Econometrics 3 (ECOM90013) Assignment 3

Let $Y_1, Y_2, \dots, Y_n$ denote a simple random sample from a population with probability density function

$$f(y) = \begin{cases} \theta y^{\theta-1}, & 0 < y < 1, \ \theta > 0 \\ 0, & otherwise. \end{cases}$$

1. **(7 marks) Show that the sample mean $\overline{Y}$ is a consistent estimator of $\theta/(\theta+1)$.**

   Hint: first derive the mean of the population and then remember that laws of large numbers are your friends.

First, we derive the mean of the population:

$$E[Y] = \int_0^1 yf(y) \ \ dy$$

$$= \int_0^1 y\theta y^{\theta-1} \ dy$$

$$= \theta \int_0^1 y^{\theta} \ dy$$

Using the power rule of integration, noting $\theta > 0$ by assumption:

$$\int_0^1 y^{k+1} \ dy = \frac{1}{k+1}$$

Hence, we get:

$$= \frac{\theta}{\theta+1}$$

Therefore:

$$E[Y] = \int_0^1 y\theta y^{\theta-1} \ dy = \frac{\theta}{\theta+1}$$

The Weak Law of Large Numbers (WLLN) can only be applied in the instances of finite variance. Hence, we confirm this:

$$E[Y^2] = E[Y^2] - E(Y)^2$$

Using the power rule again, applied to the square term, we first evaluate for first term:

$$E[Y^2] = \int_0^1 y^2 f(y)\ dy = \int_0^1 y^2 y^{\theta+1}\ dy$$

Applying the same logic as previously, this simplifies to:

$$E[Y^2] = \frac{\theta}{\theta + 2}$$

Now we can evaluate this with the second term for the variance, which is finite as there are no n terms in this expression:

$$E[Y^2] = E[Y^2] - E(Y)^2$$

$$= \frac{\theta}{\theta + 2} - \left(\frac{\theta}{\theta + 1}\right)^2$$

Although un-simplified, this expression tells us there is a finite variance. Therefore, we can apply the WLLN to it.

$$E[\bar{Y}_N] = E\left[\frac{1}{n}\sum_{i=1}^{n} y_i\right]$$

$$= \frac{1}{n}\sum_{i=1}^{n} E[y_i]$$

$$= \frac{nE[Y_i]}{n}$$

$$= \frac{\theta}{\theta + 1}$$

Therefore, as this term is bounded according to the WLLN, and it has a finite variance, we can say the sample average is a consistent estimator for the mean. Or, more formally:

$$\frac{1}{n}\sum_{i=1}^{n} Y_i \to^p \frac{\theta}{\theta + 1}$$

2. (1 mark) Derive a consistent method of moments estimator, $\tilde{\boldsymbol{\theta}}$ say, for $\boldsymbol{\theta}$.

To derive the consistent method of moments estimator, $\tilde{\theta}$, for θ we use the observed sample mean to solve for our estimator rather than computing the population mean. Therefore, we need to take our definition of the sample average and solve for our unknown.

$$\overline{Y_N} = \frac{1}{n}\sum_{i=1}^{n} Y_i = \frac{\theta}{\theta+1}$$

Now we simply need to solve for θ:

$$\overline{Y_n} = \frac{\theta}{\theta+1}$$

$$\overline{Y_n}(\theta+1) = \theta$$

$$\overline{Y_n}\theta + \overline{Y_n} = \theta$$

$$\overline{Y_n} = \theta(1-\overline{Y_n})$$

$$\theta = \frac{\overline{Y_n}}{1-\overline{Y_n}}$$

Therefore:

$$\tilde{\theta} = \frac{\overline{Y_n}}{1-\overline{Y_n}}$$

3. (1 mark) Specify the log-likelihood function for this sample.

The log likelihood function is:

$$\mathcal{L}(\theta) = log\mathcal{L}(\theta) = \log\left(\theta^n \prod_{i=1}^{n} y_i^{\theta-1}\right)$$

$$\mathcal{L}(\theta) = nlog(\theta) + (\theta-1)\sum_{i=1}^{n}\log(Y_i)$$

4. (3 marks) Derive the maximum likelihood estimator, $\widehat{\theta}$ say, for θ and prove that it is, indeed a *maximum* likelihood estimator.

To make notation simpler:

$$S = \sum_{i=1}^{n}\log(y_i)$$

Therefore:

$$(\theta) = nlog(\theta) + (\theta-1)\ S$$

First, we differentiate the log-likelihood function with respect to θ:

$$\frac{\partial log\mathcal{L}(\theta)}{\partial\theta} = \frac{n}{\theta} + S$$

Now we formally state the first order condition:

$$\left.\frac{\partial log\mathcal{L}(\theta)}{\partial\theta}\right|_{\theta=\hat{\theta}} = 0$$

Hence our maximum likelihood estimator (MLE) is given as:

$$\frac{n}{\hat{\theta}} + S = 0$$

$$n = -S\hat{\theta}$$

$$\hat{\theta} = -\frac{n}{S}$$

$$\hat{\theta} = -\frac{n}{\sum_{i=1}^{n}\log(Y_i)}$$

Now to verify this is actually a maximum by verifying the Hessian is negative:

$$\frac{\partial^2 log\mathcal{L}(\theta)}{\partial^2\theta} = -\frac{n}{\theta^2}$$

As it's assumed $\theta > 0$ this Hessian must be negative, telling us the log-likelihood is concave and the MLE is actually a maximum.

5. **(2 marks) Derive the Fisher information for the sample.**

The Fisher information is the negative expected value of the Hessian:

$$H(\theta) = \frac{\partial^2 log\mathcal{L}(\theta)}{\partial^2\theta} = -\frac{n}{\theta^2}$$

$$I(\theta) = -E(H(\theta)) = -E\left(\frac{n}{\theta^2}\right) = \frac{n}{\theta^2}$$

6. **(2 marks) Suppose that someone wishes to test the null hypothesis $\boldsymbol{H_0: \theta = 1}$ against the alternative that $\boldsymbol{H_1: \theta \neq 1}$. State the true population density function and describe in words the implication for the population when this null hypothesis when this null hypothesis is true.**

The true population density function is the Beta distribution. The implication of this particular null hypothesis is that, if true, every possible outcome has an equal chance of occurring.

Formally, the Beta distribution has the density:

$$f(y) = \frac{1}{B(\alpha, \beta)} y^{\alpha-1}(1-y)^{\beta-1}, \quad 0 < y < 1$$

When $\beta = 1$, this becomes:

$$f(y) = \frac{1}{B(\theta, 1)} y^{\theta-1}$$

Since $B(\theta, 1) = \frac{1}{\theta}$, we get:

$$f(y) = \theta y^{\theta-1}$$

7. **(12 marks) Derive the likelihood ratio, Lagrange multiplier and Wald tests for the hypotheses of Question 6. In each case provide the decision rule that you would use in practice to apply the test, including any critical value(s) you may need.**

Firstly, we provide again the log-likelihood function. For simplicity we assume $S = \sum_{i=1}^{n} \log{(Y_i)}$ throughout:

$$\mathcal{L}(\theta) = nlog(\theta) + (\theta - 1)S$$

And the MLE:

$$\hat{\theta} = -\frac{n}{S}$$

Secondly, we complete the likelihood ratio (LR) test. We begin by noting the LR statistic in this instance is:

$$LR = 2\left[\mathcal{L}(\hat{\theta}) - \mathcal{L}(1)\right]$$

Substituting in the relevant terms give is:

$$LR = 2\left[\left(nlog(\hat{\theta}) + (\hat{\theta} - 1)S\right) - \left(nlog(1) + (1-1)S\right)\right]$$

$$= \ 2\left[\left(nlog(\hat{\theta}) + (\hat{\theta} - 1)S\right) - 0\right]$$

$$LR = 2\left[\left(nlog(\hat{\theta}) + (\hat{\theta} - 1)S\right)\right]$$

The decision rule for this test depends on a chi-square distribution with one degree of freedom. Therefore, we reject the H_0if the LR test statistic is larger than the critical value at the chosen significance value. At the 5% significance level, the test statistic would need to be larger than 3.841. More formally, we would reject the significance level if:

$$LR = \ 2\left[\left(nlog\left(\hat{\theta}\right) + \left(\hat{\theta} - 1\right)S\right)\right] > \chi^2_{1,0.95} = 0.3841$$

Thirdly, we complete the Wald test. We begin by defining the Wald statistic in this instance is:

$$W = \frac{\left(\hat{\theta} - 1\right)^2}{Var(\hat{\theta})}$$

Substituting in the relevant terms (noting that $Var\left(\hat{\theta}\right) = I\left(\hat{\theta}\right) = \frac{n}{\theta^2}$ as shown in Question 5):

$$W = \frac{\left(\hat{\theta} - 1\right)^2}{\left[I\left(\hat{\theta}\right)\right]^{-1}}$$

$$= \frac{\left(\hat{\theta} - 1\right)^2}{\frac{\hat{\theta}^2}{n}}$$

$$= n\left(\frac{\hat{\theta} - 1}{\hat{\theta}}\right)^2$$

The decision rule for this test depends on a chi-square distribution with one degree of freedom. Therefore, we reject the H_0if the Wald test statistic is larger than the critical value at the chosen significance value. At the 5% significance level, the test statistic would need to be larger than 3.841. More formally, we would reject the significance level if:

$$Wald = \ n\left(\frac{\hat{\theta} - 1}{\hat{\theta}}\right)^2 > \chi^2_{1,0.95} = 0.3841$$

Finally, we complete the Lagrange Multiplier (LM) test. We begin by defining the LM test statistic in this instance:

$$LM = \frac{[\ \mathcal{L}'(\theta_0)]^2}{I(\theta_0)}$$

$$= \frac{[\ \mathcal{L}'(1)]^2}{I(1)}$$

$$= \frac{\left[\frac{n}{\theta} + S\right]^2}{\frac{n}{\theta^2}} = \frac{\left[\frac{n}{1} + S\right]^2}{\frac{n}{1^2}}$$

$$LM = \frac{[n+S]^2}{n}$$

The decision rule for this test depends on a chi-square distribution with one degree of freedom. Therefore, we reject the H_0if the LM test statistic is larger than the critical value at the chosen significance value. At the 5% significance level, the test statistic would need to be larger than 3.841. More formally, we would reject the significance level if:

$$LM = \frac{[n+S]^2}{n} > \chi^2_{1,0.95} = 0.3841$$

8. (6 marks) Without appeal to the generic properties of maximum likelihood estimators, prove that $\hat{\theta}$ is consistent for θ.

As we've already shown, the MLE is:

$$\hat{\theta} = -\frac{n}{\sum_{i=1}^{n} \log(y_i)}$$

Which can be rewritten as:

$$\hat{\theta} = \frac{1}{-\frac{1}{n}\sum_{i=1}^{n} \log(y_i)}$$

Now we need to apply the WLLN to this expression:

$$-\frac{1}{n}\sum_{i=1}^{n} \log(y_i) \to^p E[-\log(y)]$$

To calculate this expectation we need to use the probability density function provided:

$$E[-\log(y)] = \int_0^1 (-\log(y)).\theta y^{\theta-1} = \theta \int_0^1 (-\log(y)).y^{\theta-1}$$

We can evaluate the previous integral by using the substitution $y = e^{-t}$ (and thus $t = -\log(y)$:

$$\int_0^1 (-\log(y)).\theta y^{\theta-1}\ dy = \int_0^\infty t.\theta.e^{-t(\theta-1)}.e^{-t}\ dt = \int_0^\infty t.\theta.e^{-\theta t}\ dt$$

Simplifying further gives us:

$$E[-\log(y)] = \theta \int_0^{\infty} te^{-\theta t}\ dt$$

Through use of the gamma function, defined as:

$$\Gamma(z) = \int_0^{\infty} t^{z-1}e^{-t}\,dt, \qquad \Re(z) > 0$$

By setting $z = 2$, we arrive at a form similar to the previous integral:

$$\Gamma(2)\int_0^{\infty} t^{2-1}e^{-t}\,dt = 1$$

But to execute this we need to make a further substitution of $u = \theta t$ to write:

$$\int_0^{\infty} te^{-\theta t}\ dt = \frac{1}{\theta^2}\Gamma(2) = \frac{1}{\theta^2}$$

Now incorporating the full expression:

$$\int_0^{1}(-\log(y)).\theta y^{\theta-1} = \theta\int_0^{\infty} te^{-\theta t}\ dt = \theta.\frac{1}{\theta^2} = \frac{1}{\theta}$$

Therefore:

$$E[-\log(y)] = \frac{1}{\theta}$$

$$E[\log(Y)] = -\frac{1}{\theta}$$

Substituting this back into our original consistency expression:

$$-\frac{1}{E[\log(Y)]} = -\frac{1}{-\frac{1}{\theta}} = \theta$$

Hence:

$$-\frac{1}{n}\sum_{i=1}^{n}\log(y_i) \to^{p} E[-\log(y)]$$

$$\hat{\theta} \to^{p} \theta$$