Queen's University
Faculty of Arts and Sciences
Department of Mathematics and Statistics
STAT 261 Winter 2006, Midterm Exam
Professor David Steinsaltz

Thursday, March 2, 2006

PRINT YOUR NAME:

You will have 2 hours for this exam. You may use one single-sided pages of notes and a calculator. It is in your interest to show your work and explain your answers. Incorrect answers with substantially correct explanations will receive most of the credit; without such explanations, they will receive no credit.

A normal table is provided at the end.

Write your answers in the spaces provided. The point value is given for each problem. The total number of points is 100.

1) (12 pts) We have six independent observations of a random process: $(X_1, \ldots, X_6) = (4, 4, 2, 7, 5, 8)$.
a) Compute the mean and standard deviation of the observations.

b) Suppose you know that the X_i are normally distributed. What is your best estimate for the expectation and variance of X_i?

c) Suppose now you are told that the expectation of X_i is 6. What is your best estimate for the variance?

2) (12 points) We take a simple random sample of 100 single-person households from the population of Kingston, and ask their annual income. We find that the average income is $30,000, and the standard deviation is $20,000. Say which of the following statements is true or false, and explain. If you need more information to decide, say what you need, and why.
a) About 68% of the households in the sample had incomes in the range $10,000 to $50,000.

b) About 68% of the single-person households in Kingston have incomes in the range $28,000 to $32,000.

c) We can be 68% sure that the real average income of all single-person households in Kingston is in the range $28,000 to $32,000.

3) (12 points) a) What does it mean for an estimator to be consistent?

b) Give an example of an estimator that is biased but consistent.

c) Give an example of an estimator that is unbiased but not consistent.

4) (12 points) Systolic blood pressure in a certain population is normally distributed, with mean 115 mmHg and SD 12 mmHg.
a) What is the probability that an individual selected at random has blood pressure between 109 and 118?

b) Suppose we select 4 individuals at random. What is the probability that the average of their blood pressures is between 109 and 118?

5) (8 points) Statistics Canada wants to estimate the fraction of people in different cities whose incomes are below the poverty level. They will interview a random sample of 1000 people in each city. Assuming things in Toronto are roughly the same as in Kingston:

(1) The accuracy in Toronto (population 3 million) will be about the same as the accuracy in Kingston (population 100,000).
(2) The accuracy in Toronto will be quite a bit higher than the accuracy in Kingston.
(3) The accuracy in Toronto will be quite a bit lower than the accuracy in Kingston.

Choose one, and explain briefly.

6) (20 points)We have n independent observations, $X_1, \ldots, X_n$, from the distribution with density $f_\alpha(x) = (\alpha - 1)x^{-\alpha}$ on $1 \leq x < \infty$, where $\alpha > 1$ is unknown.
a) Compute a method-of-moments estimate for α.

b) Compute the maximum-likelihood estimate for α.

c) Suppose you have 100 observations, and you have computed the maximum-likelihood estimate $\hat{\alpha} = 2$. Compute an approximate 95% confidence interval for the true parameter α.

d) Suppose you have 100 observations, and you have computed the maximum-likelihood estimate $\hat{\alpha} = 2$. Describe how you would use computer simulation (bootstrap approach) to estimate a 95% confidence interval for $\hat{\alpha}$.

7) (8 points) If a penny is spun on its edge, the probability p of coming up heads differs appreciably from 1/2. We wish to estimate p to within 0.01, at a 90% confidence level by spinning the penny n times, and observing how many times it comes up heads. How large must n be?

8) (8 points) We are trying to estimate a parameter θ_0. We observe $X_1, \ldots, X_n$ which are independent observations from the distribution $\mathcal{D}$, determined by the parameter $\theta = \theta_0$. On the basis of the data, we compute an exact 95% confidence interval for θ_0.

Now, we repeat the experiment a total of 20 times, so we have 20 independently determined confidence intervals. What is the probability that the true θ_0 is in every one of the 20 intervals? If you don't have enough information to answer, explain what information you would need.

9) (8 points) The following report appeared in the Feb. 17, 2006 issue of *Science*:

> A British education researcher is causing a stir with his report indicating that U.K. children are getting a lot less sharp than they were 30 years ago.
>
> In a study submitted last month to the Economic and Social Research Council, psychologist Michael Shayer of King's College London reports that performance by children of both sexes has plummeted on a test that involves perceptions of weight and volume. Shayer compared the 1976 performance of 2350 11- and 12-year-olds in a representative sample of British schools with that of students from the years 2001-04.[. . .] In 2004, only 5.7% of boys could equal scores made by the top third in 1976.
>
> The test features questions such as whether the volume of water stays the same when it is poured into different shaped vessels. Psychologist Jim Ridgway of Durham University, U.K., calls it a "fairly robust indicator of cognitive development." Shayer blames the falling scores partly on computer games. Children, especially boys, are playing more in virtual worlds instead of "outdoors, with tools and things," he says.
>
> Durham education researcher Peter Tymms calls the findings "something to be worried about," but says they need confirmation as they are belied by rises in IQ and other test scores.

Suppose that the test scores of boys in 1976 were normally distributed, and were normalized so that the average score was 100 points and the standard deviation was 10 points. Suppose, too, that the scores of boys in 2004 remained normal, and that the standard deviation was still 10 points. On the basis of the above information, what was the average score of the boys tested in 2004? If you don't have enough information, say what additional information you would need.

Scores:

1:_____ 6:_____

2:_____ 7:_____

3:_____ 8:_____

4:_____ 9:_____

5:_____

Total:_____