Math 3339

Homework 10 (Sections 7.4, 7.6 & 7.10)

Name: James Richardson PeopleSoft ID: 1555520

Instructions:

- Homework will NOT be accepted through email or in person. Homework must be submitted through CourseWare BEFORE the deadline.
- Print out this file and complete the problems.
- Use blue or black ink or a dark pencil.
- Write your solutions in the space provided. You must show all work for full credit.
- Submit this assignment at http://www.casa.uh.edu under "Assignments" and choose **HW9**
- Total possible points: **15**

---

1. Approximately 60% of all part-time college students in the United States are female. (In other words, the population proportion of females among part-time college students is p = 0.6.)
   a. What would you expect to see in terms of the behavior of a sample proportion of females ($\hat{p}$) if random samples of size 100 were taken from the population of all part-time college students? That is give the distribution (shape, center, and spread) of the sample proportion of part-time students that are females.

      The distribution would be symmetrical/ bell shaped, a normal distribution centered at .6

   b. If a random sample of size 100 were taken, what is the probability that less than 50% are females?

      Z =(.5 - .6) / (sqrt(.6 * .4 / 100))

   c. If a random sample of size 100 were taken, there is a 95% chance that the sample proportion ($\hat{p}$) falls between what two values?

2. In 2103 a poll estimated the Well-Being index based on six domains of well-being, including life evaluation, emotional health, work environment, physical health, healthy behaviors, and basic access. The estimate was 66.2% on the national level. The following statement is from the poll:

   In 2013, for results based on 178,072 respondents, one can say with 95% confidenc
   margin of sampling error for those results is ± 0.3 percentage points.

   a. The announced poll result was 66.2 ± 0.3%. Can we be certain that the true population percent falls in this interval? Explain your answer.

   We can be 95% certain; based on our confidence we can be quite sure but not completely certain

   b. Explain to someone who knows no statistics what the announce result 66.2 ± 0.3% means.

   The true average value is within .3% more or less than the given average of 66.2%

   c. This confidence interval has the same form we have met earlier; estimate ± $z^*\sigma_{estimate}$. What is the standard deviation $\sigma_{estimate}$ of the estimated percent?

   p = .662; q = 1 - p
   sd = sqrt(p*q) = .473

3. * The article “An Evaluation of Football Helmets Under Impact Conditions” (*Amer. J. Sports Medi.*, 1984: 233- 237) reports that when each football helmet in a random sample of 37 suspension-type helmets was subjected to a certain impact test, 24 showed damage when tested in the prescribed manner.
   a. Calculate a 99% confidence interval for p, population proportion of all football helmets that show some damage.

```
n = 37; p = 24/n; conf = .99; a = 1 - conf
z = qnorm(conf + a/2)
interval = p + c(-1, 1) * z * sqrt(p*(1-p)/n)
         = (0.4464902, 0.8508070)
```

   b. What sample size would be required for the width of a 99% CI to be at most 0.10, irrespective of p?

```
> #2 * Margin of error <= .1
> #me = z * sqrt(p*(1-p)/n)
> #=> n * me^2 = z^2 * p*(1-p)
> #=> n = z^2 / me^2 * p*(1-p)
> me = .1
> n = z^2 / me^2 * p * (1-p)
> n
[1] 151.2117
```

* Problems came from Devore, Jay and Berk, Kenneth, *Modern Mathematical Statistics with Applications*, Thomson Brooks/Cole, 2007.

1. The Food and Drug Administration monitors the production line of a breakfast cereal company to determine what proportion of its boxes of cereal contain insect parts. The FDA would like to know that proportion to within 5 percentage points with 95% confidence. How many boxes of cereal should they sample?

```
> #2 * Margin of error <= .1
> #me = z * sqrt(p*(1-p)/n)
> #=> n * me^2 = z^2 * p*(1-p)#=> n = z^2 / me^2 * p*(1-p
> z = qnorm(.95 + .95/2)
Warning message:
In qnorm(0.95 + 0.95/2) : NaNs produced
> z = qnorm(.95 + .05/2)
> me = .05
> n = z^2 / me^2 * .25 #they didn't give us a value
> #                      for the proportion, used .5
> n
[1] 384.1459
```

385

2. Suppose the company has been inspected before and that previously the proportion of cereal boxes with insect parts was 0.15. The FDA wants to be as unobtrusive as possible. How many boxes of cereal should they sample?

```
> p = .15
> n = z^2 / me^2 * p*(1-p)
> n
[1] 195.9144
```

196

3. The FDA sampled 60 boxes of cereal and found 12 with insect parts. Find a 95% confidence for the true proportion using all three methods described above.

```
> p = 12 / 60
> n = z^2 / me^2 * p*(1-p)
> n
[1] 245.8534
> #they should test 246 boxes
```

1. Let p denote the proportion of all Math 3339 students who are women. On some random class day, count the number of students attending your class and the number of them who are women. At a significance level of $\alpha = 0.05$, test the null hypothesis $H_0 : p = 1/2$ against the alternative $H_1 : p < 1/2$. Assume that the students attending your class are a random sample of Math 3339 students.

* Problems came from Devore, Jay and Berk, Kenneth, *Modern Mathematical Statistics with Applications*, Thomson Brooks/Cole, 2007.

6. Describe briefly the effect of increasing the size of a sample if all facts about the population remain unchanged, on each of the following.
   a. The margin of error of a 95% confidence interval.

The margin of error is equal to z * sqrt(p(1-p)/n)

So increasing the would decrease if the size of the sample `n` increased

   b. The P-value of a test, when H_0 is false.

The sample size is always in the denominator so the p-value will decrease as n increases.

   c. The impact of an outlier in the sample data on a confidence interval or a P-value.

The impact of an outline on our computed statistic will decrease as the size of the sample increases.

∗ Problems came from Devore, Jay and Berk, Kenneth, *Modern Mathematical Statistics with Applications,* Thomson Brooks/Cole, 2007.

7. An experiment on the side effects of pain relievers assigned arthritis patients to take one of several over-the-counter pain medications. Of the 440 patients who took one brand of pain reliever, 23 suffered some "adverse symptom."
   a. If 10% of all patients suffer adverse symptoms, what would be the sampling distribution of the proportion with adverse symptoms in a sample of 440 patients?

      The proportion mean would be the same as the population mean. The population standard deviation would be equal to sqrt(.1 * .9) and thus the proportion standard deviation would be equal to sqrt(.1 * .9 / 440) * sqrt((N – 440)/(N-1)) where N is the population size.

      In the exact case where 23 of 440 patients suffered some "adverse symptom" the proportion is lower, thus the mean would be lower

   b. Does the experiment provide strong evidence that fewer than 10% of patients who take this medication have adverse symptoms? Verify that the conditions for inference are met. State the hypothesis, calculate the test statistic, then obtain and interpret the P-value.

∗ Problems came from Devore, Jay and Berk, Kenneth, *Modern Mathematical Statistics with Applications*, Thomson Brooks/Cole, 2007.

8. Explain why we cannot use the z test for a proportion in these situations.
   a. You want to know if people have predictive psychic abilities when answers their phones. You ask 10 subjects to guess which of 5 possible friends is calling them before they pick up the phone in order to test they hypothesis, H_0: p = 0.2 (no psychic ability) versus H_a: p > 0.2.

We have no information about the population, neither mean or standard deviation.

   b. You bet a friend that 90% or more of the students taking the BIO 101 class know their blood types. You contact an SRS of 80 of the 1211 students taking BIO 101 this term to test the hypothesis H_0: p = 0.9 versus H_a: p > 0.9.

Again we don't have any information about the population. Since we don't have the population standard deviation, only the sample data (mean and standard deviation can be calculated from the sample data) we have to use a t statistic.

∗ Problems came from Devore, Jay and Berk, Kenneth, *Modern Mathematical Statistics with Applications,* Thomson Brooks/Cole, 2007.

9. *Here are the names of 12 orchestra conductors and their performance times in minutes for Beethoven's Ninth Symphony:

Bernstein	71.03	Furtwängler	74.38
Leinsdorf	65.78	Ormandy	64.72
Solti	74.70	Szell	66.22
Bohm	72.60	Karajan	66.90
Mazur	69.45	Rattle	69.93
Steinberg	68.62	Tennstedt	68.40

a. Check to see that normality is a reasonable assumption for the performance time distribution.

b. Compute a 95% confidence interval for the population standard deviation, and interpret the interval.
c. Supposedly, classical music is 100% determined by the composer's notation, including all timings. Based on your results, is the true or false?

∗ Problems came from Devore, Jay and Berk, Kenneth, *Modern Mathematical Statistics with Applications,* Thomson Brooks/Cole, 2007.

10. Explain what is wrong with each of the following.
    a. You can use a significance test to evaluate the hypothesis H_0: $p\hat{p}$ = 0.6 versus the two-sided alternative.
    b. The margin of error for a confidence interval used for an opinion poll takes into account that fact that people who did not answer the poll questions may have had different responses from those who did answer the questions.
    c. If the P-value for a significance test is 0.35, we can conclude that the null hypothesis has a 35% chance of being true.

∗ Problems came from Devore, Jay and Berk, Kenneth, *Modern Mathematical Statistics with Applications*, Thomson Brooks/Cole, 2007.