Quiz 10.1A AP Statistics Name:

1. You have two large bins of several thousand plastic beads. Bin A is 60% red beads; Bin B is 48% red beads. Suppose you take a random sample of 80 beads from each bin and calculate $\hat{p}_A$ = the proportion of red beads in the sample from Bin A and $\hat{p}_B$ = the proportion of red beads in the sample from Bin B.

(a) Describe the sampling distribution of $\hat{p}_A - \hat{p}_B$.

(b) What is the probability that the proportion of red beads you select from Bin B is higher than the proportion of red beads you select from Bin A?

©BFW Publishers The Practice of Statistics for AP*, 5/e

2. Just before the presidential election in November 2008, a local newspaper conducted a poll of residents of a medium-sized city and found that 120 out of a simple random sample of 250 men intended to vote for Barack Obama and 132 out of an SRS of 240 women intended to vote for Obama.

(a) Is this convincing evidence that there was a gender difference in Obama's support in this city? Support your conclusion with a test of significance, using $\alpha = 0.05$.

(b) Construct and interpret a 95% confidence interval for the difference in proportion of women and men who supported Obama in this city.

Quiz 10.1B AP Statistics Name:

1. According to the manufacturer, 20% of plain M&M's are orange, and 23% of peanut M&M's are orange. Suppose you were able to take a simple random sample of 240 of each candy type. Let $\hat{p}_1$ = the sample proportion of plain M&M's that are orange and $\hat{p}_2$ = the sample proportion of peanut M&M's that are orange.

(a) Describe the sampling distribution of $\hat{p}_1 - \hat{p}_2$.

(b) What is the probability that you select a higher proportion of plain orange M&M's than peanut orange M&M's?

©BFW Publishers

2. A state policeman has a pet theory that people who drive red cars are more likely to drive too fast. On his day off, he borrows one of the department's radar guns, parks his car in a rest area, and measures the proportion of red cars and non-red cars that are driving too fast. (He decides ahead of time to define "driving too fast" as exceeding the speed limit by more than 5 miles per hour). To produce a random sample, he rolls a die and only includes a car in his sample if he rolls a 5 or a 6. He finds that 18 of 28 red cars are driving too fast, and 75 of 205 other cars are driving too fast.

(a) Is this convincing evidence that people who drive red cars are more likely to drive too fast, as the policemen has defined it? Support your conclusion with a test of significance, using $\alpha = 0.05$.

(b) Construct and interpret a 95% confidence interval for the difference in proportion of red cars that drove too fast and other cars that drive too fast.

Quiz 10.1C AP Statistics Name:

1. A study of "adverse symptoms" in users of over-the-counter pain relief medications assigned subjects at random to one of two common pain relievers: acetaminophen and ibuprofen. In all, 650 subjects took acetaminophen, and 44 experienced some adverse symptom. Of the 347 subjects who took ibuprofen, 49 had an adverse symptom.

(a) Does the data provide convincing evidence that the two pain relievers differ in the proportion of people who experience an adverse symptom? Support your conclusion with a test of significance. Use $\alpha = 0.05$.

©BFW Publishers The Practice of Statistics for AP*, 5/e

(b) Find the margin of error for a 95% confidence interval for the difference in the proportions of people who experience adverse reactions to these two mediations. <u>You need not carry out all the steps in constructing a confidence interval</u>.

(c) In parts (a) and (b) above, you should have used two different formulas for estimating the standard deviation of the randomization distribution of $\hat{p}_1 - \hat{p}_2$. Explain why it makes sense to do this.

Chapter 10 Solutions

Quiz 10.1A

1. (a) $\mu_{\hat{p}_A-\hat{p}_B} = p_A - p_B = 0.60 - 0.48 = 0.12$;

$\sigma_{\hat{p}_A-\hat{p}_B} = \sqrt{\frac{(p_A)(1-p_A)}{n_A}+\frac{(p_B)(1-p_B)}{n_B}} = \sqrt{\frac{(0.6)(0.4)}{80}+\frac{(0.48)(0.52)}{80}} = 0.078$;

Since $n_A p_A = 48,\ n_A(1-p_A) = 32,\ n_B p_B = 38.4,\ n_B(1-p_B) = 41.6,$ all of which are at least 10, the shape of the distribution is approximately Normal.

(b) $P(\hat{p}_A - \hat{p}_B < 0) = P\left(z < \frac{0-0.12}{0.078}\right) = P(z < -1.54) = 0.0618$

2. (a) State: We wish to test $H_0: p_M - p_F = 0$ *versus* $H_a: p_M - p_F \neq 0$, where p_M and p_F are the proportion of male and female voters in this city, respectively, who supported Obama. We will use a significance level of $\alpha = 0.05$. Plan: The procedure is a two-sample z-test for the difference of proportions. Conditions: *Random:* The problem states that two SRSs were taken. *10%:* It's safe to assume that there are more than $10 \times 250 = 2500$ males voters and $10 \times 240 = 2400$ female voters in a "medium-sized city." *Large counts:* The number of successes and failures in the two groups are 120, 130, 132, and 108—all of which are at least 10. Do: $\hat{p}_M = \frac{120}{250} = 0.48,\ \hat{p}_F = \frac{132}{240} = 0.55$, and $\hat{p}_C = \frac{120+132}{250+240} = 0.514$, so

$z = \frac{(0.48-0.55)-0}{\sqrt{\frac{(0.514)(0.486)}{250}+\frac{(0.514)(0.486)}{240}}} = \frac{-0.07}{0.0452} = -1.55$; Two-tailed P-value = (2)(0.0606) = 0.1212. Conclude: A P-value of 0.1212 is greater than $\alpha = 0.05$, so we cannot reject H_0. We do not have sufficient evidence to conclude that the there is a difference in the proportion of males and females who supported Obama.

(b) State: We wish to estimate, with 95% confidence, the difference $p_M - p_F$, as defined in part (a).

Plan: We should use a 2-sample z-interval for $p_M - p_F$. The conditions were addressed in part (a).

Do: The critical z for 95% confidence is 1.96, so the interval is

$(0.48-0.55) \pm 1.96\left(\sqrt{\frac{(0.48)(0.52)}{250}+\frac{(0.55)(0.45)}{240}}\right) = -0.07 \pm 0.088$, or $(-0.158,\ 0.018)$.

Conclude: We are 95% confident that the interval from – 0.158 to 0.018 captures the true difference in proportion of male and female voters supporting Obama in this city.

Quiz 10.1B

1. (a) $\mu_{\hat{p}_1-\hat{p}_2} = p_1 - p_2 = 0.20 - 0.23 = -0.03$;

$$\sigma_{\hat{p}_1-\hat{p}_2} = \sqrt{\frac{(p_1)(1-p_1)}{n_1} + \frac{(p_2)(1-p_2)}{n_2}} = \sqrt{\frac{(0.2)(0.8)}{240} + \frac{(0.23)(0.77)}{240}} = 0.037;$$

Since $n_1p_1 = 48,\ n_1(1-p_2) = 192,\ n_2p_2 = 55.2,\ n_2(1-p_2) = 184.8,$ all of which are at least 10, the shape of the distribution is approximately Normal.

(b) $P(\hat{p}_1 - \hat{p}_2 > 0) = P\left(z > \frac{0-(-0.03)}{0.037}\right) = P(z > 0.81) = 0.2090$ **2.** (a) State: We wish to test

$H_0 : p_R - p_O = 0;\ H_a : p_R - p_O > 0$, where p_R and p_O are the proportion of red cars and other car, respectively, who are driving too fast. We will use a significance level of $\alpha = 0.05$.

Plan: The procedure is a two-sample z-test for the difference of proportions. Conditions: *Random:* The policemen chose cars randomly by rolling a die. *10%*: We can safely assume that the number of cars driving past the rest area is essentially infinite, so the 10% restriction does not apply. *Large counts:* The number of successes and failures in the two groups are 18, 10, 75, and 130—all of which are at least 10.

Do: $\hat{p}_R = \frac{18}{28} = 0.64,\ \hat{p}_O = \frac{75}{205} = 0.37$, and $\hat{p}_C = \frac{18+75}{28+205} = 0.40$, so

$$z = \frac{(0.64-0.37)-0}{\sqrt{\frac{(0.4)(0.6)}{28} + \frac{(0.4)(0.6)}{205}}} = \frac{0.27}{0.099} = 2.73;$$

One-tailed P-value = 1 – 0.9968 = 0.0032. [A 2-proportion z-test on the calculator yields $z = 2.807$ and a P-value of 0.0025] Conclude: A P-value of 0.0032 is less than $\alpha = 0.05$, so we reject H_0. We have sufficient evidence to conclude that the proportion of red cars that drive too fast on this highway is greater than the proportion of non-red cars that drive too fast.

(b) State: We wish to estimate, with 95% confidence, the difference $p_R - p_O$, as defined in part (a).

Plan: We should use a 2-sample z-interval for $p_M - p_F$. The conditions were addressed in part (a).

Do: The critical z for 95% confidence is 1.96, so the interval is

$$(0.64-0.37) \pm 1.96\left(\sqrt{\frac{(0.64)(0.36)}{28} + \frac{(0.37)(0.63)}{205}}\right) = 0.27 \pm 0.190, \text{ or } (0.080,\ 0.460).$$

Conclude: We are 95% confident that the interval from 0.070 to 0.470 captures the true difference in the proportion of red cars and non-red cars that drive too fast on this highway.

Quiz 10.1C

1. (a) State: We are testing the hypotheses $H_0: p_A - p_I = 0$ *versus* $H_a: p_A - p_I \neq 0$, where p_A and p_I are the proportions of people who experience adverse reactions to acetaminophen and ibuprofen, respectively, among all patients (similar to those in this experiment) who might receive these drugs. We will use a significance level of $\alpha = 0.05$. Plan: The procedure is a two-sample z-test for the difference of proportions. Conditions: *Random:* Problem states that subjects were assigned at random. *10%:* Since no sampling took place, the 10% condition does not apply. *Large counts:* The number of successes (adverse reactions) and failures in the two groups are 44, 606, 49, 298—all of which are at least 10. *Independent*: Random assignment means we can view these groups as independent, and adverse reactions in one person should not influence reactions in another person.

Do: $\hat{p}_A = \frac{44}{650} = 0.068$, $\hat{p}_I = \frac{49}{347} = 0.141$, $\hat{p}_C = \frac{44+49}{650+347} = 0.093$, so

$$z = \frac{(0.068 - 0.141) - 0}{\sqrt{\frac{(0.093)(0.907)}{650} + \frac{(0.093)(0.907)}{347}}} = -3.78.$$

This test statistic is so large that by Table A the P-value is very close to 0 (by calculator, P-value = 0.00014). Conclude: A P-value of 0.00014 is much less than $\alpha = 0.05$, so we reject H_0. We have sufficient evidence to conclude that the proportions of people who experience adverse reactions to acetaminophen and ibuprofen are different. (b) The critical z for 95% confidence is 1.96, so the margin of error is $1.96\left(\sqrt{\frac{(0.068)(0.932)}{650} + \frac{(0.141)(0.859)}{347}}\right) = 0.041$.

(c) When we conduct a significance test, we assume that the null hypothesis is true. In this case, $H_0: p_A - p_I = 0$, so we are assuming that $p_A = p_I$. This means we are assuming the two samples are independent samples that estimate the same quantity. It thus makes sense to combine this information into single pooled estimate $\hat{p}_C = \frac{44+49}{650+347} = 0.093$ when estimating $\sigma_{\hat{p}_A - \hat{p}_I}$. When constructing a confidence interval, we are not making the assumption that $p_A = p_I$, so we should not pool the data.