STAT 400
UIUC

Stepanov
Dalpiaz

Practice Problems #11

SOLUTIONS

The following are a number of practice problems that may be *helpful* for completing the homework, and will likely be **very useful** for studying for exams.

1 – 2. A store sells "16-ounce" boxes of *Captain Crisp* cereal. A random sample of 9 boxes was taken and weighed. The results were the following (in ounces):

15.5 16.2 16.1 15.8 15.6 16.0 15.8 15.9 16.2

Assume the weight of cereal in a box is normally distributed.

Recall: $\overline{x} = 15.9$, $s = 0.25$.

1. a) The company that makes *Captain Crisp* cereal claims that the average weight of its box is at least 16 ounces. Use a 0.05 level of significance to test the company's claim.

Claim : $\mu \geq 16$. $\mathbf{H_0 : \mu \geq 16}$ vs. $\mathbf{H_1 : \mu < 16}$. Left - tail.

$\overline{X} = 15.9$. $s = 0.25$. $n = 9$. $\alpha = 0.05$.

σ is unknown. $n = 9$ – small.

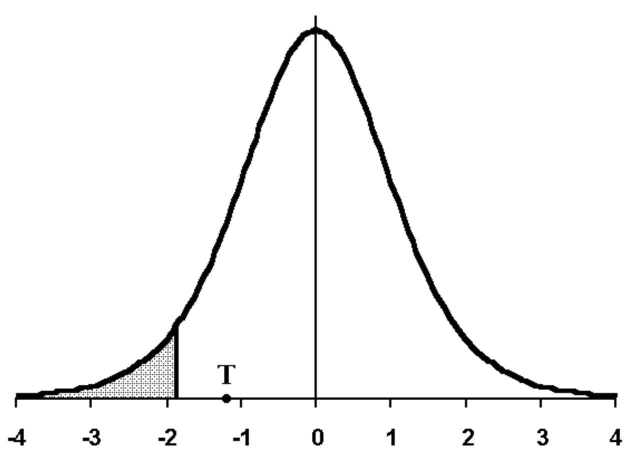
$$T = \frac{\overline{X} - \mu_0}{s / \sqrt{n}} = \frac{15.9 - 16}{0.25 / \sqrt{9}} = \mathbf{-1.2}.$$

Rejection Region: $T < -t_\alpha$

number of degrees of freedom

$n - 1 = 9 - 1 = 8$.

$-t_{0.05} = -1.860$.



The value of the test statistic is **not** in the Rejection Region.

Do NOT Reject H_0 at $\alpha = 0.05$.

b) What is the p-value of the test in part (a)? (You may give a range.)

$t_{0.10} = 1.397 > 1.2 > t_{0.25} = 0.706$. **0.10 < p-value < 0.25**.

c) Use a computer to find the p-value of the test in part (a).

"Hint": EXCEL =TDIST(t , degrees of freedom , 1)
gives area to the right of t.
=TDIST(t , degrees of freedom , 2)
gives $2 \times$ (area to the right of t).

OR R `> pt( t ,` degrees of freedom)
gives area to the left of t.

=TDIST(1.2,8,1)
0.132234

```
> 1 - pt(1.2, df = 8)
[1] 0.1322336
```

2. d) Some customers believe that the variance of the weight of boxes with *Captain Crisp* cereal is more than 0.05. Use $\alpha = 0.10$ to test the customers' claim.

Claim : $\sigma^2 > 0.05$. $H_0 : \sigma^2 \le 0.05$ vs. $H_1 : \sigma^2 > 0.05$. Right - tail.

$$\chi^2 = \frac{(n-1)\cdot s^2}{\sigma_0^2} = \frac{(9-1)\cdot 0.0625}{0.05} = \mathbf{10}.$$

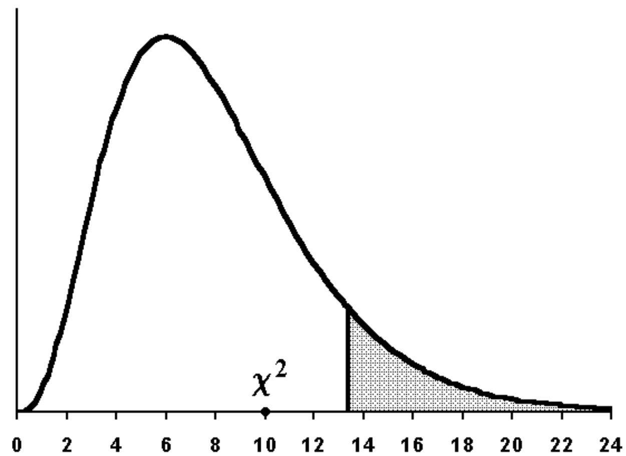
Rejection Region: $\chi^2 > \chi^2_\alpha$

number of degrees of freedom

$n - 1 = 9 - 1 = 8.$

$\chi^2_{0.10} = 13.36.$

The value of the test statistic is **not** in the Rejection Region.



Do NOT Reject H_0 at $\alpha = 0.10$.

e) What is the p-value of the test in part (d)? (You may give a range.)

$\chi^2_{0.10} = 13.36 > 10 > \chi^2_{0.90} = 3.490.$ **0.10 < p-value < 0.90**.

f) Use a computer to find the p-value of the test in part (d).

"Hint": EXCEL =CHIDIST(x , degrees of freedom)
gives area to the right of x.

OR R > pchisq(x , degrees of freedom)
gives area to the left of x.

```
=CHIDIST(10,8,1)
0.265026
```

```
> 1 - pchisq(10,df = 8)
[1] 0.2650259
```

3. Suppose that the amount of cereal dispensed into a box is normally distributed. If the mean amount dispensed in a box is "too small," then the proportion of "underfilled" boxes (boxes with less than 16 ounces of cereal in them) is too large. However, if the mean is "too large," then the company loses money "overfilling" the boxes. The CEO of the company that makes *Captain Crisp* cereal, Mr. Statman, is concerned that the machines that dispense cereal into boxes do not have the proper ("optimal") setting for the mean amount dispensed. A random sample of 196 boxes was obtained, the sample mean amount of cereal in these 196 boxes was 16.07 ounces, the sample standard deviation was 0.21 ounces.

a) Construct a 95% confidence interval for the current mean amount of cereal dispensed into a box.

$\overline{X} = 16.07.$ $s = 0.21.$ $n = 196.$

σ is unknown. $n = 196$ – large. The confidence interval : $\overline{X} \pm z_{\alpha/2} \frac{s}{\sqrt{n}}$.

95% confidence level $\alpha = 0.05.$ $\alpha/2 = 0.025.$ $z_{\alpha/2} = 1.960.$

$16.07 \pm 1.960 \frac{0.21}{\sqrt{196}}$ **16.07 ± 0.0294** **(16.0406 ; 16.0994)**

b) Mr. Statman wants to test whether the mean amount dispensed is 16.1 ounces (which he considers to be the "optimal" value for the mean) or not. Perform the appropriate test using a 10% level of significance.

Claim: $\mu = 16.1$. $\mathbf{H_0 : \mu = 16.1}$ vs. $\mathbf{H_1 : \mu \neq 16.1}$. 2 - tail.

$\overline{X} = 16.07$. $s = 0.21$. $n = 196$. $\alpha = 0.10$.

σ is unknown. $n = 196$ – large.

$$Z = \frac{\overline{X} - \mu_0}{s / \sqrt{n}} = \frac{16.07 - 16.1}{0.21 / \sqrt{196}} = \mathbf{-2.00}.$$

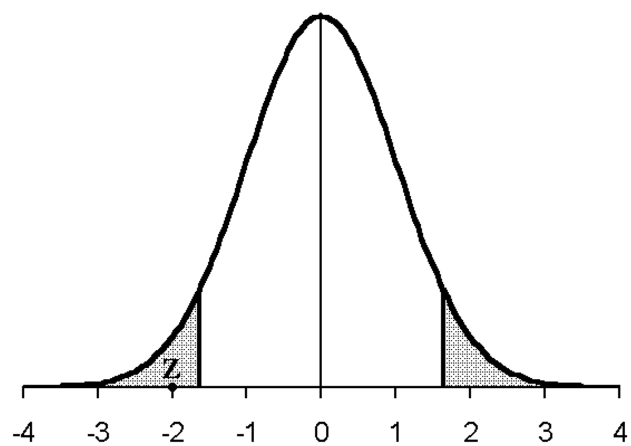
Rejection Region:

$$Z < -z_{\alpha/2} \quad \text{or} \quad Z > z_{\alpha/2}$$

$-z_{0.05} = -1.645$. $z_{0.05} = 1.645$.

The value of the test statistic **is** in the Rejection Region.

Reject $\mathbf{H_0}$ at $\boldsymbol{\alpha = 0.10}$.

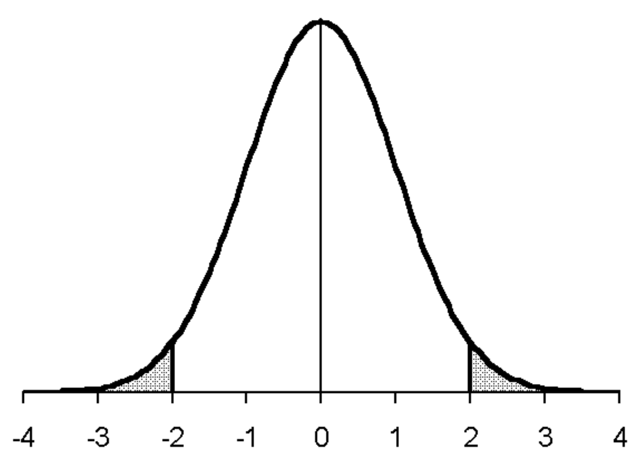


c) Find the p-value of the test in (b).

p - value $= 2 \cdot P(Z < -2.00)$

$= 2 \cdot 0.0228$

$= \mathbf{0.0465}$.



4. After numerous complaints about the weight of the cereal in a box being less than 16 ounces, the company that makes *Captain Crisp* cereal decided to check whether those complaints were valid. A random sample of 400 boxes was obtained, 94 of those 400 had less than 16 ounces of cereal in them.

a) Construct a 95% confidence interval for the overall proportion of boxes of *Captain Crisp* cereal that have less than 16 ounces of cereal.

The confidence interval : $\hat{p} \pm z_{\alpha/2} \cdot \sqrt{\dfrac{\hat{p} \cdot (1-\hat{p})}{n}}$.

$n = 400.$ $x = 94.$ $\hat{p} = \dfrac{x}{n} = \dfrac{94}{400} = 0.235.$

$\alpha = 0.05.$ $\alpha/2 = 0.025.$ $z_{\alpha/2} = 1.960.$

$0.235 \pm 1.960 \cdot \sqrt{\dfrac{(0.235) \cdot (0.765)}{400}}$ $\mathbf{0.235 \pm 0.042}$ **(0.193 , 0.277)**

b) The CEO of the company that makes *Captain Crisp* cereal wants to believe that the overall proportion of boxes of *Captain Crisp* cereal that have less than16 ounces of cereal is at most 20%. Let p be the overall proportion of boxes of *Captain Crisp* that have less than 16 ounces of cereal. Perform the appropriate test using a 10% level of significance.

Claim : $p \le 0.20.$

$H_0 : \boldsymbol{p \le 0.20}$ vs. $H_1 : \boldsymbol{p > 0.20}.$ Right - tail.

$n = 400.$ $x = 94.$ $\alpha = 0.10.$

$\hat{p} = \dfrac{x}{n} = \dfrac{94}{400} = 0.235.$

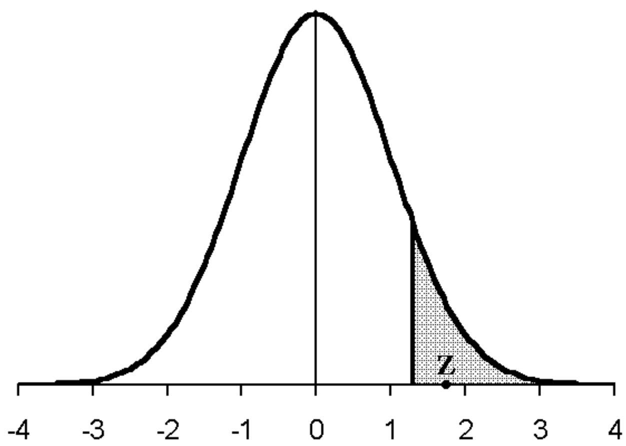
$$Z = \frac{\hat{p} - p_0}{\sqrt{\dfrac{p_0 \cdot (1-p_0)}{n}}} = \frac{0.235 - 0.20}{\sqrt{\dfrac{0.20 \cdot 0.80}{400}}} = \mathbf{1.75}.$$

Rejection Region: $Z > z_{\alpha}$

$z_{0.10} = 1.282.$

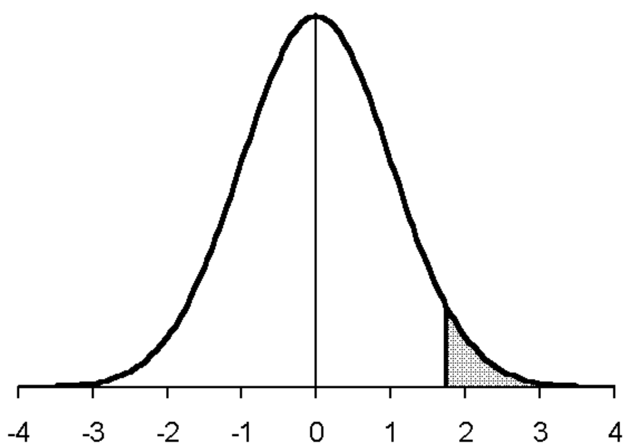
The value of the test statistic **is** in the Rejection Region.

Reject H_0 at $\alpha = 0.10$.



c) Find the p-value of the test in (b).

$$\begin{aligned} \text{p - value} &= P(Z > 1.75) \\ &= 1 - \Phi(1.75) \\ &= 1 - 0.9599 \\ &= \mathbf{0.0401}. \end{aligned}$$



d) Using the p-value from part (c), state your decision (Reject H_0 or Do Not Reject H_0) at $\alpha = 0.05$?

p - value > α $\Rightarrow$ Do NOT Reject H_0

p - value < α $\Rightarrow$ Reject H_0

Since 0.0401 < 0.05, **Reject H_0 at $\alpha = 0.05$**.

= =

5. Researchers are worried that there is excess chlorine in the drinking water supply. They collected 25 independent samples of drinking water and the amount of chlorine in each was measured. The average amount of chlorine in the water samples was 4.2 mg per liter with sample standard deviation of 0.6 mg per liter. The FDA recommends a mean of at most 4.0 mg per liter of water. The researchers want to know if this limit is exceeded in the water supply. Assume the chlorine content measurements are approximately normally distributed.

a) State the null and alternative hypothesis for this test in terms of the relevant parameter.

Claim: "mean of at most 4.0" $\mu \le 4$,

"this limit is exceeded" $\mu > 4$.

$\mathbf{H_0 : \mu \le 4}$ vs. $\mathbf{H_1 : \mu > 4}$. Right - tail.

b) Use $\alpha = 0.05$ to perform the appropriate test. Report the value of the test statistic, the critical value(s), and state your decision.

Test Statistic: $T = \dfrac{\overline{X} - \mu_0}{s / \sqrt{n}} = \dfrac{4.2 - 4}{0.6 / \sqrt{25}} = \mathbf{1.667}$.

Rejection Region: $T > t_{0.05}(24) = 1.711$.

The value of the test statistic is **not** in the Rejection Region.

Do NOT Reject $\mathbf{H_0}$ at $\boldsymbol{\alpha} = 0.05$.

c) Using the t distribution table only, what is the p-value of the test in part (b)? (You may give a range.)

$t_{0.05} = 1.711 > 1.667 > t_{0.10} = 1.318$. **0.05 < p-value < 0.10**.

6. According to Mendelian genetics, a recessive trait will appear in 25% of the population. In order to test whether a particular recessive trait is consistent with the Mendelian model for a specific hybrid plant, the botanist produces a (random) sample of 75 offspring and counts the number of plants with this recessive trait. Indeed, she believes that this recessive trait will appear more frequently in these plants.

a) State the null and alternative hypothesis for this test in terms of the relevant parameter.

Claim: "appear more frequently" $p > 0.25$.

$\mathbf{H_0 : p \le 0.25}$ vs. $\mathbf{H_1 : p > 0.25}$. Right - tail.

b) The botanist observes 27 offspring which exhibited the recessive trait of interest. Calculate the p-value for the appropriate test. What should the botanist conclude at an $\alpha = 0.05$ significant level?

$$\hat{p} = \frac{27}{75} = 0.36.$$

Test Statistic: $$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 \cdot (1 - p_0)}{n}}} = \frac{0.36 - 0.25}{\sqrt{\frac{0.25 \cdot 0.75}{75}}} = \mathbf{2.20}.$$

P-value: Right tail. P-value $= P(Z \ge 2.20) =$ **0.0139**.

Reject $\mathbf{H_0}$ **at** $\boldsymbol{\alpha = 0.05}$. $0.0139 < 0.05$

c) Use a computer to find the exact probability of observing 27 or more offspring which exhibited the recessive trait of interest in a sample of 75 if we assume the null hypothesis in part (a) is true.

"Hint": EXCEL `=BINOM.DIST( x , n , p , 1 )`
gives probability of less than or equal to x.

OR R `> pbinom( x , n , p )`
gives probability of less than or equal to x.

=1-BINOM.DIST(26,75,0.25,1)
0.022305

```
> 1- pbinom(26,size = 75, p = 0.25)
[1] 0.02230519
```

7 – 8. In the past, the average guest check at a local restaurant was $17.85. After the menu has been redesigned, a random sample of 20 guest checks was taken, the sample mean was $19.35 with the sample standard deviation of $3.88. Assume that the guest check amounts are approximately normally distributed.

7. a) Construct a 95% confidence interval for the new overall average guest check.

$\overline{X} = 19.35$, $s = 3.88$, $n = 20$.

$n - 1 = 19$ degrees of freedom.

$t_{0.025}(19) = 2.093$ $\quad\quad 19.35 \pm 2.093 \dfrac{3.88}{\sqrt{20}}$

$\mathbf{19.35 \pm 1.816}$ or $\mathbf{(17.534, 21.166)}$

b) Is there enough evidence that the average guest check has changed? Find the p-value of the appropriate test.

$H_0: \mu = 17.85$ vs. $H_1: \mu \neq 17.85$. 2-tailed.

Test Statistic: $T = \dfrac{\overline{X} - \mu_0}{s/\sqrt{n}} = \dfrac{19.35 - 17.85}{3.88/\sqrt{20}} = \mathbf{1.729}$.

$t_{0.05}(19) = 1.729$ $\quad\quad$ p-value $= 2 \times 0.05 = \mathbf{0.10}$.

8. c) Construct a 95% confidence interval for the overall standard deviation of guest check amounts.

$\chi^2_{0.025}(19) = 32.85$, $\chi^2_{0.975}(19) = 8.907$.

$$\left(\sqrt{\frac{19 \cdot 3.88^2}{32.85}}, \sqrt{\frac{19 \cdot 3.88^2}{8.907}}\right) \qquad \left(\sqrt{8.707}, \sqrt{32.113}\right)$$

$$\mathbf{(2.951, 5.667)}$$

d) Is there enough evidence that the overall standard deviation of guest checks is different than \$3.00 at a 5% level of significance?

That is, test $H_0: \sigma = 3$ vs. $H_1: \sigma \neq 3$. Report the value of the test statistic, the critical value(s), and state your decision.

$H_0: \sigma = 3$ vs. $H_1: \sigma \neq 3$. 2-tailed.

$$\chi^2 = \frac{(n-1)\cdot s^2}{\sigma_0^2} = \frac{(20-1)\cdot 3.88^2}{3^2} = \mathbf{31.7815}.$$

Rejection Region: $\chi^2 < \chi^2_{1-\alpha/2}$ or $\chi^2 > \chi^2_{\alpha/2}$

$n - 1 = 19$ degrees of freedom. $\chi^2_{0.975} = \mathbf{8.907}$, $\chi^2_{0.025} = \mathbf{32.85}$.

The value of the test statistic is **not** in the Rejection Region.

Do NOT Reject H_0 at $\alpha = 0.05$.

9 – 10. In a random sample of 120 male customers at *Burger Queen*, 84 ordered fries with their burgers.

9. a) Construct a 95% confidence interval for the overall proportion of male customers who order fries with their burgers.

The confidence interval: $\hat{p} \pm z_{\alpha/2} \cdot \sqrt{\dfrac{\hat{p} \cdot (1 - \hat{p})}{n}}$. $\qquad \hat{p} = \dfrac{84}{120} = 0.70.$

95% confidence level, $z_{\alpha/2} = z_{0.025} = 1.96.$

$0.70 \pm 1.96 \cdot \sqrt{\dfrac{0.70 \cdot 0.30}{120}}$ $\qquad$ **0.70 ± 0.082** $\qquad$ **$(\,0.618\,,\,0.782\,)$**

b) Find the p-value of the test $H_0: \; p = 0.77$ vs. $H_1: \; p \neq 0.77$, where p is the proportions of male customers who order fries with their burgers.

Test Statistic: $Z = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0 \cdot (1 - p_0)}{n}}} = \dfrac{0.70 - 0.77}{\sqrt{\dfrac{0.77 \cdot 0.23}{120}}} = \mathbf{-1.82}.$

P-value: Two – tailed. P-value $= 2 \times P(\,Z \leq -1.82\,)$

$= 2 \times 0.0344 = \mathbf{0.0688}.$

10. Suppose also that in a random sample of 80 female customers, 48 ordered fries with their burgers.

c) Construct a 95% confidence interval for the difference between proportions of male and female customers who order fries with their burgers.

$\hat{p}_M = \frac{84}{120} = 0.70.$ $\hat{p}_F = \frac{48}{80} = 0.60.$

$$\left(\hat{p}_M - \hat{p}_F\right) \pm z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}_M \cdot \left(1 - \hat{p}_M\right)}{n_M} + \frac{\hat{p}_F \cdot \left(1 - \hat{p}_F\right)}{n_F}}$$

95% confidence level, $z_{\alpha/2} = z_{0.025} = 1.96.$

$$\left(0.70 - 0.60\right) \pm 1.96 \cdot \sqrt{\frac{0.70 \cdot 0.30}{120} + \frac{0.60 \cdot 0.40}{80}}$$

$\mathbf{0.10 \pm 0.135}$ $\mathbf{\left(-0.035\,,\ 0.235\right)}$

d) Find the p-value of the test $H_0:\ p_M = p_F$ vs. $H_1:\ p_M > p_F$.

$$\hat{p} = \frac{Y_1 + Y_2}{n_1 + n_2} = \frac{84 + 48}{120 + 80} = \frac{132}{200} = 0.66.$$

Test Statistic: $Z = \dfrac{0.70 - 0.60}{\sqrt{0.66 \cdot 0.34 \cdot \left(\dfrac{1}{120} + \dfrac{1}{80}\right)}} = \mathbf{1.46}.$

P-value: Right – tailed.

P-value = (area of the right tail) = P(Z ≥ 1.46) = **0.0721**.