Stat 213 Review
Note: just because it's not on the review does not mean that it won't be on the final.
Review all the quizzes, assignments and class notes.

1. A psychologist wished to test a new method of improving memorization by university studet6ns. A sample of 80 students were taught the new technique and then asked to memorize a list of 100 word phrases. The number of word phrases recited correctly, were recorded for each student.
   (a) What is the population of interest in the psychologist's test?
   (b) What is the variable being measured in this test?
   (c) Is the variable I) discrete or ii) continuous?

   The results of psychologist's test are summarized in the frequency distribution below.

Number of correct Phrases	Frequency	Cumulative Relative Frequency (%)
30-39	4	5.0
40-49	4	
50-59		20.0
60-69	12	
70-79	28	70.0
80-89		90.0
90-99		

   (d) Complete the frequency distributions by filling in the blank spaces in the table above.
   (e) What is the class mark (midpoint) of the fifth class?
   (f) What is the class width?
   (g) (i) What are the boundaries of the last class?
       (ii) Name one procedure in which it would be appropriate to use the class boundaries.
   (h) Give the limits of the "median class".
   (i) How many students correctly recited at least 40 but less than 70 word phrases?
   (j) Find the class with the 90th percentile.
   (k) For this distribution, would you expect the mean or the median to be larger or about the same? Explain but do not do the actual calculations.

2. A company sets different prices for a particular stereo in eight different regions of the country. The accompanying table shows the number of units sold and the corresponding prices (in hundreds of dollars).

Sales	420	380	350	400	440	380	450	420
Price	5.5	6.0	6.5	6.0	5.0	6.5	4.5	5.0

   $\sum xy = 18060$ $\sum x = 45$ $\sum x^2 = 257$ $\sum y = 3240$ $\sum y^2 = 1320200$

   (a) State the equation for the linear regression for estimating sales based on price.
       $\hat{y} = 644.5159 - 42.5806\ (x)$
   (b) What effect would you expect a $100 increase in price to have on sales? (decrease sales by 42.5806 units)
   (c) Predict the number of stereo systems sold if the price were $625. ($\hat{y} = 378.3875$ (378 or 379 units)
   (d) Calculate and interpret the correlation coefficient.(r = -.9371)
   (e) Find S^2 (162.3656)

3. Consider the following probability distribution for the length of time a car is in a parkade.

Hours	1	2	3	4	5
Probability	.10	.35	.20	.25	.10

(a) What is the probability a randomly selected car is left in the parkade for at least 3 hours? (.55)
(b) Calculate the mean number of hours that a car is parked in this facility.(2.9)
(c) Find the standard deviation in the number of hours a car is parked. (1.179)
(d) The charges at this parkade are $3 for the first hour (or part thereof) and $2 for every hour thereafter. What is the mean charge per car at this facility? ($6.8)
(e) If a random sample of four cars was taken at this parkade, what is the probability that at most one of the cars would be parked for more than 4 hours? (.9477)
(f) A random sample of 50 cars is taken at this parkade. What is the probability that the average length of time the cars were parked would exceed 3 hours? (.2743)

4. It is suggested that 70% of Canadians say they favour strict environmental controls. Two hundred Canadians are selected at random and interviewed individually. Let the random variable X be the number of respondents who favour environmental controls.
    (a) What is the probability distribution of X? Name the distribution and specify its parameters.
    (b) Give the expressions for the exact probability that between 130 and 146 respondents, inclusive, favour strict environmental controls. The expression does not have to be evaluated.
    (c) Using the appropriate method, find an approximate value for the probability in (b). (.7895)

5. A cigarette manufacturer claims that the mean nicotine content in its kin-size cigarettes is 2mg with a standard deviation of 0.3 mg.
    (a) If the nicotine content of these cigarettes is normally distributed, what is the probability that a randomly selected cigarette has nicotine content greater than 2.02 mg? (.4734)
    (b) What is the interquartile range of the nicotine content of these cigarettes, assuming it is normally distributed? (.4044)
    (c) If the nicotine content is normally distributed, what is the probability that, in a random sample of 10 cigarettes, at least one cigarette has a nicotine content greater than 2.02 mg? (.9984)
    (d) What is the probability that a random sample of 900 cigarettes will yield a mean nicotine content exceeding 2.02 mg? (.0228)
    (e) What is the "population of interest" in this question?
    (f) Was it necessary to assume that the population had a normal distribution for your answer in part (d)? Explain. (No, sample size is large)

6. A particular company always advertises that it is an "Equal Opportunity Employer". A sample of 135 recently hired employees, shows that 12 could be classified as representative of minorities. Construct a 90% confidence interval estimate of the true proportion of minorities being hired by this company. (.0486,.1292)

7. Unoccupied seats on flights cause airlines to lose revenue. A large airline wanted to estimate its average number of unoccupied seats per flight over the last year. A consultant randomly selected the records of 225 flights, noted the number of unoccupied seats for each of the sample flights, and calculated a mean of 11.6 and a standard deviation of 4.1 seats. He then estimated the true mean number of unoccupied seats to be between 11.15 and 12.05. What level of confidence should the airline have in this estimate? (use z table because the t value ~ z value for a large sample size and we can't find 224 df from the table) (.90)

8. An important problem facing strawberry growers is the control of nematodes. These organisms compete with the plants for nutrients in the soil, thereby reducing yield. For this reason , fumigation is a standard part of field preparation. In the past, the fumigants yielded an average of 4 kilograms of marketable fruit for a certain standard sized plot. Recently, a new fumigant has bee developed. It is applied to six standard plots of strawberries, and the yield of marketable fruit (in kilograms) for each plot is 4.5, 4.5, 6.5, 4.5, 5.0, and 4.0
    (a) Do these results indicate a significant increase in average yield at a 5% level of significance? (sample mean = 4.8333, s = .8756, t=2.33> 2.015 RHo)

(b) State any assumptions which were required for your procedure.

9. Thirty percent of the students in a bio-statistics course are men. Of these men, 35% are pre-med students. Further 25% of the women in the course are pre-med students. What is the probability that a member selected at random is
   (a) A woman and a pre-med student? (.175)
   (b) A pre-med student? (.28)
   (c) A woman if the student in not a pre-med student? (.7292)
   (d) A man or is not a pre-med student? (.825)
   (e) Are the events "gender" and "pre-med" independent events? Show using probabilities.

10. The following table gives a breakdown of engineers and scientist (by education level) that subscribe to a technical journal.

Type	Highest degree			
	Bachelors	Masters	Doctorate	Other
Engineer	686	196	34	26
Scientist	578	292	182	6

One of the subscribers is selected at random. All references to degrees mean highest degree.
(a) What is the probability that the person selected has a Masters degree? (488/2000)
(b) What is the probability that the person selected is an engineer and does not have a Masters degree? (746/2000)
(c) What is the probability that the person selected is a scientist or has a doctorate? (1092/2000)
(d) What proportion of the scientists that subscribe to the journal have bachelor degrees? (578/1058)
(e) Are the events "Engineer" and "other highest Degree" mutually exclusive events? Are they independent events? Explain showing probabilities.

11. A financial magazine was asked to obtain a list of 10 randomly selected upper management personnel in Canada. If 15% of upper management are women,
   (a) How many women's names would you expect on the list? (1.5)
   (b) What is the probability that there is at least one woman's name on the list? (.8031)
   (c) What is the probability that there are 6 men's names on the list? (.0401)

12. Daily sales for wind chimes at Wild Wind, a kite shop, exceed $350 fifty-seven percent of the times. What is the (approximate) probability that wind chime sales exceed $350 on at least 65 of 100 randomly selected days? (.0655)

13. A consulting firm establishes a committee to investigate the amount each contract costs over and above the amount quoted in the original contracts (called over-runs). The committee has determined that the standard deviation of over-runs is $17.5 thousand.
   (a) the average over-run for a random sample of 50 contract is $12 thousand. Determine a 90% confidence interval estimate of the true mean over-run based on this sample. (7.9288, 16.0712)
   (b) How large of a sample should they use if they want to be 95% confident that the mean over-runs is in error by nor more than $2 thousand? (295)

14. The heights (in feet) of a random sample of 20 genetically altered pine trees were included in a study by forestry researchers. The following are obtained from this data.
   $\sum x = 519$ $\sum x^2 = 16149$
   (a) A researcher believes that the genetically altered pine trees are taller on average than pine trees that are not genetically altered. The average height of normal pine trees is 22 feet. Test the researcher's belief at the 5% significance level. (t= 1.4871< 1.729, Fto RHo)
   (b) Find the 95% confidence interval for the true mean height of genetically altered pine trees. (20.3907, 31.5093)

15. In a survey conducted by a medical magazine, 240 of 4000 US doctors said they would falsify a death certificate to cover up AIDS if a patient or a patient's family asked that it be done.
    (a) Use the above data to determine a 955 confidence interval estimate for the proportion of US doctors who would do this. (.0526, .0674)
    (b) A newspaper reported that less than 4% of all US doctors would falsify a death certificate for this reason. Is this statement consistent with the findings of the medical magazine? Use the results of (a) to answer this and explain why you have come to your conclusion. If the results are not consistent, explain how they differ.
    (c) Suppose a similar study is to be carried out in Canada. How large of a sample should be used if they want to be 98% confident that the margin of error does not exceed 0.02. Any relevant information from the US sample may be made use of here as it is assumed that the populations are somewhat similar. (766)
    (d) What would be the sample size required for the Canadian study mentioned in (c) if it can not be assumed that the populations are similar? (3393.0625)
    (e) Do the data indicate that more than 4% of all US doctors would falsify a death certificate for the reason given above? Carry out the appropriate statistical test assuming that there is a 5% chance of making a Type I error. Also, determine the p-value of this test and compare it to a significance level of 5%. ($z = 6.4550 > 1.645$, RHo)
    (f) At what levels of significance would you accept that more than 4% of all US doctors falsify a death certificate for the reason given above? ($\alpha <$ p-value) p-value $< .001$

16. Accident occur at Factory A with and average of 1 a week and at Factory B with and average of 1 every two weeks. The occurrence of accidents at Factory A and Factory B are independent.
    (a) What is the probability of exactly one accident at Factory A in the next two-week period? (.27067)
    (b) In the next two weeks, what is the probability of exactly one accident at factory A and none at Factory B? (.09957)
    (c) What is the probability that each factory will have at least one accident in the next week? (.2487)

17. A restaurant offers its patrons a choice of steak, chicken, and ham; if so desired, red wine or white wine may be ordered with the main course. It is known from experience that the probability a customer will order steak, chicken, or ham are, respectively, 0.6, 0.3 and 0.10. Also, the probabilities that a customer will order red wine, white wine, or no wine after he has selected steak are 0.40, 0.10 and 0.50.;the corresponding probabilities for chicken are 0.05, 0.25 and 0.70; those for ham are 0.15, 0.20 and 0.65. Finally, the probability that the customer leaves a good tip is 0.80 if he had steak and red wine; 0.30 if he had steak and white wine; 0.60 if he had steak and no wine; 0.40 if he had chicken and red wine; 0.80 if he had chicken and white wine; 0.70 if he had chicken and no wine; 0.70 if he had ham and red wine; 0.70 if he had ham and white wine; 0.50 if he had ham and no wine. Find the probability that a customer who left a good tip did not order wine. (.5447)

18. The amount of liquid dispensed by an automatic filling machine is normally distributed with a mean of 10.5 ounces with the variance 0.25 ounces. If we purchase 2 drinks, what is the probability
    (a) That each glass contains less than 10 ounces of liquid? (.0252)
    (b) That the average contents in the glasses is less than 10 ounces? (.0793)