R300 – Advanced Econometric Methods

PROBLEM SET 2 - QUESTIONS

Due by Mon. October 19

1. Suppose that

$$x_i \sim N(\theta, \sigma_i^2)$$

for known σ_i^2. (Note that, conditional on $\sigma_1^2, \ldots, \sigma_n^2$ the data are not identically distributed. However, we can set this into a random sampling framework if x_i, σ_i^2 are i.i.d. draws from some joint distribution.)

(i) Show that the sample mean

$$\hat{x} = n^{-1} \sum_{i=1}^{n} x_i$$

is unbiased for θ. Derive its sampling variance.

(ii) Show that the estimator

$$\check{x} = \sum_{i=1}^{n} w_i x_i, \qquad w_i = \frac{1/\sigma_i^2}{\sum_{i'=1}^{n} 1/\sigma_{i'}^2}$$

is unbiased. Derive its sampling variance.

(iii) Show that

$$\text{var}_\theta(\check{x}) \leq \text{var}_\theta(\hat{x}).$$

2. Consider the finite population of the three numbers 1,2, and 3.

(i) Write down all possible random samples of size two from this population.

(ii) Compute the sample mean in each of these samples.

(iii) Compute the population mean and the mean of the sample mean. What do you conclude?

(iv) Compute the variance of the sample mean. Compare it to the population variance. What do you conclude?

3. The Pareto distribution on the open interval $[m, \infty)$ is the one-parameter distribution with density

$$\frac{\theta m^{\theta}}{x^{1+\theta}}$$

for $\theta > 0$ (so the minimum value $m \geq 0$ is known here). This distribution is popular for modelling income.

(i) Derive the score and Fisher information for θ.

(ii) Show that

$$y = \log(x/m)$$

follows an exponential distribution with *rate* parameter θ .

(iii) If $y_1, \ldots, y_n$ are random draws from an exponential with rate parameter θ then $u = \sum_{i=1}^{n} y_i$ follows a Gamma distribution with *shape* and *scale* parameters $n, 1/\theta$. Because n here is an integer the distribution is also called the Erlang distribution. Its density at u is

$$\frac{\theta^n u^{n-1} e^{-\theta u}}{(n-1)!}.$$

Verify this for $n = 2$.

(iv) If u is Gamma distributed as above then

$$v = n/u = n/\sum_i y_i$$

is distributed as Inverse Gamma with *shape* and *scale* parameters $n, n\theta$. Verify this. The Inverse Gamma with shape and scale n, β has density

$$\frac{1}{(n-1)!}\beta^n \frac{1}{v^{n+1}} e^{-\beta/v}$$

4. Use the results from the previous question to answer the following questions.

(i) The *maximum likelihood estimator* of θ in the Pareto distribution, based on a random sample of n observations $x_1, \ldots, x_n$, equals

$$\hat{\theta} = \frac{n}{\sum_{i=1}^{n} \log(x_i/m)}.$$

Is this estimator unbiased?

(ii) Can you derive a first-order approximation (in n) of the bias and variance of $\hat{\theta}$? Does the variance of $\hat{\theta}$ approach the efficiency bound?

(iii) Can you come up with an unbiased estimator? Explain.