Quiz 11.1A AP Statistics Name: ______________

1. A grocery store sells four different sizes of a popular brand of corn flakes. For the past few years the proportion of boxes they sell of each size has been quite stable: 10% Small, 15% Medium, 60% Large, and 15% Jumbo. They decide to change the pricing of the four sizes and want to see if this changes the proportion of boxes they sell of each size. To test this, a few weeks after changing the prices they take a simple random sample of 120 transactions involving corn flakes and count how many boxes of each size were sold. Here are the results.

Observed number of boxes sold for each box size

Small	Medium	Large	Jumbo
8	24	61	27

(a) We wish to carry out a test of significance to see if the distribution of sizes of cereal boxes sold has changed. State the null and alternative hypotheses for this test.

(b) Find the expected counts for each size box under the assumption that the null hypothesis is true.

Expected number of boxes sold for each box size

Small	Medium	Large	Jumbo

(c) Discuss whether the conditions for this test have been met.

(d) Find the value of the test statistic and the P-value of the test, and make the appropriate conclusion. Use $\alpha = 0.05$.

(e) Based on your answer to (d), which error is it possible that you have made, Type I or Type II? Describe that error in the context of the problem.

(f) Use the components of the chi-square statistic to perform a follow-up analysis on the impact of the new prices on the sales of different sizes of cereal boxes.

Quiz 11.1B AP Statistics Name:

1. In some countries, people believe that blood type has a strong impact on personality. For example, Type B blood is thought to be associated with passion and creativity. A statistics student at a large U.S. university decides to test this theory. Reasoning that people involved in the arts should be passionate and creative, she takes a simple random sample of students majoring or minoring in arts at her university and asks them for their blood type. Here are her results:

Observed number of performing arts majors with each blood type

Type A	Type B	Type AB	Type 0	Total
50	23	10	67	150

Assume the distribution of blood type among all U.S. residents is as follows: Type A: 42%; Type B: 10%; Type AB: 4%; Type O: 44%.

(a) The student wants to carry out a test of significance to see if the distribution of blood types among arts majors or minors is different from the U.S. distribution. State the null and alternative hypotheses for this test.

(b) Find the expected counts for each blood type under the assumption that the null hypothesis is true.

Expected number of performing arts majors with each blood type

Type A	Type B	Type AB	Type 0

(c) Discuss whether the conditions for this test have been met.

(d) Find the value of the test statistic and the *P*-value of the test, and make the appropriate conclusion. Use $\alpha = 0.05$.

(e) Based on your answer to (d), which error is it possible that you have made, Type I or Type II? Describe that error in the context of the problem.

(f) Use the components of the chi-square statistic to perform a follow-up analysis on the impact of the new prices on the sales of different sizes of cereal boxes.

Quiz 11.1C AP Statistics Name:

1. The human resources department of a very large corporation (more than 20,000 employees) suspects that people are more likely to call in sick on Monday or Friday, so they can take a long weekend. They took a random sample of 850 sick-day reports from the past few years and determined the day of the week for each report. Here are the results:

Day	Monday	Tuesday	Wednesday	Thursday	Friday
# of sick days	192	151	148	152	207

(a) Do the data provide convincing evidence that sick day calls are not evenly distribution through the week? Justify your answer with appropriate statistical evidence.

(b) Based on your answer to (a), which error is it possible that you have made, Type I or Type II? Describe that error in the context of the problem.

(c) Use components of the chi-square statistic to perform a follow-up analysis on whether the data supports the claim that employees may be using sick days to take long weekends.

Chapter 11 Solutions

Quiz 11.1A

1. (a) H_0: The distribution of sizes of all boxes sold of this brand of cereal did not change when the prices changed. H_a: The distribution of sizes of all boxes sold of this brand of cereal changed when the prices changed. (b) Expected counts: Small: $(0.10)(120)=12$, Medium: $(0.15)(120)=18$, Large: $(0.60)(120)=72$, Jumbo: $(0.15)(120)=18$. (c) *Random*: the data come from a simple random sample of sales records. *10%:* We must assume that there were more than 1200 boxes of cereal sold. *Large counts:* All the expected counts (see part (b)) are at least 5.

(d) $\chi^2 = \frac{(8-12)^2}{12}+\frac{(24-18)^2}{18}+\frac{(61-72)^2}{72}+\frac{(27-18)^2}{18}=9.514$; df = 3; P-value 0.023 (Using Table C, 0.02< P-value < 0.025). Since the P-value is less than $\alpha = 0.05$, we can reject H_0. There is convincing evidence that the distribution of size of boxes sold changed after the prices changed. (e) It's possible that we have made a Type I error, which is concluding that the distribution of box sizes sold has changed when it has not. (f) Components of the chi-square statistic: Small: 1.33, Medium: 2.00, Large: 1.68, Jumbo: 4.5. The observed count of jumbo-sized boxes was much larger than the expected count, so it appears than an increase in sales of jumbo-sized boxes was the biggest impact of the price changes.

Quiz 11.1B

1. 1. (a) H_0: The distribution of blood types among all performing arts students in the population is the same as the distribution among all U.S. residents. H_a: The distribution of blood types among all performing arts students in the population is different from the distribution among all U.S. residents. (b) Expected counts: Type A: $(0.42)(150)=63$, Type B: $(0.10)(150)=15$, Type AB: $(0.04)(150)=6$, Type O: $(0.44)(150)=66$. (c) *Random*: the data come from a simple random sample of arts majors and minors. *10%:* We must assume that there are more than 1500 students either majoring or minoring in arts at this large university. *Large counts*: All the expected counts (see part (b)) are at least 5.

(d) $\chi^2 = \frac{(50-63)^2}{63}+\frac{(23-15)^2}{15}+\frac{(10-6)^2}{6}+\frac{(67-66)^2}{66}=9.631$; df = 3; P-value 0.022 (Using Table C, 0.02< P-value < 0.025). Since the P-value is less than $\alpha = 0.05$, we can reject H_0. There is convincing evidence that the distribution of blood type for performing arts students is different from the U.S. distribution. (e) It's possible that we have made a Type I error, which is concluding that the distribution of blood types for arts majors and minors is different from the U.S. distribution when it is not different.

(f) Components of the chi-square statistic: Type A: 2.68, Type B: 4.27, Type AB: 2.67, Type O: 0.02. Type B blood made the largest contribution to the chi-square statistic, because the observed count of arts students with Type B blood and Type AB blood were higher than expected by the null hypothesis. At the same time, the observed number of students with Type A blood was lower than expected. This supports the claim that arts students are more likely to have Type B blood (perhaps Type AB as well).

Quiz 11.1C

1. (a) <u>State</u>: We are testing the hypothesis H_o: The proportion of all the company's sick day reports is the same for all five workdays, against H_a: At least one day's proportion of all sick day reports is different from the others. [Or: $H_0 : p_m = p_t = p_w = p_h = p_f$, where each p is the proportion of all the company's sick day calls on a day of the week, and H_a: at least of value of p is not equal to the others. We will use a significance level of $\alpha = 0.05$. <u>Plan</u>: The procedure is a chi-square goodness-of-fit test. Conditions: *Random*: the data come from a simple random sample of sick day reports. *10%*: We must assume that there were more than 8500 sick-day reports in the past few years at this company. This is not an unreasonable assumption if the company has 20,000 employees. *Large counts:* All expected counts are $850 \div 5 = 170$ which is greater than 5 in each case.

<u>Do</u>: $\chi^2 = \frac{(192-170)^2}{170} + \frac{(151-170)^2}{170} + \frac{(148-170)^2}{170} + \frac{(152-170)^2}{170} + \frac{(207-170)^2}{170} = 17.78$,

$df = 4$; P-value = 0.0014. <u>Conclude</u>: Since the P-value is much smaller than $\alpha = 0.05$, we can reject H_0. There is convincing evidence that the proportion of all sick-day reports is not the same for all five workdays. (b) It's possible that we have made a Type I error, which is concluding that the sick day calls are not evenly distributed through the week when they are. (c) Individual components of the chi-square statistic: Monday: 2.85; Tuesday: 2.12; Wednesday: 2.85; Thursday: 1.91; Friday: 8.05. Observed sick-day reports on Friday were much larger than expected, contributing nearly half the value of the chi-square statistic. This supports the claim that more people call in sick on Monday or Friday.