Econometrics Assignment 10

BRODARD Lionel, MARCHAL Antoine
GIRO LARRAZ Tomas and SCHONENBERGER Ivan

December 1, 2018

Exercise 1: Method of moment estimator for the normal distribution. (a)

$$\hat{\theta}_{MM} = \frac{1}{n}\sum_{i=1}^{n} Y_i$$

$$\hat{\tau^2}_{MM} = \frac{1}{n}\sum_{i=1}^{n} Y_i^2 - (\frac{1}{n}\sum_{i=1}^{n} Y_i)^2 = \frac{1}{n}\sum_{i=1}^{n} Y_i^2 - \overline{Y}^2 = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \overline{Y})^2$$

To compute the bias, we need to compute the moment expectation.

$$E[\hat{\theta}_{MM}] = \theta$$

$$E[\hat{\tau^2}_{MM}] = \tau^2(\frac{n-1}{n})$$

The estimators for the mean is unbiased but the variance estimator is biased.
(b) To check the consistency, we need to find the probability limit of the estimators. To apply the weak law of large number, we need that $E[Y_i^2] < \infty$ and that $\frac{1}{n^2}\sum_{i=1}^{n} Var(Y_i^2) \longrightarrow 0$:

$$E[Y_i] = \theta$$

$$E[Y_i^2] = \theta^2 + \tau^2 < \infty$$

$$Var[Y_i^2] = E[Y_i^4] - E[Y_i^2]^2 = \theta^4 + 6\theta^2\tau^2 + 3\tau^4 - (\theta^2 + \tau^2)^2 = 4\theta^2\tau^2 + 2\tau^4$$

So:

$$\frac{1}{n^2}\sum_{i=1}^{n} Var(Y_i^2) = \frac{n(\theta^2\tau^2 + 2\tau^4)}{n^2} = \frac{\theta^2\tau^2 + 2\tau^4}{n} \longrightarrow 0$$

Then:

$$p\lim_{n\to\infty} \hat{\theta}_{MM} = p\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} Y_i = \theta$$

Both estimator are consistent

Exercise 2: Method of moments and the OLS estimator. **(a)**

$$E[X_i(Y_i - X_i'\beta)] = 0 \implies \beta = E[X_iX_i']^{-1}E[X_iY_i]$$

Therefore the estimated β with the method of moment is:

$$\hat{\beta}_{MM} = (\frac{1}{n}\sum_{i=1}^{n} X_iX_i')^{-1} * \frac{1}{n}\sum_{i=1}^{n}(X_iY_i) = \hat{\beta}_{OLS}$$

And is the same as the OLS estimated β.

(b)

We can check the limit of $\hat{\beta}_{MM}$ with the Weak Law of Large Number:

$$p\lim_{n\to\infty}(\frac{1}{n}\sum_{i=1}^{n} X_iX_i')^{-1} = E[X_iX_i']^{-1}$$

$$p\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} X_iY_i = E[X_iY_i]$$

$$\implies \hat{\beta}_{MM} = \beta_{MM}$$

Exercise 3: No first moment (30 points) . **(a).** Let X_1 and X_2 be two independent standard normal random variables and $Y = X_1/X_2$. It can be shown that the density function of Y is

$$f(y) = \frac{1}{\pi(1+y^2)} \tag{1}$$

It can be shown that Y has no first moment, namely $E[Y] = \int_{-\infty}^{+\infty} yf(y)dy$ is not defined. Indeed, we show that $E[|Y|]$ doesn't converge:

$$\int_{-\infty}^{\infty} |y| \frac{1}{\pi(1+y^2)}dy = 2\int_{0}^{\infty} \frac{y}{2\pi(1+y^2)}dy = \frac{1}{\pi}[\ln(1+y^2)]_0^{\infty} = \infty$$

Thus the mean $E[Y]$ is undefined.

(b). Considering Cauchy distributed variables, as can be seen from fig. 1, the mean does not seems to converge as the sample size increase.

c. As can be seen from fig. 2 & 3, in the case of variables following standard normal distribution or student distribution for which the first moment exist, the sample mean given by $\frac{1}{m}\sum_{i=1}^{m} Y_i$ is a consistent estimator that converge to the expectation of Y (namely zero in those case).

Exercise 4: . All the following results are obtained in the code with the seed 123.

(a)

$$\begin{aligned}
E(e^{tX}) &= \int_{-\infty}^{\infty} e^{tx}\frac{1}{\sqrt{2\pi\sigma^2}}e^{-(x-\mu)^2/2\sigma^2}dx \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}}e^{-(x^2-2x(\mu+t\sigma^2)+\mu^2)/2\sigma^2}dx \\
&= e^{((\mu+t\sigma^2)^2-\mu^2)/2\sigma^2}\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}}e^{-(x-(\mu+t\sigma^2))^2/2\sigma^2}dx \\
&= e^{\mu t}e^{\sigma^2t^2/2}
\end{aligned}$$

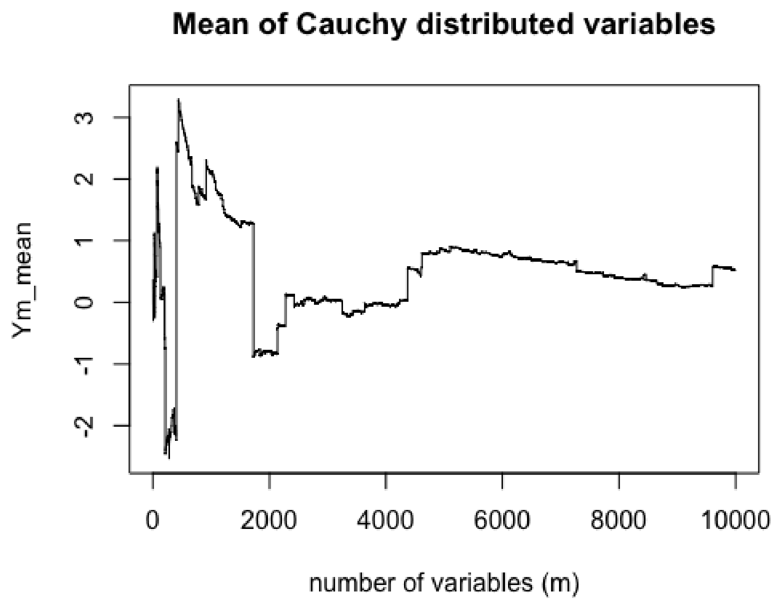

Figure 1: Sample mean of Cauchy variables.

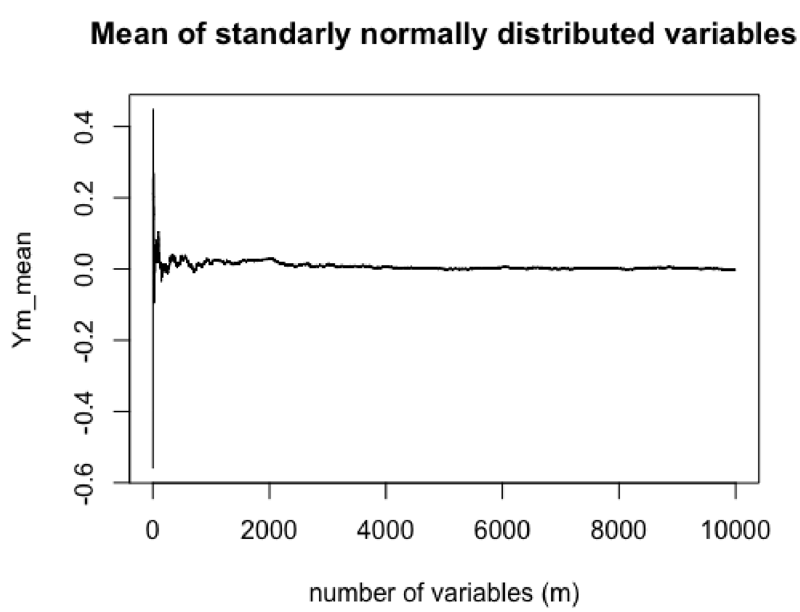

Figure 2: Mean of sample of standardly normally distributed variables.

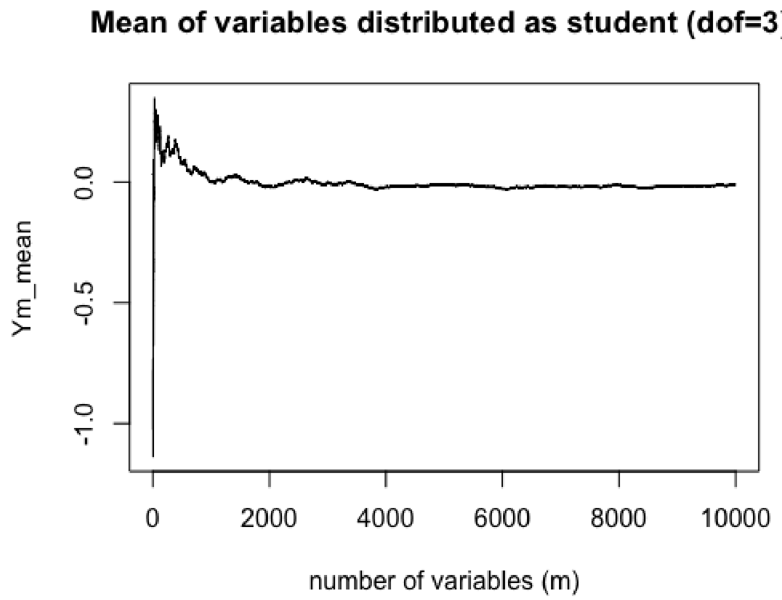

Figure 3: Mean of sample of variables following a student distribution with three degrees of freedom.

(b)

$$E(Y) = M_X(1) = exp(\mu + \sigma^2/2)$$
$$E(Y^2) = M_X(2) = exp(2\mu + 2\sigma^2)$$
$$Var(Y) = exp(2\mu + 2\sigma^2) - exp(2\mu + \sigma^2) = (exp(\sigma^2) - 1)exp(2\mu + \sigma^2)$$

(c)

$$\bar{y} = exp(\hat{\mu} + \hat{\sigma}^2/2)$$
$$\overline{y^2} = exp(2\hat{\mu} + 2\hat{\sigma}^2)$$
$$ln(\bar{y}) = \hat{\mu} + \hat{\sigma}^2/2$$
$$ln(\overline{y^2}) = 2\hat{\mu} + 2\hat{\sigma}^2$$

so,

$$\boxed{\hat{\sigma}^2 = ln(\overline{y^2}) - 2ln(\bar{y})}$$
$$\boxed{\hat{\mu} = 2ln(\bar{y}) - ln(\overline{y^2})/2}$$

(d)

$$E(1/Y) = M_X(-1) = exp(-\mu + \sigma^2/2)$$
$$\boxed{\overline{y^{-1}} = exp(-\hat{\mu} + \hat{\sigma}^2/2)}$$

(e) Generate a random sample $Y_1, ..., Y_n$ for $n = 200$ with parameters $\mu = 1.2$ and $\sigma^2 = 0.8$. Compute the method of moment estimates of μ and σ^2 in the following cases and compare the results:

(i) In the exactly identified case, we directly compute the estimators with the equations derived in (c). We obtain $\hat{\mu} = 1.129$ and $\hat{\sigma}^2 = 0.909$

(ii) In the overidentified case using the additional moment equation obtained in (d), we make a first optimization: we minimize $q = \sum_{i=0}^{3} \bar{\boldsymbol{m}}_{\boldsymbol{i}}(\boldsymbol{Y}, \boldsymbol{\theta})^2$. We obtain $\hat{\mu} = 1.140$ and $\hat{\sigma}^2 = 0.898$. With those parameters fixed, we obtain the $\hat{\boldsymbol{\Phi}}^{\boldsymbol{-1}}$ as defined in the lecture slides of chapter 9, p.17 which we will use in the next step. Then we proceed to a second optimization by minimizing $q = \bar{\boldsymbol{m}}^{\boldsymbol{T}} \boldsymbol{W} \bar{\boldsymbol{m}}$ with $\boldsymbol{W} = \hat{\boldsymbol{\Phi}}^{\boldsymbol{-1}}$. The new estimators obtained are $\hat{\mu} = 1.199$ and $\hat{\sigma}^2 = 0.668$. With those in place, we then compute the final $\hat{\boldsymbol{\Phi}}^{\boldsymbol{-1}}$ matrix which we will use to construct our tests in the next parts of the exercise

(f) We now perform a test of overidentification in this case having added another moment equation. To do so we construct the following statistic

$$nq = \sqrt{n}\overline{\mathbf{m}}(\hat{\boldsymbol{\theta}})^T \hat{\boldsymbol{\Phi}}^{\boldsymbol{-1}} \sqrt{n}\overline{\mathbf{m}}(\hat{\boldsymbol{\theta}}) \sim \chi^2_{L-K} = \chi^2_1$$

Since $L = 3$ and $K = 2$. We then get a value of 0.9875889 for our test statistic which is below the critical 95% quantile of the χ^2_1 distribution of 3.841 and so we fail the reject the null hypothesis and conclude that the parameters are not overidentified.

(g) In this question we test the hypothesis $H_0 : \mu = \sigma^2$. To do so we minimize the same criterion function as in part e)(ii), but now we fix $\mu = \sigma^2$. This means we now have only one parameter to estimate. Using the same two step procedure as in part e)(ii), we obtain $\hat{\boldsymbol{\theta}}_{\boldsymbol{R}}$ and $\hat{\boldsymbol{\Phi}}^{\boldsymbol{-1}}$. We now construct the following test statistic:

$$nq = \sqrt{n}\overline{\mathbf{m}}(\hat{\boldsymbol{\theta}_{\boldsymbol{R}}})^T \hat{\boldsymbol{\Phi}}^{\boldsymbol{-1}} \sqrt{n}\overline{\mathbf{m}}(\hat{\boldsymbol{\theta}_{\boldsymbol{R}}}) \sim \chi^2_{L-(K-J)} = \chi^2_2$$

with $L = 3$ $K = 2$, and $J = 1$. We obtain a value of 52.500 for the test statistic which is above 5.991 the 95% quantile of the χ^2_2 distribution and so we reject the null hypothesis and conclude that $\mu \neq \sigma^2$, which is consistent since in this simulated example we have not set both parameters to the same value.

(h) We test the efficiency of the overidentified GMM estimator when using different weight matrices. We will in one case use $\boldsymbol{W}_1 = \boldsymbol{I}$ and in another $\boldsymbol{W}_2 = \hat{\boldsymbol{\Phi}}^{-1}$. We obtain both set of estimators by doing each time the single two step procedure described above. After the first step we have the estimators for the first case, and after the second step we have our estimators for the second case. We simulate $R = 1000$ times different samples of lengths n=20,200,2000,20000 and we take the mean of the estimators obtained in each experiment for each case (once with W_1, once with W_2). Doing 1000 experiments for each different allows us to have more stable values and get a better picture of the convergence taking place.

We summarize our findings in the following table:

n	μ	σ^2
20	1.252020	0.6283610
200	1.218505	0.7703820
2000	1.201337	0.7947538
20000	1.202017	0.7980583

Table 1: For $\boldsymbol{W} = \boldsymbol{I}$

n	μ	σ^2
20	1.196316	0.5762504
200	1.206702	0.7451508
2000	1.199337	0.7914557
20000	1.200616	0.7994254

Table 2: For $\boldsymbol{W} = (\hat{\boldsymbol{\Phi}})^{-1}$

We see that in both cases when the sample size increases our estimators converge to the true value of μ and σ^2. This is because both $\hat{\boldsymbol{\Phi}}^{-1}$ and $\boldsymbol{I}$ are positive definite matrices and consequently our estimation procedure yields in both cases consistent estimators. We also see that by using $\hat{\boldsymbol{\Phi}}^{-1}$ when estimating our parameters converge more quickly to the correct value. This is because $\hat{\boldsymbol{\Phi}}^{-1}$ is an estimator of the optimal weighing matrix.