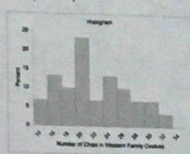
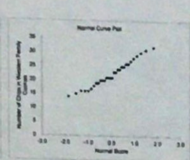
Brock Francom
A02052161

Homework More About Hypothesis Testing

1. Chips Ahoy cookies are supposed to have 1000 chocolate chips per bag. There used to be 48 cookies per bag. This would mean an average of 20.8 chocolate chips per cookie. Western Family claims that their cookies are comparable to Chips Ahoy cookies. We think that the Western Family brand has at least if not more chocolate chips per cookie than the Chips Ahoy.

   In a previous class, students took a sample of Western Family Chewy Chocolate Chip Cookies and counted the chocolate chips. Here is the data.

   Data: 21, 19, 23, 16, 20, 21, 17, 20, 14, 23, 25, 25, 15, 20, 29, 24, 16, 31, 18, 26, 21, 28, 21, 27, 32, 24, 19, 16, 30, 26

   The sample size is $n = 30$.

   Conduct a hypothesis test with $\alpha = .05$ to determine if the population mean number of chocolate chips of all Western Family Chewy Chocolate Chip Cookies is greater than 20.8. Assume the population standard deviation is $\sigma = 5$.

   (a) Specify the level of significance α.

   $\alpha = .05$

   (b) Check the conditions.

   $n = 30 \geq 30$ ✓ normal

   (c) State the null hypothesis H_0 and the alternative hypothesis H_a.

   $H_0: \mu = 20.8$

   $H_A: \mu > 20.8$

   (d) Collect the sample data and compute the value of the test statistic. The sample mean is $\overline{x} = 22.23$.

   $$Z = \frac{22.23 - 20.8}{5/\sqrt{30}} = 1.57$$

   (e) Calculate the P-value.

   $P\text{val} = 1 - \Phi(1.57) = .0582$

   (f) Reject H_0 if the P-value is less than α.

   Fail to reject H_0

   (g) Interpret your statistical results in real world terms.

   we didn't find evidence that the population mean # of Chocolate Chips ~~cook~~ is greater than 20.8.

   (h) Interpret the p-value.

   If population mean is 20.8 chips there is 5.82% Chance that we pick a sample mean ≥ 22.23

2. Let's consider the chocolate chip cookies again.

   We want to know if the mean number of chocolate chips per cookie is less than 20.8.

   We set $\alpha = .05$.

   (a) What are the hypotheses in words?

   Ho: mean # chocolate chips is 20.8

   HA: mean # Chocolate Chips is less than 20.8

   (b) What are the hypotheses in symbols?

   $H_0: \mu = 20.8$

   $H_A: \mu < 20.8$

   (c) What would be a Type I error in this situation?

   mean = 20.8, but we decide that it is less.

   (d) What is the probability of a Type I error?

   α, .05

   (e) What would be a Type II error in this situation?

   mean $\neq$ 20.8, but we decide that it is 20.8.

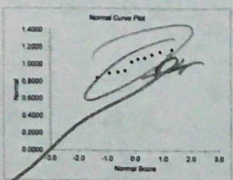
   (f) Do you think you will be more likely to reject the null hypothesis if the true mean is $\mu = 10$ or if the true mean is $\mu = 20$?

   $\mu = 10$, pvals will be much smaller.

   (g) Do you think you will be very likely to reject the null hypothesis if the true mean is $\mu = 70$?

   no!

3. I worked at KFC/TacoBell the summer before I started college. Our managers would become upset if any orders at the drive through took longer than a minute. Of course some orders were quick and some orders were longer. It is believed that the population standard deviation waiting time is $\sigma = .3$ minutes. Suppose the current manager believes that his workers are taking longer than 1 minute. He records the time on 10 random orders. The sample mean is 1.17 minutes. Obviously 1.17 is more than 1. But he wants to know if he just picked a sample that had an extra big mean or if he has evidence that the population mean (for all the orders) is bigger than 1.



(a) Conduct a hypothesis test with $\alpha = .01$ to determine if the population mean waiting time is longer than 1 minute. Make sure you follow all the steps.

Basically normal

- $\alpha = .01$
- $H_0: \mu = 1$
  $H_A: \mu > 1$
- test stat:
  $Z = \frac{1.17 - 1}{.3/\sqrt{10}}$
  $= 1.79$

- p val: $= .0367$

we didn't find evidence that the mean pop waiting time is greater than 1 minute.

(b) The p-value is .0367. Interpret this p-value.

The probability of observing a sample mean of 1.17 or greater if the population is 1 minute, is only .0367

(c) Our manager decided ahead of time to use a significance level of $\alpha = .01$. His p-value is .0367.

i. Did he reject or fail to reject the null hypothesis?

Failed to reject the null Hypothesis

ii. Is it ethical for him to change his significance level after he sees the results?

No, That Changes everything.

iii. If he had chosen a significance level of .05 instead of $\alpha = .01$ ahead of time, would he reject or fail to reject the null hypothesis?

reject H_0

(d) Suppose unbeknownst to our manager a managers at several different KFCs did the same experiment. The different p-values that they found are shown below. Which manager found the strongest evidence that the $\mu > 1$.

i. .023
ii. .564
iii. .001
iv. .00034

4. Suppose we get a p-value of essentially 0 when we test $H_0 : \mu = 33$ versus $H_A : \mu < 33$. (Software tells us the p-value was 1.95×10^{-32}). What is an appropriate conclusion?

   (a) We have proved that the population mean is 33.
   (b) We have proved that the population mean is less than 33.
   (c) We found extremely strong evidence that the population mean is 33.
   (d) We found extremely strong evidence that the population mean number is less than 33.

5. We say that our results are statistically significant if we get a ___________ p-value.

   (a) big
   (b) small

6. What does statistically significant mean?

   It is unlikely to see that big of a difference in the sample, if the H_0 is true.

7. When performing significance tests at level α, if the P-value obtained is _____ α, the null hypothesis will not be rejected.
   A) larger than
   B) smaller than
   C) None of the above.

8. When performing significance tests at level α, if the P-value obtained is _____ α, the null hypothesis will be rejected.
   A) larger than
   B) smaller than
   D) None of the above.

9. If you reject the null hypothesis when in fact the null hypothesis is true it is called __________.
   A) a Type I error
   B) a Type II error

10. If you accept the null hypothesis when in fact the alternative hypothesis is true is called __________.
   A) a Type I error.
   B) a Type II error

11. A medical researcher is working on a new treatment for a certain type of cancer. In an early trial, she tries the new treatment on several subjects. Although the survival time has increased, the results are not statistically significant, even at the 0.10 significance level. Suppose, in fact, that the new treatment does increase the mean survival time in the population of all patients with this particular type of cancer. Which of the following statements is true?
   A) A Type I error has been committed.
   B) A Type II error has been committed.
   C) No error has been committed.

12. An engineer has designed an improved light bulb. The previous design had an average lifetime of 1200 hours. Using a sample of 2000 of these new bulbs, the average lifetime of this improved light bulb is found to be 1201 hours. The null hypothesis is that the new light bulb lifetime is the same as the old light bulb lifetime. Although the difference is quite small, the effect was statistically significant at the 0.05 level. Suppose that, in fact, there is no difference between the mean lifetimes of the previous design and the new design. Which of the following statements is true?
A) A Type I error has been committed.
B) A Type II error has been committed.
C) No error has been committed.

13. Ten years ago, at a small high school in Alabama, the mean SAT Math score of all high school students who took the exam was 490 with a standard deviation of 80. This year, the SAT Math scores of a random sample of 25 students who took the exam are obtained. The mean score of these 25 students is $\overline{x} = 525$. To determine if there is evidence that the scores in the district have improved, the hypotheses $H_0 : \mu = 490$ versus $H_a : \mu > 490$ are tested at the 5% significance level. The P-value is found to be 0.014. Suppose that the average SAT Math score of all high school students at this high school is in fact equal to 505. Which of the following statements is true?
A) A Type I error has been committed.
B) A Type II error has been committed.
C) No error has been committed.

14. What is power?
Probability of correctly rejecting a false null hypothesis.

15. Higher power is a good thing. How can we increase the power and lower the probability of making an error?
Increase sample size.

16. If a statistical significance is found when performing a test of significance, then practical significance will also be found.
A) True
B) False

17. A food safety inspector is called upon to investigate a restaurant with a few customer reports of poor sanitation practices. The food safety inspector uses a hypothesis testing framework to evaluate whether regulations are not being met. If he decides the restaurant is in gross violation, its license to serve food will be revoked. H_0: Resturaunt meets food regulations
H_a: Resturaunt doesn't meet regulations
(a) Write the hypotheses in words.
(b) What is a Type 1 Error in this context? restuarant deemed unsafe when it is safe
(c) What is a Type 2 Error in this context? restuarant deemed safe when unsafe.
(d) Which error is more problematic for the restaurant owner? Why? Type 1, lose money.
(e) Which error is more problematic for the diners? Why? Type 2, unsafe.
(f) As a diner, would you prefer that the food safety inspector requires strong evidence or very strong evidence of health concerns before revoking a restaurant's license? Explain your reasoning.
Strong, I like eating and don't care as much about safety.

18. Remember and make a note of these things:
- Increasing the sample size gives you a smaller p-value.
- Increasing the sample size makes it less likely to make a Type I or Type II error.
- Increasing the sample size makes it more likely that you make the correct decision (rejecting/failing to reject the null hypothesis).
- But increasing the sample size can mean that even a very small difference between the true population mean and the value of the mean in the null hypothesis can show up as statistically significant. But that difference might be so small that it isn't practically significant (meaning no one would actually care.)

Homework T Test for Mean

Instructions: For each problem, do the following for every hypothesis test. Then do any additional parts as asked.

- Check the conditions.
- Specify the level of significance α.
- State the null hypothesis H_0 and the alternative hypothesis H_a.
- Calculate the test statistic.
- Calculate degrees of freedom.
- Find the p-value.
- Reject H_0 if the P-value is less than α.
- Interpret your statistical results in real world terms.
- **Interpret the p-value.

19. In another previous class, my students took a *different* sample of 30 Chips Ahoy Cookies. Someone, not the company, had claimed they are supposed to have an average of 33 chocolate chips per cookie.

    The sample mean is $\overline{x} = 24.83$ and the sample standard deviation is $s = 3.65$. When I saw that sample mean I was very suspicious of the claim that they have a population mean of 33 chips per cookie.

    Conduct a hypothesis test to determine if the population mean is less than 33 chocolate chips per cookie. Interpret the p-value as well. Use $\alpha = .10$.

$n=30$ ✓ normal
$\alpha=.1$
$H_0 = \mu = 33$
$H_A = \mu < 33$

test statistic:

$$t = \frac{24.83-33}{3.65/\sqrt{30}} = -12.26$$

Pval: basically 0
we found extremly strong evidence that the population mean # of chips is less than 33.

20. We want to see if a mean waiting time is greater than six minutes. A sample of 85 times gives $\overline{x} = 5.46$ and $s = 2.475$ minutes. Conduct an appropriate hypothesis test. Interpret the p-value as well.

$85 \geq 30$ ✓
$\alpha = .05$
$H_0: \mu = 6$
$H_A: \mu > 6$

test stat:

$$t = \frac{5.46-6}{2.475/\sqrt{85}} = -2.0115$$

Pval: is about .975 - .98
we didnt find evidence that the mean waiting time is greater than 6 minutes.

21. "Coopetition" is when firms that both cooperate and compete with other firms. A study looked at managers and their level of external tension (measured on a 20 point scale). They took a sample of 1532 managers and found a sample mean tension of 10.82 and a sample standard deviation of 3.04.

    (a) Review: Find and interpret a 95% confidence interval for the mean external tension for all coopetition managers. $t^* = 1.98$ $\quad 10.82 \pm 1.98\left(\frac{3.04}{\sqrt{1532}}\right) \Rightarrow (10.67, 10.97)$

    (b) Based on your confidence interval, does 10.5 points seem like a plausible value for μ?

    no. not in range

    (c) Conduct a hypothesis test with $\alpha = .05$ to see if the mean external tension for *all* coopetition managers is different from 10.5 points.

    (d) Did your answers to (b) and (c) agree?

Yes.

$1532 > 30$ ✓
$\alpha = .05$
$H_0 = \mu = 10.5$
$H_A = \mu \neq 10.5$

test stat:

$$t = \frac{10.82-10.5}{3.04/\sqrt{1532}} = 4.12$$

Pval: 0

we have evidence that the average external tension for all managers is different from 10.5

22. **Review Confidence Intervals:** A survey was done of over 25,000 adults. They found a 95% confidence interval for the population mean salary of all males with post graduate degrees is \$57,050 to \$65,631.

   (a) Which *two* of the following statements are correct?

      i. 95% of males with graduate degrees will make between \$57,050 and \$65,631.

      ii. We are 95% confident that the ***population*** mean salary of males with graduate degrees is between \$57,050 and \$65,631.

      iii. We are 95% confident that the ***sample*** mean salary of males with graduate degrees is between \$57,050 and \$65,631.

      iv. Of all the samples we could have picked, 95% of them would have resulted in a confidence interval that actually contains the ***population*** mean salary.

      v. There is a 95% probability that the ***population*** mean salary of males with graduate degrees is between \$57,050 and \$65,631.

   (b) Based on the confidence interval, do you think \$60,000 is a plausible value for the population mean?

yes. It is in the interval.

23. **Funerals**
The National Funeral Directors Association reported that the average cost for a full-service funeral in 2014 was \$7180. A sample of 36 funeral homes was taken this year. (The values are in thousands of dollars.)

Hypothesis Test: Mean vs. Hypothesized Value

7.1800	hypothesized value
6.8194	mean Data
1.2649	std. dev.
0.2108	std. error
36	n
35	df
-1.71	t
.9520	p-value (one-tailed, upper)
6.3915	confidence interval 95.% lower
7.2474	confidence interval 95.% upper
0.4280	margin of error

   (a) Conduct a hypothesis test to determine if the average cost for all the funeral homes in the nation has increased.

   (b) Find and interpret the 95% confidence interval.

a) $36 \geq 30$ ✓
$\alpha = .05$
$H_0 = \mu = 7.18$
$H_A = \mu > 7.18$

test stat:
$t = -1.71$
Pvalue: .9520

we didn't find any evidence that the average cost for all funeral homes has increased.

b) (6.3915, 7.2474)

we are 95% confident that the average cost for all funeral homes is between 6.3915 and 7.2474.

24. Matt is a physical therapist student. He believes that people usually don't bend their knee far enough when instructed. He thinks they just aren't very good at estimating the angle. He takes a sample of 10 patients. He asks them to bend their knees 120°. He measures the angles and finds

$$\overline{x} = 115.2$$

$$s = 5.3$$

Conduct a test to see if the average for all people would be less than 120°. (The data looks fairly normal.)

normal given ✓

$\alpha = .05$

$H_0: \mu = 120$

$H_A: \mu < 120$

test stat: $df = 9$

$$t = \frac{115.2 - 120}{5.3/\sqrt{10}} = -2.863$$

p-val: between .005 ~~and .01~~

we found strong evidence that the average knee bend for all patients is less than 120°.