ØAMET2200 · Fall 2019
Worksheet 10

Instructor: Fenella Carpena

November 8, 2019

Exercise 1 (Adapted from Stine and Foster, Chapter 27, 4M Example 1) The US economy fell into a recession in 2008 and 2009. The slumping economy almost wiped out the domestic auto industry. How badly did the recession hit this industry? We want to use a regression model to understand how much car sales was lost because of the recession. Specifically, we want to answer the question: how much car sales would we have expected if the trend of car sales had followed the pattern before the recession?

To investigate this question, we have data on quarterly car sales from 1990:Q1 to 2007:Q4. The data contain the following three variables.

- `Time`: the variable is coded on an annual scale as follows. `Time` $= 1990+2/12$ for the first quarter of 1990, `Time` $= 1990+5/12$ for the second quarter of 1990, `Time` $= 1990+8/12$ for the third quarter of 1990, `Time` $= 1990+11/12$ for the fourth quarter of 1990. The variable is coded similarly for the years 1991-2007.
- `Quarter`: A string variable where `Q1` means first quarter, `Q2` means second quarter, etc.
- `QuarterlyCarSales`: The number of cars that are sold in a given quarter, in thousands.

A subset of the data and variables is shown below after question (d), together with the Stata output from analyzing this data.

(a) Consider the regression of `QuarterlyCarSales` on `Time`, `X1`, `X2`, and `X3`. Interpret the coefficient on `X1`. Interpret the coefficient on `Time`.

(b) Do the residuals from the regression in part (a) appear to be independent?

(c) Using the regression in part (a), what is our prediction for car sales in Quarter 1 of 2008? What about Quarter 2 of 2008?

(d) Your colleague tells you that we should also make a prediction interval for the predictions we made in part (c). Why or why not?

Time	Quarter	QuarterlyCarSales
2005.1667	Q1	1792.4
2005.4167	Q2	2125.7
2005.6667	Q3	2068.5
2005.9167	Q4	1733.2
2006.1667	Q1	1831.6
2006.4167	Q2	2169.6
2006.6667	Q3	2081.3
2006.9167	Q4	1738.2
2007.1667	Q1	1799.1
2007.4167	Q2	2129.9
2007.6667	Q3	1921.6
2007.9167	Q4	1767.8

```
. forvalues i = 1/3 { ;
  2.         gen X`i' = (Quarter == "Q`i'") ;
  3. } ;

.  regress QuarterlyCarSales X1 X2 X3 Time ;

      Source |       SS           df       MS      Number of obs   =        72
-------------+----------------------------------   F(4, 67)        =     47.93
       Model |  2182184.05         4  545546.012   Prob > F        =    0.0000
    Residual |  762545.619        67  11381.2779   R-squared       =    0.7410
-------------+----------------------------------   Adj R-squared   =    0.7256
       Total |  2944729.67        71  41475.0658   Root MSE        =    106.68

------------------------------------------------------------------------------
QuarterlyC~s |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          X1 |   53.49839   35.60744     1.50   0.138    -17.57435    124.5711
          X2 |   379.5712   35.58166    10.67   0.000     308.5499    450.5924
          X3 |   223.9661   35.56618     6.30   0.000     152.9757    294.9565
        Time |  -16.86881   2.423364    -6.96   0.000    -21.70586   -12.03175
       _cons |   35628.95   4845.379     7.35   0.000     25957.54    45300.37
------------------------------------------------------------------------------

. estat dwatson ;

Durbin-Watson d-statistic(  5,    72) =  .8580442
```

Exercise 2 Suppose that we have a quarterly time series of $GDPGR$ (i.e., the GDP Growth Rate, measured in percentage points) from 1962:Q1 to 2012:Q4. With these data, we estimate an AR(1) model using OLS, and we obtain the following results. Note that the standard errors are shown in parenthesis below the coefficient estimates.

$$\widehat{GDPGR}_t = \underset{(0.35)}{1.99} + \underset{(0.08)}{0.34}\, GDPGR_{t-1},\ \overline{R}^2 = 0.1,\ s_e = 3.16,\ n = 204$$

We also estimated an AR(2) model. The OLS regression result is

$$\widehat{GDPGR}_t = \underset{(0.40)}{1.63} + \underset{(0.08)}{0.28}\, GDPGR_{t-1} + \underset{(0.08)}{0.18}\, GDPGR_{t-2},\ \overline{R}^2 = 0.14,\ s_e = 3.11,\ n = 204$$

The following are the values of $GDPGR$ from 2012:Q1 to 2012:Q4.

Date	2012:Q1	2012:Q2	2012:Q3	2012:Q4
$GDPGR$	3.64	1.20	2.75	0.15

(a) Is a positive growth rate of GDP in one quarter correlated with positive growth in the next quarter?

(b) What is the forecast for the growth rate of GDP in 2013:Q1 that we would have made in 2012:Q4, based on the AR(1) model?

(c) What is the forecast for the growth rate of GDP in 2013:Q1 that we would have made in 2012:Q4, based on the AR(2) model?

(d) Given the above two regression results, how many lags should be in the AR model? In other words, should the AR model be AR(1) or AR(2)?

Exercise 3 Many autoregressions are mean-reverting. *Mean-reverting* means that the forecasts eventually tend back to (revert to) the mean of the time series. For example, a manager uses an AR(1) model to predict sales next week using sales in recent weeks. Sales typically run about \$250,000 per week. The estimating equation (with sales in thousands of dollars) is

$$\widehat{y}_t = 50 + 0.8 y_{t-1}.$$

(a) If sales this week are \$250,000 (the mean level), what does the equation forecast for next week?

(b) If sales this week are \$300,000 (\$50,000 above the mean), does the equation forecast sales to increase farther above the mean or to return toward the mean?

(c) In general, when does a first-order autoregression with positive slope ($0 < b_1 < 1$) predict an increase in the time series? A decrease?

Exercise 4 (Stata) This exercise considers monthly prices of shares in JC Penney (an American department store chain) from January 2000 to December 2011. The dataset for this exercise is `penney.xls`. The data contains the variable `Month` with values $1, 2, \ldots, 120$ that can be used as the time index t.

(a) Fit a polynomial trend to the time series of prices. Try polynomials of various degrees (i.e., try polynomials with t, t^2, and other powers up to t^6). Do any of these, with orders 6 or less, capture the ups and downs of prices?

(b) Fit an autoregression to the time series of prices, using three lags of the price. Are all of these lagged predictors useful? Remove predictors that do not statistically significantly improve the fit of the model and summarize your final model.

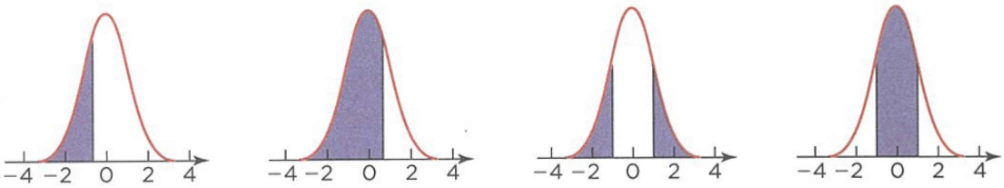
(c) Does the model you created in part (b) meet the conditions for the MRM?

(d) Generate a new variable called `return` for the "returns" on JC Penney stock. Returns are the ratio of the change in price to the price in the earlier period, that is

$$\frac{price_t - price_{t-1}}{price_{t-1}}.$$

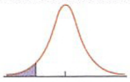
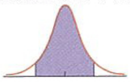
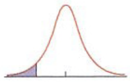
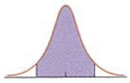
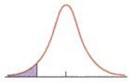
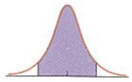
(e) Re-estimate the final regression model you chose in part (b) but using `return` as the response variable. Compare this model to the model in part (b). Which would you prefer to use for forecasting the stock price of JC Penney? Explain.

	Reject $H_0: \rho_\varepsilon = 0$ if	
n	**D is less than**	**D is greater than**
15	1.36	2.64
20	1.41	2.59
30	1.49	2.51
40	1.54	2.46
50	1.59	2.41
75	1.65	2.35
100	1.69	2.31



z	$P(Z \leq -z)$	$P(Z \leq z)$	$P(\lvert Z\rvert > z)$	$P(\lvert Z\rvert \leq z)$
0	0.50	0.50	1	0
0.0502	0.48	0.52	0.96	0.04
0.1004	0.46	0.54	0.92	0.08
0.1510	0.44	0.56	0.88	0.12
0.2019	0.42	0.58	0.84	0.16
0.2533	0.40	0.60	0.80	0.20
0.3055	0.38	0.62	0.76	0.24
0.3585	0.36	0.64	0.72	0.28
0.4125	0.34	0.66	0.68	0.32
0.4677	0.32	0.68	0.64	0.36
0.4959	0.31	0.69	0.62	0.38
0.5244	0.30	0.70	0.60	0.40
0.5828	0.28	0.72	0.56	0.44
0.6433	0.26	0.74	0.52	0.48
0.6745	0.25	0.75	0.50	0.50
0.7063	0.24	0.76	0.48	0.52
0.7388	0.23	0.77	0.46	0.54
0.7722	0.22	0.78	0.44	0.56
0.8064	0.21	0.79	0.42	0.58
0.8416	0.20	0.80	0.40	0.60
0.8779	0.19	0.81	0.38	0.62
0.9154	0.18	0.82	0.36	0.64
0.9542	0.17	0.83	0.34	0.66
0.9945	0.16	0.84	0.32	0.68
1.0364	0.15	0.85	0.30	0.70
1.0803	0.14	0.86	0.28	0.72
1.1264	0.13	0.87	0.26	0.74
1.1750	0.12	0.88	0.24	0.76
1.2265	0.11	0.89	0.22	0.78
1.2816	0.10	0.90	0.20	0.80
1.3408	0.09	0.91	0.18	0.82
1.4051	0.08	0.92	0.16	0.84
1.4758	0.07	0.93	0.14	0.86
1.5548	0.06	0.94	0.12	0.88
1.6449	0.05	0.95	0.10	0.90
1.7507	0.04	0.96	0.08	0.92
1.8808	0.03	0.97	0.06	0.94
1.9600	0.025	0.975	0.05	0.95
2.0537	0.02	0.98	0.04	0.96
2.3263	0.01	0.99	0.02	0.98
2.5758	0.005	0.995	0.01	0.99
2.8070	0.0025	0.9975	0.005	0.995
3.0902	0.001	0.999	0.002	0.998
3.2905	0.0005	0.9995	0.001	0.999
3.7190	0.0001	0.9999	0.0002	0.9998
3.8906	0.00005	0.99995	0.0001	0.9999
4.2649	0.00001	0.99999	0.00002	0.99998
4.4172	0.000005	0.999995	0.00001	0.99999

T–TABLE **Percentiles of Student's t distribution.**

$df = 1$

t	$P(T_1 \leq -t)$	$P(-t \leq T_1 \leq t)$
3.078	0.1	0.8
6.314	0.05	0.9
12.71	0.025	0.95
31.82	0.01	0.98
63.66	0.005	0.99
318.3	0.001	0.998
636.6	0.0005	0.999
6366	0.00005	0.9999

$df = 2$

t	$P(T_2 \leq -t)$	$P(-t \leq T_2 \leq t)$
1.886	0.1	0.8
2.920	0.05	0.9
4.303	0.025	0.95
6.965	0.01	0.98
9.925	0.005	0.99
22.33	0.001	0.998
31.60	0.0005	0.999
99.99	0.00005	0.9999

$df = 3$

t	$P(T_3 \leq -t)$	$P(-t \leq T_3 \leq t)$
1.638	0.1	0.8
2.353	0.05	0.9
3.182	0.025	0.95
4.541	0.01	0.98
5.841	0.005	0.99
10.21	0.001	0.998
12.92	0.0005	0.999
28.00	0.00005	0.9999

$df = 4$

t	$P(T_4 \leq -t)$	$P(-t \leq T_4 \leq t)$
1.533	0.1	0.8
2.132	0.05	0.9
2.776	0.025	0.95
3.747	0.01	0.98
4.604	0.005	0.99
7.173	0.001	0.998
8.610	0.0005	0.999
15.54	0.00005	0.9999

$df = 5$

t	$P(T_5 \leq -t)$	$P(-t \leq T_5 \leq t)$
1.476	0.1	0.8
2.015	0.05	0.9
2.571	0.025	0.95
3.365	0.01	0.98
4.032	0.005	0.99
5.893	0.001	0.998
6.869	0.0005	0.999
11.18	0.00005	0.9999

$df = 6$

t	$P(T_6 \leq -t)$	$P(-t \leq T_6 \leq t)$
1.440	0.1	0.8
1.943	0.05	0.9
2.447	0.025	0.95
3.143	0.01	0.98
3.707	0.005	0.99
5.208	0.001	0.998
5.959	0.0005	0.999
9.082	0.00005	0.9999

$df = 7$

t	$P(T_7 \leq -t)$	$P(-t \leq T_7 \leq t)$
1.415	0.1	0.8
1.895	0.05	0.9
2.365	0.025	0.95
2.998	0.01	0.98
3.499	0.005	0.99
4.785	0.001	0.998
5.408	0.0005	0.999
7.885	0.00005	0.9999

$df = 8$

t	$P(T_8 \leq -t)$	$P(-t \leq T_8 \leq t)$
1.397	0.1	0.8
1.860	0.05	0.9
2.306	0.025	0.95
2.896	0.01	0.98
3.355	0.005	0.99
4.501	0.001	0.998
5.041	0.0005	0.999
7.120	0.00005	0.9999

$df = 9$

t	$P(T_9 \leq -t)$	$P(-t \leq T_9 \leq t)$
1.383	0.1	0.8
1.833	0.05	0.9
2.262	0.025	0.95
2.821	0.01	0.98
3.250	0.005	0.99
4.297	0.001	0.998
4.781	0.0005	0.999
6.594	0.00005	0.9999

$df = 10$

t	$P(T_{10} \leq -t)$	$P(-t \leq T_{10} \leq t)$
1.415	0.1	0.8
1.895	0.05	0.9
2.365	0.025	0.95
2.998	0.01	0.98
3.499	0.005	0.99
4.785	0.001	0.998
5.408	0.0005	0.999
7.885	0.00005	0.9999

$df = 11$

t	$P(T_{11} \leq -t)$	$P(-t \leq T_{11} \leq t)$
1.397	0.1	0.8
1.860	0.05	0.9
2.306	0.025	0.95
2.896	0.01	0.98
3.355	0.005	0.99
4.501	0.001	0.998
5.041	0.0005	0.999
7.120	0.00005	0.9999

$df = 12$

t	$P(T_{12} \leq -t)$	$P(-t \leq T_{12} \leq t)$
1.383	0.1	0.8
1.833	0.05	0.9
2.262	0.025	0.95
2.821	0.01	0.98
3.250	0.005	0.99
4.297	0.001	0.998
4.781	0.0005	0.999
6.594	0.00005	0.9999

$df = 13$

t	$P(T_{13} \leq -t)$	$P(-t \leq T_{13} \leq t)$
1.350	0.1	0.8
1.771	0.05	0.9
2.160	0.025	0.95
2.650	0.01	0.98
3.012	0.005	0.99
3.852	0.001	0.998
4.221	0.0005	0.999
5.513	0.00005	0.9999

$df = 14$

t	$P(T_{14} \leq -t)$	$P(-t \leq T_{14} \leq t)$
1.345	0.1	0.8
1.761	0.05	0.9
2.145	0.025	0.95
2.624	0.01	0.98
2.977	0.005	0.99
3.787	0.001	0.998
4.140	0.0005	0.999
5.363	0.00005	0.9999

$df = 15$

t	$P(T_{15} \leq -t)$	$P(-t \leq T_{15} \leq t)$
1.341	0.1	0.8
1.753	0.05	0.9
2.131	0.025	0.95
2.602	0.01	0.98
2.947	0.005	0.99
3.733	0.001	0.998
4.073	0.0005	0.999
5.239	0.00005	0.9999

$df = 16$

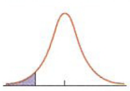
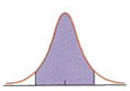
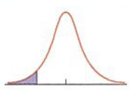
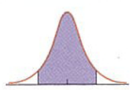
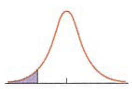
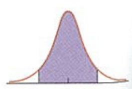
t	$P(T_{16} \leq -t)$	$P(-t \leq T_{16} \leq t)$
1.337	0.1	0.8
1.746	0.05	0.9
2.120	0.025	0.95
2.583	0.01	0.98
2.921	0.005	0.99
3.686	0.001	0.998
4.015	0.0005	0.999
5.134	0.00005	0.9999

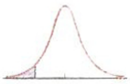
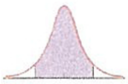
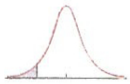
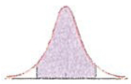
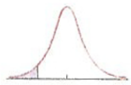
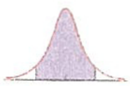
$df = 17$

t	$P(T_{17} \leq -t)$	$P(-t \leq T_{17} \leq t)$
1.333	0.1	0.8
1.740	0.05	0.9
2.110	0.025	0.95
2.567	0.01	0.98
2.898	0.005	0.99
3.646	0.001	0.998
3.965	0.0005	0.999
5.044	0.00005	0.9999

$df = 18$

t	$P(T_{18} \leq -t)$	$P(-t \leq T_{18} \leq t)$
1.33	0.1	0.8
1.734	0.05	0.9
2.101	0.025	0.95
2.552	0.01	0.98
2.878	0.005	0.99
3.610	0.001	0.998
3.922	0.0005	0.999
4.966	0.00005	0.9999

$df = 19$

t	$P(T_{19} \leq -t)$	$P(-t \leq T_{19} \leq t)$
1.328	0.1	0.8
1.729	0.05	0.9
2.093	0.025	0.95
2.539	0.01	0.98
2.861	0.005	0.99
3.579	0.001	0.998
3.883	0.0005	0.999
4.897	0.00005	0.9999

$df = 20$

t	$P(T_{20} \leq -t)$	$P(-t \leq T_{20} \leq t)$
1.325	0.1	0.8
1.725	0.05	0.9
2.086	0.025	0.95
2.528	0.01	0.98
2.845	0.005	0.99
3.552	0.001	0.998
3.850	0.0005	0.999
4.837	0.00005	0.9999

$df = 22$

t	$P(T_{22} \leq -t)$	$P(-t \leq T_{22} \leq t)$
1.321	0.1	0.8
1.717	0.05	0.9
2.074	0.025	0.95
2.508	0.01	0.98
2.819	0.005	0.99
3.505	0.001	0.998
3.792	0.0005	0.999
4.736	0.00005	0.9999

$df = 24$

t	$P(T_{24} \leq -t)$	$P(-t \leq T_{24} \leq t)$
1.318	0.1	0.8
1.711	0.05	0.9
2.064	0.025	0.95
2.492	0.01	0.98
2.797	0.005	0.99
3.467	0.001	0.998
3.745	0.0005	0.999
4.654	0.00005	0.9999

$df = 26$

t	$P(T_{26} \leq -t)$	$P(-t \leq T_{26} \leq t)$
1.315	0.1	0.8
1.706	0.05	0.9
2.056	0.025	0.95
2.479	0.01	0.98
2.779	0.005	0.99
3.435	0.001	0.998
3.707	0.0005	0.999
4.587	0.00005	0.9999

$df = 28$

t	$P(T_{28} \leq -t)$	$P(-t \leq T_{28} \leq t)$
1.313	0.1	0.8
1.701	0.05	0.9
2.048	0.025	0.95
2.467	0.01	0.98
2.763	0.005	0.99
3.408	0.001	0.998
3.674	0.0005	0.999
4.530	0.00005	0.9999

$df = 30$

t	$P(T_{30} \leq -t)$	$P(-t \leq T_{30} \leq t)$
1.31	0.1	0.8
1.697	0.05	0.9
2.042	0.025	0.95
2.457	0.01	0.98
2.75	0.005	0.99
3.385	0.001	0.998
3.646	0.0005	0.999
4.482	0.00005	0.9999

$df = 32$

t	$P(T_{32} \leq -t)$	$P(-t \leq T_{32} \leq t)$
1.309	0.1	0.8
1.694	0.05	0.9
2.037	0.025	0.95
2.449	0.01	0.98
2.738	0.005	0.99
3.365	0.001	0.998
3.622	0.0005	0.999
4.441	0.00005	0.9999

$df = 34$

t	$P(T_{34} \leq -t)$	$P(-t \leq T_{34} \leq t)$
1.307	0.1	0.8
1.691	0.05	0.9
2.032	0.025	0.95
2.441	0.01	0.98
2.728	0.005	0.99
3.348	0.001	0.998
3.601	0.0005	0.999
4.405	0.00005	0.9999

$df = 36$

t	$P(T_{36} \leq -t)$	$P(-t \leq T_{36} \leq t)$
1.306	0.1	0.8
1.688	0.05	0.9
2.028	0.025	0.95
2.434	0.01	0.98
2.719	0.005	0.99
3.333	0.001	0.998
3.582	0.0005	0.999
4.374	0.00005	0.9999

$df = 40$

t	$P(T_{40} \leq -t)$	$P(-t \leq T_{40} \leq t)$
1.303	0.1	0.8
1.684	0.05	0.9
2.021	0.025	0.95
2.423	0.01	0.98
2.704	0.005	0.99
3.307	0.001	0.998
3.551	0.0005	0.999
4.321	0.00005	0.9999

$df = 50$

t	$P(T_{50} \leq -t)$	$P(-t \leq T_{50} \leq t)$
1.299	0.1	0.8
1.676	0.05	0.9
2.009	0.025	0.95
2.403	0.01	0.98
2.678	0.005	0.99
3.261	0.001	0.998
3.496	0.0005	0.999
4.228	0.00005	0.9999

$df = 60$

t	$P(T_{60} \leq -t)$	$P(-t \leq T_{60} \leq t)$
1.296	0.1	0.8
1.671	0.05	0.9
2.000	0.025	0.95
2.390	0.01	0.98
2.660	0.005	0.99
3.232	0.001	0.998
3.460	0.0005	0.999
4.169	0.00005	0.9999

$df = 75$

t	$P(T_{75} \leq -t)$	$P(-t \leq T_{75} \leq t)$
1.293	0.1	0.8
1.665	0.05	0.9
1.992	0.025	0.95
2.377	0.01	0.98
2.643	0.005	0.99
3.202	0.001	0.998
3.425	0.0005	0.999
4.110	0.00005	0.9999

$df = 100$

t	$P(T_{100} \leq -t)$	$P(-t \leq T_{100} \leq t)$
1.290	0.1	0.8
1.660	0.05	0.9
1.984	0.025	0.95
2.364	0.01	0.98
2.626	0.005	0.99
3.174	0.001	0.998
3.390	0.0005	0.999
4.053	0.00005	0.9999

$df = 125$

t	$P(T_{125} \leq -t)$	$P(-t \leq T_{125} \leq t)$
1.288	0.1	0.8
1.657	0.05	0.9
1.979	0.025	0.95
2.357	0.01	0.98
2.616	0.005	0.99
3.157	0.001	0.998
3.370	0.0005	0.999
4.020	0.00005	0.9999

$df = 150$

t	$P(T_{150} \leq -t)$	$P(-t \leq T_{150} \leq t)$
1.287	0.1	0.8
1.655	0.05	0.9
1.976	0.025	0.95
2.351	0.01	0.98
2.609	0.005	0.99
3.145	0.001	0.998
3.357	0.0005	0.999
3.998	0.00005	0.9999

$df = \infty$

t	$P(Z \leq -t)$	$P(-t \leq Z \leq t)$
1.282	0.1	0.8
1.645	0.05	0.9
1.960	0.025	0.95
2.326	0.01	0.98
2.576	0.005	0.99
3.090	0.001	0.998
3.291	0.0005	0.999
3.891	0.00005	0.9999