UNIVERSITY OF CAMBRIDGE
FACULTY OF ECONOMICS

M.Phil. in Economics
M.Phil. in Economic Research

Subject M300 Econometric Methods

Exercise Sheet 3

1. A random sample (y_i), $(i = 1, ..., n)$, is taken from the exponential distribution
$$f(y; \theta_0) = \frac{1}{\theta_0} \exp(-y/\theta_0).$$

   (a) What is the maximum likelihood (ML) estimator of θ_0?

   Suggested solution to (a) The joint density of $y_1, ..., y_n$ evaluated at the actual data values is
$$\prod_{i=1}^{n} f(y_i; \theta_0) = \prod_{i=1}^{n} \frac{1}{\theta_0} \exp(-y_i/\theta_0) = \frac{1}{\theta_0^n} \exp(-\sum_{i=1}^{n} y_i/\theta_0).$$
   Therefore, the log-likelihood function (well-defined for $\theta > 0$) is
$$\begin{aligned} \log L(\theta; y) &= \log \left(\frac{1}{\theta^n} \exp(-\sum_{i=1}^{n} y_i/\theta) \right) \\ &= -\sum_{i=1}^{n} y_i/\theta - n \log \theta. \end{aligned}$$
   Differentiating with respect to θ, we get
$$\frac{d}{d\theta} \log L(\theta; y) = \sum_{i=1}^{n} y_i/\theta^2 - n/\theta.$$
   Note that $\sum_{i=1}^{n} y_i/\theta^2 - n/\theta > 0$ if $\theta < \sum_{i=1}^{n} y_i/n = \bar{y}$ and $\sum_{i=1}^{n} y_i/\theta^2 - n/\theta < 0$ if $\theta > \bar{y}$. Hence, the maximum of $\log L(\theta; y)$ is achieved at $\bar{y}$, and
$$\hat{\theta}_{ML} = \bar{y}.$$

   (b) Obtain the (scalar) information matrix. [*Hint*: $E[y] = \theta_0$ and $var[y] = \theta_0^2$.]

Suggested solution to (b) Let us differentiate $\log L(\theta; y)$ twice. We get

$$\frac{d^2}{d\theta^2} \log L(\theta; y) = -2 \sum_{i=1}^{n} y_i/\theta^3 + n/\theta^2.$$

We have

$$\begin{aligned}
\mathcal{I}(\theta_0) &= -E\left(\frac{d^2}{d\theta^2} \log L(\theta_0; y)\right) = -\left(-2\sum_{i=1}^{n} Ey_i/\theta_0^3 + n/\theta_0^2\right) \\
&= -\left(-2\sum_{i=1}^{n} \theta_0/\theta_0^3 + n/\theta_0^2\right) = -\left(-2\frac{n}{\theta_0^2} + \frac{n}{\theta_0^2}\right) = \frac{n}{\theta_0^2}
\end{aligned}$$

(c) Find the asymptotic distribution of the ML estimator of θ_0.

Suggested solution to (c) We know that

$$\sqrt{n}\left(\hat{\theta}_{ML} - \theta_0\right) \xrightarrow{d} N\left(0, n\mathcal{I}(\theta_0)^{-1}\right).$$

But $\mathcal{I}(\theta_0) = \frac{n}{\theta_0^2}$. Therefore,

$$\sqrt{n}\left(\hat{\theta}_{ML} - \theta_0\right) \xrightarrow{d} N\left(0, \theta_0^2\right).$$

2. The conditional distribution of y given x is normal with mean $x'\beta_0$ and variance σ_0^2, i.e.,

$$y|x \sim N(x'\beta_0, \sigma_0^2).$$

Suppose that β_0 is known and a random sample $\{(y_i, x_i), i = 1, ..., n\}$ is available

(a) Show that the maximum likelihood estimator of σ_0^2 is $\hat{\sigma}^2 = \sum_{i=1}^{n}(y_i - x_i'\beta_0)^2/n$.

Suggested solution to (a) The joint conditional density of $y_1, ..., y_n$ given $x_1, ..., x_n$ evaluated at the observed values of the data equals

$$\begin{aligned}
\prod_{i=1}^{n} f\left(y_i|x_i; \beta_0, \sigma_0^2\right) &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma_0} \exp\left\{-\frac{(y_i - x_i'\beta_0)^2}{2\sigma_0^2}\right\} \\
&= \frac{1}{(2\pi)^{n/2}(\sigma_0^2)^{n/2}} \exp\left\{-\frac{\sum_{i=1}^{n}(y_i - x_i'\beta_0)^2}{2\sigma_0^2}\right\}.
\end{aligned}$$

Therefore, the log-likelihood function

$$\begin{aligned}
\log L\left(\sigma^2; y, x\right) &= -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\left(\sigma^2\right) \\
&\quad -\frac{\sum_{i=1}^{n}(y_i - x_i'\beta_0)^2}{2\sigma^2},
\end{aligned}$$

where β_0 is a known vector, by assumption. Differentiating the lod-likelihood with respect to σ^2, we get

$$\frac{d}{d\sigma^2}\log L\left(\sigma^2; y, x\right) = -\frac{n}{2\sigma^2} + \frac{\sum_{i=1}^n \left(y_i - x_i'\beta_0\right)^2}{2\sigma^4}$$

Equating to zero, and solving, we get

$$\hat{\sigma}^2_{ML} = \frac{\sum_{i=1}^n \left(y_i - x_i'\beta_0\right)^2}{n}$$

(b) What is the (conditional) Cramér-Rao lower bound for unbiased estimation of σ_0^2?

Suggested solution to (b) The conditional Cramer-Rao lower bound equals the inverse of the conditional Fisher information matrix. Differentiating $\log L\left(\sigma^2; y, x\right)$ twice, we get

$$\frac{d^2}{d\left(\sigma^2\right)^2}\log L\left(\sigma^2; y, x\right) = \frac{n}{2\sigma^4} - 2\frac{\sum_{i=1}^n \left(y_i - x_i'\beta_0\right)^2}{2\sigma^6}$$

Therefore

$$\begin{aligned} -E\left(\frac{d^2}{d\left(\sigma^2\right)^2}\log L\left(\sigma_0^2; y, x\right)\right) &= -\frac{n}{2\sigma_0^4} + 2\frac{\sum_{i=1}^n E\left(y_i - x_i'\beta_0\right)^2}{2\sigma_0^6} \\ &= -\frac{n}{2\sigma_0^4} + 2\frac{n\sigma_0^2}{2\sigma_0^6} = \frac{n}{2\sigma_0^4}, \end{aligned}$$

and hence, $\mathcal{I}\left(\sigma_0^2\right) = \frac{n}{2\sigma_0^4}$ and the Cramer-Rao lower bound is

$$CR = \frac{2\sigma_0^4}{n}$$

(c) Describe the likelihood ratio test for the null hypothesis $H_0 : \sigma^2 = \sigma_0^2$.

Suggested solution to (c) Setting $\sigma^2 = \hat{\sigma}^2_{ML} = \frac{\sum_{i=1}^n \left(y_i - x_i'\beta_0\right)^2}{n}$ in

$$\begin{aligned} \log L\left(\sigma^2; y, x\right) &= -\frac{n}{2}\log\left(2\pi\right) - \frac{n}{2}\log\left(\sigma^2\right) \\ &\quad - \frac{\sum_{i=1}^n \left(y_i - x_i'\beta_0\right)^2}{2\sigma^2}, \end{aligned}$$

we get

$$\log L\left(\hat{\sigma}^2_{ML}; y, x\right) = -\frac{n}{2}\log\left(2\pi\right) - \frac{n}{2}\log\left(\frac{\sum_{i=1}^n \left(y_i - x_i'\beta_0\right)^2}{n}\right) - \frac{n}{2}.$$

Under the null,

$$\log L\left(\sigma_0^2; y, x\right) = -\frac{n}{2}\log\left(2\pi\right) - \frac{n}{2}\log\left(\sigma_0^2\right) - -\frac{\sum_{i=1}^n \left(y_i - x_i'\beta_0\right)^2}{2\sigma_0^2}.$$

The likelihood ratio statistic

$$LR = 2\left(\log L\left(\hat{\sigma}^2_{ML}; y, x\right) - \log L\left(\sigma_0^2; y, x\right)\right).$$

I would reject the null if LR is larger than a critical value of the $\chi^2(1)$ (one degree of freedom here corresponds to testing just one restriction).

3. The Pareto distribution is sometimes used to characterize the distribution of incomes above a certain threshold value $\underline{y} > 0$. The probability density function of the Pareto distribution is

$$f(y; \alpha_0) = \alpha_0 \underline{y}^{\alpha_0} / y^{\alpha_0 + 1}, \alpha_0 > 0, \underline{y} \leq y \leq \infty.$$

(a) i. Show that the maximum likelihood estimator of α_0 obtained from the random sample y_i, $(i = 1, ..., n)$, is

$$\hat{\alpha} = \frac{n}{\sum_{i=1}^n \log(y_i/\underline{y})}.$$

Suggested solution to (ai) The joint density of $y_1, ..., y_n$ evaluated at the observed values of the data is

$$\begin{aligned} \prod_{i=1}^n f(y_i; \alpha_0) &= \prod_{i=1}^n \left(\alpha_0 \underline{y}^{\alpha_0} / y_i^{\alpha_0+1}\right) \\ &= \frac{\alpha_0^n \underline{y}^{n\alpha_0}}{\prod_{i=1}^n y_i^{\alpha_0+1}} \end{aligned}$$

Therefore, the log-likelihood function is

$$\log L(\alpha; y) = n \log(\alpha) + n\alpha \log\left(\underline{y}\right) - (\alpha + 1) \sum_{i=1}^n \log(y_i).$$

Differentiating with respect to α, we get

$$\frac{d}{d\alpha} \log L(\alpha; y) = \frac{n}{\alpha} + n \log\left(\underline{y}\right) - \sum_{i=1}^n \log(y_i)$$

Equating this to zero and solving, we obtain

$$\begin{aligned} \hat{\alpha} &= \frac{n}{\sum_{i=1}^n \log(y_i) - n \log\left(\underline{y}\right)} \\ &= \frac{n}{\sum_{i=1}^n \log\left(y_i/\underline{y}\right)} \end{aligned}$$

ii. How would you estimate the variance of $\hat{\alpha}$?
Suggested solution to (aii) To estimate the variance of $\hat{\alpha}$, I would compute the Fisher information matrix at $\hat{\alpha}$, and invert

it. This will approximate the variance of $\hat{\alpha}$. Specifically, differentiating $\log L(\alpha; y)$ twice with respect to α, we get

$$\frac{d^2}{d\alpha^2} \log L(\alpha; y) = -\frac{n}{\alpha^2}$$

Hence

$$\mathcal{I}(\hat{\alpha}) = \frac{n}{\hat{\alpha}^2}$$

and the variance of $\hat{\alpha}$ can be approximated by $\frac{\hat{\alpha}^2}{n} = \frac{1}{\sum_{i=1}^{n} \log(y_i/\underline{y})}$.

(b) How would you test the null hypothesis $H_0 : \alpha = \alpha_0$ using

i. a Wald test;
Suggested solution to (bi). To perform the Wald test, I would compute W statistic

$$\begin{aligned} W &= n(\hat{\alpha} - \alpha_0)' \, n^{-1} \widehat{\mathcal{I}}(\hat{\alpha}) (\hat{\alpha} - \alpha_0) \\ &= n \frac{(\hat{\alpha} - \alpha_0)^2}{\hat{\alpha}^2} \end{aligned}$$

and compare it with a critical value of the $\chi^2(1)$ distribution (here one degree of freedom is used because there is only one restriction imposed by the null hypothesis). If W is above the critical value, I would reject the null. (Alternatively, use $\mathcal{I}(\alpha_0) = \frac{n}{\alpha_0^2}$ instead of $\widehat{\mathcal{I}}(\hat{\alpha})$ in the formula for W).

ii. a Lagrange multiplier test;
Suggested solution to (bii) To perform the Lagrange Multiplier test, I would compute LM statistic

$$\begin{aligned} LM &= \frac{d}{d\alpha'} \log L(\alpha_0; y) \left(\widehat{\mathcal{I}}(\alpha_0)\right)^{-1} \frac{d}{d\alpha} \log L(\alpha_0; y) \\ &= \left(\frac{n}{\alpha_0} + n \log(\underline{y}) - \sum_{i=1}^{n} \log(y_i)\right)^2 \frac{\alpha_0^2}{n} \\ &= \left(\frac{n}{\alpha_0} - \sum_{i=1}^{n} \log\left(\frac{y_i}{\underline{y}}\right)\right)^2 \frac{\alpha_0^2}{n} \end{aligned}$$

and compare it with a critical value of the $\chi^2(1)$ distribution. If LM is above the critical value, I would reject the null.

iii. a likelihood ratio test?
Suggested solution to (biii) To perform the Likelihood Ratio

test, I would compute LR statistic

$$\begin{aligned}
LR &= 2\left(\log L\left(\hat{\alpha}; y\right) - \log L\left(\alpha_0; y\right)\right) \\
&= 2n\log\left(\hat{\alpha}\right) + 2n\hat{\alpha}\log\left(\underline{y}\right) - 2\left(\hat{\alpha}+1\right)\sum_{i=1}^{n}\log\left(y_i\right) \\
&\quad -2n\log\left(\alpha_0\right) - 2n\alpha_0\log\left(\underline{y}\right) + 2\left(\alpha_0+1\right)\sum_{i=1}^{n}\log\left(y_i\right) \\
&= 2n\log\left(\frac{\hat{\alpha}}{\alpha_0}\right) + 2\left(n\log\left(\underline{y}\right) - \sum_{i=1}^{n}\log\left(y_i\right)\right)\left(\hat{\alpha}-\alpha_0\right) \\
&= 2\left(n\log\left(\frac{\hat{\alpha}}{\alpha_0}\right) - \sum_{i=1}^{n}\log\left(\frac{y_i}{\underline{y}}\right)\left(\hat{\alpha}-\alpha_0\right)\right)
\end{aligned}$$

and compare it with a critical value of the $\chi^2(1)$ distribution. If LR is above the critical value, I would reject the null.

iv. Compare and contrast the three tests.
Suggested solution to (biv) For this particular problem, LM test does not require any estimation because the null hypothesis completely specify the parameter. No estimation means less work, which is nice. In general, LM and W are not invariant under re-parameterization. LR is invariant under reparametrization. It also may have better finite sample properties. Asymptotically, all three tests are equivalent.

4. Consider a linear regression

$$y_i = \beta x_i + u_i,$$

where x_i is a scalar explanatory variable distributed as $N(0,1)$ and $u_i|x_i \sim N(0, x_i^2)$. The constant is not included in the regression for simplicity. Assume that (y_i, x_i) are i.i.d.

(a) Show that the following moment conditions are satisfied

$$\begin{aligned}
E\left(u_i x_i\right) &= 0, \\
E\left(u_i x_i^3\right) &= 0.
\end{aligned}$$

Suggested solution to (a) Since $E(u_i|x_i) = 0$, u_i has zero unconditional mean and is uncorrelated to any functions of x_i, including x_i and x_i^3. This implies that the above moment conditions are satisfied.

(b) Find the optimal weighting matrix Ω^{-1} in the GMM based on the above two moment conditions. (Hint: the $(2k)$-th raw moment of $N(0,1)$ is $(2k)!/\left(k!2^k\right)$)

Suggested solution to (b) We have

$$\Omega = E\left(\left(\begin{array}{c} u_i x_i \\ u_i x_i^3 \end{array}\right)\left(\begin{array}{cc} u_i x_i & u_i x_i^3 \end{array}\right)\right) = E\left(\begin{array}{cc} u_i^2 x_i^2 & u_i^2 x_i^4 \\ u_i^2 x_i^4 & u_i^2 x_i^6 \end{array}\right).$$

Using the law of iterated expectations, we get

$$\begin{aligned}
\Omega &= E\left(\begin{array}{cc} E\left(u_i^2|x_i\right) x_i^2 & E\left(u_i^2|x_i\right) x_i^4 \\ E\left(u_i^2|x_i\right) x_i^4 & E\left(u_i^2|x_i\right) x_i^6 \end{array}\right) \\
&= E\left(\begin{array}{cc} x_i^4 & x_i^6 \\ x_i^6 & x_i^8 \end{array}\right) = \left(\begin{array}{cc} (4)!/\left(2!2^2\right) & (6)!/\left(3!2^3\right) \\ (6)!/\left(3!2^3\right) & (8)!/\left(4!2^4\right) \end{array}\right) \\
&= \left(\begin{array}{cc} 3 & 15 \\ 15 & 105 \end{array}\right).
\end{aligned}$$

Therefore,

$$\Omega^{-1} = \frac{1}{90}\left(\begin{array}{cc} 105 & -15 \\ -15 & 3 \end{array}\right) = \frac{1}{30}\left(\begin{array}{cc} 35 & -5 \\ -5 & 1 \end{array}\right).$$

(c) Compute the variance of the asymptotic distribution of the optimal $\hat{\beta}_{GMM}$.

Suggested solution to (c) The variance of the asymptotic distribution of the optional $\hat{\beta}_{GMM}$ equals $\left(D'\Omega^{-1}D\right)^{-1}$, where

$$D = E\frac{d}{d\beta}\left(\begin{array}{c} \left(y_i - \beta x_i\right) x_i \\ \left(y_i - \beta x_i\right) x_i^3 \end{array}\right) = -E\left(\begin{array}{c} x_i^2 \\ x_i^4 \end{array}\right) = -\left(\begin{array}{c} 1 \\ 3 \end{array}\right).$$

That is,

$$\sqrt{n}\left(\hat{\beta}_{GMM} - \beta\right) \xrightarrow{d} N\left(0, \left(D'\Omega^{-1}D\right)^{-1}\right).$$

We have,

$$D'\Omega^{-1}D = \left(\begin{array}{cc} 1 & 3 \end{array}\right)\frac{1}{30}\left(\begin{array}{cc} 35 & -5 \\ -5 & 1 \end{array}\right)\left(\begin{array}{c} 1 \\ 3 \end{array}\right) = \frac{14}{30}.$$

Hence,

$$\sqrt{n}\left(\hat{\beta}_{GMM} - \beta\right) \xrightarrow{d} N\left(0, 30/14\right)$$

(d) What is the asymptotic distribution of the OLS estimator of β? Compare with (c).

Suggested solution to (d) For OLS, we have

$$\sqrt{n}\left(\hat{\beta}_{OLS} - \beta\right) \xrightarrow{d} N\left(0, E\left(x_i^2\right)^{-1}\left(Eu_i^2 x_i^2\right) E\left(x_i^2\right)^{-1}\right)$$

Since $E\left(x_i^2\right) = 1$ and $Eu_i^2 x_i^2 = E\left(E\left(u_i^2|x_i\right) x_i^2\right) = Ex_i^4 = 3$, we get

$$\sqrt{n}\left(\hat{\beta}_{OLS} - \beta\right) \xrightarrow{d} N\left(0, 3\right).$$

The asymptotic variance of OLS is larger than that of GMM. Note that there is heteroskedasticity in this problem. Of course, under homoskedasticity, OLS would be efficient.

5. The file ccapm.dta is the STATA file that contains 238 observations on the real USA consumption ratio c_{t+1}/c_t (cratio), the real gross return on Treasury bills R_{t+1} (rrate), and the real value weighted returns e_{t+1} (e). This is the adjusted Hansen and Singleton (1982) data set, used in their paper "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models" (Econometrica 50, 1269-1286). Consider the first order conditions of the Consumption Capital Asset Pricing Model

$$E_t\left[\beta\left(c_{t+1}/c_t\right)^{-\gamma}R_{t+1}-1\right]=0.$$

The parameters are the discount factor, β, and the relative risk aversion coefficient γ.

(a) Estimate parameters β and γ by GMM using moment conditions

$$\begin{aligned} E\left[\beta\left(c_{t+1}/c_t\right)^{-\gamma}R_{t+1}-1\right] &= 0, \\ E_t\left[\left(\beta\left(c_{t+1}/c_t\right)^{-\gamma}R_{t+1}-1\right)c_t/c_{t-1}\right] &= 0, \\ E_t\left[\left(\beta\left(c_{t+1}/c_t\right)^{-\gamma}R_{t+1}-1\right)R_t\right] &= 0, \\ E_t\left[\left(\beta\left(c_{t+1}/c_t\right)^{-\gamma}R_{t+1}-1\right)e_t\right] &= 0. \end{aligned}$$

Use command gmm in STATA to do the estimation. Specify starting values $\beta=1$ and $\gamma=1$.

Suggested solution to (a) See attached STATA log file

(b) Test the hypothesis that $\beta=0.98$ at 95% significance level.

Suggested solution to (b) See attached STATA log file

(c) What is the value of Hansen's J-statistic? Do you reject the overidentifying restrictions? Explain.

Suggested solution to (c) See attached STATA log file.

Suggested solution to 5a)

```
. gen cratio1=cratio[_n-1]
(1 missing value generated)

. gen rrate1=rrate[_n-1]
(1 missing value generated)

. gen e1=e[_n-1]
(1 missing value generated)

. gmm ({b=1}*(cratio^(-{g=1}))*rrate-1), inst(cratio1
rrate1 e1)

Step 1
Iteration 0:   GMM criterion Q(b) =    .0000144
Iteration 1:   GMM criterion Q(b) =  .00001312
Iteration 2:   GMM criterion Q(b) =  .00001312

Step 2
Iteration 0:   GMM criterion Q(b) =  .00603693
Iteration 1:   GMM criterion Q(b) =  .00534195
Iteration 2:   GMM criterion Q(b) =  .00534195

GMM estimation

Number of parameters =   2
Number of moments    =   4
Initial weight matrix: Unadjusted        Number of obs  =     237
GMM weight matrix:     Robust

-----------------------------------------------------------------
    |               Robust
    |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
----+------------------------------------------------------------
 /b |   .9976933   .0043145   231.24   0.000      .989237     1.00615
 /g |   .7144645   1.760412     0.41   0.685    -2.735879    4.164808
-----------------------------------------------------------------
Instruments for equation 1: cratio1 rrate1 e1 _cons
```

Suggested solution to 5b)

The value of the t-statistic is t=(.9976933-0.98)/.0043145=4.1
This is larger than 1.96 by absolute value. Hence, reject the null at 5% level.

Suggested solution to 5c)

```
. disp e(J)
1.2660431
```

We do not reject the overidentifying restrictions at 5% significance level.
Indeed, under the null, J is distributed as chi2(2). The 95% critical value of this distribution is 5.9915. This is larger than 1.2660431.

6. Suppose that ε_t is strict white noise, that is $\{\varepsilon_t, t = ... - 2, -1, 0, 1, 2, ...\}$ is an i.i.d. sequence such that $E\varepsilon_t = 0$ and $Var(\varepsilon_t) = \sigma^2$. Determine the autocovariance function $\gamma(r,s) = Cov(y_r, y_s)$ of the following time series

$$y_t = 10 + \varepsilon_t + 0.4\varepsilon_{t-1} - 0.2\varepsilon_{t-2}, (t = ... - 2, -1, 0, 1, 2, ...).$$

Suggested solution to 6. We have $y_r = 10 + \varepsilon_r + 0.4\varepsilon_{r-1} - 0.2\varepsilon_{r-2}$ and $y_s = 10 + \varepsilon_s + 0.4\varepsilon_{s-1} - 0.2\varepsilon_{s-2}$. Therefore,

$$\begin{aligned}
\gamma(r,s) &= Cov(y_r, y_s) = Cov(\varepsilon_r + 0.4\varepsilon_{r-1} - 0.2\varepsilon_{r-2}, \varepsilon_s + 0.4\varepsilon_{s-1} - 0.2\varepsilon_{s-2}) \\
&= Cov(\varepsilon_r, \varepsilon_s) + 0.4Cov(\varepsilon_r, \varepsilon_{s-1}) - 0.2Cov(\varepsilon_r, \varepsilon_{s-2}) \\
&\quad +0.4Cov(\varepsilon_{r-1}, \varepsilon_s) + 0.16Cov(\varepsilon_{r-1}, \varepsilon_{s-1}) - 0.08Cov(\varepsilon_{r-1}, \varepsilon_{s-2}) \\
&\quad -0.2Cov(\varepsilon_{r-2}, \varepsilon_s) - 0.08Cov(\varepsilon_{r-2}, \varepsilon_{s-1}) + 0.04Cov(\varepsilon_{r-2}, \varepsilon_{s-2}).
\end{aligned}$$

Since ε_t is strict white noise, $Cov(\varepsilon_t, \varepsilon_{t-k}) = 0$ for all $k \neq 0$. In particular, if $|s-r| > 2$, all the covariances in the second, third and fourth lines of the above displayed equation are zero, and thus, $\gamma(r,s) = 0$. If $s - r = -2$, all the covariances are zero except $Cov(\varepsilon_{r-2}, \varepsilon_s) = \sigma^2$, and thus $\gamma(r,s) = -0.2\sigma^2$. If $s - r = -1$, all the covariances are zero except $Cov(\varepsilon_{r-1}, \varepsilon_s) = Cov(\varepsilon_{r-2}, \varepsilon_{s-1}) = \sigma^2$, and thus $\gamma(r,s) = 0.32$. Similarly, considering the other cases when $|s-r| \leq 2$, we get

$$\gamma(r,s) = \begin{cases} 0 & \text{if } |s-r| > 2 \\ -0.2\sigma^2 & \text{if } |s-r| = 2 \\ 0.32\sigma^2 & \text{if } |s-r| = 1 \\ 1.2\sigma^2 & \text{if } s = r \end{cases}$$

7. The random process $\{y_t, t = 1, 2, ...\}$ is generated by the following equation

$$y_t = \rho y_{t-1} + \varepsilon_t,$$

where $y_0 = 0$ and ε_t, $t = 1, 2, ...$ are i.i.d. normal random variables with mean zero and constant positive variance. The predicted values $\hat{y}_t$ from the OLS regression of y_t, $t = 1, ..., T$, on y_{t-1} are

$$\hat{y}_t = \underset{(0.03)}{0.9}\, y_{t-1},$$

where the standard error is given in the parentheses, and the sample size $T = 250$.

(a) Test the null hypothesis that $\rho = 1$ against the alternative of stationarity at 5% significance level.

Suggested solution to (a) The $t-$statistic for testing the null hypothesis equals $t = \frac{0.9-1}{0.03} \approx -3.33$. The 5% critical value of the Dickey-Fuller distribution is -1.95. Since $t < -1.95$, we reject the null of the unit root.

(b) How would you test the null hypothesis $\rho_1 + \rho_2 = 1$ if the process $\{y_t, t = 1, 2, ...\}$ were generated by equation $y_t = \rho_1 y_{t-1} + \rho_2 y_{t-2} + \varepsilon_t$?

Suggested solution to (b) First, transform the model to the form

$$y_t = (\rho_1 + \rho_2)\, y_{t-1} - \rho_2 \Delta y_{t-1} + \varepsilon_t,$$

where $\Delta y_t = y_t - y_{t-1}$. Using OLS regression for the transformed model, compute the t-statistic for the test of the null $\rho_1 + \rho_2 = 1$. Compare it to -1.95, which is the 5% critical value of the Dickey-Fuller distribution. If $t < -1.95$, reject the null. This is the augmented Dickey-Fuller test.