Statistics 495, Applied Statistics for Industry I

Final Exam, Fall 2000

Name:______________________________

Site: ______________________________

INSTRUCTIONS: You will have 2 hours to complete the exam. There are 7 questions worth a total of 150 points. Not all questions have the same point value so gauge your time appropriately. Read the questions carefully and completely. Answer each question and show work in the space provided on the exam. Turn in the **entire** exam when you are done or when time is up. For essay questions, think before you write.

1. **[20 pts]** A study is conducted to investigate heat transfer in article molds used in making drinking glasses. Variability in mold temperature is of concern. There are two separate factories in the study. At each factory, 2 disks are selected. Two molds are selected from each disk and 5 determinations of temperature are made on each mold. Below are the data for Factory 1 and the associated Analysis of Variance table.

Factory 1			
Disk 1		Disk 2	
Mold 1	Mold 2	Mold 1	Mold 2
458	471	471	470
466	467	469	468
464	467	473	472
471	472	471	465
463	469	470	472

Source	df	Sum of Squares	Mean Square	Expected Mean Square
Disk	1	54.45	54.45	$\sigma^2_{Det} + 5\sigma^2_{Mold} + 10\sigma^2_{Disk}$
Mold	2	62.50	31.25	$\sigma^2_{Det} + 5\sigma^2_{Mold}$
Determination	16	154.00		σ^2_{Det}
Total	19	270.95		

(a) **[9]** Compute the estimates of the variance components and give their associated percentage contributions.

(b) **[5]** Give an interpretation of your findings. Where should improvement efforts be focused?

(c) **[6]** At Factory 2 the estimated variance components are:

- $\hat{\sigma}^2_{Disk} = 3.06$
- $\hat{\sigma}^2_{Mold} = 8.18$
- $\hat{\sigma}^2_{Det} = 4.55$

Do the two factories appear to have the same problems? Explain briefly. What recommendation would you make for this factory?

2. **[15 pts]** A textile manufacturer makes fabric for vehicle upholstery. Each bolt of fabric is inspected and graded according to the number of flaws. The top graded fabric (fewest flaws) is shipped to the company that does the vehicle upholstery. The lower graded fabric is sold at a greatly reduced price to other buyers. An engineer at the textile plant constructs a U-chart for the rate of flaws and discovers that there are numerous out-of-control alarms. The engineer goes to her manager with her findings and asks for resources to identify and eliminate the special causes affecting the process. The manager refuses, saying: "We are shipping fabric that meets our customers' specifications. We don't have a problem." Is there a problem? Discuss this situation.

3. **[20 pts]** It's December and time to dust off those strings of lights to decorate the house or apartment. The first thing to deal with are the burned out bulbs in the strings of lights that need to be replaced.

   (a) **[5]** You see in a newspaper add that someone has a box full of old bulbs but doesn't say if they all work or not. You go to the address and ask to look at the bulbs. You select 50 bulbs at random (its a big box) and test them to see if they light. You construct a confidence interval on the proportion of bulbs that do not light. Based on the confidence interval you will decide how much you are willing to pay to purchase the entire box. Is this an analytic or enumerative study? Explain briefly.

   (b) **[5]** The fellow with the box of bulbs turns your offer down so you decide to go directly to the source. You want to make sure that the factory that produces replacement bulbs knows what it is doing, statistically. You find out from their Web site that they select subgroups of size 50 every hour and chart the proportion that do not light on a p-chart. Based on the p-chart they will make adjustments to the process. Is this an analytic or enumerative study? Explain briefly.

   (c) **[10]** What are the similarities and differences between the statistical methods used in each of the studies mentioned above? Explain briefly. In your explanation it may be helpful to refer to the formulas used to construct the confidence interval in (a) and the control chart in (b).

4. **[15 pts]** A company uses mechanical switches in one of its products. The company wants the switches to withstand 50,000 on/off cycles without failure. A new supplier ships a lot of 20,000 switches. A simple random sample of 200 switches is put on test. Because of the wear on the switches at the end of the test all 200 switches are thrown away. Of the 200 switches tested, 10 fail before 50,000 on/off cycles while the remaining 190 withstand 50,000 on/off cycles.

   (a) **[5]** For this problem is sampling done with or without replacement? Explain briefly.

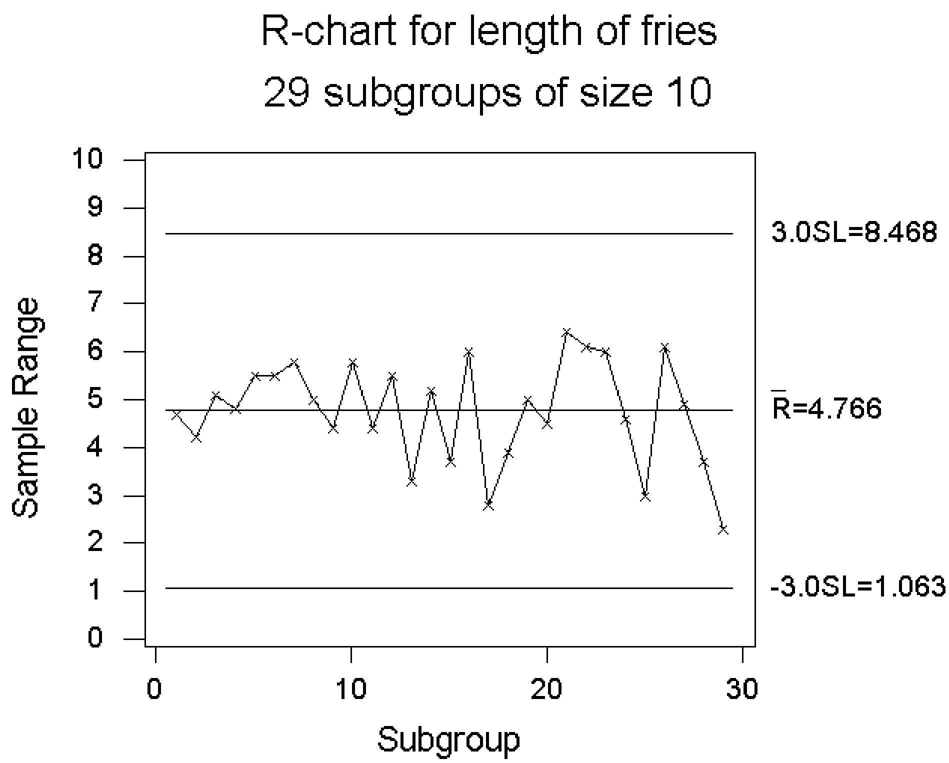
   (b) **[4]** Based on the sample what is the approximate **total number** of parts in the entire lot that will withstand 50,000 on/off cycles?

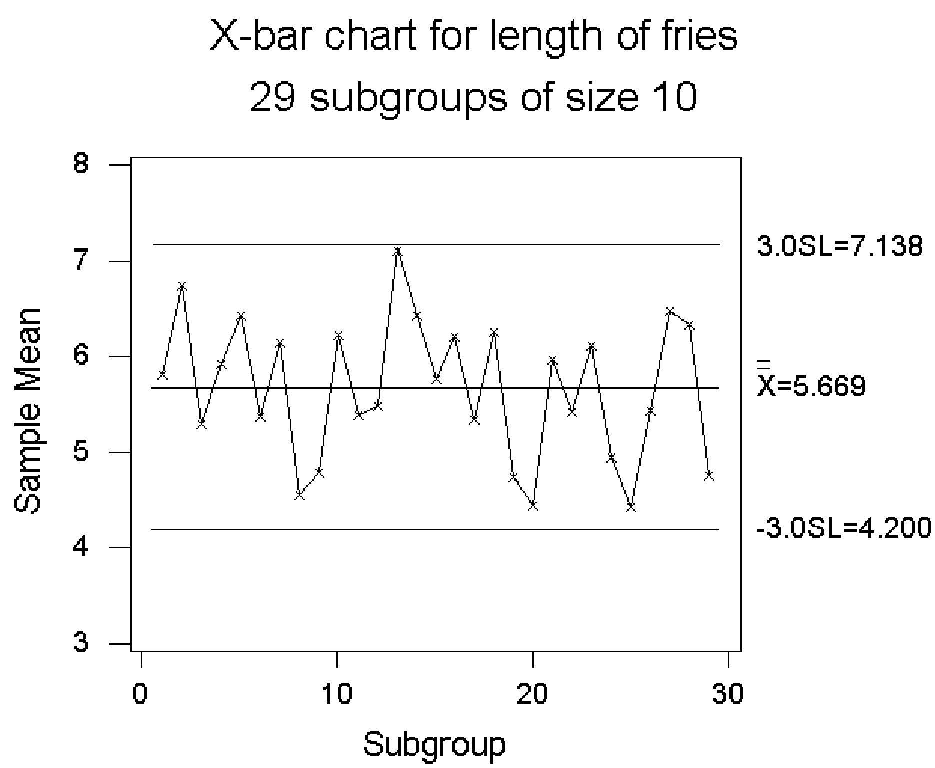
   (c) **[6]** Give an approximate 95% confidence interval for the **total number** of parts in the lot that will withstand 50,000 on/off cycles?

5. **[50 pts]** A famous fast food restaurant chain prides itself on providing consistent product across the country. In order to do this, it uses techniques of statistical process monitoring. One product is french fries. Batches of french fries are cooked, salted and put in serving size packages. Periodically, a package is selected and 10 french fries (a subgroup) are selected and measured. One of the measurements is the length of the french fry.

   (a) **[7]** If we wished to use the 10 french fries to estimate the mean length of french fries in that batch, would this be an enumerative or analytic study? Explain briefly. Be sure to comment on what assumption must be made about the selection of the 10 french fries.

   (b) **[6]** On the next page you will find control charts ($\overline{X}$ and R based on 29 subgroups of size 10) for the length of french fries. Comment on the stability of the process. Note: The control charts were constructed so that any violation of Nelson's eight rules would show up with a star and the number of rule that was violated.


R-chart for length of fries
29 subgroups of size 10
Sample Range
Subgroup
3.0SL=8.468
R=4.766
-3.0SL=1.063
0
1
2
3
4
5
6
7
8
9
10
0
10
20
30

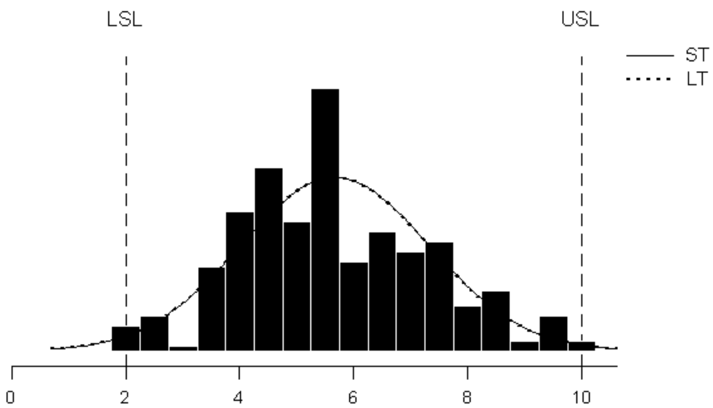

X-bar chart for length of fries
29 subgroups of size 10
Sample Mean
Subgroup
3.0SL=7.138
X=5.669
-3.0SL=4.200
3
4
5
6
7
8
0
10
20
30

Below is a partial output of the Minitab Capability Analysis using the overall standard deviation, s as the estimate of the process standard deviation. The Lower Specification Limit (LSL) is 2 cm, and the Upper Specification Limit (USL) is 10 cm. Refer to this output for the next several questions.

Capability analysis of length of fries: overall s

Process Data	
USL	10.0000
Target	6.0000
LSL	2.0000
Mean	5.6690
Sample N	290
StDev (ST)	1.65760
StDev (LT)	1.65755

Potential (ST) Capability	
Cp	0.80
CPU	0.87
CPL	0.74
Cpk	0.74
Cpm	0.79



Overall (LT) Capability	
Pp	0.80
PPU	0.87
PPL	0.74
Ppk	0.74

Observed Performance	
PPM < LSL	3448.28
PPM > USL	3448.28
PPM Total	6896.55

Expected ST Performance	
PPM < LSL	
PPM > USL	
PPM Total	

Expected LT Performance	
PPM < LSL	
PPM > USL	
PPM Total	

(c) **[5]** Report the value of $\hat{C}_p$ and give an interpretation of this value.

(d) **[5]** Report the value of $\hat{C}_{pk}$ and give an interpretation of this value.

(e) **[4]** What is the value of k for this problem?

(f) **[8]** Assuming a normal distribution for the distribution of the length of french fries compute the expected percentage above the USL=10 cm. Also compute the expected percentage below the LSL=2cm.

(g) **[6]** Is the normal distribution assumption reasonable for these data? Explain briefly. What does this say about the expected percentages computed in (f)?

(h) **[6]** What is the estimate of the process standard deviation σ based on $\overline{R}$? How does this compare to the overall standard deviation? What does this comparison tell you about the process?

(i) **[3]** If we used the estimate of the process standard deviation based on $\overline{R}$ to calculate $\hat{C}_p$ and $\hat{C}_{pk}$, would they be larger, smaller or about the same as the values you reported in (c) and (d)?

6. **[20 pts]** A lot of 200 parts arrives at the loading dock. Before accepting the lot you can test some parts. Testing is destructive and you have to pay for the parts you test, even if you decide not to accept the lot. Suppose that there are 30 defective parts in the lot.

   (a) **[4]** A random sample of 3 parts is selected and tested. If **one or more** parts are defective then the lot will be rejected. What is the probability the lot is rejected? Hint: It may be easier to compute the probability that the lot is accepted and use this to determine the probability that the lot is rejected.

   (b) **[9]** Another way to proceed is to reject the lot when you get the first defective part. Find the probability that you reject the lot on:

   - the first part tested?

   - the second part tested?

   - the third part tested?

   (c) **[5]** How do the two methods compare in terms of making a decision on the lot by testing at most 3 parts?

   (d) **[2]** Which method (a) or (b) would be less costly in terms of testing and paying for tested parts?

7. **[10]** In the space below (I will not read beyond the line below) give the most important **concept** or "big idea" from this course. Briefly explain why it is important and comment on how it will affect the way you do your job. Focus on concepts rather than techniques. There is not a "correct" answer to this question. You will be graded on how well you express yourself. **Think before you write.**

Do not write in this space