

STA362: Exam 2 (Fall 2001)

Name key

Instructions: You must show all work to receive full credit, except in cases where calculator stat functions are used. Be sure your answers are in the context of the problems. Clearly articulate your written responses. Completeness and conciseness will benefit you. Point values in <>.

1. The manufacturer of a power supply is interested in the variability of output voltage. He has tested 12 units, chosen at random, with the following results:

5.34	5.65	4.76	5.00	5.55	5.54	$n = 12$
5.07	5.35	5.44	5.25	5.35	4.61	$S = 0.3222$

- a) Estimate the variance of the population of output voltages with 90% confidence. Interpret the result in context. <10>

$$\left(\frac{(n-1)s^2}{\chi^2_{\alpha/2}}, \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}} \right) \rightarrow \left(\frac{11(.3222)^2}{19.675}, \frac{11(.3222)^2}{4.575} \right) \rightarrow (.058, .249)$$

We are 90% confident that the true variance in output voltages is somewhere between .058 and .249.

- b) What assumption(s), if any, are required to validate the confidence interval result given above? How might you go about checking this/these assumption(s)? <5>

We must assume the pop'n of output voltages is normal. We can check this using a normal quantile-quantile plot of the sample values.

- c) We have found a $100(1 - \alpha)\%$ CI for σ^2 by using the method of pivotal quantities on the expression

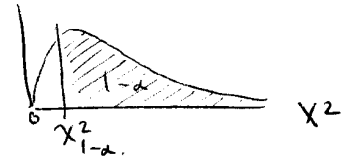
$$P(\chi^2_{1-\alpha/2} \leq \frac{(n-1)s^2}{\sigma^2} \leq \chi^2_{\alpha/2}) = 1 - \alpha$$

Such an interval is called a "two-sided" interval because we are "centering" the confidence probability on the sampling distribution of the sample variance. However, the following, for example, is also true:

$$P\left(\frac{(n-1)s^2}{\sigma^2} \leq \chi^2_{\alpha}\right) = 1 - \alpha$$

Such an interval is called a "one-sided" $100(1 - \alpha)\%$ confidence interval for σ^2 . Using an appropriate starting expression and the method of pivotal quantities, derive the general form of a one-sided $100(1 - \alpha)\%$ confidence interval for σ^2 that has the form $[0, \text{upper endpoint})$. <15>

$$\begin{aligned} P\left(\chi^2_{1-\alpha} < \frac{(n-1)s^2}{\sigma^2} < \infty\right) &= 1 - \alpha \\ \Rightarrow P\left(\frac{1}{\chi^2_{1-\alpha}} > \frac{\sigma^2}{(n-1)s^2} > \frac{1}{\infty}\right) &= 1 - \alpha \\ \Rightarrow P\left(\frac{1}{\chi^2_{1-\alpha}} > \frac{\sigma^2}{(n-1)s^2} > 0\right) &= 1 - \alpha. \\ \Rightarrow P\left(0 < \frac{\sigma^2}{(n-1)s^2} < \frac{1}{\chi^2_{1-\alpha}}\right) &= 1 - \alpha. \\ \Rightarrow P\left(0 < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{1-\alpha}}\right) &= 1 - \alpha. \end{aligned}$$



So, the general form is $[0, \frac{(n-1)s^2}{\chi^2_{1-\alpha}})$.

2. 120 Miami University Oxford undergraduates were surveyed whether or not they feel that racial prejudice is a problem on their campus. The possible responses were Y = it is a problem or N = it is not a problem. Here are the collected data:

Y- N N N Y- Y- N N Y- Y- N N Y- Y- N N N Y- N Y-
 N N N N N N N N Y- N N N Y- Y- N Y- Y- N Y- N
 N Y- Y- Y- Y- N Y- Y- N N N Y- Y- Y- Y- N N N Y- N
 N N N Y- N Y- N N N N N N N N Y- N N Y N N
 N N Y- N Y- N N N Y- Y- N N N N Y- N Y- Y- Y- N
 N N N N Y- N N N N Y- Y- Y- Y- N Y- N N N Y- N

- a) Estimate the true proportion of all MU undergraduates who feel that racial prejudice is a problem on campus. Use a 90% confidence level, and interpret the interval in context. <10>

$$\hat{p} = \frac{44}{120} = .367 \rightarrow \text{CI: } .367 \pm 1.645 \sqrt{\frac{.367(1-.367)}{120}}$$

$$\rightarrow .367 \pm .072$$

$$\rightarrow (.295, .439)$$

We can be 90% confident that between 29.5% to 43.9% of all MU undergrads feel racial prejudice is a problem on campus.

- b) Can you be 90% confident that those who feel that racial prejudice is a problem on campus constitute only a minority of the MU undergraduate population? Use your result from part a) to justify and/or explain your response. <5>

Yes; because the CI is entirely below .5.

- c) How many students should be randomly polled to estimate, correctly to within $\pm 4\%$, the true percentage that feel racial prejudice is a problem while maintaining 90% confidence? <5>

Using the above sample's \hat{p} as a preliminary estimate, we'd need

$$n = \frac{(1.645)^2 (.367)(1-.367)}{(0.04)^2} \approx \underline{\underline{393}}$$

- d) Suppose the same survey was given to a sample of 186 Ohio University students regarding racial prejudice on the OU campus. In that sample, 71 students felt racial prejudice is a problem there. Use a confidence interval to determine if there is reason to believe that a higher proportion of all OU students feel that racial prejudice is a problem at OU than do MU students for Oxford. Use 90% confidence. <15>

$$\hat{p}_{\text{mu}} = .367 \quad \hat{p}_{\text{ou}} = \frac{71}{186} = .382$$

$$\text{CI for proportion difference: } (.382 - .367) \pm 1.645 \sqrt{\frac{(.382)(.618)}{186} + \frac{(.367)(.633)}{120}}$$

$$\rightarrow (.015) \pm .093$$

$$\rightarrow (-, +)$$

Since the CI for $p_{\text{mu}} - p_{\text{ou}}$ contains 0, there is no reason to believe that a higher proportion of OU students feel racial prejudice is a problem than do MU students.

3. Data were collected on 57 males and 35 females that were chosen randomly; suppose the samples were chosen independently from one another. Summary statistics on sample pulse rates are shown below.

	N	MEAN	MEDIAN	STDEV	MIN	MAX	Q1	Q3
M_pulse	57	70.42	70.00	9.95	48.00	92.00	63.00	75.00
F_pulse	35	76.86	78.00	13.62	58.00	100.00	66.00	86.00

= indep samples.

An analyst looking at stem and leaf plots of the male and female pulse rates decided that it was reasonable to assume the data came from normal populations.

- a) Is there convincing evidence that the true mean pulse rate is higher for females than for males? Use a 95% confidence interval, and state your conclusion in context. <15>

$$S_p^2 = \frac{(57-1)(9.95)^2 + (35-1)(13.62)^2}{57 + 35 - 2} = 131.68$$

$$\begin{aligned} \text{CI (F-M)}: (76.86 - 70.42) \pm 2.000 \sqrt{131.68 \left(\frac{1}{57} + \frac{1}{35} \right)} \\ \rightarrow 6.44 \pm 4.93 \\ \rightarrow (1.51, 11.37) \end{aligned}$$

Yes, there is reason to believe this: we can be 95% confident that the true mean female pulse rate is between 1.51 to 11.37 bpm higher.

- b) Pulse rates for the same person vary over time. Suppose these 92 subjects had their pulse recorded twice -- once in the morning and once in the afternoon. If you want to consider the difference in mean pulse rates between morning and afternoon, is it appropriate to use the paired difference approach or the independent samples approach? Why? <5>

Paired difference -- each individual, when measured twice, creates a pair of 'matched' observations on the basis of all personal characteristics.

- c) Estimate the true mean pulse rate for females with 99% confidence. Interpret. <10>

$$\begin{aligned} \text{CI: } 76.86 \pm 2.75 \left(\frac{13.62}{\sqrt{35}} \right) \rightarrow 76.86 \pm 6.33 \\ \rightarrow (70.53, 83.19) \end{aligned}$$

We are 99% confident that the true mean female pulse rate is between 70.53 bpm to 83.19 bpm.

4. Briefly discuss the concept of maximum likelihood estimation and why it can be a useful method in some sampling situations. <5>

The idea is to find the estimated value $\hat{\theta}$ of a parameter θ that MAXIMIZES the probability of obtaining the chosen sample. This probability is given by the likelihood function

$$L(x_1, \dots, x_n; \theta) = f(x_1, \theta) f(x_2, \theta) \dots f(x_n, \theta).$$

This approach can be useful when it is unknown what pop'n characteristic is reflected in the parameter θ . We must, however, know the form of the pop'n dist $f(x, \theta)$.