Statistics – Regression Analysis

Michael D. Schmidt

- Course Number and Name: 3065 Business Statistics
- Program/Major: MBA / Management
- Submission Date: July 12, 2003
- Date Course was Started: June 10, 2003
- Approach to Course: Practical
- Number of Words in the Course Paper: 2031
- Other Documents Submitted for this Course: Journal
- Number of Hours Spent on this Course: 71
- Advisor: Dr. Mark Spence
- Date of Last Edit/Editor: July 13, 2003 / Laurel Barley
- English Spelling Used: USA English
- Permission to Publish on the Rushmore Website: Yes
- Resources: *Statistics for Business and Economics* by David R. Anderson, Dennis J. Sweeney and Thomas A. Williams, *Basic Statistics A Real World Approach* by Vincent E. Cangelosi, Phillip H. Taylor and Phillip R. Rice, *Statistics for Business and Economics* by James McClave, George Benson and Terry Cincich.
- Reasons for taking this course: I want to learn how to apply statistical methods for business applications. My goal is to gain a basic understanding of statistical calculations and apply the appropriate method(s) for predicting hiring goals for the technical staffing industry.

Executive Summary

Statistics involves facts and figures, such as the daily stock market results, but it involves more than just reporting facts and figures and can be defined more broadly as "the art and science of collecting, analyzing, presenting, and interpreting data."[1] The reason for collecting, analyzing, presenting and interpreting data is to gain understanding for the purpose of making sound business decisions. My goal for this course was to learn basic statistical methods and determine which method(s) to apply in analyzing data related to my company's business, the technical staffing industry. The data studied relates to the hiring of temporary employees and the goal was to predict hiring levels based on past historical data. Regression analysis was determined to be the statistical approach and is presented in the following:

[1] Anderson, Sweeney and Williams, *Statistics for Business and Economics,* 3.

1 The Project

1.1 The Project data set

The data used in this project is comprised of two variables – the number of job orders and the number of hires. A 'job order' is defined as a request from a client for one employee to be hired and work at the client's facility. A 'hire' is defined as one employee that is recruited and put on the company's payroll to work at a client's facility. The data set is a record of the weekly numbers of job orders and hires over a period of 60 consecutive weeks. See Appendix A.

1.2 The Project objective

The objective of this project is to determine if there is a correlation between the number of job orders and the resulting number of hires. Once a relationship is established and the regression equation is determined, I want to predict the number of job orders, the independent variable, required for obtaining a specific number of hires, the dependent variable. In other words, I will have the means to answer the question, "how many hires would result from *y* number of job orders?"

1.3 The Project Approach

Correlation and regression analysis involves determining and measuring the relationship between two or more variables. Regression deals with determining a quantitative expression to describe the relationship of the variables. The purpose of correlation is to determine the degree of the relationship.[2]

The approach used will involve the following steps:

1. Plot scatter diagram
2. Compute regression equation
3. Compute standard error of estimate
4. Test hypothesis concerning β
5. Compute correlation coefficient
6. Compute confidence intervals

[2] Cangelosi, Taylor and Rice, *Basic Statistics*, p.315.

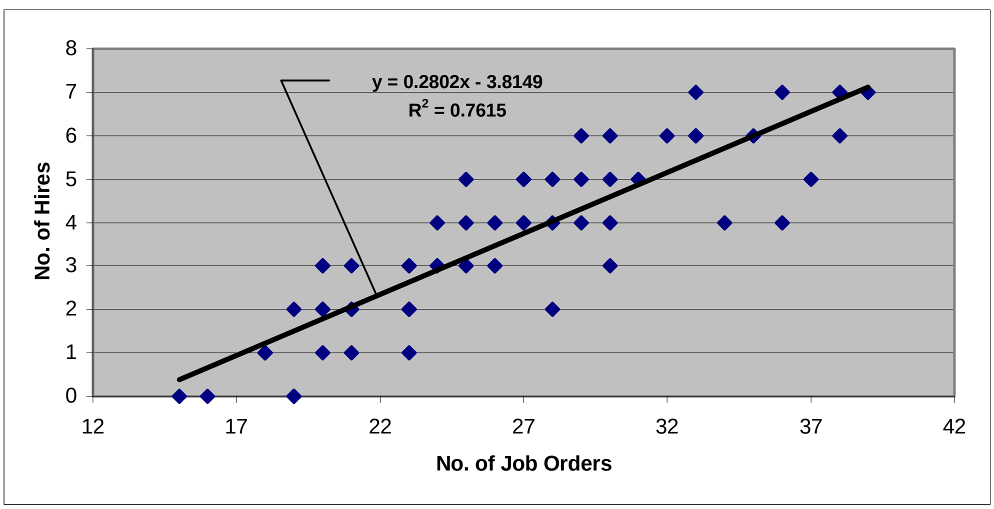
2 The Solution

2.1 Scatter Diagram

The purpose of the scatter diagram is to provide a picture of any relationship between the independent and dependent variables.[3] The plot of the data below indicates that there is a possible positive linear relationship.

2.2 Estimated Regression Equation

The estimated regression equation represents the equation of the line that best fits the data.



[3] Cangelosi, Taylor and Rice, *Basic Statistics*, p.318.

$y = \beta_0 + \beta_1 x + \varepsilon$ **Simple Linear Regression Model**

$E(y) = \beta_0 + \beta_1 x$ **Simple Linear Regression Equation**

$\hat{y} = b_0 + b_1 x$ **Estimated Simple Linear Regression Equation**

where: $\beta_0 = b_0 = y$ intercept of the regression line
$\beta_1 = b_1 =$ the slope of the regression line
$\varepsilon =$ error term (assumed to have constant variance with a mean of zero)

$b_0 = \bar{y} - b_1 \bar{x}$ where: $x_i =$ any and all values for the independent variable
$y_i =$ any and all values for the dependent variable
$\bar{x} =$ mean for all x_i
$\bar{y} =$ mean for all y_i
$n =$ total number of observations

$$b_1 = \frac{\sum x_i y_i - \frac{\sum x_i \sum y_i}{n}}{\sum x_i^2 - \frac{\left(\sum x_i\right)^2}{n}}$$

$\bar{x} = 1563/60 = 26.05$

$\bar{y} = 209/60 = 3.483$

$$b_1 = \frac{6056 - (209)(1563)/60}{42899 - (1563)^2/60} = 0.2802$$

$b_0 = 3.483 - (0.2802)(26.05) = -3.816$

$\hat{y} = b_0 + b_1 x \Rightarrow \hat{y} =$ 0.2802x - 3.816 **Regression Equation**

2.3 Testing for Significance

The estimated regression equation was determined. It approximated the relationship between the number of job orders and number of hires. The question is how well does the estimated regression equation fit the data? There are several tests that can be used. They are the coefficient of determination, t-test and F-test.

2.3.1 The Coefficient of Determination

The coefficient of determination provides a measure of the goodness of fit for the estimated regression equation. How well does the estimated regression equation fit the data?[4]

$$SSE = \sum (y_i - \hat{y}_i)^2$$

Sum of Squares due to Error – measures the variability of the actual observations about the estimated regression line. SSE can be thought of as the unexplained portions of SST.

$$SST = \sum (y_i - \bar{y})^2$$

Total Sum of Squares

$$SST = \sum y_i^2 - \frac{\left(\sum y_i\right)^2}{n}$$

Computational Formula for SST

$$SST = 953 - (209)^2 / 60 = 224.98$$

$$SSR = \sum (\hat{y}_i - \bar{y})^2$$

Sum of Squares due to Regression – can be thought of as the *explained* portion of SST

$$SSR = \frac{\left[\sum x_i y_i - \left(\sum x_i \sum y_i\right)/n\right]^2}{\sum x_i^2 - \left(\sum x_i\right)^2 / n}$$

Computational Formula for SSR

$$SSR = \frac{\left[6056 - (209)(1563)/60\right]^2}{42899 - (1563)^2 / 60} = 171.33$$

$$r^2 = \frac{SSR}{SST}$$

Coefficient of Determination – interpreted as the percentage of the total sum of squares that can be *explained* by using the regression equation.

$$r^2 = \frac{171.33}{224.98} = 0.7615$$

$$r_{xy} = \sqrt{r^2} = \sqrt{0.7615} = 0.8727$$

Correlation Coefficient

---

[4] Anderson, Sweeney and Williams, *Statistics for Business and Economics,* 558.

Conclusion

A coefficient of determination of 0.7615 (76.15%) shows a relatively good fit for the estimated regression equation. The higher the percentage, the better the 'goodness of fit.'

A correlation coefficient of 0.8727 (87.27%) indicates a strong relationship between *x* and *y*.

2.3.2 Standard Error of the Estimate

$$s^2 = MSE = \frac{SSE}{n-2}$$ **Mean Square Error**

$$s = \sqrt{MSE} = \sqrt{\frac{SSE}{n-2}}$$ **Standard Error of the Estimate**

$SST = SSR + SSE$ and $SSE = SST - SSR$

$$SSE = 224.98 - 171.33 = 53.65$$

$$s = \sqrt{\frac{53.65}{60-2}} = \sqrt{0.925} = 0.9618$$

2.3.3 t-test for Significance in Linear Regression

The purpose of the t-test is to determine if $\beta_1 \neq 0$ and reject the null hypothesis. This will support the premise that job orders and hires are linearly related.

Model for t-test

$H_0 : \beta_1 = 0$

$H_a : \beta_1 \neq 0$

Rejection Rule $\Rightarrow$ Reject H_0 if $t < -t_{a/2}$ or if $t > t_{a/2}$

where: $t = \frac{b_1}{s_{b_1}}$ and, $t_{a/2}$ will be found using the table in Appendix B

Test statistic $\Rightarrow$ $t = \frac{b_1}{s_{b_1}}$

where: $b_1 =$ slope

$$s_{b_1} = \frac{s}{\sqrt{\sum x_i^2 - \left(\sum x_i\right)^2 / n}}$$ Estimated Standard Deviation of b_1

Actual t-test for data with α = .05

$H_0 : \beta_1 = 0$
$H_a : \beta_1 \neq 0$
Rejection Rule ⇒ Reject H_0 if $t < -2.00$ or if $t > 2.00$ *

$$s_{b_1} = \frac{0.9618}{\sqrt{42899 - (1563)^2 / 60}} = \frac{0.9618}{46.721} = 0.0206$$

$$t = \frac{0.2802}{0.0206} = 13.602$$

Conclusion
Reject H_0 because 13.602 > 2.00 and conclude that $\beta_1 \neq 0$
Therefore, the statistical evidence is sufficient to conclude that there is a significant relationship between job orders and hires.

* Using the table in Appendix B we find that the one-tail t-value corresponding to α/2 = .025 (with α = .05) and n - 2 = 60 - 2 = 58 degrees of freedom is $t_{.025} = 2.00$.

2.3.4 F-Test for Significance in Linear Regression

An F test can be used to test for significance in regression. With only one independent variable, the F test will produce the same conclusion as the t test. Only the F test can be used for determining overall significant relationship when two or more variables are involved.

Model for F-Test

$H_0 : \beta_1 = 0$
$H_a : \beta_1 \neq 0$
Rejection Rule ⇒ Reject H_0 if $F > F_a$
where: F_a will be found using the table in Appendix C

Test statistic ⇒ $F = \frac{MSR}{MSE}$

where: $MSR = SSR$ / number of independent variables
$= SSR$ **Mean Square due to Regression**
$MSE = SSE / n - 2$ **Mean Square Error**

Actual F-test for data with α = .05

$H_0 : \beta_1 = 0$
$H_a : \beta_1 \neq 0$
Rejection Rule ⇒ Reject H_0 if $F >$ 4.007 *

$$F = \frac{171.33}{\left(53.65/(60-2)\right)} = \frac{171.33}{0.925} = 185.22$$

Conclusion
Reject H_0 because 185.22 > 4.007 and conclude that $\beta_1 \neq 0$

* From Appendix C we find that the *F* value corresponding to $\alpha = .05$ with one degree of freedom in the numerator and (n - 2) = (60 - 2) = 58 degrees of freedom in the denominator is $F_{.05} = 4.007$

This leads to the conclusion that the relationship between *hires* and *job orders* is statistically significant. Anderson, Sweeney and Williams point out that some caution should be taken regarding the interpretation of significance tests. "Rejecting the null hypothesis $H_0 : \beta_1 = 0$ and concluding that the relationship between x and y is significant does not enable us to conclude that a cause-and-effect relationship is present between x and y. Concluding a cause-and-effect relationship is warranted only if the analyst has some type of theoretical justification that the relationship is in fact causal."[5,6]

2.4 Using Estimated Regression Equation for Estimation and Prediction

2.4.1 Point estimation

The estimated regression equation can be used to determine a point estimate of the mean value of *y* (hires) for a specific value of *x* (job orders). Example point estimate if *x* = 25:

$$\hat{y}_{25} = 0.2802x - 3.816$$
$$\hat{y}_{25} = (0.2802)(25) - 3.816$$
$$\hat{y}_{25} = 7.005 - 3.816$$
$$\hat{y}_{25} = 3.2$$

[5] Anderson, Sweeney and Williams, *Statistics for Business and Economics*, p.573.

[6] McClave, Benson and Sincich, *Statistics for Business and Economics*, p.478-483,492.

Point estimates do not provide any suggestion of the precision associated with an estimate, so interval estimates are calculated, of which there are two types:

1. ***Confidence Interval*** (mean value of y for a given value of x)
2. ***Prediction Interval*** (individual value of y to a given value of x)

2.4.2 Confidence Interval

Model for Confidence Interval

For regression equation $\hat{y}_p = b_0 + b_1 x_p$ where: $\hat{y}_p =$ point estimate for a given value of x_p

$x_p =$ specific value of independent variable x.

The confidence interval estimate of $E(y_p)$ is the expected value of the dependent variable y corresponding to a given x_p and is expressed as:

$$\hat{y}_p \pm t_{\alpha/2} s_{\hat{y}_p} \quad \text{or} \quad \hat{y}_p - t_{\alpha/2} s_{\hat{y}_p} \le E(y_p) \le \hat{y}_p + t_{\alpha/2} s_{\hat{y}_p}$$

where: $s_{\hat{y}_p} =$ estimated standard deviation of $\hat{y}_p$

$$s_{\hat{y}_p} = s\sqrt{\frac{1}{n} + \frac{(x_p - \bar{x})^2}{\sum x_i^2 - \left(\sum x_i\right)^2 / n}}$$

$t_{\alpha/2}$ will be found using the table in Appendix B based on a t-distribution with n – 2 degrees of freedom and confidence coefficient 1 - α.

Actual 95% Confidence Interval for data

We are striving for a 95% confidence interval estimate (α = .05) of the mean number of job orders needed for obtaining x_p = 25 job orders.

$$\hat{y}_p \pm t_{\alpha/2} s_{\hat{y}_p} \quad \text{where: } t_{\alpha/2} = t_{.025} = 2.00$$

$$s_{\hat{y}_p} = 0.9618\sqrt{\frac{1}{60} + \frac{(25 - 26.05)^2}{42899 - (1563)^2 / 60}} = 0.9618\sqrt{0.1717} = 0.126$$

$\hat{y}_{25} =$ 0.2802x – 3.816

$\hat{y}_{25} =$ (0.2802)(25) – 3.816

$\hat{y}_{25} =$ 3.2

Confidence Interval = $CI = \hat{y}_4 \pm t_{\alpha/2} s_{y_{25}}$

CI = 3.2 ± (2.00)(0.126)

CI = 3.2 ± 0.25

Conclusion

For 25 job orders, the 95% confidence interval for the expected number of hires is:

$3.0 \le E(y_{25}) \le 3.5$

It is unrealistic to have less than a whole hire. Therefore, I will conclude that 3 or 4 hires will result from 25 job orders.

2.4.3 Prediction Interval

Model for Prediction Interval

$\hat{y}_p \pm t_{\alpha/2} s_{ind}$ or $\hat{y}_p - t_{\alpha/2} s_{ind} \le E(y_p) \le \hat{y}_p + t_{\alpha/2} s_{ind}$

where: s_{ind} = estimated standard deviation of individual value of $\hat{y}_p$

$$s_{ind} = s\sqrt{1 + \frac{1}{n} + \frac{(x_p - \bar{x})^2}{\sum x_i^2 - \left(\sum x_i\right)^2 / n}}$$

$t_{\alpha/2}$ will be found using the table in Appendix B based on a t-distribution with n – 2 degrees of freedom and confidence coefficient 1 - α.

Actual 95% Prediction Interval for data

We are striving for a 95% prediction interval estimate (α = .05) of the mean number of hires achieved from x_p = 25 job orders.

$\hat{y}_p \pm t_{\alpha/2} s_{ind}$ where: $\hat{y}_{25} = 3.2$

$t_{\alpha/2} = t_{.025} = 2.00$

$$s_{ind} = 0.9618\sqrt{1 + \frac{1}{60} + \frac{(25 - 26.05)^2}{42899 - (1563)^2 / 60}} = 0.9618\sqrt{1.01717} = 0.970$$

Prediction Interval = $PI = \hat{y}_4 \pm t_{\alpha/2} s_{ind_4}$

PI = 3.2 ± (2.00)(0.97)

CI = 3.2 ± 1.94

Conclusion

For 25 job orders, the 95% prediction interval for expected number of hires is:

$1.3 \leq y_{25} \leq 5.1$

It is unrealistic to have less than a whole hire. Therefore, the prediction interval concludes that 1 to 5 hires can be expected from 25 job orders.

The prediction interval range is wider than the confidence interval because it is easier to estimate the mean value of *y* more precisely than we can predict any one individual value of *y*.[7]

3 Using MS-Excel

There are software packages on the market that calculate statistical formulas like those in this paper for regression analysis more quickly than it could be done manually. Microsoft Excel has an add-in statistical program called Analysis ToolPak that I used to compare its results to the manually calculated results in this paper. To check if this add-in has been installed on your computer, first open Excel, then click on Tools, and then click on Add-ins. If the *Analysis ToolPak* box is not checked do so and click 'OK' and then follow the prompts. Once installed, Analysis ToolPak can be accessed in Excel by clicking on Tools and then clicking on Data Analysis.

It was not difficult to learn how to use Analysis ToolPak for regression analysis. For a good tutorial on Analysis ToolPak for regression analysis see *How to Run Regression with Excel* by Naoya Kaneko at http://www.geocities.com/naoyakaneko/regression/index.html.

Excel also has the ability to do a scatter plot through its Chart Wizard and picking the chart type called XY (Scatter). Once the scatter plot is completed, right clicking on any data point will bring up a pop-up box that include the option called '*Add Trendline*.' '*Linear'* line-type should be chosen and under the 'options' tab the choice of '*Display equation on chart*' can be checked. This will produce the scatter plot that is show in section 2.2 of this paper.

[7] Anderson, Sweeney and Williams, *Statistics for Business and Economics*, p.579-80.

MS-Excel Regression Analysis Output

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.872659559
R Square	0.761534707
Adjusted R Square	0.757423236
Standard Error	0.961775626
Observations	60

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	171.332617	171.332617	185.221977	1.0486E-19
Residual	58	53.6507166	0.92501235		
Total	59	224.983333			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	-3.814867413	0.55043974	-6.93058141	3.8334E-09	-4.91669143	-2.71304
X Variable 1	0.280161257	0.02058552	13.6096281	1.0486E-19	0.23895489	0.321368

Conclusion: The calculated results and the Excel summary output above compare very closely.

4 Conclusion

I accomplished my goal for this course and I am pleased with the results of the project. I learned about regression analysis and how to apply it to the business in which I am involved. The results give me a real tool that can be used in my work. The project presented in this paper dealt with correlated data. However, in studying statistics for this course I understand that other data involved in my business have more complex variables that would require analysis beyond my capability. As a side note, I was introduced to MS-Word's equation editor, which is what I used to produce the equations in this paper – I found it a lot of fun to do.

Citations

Books

Anderson, David R., Sweeney, Dennis J., and Williams, Thomas A. *Statistics for Business and Economics.* Ohio: South-Western College Publishing, 1999.

Cangelosi, Vincent E., Taylor, Phillip H., and Rice, Philip F. *Basic Statistics A Real World Approach.* Third edition. Minnesota: West Publishing Co., 1983.

McClave, James T., Benson, P. George, and Sincich, Terry. *Statistics for Business and Economics.* New Jersey: Prentice-Hall, Inc., 2001.

Websites

http://erdos.math.unb.ca/~knight/

http://www.itl.nist.gov/div898/handbook/eda/section3/eda3673.htm - ONE-05-11-20

http://www.geocities.com/naoyakaneko/regression/index.html

Appendix A

	Y	X	Y2	X2	XY
Week	**No.of Hires**	**Job Orders**			
1	3	20	9	400	60
2	2	19	4	361	38
3	3	26	9	676	78
4	4	36	16	1296	144
5	5	37	25	1369	185
6	4	30	16	900	120
7	4	25	16	625	100
8	3	23	9	529	69
9	3	20	9	400	60
10	2	21	4	441	42
11	4	27	16	729	108
12	3	25	9	625	75
13	3	24	9	576	72
14	4	28	16	784	112
15	6	30	36	900	180
16	7	36	49	1296	252
17	4	29	16	841	116
18	5	25	25	625	125
19	5	27	25	729	135
20	6	29	36	841	174
21	4	27	16	729	108
22	3	21	9	441	63
23	2	20	4	400	40
24	1	18	1	324	18
25	2	21	4	441	42
26	3	24	9	576	72
27	5	29	25	841	145
28	5	30	25	900	150
29	4	26	16	676	104
30	7	38	49	1444	266
31	2	23	4	529	46
32	1	20	1	400	20
33	1	18	1	324	18
34	1	21	1	441	21
35	5	27	25	729	135
36	7	33	49	1089	231
37	6	32	36	1024	192
38	4	34	16	1156	136
39	6	33	36	1089	198
40	3	26	9	676	78
41	1	23	1	529	23
42	0	19	0	361	0
43	0	15	0	225	0
44	0	16	0	256	0
45	2	23	4	529	46
46	2	28	4	784	56
47	6	33	36	1089	198
48	6	38	36	1444	228
49	7	39	49	1521	273
50	3	30	9	900	90
51	2	23	4	529	46
52	1	18	1	324	18
53	2	20	4	400	40
54	5	28	25	784	140
55	3	23	9	529	69
56	6	35	36	1225	210
57	5	31	25	961	155
58	4	24	16	576	96
59	2	20	4	400	40
60	0	19	0	361	0
SUM =	**209**	**1563**	**953**	**42899**	**6056**

Appendix B

PERCENTAGE POINTS OF THE T DISTRIBUTION

Tail Probabilities

		One Tail	0.10	0.05	0.025	0.01	0.005	0.001	0.0005	
		Two Tails	0.20	0.10	0.05	0.02	0.01	0.002	0.001	
D	1		3.078	6.314	12.71	31.82	63.66	318.3	637	1
E	2		1.886	2.920	4.303	6.965	9.925	22.330	31.6	2
G	3		1.638	2.353	3.182	4.541	5.841	10.210	12.92	3
R	4		1.533	2.132	2.776	3.747	4.604	7.173	8.610	4
E	5		1.476	2.015	2.571	3.365	4.032	5.893	6.869	5
E	6		1.440	1.943	2.447	3.143	3.707	5.208	5.959	6
S	7		1.415	1.895	2.365	2.998	3.499	4.785	5.408	7
	8		1.397	1.860	2.306	2.896	3.355	4.501	5.041	8
O	9		1.383	1.833	2.262	2.821	3.250	4.297	4.781	9
F	10		1.372	1.812	2.228	2.764	3.169	4.144	4.587	10
	11		1.363	1.796	2.201	2.718	3.106	4.025	4.437	11
F	12		1.356	1.782	2.179	2.681	3.055	3.930	4.318	12
R	13		1.350	1.771	2.160	2.650	3.012	3.852	4.221	13
E	14		1.345	1.761	2.145	2.624	2.977	3.787	4.140	14
E	15		1.341	1.753	2.131	2.602	2.947	3.733	4.073	15
D	16		1.337	1.746	2.120	2.583	2.921	3.686	4.015	16
O	17		1.333	1.740	2.110	2.567	2.898	3.646	3.965	17
M	18		1.330	1.734	2.101	2.552	2.878	3.610	3.922	18
	19		1.328	1.729	2.093	2.539	2.861	3.579	3.883	19
	20		1.325	1.725	2.086	2.528	2.845	3.552	3.850	20
	21		1.323	1.721	2.080	2.518	2.831	3.527	3.819	21
	22		1.321	1.717	2.074	2.508	2.819	3.505	3.792	22
	23		1.319	1.714	2.069	2.500	2.807	3.485	3.768	23
	24		1.318	1.711	2.064	2.492	2.797	3.467	3.745	24
	25		1.316	1.708	2.060	2.485	2.787	3.450	3.725	25
	26		1.315	1.706	2.056	2.479	2.779	3.435	3.707	26
	27		1.314	1.703	2.052	2.473	2.771	3.421	3.690	27
	28		1.313	1.701	2.048	2.467	2.763	3.408	3.674	28
	29		1.311	1.699	2.045	2.462	2.756	3.396	3.659	29
	30		1.310	1.697	2.042	2.457	2.750	3.385	3.646	30
	32		1.309	1.694	2.037	2.449	2.738	3.365	3.622	32
	34		1.307	1.691	2.032	2.441	2.728	3.348	3.601	34
	36		1.306	1.688	2.028	2.434	2.719	3.333	3.582	36
	38		1.304	1.686	2.024	2.429	2.712	3.319	3.566	38
	40		1.303	1.684	2.021	2.423	2.704	3.307	3.551	40
	42		1.302	1.682	2.018	2.418	2.698	3.296	3.538	42
	44		1.301	1.680	2.015	2.414	2.692	3.286	3.526	44
	46		1.300	1.679	2.013	2.410	2.687	3.277	3.515	46
	48		1.299	1.677	2.011	2.407	2.682	3.269	3.505	48
	50		1.299	1.676	2.009	2.403	2.678	3.261	3.496	50
	55		1.297	1.673	2.004	2.396	2.668	3.245	3.476	55
	60		1.296	1.671	2.000	2.390	2.660	3.232	3.460	60
	65		1.295	1.669	1.997	2.385	2.654	3.220	3.447	65
	70		1.294	1.667	1.994	2.381	2.648	3.211	3.435	70

source: http://erdos.math.unb.ca/~knight/

Apendix C

Upper critical values of the F distribution

for ν_1 numerator degrees of freedom and ν_2 denominator degrees of freedom
5% significance level

$F_{.05}(\nu_1, \nu_2)$

ν_2 \ ν_1	1	2	3	4	5	6	7	8
1	161.448	199.500	215.707	224.583	230.162	233.986	236.768	
2	18.513	19.000	19.164	19.247	19.296	19.330	19.353	
3	10.128	9.552	9.277	9.117	9.013	8.941	8.887	
4	7.709	6.944	6.591	6.388	6.256	6.163	6.094	
5	6.608	5.786	5.409	5.192	5.050	4.950	4.876	
6	5.987	5.143	4.757	4.534	4.387	4.284	4.207	
7	5.591	4.737	4.347	4.120	3.972	3.866	3.787	
8	5.318	4.459	4.066	3.838	3.687	3.581	3.500	
9	5.117	4.256	3.863	3.633	3.482	3.374	3.293	
10	4.965	4.103	3.708	3.478	3.326	3.217	3.135	
11	4.844	3.982	3.587	3.357	3.204	3.095	3.012	
12	4.747	3.885	3.490	3.259	3.106	2.996	2.913	
13	4.667	3.806	3.411	3.179	3.025	2.915	2.832	
14	4.600	3.739	3.344	3.112	2.958	2.848	2.764	
15	4.543	3.682	3.287	3.056	2.901	2.790	2.707	
16	4.494	3.634	3.239	3.007	2.852	2.741	2.657	
17	4.451	3.592	3.197	2.965	2.810	2.699	2.614	
18	4.414	3.555	3.160	2.928	2.773	2.661	2.577	
19	4.381	3.522	3.127	2.895	2.740	2.628	2.544	
20	4.351	3.493	3.098	2.866	2.711	2.599	2.514	
21	4.325	3.467	3.072	2.840	2.685	2.573	2.488	
22	4.301	3.443	3.049	2.817	2.661	2.549	2.464	
23	4.279	3.422	3.028	2.796	2.640	2.528	2.442	
24	4.260	3.403	3.009	2.776	2.621	2.508	2.423	
25	4.242	3.385	2.991	2.759	2.603	2.490	2.405	
26	4.225	3.369	2.975	2.743	2.587	2.474	2.388	
27	4.210	3.354	2.960	2.728	2.572	2.459	2.373	
28	4.196	3.340	2.947	2.714	2.558	2.445	2.359	
29	4.183	3.328	2.934	2.701	2.545	2.432	2.346	
30	4.171	3.316	2.922	2.690	2.534	2.421	2.334	
35	4.121	3.267	2.874	2.641	2.485	2.372	2.285	
40	4.085	3.232	2.839	2.606	2.449	2.336	2.249	
45	4.057	3.204	2.812	2.579	2.422	2.308	2.221	
50	4.034	3.183	2.790	2.557	2.400	2.286	2.199	
55	4.016	3.165	2.773	2.540	2.383	2.269	2.181	
56	4.013	3.162	2.769	2.537	2.380	2.266	2.178	
57	4.010	3.159	2.766	2.534	2.377	2.263	2.175	
58	4.007	3.156	2.764	2.531	2.374	2.260	2.172	
59	4.004	3.153	2.761	2.528	2.371	2.257	2.169	
60	4.001	3.150	2.758	2.525	2.368	2.254	2.167	
65	3.989	3.138	2.746	2.513	2.356	2.242	2.154	
70	3.978	3.128	2.736	2.503	2.346	2.231	2.143	
75	3.968	3.119	2.727	2.494	2.337	2.222	2.134	
80	3.960	3.111	2.719	2.486	2.329	2.214	2.126	

source: http://www.itl.nist.gov/div898/handbook/eda/section3/eda3673.htm - ONE-05-11-20