ECON 4130
2006 H
HG

Written Paper II 2006H (compulsary)

(To be submitted (in English or Norwegian), Monday 23 October, at Ekspedisjonskontoret 12th floor ES.)

Exercise 1

Section a. of the exercise at the end of the STATA tutorial.

Exercise 2

a. Suppose that X is Pareto distributed with *pdf*

(1) $$f(x) = \begin{cases} \theta b^{\theta} \dfrac{1}{x^{\theta+1}} & \text{for } x > b \\ 0 & \text{otherwise} \end{cases}$$

where $\theta > 0$ and $b > 0$ are parameters. Show that[1]

$$\mathrm{E}(X^r) = \begin{cases} b^r \dfrac{\theta}{\theta - r} & \text{for } 0 < r < \theta \\ \text{does not exist} & \text{for } r \geq \theta \end{cases}$$

b. Let $X_1, X_2, \ldots, X_n$ be *iid* with $X_i \sim \exp(\theta)$ distributed with pdf

$$f(x) = \begin{cases} \theta e^{-\theta x} & \text{for } x > 0 \\ 0 & \text{otherwise} \end{cases}$$

Show that $\bar{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$ is gamma distributed: $\bar{X} \sim \Gamma(n, n\theta)$

[**Hint:** Show first, using the *mgf* for the exponential distribution, that $Y = \sum_{i=1}^{n} X_i$ is gamma distributed.]

---

[1] Note that the result implies that the moment generating function (*mgf*) for the pareto distribution does not exist in an open interval containing 0, which would have implied that moments of all orders exist.

Exercise 3

a. Let $X_1, X_2, \ldots, X_n$ be *iid* and pareto distributed as in (1) with parameters (b, θ), where *b* is known. The preferred estimator for θ based on $X_1, X_2, \ldots, X_n$, is the so called maximum likelihood estimator (mle) that will be discussed later in the course:

$$\hat{\theta} = \frac{n}{\sum_{i=1}^{n} \ln\left(X_i/b\right)} = \frac{1}{\overline{Y}} \qquad \text{where} \quad \overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i \quad \text{and} \quad Y_i = \ln\left(X_i/b\right)$$

Show that $\hat{\theta}$ is a consistent estimator for θ by using the law of large numbers.

b. Show that the exact distribution of $\hat{\theta}$, for any *n*, has the following pdf:

(2)
$$f_{\hat{\theta}}(t) = \begin{cases} \dfrac{(n\theta)^n}{\Gamma(n)} t^{-n-1} e^{-\frac{n\theta}{t}} & \text{for } t > 0 \\ 0 & \text{otherwise} \end{cases}$$

[**Hint:** Use the result in Ex. 2b, and remember the supplementary exercise 2d.]

c. It is rather seldom that we can find (like here) the exact distribution of a mle estimator since these estimators are often complicated. Luckily the general theory for mle's provides approximate distributions (usually normal) for mle's that are valid in large samples. For example, in the present case, the general theory (details on derivation are given later in the course) states that $\hat{\theta}$ is approximately $N(\theta,\ \theta^2/n)$ distributed for large n. We are now in the fortunate position that we can study the quality of the approximation result by comparing the exact with the approximate distribution for various *n*.

Make graphs, using e.g. Stata, for $n = 5, 10, 25, 80$ respectively. For each given *n*, plot both the exact pdf in (2) and the approximate normal density for $\hat{\theta}$ in the same graph. Use the value $\theta = 3.8$ (which corresponds to the estimate for Norwegian female incomes above b = 250 000 in Norway 1998 given in supplementary exercise 2, 3). Comment on the results.

[**Hint:** Choose an interval, from 0 to 8 for example, that covers most of the variation. Make a column of arguments, for example 0,1, 0,2, 0,3, ….., 8,0 (use the fill command and the expression 0.1(0.1)8). (Remember to specify first the length of the column by the command: set obs 80). Then calculate the values of the two densities for each arguments and the four *n* values, giving altogether 8 columns. Then plot, using the *line* option, the two corresponding densities in the same graph for each *n*. The easiest is to use "Overlaid twoway graphs" from the Graphics menu.

When calculating the density in (2), it is convenient to calculate $\ln(f_{\hat{\theta}}(t))$ first and then $\exp[\ln(f_{\hat{\theta}}(t))]$. Note that Stata has a function lngamma(x) that calculates the log of the gamma function. The normal density you can calculate with the function *normalden* (use the command "help normalden" for info on this, or "help functions" for more on functions.]