Name_____   Student # _____

STAT 201
Final Examination Version A

Richard Lockhart   15 December 2004

Instructions:

1. This is an open book exam.

2. You may use a calculator (with no wireless communications ability).

3. Sometimes, to do the problem, you will need to make assumptions. You should be clear and explicit about what assumptions you need to make the technique you are using reasonable. Marks will be deducted for failing to make such assumptions clear.

4. The exam is out of 50.

5. There are 9 pages including this one.

1. On the next page are a histogram for the gestation period (in days) of 1170 babies and a scatter plot of the weights (in ounces) of these babies against the gestation period (in days) for the babies. The birth weights average 119.5 ounces with a standard deviation of 18.35 ounces. The gestation periods average 279.2 days with a standard deviation of 15 days. The correlation between birth weight and gestation period is 0.43.

   (a) Approximately what fraction of gestation periods were more than 295 days? [4 marks]

   (b) Predict the average birth weight for babies with gestation period 295 days. [4 marks]
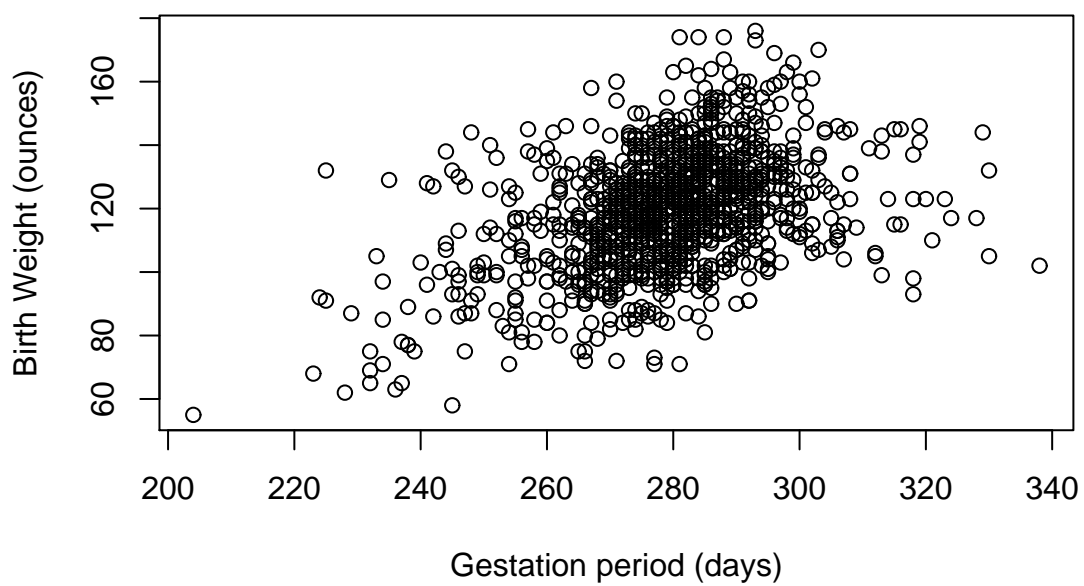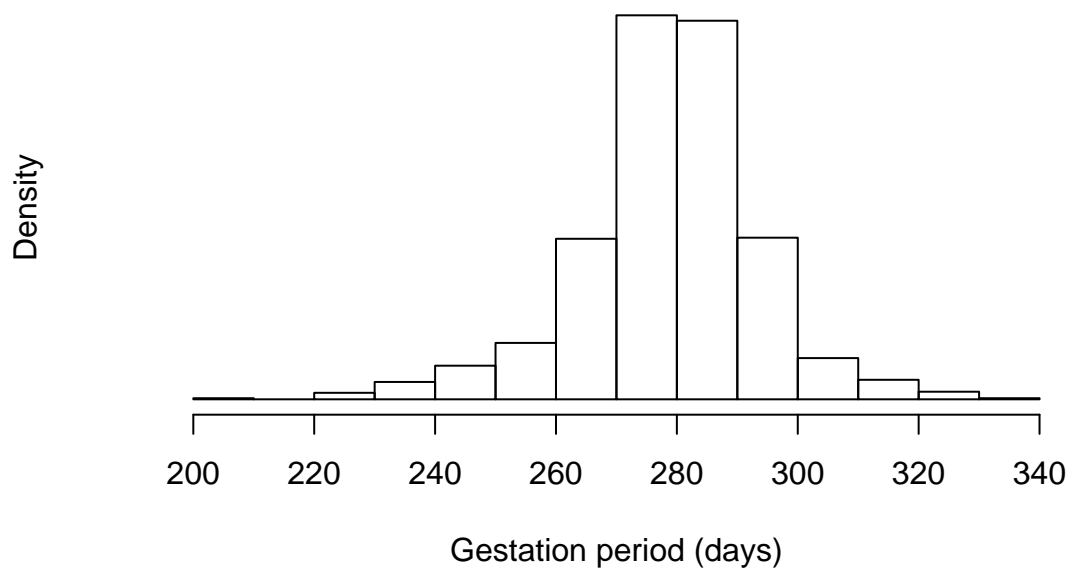
(c) Of the 1170 mothers 459 smoked and 711 did not. I drew a sample of 20 of the smokers and an independent sample of 20 of the non smokers. It is widely reported that birth weights are lower among smoking mothers than among non-smoking mothers. Treat the two samples of twenty as independent simple random samples from some large population and assess the evidence in this data set for this assertion. For the sample of 20 smokers the mean birth weight is 112.1 ounces with a standard deviation of 16.1 ounces. For the non-smokers the mean was 131 ounces with a standard deviation of 15.8 ounces. [5 marks]

2. Michelson's last 20 measurements of the speed of light averaged 299831.5 km/sec with a standard deviation of 54.2 km/sec. Find a 96% confidence interval for the speed of light. [5 marks]

3. A statistician working for a big box retailer looks at daily sales records for one store. S/he notices that there is a correlation of 0.8 between daily sales of tortilla chips and daily sales of televisions. S/he argues that the store should consider further discounts on chips, in order to drive up sales of both tortilla chips and of the much more lucrative televisions. Identify a pitfall in the reasoning and provide an alternative explanation of the observed correlation. [2 marks]

4. Moths are trapped in each of three parks. In each park the trapped moths are examined to see which of 3 species they belong to.

The data are:

		Park Number 1	2	3
	A	40	56	87
Species	B	21	23	14
	C	62	41	38
	Total	123	120	139

Carry out a χ^2 test to see if there is a difference between the parks in distribution of moth species, that is, to see if there is a relationship between park and species of moth.

As part of your answer you should complete the following tables on the next page. [5 marks]

Expected Cell Counts

		Park Number 1	2	3
	A	58.92	57.49	66.59
Species	B	18.68		
	C	45.40		

Components of X^2

		Park Number 1	2	3
	A		0.04	6.26
Species	B	0.29		2.39
	C	6.07	0.24	3.45

5. Give a 98% confidence interval for the difference between the proportion of moths in Park 1 which are of species A and the proportion of moths in Park 2 which are of species A. [5 marks]

6. A researcher is investigating the concentration of a certain metal in river water. S/he is concerned about the difference in this concentration between water near the surface of the river and water near the bottom of the river. S/he picks 6 locations on the river at random. At each location s/he takes two samples: one at the surface and one at the bottom of the river. The concentration of the metal is measured at each site. Is there a difference in the metal concentration between the surface and the river bottom? [5 marks]

Here are the data

Site	Bottom	Surface	Difference
1	0.430	0.415	0.015
2	0.266	0.238	0.028
3	0.567	0.390	0.177
4	0.531	0.410	0.121
5	0.707	0.605	0.102
6	0.716	0.609	0.107
Mean	0.5362	0.4445	0.09167
SD	0.1713	0.1418	0.0607

7. An experiment was conducted in which black bears were moved various distances from their home and then released. Each bear was recorded as eventually returning to its home or not. A total of 54 bears were moved about 100 km and 27 of these returned home. A total of 36 bears were moved over 200 km and of these 12 returned home. Test the hypothesis that the return rate is the same for the two groups of bears. [5 marks]

8. A researcher suspects that blood pressure measurements made by doctors are increased by the fact that they wear white lab coats. To test the theory s/he measures the blood pressure of a sample of 1000 patients. Each patient is measured on two separate occasions, once when the doctor is wearing the lab coat, once not. Wearing the coat the mean is 75 with a standard deviation of 10; without the coat the mean is 74.8 also with and SD of 10. The doctor does a two sample t test and finds no significant difference between the two measurements. Studying the results later s/he picks out the cases where the blood pressure is over 90 when measured wearing the lab coat and carries out a two independent samples t test on these measurements. The difference is now significant (one sided test). The doctor modifies his/her theory to say that there is a group of patients on whom the white coat has a clear effect. Identify two flaws in this line of argument. [3 marks]

9. I have a box with 3 red and 1 green billiard balls in it. I have a bag with 1 red and 3 green balls. I pick one ball from the box without looking and 1 ball from the bag without looking. What is the chance that I get two green balls?   [2 marks]

10. If I pick a card from a well shuffled deck there is a chance of 1/13 that I select an Ace. Suppose I repeat the process of shuffling and picking 1600 times and get 144 Aces. Should I suspect that the chance of getting an Ace is not 1/13?   [5 marks]