Steve Sawin
Statistics 3352

Test II

You may use two pages of notes, calculator and computer. Please sign your name below indicating that you have not offered or received help on this test to or from any person, nor used any resource other than your notes, me, and the Excel files on my website.

____

1. [8pt] As their final project the 50 graduating PhD students in statistics each have to do a survey and give a 90% confidence interval for the difference in average income between urban residents and rural residents. Of course this is well within their expertise and they are motvated so they all do this well. In fact the difference in average incomes is $26,000, What does the 90% tell you precisely about this particular situation, in a careful grammatically correct English sentence?

2. [10pt] The average car in America drives 13,500 miles in a year with a standard deviation of 6,000 miles. If you took a simple random sample of 100 cars, what would be the mean and standard deviation of the average annual miles in the sample?

3. In a recent quiz in one section of the 500 student class Meta-Theory of Quantitative Measurement, 3 students got a 10, 6 students got a 9, 4 got an 8, 2 got a 7, 1 student got a 5, 1 got a 4 and 1 student got 0. This is an 18 student class and I got that the average was 7.72.

   (a) [11pt] Treat this section as a simple random sample of the entire class, and give a 99% confidence interval for the average grade in the class.

   (b) [4pt] Why is the normality assumption probably not met in this case?

4. Consider a random sample of size n from a population with an exponential distribution with parameter θ.

   (a) [5pt] Find the pdf $f_\theta(x_1, \ldots, x_n)$ for a sample of size n form this population.

   (b) [12pt] Find $\ln(f_\theta(x_1, \ldots, x_n)$ and $\partial \ln(f_\theta(x_1, \ldots, x_n))/\partial\theta$.

   (c) [5pt] Find the value of θ that maximizes $\ln(f_\theta(x_1, \ldots, x_n)$

(Over)

(d) [5pt] Use the previous page to identify the maximum likelihood estimator for θ in an exponential distribution.

(e) [8pt] Check that the estimator you found in the previous part (which by the way was $\overline{X}$) is an unbiased estimator for θ in an exponential distribution.

(f) [8pt] Suppose a certain brand of tires has a useful life (in miles) that can be represented by an exponential distribution, but with an unknown expected lifespan. Use the random sample of 5 tires below to estimate the expected lifetime of this brand of tire using the maximum likilihood *unbiased* estimator you found above

$$35200 \qquad 41000 \qquad 44700 \qquad 38600 \qquad 41500.$$

5. In a simple random sample of 150 TV viewers, 46 watched the Superbowl this year.

   (a) [10pt] Find a 90% confidence interval for the proportion of all TV viewers that watched the Superbowl this year, showing your calculations (not just quoting the template!). Summarize your answer in a sentence.

   (b) [7pt] Check all three assumptions.

   (c) [7pt] Suppose instead of taking a SRS, you sampled by asking all of your coworkers at ESPN. Identify a plausible source of sampling bias!

out of 100 points

Some Helpful Information

- If X is normal with a mean μ and variance σ^2, then $\overline{X}$ is normal with mean μ and variance σ^2/n, S^2 is indpendent and $(n-1)S^2/\sigma^2$ has a χ^2 distribution with $n-1$ degrees of freedom.
- a χ^2 distribution with ν degrees of freedom has mean ν and standard deviation 2ν.
- An exponential distribution with parameter θ and pdf $\frac{1}{\theta}e^{-x/\theta}$, mean θ and variance θ^2.
- The C confidence interval for proportion is $\hat{\theta} \pm z^*_C\sqrt{\hat{\theta}(1-\hat{\theta})/n}$ and $z^*_{.9} = 1.645$, $z^*_{.95} = 1.96$, $z^*_{.99} = 2.573$,
- The maximum likelihood estimate for θ is the value of θ that maximizes $f_\theta(x_1, \ldots, x_n)$.
- $\hat{\Theta}$ is an unbiased estimator for θ if

$$E\left(\hat{\theta}\right) = \theta.$$