BAN 602: Quantitative Fundamentals

Lecture 4: Hypothesis Testing and Inference about Two Population Means or Proportions

Agenda

1. Big Picture of Chapter 9: Hypothesis Tests
2. Big Picture of Chapter 10: Inferences about Two Population Means or Proportions
3. Example 1 (population mean)
4. Example 2 (two population proportions)

Big Picture of Chapter 9: Hypothesis Tests

Develop Null and Alternative Hypotheses	Type I and Type II Errors	Probability of Type II Error and Power
● Types of Tests ○ Lower-tailed test ○ Upper-tailed test ○ Two-tailed test ● Hypothesis Tests of a Population Mean ○ Sigma is known ○ Sigma is unknown ● Hypothesis Tests of a Population Proportion	● Type I Error ○ Mistakenly conclude that the null hypothesis is false ● Type II Error ○ Mistakenly conclude that the alternative hypothesis is false ● P-value vs. critical values	● Beta = pr(Type II error) =pr(conclude Ha is false | Ha is true) = pr(sample mean or proportion is within acceptance range | Ha is true) ● Power = 1 - beta

Big Picture of Chapter 10: Inference about Means and Proportions with Two Populations

Inference about the difference b/t two populations means with known σ's (independent simple random samples)	Inference about the difference b/t two populations means with unknown σ's (independent simple random samples)	Inference about the difference b/t two populations means: matched samples	Inference about the difference between two populations proportions
● Interval estimation ○ Normal ○ SE = $\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$ ● Hypothesis tests ○ Normal (z test statistic) ○ SE = $\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$	● Interval estimation ○ t distribution (df) ○ SE = $\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$ ● Hypothesis tests ○ t distribution (df) ○ SE = $\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$	● degenerate into inference about one population mean	● Interval estimation ○ SE = $\sqrt{\frac{\bar{p}_1(1-\bar{p}_1)}{n_1}+\frac{\bar{p}_2(1-\bar{p}_2)}{n_2}}$ ● Hypothesis tests ($p_1=p_2=p$) ○ p known: SE=$\sqrt{p(1-p)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$ when p is known ○ p unknown: replace p with $\bar{p}=\frac{n_1\bar{p}_1+n_2\bar{p}_2}{n_1+n_2}$.

Hypotheses, Type I/II Errors, and False Positive/Negative

- We consider “+” or “positive” as the “disease” or “guilty”, and “-” or “negative” as “no-disease” or “innocent”. We consider “-” as the null hypothesis and “+” as the alternative hypothesis

		Truth	
		- or null	+ or alternative
Prediction	- or null	TN (true negative)	FN (false negative) Type II Error Prob(FN) = beta
	+ or alternative	FP (false positive) Type I Error Prob(FP) = p-value	TP (true positive) Prob(TP) = power

Example 1: Hypothesis Test about a Population Mean

A consumer research group is interested in testing an automobile manufacturer's claim that a new economy model will travel at least 25 miles per gallon of gasoline.

- What would be the population of our interest? What would be the random variable that describes the population. What do we know about this random variable? (In-class exercise question 1)

Example 1

A consumer research group is interested in testing an automobile manufacturer's claim that a new economy model will travel at least 25 mpg.

- What would be the population of our interest? What would be the random variable that describes the population. What do we know about this rv?

  The population is the mpg's of all cars of this new model.

  X represents the mpg of a randomly selected car of this model.

  $X \sim A\ Distribution(\mu, \sigma)$

Example 1

- The focus of this study is the population mean μ and investigates whether the auto maker's claim is true. To that end, we turn to our random sampling process and the resulting sampling distribution of sample mean, $\bar{X}$. Please describe this sampling distribution (central limit theorem). (In-class exercise question 2)

Example 1

- The focus of this study is the population mean μ and investigates whether the auto maker's claim is true. To that end, we turn to our random sampling process and the resulting sampling distribution of sample mean, $\bar{X}$. Please describe this sampling distribution (central limit theorem).
    - $\bar{X} \sim Normal\left(\mu_{\bar{X}} = \mu,\ \sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}\right).$

Example 1

- What would be appropriate hypotheses for testing the automaker’s claim? (In-class exercise question 3)

Example 1

- What would be appropriate hypotheses for testing the automaker's claim? (In-class exercise question 3)

H_0: $\mu \geq 25$

H_a: $\mu < 25$. (this is called lower-tailed test)

Example 1

- Suppose a test on 36 cars of this model indicates an average of 24 mpg, with a **sample** standard deviation of 3 mpg. What would be an appropriate sample (test) statistic for our hypothesis testing? Describe the sampling distribution of this sample (test) statistic. (in-class exercise question 4)

Example 1

- Suppose a test on 36 cars of this model indicates an average of 24 mpg, with a **sample** standard deviation of 3 mpg. What would be an appropriate sample (test) statistic for our hypothesis testing? Describe the sampling distribution of this sample (test) statistic. (in-class exercise question 4)

$$t\ statistic = \frac{\bar{X}-\mu}{s/\sqrt{n}} \sim t\ distribution(df = n-1)$$

Example 1

- Suppose the null hypothesis is true. That is, the population mean μ = 25 mpg (ignore the part μ > 25 for the time being). What is the probability that we observe a sample mean of 24 mpg or lower given sample size n is 36?

Example 1

- Suppose the null hypothesis is true. That is, the population mean μ = 25 mpg (ignore the part μ > 25 for the time being). What is the probability that we observe a sample mean of 24 mpg or worse given sample size n is 36?

$$\Pr(\bar{X} \leq 24|\mu = 25)$$
$$= \Pr(\bar{X} - \mu \leq 24 - \mu|\mu = 25)$$
$$= Pr\left(\frac{\bar{X}-\mu}{s/\sqrt{n}} \leq \frac{24-\mu}{s/\sqrt{n}}|\mu = 25\right)$$
$$= Pr\left(t\ statistic = \frac{\bar{X}-25}{3/6} \leq \frac{-1}{3/6}\right)$$
$= pt(-2, df = 35) = 2.67\%$. Note that $pnorm(-2) = 2.28\%$

Example 1

- Now we know that the probability that we observe a sample mean of 24 mpg or worse is 2.67%. How can we interpret this result? (In-class exercise question 5)

Example 1

- Now we know that the probability that we observe a sample mean of 24 mpg or worse is 2.67%. How can we interpret this result?

1. If we accept the null hypothesis as truth, then there is at most 2.67% chance that we observe a sample mean of 24 mpg or worse. (The “at most” part takes care of the “$\mu > 25$” part of the null hypothesis.)
2. If we reject the null hypothesis, then there is at most 2.67% chance that we make a Type I error.

Example 1

- Now we know that the probability that we observe a sample mean of 24 mpg or worse is 2.67%. How can we interpret this result?

3. This probability is called p-value.
   - This is the probability that, if we assume the null hypothesis is true, we observe a result as extreme as or more extreme than the sampling result or in this example, observed sample mean (24 mpg).
   - Small p-value indicates more evidence against the null hypothesis. So, we are more confident to reject the null hypothesis.
   - And p-value is also called observed level of significance.

Example 1

- How are we going to draw conclusions, then? This is where α, the significance level, comes into play. Let's say, we decide that α = 5%. How can we interpret this significance level?

Example 1

- How are we going to draw conclusions, then? This is where α, the significance level, comes into play. Let's say, we decide that α = 5%. How can we interpret this significance level?

If there is a chance of α or better that the null hypothesis is true, then we do not reject it. In this example, the significant level of 5% means that if there is 5% or better chance that H_0 is true, we do not reject it.

The actual p-value is 2.67%, i.e., there is less than 3% chance H_0 is true. We thus reject the null hypothesis and conclude that the gas mileage of this new model is less than 25 mpg. (Some people insist that we use "do not accept" instead of "reject").

Example 1

- Now that we know the significance level α = 5%. We can use another approach, critical value approach, to draw our conclusions. More specifically, given α = 5%, what is the range of sampling result (sample mean) in which we accept H_0? This is called acceptance region. The opposite to the acceptance range is rejection region. The threshold value between the two ranges is called the critical value.
- If the sampling result falls within the acceptance region, we accept H_0; if the sampling result falls within the rejection region, we reject H_0.
- How can we find this critical value and corresponding acceptance/rejection region? (in-class exercise question 6)

Example 1

- How can we find this critical value and corresponding acceptance/rejection region? (in-class exercise question 6)

$Pr(\bar{X}\ as\ extreme\ as\ or\ more\ extreme\ than\ critical\ value|H_0\ is\ true) = \alpha.$

In our example, $Pr(\bar{X}\ \leq \bar{X}_c = critical\ value|\mu = 25) = \alpha.$

$Pr\left(\frac{\bar{X}-\mu}{s/\sqrt{n}} \leq \frac{\bar{X}_c-\mu}{s/\sqrt{n}}\,|\mu = 25\right) = \alpha\ , Pr\left(t\ statistic \leq \frac{\bar{X}_c-25}{0.5}\right) = \alpha = 0.05$

$$\frac{\bar{X}_c - 25}{0.5} = qt(0.05, 35) \Longrightarrow \bar{X}_c = 25 + qt(0.05, 35) * 0.5 = 25 - 1.69 * 0.05 = 24.155$$

- $qt(0.05, 35)$ sometimes called the critical value of test statistic.
- $\bar{X}_c$ sometimes called the critical value of sample mean.

Example 1

- Now we know $\bar{X}_c = 24.155$. What are the corresponding acceptance/rejection region? How do we draw conclusions?

Example 1

- Now we know $\bar{X}_c = 24.155$. What are the corresponding acceptance/rejection region? How do we draw conclusions?

  $\text{rejection region: } \bar{X} < 24.155$

  $\text{acceptance region: } \bar{X} > 24.155$

The observed sampling result (in this case the observed sample mean) is 24, which is in the rejection region. We thus reject the null hypothesis.

Example 1

- What if this were upper-tailed or two-tailed test? How can we determine the critical value and corresponding acceptance/rejection region? How can we determine the p-value?

Example 1

- What if this were upper-tailed or two-tailed test? How can we determine the critical value and corresponding acceptance/rejection region? How can we determine the p-value?

Upper-tailed test

p-value = prob($\bar{X} \geq$ observed sample mean$|$ H_0 *is true*):

- $\bar{X}$ is the random variable sample mean;
- H_0 *is true* can be interpreted as $\mu = \mu_0$.

critical value: prob($\bar{X} \geq \bar{X}_c \,|\mu = \mu_0) = \alpha$, or $Pr\left(\frac{\bar{X}-\mu_0}{se} \geq \frac{\bar{X}_c-\mu_0}{se}\right) = \alpha$. thus $\bar{X}_c = \mu_0 + F^{-1}(1-\alpha) * se$

Example 1

- What if this were upper-tailed or two-tailed test? How can we determine the critical value and corresponding acceptance/rejection range? How can we determine the p-value?

Two-tailed test

p-value $= \text{prob}(\bar{X} \leq -|observed\ sample\ mean| \text{ or } |sampling\ result| \leq \bar{X}|\mu = \mu_0)$

$= 2 * F\left(\frac{-|observed\ sample\ mean| - \mu_0}{se}\right)$.

critical value: $\text{prob}(\bar{X} \leq -|\bar{X}_c| \text{ or } |\bar{X}_c| \leq \bar{X}\ |\mu = \mu_0) = \alpha$, or $Pr\left(\frac{\bar{X}-\mu_0}{se} \leq\right.$

Example 1

- How can we interpret Type II error in this example? (in-class exercise question 7)

Example 1

- How can we interpret Type II error in this example? (in-class exercise question 7)

  In general, Type II error = false negative: we mistakenly conclude that H_0 is true when it is indeed false. Therefore, to discuss Type II error, H_0 must be false to begin with.

  In this example, Type II error is, we mistakenly conclude that the gas mileage is at least 25 mpg when it is indeed less than 25.

Example 1

- Recall that the probability of making Type II error is called β and 1- β is called power. How can we actually compute β?

Example 1

- Recall that the probability of making Type II error is called β and 1- β is called power. How can we actually compute β?
  $\beta = prob(conclude\ H_0 \text{ is true} | H_0 \text{ is false}).$

  Be very careful what "$conclude\ H_0$ is true" and what "H_0 is false" mean in upper/lower/two-tailed tests.

  In our example of lower-tailed test, "$conclude\ H_0$ is true" means that the sampling result is greater than $\bar{X}_c$; "H_0 is false" means that $\mu < \mu_0 = 25$.

Example 1

- Recall that the probability of making Type II error is called β and 1- β is called power. How can we actually compute β?
  In order for us to actually compute, we must specify an actual value for μ. For example, $\mu = 24$. Note that this specified value must be less than 25. Otherwise, we won't commit Type II error.

  For every single specified value of μ, we can compute its corresponding β and power. In fact, we can plot a power curve with respect to all possible values of μ on the horizontal axis and power on vertical axis.

Example 1

- In our example, let’s assume the true mean is 24. How can we actually compute β?

Example 1

- In our example, let's assume the true mean is 24. How can we actually compute β?

  $\beta = prob(conclude\ H_0 \text{ is true}|H_0 \text{ is false}).$

  This is the definition of beta or probability of making Type II error.

  $= prob(\bar{X} \geq \bar{X}_c|\mu = 24)$

  Recall that $\bar{X} \geq \bar{X}_c$ is the acceptance range. Only true for this lower-tailed test example.

  $= prob\left(\frac{\bar{X}-\mu}{se} \geq \frac{\bar{X}_c-24}{se}\right) = 1 - F\left(\frac{\bar{X}_c-24}{se}\right) = 1 - pt\left(\frac{\bar{X}_c-24}{se}, df = n - 1\right).$

  Even though p-value is much more convenient and thus popular than critical value in hypothesis testing, the critical values are essential for computing beta or power.

Tradeoff Between Type I and II Errors

- Given sample size, if we prefer less Type I error, then we pay a price of higher probability of making Type II error; vice versa.
- Increase in sample size can decrease both p-value and beta.

Quiz 4

1. Hypotheses.
2. Computer se.
3. R code for computing p-value and critical value of test statistic.
4. Draw conclusions on hypothesis test.
5. R code for computing beta.
6. Understand p-value, Type I/II error, alpha, beta, and power.

Example 2: Inference about Two Population Proportions

The Professional Golf Association (PGA) measured the putting accuracy of professional golfers playing on the PGA Tour and the best amateur golfers playing in the World Amateur Championship. A sample of 1075 6-foot putts by professional golfers found 688 made puts. A sample of 1200 6-foot putts by amateur golfers found 696 made putts.

- population 1: professional; population 2: amateur
- $p_1, \bar{p}_1, p_2, \bar{p}_2$
- $\bar{p}_1 \sim normal(p_1, \sqrt{\frac{p_1(1-p_1)}{n_1}}), \bar{p}_2 \sim normal(p_2, \sqrt{\frac{p_2(1-p_2)}{n_2}})$

Example 2: Inference about Two Population Proportions

Estimate the proportion of made 6-foot putts by professional golfers. Estimate the proportion of made 6-foot putts by amateur golfers. Which group had a better putting accuracy?

- Point estimate and interval estimate for p_1 and p_2, respectively.

Example 2: Inference about Two Population Proportions

What is the point estimate of the difference between the proportions of the two populations? Interval estimate?

- Population parameter: $p_1 - p_2$, sample parameter: $\bar{p}_1 - \bar{p}_2$.
- Sampling distribution of difference of sample proportions
- $\bar{p}_1 - \bar{p}_2 \sim normal(p_1 - \ p_2, \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}})$
- We don't know p_1, p_2. So, replace $\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$ with $\sqrt{\frac{\bar{p}_1(1-\bar{p}_1)}{n_1} + \frac{\bar{p}_2(1-\bar{p}_2)}{n_2}}$.

Example 2: Inference about Two Population Proportions

What is the point estimate of the difference between the proportions of the two populations? Interval estimate?

- Population parameter: $p_1 - p_2$, sample parameter: $\bar{p}_1 - \bar{p}_2$.
- Sampling distribution of difference of sample proportions: $\bar{p}_1 - \bar{p}_2 \sim normal(p_1 - p_2, \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}})$

Example 2: Inference about Two Population Proportions

What is the point estimate of the difference between the proportions of the two populations? Interval estimate?

- For hypothesis test, $p_1 = p_2 = p$, if p is known, then SE=$\sqrt{p(1-p)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$.

  Otherwise replace p $\sqrt{\frac{p_1(1-p_1)}{n_1}+\frac{p_2(1-p_2)}{n_2}}$ with $\bar{p} = \frac{n_1\bar{p}_1+n_2\bar{p}_2}{n_1+n_2}$.