LECTURE 20

11/9/23

- confidence interval for one population proportion:

$$\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

but the intervals weren't big enough.

- Agresti-Coull Adjustment: add 4 more Bernoulli trials, 2 of which are successes.

$\tilde{n} = n+4$ $\quad \tilde{p} = \frac{x+2}{n+4} = \frac{x+2}{\tilde{n}}$

So, CI is

$$\tilde{p} \pm z_{\alpha/2}\underbrace{\sqrt{\frac{\tilde{p}(1-\tilde{p})}{\tilde{n}}}}_{\text{standard error to estimate SD}}$$

$SD(\tilde{p}) = \sqrt{\frac{p(1-p)}{n}}$ = unknown b/c population proportion unknown.

- If lower bound of CI < 0, make it a 0
  If upper bound of CI > 1, make it a 1
- When we do this adjustment, we only have to check random sample as $\tilde{p}$ is always gonna be Normal

Example: (Modified Exercise 5.2.1, page 342) In a simple random sample of 70 automobiles, 28 of them were found to have emissions levels that exceed a state standard.

a. Find a 95% confidence interval for the proportion of automobiles in the state whose emission levels exceed the standard.

Step 1: Let p be the population proportion of all automobiles in the state whose emission levels exceed the state standard.

Step 2: Assumption: random sample with n < 5% population size.
Conditions: The problem states that it's random.
70 is 5% of 1400, and it's reasonable that there are more than 1400 autos in state.

Step 3:

$\tilde{p} \pm z_{\alpha/2}\sqrt{\frac{\tilde{p}(1-\tilde{p})}{\tilde{n}}} = 0.4054 \pm 1.96\sqrt{\frac{0.4054(1-0.4054)}{74}} \approx 0.4054 \pm 0.1119$

$= (0.2935, 0.5173)$

$\tilde{n} = n+4 = 74$

$\tilde{p} = \frac{28+2}{\tilde{n}} = \frac{30}{74} \approx 0.4054$

$95\% \Rightarrow \alpha = 0.05 \Rightarrow \frac{\alpha}{2} = 0.025$

$z_{\frac{\alpha}{2}} = z_{0.025} = 1.96$

Step 4:

With 95% confidence, we estimate that the population proportion of all automobiles in the state whose emission levels exceed the state standard is between 29.35% and 51.73%.

b. How many automobiles must be sampled to specify the proportion that exceed the standard to within ±0.03 (m) with 95% confidence?

$ME = z_{\frac{\alpha}{2}}\sqrt{\frac{\tilde{p}(1-\tilde{p})}{\tilde{n}}} \leq m$

$\frac{\tilde{p}(1-\tilde{p})}{n+4} \leq \left(\frac{m}{z_{\alpha/2}}\right)^2$

$\frac{\tilde{p}(1-\tilde{p})(z_{\alpha/2})^2}{m^2} \leq n+4$

$\boxed{n \geq \tilde{p}(1-\tilde{p})\left(\frac{z_{\alpha/2}}{m}\right)^2 - 4}$

Case 1: Some previous studies provide us w/ $\tilde{p}$, so we use our formula.

Case 2: no studies have been done, so we don't have any estimate of p. In this case, we replace $\tilde{p}(1-\tilde{p})$ w/ its max. value, 0.25

$n \geq \frac{30}{74}\left(1-\frac{30}{74}\right)\left(\frac{1.96}{0.03}\right)^2 - 4 = 1024.92$ $\boxed{n = \text{at least } 1025}$

$n \geq \frac{1}{4}\left(\frac{1.96}{0.03}\right)^2 - 4 = 1063.11$ $\boxed{n = \text{at least } 1064}$

LECTURE 21

11/14/23

Inference about $\mu_1 - \mu_2$

– Assumptions:

independence <u>within</u> the samples

1. $X_1, X_2, \ldots, X_{n_x}$ is a random sample of size n_x from a normally distributed population w/ mean μ_x and var σ_x^2
2. $Y_1, Y_2, \ldots, Y_{n_Y}$ is a random sample of size n_Y from a normally distributed population w/ mean μ_Y and var σ_Y^2
3. X and Y samples are independent of one another.

↳ independent b/w the samples

– If $X \sim$ Normal and $Y \sim$ Normal, THEN $\bar{X} - \bar{Y} \sim$ exactly Normal

$$\bar{X} - \bar{Y} \sim N\left(\mu_x - \mu_Y, \frac{\sigma_x^2}{n_x} + \frac{\sigma_Y^2}{n_Y}\right)$$

– If $X \sim$ not Normal, but $n_x \geq 30$ OR $Y \sim$ not Normal but $n_Y \geq 30$, then use CLT to conclude that $\bar{X} - \bar{Y} \sim$ approx. Normal

$$\bar{X} - \bar{Y} \sim N\left(\mu_x - \mu_Y, \frac{\sigma_x^2}{n_x} + \frac{\sigma_Y^2}{n_Y}\right) \text{ (approximately)}$$

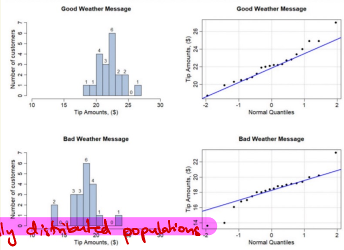
→ both X and Y need to be Normal/approx. Normal.

– If we know σ_x and σ_Y, then we can use z-test otherwise, we use t-test

* random assignment in one population ensures independence

Example: (Source: Exercise 21.37, page 501, Moore, Notz, and Fligner, *Introduction to the Practice of Statistics*, 9e, 2021) Researchers gave 40 index cards to a waitress at an Italian restaurant in New Jersey. Before delivering the bill to each customer, the waitress selected one of the cards without reading it and wrote on the bill the same message that was printed on the card. Twenty of the cards had the message, "The weather is supposed to be really good tomorrow. I hope you enjoy the day!" GW Another 20 cards contained the message, "The weather is supposed to be not so good tomorrow. I hope you enjoy the day anyway!" BW After the customers left, the waitress recorded the amount of the tip (percent of bill) before taxes.

	mean	sd	0%	25%	50%	75%	100%	n
BadWeather	18.19	2.105	13.6	17.375	18.45	19.125	23.2	20
GoodWeather	22.22	1.955	18.7	20.750	22.15	22.950	27.0	20



both look approx. Normal

Step 1: Let $\mu_B - \mu_G$ be the difference b/w the population mean amount of the tip (as % of the bill) for customers who receive a message about bad weather and the population mean amount of the tip (as % of the bill) for customers who receive a message about good weather for population of all customers.

Step 2:

State Assumptions: two independent random samples from Normally distributed populations

Check Conditions:

1. Independence between

Customers randomly assigned to the treatment, so we have independent groups.

2. Independence within

Not stated if the customers were randomly selected. Reasonable to assume that one person's tip doesn't affect another person's tip. 20 = 5% of 400, so reasonable to assume that >400 customers got good and >400 got bad. Proceed w/ caution

3. Underlying population dist. both Normal.

We aren't told that the population distributions are Normal. Looking at the data, both samples look approx. Normal and symmetric. The qq plots also look about Normal as the line is straight. So, reasonable to think that the tip amounts came from a Normally distributed population.

- Three possible hypotheses:

H_0:	$\mu_x - \mu_y \leq \Delta_0$	$\mu_x - \mu_y = \Delta_0$	$\mu_x - \mu_y \geq \Delta_0$
H_A:	$\mu_x - \mu_y > \Delta_0$	$\mu_x - \mu_y \neq \Delta_0$	$\mu_x - \mu_y < \Delta_0$

- When σ_x or σ_y is unknown,

$$T = \frac{(\bar{x}-\bar{y}) - \Delta_0}{\sqrt{\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}}}$$

- use technology to compute df.

hypotheses:

H_0: $\mu_B - \mu_G = 0$

H_A: $\mu_B - \mu_G \neq 0 \longleftrightarrow \mu_B \neq \mu_G$

Step 3:

$$t = \frac{(\bar{x}-\bar{y}) - 0}{\sqrt{\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}}} = \frac{(18.19 - 22.22)}{\sqrt{\frac{2.105^2}{20} + \frac{1.955^2}{20}}} = -6.274$$

Calculate p-value using df = 37 (round down) Assuming H_0 true

H_A is two-sided

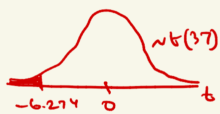
$\frac{1}{2}$ p-value $= P[\bar{x} - \bar{y} \leq -4.03$ assuming H_0 is true$]$

$= P[t < -6.274]$

< 0.0005 (by table)

p-value $< 0.0010 \leq \alpha$ for $\alpha = 0.01$, 0.05 or 0.10

reject H_0



Step 4:

At $\alpha = 0.01$ significance level, there is sufficient evidence ($p < 0.001$) to suggest that the population mean amount of tip (as % of bill) left by customers who got the bad massage is different than the population mean amount of tip (as % of bill) left by customers who got the good massage in the population of all customers.

LECTURE 22

11/16/23

- Confidence Interval:

$$(\bar{x}-\bar{y}) \pm t_{df,\frac{\alpha}{2}}\sqrt{\frac{s_x^2}{n_x}+\frac{s_y^2}{n_y}}$$

- can calculate upper/lower bound by only doing +/- and changing $\frac{\alpha}{2}$ to α
- If there is a difference, what is a reasonable range of values for the difference in mean tipping percentage?

Step 1/2 : same as HT

Step 3 : 95% CI for $\mu_B-\mu_G$:

$$(\bar{x}-\bar{y}) \pm t_{37,0.025}\sqrt{\frac{s_x^2}{n_x}+\frac{s_y^2}{n_y}}$$

$$= (18.19-22.22) \pm 2.030\sqrt{\frac{1.955^2}{20}+\frac{2.105^2}{20}}$$

$$= -4.03 \pm 1.304 = (-5.334, -2.726)$$

$$-5.334 < \mu_B-\mu_G < -2.726$$

$$\mu_G-5.334 < \mu_B < \mu_G-2.726$$

Step 4:

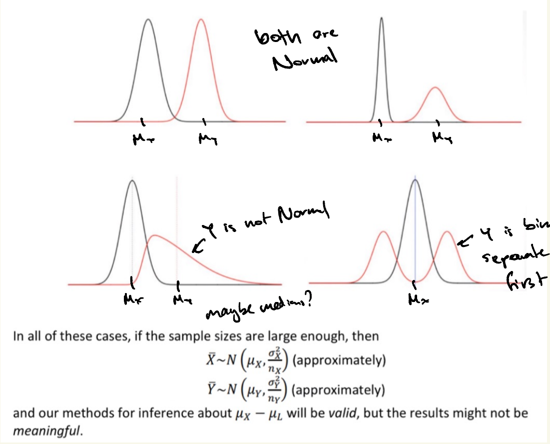
With 95% confidence, we estimate that the population mean amount of tip in % of the bill left by the customers who get the bad message will be b/w 5.334% lower and 2.726% lower than the population mean amount of tip as % of the bill left by customers who get the good message for the population of all customers.

- pool when two population parameters are the same, but their common value is unknown. The we estimate it using data from both samples.

for H_0: $\mu_1-\mu_2 \leq 0$ $\quad \mu_1-\mu_2 > 0$

$\sigma_1^2 = \sigma_2^2 = \sigma \leftarrow$ with s_{pooled} computed using $X_1, \dots, X_{n_x}$ and $Y_1, \dots, Y_{n_y}$

- we won't use pooling for means



INFERENCE OF μ_D

- used in experiment that pairs together both samples, making matched pairs.

$$\begin{array}{lllll} & x_1 & x_2 & \dots & x_n \\ (-) & y_1 & y_2 & \dots & y_n \\ \hline & d_1 & d_2 & \dots & d_n \rightarrow \bar{d} \end{array} \qquad \mu_D$$

- parameter of interest is μ_D
- Assumptions (same as one population mean)
  1. Differences $D_1, \dots, D_n$ are a random sample of differences
  2. from Normal population of differences
  3. σ_D^2 = known $\longrightarrow$ not satisfied $\rightarrow$ s_D^2 $\rightarrow$ t-test
- Hypotheses:

$$H_0: \mu_D \geq \Delta_0 \qquad \mu_D = \Delta_0 \qquad \mu_D \leq \Delta_0$$

$$H_A: \mu_D < \Delta_0 \qquad \mu_D \neq \Delta_0 \qquad \mu_D > \Delta_0$$

- test statistic:

$$T = \frac{\bar{D} - \Delta_0}{s_D/\sqrt{n}}$$

can use CLT here, just like one μ

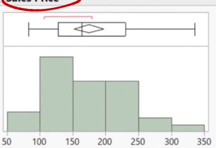
- Confidence Interval:

$$\bar{d} \pm t_{n-1,\alpha/2} \frac{s_D}{\sqrt{n}}$$

can calculate upper/lower bound through making it α

Example: Real estate is typically reassessed annually for property tax purposes. This assessed value of a home, however, is not necessarily the same as the fair market value of the property (i.e., how much the property can be sold for). The sales price and assessed value for each of 34 randomly selected recently sold single-family homes in a Midwestern city were recorded and use for the following output. Both variables are measured in thousands of dollars. Does it appear that sales prices are, on average, higher than assessed values?

Sales Price



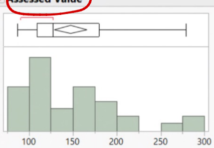
Quantiles

100.0%	maximum	335
99.5%		335
97.5%		335
90.0%		270
75.0%	quartile	230.125
50.0%	median	163
25.0%	quartile	127.175
10.0%		95.75
2.5%		82.5
0.5%		82.5
0.0%	minimum	82.5

Summary Statistics

Mean	174.82647
Std Dev	64.528529
Std Err Mean	11.066551
Upper 95% Mean	197.34154
Lower 95% Mean	152.3114
N	34

Assessed Value



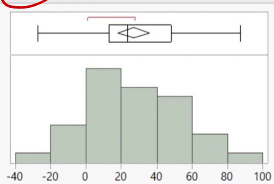
Quantiles

100.0%	maximum	279
99.5%		279
97.5%		279
90.0%		231.85
75.0%	quartile	179.9
50.0%	median	127.6
25.0%	quartile	108.6
10.0%		91.55
2.5%		87
0.5%		87
0.0%	minimum	87

Summary Statistics

Mean	147.54118
Std Dev	53.287921
Std Err Mean	9.138803
Upper 95% Mean	166.13421
Lower 95% Mean	128.94814
N	34

Diff (SP-AV) Sales − Assessed



Quantiles

100.0%	maximum	87.6
99.5%		87.6
97.5%		87.6
90.0%		64.85
75.0%	quartile	48.25
50.0%	median	23.35
25.0%	quartile	12.825
10.0%		-5.15
2.5%		-27.5
0.5%		-27.5
0.0%	minimum	-27.5

Summary Statistics

Mean	27.285294
Std Dev	25.548577
Std Err Mean	4.3815448
Upper 95% Mean	36.199614
Lower 95% Mean	18.370974
N	34

distribution is unimodal, almost symmetric, but slightly skewed left $\Rightarrow$ median, IQR

Sample mean difference of $\bar{d} = 27.29$

Sample SD of differences is $s_D = 25.55$

No apparent outlier.

*We have sales price and assessed value for each home, which are dependent samples, so we should do matched pairs.

Step 1: Let μ_D be the population mean of the differences between sales price and assessed value of a single family home in this Midwestern city for the population of all such homes.

$$H_0: \mu_D \leq 0 \qquad H_A: \mu_D > 0$$

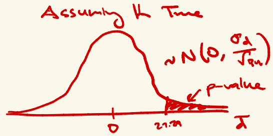
Step 2:

State Assumptions: The differences are a random sample of differences from a Normally distributed population of differences.

Check Conditions:

1. stated that it's a random sample
   Also $34 = 5\%$ of 680, it is reasonable to assume that population size > 680.
2. Since $n = 34 \geq 30$, by CLT, the sampling dist. of $\bar{D}$ will be approx. Normal.
3. σ_D unknown $\longrightarrow$ use t-test

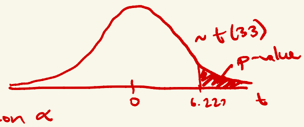
Step 3



$$t = \frac{\bar{d} - \Delta_0}{s_d/\sqrt{n}} = \frac{27.285294 - 0}{25.548574/\sqrt{34}} = 6.227$$

$df = 34 - 1 = 32$ (use 30 w/ table)

$H_A: \mu_d > 0$ right-sided



p-value: $P[\bar{d} \geq 27.289294$ assuming $\mu_d = 0]$

$= P[T \geq 6.227] < 0.0005 \leq \alpha$ for common α

$\downarrow$

reject H_0

Step 4:

for 5% level of significance there is sufficient evidence (p-value < 0.0005) to suggest that the sales price is higher than the assessed value, on average, in the population of all single-family homes in this city

We will compute a 95% LCB for μ_d

Steps 1, 2 - same as H T

Step 3

$$\bar{d} - t_{n-1,\alpha}\frac{s_d}{\sqrt{n}} = 27.285294 - 1.697\,\frac{25.548577}{\sqrt{34}} = 19.8498$$

$n - 1 = 33$

$\alpha = 0.05$

$t_{33,0.05} \approx 1.697$

Step 4

With 95% confidence, we estimate that the sales price, on average, is \$19,850 higher than the assessed value in the population of all single-family homes in this Midwestern city.

LECTURE 23

11/21/23

Inference of $p_x - p_y$

- difference of 2 population prop.
- similar to inference for one p.
- Now, we need independent, random samples from 2 Bernoulli populations to have all of them independent.
- 3 possible hypotheses:

H_0	$p_x \geq p_y$	$p_x = p_y$	$p_x \leq p_y$
H_A:	$p_x < p_y$	$p_x \neq p_y$	$p_x > p_y$

- we use $\hat{p}_x - \hat{p}_y$ to estimate $p_x - p_y$

$$E[\hat{p}_x - \hat{p}_y] = p_x - p_y$$

$$Var[\hat{p}_x - \hat{p}_y] = \frac{p_x(1-p_x)}{n_x} + \frac{p_y(1-p_y)}{n_y}$$

- distribution of $\hat{p}_x - \hat{p}_y$ will be approx. Normal when we expect ≥ 10 successes and ≥ 10 failures in each sample, assuming H_0 is true.
- Under H_0, $p_x = p_y = p$, so we use both $\hat{p}_x$ and $\hat{p}_y$ to estimate p.
  - ↳ so we find <u>pooled</u> proportion $\hat{p}$!

$$\hat{p}_{pooled} = \hat{p} = \frac{X+Y}{n_x + n_y}$$

- test statistic:

$$z = \frac{(\hat{p}_x - \hat{p}_y) - 0}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_x} + \frac{1}{n_y}\right)}}$$

→ $SD(\hat{p}_x - \hat{p}_y)$ under H_0 estimated with $SE(\hat{p}_x - \hat{p}_y)$.

$$\hat{p}_x = \frac{X}{n_x} \qquad \hat{p}_y = \frac{Y}{n_y}$$

- to check that $(\hat{p}_x - \hat{p}_y)_0 \sim$ approx. Normal,
  1. $n_x \hat{p} \geq 10$ $\quad n_x(1-\hat{p}) \geq 10$
  2. $n_y \hat{p} \geq 10$ $\quad n_y(1-\hat{p}) \geq 10$
- For CI of 2 p, we need Agresti-Coulli adjustment:

$$\tilde{n}_x = n_x + 2, \quad \tilde{n}_y = n_y + 2$$

$$\tilde{p}_x = \frac{X+1}{n_x+2}, \quad \tilde{p}_y = \frac{Y+1}{n_y+2}$$

$$(\tilde{p}_x - \tilde{p}_y) \pm z_{\alpha/2}\sqrt{\frac{\tilde{p}_x(1-\tilde{p}_x)}{\tilde{n}_x} + \frac{\tilde{p}_y(1-\tilde{p}_y)}{\tilde{n}_y}}$$

- If UB < -1, LB $= -1$; UB > 1, UB $= 1$

- HT assumptions:
  1. Two independent random samples
  2. expected successes, failures ≥ 10

- CI assumptions:
  1. Two independent random samples

Example: A Harris Interactive survey asked adults from the United States and Germany about the importance of brand names when buying clothes. Of the 2309 U.S. adults surveyed, 600 said brand names were important, compared with 233 of the 1058 German adults. Do these results suggest that the population proportions of adults who think brand names are important in the U.S. and Germany differ?
(Source: *Statistics and Probability with Applications* (3e), Starnes and Tabor, page 625)

Step 1:

Let $p_x - p_y$ denote the difference in the population proportion of adults who think brand names are important in the population of all adults in the US and the population proportion of adults who think that brand names are important in the population of all adults in Germany.

$H_0: p_x = p_y \qquad H_A: p_x \neq p_y$

Step 2:

State Assumptions: We have independent random samples from each of the 2 Bernoulli populations. $\hat{p}_x - \hat{p}_y \sim$ approx. Normal under H_0

Check Conditions:

Independence between: Reasonable to believe that the opinions of US adults surveyed did not affect opinions of German adults surveyed and vice versa

Independence within: Not stated that random samples. Reasonable to believe that the opinions of US adults surveyed did not affect opinions of other US adults surveyed and same for the Germans.

Is $\hat{p}_x - \hat{p}_y$ Normal under H_0

1. $\hat{p} = \frac{X+Y}{n_x+n_y} = \frac{600+233}{2309+1058} = \frac{833}{3367} \approx 0.247$

$n_x \hat{p} = 2309(0.247) \approx 570.3 \geq 10$ ✓

$n_x(1-\hat{p}) = 2309(1-0.247) \approx 1738.7 \geq 10$ ✓

$n_y \hat{p} = 1058(0.247) = 261.3 \geq 10$ ✓

$n_y(1-\hat{p}) = 1058(1-0.247) \approx 796.7 \geq 10$ ✓

Yes, assuming H_0 is true, normal approximation is appropriate for $\hat{p}_x - \hat{p}_y \rightarrow$ z-test can be used.

Step 3:

$$z = \frac{(\hat{p}_x - \hat{p}_y) - 0}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_x}+\frac{1}{n_y}\right)}} = \frac{\frac{600}{2309} - \frac{233}{1058}}{\sqrt{\frac{833}{3367}\left(1-\frac{833}{3367}\right)\left(\frac{1}{2309}+\frac{1}{1058}\right)}} = 2.47$$

$H_A: p_x \neq p_y$ two-sided.

$\frac{1}{2}$ p-value $= P[\hat{p}_x - \hat{p}_y \geq 0.0396$ assuming $p_x = p_y] = P[z \geq 2.47] = 1 - P[z < 2.47] = 0.0068$

p-value $= 0.0136 \leq \alpha = 0.05 \Rightarrow$ reject H_0

Step 4

At 5% significance, there is sufficient evidence ($p = 0.0136$) to suggest that the population prop. of adults in US who think that brand names are important when buying clothes is different than German adults who think that brand names are important.

Example: (Exercise 5.5.3a, page 362) Angioplasty is a medical procedure in which an obstructed blood vessel is widened. In some cases, a wire mesh tube, called a stent, is placed in the vessel to help it remain open. The article "Long-term Outcomes of Patients Receiving Drug-eluting Stents" (A. Philpott, D. Southern, et al., *Canadian Medical Association Journal*, 2009:167–174) presents the results of a study comparing the effectiveness of a bare metal stent with one that has been coated with a drug designed to prevent reblocking of the vessel. A total of 5320 patients received bare metal stents, and of these, 841 needed treatment for reblocking within a year. A total of 1120 received drug coated stents, and 134 of them required treatment within a year. Find a 98% confidence interval for the difference between the proportions for drug coated stents and bare metal stents.

Step 1

Let $p_X - p_Y$ be the difference in population proportions of patients who need treatment for reblocking within a year in the population of all patients w/ bare metal stents and the population proportion of patients who need reblocking within a year in the population of patients w/ drug stents.

Step 2

State Assumptions: Independent random samples from each of the 2 Bernoulli pop.

Check Conditions:

Independence b/w: doesn't depend on other group

Independence within: doesn't depend on each others within each group.

Step 3

$\tilde{n}_X = n_X + 2 = 5320 + 2 = 5322$, $\quad \tilde{p}_X = \frac{X+1}{\tilde{n}_X} = \frac{842}{5322}$

$\tilde{n}_Y = n_Y + 2 = 1120 + 2 = 1122$, $\quad \tilde{p}_Y = \frac{Y+1}{\tilde{n}_Y} = \frac{135}{1122}$

$98\%\ CI \Rightarrow \frac{\alpha}{2} = 0.01$

$z_{0.01} = 2.326$

$$(\tilde{p}_X - \tilde{p}_Y) \pm z_{\alpha/2}\sqrt{\frac{\tilde{p}_X(1-\tilde{p}_X)}{\tilde{n}_X} + \frac{\tilde{p}_Y(1-\tilde{p}_Y)}{\tilde{n}_Y}} = \ldots = 0.038 \pm 0.0254 = (0.0126, 0.0634)$$

Step 4

With 98% confidence, we estimate that the population proportion of all patients who needed treatment for reblocking within a year of receiving a bare metal stent will be between 1.26% higher and 6.34% higher than the population proportion of all patients who needed reblocking within a year of getting a drug-coated stent.

Other Considerations

- practically significant ≠ statistically significant

Example: A current manufacturing process produces outputs with an average breaking strength of 50 N. A new (more expensive) process is proposed. We have a random sample of 1000 measurements from the new process. These measurements have mean 50.1 N and standard deviation 1 N. The new process will be used if it can be shown that it produces stronger outputs. In this case, we have the following,

$H_0: \mu \leq 50$ versus $H_1: \mu > 50$, test statistic: $t = 3.162$, P-value = 0.0008

We have a very small P-value, which gives us strong evidence against the null hypothesis. We conclude that the new process is better.

But... Our point estimate of the breaking strength for the new material is only 50.1 N, which is very close to the mean value for the current material, 50 N. Even though it is very likely that the new material is stronger, an increase in average strength of 0.1 N may not be worth the cost of converting to the new (more expensive) process.

Moral: Practical significance must be judged in the context of the problem.

- more important than passing the α threshold is scientific reasoning.
- If H_0 is true, but we reject it, then Type I error.
- If H_0 is false, but we fail-to-reject it, then Type II error.
-

	Reject H_0	fail-to-reject H_0
H_0 T	Type I	correct
H_0 F	correct	Type II

- P(Type I) and P(Type II) are inversely correlated
- when choosing α, we need to consider which error is worse and minimize the Probability of that

Example: When a parachute is inspected, the inspector is looking for anything that might indicate that the parachute will not open.

H_0: the parachute will open

H_1: the parachute will not open

Identify the Type I and Type II errors for this situation. Which error is worse?

Type I: Conclude that parachute won't open when it will.
Waste of money as a good parachute will be thrown out.

Type II: Conclude that parachute will open when it won't.
Severe injury and maybe death.

Here, we need to minimize Type II error.

Controlling Type I Errors

Say we are going to collect a large sample of data and will perform an α-level test of

$$H_0: \mu \le \mu_0 \text{ versus } H_1: \mu > \mu_0$$

Question: What is the probability that we will make a Type I error?

$P[\text{Type I}] = P[\text{having } \bar{x} \text{ where } p \le \alpha \text{ when } H_0 \text{ true}]$

$= P[\bar{x} \ge \bar{x}_a \text{ when } H_0 \text{ true}]$

$= \alpha$

$P[\text{Type II}] = P[\text{having } \bar{x} \text{ where } p > \alpha \text{ when } H_A \text{ true}]$

(depends on H_A)

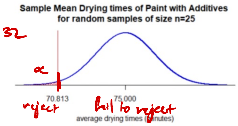
Example: The drying time of a certain type of paint under specified test conditions is known to be normally distributed with mean 75 minutes and standard deviation 9 minutes. Chemists have proposed a new additive designed to decrease drying time. The company will collect 25 samples of paint with the additive. Because of the expense associated with the additive, evidence should strongly suggest an improvement in average drying time before such a conclusion is adopted, and it has been decided that a significance level of $\alpha = 0.01$ will be used when analyzing the results of the experiment.

$H_0: \mu \ge 75 \quad H_A: \mu < 75$

$P[\bar{x} \le \bar{x}_a = c] = P[Z \le -z_\alpha] = \alpha$

$\alpha = 0.01 \Rightarrow z_1 = -z_{99} = -2.326$

$c = 75 - 2.326 \frac{9}{\sqrt{25}} = 70.8132$



$\beta = P[\text{Type II}]$

$= P[\bar{x} > 70.8132 \mid \mu < 75]$

Question: What is the probability of a Type II error if the mean drying time with the additive is $\mu = 72$ minutes?

$\beta = P[\bar{x} > 70.8132 \text{ when } \mu = 72]$

$= P\left[z \ge \frac{70.8132 - 72}{\frac{9}{\sqrt{25}}}\right] = 0.7454$