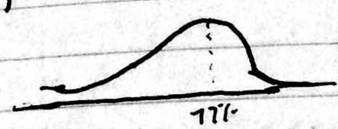
Unit 5: Problem Set

(6.2)

a) False. It does not satisfy the success/failure condition.

~~B)~~

b)



False, the distribution would be left-skewed because we would expect $\hat{p}$ to be near .77, which would cause the data to be left-skewed.

c)

~~SE~~ $SE_{\hat{p}} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

$SE_{\hat{p}} = \sqrt{\frac{.77(1-.77)}{60}} = .0543$

$\frac{.85-.77}{.0543} = 1.47$

False

when $\hat{p} = .85$, it is only 1.47 standard errors away from the mean, which is not unusual.

d)

$SE_{\hat{p}} = \cancel{.0543} \sqrt{\frac{.77(1-.77}{120}} = .0384$

$\frac{.85-.77}{\cancel{.0543}\ .0384} = 2.08$

True. $\hat{p} = .85$ is 2.08 standard errors away from the mean which is considered unusual

6.4

③ a) False. Since the sample size is only 12 families, it does not match the independence condition nor the success/failure condition. Therefore, we cannot determine if the distribution is left skewed.

b) False. The sample size needs only to be at least 30 to meet the independence condition.

c)

$$SE_{\hat{p}} = \sqrt{\frac{.25(1-.25)}{50}} = .0612$$

$$\frac{.20-.25}{.0612} = -.817$$

False. $SE_{\hat{p}} = .0612$ and $\hat{p} = .20$ is only .817 ~~away f~~ SE's away from the mean which would not ~~bee~~ be considered unusual.

d)

$$SE_{\hat{p}} = \sqrt{\frac{.25(1-.25)}{150}} = .0354$$

$$\frac{.20-.25}{.0354} = 1.41$$

False. $\hat{p} = .20$ is 1.41 SE's away from the mean which is not considered unusual c)

$1/\cancel{\sqrt{.33}} \cancel{= \sqrt{3}} = 1/\sqrt{3} = .577$

False. Tripling the sample size will reduce the standard error by $1/\sqrt{3}$ (.577).

6.8)

a) Margin of error = $2 * \sqrt{\frac{p(1-p)}{n}}$

$1.96 \times \sqrt{\frac{.66(1-.66)}{1,018}} = .0291$

The margin of error is 2.91% which rounds up to 3%.

b)

~~point estimate~~ $\hat{p} \pm 2^* SE_{\hat{p}}$

$.66 \pm 1.96 \left(\sqrt{\frac{.66(1-.66)}{1018}}\right) = .66 \pm .0291$

(.6309, .6891)

The confidence interval, based on a 95% confidence level, is between 63.09% and 68.91%. therefore the poll does not provide convincing evidence that more than 70% of drivers should be required to take a road test after ~~t~~urning 65.

Conditions:

Independence: Less than 10% of the population

Success/Failure condition met

6.12)

~~A=24~~ $n = 1250$

$\hat{p} = .48$

<u>Conditions</u>

Indepence: Less than 10% of the population size, assume random sampling

Success/failure: At least 10 success + 10 failures

a) ~~48% is a sample statistic because~~

48% is a population parameter because we have a sample of 1250 respondents and we can measure this 48% from the sample

b)

$\hat{p} \pm z^* SE_{\hat{p}}$

$.48 \pm 1.96 \times \sqrt{\frac{.48(1-.48)}{1250}}$

$= (.452, .508)$

We are 95% confidence that approximatley ~~45%~~ 45.2% to 50.8% of U.S. Residents believe that marijuana should be made legal.

c)

$n(1-p) = 1250(1-.48) = 650$

$np = 1250(.48) = 600$

$\left.\begin{matrix} 650 \\ 600 \end{matrix}\right\}$ success/failure condition is met

$n \leq 10\%$ of the population as well as $n \geq 30$

The distribution is normal

a) With the confidence interval we calculated, (.452, .508), it is not justified to say that a majority of voters want to legalize marijuana.

(6.16)

$n = 331$

$\hat{p} = .48$

H_0: ~~p=.48~~ ~~p=.50~~ $p = .48$

H_A: ~~p=.48~~ $p > \cancel{.50}\ \cancel{.48}\ .50$

Independence: It is a random sample and the sample size is less than 10% of the population.

$np = 159$, $n(1-p) = 172$ → the sucess/failure condition is also met.

we can continue with the hypothesis test

~~Z statistic:~~

$$\frac{\text{point estimate} - \text{null}}{SE} = \frac{.50 - .48}{\sqrt{\frac{.48(1-.48)}{331}}} = \frac{.50 - .48}{.0275} = .7$$

$Z = .727 \rightarrow$ p-value $= .767$

Since the p-value is greater than $\alpha = .05$, we fail to reject the null hypothesis that a minority of Americans don't go to college because they cannot afford it.

b)

$\hat{p} \pm z^{*}SE$

assume $\alpha = .05$

$.48 \pm 1.96(.0275) =$

$(.426, .534)$

I would expect the confidence interval to include .5.

6.20)

margin of error $= z^{*}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

$1.96\sqrt{\frac{.48(1-.48)}{n}} = .02$

$n = 2397$

We would need to survey 2,397 Americans in order to limit the margin of error to 2%.

6.26) college grads

$n = 1{,}099$; $\hat{p} = 33\%$ watch daily show

high-school

$n = 1{,}110$; $\hat{p} = 22\%$ watch the daily show

95% confidence ($\alpha = .05$) has interval of the difference at $(.07, .15)$

$\longrightarrow$

a) TRue. Since we are calculating the difference of ($p_{college} - p_{hs}$) then the interval describes the difference as 7% - 15% less of highschool graduates watch the daily show than college grads.

b) False. See explanation in A

c) False. The 95% confident interval means that we expect 95% of the population to fall between these intervals.

d) False. The range would be narrow thus causing us to be less confident

e) ~~True~~ False. ~~It would~~ The interval cannot be negative.

6.28)

$n = 11,545$; $\hat{p} = .080$
$n = 4,691$; $\hat{p} = .088$

Independence: less than 10% of the population
sucess/failure satisfied

$$SE = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}} = \frac{.08(1-.08)}{11545} + \frac{.088(1-.088)}{4691}$$

$= .0048$

$.0048 \times 1.96 \rightarrow (-.0174, .0014)$

we are 95% confident that the proportions of Califorians who are sleep deprived ~~are~~ is 1.74% less to .14% more than Orgenians.

6.30a)

$H_0 = p_{cali} - p_{oregon} = 0$
$H_A =$ " $\neq 0$

Independece: Random sample, represents less than 10% of the population
Success/failure condition is satisfied

$$SE = \sqrt{\frac{.082(1-.082)}{11545} + \frac{.082(1-.082)}{4691}} = .00475$$

$$\frac{-.008 - 0}{.00475} = -1.684 \rightarrow \text{p-value of } .0465$$

$.0465 \cdot 2 = .093$
2 tailed

Since the p-value of .093 is greater than $\alpha = .05$, we would fail to reject the null

b) A type II error could have been made, which is when we fail to reject the null when the alternative is actually true

6.36)

a)

		yes	NO	total
Employment	yes	717	47057	47,774
	no	146	5709	5,855
	total	863	52766	

b)

H_0: $p_{diabetes} = p_{no\ diabetes}$

H_A: $p_{diabetes} \neq p_{no\ diabetes}$

c) ~~state~~ the p-value is statistically significant because it is small

6.44)

H_0: Barking deer ~~are~~ will forage in each habitat
H_A: Barking deer only forage in certain habitats

b) Chi-square test

c) Independence ✓
Sample size ✓

(.048)(426) = 20
(.147)(426) = 63
(.396)(426) = 169
(.409)(426) = 174

d) $X^2 = 276.8$
df = 3

Since the p-value is much less than .001, we reject the null hypothesis

6.48)

a) Chi-squared test for independence

b) H_0: ~~Fee~~ Depression is not caused by coffee intake

H_A: Depression is related to coffee intake

c) $\frac{2607}{50739} = .051 \rightarrow$ women who suffer from depression

$\frac{48132}{50739} = .949 \rightarrow$ women who don't suffer from depression

d)

$$\frac{2607 - 6617}{50739} = 340$$

$$\frac{(373-340)^2}{340} = 3.2$$

e)

$df = (R-1)(C-1) = (4)(1) = 4$

$\rightarrow$ p-value of less than .0001

f) Since our p-value is so small, we reject the null hypothesis.

6.52)

I + IV

6.54)

$H_0: p = .31$

$H_A: p > .31$

b) does not meet sucess/failure condition

6.56) $H_0: p_{treatment} = p_{control}$

$H_A: p_{treatment} > p_{control}$

b) $\frac{10}{34} - \frac{4}{16} = .0441$

c) The p-value is between .048 + .26. under a significance level of .05, we fail to reject the null.