FACULTY OF ECONOMICS STUDY AIDS 2018

ECT1 Paper 3 Quantitative Methods in Economics

The Faculty Board has agreed to release outline solutions to the 2018 examinations as a study aid for exam revision. They are abridged solutions, and not 'definitive', and should therefore not be considered as an exemplar for 'complete' answers.

Also note that the Faculty will not respond to any queries regarding these solutions.

Revised 28/5/19

Revised 26/3/2020

SECTION A

1. (a) $f(x) = \ln(1-x)$. Differentiating it successively,

$$f^{(n)}(x) = -\frac{(n-1)!}{(1-x)^n} \quad \text{for } n \geq 1$$

Evaluated at $x = 0$, we have

$$f(0) = 0 \qquad \text{and} \quad f^{(n)}(0) = -(n-1)! \quad \text{for } n \geq 1$$

leading to the Taylor series

$$\sum_{n=0}^{\infty} f^{(n)}(0)\frac{x^n}{n!} = -\sum_{n=1}^{\infty} \frac{x^n}{n} = -x - \frac{x^2}{2} - \frac{x^3}{3} - \frac{x^4}{4} - \cdots$$

(b) The nth Taylor polynomial will have its error term bounded by

$$\max_{z\in[0,x]} \left| f^{(n+1)}(z)\frac{x^{n+1}}{(n+1)!} \right| = \max_{z\in[0,x]} \left| \frac{n!}{(1-z)^{n+1}} \frac{x^{n+1}}{(n+1)!} \right| = \frac{1}{n+1}\left(\frac{x}{1-x}\right)^{n+1}$$

$x \in (0, 0.1)$, therefore $x/(1-x) \leq 1/9$. For $n \geq 4$, the above expression is less than 0.0001.

(c) $\ln(0.99) = \ln(1 - 0.01)$. An upper bound on the error term for $x = 0.01$ is

$$\frac{1}{n+1}\left(\frac{0.01}{0.99}\right)^{n+1}$$

For $n = 2$, the above error term is less than 10^{-6}. So, a satisfactory estimate for $\ln(0.99)$ is

$$-0.01 - \frac{0.01^2}{2} = -0.01 - 0.00005 = -0.01005$$

2. (a) $\int xe^{3x}\, dx$. Use integration by parts with $u = x$ and $e^{3x}dx = dv$, so the integral becomes

$$uv - \int vdu = x\frac{1}{3}e^{3x} - \int \frac{e^{3x}}{3}dx = e^{3x}\left(\frac{3x-1}{9}\right)$$

(b) $\displaystyle\int_e^{e^2} \frac{dx}{x(\ln x)^3}$

Use change of variables with $u = \ln x$, so

$$\int_e^{e^2} \frac{dx}{x(\ln x)^3} = \int \frac{du}{u^3} = -\frac{1}{2}u^{-2} = -\frac{1}{2}(\ln x)^{-2}\Big|_e^{e^2} = \frac{1}{2} = \frac{1}{8} = \frac{3}{8}$$

(c)

$$\int \frac{2x}{x^2+6x+5}\, dx = \int \frac{2(x+3)-6}{(x+3)^2-4}\, dx$$

Set $u = x+3$ to simplify it as

$$\begin{aligned}
\int \frac{2u-6}{u^2-4} du &= \int \frac{2udu}{u^2-4} - \int \frac{6du}{u^2-4} \\
&= \ln|u^2-4| - \frac{6}{4} \int \left(\frac{1}{u-2} - \frac{1}{u+2} \right) du \\
&= \ln|x^2+6x+5| - \frac{3}{2} \ln\left|\frac{x+1}{x+5}\right| \\
&= \frac{5}{2} \ln|x+5| - \frac{1}{2} \ln|x+1|
\end{aligned}$$

Alternatively

$$\begin{aligned}
\int \frac{2x}{x^2+6x+5}\, dx &= \int \frac{2x+6}{x^2+6x+5}\, dx - 6 \int \frac{dx}{(x+1)(x+5)} \\
&= \ln|x^2+6x+5| - \frac{6}{4} \int \left(\frac{1}{x+2} - \frac{1}{x+5} \right) dx \\
&= \ln|x^2+6x+5| - \frac{3}{2} \ln\left|\frac{x+1}{x+5}\right| \\
&= \frac{5}{2} \ln|x+5| - \frac{1}{2} \ln|x+1|
\end{aligned}$$

3. (a) The matrix has the columns Te_1, Te_2, Te_3 where e_i are the unit base vectors. Noting that $e_1 = v_1, e_2 = v_2 - 1$, and $e_3 = v_3 - v_2$

$$\begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

(b) T rotates e_1 and e_2 by 90 degrees clockwise on the horizontal plane, and flips e_3 upside down in the third dimension. Hence A^2 would correspond to rotating e_1 and e_2 by 180 degrees clockwise, while fixing e_3

$$A^2 = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Note that A^4 will then be the identity matrix since e_1 and e_2 are rotated 360 degrees in their plane, and e_3 is flipped an even number of times, and hence fixed. Thus $A^5 = A$.

(c) $A^{-1} = A^3$, hence T is non-singular. The matrix representation of T^{-1} is A^3:

$$\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

4. (a) $\max_{K,L}\ f(K,L)p - Kr - Lw \quad = \quad \max_{K,L}\ K^\alpha L^\beta p - Kr - Lw$

(b) $c(q) = \min_{K,L}\ Kr + Lw$ subject to $f(K,L) = q$

(c) The FOC associated with solving

$$\max_{K,L} \pi(K,L) \ = \ \max_{K,L} 4K^{0.25}L^{0.25} - 4K - 0.25L$$

imply $MP_L/w = MP_K/r$, i.e., $f_L/w = f_K/r$, so

$$\frac{0.25K^{0.25}L^{-0.75}}{0.25} = \frac{0.25K^{-0.75}L^{0.25}}{4} \quad \Longrightarrow \quad 16K = L$$

Plugging $L = 16K$ in $\pi_K = K^{-0.75}L^{0.25} - 4 = 0$ we get

$$2K^{-0.5} = 4 \quad \Longrightarrow \quad K^* = 0.25 \qquad L^* = 4$$

resulting in

$$\begin{aligned} \pi^* &= 4(0.25 \times 4)^{0.25} - 0.25 \times 4 - 8 \times 0.25 \\ &= 2 \end{aligned}$$

(d) The approximate monetary benefit of innovations A and B are given by

$$\frac{d\pi^*}{d\alpha} \times 0.01 \qquad \text{and} \qquad \frac{d\pi^*}{d\beta} \times 0.01$$

So, we will compare

$$\frac{d\pi^*}{d\alpha} = \left.\frac{\partial \pi}{\partial \alpha}\right|_{\text{at the optimal solution}} \qquad \text{with} \qquad \frac{d\pi^*}{d\beta} = \left.\frac{\partial \pi}{\partial \beta}\right|_{\text{at the optimal solution}}$$

Partial differentiation with respect to α and β reveals

$$\frac{\partial \pi}{\partial \alpha} = \frac{\partial(K^\alpha L^\beta p - Kr - Lw)}{\partial \alpha} = (\ln K)K^\alpha L^\beta p$$

$$\frac{\partial \pi}{\partial \beta} = \frac{\partial(K^\alpha L^\beta p - Kr - Lw)}{\partial \beta} = (\ln L)K^\alpha L^\beta p$$

Since $L^* = 16K^*$, clearly B improves the profit better.
Plugging in $\alpha = \beta = 0.25$, $w = 0.25$, $r = 4$, $p = 4$, $K^* = 0.25$ and $L^* = 4$, the approximate monetary value of innovation B can be computed as

$$\ln 4 \times (0.25 \times 4)^{0.25} \times 4 \times 0.01 \quad = \frac{\ln 4}{25} \quad \approx \quad 0.055$$

SECTION B

5. (a) The price elasticity of demand is $f'(p)p/f(p)$. The price elasticity of supply is $g'(p)p/g(p)$. These are given to be equal to ap^2 and bp^2, respectively. Hence we have two differential equations

$$\frac{f'(p)p}{f(p)} = ap^2 \qquad \text{and} \qquad \frac{g'(p)p}{g(p)} = bp^2$$

which can be rewritten as

$$\frac{f'(p)}{f(p)} = ap \qquad \text{and} \qquad \frac{g'(p)}{g(p)} = bp$$

Integrating with respect to p, we obtain

$$\ln f(p) \;=\; \frac{ap^2}{2} + \alpha \qquad \text{and} \qquad \ln g(p) \;=\; \frac{bp^2}{2} + \beta$$

where α and β are constants. Taking the exponential yields

$$f(p) \;=\; A\exp(ap^2/2) \qquad \text{and} \qquad g(p) \;=\; B\exp(bp^2/2)$$

where A and B are positive constants.

We are told that $f(1) = g(1) = 100$, which imply

$$A \;=\; 100\exp(-a/2) \qquad \text{and} \qquad B \;=\; 100\exp(-b/2)$$

and therefore

$$f(p) \;=\; 100\exp(-a/2+ap^2/2) \qquad \text{and} \qquad g(p) \;=\; 100\exp(-b/2+bp^2/2)$$

(b) Simply imposing $D_t = S_t$, i.e., $f(p_t) = g(p_{t-1})$ yields

$$\exp\left(-\frac{a}{2} + \frac{ap_t^2}{2}\right) \;=\; \exp\left(-\frac{b}{2} + \frac{bp_{t-1}^2}{2}\right)$$

which implies

$$a(1-p_t^2) \;=\; b(1-p_{t-1}^2)$$

A steady state p^* implies $p_t = p_{t-1} = p^*$, and hence

$$(a-b)(1-(p^*)^2) \;=\; 0$$

Demand is downward sloping, so $a < 0$, whereas supply is upward sloping, so $b > 0$. In particular $a - b \neq 0$, and therefore $(p^*)^2 = 1$, and $p^* = 1$ since price cannot be negative.

(c) In order to make the above difference equation look like a first order linear autonomous difference equation, we can set $q_t = 1 - p_t^2$, so

$$aq_t = bq_{t-1}$$

and therefore $q_t = (b/a)^t q_0$, which means

$$1 - p_t^2 = (b/a)^t (1 - p_0^2)$$

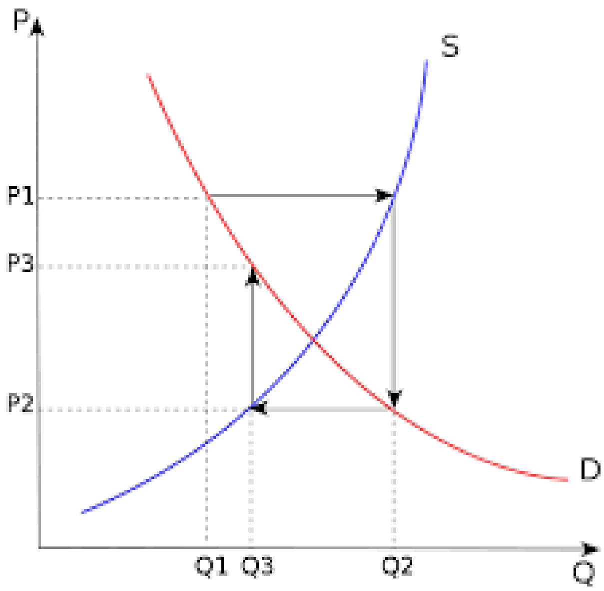
and therefore

$$p_t = \left(1 - \left(\frac{b}{a}\right)^t (1 - p_0^2)\right)^{1/2} = \left(\left(\frac{b}{a}\right)^t p_0^2 + 1 - \left(\frac{b}{a}\right)^t\right)^{1/2}$$

(d) p_t converges if and only if $|b/a| < 1$ and the limit is the unique stationary point $p^* = 1$.

The stationary price $p^* = 1$ is stable if there is convergence $p_t \to p^*$, and this happens if and only if $|b| < |a|$. The demand and the supply functions have the "same shape", and when $|b| < |a|$, demand is more elastic than supply. If supply were to be more elastic than demand ($|a| < |b|$) the prices would go further and further away from p^* as $t \to \infty$.

Note also that $a < 0 < b$ since we are talking about an ordinary good: demand is downward sloping, supply is upward sloping. The price will fluctuate about p^*. For $b < |a|$, the movement of price-quantity pairs on supply and demand curves lead to the so-called cobweb picture spiralling around the intersection point of the demand and supply curves (which happens at $(1, 100)$.)



6. (a) The FOC for f are

$$\begin{aligned} f_x &= 4x - 4y - 6 &= 0 \\ f_y &= 10y - 4x + 4 &= 0 \end{aligned}$$

which lead to a unique solution $(7/6, 1/3)$.

The Hessian matrix is $\begin{bmatrix} 4 & -4 \\ -4 & 10 \end{bmatrix}$ which is positive definite everywhere. Thus f is convex on $\mathbb{R}^2$, and therefore the unique stationary point $(7/6, 1/3)$ is the global minimum.

Moreover, f does not admit a maximum. If it did, the FOC would identify that point since all points in the domain are interior, and f is differentiable.

(b) D is a compact set. This is because $(x-y)^2 + \frac{3}{2}y^2 \leq 87$ implies that $|y| < 10$ and $|x - y| < 10$ for all points $(x, y) \in D$. That means $|x| < 20$ for all such points. Thus D is bounded. It contains the boundary points (thanks to the inequality being weak), thus D is closed. Since f is continuous, it admits both its maximum and minimum values on D.

(c) The point $(7/6, 1/3)$, where f takes its minimum value over $\mathbb{R}^2$ also lies in D, hence it is also the minimum point in domain D.

We know f satisfies the FOC only at one point where it achieves its minimum value. Hence, its maximum value on D cannot be in the interior. (For otherwise, it ill have to satisfy the FOC.) Hence, its maximum value on D is achieved on the boundary described by $(x - y)^2 + \frac{3}{2}y^2 = 87$.

In order to identify the maximum point in D, set up the Lagrangian

$$\mathcal{L} = f(x, y) + \lambda \left[(x - y)^2 + \frac{3}{2}y^2 - 87 \right]$$

which lead to the FOC

$$\begin{aligned} \mathcal{L}_x : \quad & 4x - 4y - 6 + \lambda[2x - 2y] = 0 \implies (2x - 2y)(2 + \lambda) = 6 \\ \mathcal{L}_y : \quad & 10y - 4x + 4 + \lambda[5y - 2x] = 0 \implies (5y - 2x)(2 + \lambda) = -4 \\ \mathcal{L}_\lambda : \quad & (x - y)^2 + \frac{3}{2}y^2 - 87 = 0 \end{aligned}$$

First two equations imply $\frac{x-y}{5y-2x} = -\frac{3}{4}$ which yield $2x = 11y$, and hence $x = \frac{11}{2}y$. Plugging this into the third equation

$$\frac{81}{4}y^2 + \frac{3}{2}y^2 = 87 \implies y^2 = 4 \implies y = -2 \text{ or } y = 2$$

So, we have identified $(-11, -2)$ and $(11, 2)$ as the candidate points where f might take its maximum value on domain D. Not hard to see that it will achieve a higher value at $(-11, -2)$.

SECTION C

7. (a) $f(X) = 2x \quad 0 \leq x \leq 1$ and $f(Y) = \frac{1+y}{4} \quad 0 \leq y \leq 2$

   (b) $E(X) = \int_0^1 x.2x dx = \left[\frac{2x^3}{3}\right]_0^1 = \frac{2}{3}$. $E(Y) = \int_0^2 y.\frac{(1+y)}{4} dy = \left[\frac{y^2}{8} + \frac{y^3}{12}\right]_0^2 = \frac{28}{24}$.

   (c) $f_{XY}(x, y) = f_X(x) f_Y(y)$ so indep

8. (a) 56%

   (b) Let A be the random variable that takes on two values with $a1$ being the event "bag with mostly white balls chosen" and $a2$ the event "bag with mostly black balls chosen" . We know that $P(a1) = P(a2) = \frac{1}{2}$ since we choose the bag at random. Let B be the event "4 white balls and one black ball chosen from 5 selections". Now

$$P(B \mid a1) = ^5C_1 \left(\frac{3}{4}\right)^4 \left(\frac{1}{4}\right) = \frac{405}{1024}$$

and

$$P(B \mid a2) = ^5C_1 \left(\frac{3}{4}\right) \left(\frac{1}{4}\right)^4 = \frac{15}{1024}$$

then

$$\begin{aligned} P(a1 \mid B) &= \frac{P(B \mid a1) P(a1)}{P(B)} \\ &= \frac{P(B \mid a1) P(a1)}{P(B \mid a1) P(a1) + P(B \mid a2) P(a2)} \\ &= \frac{405}{420} = 0.964 \end{aligned}$$

9. Mainly bookwork

   (a) Divide Ω into A and R then do not reject H_0 if $t \in A$ and reject H_0 if $t \in R$. We require $A \cap R = \phi$ and $A \cup R = \Omega$. Ensures we always either reject or not reject H_0

   (b) Type I is reject H_0 when H_0 true and Type II is do not reject H_0 when H_1 true. Size is Prob(Type I error) and power is $1 - Prob$ (Type II error)

   (c) If we set size=0 (ie never reject null) we can expect severe Type II errors.

   (d) For n independent tests probability of not rejecting each correct null is $(1 - \alpha)$ so prob of not rejecting any of them is $(1 - \alpha)^n$. Hence prob of rejecting at least one correct null is $1 - (1 - \alpha)^n$. For small α this is approximately $n\alpha$. So if we reduce the size of each test to α/n the overall size (ie the probability we do not reject any null when they are all true) will be approximately $1 - \left(1 - n\left(\frac{\alpha}{n}\right)\right) = \alpha$

10. (a) Consider the regression of wages on union status.

$$\hat{wage} = \underset{(0.245)}{8.659} + \underset{(0.573)}{2.140}union$$

What does the coefficient on union mean?

ANSWER: The coefficient on union is the difference in sample means between unionized and non-unionized workers. Unionized workers (for whom union=1) make \$2.14 more than non-unionized workers per hour.

(b) Test the null hypothesis that on average, the wages of unionized and non-unionized workers are the same, against the alternative that they are not.

ANSWER: To test this hypothesis we look at the t statistic of the coefficient on union. Since this statistic is equal to $2.140/0.573 = 3.73$, which in absolute value > 1.96, we reject the null hypothesis and conclude that the wages of unionized and non-unionized workers are, on average, different.

(c) Next, consider the regression of wages on union status and experience below:

$$\hat{wage} = \underset{(.394)}{8.130} + \underset{(0.577)}{2.015}union + \underset{(0.018)}{0.031}exper$$

Based on this regression, write the expression for the fitted least squares line for unionized workers.

ANSWER:

$$\hat{wage} = 10.14 + 0.03X_2$$

Similarly write the expression for the fitted least squares line for non-unionized workers.

ANSWER:

$$\hat{Y} = 8.13 + 0.03exper$$

(d) If two people have the same labor market experience, but one is unionized and the other is not, what is their predicted average wage difference?

ANSWER: Given that these two people have the same value for exper, their predicted wages differ by the intercepts of the two fitted lines, (i.e., 10.14 and 8.13). The unionized worker earns, on average, \$ 2.01 per hour more than the non-unionized worker in this sample. Notice that this predicted average wage differential is just the coefficient on the union dummy.

(e) List the assumptions under which the least squares estimators are unbiased. Do you think the coefficient estimate on the union variable in part (a) is unbiased? Explain why or why not.

ANSWER - Bookwork

The assumptions of the classical regression model: for a bivariate population

(i) the conditional expectation function is linear, $E(Y|X) = \alpha + \beta X$ with constant conditional variances $V(Y|X) = \sigma^2$

(ii) we draw stratified random sample, with stratification on fixed X's, (i.e, we have a fixed sample of X's and we randomly draw Y's for each subpopulation)

(iii) $E(Y_i) = \alpha + \beta X_i$

(iv) $V(Y_i) = \sigma^2$

(v) $C(Y_h, Y_i) = 0$ (i.e., the draws on Y_i are independent of each other) for $i = 1, 2, ..., n$.

The estimate of the coefficient on union in the short regression is likely to be biased.

Little coverage of beyond bivariate regression but good students will be able to say something sensible here

Suppose we have observations (Y, X_1, X_2) and the true population model that describes their relationship is

$$E(Y|X_1, X_2) = \beta_0 + \beta_1 X_1 + \beta_2 X_2$$

$$\begin{aligned} \text{Define} \hat{Y} &= b_0 + b_1 X_1 + b_2 X_2 && \text{as the long regression} \\ \tilde{Y} &= a + b X_1 && \text{as the short regression, and} \\ \hat{X}_2 &= c_0 + c X_1 && \text{as the auxiliary regression.} \end{aligned}$$

The omitted variable rule implies that the slope parameter in the short regression, b, is a mixture of the parameters in the long regression, b_1 and b_2.

$$b = b_1 + c b_2$$

Thus, b is a biased estimator for the true population parameter β_1.

SECTION D

11. (a) $P(X_1 = x, X_2 \leq x) = P(X_1 = x) P(X_2 \leq x) = f(x)F(x)$ by independence. So $P(max\{X_1, X_2\} = x) = P(X_1 = x, X_2 \leq x) + P(X_1 \leq x, X_2 = x) = 2f(x)F(x)$

(b) $F(x) = x \quad 0 \leq x \leq 1$ hence pdf of Y is $f(y) = 2y \quad 0 \leq y \leq 1$

(c) $E(Y) = \int_0^1 yf(y)dy = \int_0^1 2y^2 dy = \left[\frac{2y^2}{3}\right]_0^1 = \frac{2}{3}$. $E(Y^2) = \int_0^1 y^2 f(y)dy = \int_0^1 2y^3 dy = \left[\frac{2y^4}{4}\right]_0^1 = \frac{2}{4}$ and the result follows.

(d) For each student their score is a random variable with $E(Y) = \frac{2}{3}$ and $Var(Y) = \frac{1}{18}$. If scores are independent and we add 90 of these we get (by central limit theorem) an approximately normally distributed random variable $N\left(90 \times \frac{2}{3}, \frac{90}{18}\right) = N(60, 5)$. The probability this exceds 60 is 50% and the probability it exceeds 65 is $1 - \Phi\left(\sqrt{5}\right)$ which is 1.27%

(e) In this situation $Y = max\{X_1, X_2 \ldots, X_N\}$ will have pdf $Ny^{N-1} \quad 0 \leq y \leq 1$ and so $E(Y) = \frac{N}{N+1}$ and $Var(Y) = \frac{N}{N+2} - \left(\frac{N}{N+1}\right)^2 = \frac{N}{(N+2)(N+1)^2}$ so the sum of their scores will be (approximately) distributed $N\left(90 \times \frac{N}{N+1}, \frac{90N}{(N+2)(N+1)^2}\right)$ so prob sum is below 60 or below 65 tends to zero as N increases.

12. (a) From Lecture notes

Define $e_i = Y_i - \widehat{Y}_i$

The least squares criterion function:

$$\begin{aligned} \min_{\widehat{\alpha}} \text{SSR} &= \min_{\widehat{\alpha}} \sum_{i=1}^{n} e_i^2 \\ &= \min_{\widehat{\alpha}} \sum_{i=1}^{n} (Y_i - \widehat{\alpha})^2 \end{aligned}$$

Set the derivative equal to zero to get the FOCs

$$\begin{aligned} \frac{\partial SSR}{\partial \widehat{\alpha}} = 0 &= -2 \sum_{i=1}^{n} e_i \\ 0 &= \sum_{i=1}^{n} e_i \\ 0 &= \sum_{i=1}^{n} (Y_i - \widehat{\alpha}) \\ 0 &= \sum_{i=1}^{n} Y_i - \sum_{i=1}^{n} \widehat{\alpha} \\ 0 &= \sum_{i=1}^{n} Y_i - n\widehat{\alpha} \\ n\widehat{\alpha} &= \sum_{i=1}^{n} Y_i \\ \widehat{\alpha} &= \frac{1}{n} \sum_{i=1}^{n} Y_i \end{aligned}$$

The value of $\widehat{\alpha}$ that minimizes deviations is the sample mean of Y

(b) Standard two sample test

(c) Difference in sample means (bookwork)

(d) $\widehat{\lambda} = \bar{Y}^B$, $\widehat{\beta} = \bar{Y}^A - \bar{Y}^B$

(e) Use of multiple regression to allow for other determinants of Y correlated with D

(f) $\mathrm{Cov}(D, \varepsilon) = 0$, random allocation of individuals to populations A and B