PW #8: Inference for Proportions +
Type I and Type II Errors

(1) Method of Inference: 95% confidence interval for proportions

Parameter of Interest: p = the true proportion of orders shipped on time

Conditions: SRS - stated in problem

Independence: all orders shipped $\geq 10(100)$

condition met

Normal - $86 \geq 10$ $14 \geq 10$ Condition met

Calculations: $\hat{p} = .86$ $\hat{q} = .14$ $z^*_{.95} = 1.96$ $n = 100$ $CL = .95$

$.86 \pm 1.96\left(\sqrt{\frac{(.86)(.14)}{100}}\right)$ $(.792, .928)$

$.86 \pm 1.96(.0347) = .86 \pm .068$

Conclusion: We are 95% confident the true proportion of orders shipped on time for this company is in the interval 79.2% to 92.8%.

(2) "Method of difference: 95% confidence interval for 2-sample proportions"

Parameters of Interest: p_1 = the true proportion of in-line skaters who <u>don't</u> <u>wear</u> wrist guards and suffer injury

p_2 = the true proportion of in-line skaters who <u>do</u> wear wrist guards and suffer injury

Conditions:	<u>No Wrist Guards</u>	<u>Wrist Guards</u>
SRS	Stated in problem	Stated in problem
Independence	$45 \geq 10$ $63 \geq 10$ Condition Met	$6 \geq 10$ * Proceed with caution $47 \geq 10$

Calculation: $\hat{p}_1 = \frac{45}{108} = .4167$ $\quad \hat{p}_2 = \frac{6}{53} = .1132$ $\quad z^*_{.95} = 1.96$

$\hat{q}_1 = \frac{63}{108} = .5833$ $\quad \hat{q}_2 = \frac{47}{53} = .8868$ $\quad CL = .95$

$n_1 = 53$ $\quad n_2 = 108$

$$(.4167 - .1132) \pm 1.96\left(\sqrt{\frac{(.4167)(.5833)}{53} + \frac{(.1132)(.8868)}{108}}\right)$$

$$.3035 \pm 1.96(.0648)$$

$$.3035 \pm .1270 \qquad (.1773, .4296)$$

We are 95% confident the true difference in proportion of injury suffered while in-line skating between skaters who don't wear wrist guards and skaters who do wear wrist guards is in the interval 17.7% to 43.0%. Since there is evidence that it is plausible injuries could occur ~~more than~~ 40% of the time when not wearing wrist guards, I would recommend wearing wrist guards when in-line skating.

(3) Method of Inference: 90% confidence interval for 2 sample proportion

Parameters of Interest: p_1 = the true proportion of dogs exposed to herbicide having lymphoma

p_2 = the true proportion of dogs not exposed to herbicide having lymphoma

Conditions:	Exposed	Not Exposed
SRS:	not stated, no reason to assume otherwise	not stated, no reason to assume otherwise
Independence:	all dogs $\geq 10(827)$ condition met	all dogs $\geq 10(130)$ condition met
Normal:	$473 \geq 10$ $354 \geq 10$ condition met	$19 \geq 10$ $111 \geq 10$ Condition Met

Calculations: $\hat{p}_1 = \frac{473}{827} = .572$ $\quad \hat{p}_2 = \frac{19}{130} = .146$ $\quad z^*_{.90} = 1.645$

$\hat{q}_1 = \frac{354}{827} = .428$ $\quad \hat{q}_2 = \frac{111}{130} = .854$ $\quad CL = .90$

$n_1 = 827$ $\quad n_2 = 130$

$$(.572 - .428) \pm 1.645\left(\sqrt{\frac{(.572)(.428)}{827} + \frac{(.146)(.854)}{130}}\right)$$

$$(.572 - .428) \pm 1.645(.0354)$$

$$(.368, .484)$$

We are 90% confident the true difference in proportions of dogs exposed to herbicide having lymphoma and dogs not exposed to herbicide having lymphoma is in the interval 36.8% to 48.4%.

(4) Method of Inference: 2 sample z-test for proportions

Parameters of Interest: p_1 = proportion of voters who favor the candidate in first poll

p_2 = proportion of voters who favor the candidate in second poll

Conditions:	First Poll	Second Poll
SRS	not stated, no reason to assume otherwise	not stated, no reason to assume otherwise
Independence	all voters in first poll $\geq 10(800)$ condition met	all voters in second poll $\geq 10(1000)$ condition met
Normal	$460 \geq 10$ $340 \geq 10$ condition met	$520 \geq 10$ $480 \geq 10$ condition met

Hypothesis:

H_0: no difference in proportion of voters in first and second poll who favor the candidate

Ha: the proportion of voters in the first poll who favor the candidate is higher than the proportion of voters in the second poll who favor the candidate

$H_0: p_1 - p_2 = 0$

$H_a: p_1 - p_2 > 0$

Calculations:

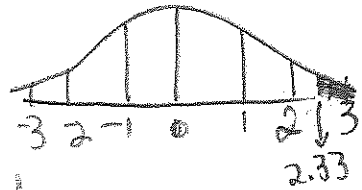
$\hat{p}_1 = \frac{460}{800} = .575$ $\quad \hat{p}_2 = \frac{520}{1000} = .52$ $\quad \alpha = .10$

$\hat{q}_1 = \frac{340}{800} = .425$ $\quad \hat{q}_2 = \frac{480}{1000} = .48$

$n_1 = 800$ $\quad n_2 = 1000$

④ cont'd

$$P\left(Z > \frac{(.575 - .52) - 0}{\sqrt{\frac{(.575)(.425)}{800} + \frac{(.52)(.48)}{1000}}}\right) = P(Z > 2.33) = .0099$$



Since our p-value of 0.0099 is smaller than the significance level of $\alpha = .10$, we have evidence to reject the null. We can conclude the candidate's popularity has decreased from the first poll to the second poll for sample sizes $n_1 = 800$ and $n_2 = 1000$

⑤ Method of Inference: 2 sample z-test for proportions

Parameters of Interest: p_1 = the true proportion of cars with defects using first procedure

p_2 = the true proportion of cars with defects using second procedure

$p_1 - p_2$ = difference in proportion of cars with defects

Conditions:

	1st procedure	2nd procedure
SRS	not stated, no reason to assume otherwise	not stated, no reason to assume otherwise
Independence	all cars produced by 1st procedure $\geq 350(10)$ Condition met	all cars produced by 2nd procedure $\geq 500(10)$ Condition met
Normal	$28 \geq 10$, $322 \geq 10$ condition met	$32 \geq 10$, $468 \geq 10$ Condition met

Hypothesis:

Ho: there is no difference in proportion of defects from 1st procedure and 2nd procedure

Ha: there is a difference in proportion of defects from 1st procedure and 2nd procedure

Ho: $p_1 - p_2 = 0$ Ha: $p_1 - p_2 \neq 0$ *Note "two-tailed"

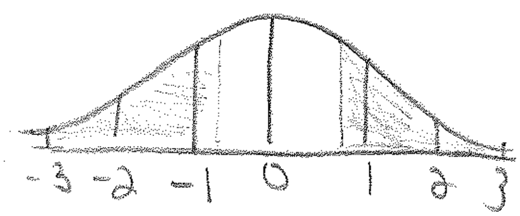
⑤ cont'd

Calculations: $\hat{p}_1 = \frac{28}{350} = .08$ $\hat{p}_2 = \frac{32}{500} = .064$ $\alpha = .10$

$\hat{q}_1 = \frac{322}{350} = .92$ $\hat{q}_2 = \frac{418}{500} = .936$

$n = 350$

$$2P\left(z > \frac{(.08 - .064) - 0}{\sqrt{\frac{(.08)(.92)}{350} + \frac{(.064)(.936)}{500}}}\right) = 2P(z > 0.896) = .370$$



Since our p-value of .37 is greater than our significance level of $\alpha = .10$, we have evidence to fail to reject the null. We can conclude that there is no difference in the proportion of defects from 1st procedure to 2nd procedure.

⑥ Method of Inference: 1 sample z-test for proportions

Parameter of Interest: p = the true proportion of people who preferred fresh brewed coffee

Conditions: SRS - not stated, no reason to assume otherwise

Independence - all coffee drinkers $\geq 10(50)$ Condition met

Normal - $19 \geq 10$ $31 \geq 10$ Condition met

Hypothesis: H_0: the true proportion of coffee drinkers who prefer fresh brewed coffee is .5

H_a: the true proportion of coffee drinkers who prefer fresh brewed coffee is more than .5

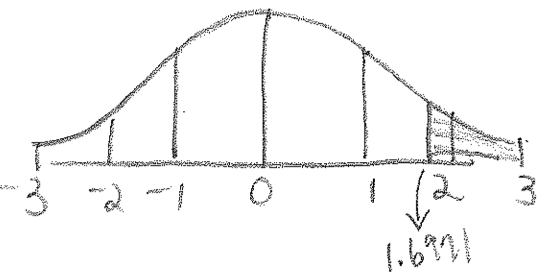
(6) cont'd

$H_0: p = .5$

$H_a: p > .5$

Calculations: $\hat{p} = \frac{31}{50} = .62$ $\hat{q} = \frac{19}{50} = .38$ $n = 50$ $\alpha = .05$

$$P\left(z > \frac{.62 - .5}{\sqrt{\frac{(.5)(.5)}{50}}}\right) = P(z > 1.697) = .0448$$



Since our p-value of .0448 is less than the significance level $\alpha = .05$, we have evidence to reject the null. We can conclude that a majority of people do prefer the taste of fresh brewed coffee.

(7) H_0: pregnant woman does not have gestational diabetes

H_a: pregnant woman does have gestational diabetes

Type I error: screening determines pregnant woman does have gestational diabetes when in fact she does not.

Consequences: • more tests will be run on the woman ($$)

• unnecessary stress on woman which will carry over to stress on baby

• prescribed medication and/or bedrest when not needed

Type II error: screening determines woman does not have gestational diabetes, when in fact she does

Consequence: medication not prescribed which endangers health (and possibly life) of both mother and baby

I would consider Type II more severe in this case. Death is a more serious consequence than money or stress.