Unit 5 Review – Short Response Answers

10.

a) Randomization condition: The 325 male students are probably representative of all males.
10% condition: 325 male students are less than 10% of the population of males.
Success/Failure condition: $np = (325)(0.08) = 26$ and $nq = (325)(0.92) = 299$ are both greater than 10, so the sample is large enough.

Since the conditions have been satisfied, a Normal model can be used to model the sampling distribution of the proportion of colorblind men among 325 students.

b) $\mu_{\hat{p}} = p = 0.08$

$$\sigma(\hat{p}) = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(0.08)(0.92)}{325}} \approx 0.015$$

11.

(a) $\mu_{\bar{x}} = 12$ and $\sigma_{\bar{x}} = \dfrac{0.4}{\sqrt{50}} \approx 0.0566$. **(b)** Since $n = 50$, which is greater than 30, we can use the Normal probability distribution. $P(\bar{x} < 11.9) = P\left(z < \dfrac{11.9 - 12}{0.0566}\right) = P(z < -1.77) = 0.0384$

12.

(a) Randomization condition: Gallup randomly selected 537 American adults.
10% condition: 537 results is less than 10% of all American adults.
Success/Failure condition: $n\hat{p} = (537)(0.47) = 252$ and $n\hat{q} = (537)(0.53) = 285$ are both greater than 10, so the sample is large enough.

Since the conditions are met, we can use a one-proportion z-interval to estimate the percentage of American adults who favor the death penalty.

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}\hat{q}}{n}} = (0.47) \pm 1.960\sqrt{\frac{(0.47)(0.53)}{537}} = (42.7\%, 51.1\%)$$

We are 95% confident that between 42.7% and 51.1% of Americans favor the death penalty.

b) Since the interval extends above 50%, it is plausible that the death penalty still has majority support.

c)

$$ME = z^* \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

$$0.02 = 2.326\sqrt{\frac{(0.50)(0.50)}{n}}$$

$$n = \frac{(2.326)^2(0.50)(0.50)}{(0.02)^2}$$

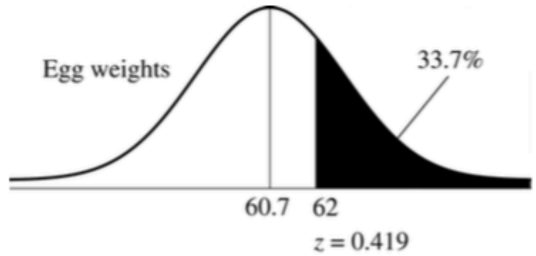
$$n \approx 3382 \text{ people}$$

We do not know the true proportion of American adults in favor of the death penalty, so use $\hat{p} = \hat{q} = 0.50$, for the most cautious estimate. In order to determine the proportion of American adults in favor of the death penalty to within 2% with 98% confidence, we would have to sample at least 3382 people.

13.

Eggs.

a) According to the Normal model, approximately 33.7% of these eggs weigh more than 62 grams.



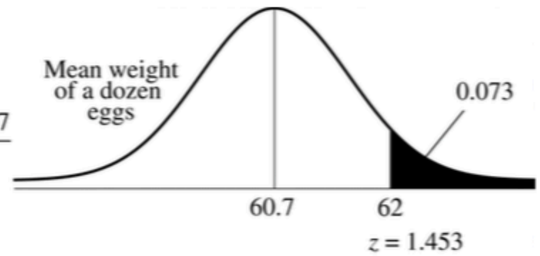
b) **Randomization condition:** The dozen eggs are selected randomly.
10% condition: The dozen eggs are less than 10% of all eggs.

The mean egg weight is $\mu = 60.7$ grams, with standard deviation $\sigma = 3.1$ grams. Since the distribution of egg weights is Normal, we can model the sampling distribution of the mean egg weight of a dozen eggs with a Normal model, with $\mu_{\bar{y}} = 60.7$ grams and standard deviation $\sigma(\bar{y}) = \frac{3.1}{\sqrt{12}} \approx 0.895$ grams.

According to the Normal model, the probability that a randomly selected dozen eggs have a mean greater than 62 grams is approximately 0.073.

$$z = \frac{62 - 60.7}{\frac{3.1}{\sqrt{12}}}$$

$z \approx 1.453$



14.

Twins.

H$_0$: The proportion of preterm twin births in 1990 is the same as the proportion of preterm twin births in 2000. $(p_{1990} = p_{2000}$ or $p_{1990} - p_{2000} = 0)$

H$_A$: The proportion of preterm twin births in 1990 is the less than the proportion of preterm twin births in 2000. $(p_{1990} < p_{2000}$ or $p_{1990} - p_{2000} < 0)$

Randomization condition: Assume that these births are representative of all twin births.
10% condition: 43 and 48 are both less than 10% of all twin births.
Independent samples condition: The samples are from different years, so they are unlikely to be related.
Success/Failure condition: $n\hat{p}$ (1990) = 20, $n\hat{q}$ (1990) = 23, $n\hat{p}$ (2000) = 26, and $n\hat{q}$ (2000) = 22 are all greater than 10, so both samples are large enough.

Since the conditions have been satisfied, we will perform a two-proportion z-test. We will model the sampling distribution of the difference in proportion with a Normal model with mean 0 and standard deviation estimated by:

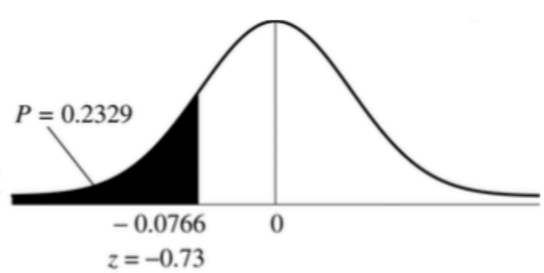
$$SE_{\text{pooled}}(\hat{p}_{1990} - \hat{p}_{2000}) = \sqrt{\frac{\hat{p}_{\text{pooled}}\hat{q}_{\text{pooled}}}{n_{1900}} + \frac{\hat{p}_{\text{pooled}}\hat{q}_{\text{pooled}}}{n_{2000}}} = \sqrt{\frac{\left(\frac{46}{91}\right)\left(\frac{45}{91}\right)}{43} + \frac{\left(\frac{46}{91}\right)\left(\frac{45}{91}\right)}{48}} \approx 0.1050.$$

The observed difference between the proportions is:
0.4651 – 0.5417 = – 0.0766

Since the P-value = 0.2329 is high, we fail to reject the null hypothesis. There is no evidence of an increase in the proportion of preterm twin births from 1990 to 2000, at least not at this large city hospital.

$$z = \frac{-0.0766 - 0}{0.1050}$$

$z \approx -0.73$



15.

Method 2: Use CLT

$N\left(30, \frac{2}{\sqrt{4}}\right)$ $\quad P\left(\bar{x} > \frac{126}{4}\right) = P(\bar{x} > 31.5) = P\left(z > \frac{31.5-30}{1}\right)$

Pumpkin pie A can of pumpkin pie mix contains a mean of 30 ounces and a standard deviation of 2 ounces. The contents of the cans are normally distributed. What is the probability that four randomly selected cans of pumpkin pie mix contain a total of more than 126 ounces?

$= P(z > 1.5)$

$= 0.067$

Method 1: P = one can of pumpkin pie mix

T = four cans of pumpkin pie mix

we are told that the contents of the cans are normally distributed, and can assume that the content amounts are independent from can to can.

$E(T) = E(P_1 + P_2 + P_3 + P_4) = E(P_1) + E(P_2) + E(P_3) + E(P_4)$

$= 30 + 30 + 30 + 30 = 120$ ounces

Since the content amounts are independent,

$Var(T) = Var(P_1 + P_2 + P_3 + P_4) = Var(P_1) + Var(P_2) + Var(P_3) + Var(P_4) = (2^2)(4) = 16$

$\sqrt{Var(T)} = \sqrt{16} = 4$ ounces

model T w/ $N(120, 4)$

$z = \frac{126-120}{4} = 1.5$

$P(T > 126) = P(z > 1.5) = 0.067$

there is a 6.7% chance that 4 randomly selected cans of pumpkin pie mix contain more than 126 ounces.

16.

$\hat{p} = 0.58$

$n = 1150$

$ME = 0.03$

$0.03 = z^* \sqrt{\frac{(0.58)(0.42)}{1150}}$

$z = 2.06$

→ Confidence level is 96%

17. Depression

If birth weigh was not a risk factor for susceptibility to depression, an observed difference in incidence of depression this large (or larger) would occur in only 2.48% of such samples.

P-value = P(observed statistic value | Ho)

18. Truckers

a. Based on the results of this inspection station, construct and interpret a 95% confidence interval for the proportion of truck drivers that have driven too many hours in a day.

Conditions: Random: we assume that truck's driving times do not influence other trucker's driving times.

10%: This sample of 348 truckers is less than 10% of all truckers.

S/F: $n\hat{p} = 348\left(\frac{49}{348}\right) = 49$

$n\hat{q} = (348)\left(\frac{299}{348}\right) = 299 \quad \geq 10$

we can use a Normal model for the sampling distribution of the proportion

b. Explain the meaning of "95% confidence" in part A.

If we repeated the sampling and created many confidence intervals many times we would expect about 95% of those intervals to contain the actual proportion of truck drivers that have driven too many hours in a day.

$n = 348$

$\hat{p} = 0.14$

$SE(\hat{p}) = \sqrt{\frac{\hat{p}\hat{q}}{n}} = \sqrt{\frac{(0.14)(0.86)}{348}} = 0.0186$

$\hat{p} \pm z^* \times SE(\hat{p})$

$0.14 \pm 1.96\,(0.0186) = 0.14 \pm 0.0365$

$(0.1035,\ 0.1765)$

we are 95% confident that between 10.4% and 17.7% of truck drivers have driven too many hours in a day.