A5

Stat 301

Summer 2019

(1) *Carbonated Beverages.* The U.S. Department of Agriculture reports that the mean American consumption of carbonated beverages per year is more than 52 gallons. From previous studies, it is known that the standard deviation of yearly American consumption of carbonated beverages is 20 gallons. A random sample of 30 Americans was taken and the sample mean of those 30 was found to be 69 gallons.
   (a) Estimate the true mean μ American yearly consumption of carbonated beverages with 90% confidence.
   (b) Is the report from the USDA that American consume more than 52 gallons per year feasible based on the results of the CI? Briefly explain your answer.
   (c) If we want to take another sample with a bound half as big as the one we calculated in part a, what sample size should be taken to achieve that?
   (d) Is there sufficient evidence that the mean consuption of carbonated beverages is significantly higher than what the FDA recommends? Complete a hypothesis test; show *all* steps.
   (e) Based on the results of the hypothesis test, what kind of error could have been made (in context)?

(2) *There is no Dana, only Zeul (Who you gonna call?).* In November of 2005 the Harris Poll asked 889 randomly selected US adults, "Do you believe in ghosts?" 29% said they did.
   (a) In constructing confidence intervals, would we use $z^\star$ or $t^\star$ in this situation? Briefly explain why you would use one instead of the other.
   (b) Estimate p, the true proportion of US adults that believe in ghosts, with 90% confidence. Interpret the interval in context of the data.
   (c) Suppose, using the information from the survey (the 29% that believe in ghosts) that a new survey is to be taken and the new bound is to be 2%. What sample size will be required?
   (d) Previous polls have the percent of people that believe in ghosts at 33%. Is there sufficient evidence that the proportion of people who believe in ghosts is less that what previous polls indicate? Complete a hypothesis test; show *all* steps, use the **pvalue approach**.
   (e) Based on the results of the hypothesis test, what kind of error could have been made (in context)?

(3) *Using the t table.* Find the degrees of freedom (*df*) and the value of $t^\star$ for the given sample size and confidence level or significance level (α). [Hint: if it states 'CL', that means that α is divided by 2. If it says '$\alpha =$', then you do not divide α by 2.]
   (a) $n = 6$,$CL = 90\%$
   (b) $n = 21$,$CL = 98\%$
   (c) $n = 29$,$CL = 95\%$
   (d) $n = 12$,$CL = 99\%$
   (e) $n = 6$,$\alpha = 0.10$
   (f) $n = 21$,$\alpha = 0.01$
   (g) $n = 40$,$\alpha = 0.05$

(4) *It ain't easy bein' green.* A dealer in recycled paper places empty trailers at various sites. The trailers are gradually filled by individuals who bring in old newspapers and magazines, and are picked up on several schedules. One such schedule involves pickup every second week. This schedule is desirable if the average amount of recycled paper is more than 1600 cubic feet per 2-week period. The dealer's records for eighteen 2-week periods show the following volumes (in cubic feet) at a particular site. The mean and standard deviation are as follows: $\bar{X} = 1721.6$ and $s = 154.5$
   (a) In constructing confidence intervals, would we use $z^\star$ or $t^\star$ in this situation? Briefly explain why you would use one instead of the other.
   (b) Estimate the true mean weight of recycled paper with 95% confidence. Interpret.

(c) Is there sufficient evidence that the mean amount of recycled paper is more than 1600 cubic feet per 2 week period? Conduct a hypothesis test.
(d) State the kind of error could have been made in context of the problem.

```
recycle=c(1935,1556,1752,1969,1804,1842,1994,1810,1827,1725,2003,1499,1809,1795,1622,1620,1777,2035)
```

(5) In an experiment designed to study effects of illumination level on task performance[1], subjects were required to insert a fine-tipped probe into the eyeholes of ten needles in rapid succession both for a low light level with a black background and a higher level with a white background. Each observation is the time each subject took for task completion.
(a) Estimate the true mean difference in completion time with 95% confidence. Interpret.
(b) Does the data indicate that the higher level of illumination decreases the time for task completion? Conduct hypothesis test.
(c) State the kind of error could have been made in context of the problem.

```
tasks
```

```
  subject black white     di
1       1 25.85 18.23  -7.62
2       2 28.84 20.84  -8.00
3       3 32.05 22.96  -9.09
4       4 25.74 19.68  -6.06
5       5 20.89 19.50  -1.39
6       6 41.05 24.98 -16.07
7       7 25.01 16.61  -8.40
8       8 24.96 16.07  -8.89
9       9 27.47 24.59  -2.88
```

```
rbind(xbar.d,s.d,n)
```

```
         [,1]
xbar.d -7.600
s.d     4.178
n       9.000
```

(6) A study of 584 longleaf pine trees in the Wade Tract in Thomas County, Georgia had several purposes. To see if there is a difference in their sizes (in diameters) in two separate areas of the Wade Tract (northern and southern areas), a random sample of 30 trees from the northern area and 30 trees from the southern area was taken; summary statistics provided below.
(a) Estimate the true difference in mean tree sizes between the northern and southern parts of the Wade Tract with 95% confidence. Interpret.
(b) Is there a significant difference in the mean diameter of trees in the north versus the trees in the south? Conduct hypothesis test.
(c) State the kind of error could have been made in context of the problem.
The summary statistics are as follows:

	n_i	$\overline{x}$	s_i
North	30	23.7	17.5
South	30	34.53	14.258

[1] "Performance of Complex Tasks Under Different Levels of Illumination", *J. Illuminating Eng.*, 1976: 235-242.