FACULTY OF ECONOMICS STUDY AIDS 2017

ECT1 Paper 3 Quantitative Methods in Economics

The Faculty Board has agreed to release outline solutions to the 2017 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.

SECTION A - Answer **ALL FOUR** questions from this Section.

1. (a) Find the Taylor series expansion for

$$f(x) = x^5 + x - 2$$

centred on $x = 1$.

(b) Let Y_t be national product at time t, C_t be total expenditure and I investment. Suppose $C_t = \alpha + \beta Y_t$, $\dot{Y}_t = \gamma(C_t + I - Y_t)$, I is a constant, and α, β and γ are positive constants.

i. Derive a differential equation for Y_t.

ii. Determine the value of any stationary point.

iii. Find a solution for Y_t if $Y_2 = 3$.

iv. Determine and interpret the conditions under which the system is stable.

2. Let $g(x) = (x-3)(x-2)(x-1)$

(a) Determine the values of x for which $g(x) > 0$, and also those for which $g(x) < 0$.

(b) Expand $g(x)$ expressing it in terms of x^3 and x^2 etc.

(c) Find the second order partial differentials of the function

$$f(x, y) = \frac{1}{20}\lambda x^5 - \frac{1}{2}\lambda x^4 + \frac{11}{6}\lambda x^3 - 3\lambda x^2 + 117x + \frac{\alpha}{\lambda}y^2 + 2456$$

(d) Using the results of (a) and (b) (or otherwise) determine values of x, α and λ for which the function $f(x, y)$ given in (c) is convex.

3. (a) "Matrix multiplication is not commutative in general, although it is associative; matrix addition is distributive."

Explain the above sentence. Is it correct?

(b) A and B are square matrices with $AB + BA = A$

Show that:

i. $A^2B - BA^2 = 0$

ii. $A^3B + BA^3 = A^3$

(c) Suppose

$$A = \begin{pmatrix} 3 & 1 & 2 \\ 4 & 0 & 6 \\ 1 & 1 & 1 \end{pmatrix} \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad \mathbf{y} = \begin{pmatrix} 29 \\ 46 \\ 12 \end{pmatrix} \quad \text{and } A\mathbf{x} = \mathbf{y}$$

Using Cramer's rule (or otherwise), determine x_2.

4. Evaluate the following integrals:

(a)

$$\int_0^1 \left(\frac{e^{3x}}{(1+e^x)^2} + \frac{2e^x}{(1+e^x)} \right) dx$$

(b)

$$\int_1^5 (x-1)(5-x)^{\frac{1}{2}}\, dx$$

SECTION B - Answer **ONE** question from this Section.

5. A monopolist can set prices differently in two isolated markets. The demand curves for each market and their total costs are given as:

$$\begin{array}{ll} \text{Market 1} & q_1 = 42.4 - 0.4p_1 \\ \text{Market 2} & q_2 = 20.6 - 0.1p_2 \end{array}$$

and

$$TC = 300 + 6\,(q_1 + q_2)$$

where q_i, and p_i are respectively quantity demanded and price of the commodity provided for market $i = 1, 2$. TC is total cost.

(a) Determine the maximum profit that can be achieved with price discrimination.

(b) Determine the price elasticity of demand in both markets at the profit maximising level of output.

(c) Determine the maximum profit that can be achieved without price discrimination.

(d) Compare and interpret your findings in (a) and (c).

(e) If a temporary quota on production is introduced at $Q = q_1 + q_2 = 8$ determine the maximised profits.

(f) Determine the price elasticity of demand at the point of maximal profits determined in (e).

6. An individual's utility from consuming x_1 units of good number 1 and x_2 units of good number 2 is:

$$U(x_1, x_2) = \alpha \ln(x_1 - a) + \beta \ln(x_2 - b)$$

where a, b, α and β are positive constants, $0 < \alpha,\ \beta < 1$.

(a) Determine the marginal rate of substitution of x_1 for x_2.

(b) If the utility level reached is constrained by the budget equation $px_1 + qx_2 = M$, where p and q are prices of units of goods 1 and 2 respectively, and if $pa + qb < M$, determine the demand curve for each good.

(c) How does an optimised value of utility change if M increases slightly. Simplify your answer as much as possible.

(d) Show that if $\alpha + \beta = 1$ then the demand curves can be expressed as:

$$\begin{aligned} px_1 &= \alpha M + pa - \alpha(pa + qb) \\ qx_2 &= \beta M + qb - \beta(pa + qb) \end{aligned}$$

(e) For either demand curve derived in (d), verify that the price elasticity of demand is negative.

Section C - Answer **ALL FOUR** questions from this Section.

7. X and Y are discrete random variables whose joint probability mass function is given (partially) in the table below. The marginal distributions are such that $P(X=-1) = 1/3$ and $P(Y=-1) = P(Y=0) = 1/4$.

		Y		
		-1	0	$+1$
X	-1	$1/48$		
	$+1$			$1/4$

(a) Complete the table and explain whether X and Y are independent random variables.

(b) Write out the probability mass functions for the random variables $\frac{Y}{X}$ and $\frac{1}{X}$.

(c) Calculate $E(X)$, $E(Y)$, $E\left(\frac{Y}{X}\right)$ and $E\left(\frac{1}{X}\right)$.

(d) Hence or otherwise calculate $Cov\left(Y, \frac{1}{X}\right)$.

8. (a) You purchase a certain product. The manual states that the lifetime, T, of the product, defined as the amount of time (in years) the product works properly until it breaks down, satisfies

$$P(T \geq t) = exp\left(-\frac{t}{5}\right), \text{ for all } t \geq 0$$

If you use it for two years without any problems, what is the probability that it breaks down in the following year (i.e. between $t = 2$ and $t = 3$)?

(b) According to the Centre for Disease Control (CDC), women who smoke are about 13 times more likely to develop lung cancer than women who do not smoke. They also report that in 2015 13.6% of women were smokers. If you learn that a woman has been diagnosed with lung cancer, and you know nothing else about her, what is the probability that she is a smoker?

9. A continuous random variable X has probability density function given by

$$f(x) = \frac{1}{\theta}, \quad \text{for } 0 \leq x \leq \theta$$

where $\theta > 0$ is some parameter.

(a) Show that this function satisfies the definition of a probability density function.

(b) Calculate $E(X)$ and $E\left(X^2\right)$ and hence $Var\left(X\right)$.

(c) What is the cumulative distribution function for the random variable X?

(d) Given an IID random sample $X_1, \ldots, X_n$ of observations from $f(x)$, calculate the expected value and variance of the sample average given by

$$\bar{X} = \frac{1}{N} \sum_{i=1}^{N} X_i$$

(e) Give an unbiased estimator for θ. Explain why it is unbiased and calculate its variance.

10. Consider the linear population regression model $y = \beta_0 + \beta_1 x + u$.

(a) State the Gauss Markov assumptions.

Let $z = log(1 + x^2)$. Define an estimator for β_1 as

$$\tilde{\beta}_1 = \frac{\sum_{i=1}^{N}(z_i - \overline{z})y_i}{\sum_{i=1}^{N}(z_i - \overline{z})x_i}$$

(b) Assume that the Gauss-Markov assumptions hold. Show that $\tilde{\beta}_1$ is an unbiased estimator for β_1.

(c) Next, assume that the assumption of constant conditional variance is violated. Explain how this affects your answer to part (b).

Section D - Answer **ONE** question from this Section.

11. Workers in a firm sell a product. The value of a worker's sales is a random variable X with exponential distribution. The exponential distribution has probability density function $f(x) = \frac{1}{\beta} e^{-x/\beta}$, where $x \geq 0$ and $\beta > 0$ is a parameter.

(a) For the exponential distribution calculate $E(X)$ and $E\left(X^2\right)$ and hence $Var(X)$.

(b) Initially the firm offers a bonus payment to any worker whose sales exceed 2. What fraction of the workforce (as a function of β) will get the bonus?

(c) The firm observes average sales of a sample of 100 workers to be $\bar{x}$. Stating clearly any assumptions you make, what can you say about the distribution of $\bar{x}$?

(d) For a test of a hypothesis H explain what is meant by

i. a Type I and Type II error,

ii. the size and power of the test.

(e) The firm thinks that its workers either all have β equal to $1/3$ or all have β equal to $1/2$. Test, at the 1% level, the null hypothesis that $\beta = 1/3$ against the alternative $\beta = 1/2$ if the realised value of $\bar{x} = 0.433$.

(f) What is the power of this test against the stated alternative?

(g) The firm then decides it should only pay a bonus to 5% of the workers. If half the workforce is of each type, what threshold should be chosen? You may assume without proof that the cubic equation $x^3 - 10x - 10 = 0$ has a single positive solution at $x = 3.578$.

12. Consider the regression $lwage_i = \alpha + \beta_1 educ_i + \varepsilon_i$ where *lwage* is the natural log of the wage, and *educ* is years of education. An economist estimates this regression using a nationally representative sample from the population of a high income country.

Table 1

```
. reg lwage educ

      Source |       SS       df       MS              Number of obs =     864
-------------+------------------------------           F(  1,   862) =   91.61
       Model |  22.2629147     1  22.2629147           Prob > F      =  0.0000
    Residual |   WWWW        862  .243022541           R-squared     =   TTTT
-------------+------------------------------           Adj R-squared =   SSSS
       Total |  231.748345   863  .268538059           Root MSE      =  .49297

------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |    .055974   .0058481      VVVV    XXXX       YYYY        ZZZZ
       _cons |   1.195559   .0771467    15.50   0.000     1.044141    1.346976
------------------------------------------------------------------------------
```

(a) What is the estimate of β_1 in this sample? How do you interpret the estimate?

(b) Calculate VVVV, YYYY, and ZZZZ in the Table above.

(c) Given you answer to part (b), what is a likely value for XXXX? Interpret the p-value associated with the parameter estimate of the coefficient on *educ*.

(d) Calculate WWWW. Explain what Residual SS is.

(e) Provide a mathematical formula for R^2. Calculate its value for this regression.

Next the same economist uses Stata to generate the output in Table 2 below.

Table 2

```
. reg lwage

      Source |       SS       df       MS              Number of obs =     864
-------------+------------------------------           F(  0,   863) =    0.00
       Model |           0     0           .           Prob > F      =       .
    Residual |  231.748345   863  .268538059           R-squared     =  0.0000
-------------+------------------------------           Adj R-squared =  0.0000
       Total |  231.748345   863  .268538059           Root MSE      =  .51821

------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _cons |   1.916288   .0176297   108.70   0.000     1.881686    1.950891
------------------------------------------------------------------------------

. predict lwage_pred
(option xb assumed; fitted values)

. gen lwage_dev=lwage-lwage_pred
```

(f) Write out the regression function for the Stata output from Table 2. Interpret the constant term.

(g) Using the information in Tables 1 and 2, roughly sketch a scatterplot of the residuals from the regression in Table 2 in which the residuals are on the y-axis and year of education is on the x-axis. Explain any pattern one would observe.

The same economist then uses Stata again to generate the output in Table 3 below.

Table 3

```
. reg educ

      Source |       SS       df       MS              Number of obs =     864
-------------+------------------------------           F(  0,   863) =    0.00
       Model |           0     0           .           Prob > F      =      .
    Residual |  7105.74884   863  8.23377618           R-squared     =  0.0000
-------------+------------------------------           Adj R-squared =  0.0000
       Total |  7105.74884   863  8.23377618           Root MSE      =  2.8695

------------------------------------------------------------------------------
        educ |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _cons |   12.87616   .0976209   131.90   0.000     12.68456    13.06776
------------------------------------------------------------------------------

. predict educ_pred
(option xb assumed; fitted values)

. gen educ_dev=educ-educ_pred
```

Finally, the economist estimates

```
. reg lwage_dev educ_dev
```

(h) Explain what the estimated coefficient on *educ_dev* will be. Use the definition of the slope parameter in a bivariate regression to provide a derivation that justifies your answer.

(i) How does the value of R^2 in this last regression compare to that of the R^2 you calculated in part (e) above for the regression of *lwage* on *educ*? Explain.

END OF PAPER

ANSWERS

SECTION A - ANSWERS

A1. a)

$$f(x) \cong 6(x-1) + 10(x-1)^2 + 10(x-1)^3 + 5(x-1)^4 + (x-1)^5$$

i. $\dot{Y} = -\lambda(1-\beta)Y_t + \lambda(\alpha + I)$

ii. $Y^* = \frac{\alpha+I}{1-\beta}$ if $\beta \neq 1$.

iii. $Y_t = (3 - Y^*)e^{-\gamma(1-\beta)(t-2)} + Y^*$

iv. $\beta < 1$ i.e. consumption is less than the national product.

A2. a) i. $g(x) > 0$ when $1 < x < 2$ or $x > 3$; $g(x) < 0$ when $x < 1$ or $2 < x < 3$.

ii. $g(x) = x^3 - 6x^2 + 11x - 6$

iii.

$$\begin{aligned} f''_{xx}(x,y) &= \lambda(x^3 - 6x^2 + 11x - 6) = \lambda g(x) \\ f''_{yy}(x,y) &= \frac{2\alpha}{\lambda}; f''_{xy}(x,y) = 0 \end{aligned}$$

iv. For convexity $f''_{yy}(x,y) \geq 0$, so α and λ have the same sign.
We also need both:

1. determinant of Hessian $= 2\alpha(x^3 - 6x^2 + 11x - 6) = 2\alpha g(x) \geq 0$
and
2. $f''_{xx}(x,y) = \lambda(x^3 - 6x^2 + 11x - 6) = \lambda g(x) \geq 0$.
Using b)(i)
if α and $\lambda > 0$ then we need $1 < x < 2$ or $x > 3$
if α and $\lambda < 0$ then we need $x < 1$ or $2 < x < 3$.

A3. a) Not commutative. It is not always the case that for any 2 matrices, A, B:
$AB = BA$.

Associative. For any matrices, A, B and C:
$A(BC) = (AB)C$.

i.e. the order in which matrices are multiplied can be disregarded.

Distributive. For any matrices, A, B and C:
$A(B + C) = AB + AC$ and $(B + C)A = BA + CA$.

b) There are many ways to derive the results. For example:

$$AB = BA = A \qquad (1)$$

Pre-multiply (1) by A:	$A^2B + ABA = A^2$	(2)
Post-multiply (1) by A:	$ABA + BA^2 = A^2$	(3)
Subtract (3) from (2):	$\mathbf{A^2B - BA^2 = 0}$	(4)
Pre-multiply (4) by A:	$A^3B - ABA^2 = 0$	(5)
Post-multiply (4) by A:	$A^2BA - BA^3 = 0$	(6)
Subtract (6) from (5):	$A^3B + BA^3 - (A^2BA + ABA^2) = 0$	(7)
Post multiply (2) by A:	$A^2BA + ABA^2 = A^3$	(8)
Substitute (8) into (7)	$\mathbf{A^3B + BA^3 = A^3}$	

c) $x_2 = 2$.

A4. i) $1 + e - \frac{1}{1+e} - \frac{3}{2}$

ii) $\frac{128}{15}$

SECTION B - ANSWERS

B.1 i) $q_1^* = 20$; $p_1^* = 56$; $q_2^* = 10$; $p_2^* = 106$; $\pi^* = 1700$.

ii) $\varepsilon_{p1} = -1.12$; $\varepsilon_{p2} = -1.06$.

iii) $Q = 63 - 0.5p$; or $p = 126 - 2Q$

$Q^* = 30$; $p* = 66$; $\pi^* = 1500$

iv) With price discrimination the monopolist has more choice. So we would expect profits to be at least as large. As it happens they appear to increase by 1700-1500 = £200. Q is the sum of two demand curves with different slopes. As a result there will be a kink. Market 1 clearly has the flatter curve (if prices are on the vertical access). When output in Market 1 is 0 the price given by that curve is 106. When price is 106 in Market 2, then output level $q_2 = 20.6 - 0.1 * 106 = 20.6 - 10.6 = 10$. So for output levels below 10, the total demand curve will coincide with the demand curve for market 2. We have found though that an with aggregated demand curve (non-price discrimination) profit maximising output is to the right of the kink level at $Q^* = 30 > 10$.

v) The fixed output level is 8, which is below the level at the kink. So maximised profits will be determined in the demand curve for market 2.

$Q = 8$; $p^* = 126$; $\pi^* = 660$.

vi) Although Q is fixed at 8, there is still a demand curve. If prices were to change then demand would change according to the parameters of the demand in Market 2.

So at $Q = 8$ $\varepsilon_p = 0.1(126/8) = -1.575$.

B2. i)

$$\frac{dx_2}{dx_1} = \frac{\alpha(x_2 - b)}{\beta(x_1 - \alpha)} \quad \text{if } x_1 \neq a$$

ii)

$$x_1^* = \frac{M - qb + pa\frac{\beta}{\alpha}}{p(1 + \frac{\beta}{\alpha})}; \; x_2^* \frac{M - pa + qb\frac{\beta}{\alpha}}{q(1 + \frac{\alpha}{\beta})}$$

iii)

$$\frac{\partial U^*}{\partial M} = \lambda(\text{lagrange multiplier}) \; = \frac{\alpha + \beta}{M - (qb + pa)} > 0$$

iv)

$$\begin{aligned} x_1^* &= \frac{M - qb + pa\frac{\beta}{\alpha}}{p(1 + \frac{\beta}{\alpha})} = \frac{\frac{M\alpha - qb\alpha + pa\beta}{\alpha}}{p(\frac{\alpha+\beta}{\alpha})} = \frac{M\alpha - qb\alpha + pa\beta}{p} \\ &= \frac{M\alpha - qb\alpha + pa(1 - \alpha)}{p} = \frac{M\alpha - pa + \alpha(pa + qb)}{p} \end{aligned}$$

$$\begin{aligned} \text{So } px_1^* &= M\alpha + pa - \alpha(pa + qb) \\ \text{By symmetry } qx_2^* &= M\beta + qb - \beta(pa + qb) \end{aligned}$$

v)

$$\frac{p.\partial x_1^*}{\partial p} + x_1^* = (a - \alpha a)$$

$$\text{Elasticity } \varepsilon = \frac{p.\partial x_1^*}{x_1^* \partial p} = \frac{a(1-\alpha)}{x_1^*} - 1$$

But

$$x_1^* = \frac{M\alpha + pa - \alpha(pa + qb)}{p} > \frac{M\alpha + pa - \alpha(M)}{p} = a$$

$$\varepsilon < \frac{a(1-\alpha)}{a} - 1 = -\alpha \text{ so } \varepsilon < 0.$$

SECTION C

1.

		Y: -1	0	$+1$	
X	-1	$1/48$	$3/48$	$1/4$	$1/3$
	$+1$	$11/48$	$9/48$	$1/4$	$2/3$
		$1/4$	$1/4$	$1/2$	

(a) Not independent as $P\left((X=1, Y=1\right) = 1/4 \neq \frac{2}{3}.\frac{1}{2} = P\left(X=1\right)P\left(Y=1\right)$ etc.

(b) PMFs $\frac{Y}{X}$:

	$Prob$
-1	$23/48$
0	$1/4$
+1	$13/48$

and $\frac{1}{X}$:

	$Prob$
-1	$1/3$
+1	$2/3$

(c) $E\left(X\right) = 1/3$, $E\left(Y\right) = 1/4$, $E\left(\frac{Y}{X}\right) = -10/48$ and $E\left(\frac{1}{X}\right) = 1/3$

(d) No relationships as $E\left(g\left(X\right)\right) \neq g\left(E\left(X\right)\right)$ and $E\left(XY\right) \neq E\left(X\right)E\left(Y\right)$ if X and Y not independent.

(e) $Cov\left(Y, \frac{1}{X}\right) = E\left(Y.\frac{1}{X}\right) - E\left(Y\right)E\left(\frac{1}{X}\right) = -10/48 - 1/4.1/3 = -14/48 = -.2917$

2.

(a) Define events A =breaks in 3rd year and B = does not break in first two years. We seek $P\left(A \mid B\right) = \frac{P(A\&B)}{P(B)}$. Now $P\left(B\right) = e^{-\frac{2}{5}}$ and $P\left(A\&B\right) = P\left(A\right)$ as $A \subset B$ and $P\left(A\right) = P\left(T \geq 2\right) - P\left(T \geq 3\right) = \frac{e^{-\frac{2}{5}} - e^{-\frac{3}{5}}}{e^{-\frac{2}{5}}} = 0.1813$

(b) $Prob\left(smoker \mid cancer\right) = \frac{Prob(cancer|smoker)P(smoker)}{Prob(cancer|smoker)P(smoker)+Prob(cancer|non-smoker)P(non-smoker)}$ and we know $Prob\left(cancer \mid smoker\right) = 13 * P(cancer \mid non-smoker)$ so the answer is $\frac{13\times.136}{13\times.136+(1-.136)}$=.6717 ie about 67%

3.

(a) Everywhere positive and $\int_0^\theta \frac{1}{\theta}dx = \frac{1}{\theta}\left[x\right]_0^\theta = 1$

(b) $E(X) = \int_0^\theta \frac{x}{\theta} dx = \frac{1}{\theta}\left[\frac{x^2}{2}\right]_0^\theta = \frac{\theta}{2}$, $E\left(X^2\right) = \int_0^\theta \frac{x^2}{\theta} dx = \frac{1}{\theta}\left[\frac{x^3}{3}\right] = \frac{\theta^2}{3}$. Hence $Var(X) = \frac{\theta^2}{3} - \left(\frac{\theta}{2}\right)^2 = \frac{\theta^2}{12}$

(c) $F(x) = \begin{cases} 0 & x < 0 \\ \frac{x}{\theta} & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases}$

(d) $E(\bar{x}) = \frac{\theta}{2}$, $Var(\bar{x}) = \frac{\theta^2}{12N}$

(e) Many possible. $2X_1$ will do ($E(2X_1) = 2E(X_1) = 2\frac{\theta}{2} = \theta$).
$Var(2X_1) = 4Var(X_1) = \frac{\theta^2}{3}$. Or $2\bar{X}$ will also do.

4. (a) For a bivariate population, we assume

i. A linear conditional expectation function: $E(Y|X) = \alpha + \beta X$ or linear population regression function: $Y = \alpha + \beta X + u$

ii. The X's are non-stochastic variables whose values are fixed and not all identical.

iii. The expected error is zero. $E(u) = 0$.

iv. Constant conditional variance. $V(Y|X) = \sigma^2$

v. The Y_i's are independent. $C(Y_h, Y_i) = 0$

(b) Since

$$\tilde{\beta}_1 = \frac{\sum_{i=1}^N (z_i - \overline{z})y_i}{\sum_{i=1}^N (z_i - \overline{z})x_i}$$

we have that $\tilde{\beta}_1 = \sum_{i=1}^N w_i y_i$ where $w_i = (z_i - \overline{z})/SST_{zx}$ and $SST_{zx} \equiv \sum_{i=1}^N (z_i - \overline{z})x_i$. It follows that

$$\tilde{\beta}_1 = \sum_{i=1}^N w_i y_i = \sum_{i=1}^N w_i(\beta_0 + \beta_1 x_i + u_i) = \beta_1 + \sum_{i=1}^N w_i u_i$$

Now conditional on the z_i and the x_i in the sample we quickly see

$$E[\tilde{\beta}_1] = E[\beta_1 + \sum_{i=1}^N w_i u_i] = \beta_1 + \sum_{i=1}^N w_i E[u_i] = \beta_1$$

(c) The assumption of constant conditional variance is not used in the proof of unbiasedness. The estimator is unbiased.

SECTION D

5.

(a) $E(X) = \int_0^\infty \frac{x}{\beta} exp\left(-\frac{x}{\beta}\right) dx = \left[-xe^{-x/\beta}\right]_0^\infty + \int_0^\infty exp\left(-\frac{x}{\beta}\right) dx = \left[-\beta e^{-x/\beta}\right]_0^\infty = \beta$ and $E\left(X^2\right) = \int_0^\infty \frac{x^2}{\beta} exp\left(-\frac{x}{\beta}\right) dx = 2\beta^2$ and hence $Var\left(X\right) = \beta^2$

(b) $\int_2^\infty \frac{1}{\beta} exp\left(-\frac{x}{\beta}\right) dx = \frac{1}{\beta}\left[-\beta e^{-x/\beta}\right]_2^\infty = e^{-2/\beta}$

(c) Random sampling plus CLT implies the approximation $\bar{x} \sim N\left(\beta, \frac{\beta^2}{N}\right)$

(d) Bookwork

(e) Under H_0 we have $\bar{x} \sim N\left(.3333, \frac{1/9}{100}\right)$ so $t = \frac{.433-.333}{\sqrt{\frac{1}{900}}} = 3$ critical is 2.33 so reject H_0:$\beta = 1/3$.

(f) Would not reject H_0 at 1% level if $\frac{\bar{x}-.333}{\frac{1}{30}} \leq 2.33$ ie if $\bar{x} \leq 0.4107$.
So Power$=1 - Prob\left(Accept\, H_0 \mid H_1\, true\right) = 1 - Prob\left(\bar{x} \leq .4107 \mid \bar{x} \sim N\left(.5, \frac{1}{400}\right)\right) =$ .9629

(g) We seek τ such that $\frac{e^{-3\tau}+e^{-2\tau}}{2} = 0.05$ or $10 + 10e^\tau = e^{3\tau}$ or $\left(e^\tau\right)^3 - 10\left(e^\tau\right) - 10 = 0$. Solution is $e^\tau = 3.578$ or $\tau = 1.275$.

6. (a) $\hat{\beta} = .055974$. Loosely, an additional year of education raises the wage by 5.5974%. Some students will provide the exact log transformation which implies an additional year of education raises wages by 5.75%. Both are acceptable.

(b) VVVV= 9.57, YYYY= .0444957, ZZZZ= .0674522.

(c) XXXX=0.00. The p-value gives the smallest level of significance at which we can reject the null that the coefficient on educ is 0. We can reject the null at any conceivable significance level.

(d) WWWW= 209.48543. The residual sum of squares or error sum of squares is the portion of variation in y that is NOT explained by the model (the education variable in this case).

(e) R^2 = Model sum of squares / Total sum of squares = Regression sum of squares/ Total sum of squares. $R^2 = 0.0961$.

(f) $lwage_i = \alpha + \varepsilon_i$. a or $\hat{\alpha}$ is the sample average of lwage.

(g) At high levels of education there will be lots of positive residuals and at low levels of education there will be negative residuals.

(h) $\hat{\beta}_1 = .055974$. The regression in table 1 is:

$$lwage_i = \alpha + \ \beta_1 educ_i + u_i$$

which we can rewrite as

$$y_i = \alpha + \ \beta_1 x_i + u_i$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

Thus OLS demeans both the "x" and "y" variables. Demeaning them again changes nothing in the formula for $\hat{\beta}_1$ and so $\hat{\beta}_1$ does not change.

(i) The value of R^2 in the demeaned regression is identical to that in the original regression. Demeaning the variables does not change the variation in each of the variables. Thus, the analysis of variance is unaffected.

```
. reg lwage_dev educ_dev

      Source |       SS       df       MS              Number of obs =     864
-------------+------------------------------           F(  1,   862) =   91.61
       Model |  22.2629147     1  22.2629147           Prob > F      =  0.0000
    Residual |   209.48543   862  .243022541           R-squared     =  0.0961
-------------+------------------------------           Adj R-squared =  0.0950
       Total |  231.748345   863  .268538059           Root MSE      =  .49297

------------------------------------------------------------------------------
   lwage_dev |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    educ_dev |    .055974   .0058481     9.57   0.000     .0444957    .0674522
       _cons |   7.45e-08   .0167713     0.00   1.000    -.0329173    .0329174
------------------------------------------------------------------------------

. reg lwage educ_dev

      Source |       SS       df       MS              Number of obs =     864
-------------+------------------------------           F(  1,   862) =   91.61
       Model |  22.2629147     1  22.2629147           Prob > F      =  0.0000
    Residual |   209.48543   862  .243022541           R-squared     =  0.0961
-------------+------------------------------           Adj R-squared =  0.0950
       Total |  231.748345   863  .268538059           Root MSE      =  .49297

------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    educ_dev |    .055974   .0058481     9.57   0.000     .0444957    .0674522
       _cons |   1.916288   .0167713   114.26   0.000     1.883371    1.949206
------------------------------------------------------------------------------
```