HW#8 Due Fri Mar 1, 2019 at the beginning of class.

Problem 1. Consider the probability distributions: normal, t, χ2 and F.
(a) In words, describe the relation between random variables with these distributions. For example, a Bernoulli(p) and Binomial(n,p) distributions are related because they both derive from the Bernoulli process, with Bernoulli(p) random variables being the # "successes" in 1 trial, and Binomial (n,p) random variables may be thought of as the total # "successes" in n independent trials.
(b) If these (and the Gamma distribution) are not already on your page/cards of Probability Notes, then add them now.

Problem 2.
Read this problem VERY CAREFULLY & be sure you understand the scenario & your task.

Age of these Boeing 737-300 jets is of concern. In a front page Seattle Times article Tues. Apr. 5, 2011, a "histogram" was given depicting the numbers of 737-300s delivered by year. (See next page.) The total number of these jets is 1,113. Some summary statistics are as follows.

Mean = 20.59 years Variance = 17.15 years^2 SD = 4.14 years

A statistics student viewed this an excellent opportunity to practice confidence interval computations and interpretation. She poses for herself the following problem and works the following solution. Your task is to critique her work. Check that she has done everything correctly, has suitably justified her steps, and that you agree with her conclusion. If you find anything questionable or incorrect, then explain how you would do it differently.

" Clearly this distribution is not normal, but If a simple random sample of 100 of these jets is taken, that is a large enough sample to use a normal curve based procedure for figuring a CI for a mean, μ. I wonder how the non-normality of the underlying distribution impacts the coverage probability for confidence intervals here.
If I assume that we are given that the population sd is 4.14 years, then I would find an approximate level 95% confidence interval as follows.

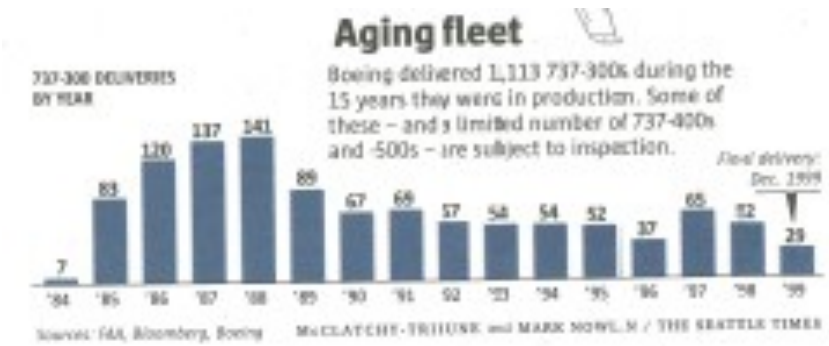
$$\bar{X} \pm 1.96\ SE(\bar{X})$$

I can use the multiplier 1.96 because the sample size (100) is large, and the sd is given. So, filling in the numbers, I get

$$\bar{X} \pm 1.96\ (4.14 / \sqrt{100}) = \bar{X} \pm 0.81$$

Now we can check the probability that a CI constructed in this way actually ends up covering the population parameter (mean) value (20.59 years).
The CI won't cover μ (20.59) if I observe a mean that is more than 0.81 from μ. That means if I get an observation below 20.59-0.81 (= 19.78 years) or above 20.59+0.81 (=21.40 years).
Looking at the histogram from the newspaper, that would happen for planes delivered 19 or fewer years ago (i.e. 1992 or later, since 2011-1992 = 19 years) and for planes delivered 22 or more years ago (i.e. 1989 or before).
The number of such planes is (57+54+54+52+37+65+52+29) plus (7+83+120+137+141+89) = 977. There were 1,113 planes total, so P(the 95% confidence interval misses the true mean μ) is 997/1,113 = 0.88. So, what I would get using this normal-based confidence interval procedure for a level 95% CI actually gets me an 88% CI. That difference must be due to the non-normality of the population distribution.

I guess it's close, but not perfect."



(The years on this graph are ’84 to ’99. The #s over the bars are 7, 83, 120, 137, 141, 89, 67, 69, 57, 54, 54, 52, 37, 65, 52 & 29)

Problem 3.
Let be an iid random sample of size from $X_1, X_2, ..., X_n$ be an iid random sample of size n from the the Uniform(0, θ) distribution.

a. Find the cdf of X_{max}.
b. Find a pivot, Q, such that $P(Q \leq t)$ does not depend on θ .
c. For what value c is $P(Q \leq c) = 0.95$?
d. A 95% confidence interval for θ can take the form (X_{max}, b) where b is a function of X_{max} and c. Recall that $P(X_{max} < \theta) = 1$. Find b.

Problem 4.
Write clear explanations of the following to someone who is “intellectually curious”, and currently is taking an introductory statistics course (for example, STAT 311 or AP Stat), but struggling to understand confidence intervals. (An “intellectually curious” person will ask you lots of questions along the way about why you are doing these steps, and how things fit together.)

(a) Explain how one constructs an approximate level 95% confidence interval for the mean of a normal distribution, with known variance & large sample size – step by step, with each step explained AND justified. Be sure to explain the required assumptions & what they mean, and to clearly remind your reader of what is random, what is fixed, and what is “something else”. Do NOT merely give computational formulae. You will have to use words. Connect your explanation with a realistic example (i.e. not a “toy problem”). Use actual variable names & numbers (but not any of the ones we already have discussed in class).
Be careful.

(b) For each of our “yes/no” interpretations from the “Interpreting Confidence Intervals” page in our current lecture notes set, explain clearly why it is “true” or “false”. Write clear, complete sentences, using your own words.

Problem 5.
Based on what you learned on our NOAA field trip, write about at least two “take aways” from our visit. What did you learn? Did it expand your thinking? If so, in what way? If not, what would have been more informative? Explain. [Exercise your verbal skills here!]
If you did not join us on our NOAA visit, then base your response here on our in-class discussion about our trip.