STA380 Time Series Analytics | Homework on Random Samples
Submitted by: Ruchi Sharma [rs58898]

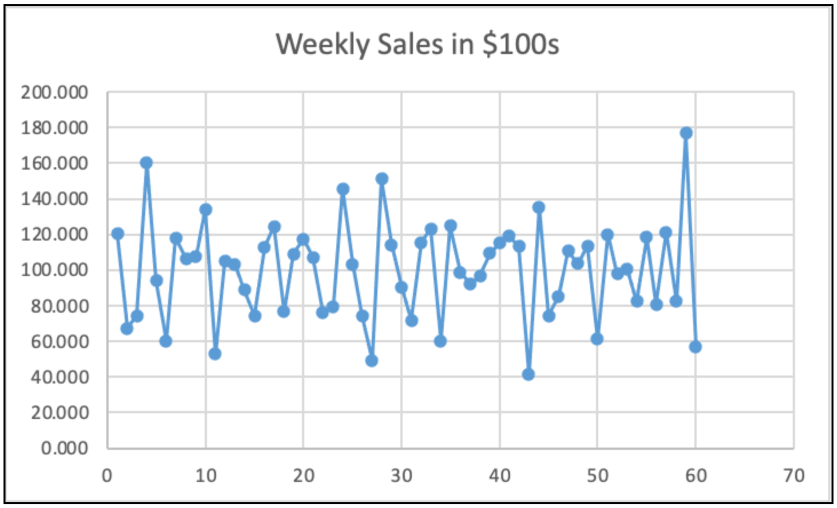
<u>Firm 01</u>

<u>Ques 01:</u> Test the data visually for conformity with the L specification of the Random Sample model. Explain clearly whether you think L applies and why (no hedging – you must indicate a definite yes or no.)

Yes, L applies.

By making a plot of the weekly sales over time and examining the general trend from early to late, we can visualize that the time series appears to be leveled.

The trend in means remains roughly leveled indicating consistency with L specification of the Random Sample Model.

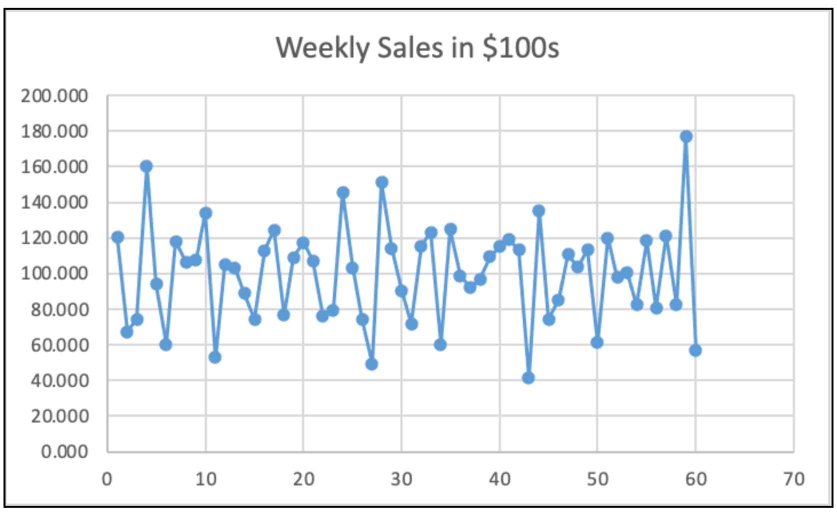


<u>Ques 02:</u> Test the data visually for conformity with the E specification of the Random Sample model. Explain clearly whether you think E applies and why (no hedging – you must indicate a definite yes or no.)

Yes, E applies.

By making a plot of the weekly sales over time and examining the general pattern of up-and-down variation around the overall level from early to late, we can visualize that the time series is consistent with E.

The pattern of variation appears to be roughly constant from start to end.



<u>Ques 03:</u> Test the data quantitatively for conformity with the "I" specification of the Random Sample model. Explain clearly whether you think "I" applies and why (no hedging –you must indicate a definite yes or no.)

No, I doesn't apply.

Performing the Quantitative Test for I:

1. Take the lag(1) shifted sales.
2. Calculate lag(1) autocorrelation for doing a quantitive test by using CORREL(Series1, Series2) in Excel.
   AutoCorr = -0.389
3. Check if AutoCorr is close to 0. If the value is close to zero then I is satisfied.
4. The calculation below shows that we reject the null hypothesis that AutoCorr is zero. Thus, I is NOT satisfied.

Week	Weekly Sales in $1000s	Lag 1 Weekly Sales in $1000s
1	120.747	
2	67.379	120.747
3	74.456	67.379
4	160.120	74.456
5	94.002	160.120
6	60.157	94.002
7	118.043	60.157
8	106.523	118.043
9	107.573	106.523
10	133.986	107.573

… total 60 Weeks

Lag 1 Autocorrelation (Week 2 to 60)
-0.38939791
Number of Pairs = 59
2/SQRT(N)
0.260377822

Is AutoCorr close to 0?
HO: AutoCorr = 0
HA: AutoCorr <> 0
Lag 1 Autocorrelation (Week 2 to 60)
-0.38939791
Number of Pairs = 59
StDev(AutoCorr) = 1/SQRT(N)
0.130188911
T-Stat(Z)
-2.991022098
Since |T(Z)| > 2, we reject HO

Ques 04: Suppose that the Random Sample model is valid for these data. In addition, suppose that weekly sales are normally distributed. Forecast sales for week 61 and give an interval in which you can have approximately 95% confidence that actual sales for week 61 will lie.

Assuming the Random Sample model is valid, the forecast for the Week 61 is going to be the Mean Value from Week 1 to Week 60. Thus, Forecasted Sales for Week 61 = $100034 with 95% confidence interval (44.840, 155.229) in $100s.

Mean (=AVERAGE) from Week 1 to Week 60
100.034
Sample Standard Deviation (=STDEV.S)
27.59736171
2*StDev
55.19472343
For 95% (-2*StDev, +2*StDev)
44.840
155.229

Ques 05: Suppose that the Random Sample model is valid for these data. In addition, suppose that weekly sales are normally distributed. Forecast sales for week 62 and give an interval in which you can have approximately 95% confidence that actual sales for week 62 will lie.

Assuming the Random Sample model is valid, the forecast for the Week 62 is going to be the Mean Value from Week 1 to Week 61. Thus, Forecasted Sales for Week 61 = $ 100034 with 95% confidence interval (45.301, 154.767) in $100s.

Mean (=AVERAGE from Week 1 to 61)
100.034
Sample Standard Deviation (=STDEV.S)
27.36641739
2*StDev
54.73283478
For 95% (-2*StDev, +2*StDev)
45.301
154.767

<u>Ques 06:</u> Suppose that the Random Sample model is valid for these data. Estimate mean sales per week and give an interval in which you can have approximately 95% confidence that the actual mean of the distribution of weekly sales will lie.

Mean sales = $ 100034, and 95% confidence interval = (93.051, 107.017) in $100s

Mean (=AVERAGE) from Week 1 to Week 60
100.034
Sample Standard Deviation (=STDEV.S)
27.36641739
Std Error (StDev/Sqrt(N))
3.562804077
For 95% (-2*SE, +2*SE)
93.051
107.017

Firm 02

<u>Ques 07:</u> Regress weekly sales as Y on the week number as X. Test the residuals visually for conformity with the L specification of the Random Sample model. Explain clearly whether you think L applies and why (no hedging – you must indicate a definite yes or no.)

Regressing weekly sales as Y on the week number as X gives the following regression output:

SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.743373065							
R Square	0.552603514							
Adjusted R Squar	0.544889782							
Standard Error	18.82253648							
Observations	60							
ANOVA								
	df	*SS*	*MS*	*F*	*Significance F*			
Regression	1	25380.80323	25380.80323	71.63892614	1.02695E-11			
Residual	58	20548.69701	354.2878794					
Total	59	45929.50023						
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	94.13237979	4.921349398	19.1273515	9.98299E-27	84.28122866	103.9835309	84.28122866	103.9835309
X Variable 1	1.187617957	0.140314393	8.463978151	1.02695E-11	0.906748183	1.468487731	0.906748183	1.468487731

Calculating Residuals as Y_Actual - Y_Predicted

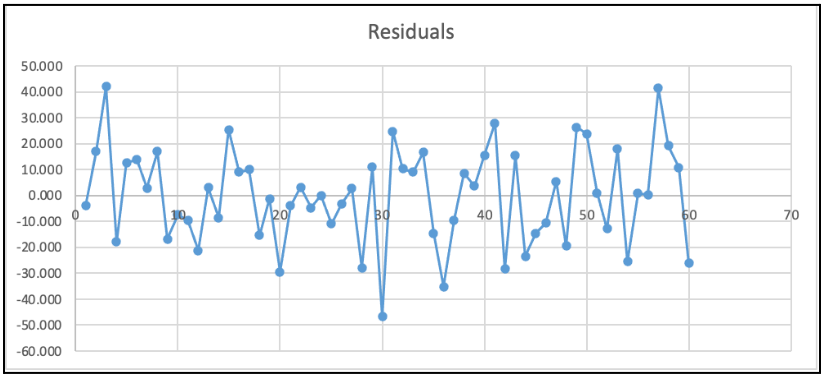
Week	Weekly Sales in $1000s	(Coeff)*X + Intercept	Residual	Lag(1) Residual
1	91.349	95.31999775	3.971	
2	113.514	96.5076157	-17.006	3.971
3	139.758	97.69523366	-42.063	-17.006
4	80.972	98.88285162	17.911	-42.063
5	112.844	100.0704696	-12.773	17.911
6	115.060	101.2580875	-13.802	-12.773
7	105.344	102.4457055	-2.898	-13.802
8	120.578	103.6333234	-16.945	-2.898
9	87.967	104.8209414	16.854	-16.945
10	98.499	106.0085594	7.510	16.854

… total 60 Weeks

Yes, L applies.

By making a plot of the residuala over time and examining the general trend from early to late, we can visualize that the time series appears to be leveled.

The trend in means remains roughly leveled indicating consistency with L specification of the Random Sample Model.

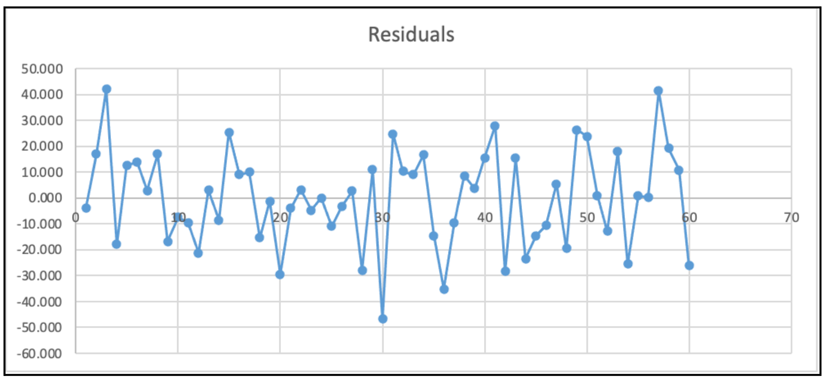


<u>Ques 08:</u> Regress weekly sales as Y on the week number as X. Test the residuals visually for conformity with the E specification of the Random Sample model. Explain clearly whether you think E applies and why (no hedging – you must indicate a definite yes or no.)

Yes, E applies.

By making a plot of residuals over time and examining the general pattern of Up-and-down variation around the overall level from early to late, we can visualize that the time series is consistent with E.

The pattern of variation appears to be roughly constant from start to end.



Ques 09: Regress weekly sales as Y on the week number as X. Test the residuals quantitatively for conformity with the "I" specification of the Random Sample model. Explain clearly whether you think "I" applies and why (no hedging – you must indicate a definite yes or no.)

Yes, I applies.

Performing the Quantitative Test for I:

1. Take the lag(1) shifted sales.
2. Calculate lag(1) autocorrelation for doing a quantitive test by using CORREL(Series1, Series2) in Excel.
   AutoCorr = -0.0761
3. Check if AutoCorr is close to 0. If the value is close to zero then I is satisfied.
4. The calculation below shows that we fail to reject the null hypothesis that AutoCorr is zero. Thus, I is satisfied.

Lag 1 Autocorrelation (Week 2 to 60)
-0.076130585
Number of Pairs = 59
2/SQRT(N)
0.260377822

Is AutoCorr close to 0?
HO: AutoCorr = 0
HA: AutoCorr <> 0
Lag 1 Autocorrelation (Week 2 to 60)
-0.076130585
Number of Pairs = 59
StDev(AutoCorr) = 1/SQRT(N)
0.130188911
T-Stat(Z)
-0.584770122
Since |T(Z)| < 2, we fail to reject HO

Ques 10: Suppose that $Y_1, Y_2, Y_3, \ldots Y_n$ is a Random Sample time series of non-degenerate random variables (i.e., the variance of each is positive [not zero]). Consider the time series of running totals of $Y_1, Y_2, Y_3, \ldots Y_n$ namely, Y_1, $Y_1 + Y_2$, $Y_1 + Y_2 + Y_3$, ... Is the time series of running totals …

(A) ... always a Random Sample?
(B) ... never a Random Sample?
(C) ... sometimes a Random Sample and sometimes not a Random Sample?

If your answer is (A) or (B), present a compelling argument for your position. Your argument need not be a mathematical proof – a clear intuitive argument will suffice. If your answer is (C), present an example of a Random Sample time series whose running totals are also a Random Sample, and a different example of a Random Sample whose running totals are not a Random Sample.

My answer is (B)

In the time series of running totals, the next point in time clearly depends on the previous point in time. Thus indicating the presence of autocorrelation which means the conditions of LIE will not be entirely satisfied. Thus, we can safely consider that it will never be a random sample.