UNIVERSITY OF OSLO
DEPARTMENT OF ECONOMICS

Assignment in: **ECON4130 – Statistics 2**

Handed out: Monday, October 29, 2007

To be delivered by: Monday, November 5, 2007 **not later than 2:00 p.m.**

Place of delivery: Department office, 12th floor

Further instructions:

- This assignment is part of the **portfolio assessment**. Candidates who have passed the portfolio assessment in a previous semester, do not have the right to hand in the assignments again. This is so, even if the candidate did not pass the exam.
- **Note:** The students can feel free to discuss with each other how to solve the problems, but each student is supposed to formulate her/his own answers. Only single-authored papers are accepted, and papers that for all practical purposes are identical will not be approved.
- If one of the assignments is not accepted, you will be given a new attempt. In order to sit in for the exam, all three assignments must be approved. If no, you will be withdrawn from the exam, so that this will not be an attempt.
- If a student believes that she or he has a good cause not to meet the deadline (e.g. illness) she or he should discuss the matter with the course teacher and seek a formal extension. Normally extension will only be granted when there is a good reason backed by supporting evidence (e.g. medical certificate).

ECON 4130 Statistics 2
Autumn 2007
Tapas Kundu

Assignment 3

Exercise 1

The pareto distribution is given by the following density function:

$$f(x \mid x_0, \theta) = \theta x_0^{\theta} x^{-\theta-1}, \quad x \geq x_0, \theta > 1$$

Assume that $x_0 > 0$ is given (so you do not have to estimate x_0; instead, when estimating θ, you can use x_0 as given) and that $X_1, \ldots X_n$ is an i.i.d. sample.

a. Find the method of moment estimate of θ.

b. Find the mle of θ.

c. Find a sufficient statistic for θ.

Exercise 2

$X_1, \ldots, X_n$ are independent and distributed as $gamma(\alpha, \lambda)$. Find the distribution of $\bar{X}$.

[Hint: Calculating mgf, first show that $\sum X_i$ follows $gamma(n\alpha, \lambda)$. Then, show that if $Y \sim gamma(\alpha, \lambda)$ then $kY \sim gamma\left(\alpha, \frac{\lambda}{k}\right)$. Combine these two results to find the distribution of $\bar{X}$.

Exercise 3

Rice exercise 46 (edition 3) **in chapter** 8.

The data for this exercise (in a STATA data file), can be downloaded from the course webside (use the "save" option and not the "open" option on the menu that appears when you click the link to the data).

Answer first *a*, *b* and *c*.

Hint for *c*: Maximum likelihood can be done in STATA by the ml-command, but somewhat involved. It is much easier to use Excel:

You don't need the whole data set in Excel, only the values of n, $\sum \ln x_i$, $\sum x_i$. Choose two cells for the arguments α and λ, of the log-likelihood, that you fill with suitable start values, e.g., the moment estimates from *b*. Choose also cells containing the three values above. Then define the log-likelihood function in a sixth cell (remember to start the function definition by an equality sign, =). Note also that the function, $\ln \Gamma(\alpha)$ is implemented in Excel under the name GAMMALN. Click the cell where the log-likelihood function is, and use the solver- module (too be found under tools on the menu) to maximize the log likelihood.

[If you have never used Excel before, you are allowed to use the values 1,6 for $\hat{\alpha}$, and 2,6 for $\hat{\lambda}$ (which are close to the mle's) for the rest of the exercise. **State this in your answer**.]

Answer d.

Hint for d: You need to use the "twoway" graph command in STATA (see help twoway). You can combine several graphs in one by separating several graph commands by double vertical lines, ||. For example, suppose the data are in column, x, and the values of a density, calculated for all values in x, are in a column, g. A graph that combines a histogram with the density plot, is for example made by the command

$$\text{twoway histogram } x,\ \text{bin}(15)\ ||\ \text{line } g\ x,\ \text{sort}$$

The option bin(15) says that the histogram shall contain 15 intervals. Choose the number of intervals yourself. The option, sort, to the line command sorts the data before plotting (try what happens without this option).

Alternatively you can use the menu "Overlaid twoway graphs" on the Graphics menu.

To calculate the gamma density, you can use the function, gammaden$(a, 1, 0, x)$, that calculates the $\Gamma(a, 1)$ density. (gammaden$(a, b, 0, x)$ does not seem to work properly for b different from 1). To calculate the $\Gamma(a, b)$ density, you can instead use the following: If $g(x)$ is the density for $\Gamma(a, 1)$, then $b \cdot g(bx)$ is the density for $\Gamma(a, b)$. Hence, the function, $b*$gammaden$(a, 1, 0, b * x)$, should do the trick.

Skip e, f:

Answer g.