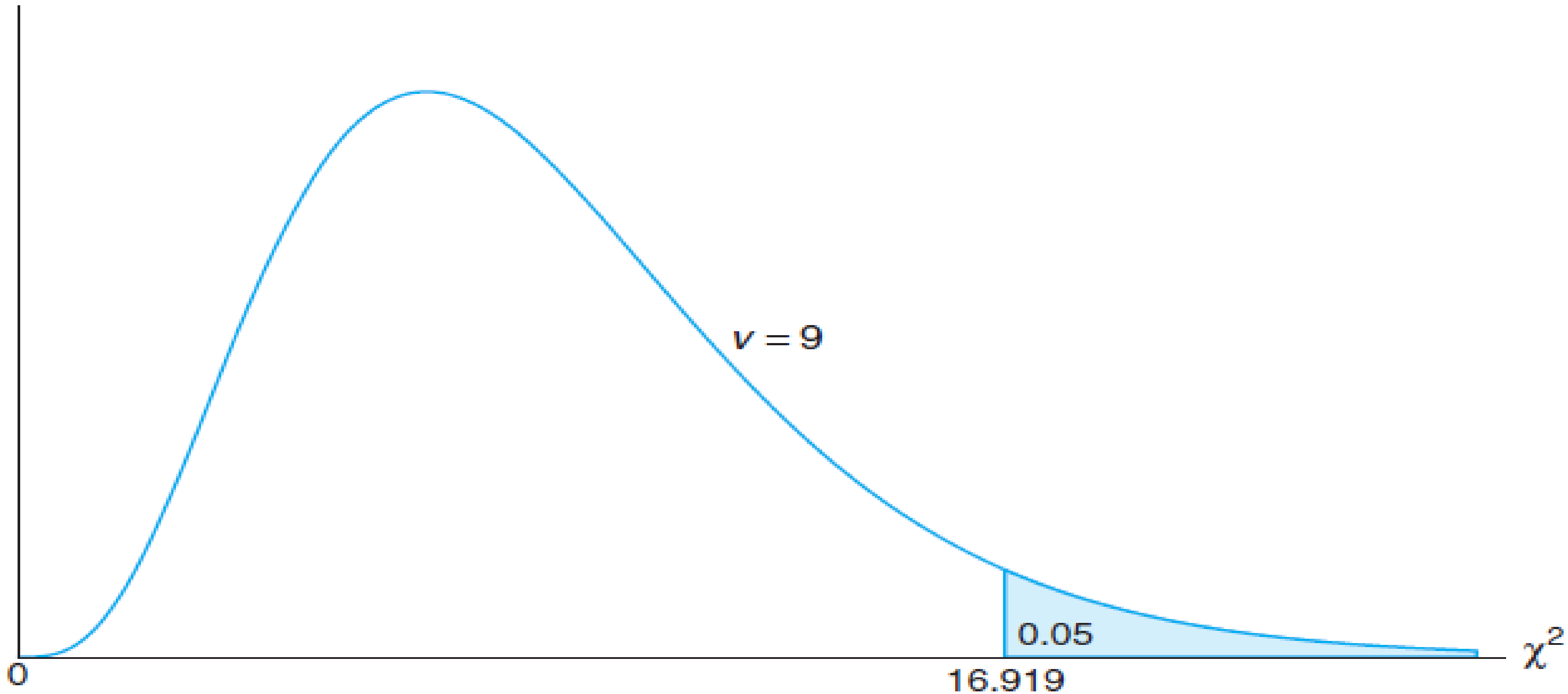
$$\chi^2 = \frac{(9)(1.44)}{0.81} = 16.0, \qquad P \approx 0.07.$$

6. Decision: The χ^2-statistic is not significant at the 0.05 level. However, based on the P-value 0.07, there is evidence that $\sigma > 0.9$.

Hypothesis Testing

Tests on Variance: single sample

Example:.



Hypothesis Testing

Tests on Variance: two sample

Example:

H0: $\sigma_1^2 = \sigma_2^2$

H1: $\sigma_1^2 < \sigma_2^2, \quad \sigma_1^2 > \sigma_2^2, \quad \text{or} \quad \sigma_1^2 \neq \sigma_2^2.$

The ratio between the two variances has F-distribution (with n1-1 and n2-1 degrees of freedom)

$$f = \frac{s_1^2}{s_2^2},$$

Hypothesis Testing

Tests on Variance: two sample

Critical regions:

One tailed tests:

$$f < f_{1-\alpha}(v_1, v_2) \text{ and } f > f_{\alpha}(v_1, v_2)$$

Two tailed tests:

$$f < f_{1-\alpha/2}(v_1, v_2) \text{ or } f > f_{\alpha/2}(v_1, v_2)$$

Hypothesis Testing

Tests on Variance: two sample

Example: Is this right in the previous example that the two variances are equal (0.1 level of significance)

An experiment was performed to compare the abrasive wear of two different laminated materials. Twelve pieces of material 1 were tested by exposing each piece to a machine measuring wear. Ten pieces of material 2 were similarly tested. In each case, the depth of wear was observed. The samples of material 1 gave an average (coded) wear of 85 units with a sample standard deviation of 4, while the samples of material 2 gave an average of 81 with a sample standard deviation of 5. Can we conclude at the 0.05 level of significance that the abrasive wear of material 1 exceeds that of material 2 by more than 2 units? Assume the populations to be approximately normal with equal variances.

Hypothesis Testing

Tests on Variance: two sample

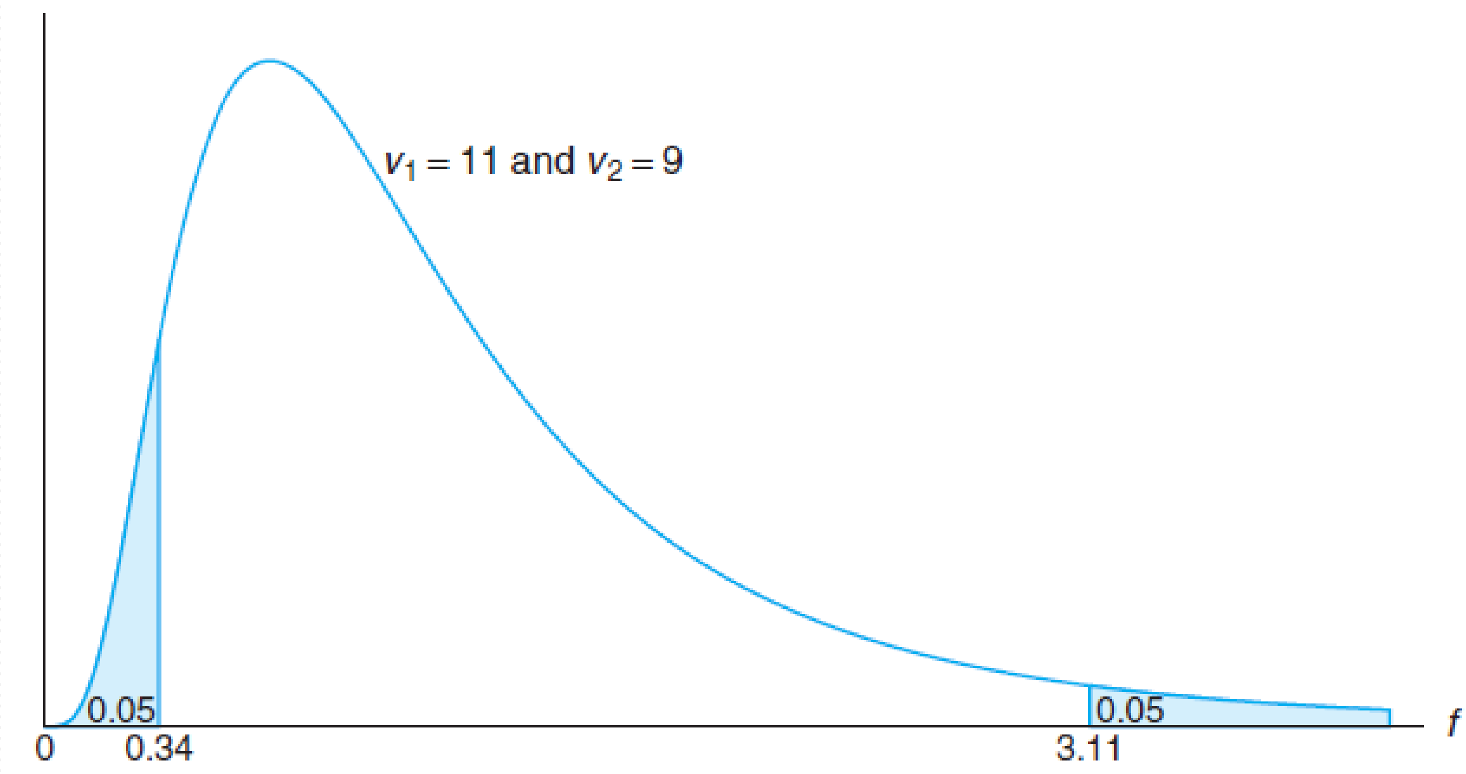
1. H_0: $\sigma_1^2 = \sigma_2^2$.
2. H_1: $\sigma_1^2 \neq \sigma_2^2$.
3. $\alpha = 0.10$.
4. Critical region: From Figure 10.20, we see that $f_{0.05}(11, 9) = 3.11$, and, by using Theorem 8.7, we find

$$f_{0.95}(11, 9) = \frac{1}{f_{0.05}(9, 11)} = 0.34.$$

   Therefore, the null hypothesis is rejected when $f < 0.34$ or $f > 3.11$, where $f = s_1^2/s_2^2$ with $v_1 = 11$ and $v_2 = 9$ degrees of freedom.
5. Computations: $s_1^2 = 16$, $s_2^2 = 25$, and hence $f = \frac{16}{25} = 0.64$.
6. Decision: Do not reject H_0. Conclude that there is insufficient evidence that the variances differ.

Hypothesis Testing

Tests on Variance: two sample



Hypothesis Testing

Goodness-of-fit Tests:

In many statistical hypothesis tests, we assume that the original distribution is (nearly) normal (for instance tests related to variances). If this assumption is violated then the tests are unreliable. Goodness of fit tests are used to check if the frequency of the occurrence of observations fit the expected frequencies from the hypothesized theoretical distribution.

Hypothesis Testing

Goodness-of-fit Tests:

Example: Throwing a dice -> uniform distribution with 120 observations.

$$f(x) = \frac{1}{6}, \quad x = 1, 2, \ldots, 6.$$

Observed and Expected Frequencies of 120 Tosses of a Die

Face:	1	2	3	4	5	6
Observed	20	22	17	18	19	24
Expected	20	20	20	20	20	20

Hypothesis Testing

Goodness-of-fit Tests:

Use Chi-square distribution for the statistics (with k-1 degree of freedom)

$$\chi^2 = \sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i},$$

Hypothesis Testing

Goodness-of-fit Tests:

H0 – the fit is good

H1 – Reject H0 (the fit is bad)

Critical region – $\chi^2 > \chi^2_\alpha$

Steps:

- Compute the statistics.
- Compare with the critical region.

Hypothesis Testing

Goodness-of-fit Tests:

Example: test the goodness-of-test for the example with α=0.05.

$$\chi^2 = \frac{(20-20)^2}{20} + \frac{(22-20)^2}{20} + \frac{(17-20)^2}{20} + \frac{(18-20)^2}{20} + \frac{(19-20)^2}{20} + \frac{(24-20)^2}{20} = 1.7.$$

Using Table A.5, we find $\chi^2_{0.05} = 11.070$ for $v = 5$ degrees of freedom. Since 1.7 is less than the critical value, we fail to reject H_0. We conclude that there is insufficient evidence that the die is not balanced.

Hypothesis Testing

Tests for independence (categorical):

Chi-square tests could also be used to test the independence between categorical (classification) variables.

Example:

we wish to determine whether the opinions of the voting residents of the state of Illinois concerning a new tax reform are independent of their levels of income. Members of a random sample of 1000 registered voters from the state of Illinois are classified as to whether they are in a low, medium, or high income bracket and whether or not they favor the tax reform. The observed frequencies are presented in Table 10.6, which is known as a **contingency table**.

Hypothesis Testing

Tests for independence (categorical):

	Income Level			
Tax Reform	Low	Medium	High	Total
For	182	213	203	598
Against	154	138	110	402
Total	336	351	313	1000

Hypothesis Testing

Tests for independence (categorical):

We need to test if there is any independence between the voters' **opinions** and their **incomes**.

L: A person selected is in the low-income level.

M: A person selected is in the medium-income level.

H: A person selected is in the high-income level.

F: A person selected is for the tax reform.

A: A person selected is against the tax reform.

Hypothesis Testing

Tests for independence (categorical):

We need to test if there is any independence between the voters' **opinions** and their **incomes**.

$$P(L) = \frac{336}{1000}, \quad P(M) = \frac{351}{1000}, \quad P(H) = \frac{313}{1000},$$

$$P(F) = \frac{598}{1000}, \quad P(A) = \frac{402}{1000}.$$

Hypothesis Testing

Tests for independence (categorical):

If H0 is true – variables are independent, i.e

$$P(L \cap F) = P(L)P(F) = \left(\frac{336}{1000}\right)\left(\frac{598}{1000}\right),$$

$$P(L \cap A) = P(L)P(A) = \left(\frac{336}{1000}\right)\left(\frac{402}{1000}\right),$$

$$P(M \cap F) = P(M)P(F) = \left(\frac{351}{1000}\right)\left(\frac{598}{1000}\right),$$

$$P(M \cap A) = P(M)P(A) = \left(\frac{351}{1000}\right)\left(\frac{402}{1000}\right),$$

$$P(H \cap F) = P(H)P(F) = \left(\frac{313}{1000}\right)\left(\frac{598}{1000}\right),$$

$$P(H \cap A) = P(H)P(A) = \left(\frac{313}{1000}\right)\left(\frac{402}{1000}\right).$$

Hypothesis Testing

Tests for independence (categorical):

How to compute the expected frequencies?

Example - the expected number of low-income voters in our sample who favor the tax reform is estimated to be

$$\left(\frac{336}{1000}\right)\left(\frac{598}{1000}\right)(1000) = \frac{(336)(598)}{1000} = 200.9$$

Hypothesis Testing

Tests for independence (categorical):

How to compute the expected frequencies?

$$\text{expected frequency} = \frac{(\text{column total}) \times (\text{row total})}{\text{grand total}}.$$

	Income Level			
Tax Reform	Low	Medium	High	Total
For	182 (200.9)	213 (209.9)	203 (187.2)	598
Against	154 (135.1)	138 (141.1)	110 (125.8)	402
Total	336	351	313	1000

Hypothesis Testing

Tests for independence (categorical):

Calculate Chi-square statistics (with (r-1)*(c-1) degree of freedom)

Calculate

$$\chi^2 = \sum_i \frac{(o_i - e_i)^2}{e_i},$$

where the summation extends over all rc cells in the $r \times c$ contingency table.

Hypothesis Testing

Tests for independence (categorical):

Example -

$$\chi^2 = \frac{(182-200.9)^2}{200.9} + \frac{(213-209.9)^2}{209.9} + \frac{(203-187.2)^2}{187.2} + \frac{(154-135.1)^2}{135.1} + \frac{(138-141.1)^2}{141.1} + \frac{(110-125.8)^2}{125.8} = 7.85,$$

$$P \approx 0.02.$$

From Table A.5 we find that $\chi^2_{0.05} = 5.991$ for $v = (2-1)(3-1) = 2$ degrees of freedom. The null hypothesis is rejected and we conclude that a voter's opinion concerning the tax reform and his or her level of income are not independent.

Further Readings

- Textbook 2 - Chapter 10.
- Textbook 1- Chapter 9-10.