Lecture 20

Type II Error and Power; Inference on Two Population Means

STAT 8010 Statistical Methods I
October 7, 2019

Whitney Huang
Clemson University

Agenda

1. Type I & Type II Errors
2. Sample Size Determination
3. Duality of Hypothesis Test with Confidence Interval
4. Inference on Two Population Means

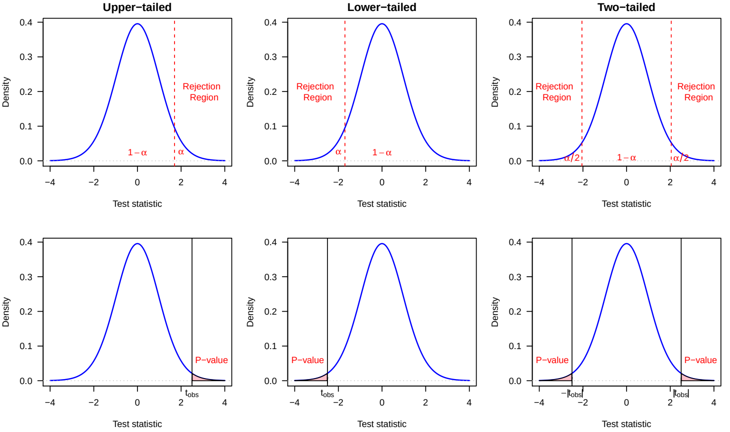
Review: Hypothesis Testing

1. State the null H_0 and the alternative H_a hypotheses
   - $H_0 : \mu = \mu_0$ vs $H_a : \mu > \mu_0 \Rightarrow$ Upper-tailed
   - $H_0 : \mu = \mu_0$ vs $H_a : \mu < \mu_0 \Rightarrow$ Lower-tailed
   - $H_0 : \mu = \mu_0$ vs $H_a : \mu \neq \mu_0 \Rightarrow$ Two-tailed
2. Compute the test statistic

   $t^* = \frac{\bar{X}_n - \mu_0}{s/\sqrt{n}}$ (σ unknown); $z^* = \frac{\bar{X}_n - \mu_0}{\sigma/\sqrt{n}}$ (σ known)
3. Identify the rejection region(s) (or compute the P-value)
4. Draw a conclusion

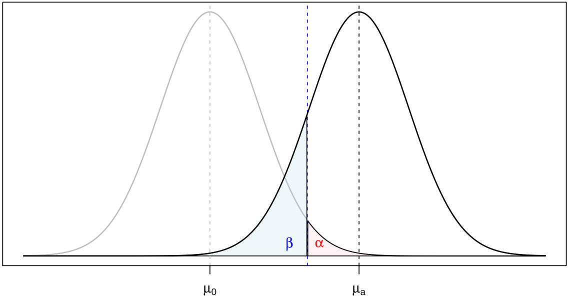
   We do/do not have enough statistical evidence to conclude H_a at α significant level

Region Region and P-Value Methods



Type I & Type II Errors

- Type I error: $\mathbb{P}(\text{Reject } H_0 | H_0 \text{ is true}) = \alpha$
- Type II error: $\mathbb{P}(\text{Fail to reject } H_0 | H_0 \text{ is false}) = \beta$



The relationship between α and β: $\alpha \downarrow \beta \uparrow$ and vice versa

Type II Error and Power

- The type II error, β, depends upon the true value of μ (let's call it μ_a)
- We use the formula below to compute β

$$\beta(\mu_a) = \mathbb{P}(z^* \leq z_\alpha - \frac{|\mu_0 - \mu_a|}{\sigma/\sqrt{n}})$$

- The power (PWR): $\mathbb{P}$(Reject $H_0 | H_0$ is false) $= 1 - \beta$. Therefore PWR$(\mu_a) = 1 - \beta(\mu_a)$

Question: What increases Power?

Sample Size Determination

Suppose that we wish to determine what sample size is required to detect the difference between a hypothesized mean and true mean $\mu_0 - \mu_a$, denoted by Δ, with a given power $1 - \beta$ and specified significance level α and known standard deviation σ. We can use the following formulas

$$n = \sigma^2 \frac{(z_\alpha + z_\beta)^2}{\Delta^2} \text{ for a one-tailed test}$$

$$n \approx \sigma^2 \frac{(z_{\alpha/2} + z_\beta)^2}{\Delta^2} \text{ for a two-tailed test}$$

Example

An existing manufacturing process produces, on average, 100 units of output per day. A pilot plant is used to evaluate a possible process change. Suppose the Company CEO wants to know if yield is increased. The CEO uses $\alpha = 0.05$ and the sample mean $(n = 25)$ is 103. Do we have sufficient evidence to conclude that the mean yield exceeds 100 if $\sigma = 10$?

1. $H_0 : \mu = 100$ vs. $H_a : \mu > 100$
2. $z_{obs} = \frac{103-100}{10/\sqrt{25}} = 1.5$
3. The cutoff value of the rejection region is $z_{0.05} = 1.645$. Therefore we do not have enough evidence to conclude that the new process mean yield exceeds 100

Example Cont'd

Suppose the true true mean yield is104.

- What is the power of the test?

Example Cont'd

Suppose the true true mean yield is104.

- What is the power of the test?

$$\begin{aligned}\beta(\mu = 104) &= \mathbb{P}\left(Z \le z_{0.05} - \frac{|100-104|}{10/\sqrt{25}}\right)\\ &= \mathbb{P}(Z \le 1.645 - 4/2) = \mathbb{P}(Z \le -0.355)\\ &= \Phi(-0.355) = 0.3613\end{aligned}$$

Therefore, the power is $1 - 0.3613 = 0.6387$

- What sample size is required to yield a power of 0.8 with a significance level of 0.05?

$$n = \sigma^2 \frac{(z_{0.05} + z_{0.2})^2}{\Delta^2} = 10^2 \frac{(1.645 + 0.8416)^2}{4^2} = 38.6324$$

Therefore, the required sample size is 39

Duality of Hypothesis Test with Confidence Interval

There is an interesting relationship between CIs and hypothesis tests. If H_0 is rejected with significance level α then the corresponding confidence interval does not contain the value μ_0 targeted in the hypotheses with the confidence level $(1-\alpha)$, and vice versa

Hypothesis testing at α level	$(1-\alpha)$-level Confidence Interval
$H_0 : \mu = \mu_0$ vs. $H_a : \mu \neq \mu_0$	$\bar{X} \pm t(\alpha/2, n-1)s/\sqrt{n}$
$H_0 : \mu = \mu_0$ vs. $H_a : \mu > \mu_0$	$(\bar{X} - t(\alpha/2, n-1)s/\sqrt{n}, \infty)$
$H_0 : \mu = \mu_0$ vs. $H_a : \mu < \mu_0$	$(-\infty, \bar{X} + t(\alpha/2, n-1)s/\sqrt{n})$

Comparing Two Population Means

- We often interested in comparing two groups (e.g.)
  - Does a particular pesticide increase the yield of corn per acre?
  - Do men and women in the same occupation have different salaries?
- The common ingredient in these questions: They can be answered by conducting statistical inferences of two populations using two (independent) samples, one from each of two populations

Notation

- Parameters:
  - Population means: μ_1, μ_2
  - Population standard deviations: σ_1, σ_2
- Statistics:
  - Sample means: $\bar{X}_1, \bar{X}_2$
  - Sample standard deviations: s_1, s_2
  - Sample sizes: n_1, n_2

Statistical Inference for $\mu_1 - \mu_2$

- Point estimate: $\bar{X}_1 - \bar{X}_2$
- Interval estimate: Need to figure out $\sigma_{\bar{X}_1 - \bar{X}_2}$
- Hypothesis Testing:
  - Upper-tailed test: $H_0 : \mu_1 - \mu_2 = 0$ vs. $H_a : \mu_1 - \mu_2 > 0$
  - Lower-tailed test: $H_0 : \mu_1 - \mu_2 = 0$ vs. $H_a : \mu_1 - \mu_2 < 0$
  - Two-tailed test: $H_0 : \mu_1 - \mu_2 = 0$ vs. $H_a : \mu_1 - \mu_2 \neq 0$

Confidence Intervals for $\mu_1 - \mu_2$

If we are willing to **assume** $\sigma_1 = \sigma_2$, then we can “pool” these two (independent) samples together to estimate the common σ using s_p:

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

We can then derive the (estimated) standard error of $\bar{X}_1 - \bar{X}_2$, which takes the following form

$$\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}} = s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

With CLT (assuming sample sizes are sufficiently large), we obtain the $(1 - \alpha) \times 100\%$ CI for $\mu_1 - \mu_2$:

$$\underbrace{\bar{X}_1 - \bar{X}_2}_{\text{point estimate}} \pm \underbrace{t(\alpha/2, n_1 + n_2 - 1)s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}_{\text{margin of error}}$$

Confidence Intervals for $\mu_1 - \mu_2$: What if $\sigma_1 \neq \sigma_2$?

- We will use s_1^2, s_2^2 as the estimates for σ_1^2 and σ_2^2 to obtain the standard error:

$$\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

- The formula for the degrees of freedom is somewhat complicated:

$$\frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{(s_1^2/n_1)^2}{n_1 - 1} + \frac{(s_2^2/n_2)^2}{n_2 - 1}}$$

- We can then construct the $(1-\alpha) \times 100\%$ CI for $\mu_1 - \mu_2$:

$$\underbrace{\bar{X}_1 - \bar{X}_2}_{\text{point estimate}} \pm \underbrace{t(\alpha/2, \text{ df calculated from above }) \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}_{\text{margin of error}}$$

Summary

In this lecture, we learned

- Type II error β and power $1-\beta$
- Sample size determination for given α, β, $\Delta = |\mu_a - \mu_0|$
- The Duality of hypothesis test with confidence interval
- Point/Interval estimate for $\mu_1 - \mu_2$

In next lecture we will learn

- Test if $\sigma_1 = \sigma_2$
- Hypothesis Testing for $\mu_1 - \mu_2$