Module 9

Xihan Qian

July, 2022

1 Suppose that we want to test whether the average height of American males is 175cm or not. Discuss whether the following notations are correct in this context. If it is not correct, explain why.

(a) $\bar{X} = 175$

(b) $\mu = 175$

Solution: (a) Incorrect. We want to test a hypothesis about the population mean μ, not the sample mean $\bar{X}$. We can never test a hypothesis about a sample mean or a sample proportion.

(b) Correct. We want to test this hypothesis about the population mean, which in this case represents the average height of American males.

2 Suppose that we conduct a hypothesis test at significance level $\alpha = 0.05$. If now α is changed to be 0.10, how does the probability of Type I and Type II error change? Do they increase, decrease, or stay the same?

Solution: P(Type I error) = α. Hence if α increases from 0.05 to be 0.10 now, the probability of Type I error would also increase. This would decrease the probability of Type II error.

3 Men typically got married at 25.9 years of age in California in the year 2000. Your classmate suspects that since then, men have been waiting longer to be married. We want to test whether the average age of when men get married has increased since 2000.

(a) Please write down the appropriate null and alternative hypotheses.

(b) We want to test our hypothesis based on a randomly selected group of 50 men who married for the first time this year. Check whether the assumptions are satisfied.

(c) What is the appropriate sampling distribution model for this?

Solution: (a) Let μ be the average of marriage for men in California. We have $H_0 : \mu = 25.9; H_A : \mu > 25.9$.

(b) This is a random sample of men from the population. We also have a sample size of $50 \geq 30$. Hence the assumptions to use the Central Limit Theorem are satisfied.

(c) The sample average age $\bar{X}$ at marriage should have an approximately normal distribution with mean μ and standard deviation $\sigma/\sqrt{50}$, where $\mu = 25.9$ is the population mean and σ is the population standard deviation.

4 Continuing Exercise 3(, if the men in our sample married with a mean of 27 years old and a standard deviation of 4.7 years,

(a) Compute the t test statistic and p-value.

(b) Explain what this p-value means.

(c) What is your final conclusion at significance level $\alpha = 0.05$?

Solution:

$$\text{test statistic: } T = \frac{\bar{X} - \mu}{SE} = \frac{27 - 25.9}{4.7/\sqrt{50}} = 1.655$$

This test has a degrees of freedom of $50 - 1 = 49$; So the p-value is $P(T > 1.655) = 0.052$ using R: $\mathrm{pt}(1.655, 49, \mathrm{lower.tail} = \mathrm{FALSE})$

(b) This means that, if the mean age at which men get married had not changed since 2000, then the average age of the men in a random sample of size 50 at the time of marriage would be at least as large as it was in our sample approximately 5.2 percent of the time.

(c) Since $0.052 > 0.05$, we conclude that we do not have enough evidence to conclude that the mean age at which men get married has increased since 2000.

5 Discuss whether the following statements are true or not.

(a) It will always be inappropriate to report a confidence interval if the sample size is only 25.

(b) If the population distribution is skewed, then the distribution of the sample mean will look more and more normal as the sample size n grows large.

Solution: (a) False. If the population distribution is normal or approximately normal, then it is reasonable to also use an approximately normal model for small sample sizes.

(b) True. l

6 Discuss whether the data associated with the following scenarios is paired data or independent data.

(a) A professor would like to compare students' midterm and final scores. A random sample of students were chosen and their midterm/final scores were recorded.

(b) In a research study, a group of 500 adults were selected randomly. They were asked if they like dogs or cats better and how many pets they own. The groups in this study are the dog lovers and cat lovers.

Solution: (a) This would be paired data since a student's midterm and final scores are correlated with one another.

(b) This would be independent data since how many pets each of them owns are most likely independent of each other.

7 In a certain test, researchers are interested in whether attending discussion sessions improves students' performance. The following data is collected from a sample of students:

Statistics	Students not attending discussion sessions	Students attending discussion sessions
$\bar{x}$	16	18.5
s	2.4	3
n	127	209

Here $\bar{x}$ denotes the sample average of final grades, s denotes the standard deviations of the sample, and n denotes the sample sizes.

Construct a 95% confidence interval to estimate the mean difference and interpret the confidence interval.

Solution: Let $\bar{x}_1$ denote the grades of the students who did not attend discussion sessions, and $\bar{x}_2$ denote the grades of the students who did attend discussion sessions. Then $\bar{x}_1 - \bar{x}_2 = 16 - 18.5 = -2.5$; The standard error can be computed as:

$$\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{2.4^2}{127} + \frac{3^2}{209}} = 0.297.$$

Here $\mathrm{df} = \min\{n_1, n_2\} - 1 = 127 - 1 = 126$.

We find the corresponding critical t value to be $t^*_{126} = 1.979$ with R code: $\mathrm{qt}(0.025, 126, \mathrm{lower.tail} = \mathrm{FALSE})$.

The margin of error can be computed to be $t^*_{126} \times SE = 1.979 \times 0.297 = 0.588$. So a 95% confidence interval is $(\bar{x}_1 - \bar{x}_2 - \mathrm{ME}, \bar{x}_1 - \bar{x}_2 + \mathrm{ME}) = (-2.5 - 0.588, -2.5 + 0.588) = (-3.088, -1.912)$

We are 95% confident that the mean difference in the exam scores of those students who didn't attend discussion sessions and those who did is between -3.088 and -1.912 points.

Solution: It is known that $P(\text{Type II error}) + \text{power} = 1$, or in other words, $\text{power} = 1 - P(\text{Type II error})$. Hence if power has increased, then the probability of Type II error would decrease.

8 In a certain test, researchers are interested in whether attending discussion sessions improves students' performance. The following data is collected from a sample of students:

Statistics	Students not attending discussion sessions	Students attending discussion sessions
$\bar{x}$	18	18.5
s	2.4	3
n	127	209

Here $\bar{x}$ denotes the sample averages, s denotes the standard deviations of the sample, and n denotes the sample sizes.

Conduct a hypothesis test and conclude whether the extra classes are improving the students' performances at significance level $\alpha = 0.05$.

Solution: Let $\bar{x}_1$ and s_1 be the mean and standard deviation of the exam scores of those students not attending discussion sessions and $\bar{x}_2$ and s_2 be the mean and standard deviation of the exam scores of those students who attend discussion sessions. Let n_1 and n_2 be the number of students in each sample respectively. Let μ_1 and μ_2 be the true mean exam scores that would be achieved by students who don't attend discussions and those who do, respectively. We test $H_0 : \mu_1 = \mu_2$ against $H_A : \mu_1 < \mu_2$. The test statistic is

$$T = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{18 - 18.5}{\sqrt{\frac{2.4^2}{127} + \frac{3^2}{209}}} \approx -1.682$$

If H_0 were to be true, T would follow a t-distribution with degrees of freedom to be $127 - 1 = 126$. We calculate the p-value to be $P(T < -1.682) \approx 0.048$. At significance level $\alpha = 0.05$, we have that $0.048 < 0.05$, and conclude that we have strong evidence that students who do attend discussion sessions would perform better than those who don't.

9 If by increasing the sample size, we increased the power, how would the probability of Type II error change?

Solution: It is known that $P(\text{Type II error}) + \text{power} = 1$, or in other words, $\text{power} = 1 - P(\text{Type II error})$. Hence if power has increased, then the probability of Type II error would decrease.

10 We are interested in the statistical power of a hypothesis test associated with whether taking weekly quizzes would improve students' exam scores. Let μ_w denote the mean exam scores of students who take weekly quizzes; and let μ_{wo} denote the mean exam scores of students who did not take weekly quizzes. The null and alternative hypotheses, associated with our power calculation, are listed below:

$H_0 : \mu_w - \mu_{wo} = 0$

$H_A : \mu_w - \mu_{wo} = 0.5$

Suppose that for the test we take sample sizes to be $n_w = 270; n_{wo} = 189$. It is also estimated that $s_w = 4.3, s_{wo} = 7.5$. Calculate the statistical power of this test at significance level $\alpha = 0.10$.

Solution: We can estimate that

$$\text{SE} = \sqrt{\frac{s_w^2}{n_w} + \frac{s_{wo}^2}{n_{wo}}} = \sqrt{\frac{4.3^2}{270} + \frac{7.5^2}{189}} = 0.605.$$

Moreover, since $df_w = 269 \geq 30$ and $df_{wo} = 188 \geq 30$, under H_0, we can model $\bar{x}_w - \bar{x}_{wo}$ as normal with mean 0 and SE = 0.605. We know that rejecting H_0 occurs when $\bar{x}_w - \bar{x}_{wo}$ lies in the rejection region, i.e., when $\bar{x}_w - \bar{x}_{wo}$ is less than $-1.645 \times 0.605 = -0.995$, or greater than $1.645 \times 0.605 = 0.995$. We also know that the alternative is normal with mean 0.5. Then we can compute that $z = \frac{0.995-0.5}{0.605} = 0.818$. Hence $P(X > 0.995) = P(Z > 0.818) \approx 0.2061$.

Therefore the statistical power of this test is 0.2061.