STATISTICS 101 - Homework 8 Answers

1. (20 pts) Read pages 341 - 349 and starting on page 350 do exercise 28 (omit part c)

   - (15 pts) Problem 28

     (a) (4 pts) We need to find $P(X < 10)$. Using the normal model, this is

     $$P(X < 10) = P\left(Z < \frac{10 - 10.2}{0.12}\right) = P(Z < -1.6666666)$$

     $z = -1.66666$ in the table is 0.04779. Approximately 4.8% of bags will be underweight.

     (b) (4 pts) We need to find $P(\overline{X} < 10)$ for a sample of size n=3.

     $$P(\overline{X} < 10) = P\left(Z < \frac{10 - 10.2}{0.12/\sqrt{3}}\right) = P(Z < -2.8867..)$$

     $z = -2.8867..$ in the table is 0.001946... If μ is really 10.2 ounces, the probablility of observing an $\overline{X}$ less than 10 ounces in a sample of size 3 is .19

     (c) (4pts) We need to find $P(\overline{X} < 10)$ for a sample of size n=24.

     $$P(\overline{X} < 10) = P\left(Z < \frac{10 - 10.2}{0.12/\sqrt{24}}\right) = P(Z < -8.164965..)$$

     $z = -8.164965.$ in the table is essentially zero. If μ is really 10.2 ounces, the probablility of observing an $\overline{X}$ less than 10 ounces in a sample of size 24 is essentially zero

2. (53 pts) Read pages 430 - 447 and starting on page 447 do exercises 8, 10, 22, 24.

   - (7 pts) Problem 8

     (a) (3 pts) We are 95% confident that the population mean age at which babies begin to crawl is between 29.2 and 31.8 weeks.

     (b) (2 pts) The width of the interval is 31.8 - 29.2 = 2.6 weeks. The margin of error is half of the width or 1.3 weeks.

     (c) (2 pts) The margin of error for a 90% confidence interval would have been smaller. Less confidence means a smaller $t^\star$ value and therefore a smaller margin of error.

   - (14 pts) Problem 10

     (a) (2 pts) The 10% condition is satisfied. 44 days is definitely less than 10% of all workdays. The nearly normal condition is satisfied since the sample size of 44 is fairly large. However, checking the data if available would be a good step.

     (b) (4 pts) The 90% CI is given by

     $$\begin{array}{rcl} \bar{y} & \pm & t^\star_{43}\frac{s}{\sqrt{n}} \\ 126 & \pm & 1.684\frac{15}{\sqrt{44}} \\ 126 & \pm & 3.81 \\ (122.19 & , & 129.81) \end{array}$$

(c) (3 pts) We are 90% confident the mean daily fees from the parking garage will be between \$122.19 and \$129.81.

(d) (3 pts) 90% of all random samples of 44 workdays will produce confidence intervals that contain the true mean daily fees from this parking garage.

(e) (2 pts) The confidence interval is completely below \$130. Because of this, there is some evidence that the consultant overestimated the mean daily income from this garage.

- (18 pts) Problem 22

  (a) (2 pts) $H_o : \mu = 26,\ H_A : \mu < 26$

  (b) (2 pts) Random sample: The trips were a random sample from the population. Nearly Normal: The sample size is large (n=50) so we can apply the t distribution without worrying about the distribution of the population.

  (c) (2 pts) The sampling distribution will be a t distribution with $n - 1 = 49$ degrees of freedom.

  (d) (6 pts) The test statistic is

$$t = \frac{\overline{y} - \mu_o}{\frac{s}{\sqrt{n}}} = \frac{25.02 - 26}{\frac{4.83}{\sqrt{50}}} = -1.43$$

  Since the alternative hypothesis is less than, the p-value is $P(t_{49} < -1.43) = P(t_{49} > 1.43)$. Using d.f. = 45, this probability is between 0.05 and 0.10.

  (e) (3 pts) The p-value is the probability of obtaining a sample mean of 25.02 or less if the population mean is really 26mpg is between 5% and 10%.

  (f) (3 pts) Since the p-value is greater than $\alpha = 0.05$, we will not reject the null hypothesis. There is little evidence to suggest that the mean mileage of cars in the fleet is less than 26 mpg.

- (14 pts) Problem 24

  (a) (2 pts) Random sample: There is no reason to believe the 6 bags will be any different than a random sample from the population. Nearly Normal: With only 6 bags, this condition is very difficult to check. There is no reason to think that the filling process would produce a skewed distribution for the weight of the potato chips. In addition, there are no large or small outliers out of the 6 bags.

  (b) (4 pts) $\overline{y} = 28.98$ and $s = 0.36$ grams.

  (c) (6 pts) For 95% confidence and d.f. = n-1 = 5, $t^\star = 2.571$. The 95% confidence interval for the population mean weight of the contents of the potato chip bags is

$$\begin{aligned} \overline{y} &\pm t^\star\left(\frac{s}{\sqrt{n}}\right) \\ 28.98 &\pm 2.571\left(\frac{0.36}{\sqrt{6}}\right) \\ 28.98 &\pm 0.38 \\ (28.60 &,\ 29.36) \end{aligned}$$

  We are 95% confident that the mean weight of the contents of Doritos bags is between 28.60 and 29.36 grams.

(d) (2 pts) The stated weight of 28.3 grams is outside the confidence interval. This is evidence that the company is attempting to fill the bags with more than the stated amount, on average.

3. (24 pts) Advertisements claim that Happy's Joes puts more pepperoni on its pizzas than competitors. In order to investigate this claim, 9 medium pizzas from Happy Joe's and 9 medium pizzas from Pizza Hut are purchased at random times over a month long period. The number of slices of pepperoni on each pizza is then counted. Below are the data.

(1) Happy Joe's	43	51	53	49	39	45	56	44	52	$\overline{X}_1 = 48$	$s_1 = 5.55$
(2) Pizza Hut	47	33	37	43	28	36	35	35	39	$\overline{X}_2 = 37$	$s_2 = 5.55$

(a) (4 pts) Construct a back-to-back stem and leaf plot to compare the two pizza parlors. Describe what you see.

Back-to-back stem and leaf plot

Happy Joe's		Pizza Hut
	2	8
9	3	3 5 5 6 7 9
9 5 4 3	4	3 7
6 3 2 1	5	

It appears the number of slices of pepperoni on Happy Joe's pizza is higher than on Pizza Hut pizza. The spread of the two samples is about the same.

(b) (15 pts) Does Happy Joe's have more pepperoni, on average, than Pizza Hut? Perform the appropriate statistical test of hypothesis. Be sure to comment on the truthfulness of Happy Joe's advertising, based on your test.

μ_1 = mean number of pepperoni slices on Happy Joe's pizza. μ_2 = mean number of pepperoni slices on Pizza Hut pizza.

For this test of hypothesis, the null and alternative hypotheses are

$$H_o : \mu_1 - \mu_2 = 0$$

$$H_a : \mu_1 - \mu_2 > 0$$

Assumptions: Both sets of pizzas are random samples of all pizzas produced at the two restaurants. The two samples are independent of each other. The distribution of both samples is not necessarily normal, but neither has a large skewness or large outliers.

The two sample t statistic is

$$t = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{48 - 37}{\sqrt{\frac{5.55^2}{9} + \frac{5.55^2}{9}}} = \frac{11}{2.62} = 4.198 \tag{1}$$

The t statistic has an approximate t distribution with 8 degrees of freedom.

From the table, the t critical values closest to 4.198 are 3.833 and 4.501. Therefore, the p-value is less than 0.005.

Since the p-value is less than any reasonable significance level α, we will reject the null hypothesis and conclude the mean number of slices of pepperoni on Happy's Joe's pizza is larger than the mean number of slices of pepperoni on Pizza Hut's pizza.

Yes, Happy Joe's claim is correct.

(c) (5 pts) Construct a 90% confidence interval for the difference between the mean number of slices of pepperoni per medium pizza.

The number of degrees of freedom for this problem is 8. According to the t table, the value of t* for 8 d.f. and a 90% CI is 1.860.

The 90% confidence interval is

$$(48 - 37) \pm t^{\star}\sqrt{\frac{5.55^2}{9} + \frac{5.55^2}{9}} = 11 \pm 1.860(2.62) = (6.13, 15.87)$$

We are 90% confident that the difference in the mean number of slices of pepperoni between Happy Joes and Pizza Hut is between 6.13 and 15.87 slices

4. (22 pts) A random sample of 20 students, 10 girls and 10 boys, is taken in a Midwest school district. The mean IQ score for the 10 girls is 103.6 with a standard deviation of 8.91. The mean IQ score for the 10 boys is 107 with a standard deviation of 5.58.

(a) (5 pts) Calculate a 95% confidence interval for the difference between the population mean IQ score for boys and the population mean IQ score for girls from this school district.

The number of degrees of freedom for this problem is 9. According to the t table, the value of t* for 9 d.f. and a 95% CI is 2.262.

The 95% confidence interval is

$$(107 - 103.6) \pm t^{\star}\sqrt{\frac{8.91^2}{10} + \frac{5.58^2}{10}} = 3.4 \pm 2.262(3.32) = (-4.11, 10.91)$$

We are 95% confident that the true difference in mean IQ scores between the boys and girls in this school district is between -4.11 and 10.91.

(b) (3 pts) We would like to test whether the two population mean IQ scores are different. The null and alternative hypotheses for this test are

$$H_o : \mu_1 - \mu_2 = 0$$
$$H_a : \mu_1 - \mu_2 \neq 0$$

Without calculating the two-sample t statistic, would you reject the null hypothesis of equal population means at the 5% level of significance? Explain your answer.

We would not reject the null hypothesis of equal population means at the 5% level of significance. The 95% confidence interval calculated in part (a) contains the value 0. Therefore, it is likely that the mean difference between the two population means is zero.

(c) (14 pts) Complete the test of hypothesis from part (b). Do your decision and conclusion agree with what you found in part (b)?

The null and alternative hypotheses for the test of hypothesis are

$$H_o : \mu_1 - \mu_2 = 0$$

$$H_a : \mu_1 - \mu_2 \neq 0$$

.

Assumptions: Each sample is a random sample. Both samples are independent. We are not given information about the population distribution, however IQ scores are normally distributed.

The two-sample t statistic is

$$t = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{107 - 103.6}{\sqrt{\frac{8.91^2}{10} + \frac{5.58^2}{10}}} = \frac{3.4}{3.32} = 1.02 \tag{2}$$

The t statistic has an approximate t distribution with 9 degrees of freedom.

From the table, the p-value for the test is greater than 0.20.

The p-value is greater than the significance level of $\alpha = 0.05$, so we will not reject the null hypothesis and conclude that the mean IQ scores for boys and girls in this school district are equal, which agrees with our prior knowledge based on the CI.