Statistics/Data Analysis

```
      name:  mainlog
       log:  C:\Users\Conor\Documents\Conor\Grad School\TA Work\Econ 103 - Econometric
> s\Stata Work\Week 9\wk9_section_log.smcl
  log type:  smcl
 opened on:  22 Feb 2018, 19:46:47

1 .
2 . // Demonstration STATA code for week 9
3 . // Principles of Econometrics 4th Edition
4 . // Covered Problems: 7.2 and 7.14
5 .
6 . set more off

7 . clear all

8 . use fullmoon.dta, clear

9 .
10. ////////////////////////////////////////////////////////////////////////////
  > ////////////////////////////////// Question 7.2 \\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
  > ////////////////////////////////////////////////////////////////////////////
  >
11. ******************************************************************************
12. * Long story short: Compare the results of the regression below, with and
13. * without the fullmoon and newmoon variables. Conduct the joint hypothesis test
14. * for whether to include either fullmoon and/or newmoon in the regression.
15. ******************************************************************************
16.
17. reg case t holiday friday saturday fullmoon newmoon

      Source |       SS           df       MS      Number of obs   =       229
-------------+----------------------------------   F(6, 222)       =      7.77
       Model |  5693.37691         6  948.896151   Prob > F        =    0.0000
    Residual |   27108.824       222   122.11182   R-squared       =    0.1736
-------------+----------------------------------   Adj R-squared   =    0.1512
       Total |  32802.2009       228  143.869302   Root MSE        =     11.05

------------------------------------------------------------------------------
       cases |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           t |   .0337998   .0110528     3.06   0.003     .0120179    .0555816
     holiday |   13.86293   6.445175     2.15   0.033     1.161376    26.56448
      friday |   6.909776   2.111318     3.27   0.001     2.748985    11.07057
    saturday |    10.5894   2.118432     5.00   0.000     6.414592    14.76421
    fullmoon |   2.454453   3.980923     0.62   0.538    -5.390782    10.29969
     newmoon |   6.405947   4.256893     1.50   0.134    -1.983144    14.79504
       _cons |   93.69583   1.559159    60.09   0.000     90.62318    96.76847
------------------------------------------------------------------------------

18. test (fullmoon = 0) (newmoon = 0)

 ( 1)  fullmoon = 0
 ( 2)  newmoon = 0

       F(  2,   222) =    1.29
            Prob > F =    0.2770

19.
20. // Can also calculate the F statistic by calculating the "restricted" regression
```

```
21. // and using either the sse equation or the r2 equation.
22.
23. // Store sse and r2 from the unrestricted regression. Also collect the degrees
24. // of freedom from the unrestricted regression
25. scalar sse_u = e(rss)

26. scalar r2_u = e(r2)

27. scalar df_u = e(df_r)

28.
29. // Run "restricted" regresion where we impose that betas for fullmoon and
30. // newmoon are zero (i.e. drop them from the regression)
31. reg case t holiday friday saturday
```

Source	SS	df	MS	Number of obs	=	229
				F(4, 224)	=	10.98
Model	5378.00978	4	1344.50245	Prob > F	=	0.0000
Residual	27424.1911	224	122.429425	R-squared	=	0.1640
				Adj R-squared	=	0.1490
Total	32802.2009	228	143.869302	Root MSE	=	11.065

cases	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
t	.0338305	.0110672	3.06	0.003	.0120214	.0556395
holiday	13.61679	6.451068	2.11	0.036	.9042466	26.32934
friday	6.84914	2.113669	3.24	0.001	2.683921	11.01436
saturday	10.34207	2.115326	4.89	0.000	6.173589	14.51056
_cons	94.02146	1.545848	60.82	0.000	90.97519	97.06772

```
32.
33. // Store sse and r2 from the restricted regression.
34. scalar sse_r = e(rss)

35. scalar r2_r = e(r2)

36.
37. // Calculate the F-stat using the sse formula and the r2 formula
38. // Compare the results to the test command from earlier
39. scalar fstat_sse = ((sse_r-sse_u)/2)/(sse_u/df_u)

40. scalar fstat_r2 = ((1-r2_r)/(1-r2_u) - 1)*(df_u/2)

41.
42. disp "F-statistic using SSE formula: " fstat_sse " (p-value = " Ftail(2,df_u,fstat_s
  > se) ")"
  F-statistic using SSE formula: 1.2913047 (p-value = .27696856)

43. disp "F-statistic using R2 formula: " fstat_r2
  F-statistic using R2 formula: 1.2913047

44.
45.
46. ////////////////////////////////////////\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
  > ////////////////////////////// Question 7.14 \\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
  > ////////////////////////////////////////\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
  >
47. clear all
```

```
48. use fair4.dta

49.
50. ******************************************************************************
51. *Part A: Consider the regression model
52. *
53. * VOTE = beta1 + beta2*GROWTH + beta3*INFLATION + beta4*GOODNEWS
54. *              + beta5*PERSON + beta6*DURATION + beta7*PARTY + beta8*WAR
55. *
56. * Discuss the anticipated effect of the dummy variables PERSON and WAR
57. ******************************************************************************
58.
59. ******************************************************************************
60. *Part B: The binary variable PARTY is somewhat different from the dummy
61. * variables we have considered. Write out the regression function E(VOTE) for
62. * the two values of PARTY. Discuss the effects of this specification.
63. ******************************************************************************
64.
65. ******************************************************************************
66. *Part C: Use the data for the period 1916-2004 to estimate the proposed model.
67. * Discuss the estimation results. Are the signs as expected? Are the estimates
68. * statistically significant? How well does the model fit the data?
69. ******************************************************************************
70.
71. reg vote growth inflation goodnews person duration party war if year >= 1916 & year
  > <=2004
```

Source	SS	df	MS		
Model	**924.332053**	**7**	**132.047436**	Number of obs	= **23**
Residual	**96.8419433**	**15**	**6.45612955**	F(7, 15)	= **20.45**
				Prob > F	= **0.0000**
Total	**1021.174**	**22**	**46.4169998**	R-squared	= **0.9052**
				Adj R-squared	= **0.8609**
				Root MSE	= **2.5409**

vote	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
growth	**.6796892**	**.1106514**	**6.14**	**0.000**	**.4438414**	**.915537**
inflation	**-.6571746**	**.2913931**	**-2.26**	**0.039**	**-1.278264**	**-.0360848**
goodnews	**1.074924**	**.2493241**	**4.31**	**0.001**	**.5435026**	**1.606346**
person	**3.298287**	**1.408083**	**2.34**	**0.033**	**.2970301**	**6.299544**
duration	**-3.329959**	**1.212408**	**-2.75**	**0.015**	**-5.914146**	**-.745772**
party	**-2.676307**	**.6264261**	**-4.27**	**0.001**	**-4.011502**	**-1.341111**
war	**5.614855**	**2.687901**	**2.09**	**0.054**	**-.11427**	**11.34398**
_cons	**47.2628**	**2.538367**	**18.62**	**0.000**	**41.8524**	**52.6732**

```
72.
73. ******************************************************************************
74. *Part D: Predict the outcome of the 2008 election using the given 2008 data
75. * for values of explanatory variables. Based on the prediction, would you have
76. * picked the outcome of the election correctly?
77. ******************************************************************************
78.
79. predict voteHat, xb

80. list voteHat if year == 2008
```

	voteHat
33.	**48.09079**

```
81.
82. ******************************************************************************
83. *Part E: Construct a 95% confidence interval for the outcome of the 2008
84. * election.
85. ******************************************************************************
86.
87. // Generate the standard errors for voteHat: prediction (stdp) and forecast (stdf)
88. predict stdp, stdp

89. predict stdf, stdf

90.
91. // Calculate t-critical value using inverse t function
92. scalar alpha = 0.05

93. scalar tval = invt(e(df_r),1-alpha/2)

94.
95. // Calculate confidence intervals using stdp and stdf
96. gen cilow_stdp = voteHat - tval*stdp

97. gen cihigh_stdp = voteHat + tval*stdp

98. gen cilow_stdf = voteHat - tval*stdf

99. gen cihigh_stdf = voteHat + tval*stdf

100
101 // Display fitted value and confidence intervals
102 list voteHat cilow_stdp cihigh_stdp cilow_stdf cihigh_stdf if year == 2008
```

	voteHat	cilow_~p	cihigh~p	cilow_~f	cihigh~f
33.	**48.09079**	**45.50838**	**50.67319**	**42.09083**	**54.09075**

```
103
104 ******************************************************************************
105 *Part F: Using data values of your choice (you must explain them), predict the
106 * outcome of the 2012 election.
107 ******************************************************************************
108
109 // Add a new row to the data set (will be all blank values)
110 scalar newOb = _N+1

111 set obs `=newOb'
  number of observations (_N) was 33, now 34

112
113 replace year = 2012 in `=newOb' // add value for year
  (1 real change made)

114
115 // Set values for known variables
116 replace party = 1                        if year == 2012 // democrat in 2012
  (1 real change made)

117 replace person = 1                       if year == 2012 // Obama re-election run in
  > 2012
  (1 real change made)
```

```
118 replace war = 0                          if year == 2012 // Not 1920, 1944, or 1948
  (1 real change made)

119 replace duration = 0            if year == 2012 // Same as 1984, 1996, 2004, etc.
  (1 real change made)

120
121 // Set values for economic variables (open to interpretation)
122 replace goodnews = 4            if year == 2012
  (1 real change made)

123 replace growth = 2                       if year == 2012
  (1 real change made)

124 replace inflation = 1.5         if year == 2012
  (1 real change made)

125
126 // Re-run regression, now including the 2008 observation
127 qui reg vote growth inflation goodnews person duration party war if year >= 1916 & y
  > ear <=2008

128
129 // Calculate fitted value and STDF
130 predict voteHat2, xb

131 predict stdf2, stdf

132
133 // Calculate new t-val (+1 to degrees of freedom from previous estimate)
134 scalar tval2 = invttail(e(df_r),alpha/2)

135
136 gen lb = voteHat2 - tval2*stdf2

137 gen ub = voteHat2 + tval2*stdf2

138
139 // Show the new fitted value and confidence interval
140 list lb voteHat2 ub if year == 2012
```

	lb	voteHat2	ub
34.	**46.54211**	**52.50974**	**58.47736**

```
141
142 //Convert log file (smcl) to pdf
```