ECOM40006/ECOM90013 Econometrics 3
Department of Economics
University of Melbourne

Week 11 Tutorial Exercise Solutions

Semester 1, 2025

1. Ask any questions that you may have about the lectures, etc. If there is still time then please attempt the following questions.

2. Find Method of Moments estimators for the parameter θ, based on a simple random sample $X_1, X_2, \ldots, X_n$, in the following models:

   (a) The Bernoulli Distribution.

   $$f(x) = \theta^x(1-\theta)^{1-x}, \quad 0 \le \theta \le 1; x \in \{0,1\}.$$

   Hint: $\mathrm{E}[X] = \theta$. In an ideal world you would prove this for yourself.

   Solution:

   We have a single parameter so we should be able to get by using the first raw moment, which we want to derive. First, we have the following identity

   $$\sum_{x\in\{0,1\}} f(x) = \theta^0(1-\theta)^{1-0} + \theta^1(1-\theta)^{1-1} = 1-\theta+\theta = 1, \tag{1}$$

   Now

   $$\mathrm{E}[X] = \sum_{x\in\{0,1\}} xf(x) = 0\times\theta^0(1-\theta)^{1-0} + 1\times\theta^1(1-\theta)^{1-1} = \theta.$$

   It follows that the Method of Moments estimator for θ is

   $$\hat{\theta} = \overline{X}, \qquad \text{where } \overline{X} = \frac{1}{n}\sum_{j=1}^{n} X_j.$$

   (b) The Geometric Distribution.

   $$f(x) = \theta(1-\theta)^x, \quad 0 < \theta \le 1; x \in \{0,1,2,\ldots\}$$

   Hint: $\mathrm{E}[X] = (1-\theta)/\theta$. In an ideal world you would prove this for yourself for which an additional hint is to differentiate both sides of the identity $\sum_{x=0}^{\infty} f(x) = 1$ with respect to θ.

   Solution:

   We have a single parameter so we should be able to get by using the first raw moment, which we want to derive. First, we have the following identity (which follows from the geometric distribution being a *proper* distribution, which is one that sums to unity):

$$\sum_{x=0}^{\infty} \theta(1-\theta)^x = 1. \tag{2}$$

As an aside, observe that the density gets its name from the following observation

$$\sum_{x=0}^{\infty} \theta(1-\theta)^x = 1 \implies \sum_{x=0}^{\infty} (1-\theta)^x = \frac{1}{\theta}, \tag{3}$$

which is simply the sum of a geometric series (and an application of the general hint below). Second, we note that, on differentiating (2) with respect to θ,

$$\frac{\mathrm{d}1}{\mathrm{d}\theta} = \frac{\mathrm{d}}{\mathrm{d}\theta} \sum_{x=0}^{\infty} \theta(1-\theta)^x = \sum_{x=0}^{\infty} \frac{\mathrm{d}}{\mathrm{d}\theta} \left[\theta(1-\theta)^x\right] = \sum_{x=0}^{\infty} \left[(1-\theta)^x - x\theta(1-\theta)^{x-1}\right].$$

That is,

$$0 = \sum_{x=0}^{\infty} \left[(1-\theta)^x - x\theta(1-\theta)^{x-1}\right] \implies \sum_{x=0}^{\infty} x\theta(1-\theta)^{x-1} = \sum_{x=0}^{\infty} (1-\theta)^x.$$

Multiplying both sides of this last equality by $(1-\theta)$ yields

$$\sum_{x=0}^{\infty} x\theta(1-\theta)^x = (1-\theta) \sum_{x=0}^{\infty} (1-\theta)^x = \frac{1-\theta}{\theta}, \tag{4}$$

where the second equality is given by (3).

Turning now to the problem at hand we see that

$$\mathrm{E}\left[X\right] = \sum_{j=0}^{\infty} x f(x) = \sum_{x=0}^{\infty} x\theta(1-\theta)^x = \frac{1-\theta}{\theta}.$$

The Method of Moments estimator is obtained on replacing the population expectation(s) by the sample averages. Thus,

$$\overline{X} = \frac{1-\hat{\theta}}{\hat{\theta}} \implies \hat{\theta} = \frac{1}{1+\overline{X}}$$

where $\overline{X}$ is as defined in the previous question.

(c) The Beta Distribution. $\theta = (\alpha, \beta)'$.

$$f(x) = \frac{1}{\mathrm{B}(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \quad \alpha > 0, \beta > 0; \mathrm{B}(\alpha,\beta) = \frac{\Gamma(\alpha)\,\Gamma(\beta)}{\Gamma(\alpha+\beta)}; 0 < x < 1.$$

Hint: Here $\mathrm{E}\left[X\right] = \alpha/(\alpha+\beta)$ and $\mathrm{E}\left[X^2\right] = \alpha(\alpha+1)/[(\alpha+\beta)(\alpha+\beta+1)]$. Ideally you should derive these values for yourself.

Solution:

Here there are two parameters and so we will need to estimating equations. First,

$$\mathrm{E}\left[X\right] = \frac{1}{\mathrm{B}(\alpha,\beta)} \int_0^1 x^{\alpha}(1-x)^{\beta-1}\,\mathrm{d}x. \tag{5}$$

From the definition of the density we know that

$$\int_0^1 x^{\alpha}(1-x)^{\beta-1}\,\mathrm{d}x = \mathrm{B}(\alpha,\beta). \tag{6}$$

Hence,

$$\mathrm{E}\left[X\right] = \frac{1}{\mathrm{B}(\alpha,\beta)}\mathrm{B}(\alpha+1,\beta) = \frac{\Gamma\left(\alpha+\beta\right)}{\Gamma\left(\alpha\right)\Gamma\left(\beta\right)}\frac{\Gamma\left(\alpha+1\right)\Gamma\left(\beta\right)}{\Gamma\left(\alpha+\beta+1\right)} = \frac{\alpha}{\alpha+\beta}, \tag{7}$$

where the final equality has exploited the factorial property of Gamma functions: $\Gamma\left(x+1\right) = x\Gamma\left(x\right), x \neq 0$. Next,

$$\begin{aligned}
\mathrm{E}\left[X^2\right] &= \frac{1}{\mathrm{B}(\alpha,\beta)}\int_0^1 x^{\alpha+1}(1-x)^{\beta-1}\,\mathrm{d}x \\
&= \frac{1}{\mathrm{B}(\alpha,\beta)}\mathrm{B}(\alpha+2,\beta) = \frac{\Gamma\left(\alpha+\beta\right)}{\Gamma\left(\alpha\right)\Gamma\left(\beta\right)}\frac{\Gamma\left(\alpha+2\right)\Gamma\left(\beta\right)}{\Gamma\left(\alpha+\beta+2\right)} \\
&= \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)}
\end{aligned}$$

Method of Moments estimators are then the solutions to the equations

$$\overline{X} = \frac{\hat{\alpha}}{\hat{\alpha}+\hat{\beta}} \tag{8a}$$

$$\overline{\overline{X}} = \frac{\hat{\alpha}(\hat{\alpha}+1)}{(\hat{\alpha}+\hat{\beta})(\hat{\alpha}+\hat{\beta}+1)}, \qquad \overline{\overline{X}} = \frac{1}{n}\sum_{j=1}^{n} X_j^2. \tag{8b}$$

From (8a) we obtain the following two expressions

$$\hat{\alpha}+\hat{\beta} = \frac{\hat{\alpha}}{\overline{X}} \qquad \text{and} \qquad \hat{\beta} = \hat{\alpha}\frac{(1-\overline{X})}{\overline{X}}.$$

Substituting the first of these expressions into (8b) yields

$$\overline{\overline{X}} = \frac{\hat{\alpha}(\hat{\alpha}+1)}{\frac{\hat{\alpha}}{\overline{X}}\left(\frac{\hat{\alpha}}{\overline{X}}+1\right)} = \frac{\overline{X}(\hat{\alpha}+1)}{\left(\frac{\hat{\alpha}+\overline{X}}{\overline{X}}\right)} = \frac{\overline{X}^2(\hat{\alpha}+1)}{\hat{\alpha}+\overline{X}} \implies \hat{\alpha} = \frac{\overline{X}(\overline{X}-\overline{\overline{X}})}{\overline{\overline{X}}-\overline{X}^2}$$

which further implies that

$$\hat{\beta} = \frac{\overline{X}(\overline{X}-\overline{\overline{X}})}{\overline{\overline{X}}-\overline{X}^2}\frac{(1-\overline{X})}{\overline{X}} = \frac{(\overline{X}-\overline{\overline{X}})(1-\overline{X})}{\overline{\overline{X}}-\overline{X}^2}.$$

As an aside, we note that

$$\overline{\overline{X}} - \overline{X}^2 = \frac{1}{n}\sum_{j=1}^{n}(X_i-\overline{X})^2 = \hat{\sigma}_X^2,$$

say, allowing us to write $\hat{\alpha} = \overline{X}(\overline{X}-\overline{\overline{X}})/\hat{\sigma}_X^2$ and $\hat{\beta} = (1-\overline{X})(\overline{X}-\overline{\overline{X}})/\hat{\sigma}_X^2$.

(d) The Pareto Distribution. $\theta = (\theta_1, \theta_2)'$

$$f(x) = \frac{\theta_1 \theta_2^{\theta_1}}{x^{\theta_1+1}}, \quad \theta_1 > 0, \theta_2 > 0; \theta_2 < x < \infty.$$

Observation: This all gets a bit messy but does simplify towards the end. As such, this question is really only for the super keen.

Solution:

Again we have two parameters and so require two estimating equations. First, we see that

$$\int_{\theta_2}^{\infty} \frac{1}{x^{\theta_1+1}}\,\mathrm{d}x = \frac{1}{\theta_1 \theta_2^{\theta_1}}.$$

So, using this result, we obtain

$$\mathrm{E}\left[X\right] = \theta_1 \theta_2^{\theta_1} \int_{\theta_2}^{\infty} \frac{x}{x^{\theta_1+1}}\,\mathrm{d}x = \theta_1 \theta_2^{\theta_1} \int_{\theta_2}^{\infty} \frac{1}{x^{\theta_1}}\,\mathrm{d}x = \frac{\theta_1 \theta_2^{\theta_1}}{(\theta_1 - 1)\theta_2^{(\theta_1-1)}} = \frac{\theta_1 \theta_2}{(\theta_1 - 1)},$$

provided $\theta_1 > 1$ or else the integral is unbounded and $\mathrm{E}\left[X\right]$ is undefined, and

$$\mathrm{E}\left[X^2\right] = \theta_1 \theta_2^{\theta_1} \int_{\theta_2}^{\infty} \frac{x^2}{x^{\theta_1+1}}\,\mathrm{d}x = \theta_1 \theta_2^{\theta_1} \int_{\theta_2}^{\infty} \frac{1}{x^{\theta_1-1}}\,\mathrm{d}x = \frac{\theta_1 \theta_2^{\theta_1}}{(\theta_1 - 2)\theta_2^{(\theta_1-2)}} = \frac{\theta_1 \theta_2^2}{(\theta_1 - 2)},$$

provided $\theta_1 - > 2$ or else the integral is unbounded. Our estimating equations are now

$$\overline{X} = \frac{\hat{\theta}_1 \hat{\theta}_2}{(\hat{\theta}_1 - 1)} \qquad \text{and} \qquad \overline{\overline{X}} = \frac{\hat{\theta}_1 \hat{\theta}_2^2}{(\hat{\theta}_1 - 2)}.$$

Simple rearrangement yields

$$\hat{\theta}_2 = \frac{(\hat{\theta}_1 - 1)\overline{X}}{\hat{\theta}_1} \implies \overline{\overline{X}} = \frac{\hat{\theta}_1 \left(\frac{(\hat{\theta}_1-1)\overline{X}}{\hat{\theta}_1}\right)^2}{(\hat{\theta}_1 - 2)} = \frac{(\hat{\theta}_1 - 1)^2 \overline{X}^2}{\hat{\theta}_1(\hat{\theta}_1 - 2)}.$$

Rearranging the quadratic expression in $\hat{\theta}_1$ we find that

$$\hat{\sigma}_X^2 \hat{\theta}_1^2 - 2\hat{\sigma}_X^2 \hat{\theta}_1 - \overline{X}^2 = 0,$$

where $\hat{\sigma}_X^2$ is as defined in the previous question. Possible roots for this equation are

$$\hat{\theta}_1 = \frac{2\hat{\sigma}_X^2 \pm \sqrt{4\hat{\sigma}_X^4 + 4\hat{\sigma}_X^2 \overline{X}^2}}{2\hat{\sigma}_X^2} = 1 \pm \sqrt{1 + \overline{X}^2/\hat{\sigma}_X^2}.$$

Noting that $\theta_1 > 0$ and that $1 + \overline{X}^2/\hat{\sigma}_X^2 > 1$ it follows that the only feasible root is

$$\hat{\theta}_1 = 1 + \sqrt{1 + \overline{X}^2/\hat{\sigma}_X^2},$$

giving

$$\hat{\theta}_2 = \frac{\overline{X}\sqrt{1 + \overline{X}^2/\hat{\sigma}_X^2}}{1 + \sqrt{1 + \overline{X}^2/\hat{\sigma}_X^2}}.$$

As a final observation, although the derivation was messy, the actual computation of the estimators is not difficult at all, being relatively simple functions of the mean and sample variance of the data.

Hint: As a general hint, remember that probability mass/density functions, $f(x)$ say, are typically of the form:

$$f(x) = \text{normalizing constant} \times \text{kernel} = c \times k(x),$$

in an obvious notation, where the kernel of the density depends on the random variable and the normalizing constant does not. Consequently, because probability mass/density functions must sum/integrate to unity we know that

$$\int_X k(x)\,\mathrm{d}\theta = \frac{1}{c},$$

where $k(x)$ denotes the kernel of the probability mass/density function of a random variable X and $\int_X k(x)\,\mathrm{d}\theta$ should be read as the sum of the probabilities for all values x in the support of $k(x)$ if X is a discrete random variable and as the integral over the support of $k(x)$ if X is a continuous random variable.

3. Please attempt at least Question 2 from the Week 10 Tutorial Exercise.

   Solution:
   See Week 10 Tutorial Exercise Solutions.