Economics 102: Analysis of Economic Data
Cameron Winter 2014 February 27
Department of Economics, U.C.-Davis

Second Midterm Exam (Version A)

Compulsory. Closed book. Total of 30 points and worth 22.5% of course grade.
Read question carefully so you answer the question.
You are to use only **simple calculations (+, -, /, *, square root) and show all workings**.
For computations final answers should be to at least four significant digits.
You may remove the formula sheet and the Stata output sheet(s) at end of exam.

Question scores

Question	$1a$	$1b$	$1c$	$1d$	$1e$	$1f$	$2a$	$2b$	$2c$	$2d$	$2e$	$2f$	$3a$	$3b$	$3c$	$4a$	$4b$	$4c$	$4d$	*Mult. choice*
Points	1	1	1	1	1	3	1	1	2	2	2	2	1	1	1	1	1	1	1	5

1.(a) Consider the time series plot for the natural logarithm of y given at the top of the last page. What is the approximate annual growth rate in the level of y? Explain your answer.

(b) An investment takes four years to double. What is the approximate annual rate of return for the investment? Explain your answer.

(c) Give the formula for a four-period moving average in the series y_t.

(d) What is a convenient source for time series data on government bond interest rates?

(e) What does the following Stata code produce: `graph twoway (line y year) (line x year)`?

(f) Suppose we have four observations with $(x, y) = (2, 6)$, $(3, 1)$, $(5, 1)$ and $(6, 4)$.
Compute the least squares intercept and slope coefficients from the formulae. **Show all workings.**

QUESTIONS 2-3 USE STATA OUTPUT GIVEN AT THE END OF THIS EXAM.
For some questions the answer is given directly in the output.
For other questions you will need to use the output plus additional computation.

2. This question uses data for hospitals in New York state in 2011.
`meancost` is the average cost of these knee replacements
`meancharge` is the average charge for knee replacements.
Note: the hospital charge is the initial price that the hospital charges. The hospital usually reeives less than this after negotiating with the patient or the patient's insurance company.

(a) How does the mean charge change when mean cost increases by one thousand dollars?

(b) Give a **95 percent** confidence interval for the population slope coefficient.

(c) Give a **99 percent** confidence interval for the population slope coefficient.

(d) The claim is made that the mean charge is not associated with the mean cost. Test this claim at significance **level 0.05**. **State clearly the null and alternative hypotheses and your conclusion.**

(e) The claim is made that the mean charge increases with the mean cost. Test this claim at significance **level 0.05**. **State clearly the null and alternative hypotheses and your conclusion.**

(f) The claim is made that the mean charge increases by \$1 with each extra \$1 of mean cost. Test this claim at significance **level 0.01**. **State clearly the null and alternative hypotheses and your conclusion.**

3. This question continues from the previous question.
If the Stata output given at the end of the exam is insufficient to answer the question then say so.

(a) Give the predicted mean hospital charge when the mean cost is $20,000. Show any workings.

(b) Suppose we regressed `meancost` on an intercept and `meancharge`. What value do you expect for the slope coefficient? Explain you answer.

(c) Suppose we regressed `meancharge` on only an intercept. What value do you expect for the intercept coefficient? Explain you answer.

4. You are given the following information following regression of y on an intercept and x:

```
Regression sum of squares = 40
Total sum of squares = 160
Number of observations = 10
```

(a) Give the R^2 for this regression. Show all workings.

(b) Give the correlation coefficient between x and y. Show all workings.

(c) Give the standard error of the residual for this regression.

(d) This part unrelated to parts a-c
Following `regress y x`, what does the following Stata command produce?
`predict z, resid`

Multiple Choice Questions (1 point each)

1. If you regress monthly data for the Dow Jones Industrial average on the Standard and Poors 500 then the R^2 is

a. between 0.00 and 0.25

b. between 0.25 and 0.50

c. between 0.50 and 0.75

d. between 0.75 and 1.00.

2. The sample correlation coefficient between y and x equals

a. the slope coefficient from regression of y on x

b. the slope coefficient from regression of x on y

c. the slope coefficient from regression of $(y - \bar{y})/s_y$ on $(x - \bar{x})/s_x$

d. all of the above

e. none of the above.

3. Assumptions necessary for the least squares estimator to be unbiased for β_2 include

a. $y = \beta_1 + \beta_2 x + u$

b. $\text{Var}[u|x] = \sigma_u^2$

c. neither a. nor b.

d. both a. and b.

4. If we rescale y by multiplying by 10 then

a. the correlation of y and x is 10 times larger

b. the covariance of y and x is 10 times larger

c. neither a. nor b.

d. both a. and b.

5. The width of a confidence interval for the slope coefficient

a. increases with increasing confidence level

b. increases with increasing sample size

c. both a. and b.

d. neither a. nor b.

Cameron: Department of Economics, U.C.-Davis

SOME USEFUL FORMULAS

Univariate Data

$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad and \qquad s_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$

$\bar{x} \pm t_{\alpha/2;n-1} \times (s_x/\sqrt{n}) \qquad and \qquad t = \frac{\bar{x}-\mu_0}{s/\sqrt{n}}$

$\text{ttail}(df, t) = \Pr[T > t]$ where $T \sim t(df)$

$t_{\alpha/2}$ such that $\Pr[|T| > t_{\alpha/2}] = \alpha$ is calculated using $\text{invttail}(df, \alpha/2)$.

Bivariate Data

$r_{xy} = \dfrac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \times \sum_{i=1}^{n}(y_i - \bar{y})^2}} = \dfrac{s_{xy}}{\sqrt{s_{xx} \times s_{yy}}}$ [Here $s_{xx} = s_x^2$ and $s_{yy} = s_y^2$].

$\widehat{y} = b_1 + b_2 x_i \qquad b_2 = \dfrac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \qquad b_1 = \bar{y} - b_2\bar{x}$

$\text{TSS} = \sum_{i=1}^{n}(y_i - \bar{y}_i)^2 \qquad \text{ErrorSS}=\sum_{i=1}^{n}(y_i - \widehat{y}_i)^2 \qquad \text{RegSS} = \text{TSS - ErrorSS}$

$R^2 = 1 - \text{ErrorSS/TSS}$

$b_2 \pm t_{\alpha/2;n-2} \times s_{b_2}$

$t = \dfrac{b_2 - \beta_{20}}{s_{b_2}} \qquad s_{b_2}^2 = \dfrac{s_e^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \qquad s_e^2 = \frac{1}{n-2}\sum_{i=1}^{n}(y_i - \widehat{y}_i)^2$

$y|x = x^* \in b_1 + b_2 x^* \pm t_{\alpha/2;n-2} \times s_e \times \sqrt{\frac{1}{n} + \frac{(x^*-\bar{x})^2}{\sum_i (x_i - \bar{x})^2} + 1}$

$\text{E}[y|x = x^*] \in b_1 + b_2 x^* \pm t_{\alpha/2;n-2} \times s_e \times \sqrt{\frac{1}{n} + \frac{(x^*-\bar{x})^2}{\sum_i (x_i - \bar{x})^2}}$

Multivariate Data

$\widehat{y} = b_1 + b_2 x_{2i} + \cdots + b_k x_{ki}$

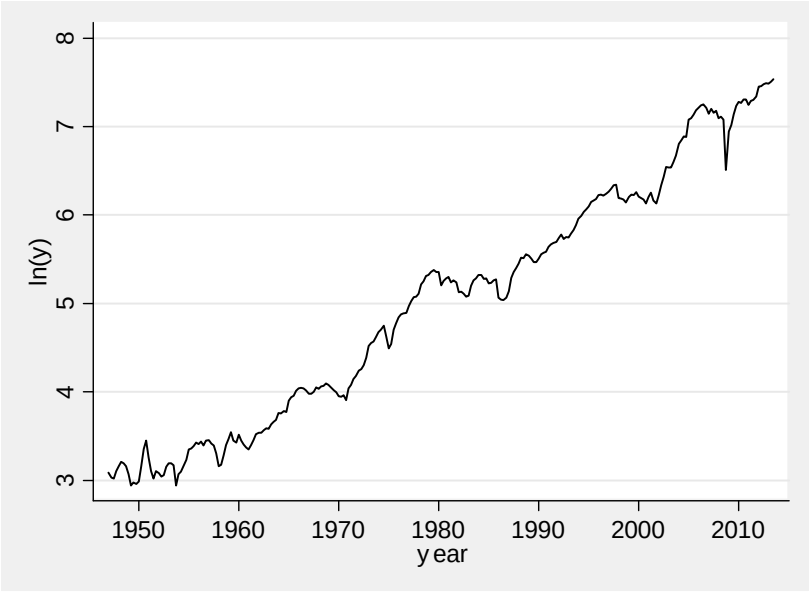
$R^2 = 1 - \text{ErrorSS/TSS} \qquad \bar{R}^2 = R^2 - \frac{k-1}{n-k}(1 - R^2)$

$b_j \pm t_{\alpha/2;n-k} \times s_{bj} \qquad and \qquad t = \dfrac{b_j - \beta_{j0}}{s_{b_j}}$

$F = \dfrac{R^2/(k-1)}{(1-R^2)/(n-k)} \qquad and \qquad F = \dfrac{(SSE_r - SSE_u)/(k-g)}{SSE_u/(n-k)}$

$\text{Ftail}(df1, df2, f) = \Pr[F > f]$ where F is $F(df1, df2)$ distributed.

F_α such that $\Pr[F > f_\alpha] = \alpha$ is calculated using $\text{invFtail}(df1, df2, \alpha)$.




```
. summarize meancharge meancost
```

Variable	Obs	Mean	Std. Dev.	Min	Max
meancharge	169	47957.27	22248.97	14953	123131
meancost	169	21007.51	10376.75	7021	86730

```
. regress meancharge meancost
```

Source	SS	df	MS
Model	3.1277e+10	1	3.1277e+10
Residual	5.1886e+10	167	310694328
Total	8.3163e+10	168	495016535

```
Number of obs =     169
F(  1,   167) =  100.67
Prob > F      =  0.0000
R-squared     =  0.3761
Adj R-squared =  0.3724
Root MSE      =   17627
```

meancharge	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
meancost	1.314908	.1310541	10.03	0.000	1.056171	1.573644
_cons	20334.33	3068.893	6.63	0.000	14275.5	26393.15

```
.
. display "t_.005;167 = " invttail(167,.005) "   t_.01;167 = " invttail(167,.01) ///
>       _n "t_.025;167 = " invttail(167,.025) "   t_.05;167 = " invttail(167,.05) ///
>          "  t_.10;167 = " invttail(167,.10)
t_.005;167 = 2.6055891   t_.01;167 = 2.3488842
t_.025;167 = 1.974271   t_.05;167 = 1.6540291  t_.10;167 = 1.2866415
```