Final Exam - Answers

Tomas Dvorak

Econ 253 Spring 2001

You have 120 minutes to answer the following 5 questions. The maximum number of points on this exam is 120. The number of points that a question is worth is the number of minutes that you should spend on that question. You may use a calculator and a double-sided sheet of paper with notes and formulas. Any collaboration is considered a violation of the honor code. Good luck!

1. (20 points) Consider an economy in which the overall unemployment rate is 18%. 20% of all individuals are college graduates, and 2% of all individuals are unemployed college graduates.

   (a) What is the probability that a randomly selected person is either unemployed or a college graduate?

   Answer: Let A be the event that a person is unemployed. Let B be the event that a person is a college graduate. We know that $P(A) = 0.18$, $P(B) = 0.2$ and $P(AB) = 0.02$. The probability that a person is either unemployed or a college graduate is $P(A + B) = P(A) + P(B) - P(AB) = 0.36$.

   (b) What is the probability that a randomly selected college graduate is unemployed?

   Answer: The probability that a randomly selected college graduate is unemployed is $P(A|B) = \frac{P(AB)}{P(B)} = 0.10$.

   (c) What is the probability that a randomly selected non-college graduate is unemployed?

   Answer: We need to find the probability of a person being unemployed given that he is not a college graduate, i.e. $P(A|notB) = \frac{P(notBA)}{P(notB)}$. The probability that a person is not a college graduate is $P(notB) = 0.8$. Now, we need to find the probability that a person is not a college graduate and is unemployed, $P(notBA)$. From the Venn diagram we see that $P(notBA) = P(A) - P(AB) = 0.16$. Hence, $P(A|notB) = \frac{0.16}{0.8} = 0.2$.

2. (20 points) An insurance company needs to estimate the mean amount claimed by its policy holders in the year 2000. A random sample of 121 policy holders reveals that the sample mean claim is \$744 and the sample standard deviation is \$300.

   (a) Compute a 95 percent confidence interval for the average amount claimed.

   Answer: Since $\frac{\overline{x}-\mu}{s/\sqrt{n}}$ is distributed t_{120}, $P(-1.98 < \frac{\overline{x}-\mu}{s/\sqrt{n}} < 1.98) = 0.95$. Hence the confidence interval is (744-54,744+54)=(690,798).

   (b) Test the hypothesis that the average claim is \$800 against the alternative that it is less than \$800.

   Answer: We test $H_0 : \mu = 800$ against the alternative $H_1 : \mu < 800$. The test statistic is $t = \frac{744-800}{27.27} = -2.05$. The critical region for this one sided test is $(-\infty, -1.658)$. The test statistic falls into the critical region and we reject the null hypothesis.

   (c) The insurance company would like to compare the mean amount claimed in 2000 with the data from 1999. The sample mean claim in 1999 was \$700, the sample standard deviation was \$250, and the sample size was 100. Is the increase in claims in 2000 statistically significant?

   Answer: This is the difference in means test. The null hypothesis is $H_0 : \mu_{1999} = \mu_{2000}$ against the alternative $H_1 : \mu_{1999} < \mu_{2000}$. The test statistic is $Z = \frac{700-744}{\sqrt{90000/121+62500/100}} = \frac{-44}{37} = -1.19$ The rejection region for this one tailed test is $(-\infty, -1.64)$ We can not reject the null hypothesis that the mean claims in 1999 and 2000 were the same and conclude that the increase in average claims in 2000 was not statistically significant.

3. (20 points) Consider the following model of a state's government expenditures on environment $ENV_i = B_1 + B_2 REV_i + u_i$ where ENV_i is the amount of environmental expenditures in state i and REV_i is the state's tax revenue. Suppose that the variance of the error term u_i is twice as high in states that have a democratic governor then in states with a non-democratic governor. Derive the transformed model and describe how you would find efficient estimates of B_1 and B_2.

Answer: We can write:
$var(u_i) = \sigma^2$ if the state has a non-democratic governor
$var(u_i) = 2\sigma^2$ if the state has a democratic governor
The transformed model is:
$Y_i = B_1 \cdot 1 + B_2 REV_i + u_i$ if the state has a non-democratic governor
$\frac{Y_i}{\sqrt{2}} = B_1 \cdot \frac{1}{\sqrt{2}} + B_2 \frac{REV_i}{\sqrt{2}} + \frac{u_i}{\sqrt{2}}$ if the state has a democratic governor
Our transformed model is:
$Y_i^* = B_1 X_i^* + B_2 REV_i^* + u_i^*$ where
$Y_i^* = Y_i$ if a non-democratic governor
$\quad = Y_i/\sqrt{2}$ if a democratic governor
$X_i^* = 1$ if a non-democratic governor
$\quad = 1/\sqrt{2}$ if a democratic governor
$REV_i^* = REV_i$ if a non-democratic governor
$\quad = REV_i/\sqrt{2}$ if a democratic governor

4. (30 points) Consider the following cross country data on economic growth and its determinants in 104 countries.

- grwth = average annual percentage growth rate of per capita income 1960-85
- y60 = log of income in 1960
- inv = log of average investment to GNP ratio
- school = log of average percentage population in school
- nonoil = 1 if country does not produce oil
- indstr = 1 if country is industrialized
- oecd = 1 if country is a member of the OECD

Consider the following regression output:

```
  Source|       SS       df       MS             Number of obs =     104
--------+------------------------------          F(  3,   100) =    7.46
   Model|  63.5425115     3  21.1808372          Prob > F      =  0.0001
Residual|  283.755419   100  2.83755419          R-squared = 0.1830
--------+------------------------------          Adj R-squared =  0.1585
   Total|  347.297931   103  3.37182457          Root MSE      =  1.6845

------------------------------------------------------------------
   grwth|      Coef.   Std. Err.       t     P>|t|  [95\% Conf. Int.]
--------+---------------------------------------------------------
  nonoil|  -1.426046   .7722031     -1.847   0.068       -2.958   .105
  indstr|   1.451058   .4206075      3.450   0.001        .6165  2.285
    oecd|   .8800666   .4272218      2.060   0.042        .0324  1.727
   _cons|   1.848703   .6876959      2.688   0.008        .4843  3.213
------------------------------------------------------------------
```

(a) What is the average per capita income growth in non-oil producing OECD countries? (Note that all OECD countries are industrialized.)

Answer: $E(grwth|nonoil = 1, industr = 1, oecd = 1) = 1.85 - 1.42 + 1.45 + 0.88 = 2.76\%$. Per capita income in non-oil producing OECD countries grew on average 2.7% per year.

(b) Is it possible to introduce an interaction between the `oecd` dummy and `indstr` dummies?

Answer: Since all OECD countries are also industrialized the interaction dummy `OECD`*`indstr` would be equal to the `oecd` dummy. Thus, there would be a perfect multicollinearity and the estimation would fail.

(c) As a researcher you hypothesise that OECD countries grew faster because they had higher investment, higher percentage of population in schools and higher initial income. Therefore, you run a regression controlling for these characteristics. The following is the computer output that you obtain:

```
  Source |       SS       df       MS              Number of obs =     104
---------+------------------------------           F(  6,    97) =   24.63
   Model |  209.673521     6  34.9455868           Prob > F      =  0.0000
Residual |   137.62441    97  1.41880835           R-squared = 0.6037
---------+------------------------------           Adj R-squared =  0.5792
   Total |  347.297931   103  3.37182457           Root MSE      =  1.1911

------------------------------------------------------------------------------
   grwth |      Coef.   Std. Err.       t     P>|t|   [95\% Conf. Interval]
---------+--------------------------------------------------------------------
  nonoil |  -2.486448   .6067651     -4.098   0.000    -3.690709    -1.28218
  indstr |   1.157416   .3498265      3.309   0.001     .4631074     1.85172
    oecd |   1.290314   .3804038      3.392   0.001      .535318     2.04531
     y60 |  -1.537046   .2005156     -7.665   0.000    -1.935014    -1.13907
     inv |   1.949172   .3144045      6.200   0.000     1.325166     2.57317
  school |    .658175    .222121      2.963   0.004     .2173263     1.09902
   _cons |   8.442761   1.780641      4.741   0.000     4.908681     11.9768
------------------------------------------------------------------------------
Durbin-Watson Statistic =  2.051045
```

Was your hypothesis correct?

Answer: No. Even after controlling for investment schooling and initial income, the `oecd` dummy is significant. Therefore, there is something else besides high investment ratios and schooling that makes OECD countries grow.

(d) What is the interpretation of the coefficient on `inv`?

Answer: Since the investment ratio is in logarithm, the interpretation of the coefficient is that if investment ratio increases by 1% the average growth will increase by 0.0194 percentage points per year.

(e) Suppose that you hypothesize that the effect of investment on economic growth is different in industrialized countries then in non-industrialized countries. Describe carefully how would you test this hypothesis.

Answer: One could introduce an interaction between the investment variable `inv` and the dummy for industrialize countries: `inv_indstr=inv*indstr` and include this variable in the regression. If the coefficient on the interaction is significant, we conclude that there is a difference in the effect of investment in industrialized and non-industrialized countries.

(f) "Countries that were relatively poor in 1960 grew faster than countries that were relatively rich." Do you agree with this statement?

Answer: Yes. The negative sign on `y90` indicates that the higher the initial income the lower the subsequent growth. The interpretation of the coefficient on the log of initial income `y60` is: if initial income increases by 1% the country will grow 0.015 percentage points slower. This result shows evidence for convergence in per capita incomes across countries.

5. (30 points) The U.S. economy has chronically low savings rate. An economist tried to investigate consumption and savings pattern using the following model:

$$ln(C_t) = B_1 + B_2 ln(PDI_t) + u_t \tag{1}$$

where C_t is annual aggregate personal consumption in billions of dollars, and PDI_t is annual aggregate personal disposable income also in billions of dollars. Answer the following questions:

(a) Describe how would you test for unitary elasticity of consumption with respect to income.

Answer: The coefficient B_2 is the elasticity of consumption with respect to income $B_2 = \frac{\Delta C/C}{\Delta PDI/PDI}$. Therefore, to test for unitary elasticity we need to test $H_0 : B_2 = 1$ against $H_1 : B_2 \neq 1$. The test statistic is $t = \frac{b_2-1}{\sigma_{b_2}}$ which is distributed t_{n-1}. The critical(rejection) region is in the tails of the t distribution.

(b) The economist estimated the model in (1) using annual data from 1950 to 2000 which are printed at the back of this exam. He obtained the Durbin Watson statistic $d = 0.24$. Does the model suffer from first order autocorrelation?

Answer: Yes, the Durbin Watson lower limit for a model with two explanatory variables ($k' = 1$) and 51 observations is 1.503. We therefore we reject the null hypothesis of no autocorrelation and accept the alternative that there is positive autocorrelation.

(c) Suppose that the economist ignores the autocorrelation and tests the hypothesis that consumption is unitary elastic with respect to income. Will his conclusions be valid? Why or why not?

Answer: In the presence of autocorrelation the estimates of standard errors of our coefficient estimates are biased and inconsistent, rendering test statistics incorrect. Therefore, if the economist ignores autocorrelation his test will be invalid.

(d) How would you obtain efficient estimates? Write down the transformed model.

Answer: If the error term suffers from first order autocorrelation it can be written as $u_t = \rho u_{t-1} + v_t$. In order to obtain efficient estimates we need to transform the model so that the error term in the transformed model is v_t:

$$(lc_t - \rho lc_{t-1}) = B_1(1-\rho) + B_2(PDI_t - \rho PDI_{t-1}) + v_t \tag{2}$$

The efficient estimates are obtained by running the regression on transformed variables.

(e) The following is the printout of the computer output from the regression of the transformed model. Here `tlc` and `tlpdi` are the transformed variables.

```
Source|       SS       df       MS            Number of obs = 50
------+------------------------------         F(  1,    48) = 8042.76
 Model|   .93868105     1   .93868105         Prob > F      =  0.0000
 Resid|   .00560214    48  .000116711         R-squared     = 0.9941
------+------------------------------         Adj R-squared =  0.9939
 Total|   .94428319    49  .019271086         Root MSE      =   .0108

---------------------------------------------------------------------
  tlc |      Coef.   Std. Err.      t    P>|t|   [95\% Conf. Interval]
------+--------------------------------------------------------------
 tlpdi|   1.018631    .0113583    89.681  0.000     .9957938    1.0414
 _cons|  -.0294716    .0105545    -2.792  0.007    -.0506929   -.00825
---------------------------------------------------------------------
Durbin-Watson Statistic =  1.795359
```

Test for unitary elasticity of consumption with respect to income.

Answer: Following the steps outlined in part (a), the test statistic is $t = \frac{b_2 - 1}{\widehat{\sigma_{b_2}}} = \frac{0.018}{0.011} = 1.636$. The critical region is $(-\infty$, -2) and (2, $\infty)$. The computed test statistic falls into the acceptance region. We thus accept the null hypothesis that consumption is unitary elastic with respect to income.

(f) Suppose that President Bush succeeds in implementing his tax cut proposal in 2001. This will increase personal disposable income in 2001 to 7 trillion dollars. What is your prediction for aggregate consumption in 2001?

Answer: Here are the pieces of information we need:

$\widehat{\rho} = 0.88$

$b_1 = \frac{-0.029}{1-0.88} = -0.24$

$b_2 = 1.019$

$e_T = ln(C_T) - b_1 - b_2 ln(PDI_T) = 8.818 + 0.24 - 1.019 \cdot 8.852 = 0.04$

The predicted logarithm of consumption is:

$$
\begin{aligned}
ln(\widehat{C_{T+1}}) &= b_1 + b_2 ln(PDI_{T+1}) + \widehat{\rho} e_T \\
&= -0.24 + 1.019 \cdot ln(7000) + 0.88 \cdot 0.04 \\
&= -0.24 + 9.022 + 0.035 = 8.817
\end{aligned}
$$

The logarithm of consumption 8.817 corresponds to the level of consumption at \$6748 billion.

	year	pdi	c	lpdi	lc
1.	1950	210.6	192.7	5.349961	5.261135
2.	1951	231.2	208.6	5.443283	5.340418
3.	1952	243.6	219.7	5.495528	5.392263
4.	1953	258.8	233.4	5.556056	5.452754
5.	1954	264.5	240.5	5.577841	5.48272
6.	1955	283.4	259	5.646859	5.556828
7.	1956	302.8	271.9	5.713072	5.605434
8.	1957	319.7	287	5.767383	5.659482
9.	1958	330.8	296.6	5.801514	5.692384
10.	1959	351.2	318.1	5.861356	5.762366
11.	1960	366.2	332.3	5.90318	5.806038
12.	1961	382.4	342.7	5.946467	5.836855
13.	1962	405.6	363.8	6.005367	5.896604
14.	1963	425.8	383.1	6.05397	5.948296
15.	1964	463	411.7	6.137727	6.020295
16.	1965	498.9	444.3	6.212406	6.0965
17.	1966	539.1	481.8	6.289901	6.177529
18.	1967	576.2	508.7	6.356455	6.231858
19.	1968	626.2	558.7	6.43967	6.325613
20.	1969	675	605.5	6.514713	6.406054
21.	1970	736.5	648.9	6.601909	6.475279
22.	1971	801.7	702.4	6.686735	6.554503
23.	1972	868.6	770.7	6.766883	6.647299
24.	1973	979	852.5	6.886532	6.748173
25.	1974	1072.3	932.4	6.977561	6.837762
26.	1975	1181.4	1030.3	7.074455	6.937605
27.	1976	1299.9	1149.8	7.170043	7.047343
28.	1977	1436	1278.4	7.269617	7.153365
29.	1978	1614.8	1430.4	7.386966	7.265709
30.	1979	1808.2	1596.3	7.500087	7.375444
31.	1980	2019.8	1762.9	7.610754	7.474716
32.	1981	2247.9	1944.2	7.717752	7.572606
33.	1982	2406.8	2079.3	7.786053	7.639787
34.	1983	2586	2286.4	7.857868	7.734734
35.	1984	2887.6	2498.4	7.968181	7.823406
36.	1985	3086.5	2712.6	8.034793	7.905663
37.	1986	3262.5	2895.2	8.090249	7.970809
38.	1987	3459.5	3105.3	8.148879	8.040866
39.	1988	3752.4	3356.6	8.230151	8.118684
40.	1989	4016.3	3596.7	8.298117	8.187772
41.	1990	4293.6	3831.5	8.364881	8.251012
42.	1991	4474.8	3971.2	8.406217	8.286823
43.	1992	4754.6	4209.7	8.466867	8.345147
44.	1993	4935.3	4454.7	8.504169	8.401715
45.	1994	5165.4	4716.4	8.549738	8.458801
46.	1995	5422.6	4969	8.59833	8.510974
47.	1996	5677.7	5237.5	8.644301	8.5636
48.	1997	5968.2	5529.3	8.694201	8.617817
49.	1998	6320	5850.9	8.751474	8.674351
50.	1999	6637.7	6268.7	8.800521	8.743324
51.	2000	6989.8	6757.3	8.852207	8.818378