KING FAHD UNIVERSITY OF PETROLEUM & MINERALS
COLLEGE OF COMPUTER SCIENCES & ENGINEERING

COMPUTER ENGINEERING DEPARTMENT

COE-587 – Performance Evaluation and Analysis
CSE-642 –Computer Systems Performance
April 24th, 2011 – Major Exam

Student Name:
Student Number:
Exam Time: 90 mins

- Do not open the exam book until instructed
- Answer all questions
- **<u>All steps must be shown</u>**
- Any assumptions made must be clearly stated

Question No.	Max Points	
1	30	
2	20	
3	40	
4	40	

Total: 130

Q.1) (30 points) On the subject of the art of workload selection and workload characterization

a) **(10 points)** Consider the following two sequences of data:

X =[77 69 85 69 62 61], and

Y = [72 72 70 65 52 51].

Compute the arithmetic mean, variance, and standard deviation for X and Y. Also compute the correlation among the variables X and Y, R_{xy}.

b) **(10 points)** If it is required to perform Principal Factor Analysis (PFA) for the data in part (a), then *specify* the correlation matrix and also the characteristic equation needed to find the eigen-values and eigen-vectors. **Note that is it NOT required to find the eigen-values and eigen-vectors**.

c) **(10 points)** What is the main purpose of the PFA?

Q.2) (20 points) On the subject of confidence interval and its applications, answer the following questions:

a) **(10 points)** A single experiment was repeated on two systems 40 times. System A was found to be superior to system B in 26 repetitions. Can we state with 99% confidence that system A is superior to system B.

b) **(10 points)** Assuming the average of the proportion is the same that for part (a), what would be the needed number of repetitions to state that system A is superior to system B with 99% confidence.

Q.3) (40 points) On the subject regression models:

Using the context of simple and other regression models answer the following questions:

a) **(20 points)** List and briefly explain the four visual tests that can be performed to verify the simple linear regression.

b) **(10 points)** Explain briefly the F-test and the involved quantities and degrees of freedom. How does this test relate to the computation of confidence intervals for the regression parameters b_i?

c) **(10 points)** What is referred to by transformations in the area of linear regression? Explain briefly how is the appropriate transformation found?

Q.4) (40 points) On the subject regression models:

The results of a multiple regression based on *nine* observations are shown in Table below. Based on these results answer the following questions:

a) **(5 points)** What percentage of variance is explained by the regression.

b) **(5 points)** Is the regression significant at the 90% coefficient and why? Indicate the degrees of freedom for the corresponding *F* parameter.

c) **(10 points)** Which variable is most significant? Why?

d) **(5 points)** Which parameters are not significant at 90%?

e) **(10 points)** What is the problem with this regression?

f) **(5 points)** What would you try next?

The solution must show all needed steps and calculations.

Table 1: Results of a multiple regression analysis

j	b_j	s_{b_j}
1	1.3	3.6
2	2.7	1.8
3	0.5	0.6
4	5.0	8.3

Intercept = 75.3

Coefficient of multiple correlation = 0.95

Standard deviation of errors = 12.0

F-value = 14.1

Tables/Formulas you may need:

A.1 AREA OF THE UNIT NORMAL DISTRIBUTION

Table A.1 lists area between 0 and z. For example, the area between $z = 0$ and $z = 1.03$ is 0.3485. Due to symmetry of the normal distribution, the area between $z = 0$ and $z = -1.03$ is also the same.

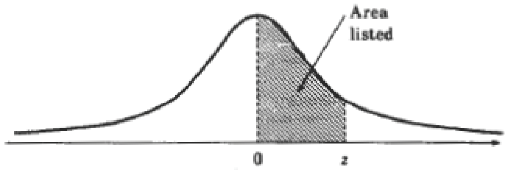

TABLE A.1 Area of the Unit Normal Distribution

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2703	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990

A.2 QUANTITIES OF THE UNIT NORMAL DISTRIBUTION

Table A.2 lists z_p for a given p. For example, for a two-sided confidence interval at 95%, $\alpha = 0.05$ and $p = 1 - \alpha/2 = 0.975$. The entry in the row labeled 0.97 and column labeled 0.005 gives $z_p = 1.960$.

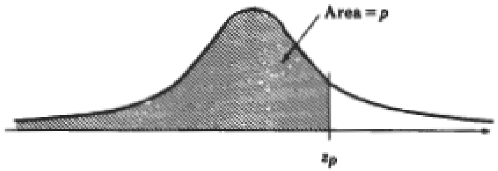

TABLE A.2 Quantiles of the Unit Normal Distribution

p	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.5	0.000	0.025	0.050	0.075	0.100	0.126	0.151	0.176	0.202	0.228
0.6	0.253	0.279	0.305	0.332	0.358	0.385	0.412	0.440	0.468	0.496
0.7	0.524	0.553	0.583	0.613	0.643	0.674	0.706	0.739	0.772	0.806
0.8	0.842	0.878	0.915	0.954	0.994	1.036	1.080	1.126	1.175	1.227

p	0.000	0.001	0.002	0.003	0.004	0.005	0.006	0.007	0.008	0.009
0.90	1.282	1.287	1.293	1.299	1.305	1.311	1.317	1.323	1.329	1.335
0.91	1.341	1.347	1.353	1.359	1.366	1.372	1.379	1.385	1.392	1.398
0.92	1.405	1.412	1.419	1.426	1.433	1.440	1.447	1.454	1.461	1.468
0.93	1.476	1.483	1.491	1.499	1.506	1.514	1.522	1.530	1.538	1.546
0.94	1.555	1.563	1.572	1.580	1.589	1.598	1.607	1.616	1.626	1.635
0.95	1.645	1.655	1.665	1.675	1.685	1.695	1.706	1.717	1.728	1.739
0.96	1.751	1.762	1.774	1.787	1.799	1.812	1.825	1.838	1.852	1.866
0.97	1.881	1.896	1.911	1.927	1.943	1.960	1.977	1.995	2.014	2.034
0.98	2.054	2.075	2.097	2.120	2.144	2.170	2.197	2.226	2.257	2.290

p	0.0000	0.0001	0.0002	0.0003	0.0004	0.0005	0.0006	0.0007	0.0008	0.0009
0.990	2.326	2.330	2.334	2.338	2.342	2.346	2.349	2.353	2.357	2.362
0.991	2.366	2.370	2.374	2.378	2.382	2.387	2.391	2.395	2.400	2.404
0.992	2.409	2.414	2.418	2.423	2.428	2.432	2.437	2.442	2.447	2.452
0.993	2.457	2.462	2.468	2.473	2.478	2.484	2.489	2.495	2.501	2.506
0.994	2.512	2.518	2.524	2.530	2.536	2.543	2.549	2.556	2.562	2.569
0.995	2.576	2.583	2.590	2.597	2.605	2.612	2.620	2.628	2.636	2.644
0.996	2.652	2.661	2.669	2.678	2.687	2.697	2.706	2.716	2.727	2.737
0.997	2.748	2.759	2.770	2.782	2.794	2.807	2.820	2.834	2.848	2.863
0.998	2.878	2.894	2.911	2.929	2.948	2.968	2.989	3.011	3.036	3.062
0.999	3.090	3.121	3.156	3.195	3.239	3.291	3.353	3.432	3.540	3.719

A.4 QUANTILES OF THE t DISTRIBUTION

Table A.4 lists $t_{[p;n]}$. For example, the $t_{[0.95;13]}$ required for a two-sided 90% confidence interval of the mean of a sample of 14 observation is 1.771.

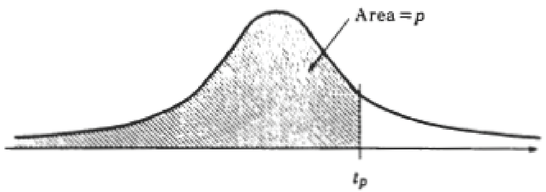

TABLE A.4 Quantiles of the t Distribution

n	p = 0.6000	0.7000	0.8000	0.9000	0.9500	0.9750	0.9950	0.9995
1	0.325	0.727	1.377	3.078	6.314	12.706	63.657	636.619
2	0.289	0.617	1.061	1.886	2.920	4.303	9.925	31.599
3	0.277	0.584	0.978	1.638	2.353	3.182	5.841	12.924
4	0.271	0.569	0.941	1.533	2.132	2.776	4.604	8.610
5	0.267	0.559	0.920	1.476	2.015	2.571	4.032	6.869
6	0.265	0.553	0.906	1.440	1.943	2.447	3.707	5.959
7	0.263	0.549	0.896	1.415	1.895	2.365	3.499	5.408
8	0.262	0.546	0.889	1.397	1.860	2.306	3.355	5.041
9	0.261	0.543	0.883	1.383	1.833	2.262	3.250	4.781
10	0.260	0.542	0.879	1.372	1.812	2.228	3.169	4.587
11	0.260	0.540	0.876	1.363	1.796	2.201	3.106	4.437
12	0.259	0.539	0.873	1.356	1.782	2.179	3.055	4.318
13	0.259	0.538	0.870	1.350	1.771	2.160	3.012	4.221
14	0.258	0.537	0.868	1.345	1.761	2.145	2.977	4.140
15	0.258	0.536	0.866	1.341	1.753	2.131	2.947	4.073
16	0.258	0.535	0.865	1.337	1.746	2.120	2.921	4.015
17	0.257	0.534	0.863	1.333	1.740	2.110	2.898	3.965
18	0.257	0.534	0.862	1.330	1.734	2.101	2.878	3.922
19	0.257	0.533	0.861	1.328	1.729	2.093	2.861	3.883
20	0.257	0.533	0.860	1.325	1.725	2.086	2.845	3.850
21	0.257	0.532	0.859	1.323	1.721	2.080	2.831	3.819
22	0.256	0.532	0.858	1.321	1.717	2.074	2.819	3.792
23	0.256	0.532	0.858	1.319	1.714	2.069	2.807	3.768
24	0.256	0.531	0.857	1.318	1.711	2.064	2.797	3.745
25	0.256	0.531	0.856	1.316	1.708	2.060	2.787	3.725
26	0.256	0.531	0.856	1.315	1.706	2.056	2.779	3.707
27	0.256	0.531	0.855	1.314	1.703	2.052	2.771	3.690
28	0.256	0.530	0.855	1.313	1.701	2.048	2.763	3.674
29	0.256	0.530	0.854	1.311	1.699	2.045	2.756	3.659
30	0.256	0.530	0.854	1.310	1.697	2.042	2.750	3.646
60	0.254	0.527	0.848	1.296	1.671	2.000	2.660	3.460
90	0.254	0.526	0.846	1.291	1.662	1.987	2.632	3.402
120	0.254	0.526	0.845	1.289	1.658	1.980	2.617	3.373

A.6 90-PERCENTILES OF THE $F(n, m)$ DISTRIBUTION

Table A.6 lists $F_{[0.90;n,m]}$. For example, the $F_{[0.90;9,18]}$ required for an F-test at 90% confidence level is 2.00.

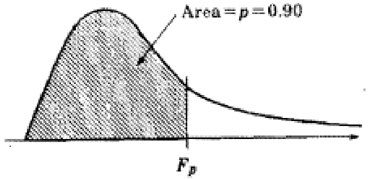

TABLE A.6 90-Percentiles of the $F(n, m)$ Distribution

m	Numerator Degrees of Freedom n: 1	2	3	4	5	6	7	8	9	10
1	39.86	49.50	53.59	55.83	57.24	58.20	58.90	59.44	59.86	60.19
2	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38	9.39
3	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24	5.23
4	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94	3.92
5	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32	3.30
6	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96	2.94
7	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72	2.70
8	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56	2.54
9	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44	2.42
10	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35	2.32
12	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21	2.19
14	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12	2.10
16	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06	2.03
18	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	2.00	1.98
20	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96	1.94
25	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89	1.87
30	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85	1.82
35	2.85	2.46	2.25	2.11	2.02	1.95	1.90	1.85	1.82	1.79
40	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79	1.76
500	2.72	2.31	2.09	1.96	1.86	1.79	1.73	1.68	1.64	1.61

m	Numerator Degrees of Freedom n: 12	14	16	18	20	25	30	35	40	500
1	60.70	61.07	61.35	61.56	61.74	62.05	62.26	62.41	62.53	63.26
2	9.41	9.42	9.43	9.44	9.44	9.45	9.46	9.46	9.47	9.49
3	5.22	5.20	5.20	5.19	5.18	5.17	5.17	5.16	5.16	5.13
4	3.90	3.88	3.86	3.85	3.84	3.83	3.82	3.81	3.80	3.76
5	3.27	3.25	3.23	3.22	3.21	3.19	3.17	3.16	3.16	3.10
6	2.90	2.88	2.86	2.85	2.84	2.81	2.80	2.79	2.78	2.72
7	2.67	2.64	2.62	2.61	2.59	2.57	2.56	2.54	2.54	2.47
8	2.50	2.48	2.45	2.44	2.42	2.40	2.38	2.37	2.36	2.29
9	2.38	2.35	2.33	2.31	2.30	2.27	2.25	2.24	2.23	2.16
10	2.28	2.26	2.23	2.22	2.20	2.17	2.16	2.14	2.13	2.06
12	2.15	2.12	2.09	2.08	2.06	2.03	2.01	2.00	1.99	1.90
14	2.05	2.02	2.00	1.98	1.96	1.93	1.91	1.90	1.89	1.80
16	1.99	1.95	1.93	1.91	1.89	1.86	1.84	1.82	1.81	1.72
18	1.93	1.90	1.87	1.85	1.84	1.80	1.78	1.77	1.75	1.66
20	1.89	1.86	1.83	1.81	1.79	1.76	1.74	1.72	1.71	1.61
25	1.82	1.79	1.76	1.74	1.72	1.68	1.66	1.64	1.63	1.52
30	1.77	1.74	1.71	1.69	1.67	1.63	1.61	1.59	1.57	1.46
35	1.74	1.70	1.67	1.65	1.63	1.60	1.57	1.55	1.53	1.41
40	1.71	1.68	1.65	1.62	1.61	1.57	1.54	1.52	1.51	1.38
500	1.56	1.52	1.49	1.46	1.44	1.39	1.36	1.33	1.31	2.16

A.7 95-PERCENTILES OF THE $F(n, m)$ DISTRIBUTION

Table A.7 lists $F_{[0.95;n,m]}$. For example, the $F_{[0.95;9,14]}$ required for an F-test at 95% confidence level is 2.70.

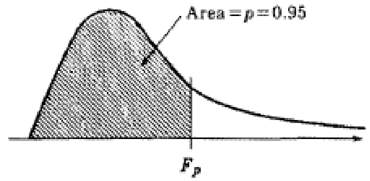

TABLE A.7 95-Percentiles of the $F(n, m)$ Distribution

m	Numerator Degrees of Freedom n: 1	2	3	4	5	6	7	8	9	10
1	161.45	199.50	215.70	224.57	230.15	233.97	236.76	238.87	240.53	241.87
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16
35	4.12	3.27	2.87	2.64	2.49	2.37	2.29	2.22	2.16	2.11
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08
500	3.86	3.01	2.62	2.39	2.23	2.12	2.03	1.96	1.90	1.85

m	Numerator Degrees of Freedom n: 12	14	16	18	20	25	30	35	40	500
1	243.90	245.35	246.45	247.31	248.00	249.25	250.09	250.68	251.13	254.05
2	19.41	19.42	19.43	19.44	19.45	19.46	19.46	19.47	19.47	19.49
3	8.74	8.71	8.69	8.67	8.66	8.63	8.62	8.60	8.59	8.53
4	5.91	5.87	5.84	5.82	5.80	5.77	5.75	5.73	5.72	5.63
5	4.68	4.64	4.60	4.58	4.56	4.52	4.50	4.48	4.46	4.36
6	4.00	3.96	3.92	3.90	3.87	3.83	3.81	3.79	3.77	3.67
7	3.57	3.53	3.49	3.47	3.44	3.40	3.38	3.36	3.34	3.23
8	3.28	3.24	3.20	3.17	3.15	3.11	3.08	3.06	3.04	2.93
9	3.07	3.03	2.99	2.96	2.94	2.89	2.86	2.84	2.83	2.71
10	2.91	2.86	2.83	2.80	2.77	2.73	2.70	2.68	2.66	2.54
12	2.69	2.64	2.60	2.57	2.54	2.50	2.47	2.44	2.43	2.30
14	2.53	2.48	2.44	2.41	2.39	2.34	2.31	2.28	2.27	2.13
16	2.42	2.37	2.33	2.30	2.28	2.23	2.19	2.17	2.15	2.01
18	2.34	2.29	2.25	2.22	2.19	2.14	2.11	2.08	2.06	1.92
20	2.28	2.22	2.18	2.15	2.12	2.07	2.04	2.01	1.99	1.84
25	2.16	2.11	2.07	2.04	2.01	1.96	1.92	1.89	1.87	1.71
30	2.09	2.04	1.99	1.96	1.93	1.88	1.84	1.81	1.79	1.62
35	2.04	1.99	1.94	1.91	1.88	1.82	1.79	1.76	1.74	1.56
40	2.00	1.95	1.90	1.87	1.84	1.78	1.74	1.72	1.69	1.51
500	1.77	1.71	1.66	1.62	1.59	1.53	1.48	1.45	1.42	2.21

A.8 99-PERCENTILES OF THE $F(n,m)$ DISTRIBUTION

Table A.8 lists $F_{[0.99;n,m]}$. For example, the $F_{[0.99;6,12]}$ required for an F-test at 99% confidence level is 4.82.

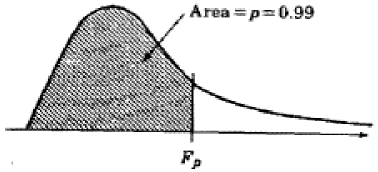

TABLE A.8 99-Percentiles of the $F(n,m)$ Distribution

	Numerator Degrees of Freedom n									
m	1	2	3	4	5	6	7	8	9	10
1	4052.18	4999.50	5403.05	5624.30	5763.37	5858.71	5928.09	5980.80	6022.21	6055.58
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39	99.40
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.34	27.23
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03	3.94
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22	3.13
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98
35	7.42	5.27	4.40	3.91	3.59	3.37	3.20	3.07	2.96	2.88
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80
500	6.69	4.65	3.82	3.36	3.05	2.84	2.68	2.55	2.44	2.36

	Numerator Degrees of Freedom n									
m	12	14	16	18	20	25	30	35	40	500
1	6106.06	6142.42	6169.85	6191.28	6208.48	6239.57	6260.40	6275.32	6286.53	6359.26
2	99.42	99.43	99.44	99.44	99.45	99.46	99.47	99.47	99.47	99.50
3	27.05	26.92	26.83	26.75	26.69	26.58	26.50	26.45	26.41	26.12
4	14.37	14.25	14.15	14.08	14.02	13.91	13.84	13.79	13.75	13.46
5	9.89	9.77	9.68	9.61	9.55	9.45	9.38	9.33	9.29	9.02
6	7.72	7.60	7.52	7.45	7.40	7.30	7.23	7.18	7.14	6.88
7	6.47	6.36	6.28	6.21	6.16	6.06	5.99	5.94	5.91	5.65
8	5.67	5.56	5.48	5.41	5.36	5.26	5.20	5.15	5.12	4.86
9	5.11	5.01	4.92	4.86	4.81	4.71	4.65	4.60	4.57	4.31
10	4.71	4.60	4.52	4.46	4.41	4.31	4.25	4.20	4.17	3.91
12	4.16	4.05	3.97	3.91	3.86	3.76	3.70	3.65	3.62	3.36
14	3.80	3.70	3.62	3.56	3.51	3.41	3.35	3.30	3.27	3.00
16	3.55	3.45	3.37	3.31	3.26	3.16	3.10	3.05	3.02	2.75
18	3.37	3.27	3.19	3.13	3.08	2.98	2.92	2.87	2.84	2.57
20	3.23	3.13	3.05	2.99	2.94	2.84	2.78	2.73	2.69	2.42
25	2.99	2.89	2.81	2.75	2.70	2.60	2.54	2.49	2.45	2.17
30	2.84	2.74	2.66	2.60	2.55	2.45	2.39	2.34	2.30	2.01
35	2.74	2.64	2.56	2.50	2.44	2.35	2.28	2.23	2.19	1.89
40	2.66	2.56	2.48	2.42	2.37	2.27	2.20	2.15	2.11	1.80
500	2.22	2.12	2.04	1.97	1.92	1.81	1.74	1.68	1.63	2.30

A.9 QUANTILES OF THE K-S DISTRIBUTION

Table A.9 lists quantiles $K_{[p;n]}$ of the K-S distribution. For example, the $K_{[0.99;12]}$ required for a K-S test at 99% confidence level is 1.4521.

TABLE A.9 Quantiles of the K-S Distribution

	p										
n	0.00	0.01	0.05	0.10	0.20	0.50	0.80	0.90	0.95	0.99	0.995
1	0.0050	0.0100	0.0500	0.1000	0.2000	0.5000	0.8000	0.9000	0.9500	0.9900	0.9950
2	0.0067	0.0135	0.0673	0.1296	0.2416	0.5176	0.7818	0.9670	1.0980	1.2728	1.3142
3	0.0081	0.0162	0.0792	0.1471	0.2615	0.5147	0.8187	0.9783	1.1017	1.3589	1.4359
4	0.0093	0.0186	0.0879	0.1590	0.2726	0.5110	0.8248	0.9853	1.1304	1.3777	1.4685
5	0.0103	0.0207	0.0947	0.1675	0.2793	0.5245	0.8277	0.9995	1.1392	1.4024	1.4949
6	0.0113	0.0226	0.1002	0.1739	0.2834	0.5319	0.8343	1.0052	1.1463	1.4144	1.5104
7	0.0121	0.0243	0.1048	0.1787	0.2859	0.5364	0.8398	1.0093	1.1537	1.4246	1.5235
8	0.0130	0.0259	0.1086	0.1826	0.2874	0.5392	0.8431	1.0135	1.1586	1.4327	1.5324
9	0.0137	0.0275	0.1119	0.1856	0.2881	0.5411	0.8455	1.0173	1.1624	1.4388	1.5400
10	0.0145	0.0289	0.1147	0.1880	0.2884	0.5426	0.8477	1.0202	1.1658	1.4440	1.5461
11	0.0151	0.0303	0.1172	0.1900	0.2883	0.5439	0.8498	1.0225	1.1688	1.4484	1.5512
12	0.0158	0.0314	0.1193	0.1916	0.2879	0.5453	0.8519	1.0246	1.1714	1.4521	1.5555
13	0.0164	0.0324	0.1212	0.1929	0.2903	0.5468	0.8537	1.0265	1.1736	1.4553	1.5593
14	0.0171	0.0333	0.1229	0.1940	0.2925	0.5486	0.8551	1.0282	1.1755	1.4581	1.5626
15	0.0176	0.0342	0.1244	0.1948	0.2944	0.5500	0.8564	1.0298	1.1773	1.4606	1.5655
16	0.0182	0.0351	0.1257	0.1955	0.2961	0.5512	0.8576	1.0311	1.1789	1.4629	1.5681
17	0.0188	0.0359	0.1269	0.1961	0.2975	0.5523	0.8587	1.0324	1.1803	1.4649	1.5704
18	0.0193	0.0367	0.1280	0.1965	0.2987	0.5532	0.8597	1.0335	1.1816	1.4667	1.5725
19	0.0198	0.0374	0.1290	0.1968	0.2998	0.5540	0.8607	1.0346	1.1828	1.4683	1.5744
20	0.0203	0.0381	0.1298	0.1971	0.3007	0.5547	0.8616	1.0355	1.1839	1.4699	1.5761
21	0.0208	0.0387	0.1306	0.1973	0.3015	0.5554	0.8624	1.0365	1.1850	1.4712	1.5777
22	0.0213	0.0394	0.1313	0.1974	0.3023	0.5561	0.8631	1.0373	1.1859	1.4725	1.5792
23	0.0217	0.0400	0.1320	0.1974	0.3030	0.5567	0.8639	1.0381	1.1868	1.4737	1.5806
24	0.0221	0.0405	0.1326	0.1974	0.3035	0.5573	0.8645	1.0388	1.1876	1.4748	1.5816
25	0.0225	0.0411	0.1331	0.1974	0.3041	0.5579	0.8651	1.0395	1.1884	1.4758	1.5829
26	0.0228	0.0416	0.1336	0.1977	0.3046	0.5585	0.8657	1.0402	1.1891	1.4768	1.5840
27	0.0231	0.0421	0.1340	0.1985	0.3050	0.5590	0.8663	1.0408	1.1898	1.4777	1.5850
28	0.0235	0.0426	0.1344	0.1992	0.3054	0.5595	0.8668	1.0414	1.1905	1.4786	1.5860
29	0.0238	0.0431	0.1348	0.2000	0.3058	0.5600	0.8673	1.0419	1.1911	1.4794	1.5869
30	0.0241	0.0435	0.1351	0.2006	0.3062	0.5605	0.8678	1.0424	1.1916	1.4801	1.5878

Box 12.1 Summarizing Observations

Given: A sample $\{x_1, x_2, \ldots, x_n\}$ of n observations.

1. Sample arithmetic mean: $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$
2. Sample geometric mean: $\dot{x} = \left(\prod_{i=1}^{n} x_i\right)^{1/n}$
3. Sample harmonic mean: $\ddot{x} = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \cdots + \frac{1}{x_n}}$
4. Sample median: $\begin{cases} x_{((n-1)/2)} & \text{if } n \text{ is odd} \\ 0.5(x_{(n/2)} + x_{((1+n)/2)}) & \text{otherwise} \end{cases}$

   Here $x_{(i)}$ is the ith observation in the sorted set.
5. Sample mode = observation with the highest frequency (for categorical data).
6. Sample variance: $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$
7. Sample standard deviation: $s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$
8. Coefficient of variation = $s/\bar{x}$
9. Coefficient of skewness = $\frac{1}{ns^3}\sum_{i=1}^{n}(x_i - \bar{x})^3$
10. Range: Specify the minimum and maximum.
11. Percentiles: $100p$-percentile $x_p = x_{(\lfloor 1+(n-1)p \rfloor)}$.
12. Semi-interquartile range SIQR = $\frac{Q_3 - Q_1}{2} = \frac{x_{0.75} - x_{0.25}}{2}$
13. Mean absolute deviation = $\frac{1}{n}\sum_{i=1}^{n}|x_i - \bar{x}|$

Box 13.1 Confidence Intervals

1. Given: A sample $\{x_1, x_2, \ldots, x_n\}$ of n observations.
   $\bar{x}$ = sample mean; s = sample standard deviation

   (a) Standard error of the sample mean: $\sigma_{\bar{x}} = \frac{s}{\sqrt{n}}$

   (b) $100(1-\alpha)\%$ two-sided confidence interval for the mean:
   $\bar{x} \mp z_{1-\alpha/2} s/\sqrt{n}$
   If $n \le 30$†: $\bar{x} \mp t_{[1-\alpha/2;n-1]} s/\sqrt{n}$

   (c) $100(1-\alpha)\%$ one-sided confidence interval for the mean:
   $(\bar{x}, \bar{x} + z_{1-\alpha} s/\sqrt{n})$ or $(\bar{x} - z_{1-\alpha} s/\sqrt{n}, \bar{x})$
   If $n \le 30$†: $(\bar{x}, \bar{x} + t_{[1-\alpha;n-1]} s/\sqrt{n})$ or $(\bar{x} - t_{[1-\alpha;n-1]} s/\sqrt{n}, \bar{x})$

2. To compare two systems using unpaired observations:

   (a) The standard error of the mean difference: $s = \sqrt{\frac{s_a^2}{n_a} + \frac{s_b^2}{n_b}}$

   (b) The effective number of degrees of freedom:

   $$\nu = \frac{(s_a^2/n_a + s_b^2/n_b)^2}{\frac{1}{n_a+1}\left(\frac{s_a^2}{n_a}\right)^2 + \frac{1}{n_b+1}\left(\frac{s_b^2}{n_b}\right)^2} - 2$$

   (c) The confidence interval for the mean difference:
   $(\bar{x}_a - \bar{x}_b) \mp t_{[1-\alpha/2;\nu]} s$

3. If n_1 of the n observations belong to a certain class, the following statistics can be reported for the class:

   (a) Proportion of the observations in the class: $p = \frac{n_1}{n}$

   (b) $100(1-\alpha)\%$ two-sided confidence interval for the proportion‡:
   $p \mp z_{1-\alpha/2}\sqrt{\frac{p(1-p)}{n}}$

   (c) $100(1-\alpha)\%$ one-sided confidence interval for the proportion‡:

   $$\left(p, p + z_{1-\alpha}\sqrt{\frac{p(1-p)}{n}}\right) \quad \text{or} \quad \left(p - z_{1-\alpha}\sqrt{\frac{p(1-p)}{n}}, p\right)$$

† Only for samples from normal populations.
‡ Provided $np \ge 10$.

Box 14.1 Simple Linear Regression

1. Model: $y_i = b_0 + b_1 x_i + e_i$
2. Parameter estimation: $b_1 = \dfrac{\Sigma xy - n\bar{x}\,\bar{y}}{\Sigma x^2 - n(\bar{x})^2}$

   $b_0 = \bar{y} - b_1\bar{x}$
3. Allocation of variation: $\text{SSY} = \sum_{i=1}^{n} y_i^2$

   $\text{SS0} = n\bar{y}^2$

   $\text{SST} = \text{SSY} - \text{SS0}$

   $\text{SSE} = \Sigma y^2 - b_0\Sigma y - b_1\Sigma xy$

   $\text{SSR} = \text{SST} - \text{SSE}$
4. Coefficient of determination $R^2 = \dfrac{\text{SSR}}{\text{SST}} = \dfrac{\text{SST} - \text{SSE}}{\text{SST}}$
5. Standard deviation of errors $s_e = \sqrt{\dfrac{\text{SSE}}{n-2}}$
6. Degrees of freedoms: $\text{SST} = \text{SSY} - \text{SS0} = \text{SSR} + \text{SSE}$

   $n-1 = n - 1 = 1 + (n-2)$
7. Standard deviation of parameters: $s_{b_0} = s_e\left[\dfrac{1}{n} + \dfrac{\bar{x}^2}{\Sigma x^2 - n\bar{x}^2}\right]^{1/2}$

   $s_{b_1} = \dfrac{s_e}{\left[\Sigma x^2 - n\bar{x}^2\right]^{1/2}}$
8. Prediction: Mean of future m observations:

   $\hat{y}_p = b_0 + b_1 x_p$

   $s_{\hat{y}_p} = s_e\left[\dfrac{1}{m} + \dfrac{1}{n} + \dfrac{(x_p - \bar{x})^2}{\Sigma x^2 - n\bar{x}^2}\right]^{1/2}$
9. All confidence intervals are computed using $t_{[1-\alpha/2;n-2]}$.

Box 15.1 Multiple Linear Regression

1. Model: $y_i = b_0 + b_1 x_{1i} + b_2 x_{2i} + \cdots + b_k x_{ki} + e_i$

   or in matrix notation: $\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e}$

   where $\mathbf{b}$ = A column vector with $k+1$ elements = $\{b_0, b_1, \ldots, b_k\}$

   $\mathbf{y}$ = A column vector of n observed values of $y = \{y_1, \ldots, y_n\}$

   $\mathbf{X}$ = An n row by $k+1$ column matrix whose $(i, j+1)$th element $X_{i,j+1} = 1$ if $j = 0$ else x_{ij}
2. Parameter estimation: $\mathbf{b} = (\mathbf{X}^T\mathbf{X})^{-1}(\mathbf{X}^T\mathbf{y})$
3. Allocation of variation: $\text{SSY} = \sum_{i=1}^{n} y_i^2$ $\quad \text{SS0} = n\bar{y}^2$

   $\text{SST} = \text{SSY} - \text{SS0}$ $\quad \text{SSE} = \{\mathbf{y}^T\mathbf{y} - \mathbf{b}^T\mathbf{X}^T\mathbf{y}\}$

   $\text{SSR} = \text{SST} - \text{SSE}$
4. Coefficient of determination: $R^2 = \dfrac{\text{SSR}}{\text{SST}} = \dfrac{\text{SST} - \text{SSE}}{\text{SST}}$
5. Coefficient of multiple correlation $R = \sqrt{\dfrac{\text{SSR}}{\text{SST}}}$
6. Degrees of freedom: $\text{SST} = \text{SSY} - \text{SS0} = \text{SSR} + \text{SSE}$

   $n-1 = n - 1 = k + (n-k-1)$
7. Analysis of variance: $\text{MSR} = \dfrac{\text{SSR}}{k}$; $\text{MSE} = \dfrac{\text{SSE}}{n-k-1}$

   Regression is significant if MSR/MSE is greater than $F_{[1-\alpha;k,n-k-1]}$.
8. Standard deviation of errors: $s_e = \sqrt{\text{MSE}}$
9. Standard deviation of parameters: $s_{b_j} = s_e\sqrt{C_{jj}}$

   where C_{jj} is the jth diagonal term of $\mathbf{C} = (\mathbf{X}^T\mathbf{X})^{-1}$
10. Prediction: Mean of m future observations

    $\hat{y}_p = b_0 + b_1 x_{1p} + b_2 x_{2p} + \cdots + b_k x_{kp}$

    Or in vector notation: $\hat{y}_p = \mathbf{x}_p^T\mathbf{b}$; here, $\mathbf{x}_p^T = (1, x_{1p}, x_{2p}, \ldots, x_{kp})$
11. Standard deviation of predictions: $s_{\hat{y}_p} = s_e\sqrt{\left\{\dfrac{1}{m} + \mathbf{x}_p^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{x}_p\right\}}$
12. All confidence intervals are computed using $t_{[1-\alpha/2;n-k-1]}$.
13. Correlations among predictors:

    $$R_{x_1x_2} = \frac{\sum x_{1i}x_{2i} - n\bar{x}_1\bar{x}_2}{\left[\sum x_{1i}^2 - n\bar{x}_1^2\right]^{1/2}\left[\sum x_{2i}^2 - n\bar{x}_2^2\right]^{1/2}}$$