

## Practice Midterm 2

**Instructions** (please read the following carefully before starting the exam):

- **Do not open** the exam unless you are explicitly instructed to do so.
- You may use **one** formula sheet (with writing on front and back), a standard scientific calculator (with **no** internet access), and pencil(s) - all fully functioning (no sharing or borrowing is allowed). **No** cell phones or any other electronic devices are allowed.
- There is **NO partial credit**. **Only ONE answer is correct** for each of the questions. If you mark two or more options in the Scantron for any question, you will get 0 points for that question *irrespective* of whether one of those marked options is correct or not. **Advice:** Do **not** be hasty! Read **all** options *carefully* before choosing your final answer!
- **Mark your answers clearly** (and fully) with a number 2 pencil on your Scantron.
- **Mark which form** you have, **A or B**, on your Scantron.
- Also, **clearly bubble in the following on your Scantron:** *department* (**STAT**), the *course number* (**212**), the *section number* (**501**), your **name and UIN** on the Scantron.
- You may write on your exam (and use the blank side of each page for scratch work). But none of this work will count. Your score will be based solely on the Scantron.
- No cheating or discussing or helping others in any way will be tolerated. If detected, everyone involved will get a grade of 0 on the exam. You **must** work alone.
- **Turn in both your exam and Scantron** when you are done. Bring your **TAMU ID**.
- Good luck!

1. Data for 146 LPGA golfers from the year 2009 were collected. The following variables from this data set were considered for a regression analysis:

$y$  = scoring average    $x_1$  = average driving distance    $x_2$  = % fairways hit

$x_3$  = % greens hit    $x_4$  = average putts per round    $x_5$  = % sand saves

$x_6$  = no. of tournaments    $x_7$  = putts per greens hit    $x_8$  = no. of tournaments completed.

The accompanying output (available at the back of this exam as a 3-page document) shows information and plots obtained from regression models fit to these data in an “R” session.

**Use the information and the attachment above to answer the following 7 questions (i.e. the current one and the next 6 questions).**

The value of  $R^2$  for the model containing independent variables  $x_3$ ,  $x_4$  and  $x_6$  is closest to:

- (a) 0.655.
- (b) 0.941.
- (c) 0.902.
- ☒ (d) 0.945.
- (e) 0.952.

Using the leaps output, the model containing  $x_3$ ,  $x_4$  and  $x_6$  is the sixth one given, and the sixth  $R^2$  value is 0.945.

2. Considering AIC, BIC,  $R^2$  and the principle of parsimony, the best choice of the model is:

- (a) Probably the one containing only  $x_3$ .
- (b) Probably the one containing all 8 independent variables.
- ☒ (c) Probably the one containing  $x_1$ ,  $x_3$ ,  $x_4$ ,  $x_5$ ,  $x_6$  and  $x_8$ .
- (d) Any of the models that contain  $x_8$ .
- (e) A secret that Dr. Guha won't reveal.

The six-variable model containing  $x_1$ ,  $x_3$ ,  $x_4$ ,  $x_5$ ,  $x_6$  and  $x_8$  has the smallest BIC value and to two decimal places it has the same  $R^2$  as the full model, which has the smallest AIC value.

3. A sports reporter decides to use the simple model containing only  $x_3$ ,  $x_4$  and  $x_8$  to predict next year's scoring average for a promising rookie, Jane Doe. Assume that Jane's percentage of greens hit, average putts per round and number of tournaments completed are 70, 29.5 and 18, respectively. The prediction of her scoring average would be:

- (a) 70.1.
- (b) 69.1.
- ☒ (c) 71.2.
- (d) 72.1.
- (e) 73.5.

The prediction using the three-variable model with  $x_3$ ,  $x_4$  and  $x_8$  is:

$$\hat{y} = 61.240140 - 0.195544(70) + 0.820476(29.5) - 0.032671(18) = 71.2.$$

4. It is of interest to test the hypothesis that the variables  $x_2$ ,  $x_6$  and  $x_7$  are not needed in the same model with the other five independent variables. If we test this null hypothesis using  $\alpha = 0.05$ , then which of the following is correct?

- ☒ (a) The  $F$ -statistic is 4.245 and we would conclude that  $x_2$ ,  $x_6$  and  $x_7$  are needed in the model if  $4.245 > F_{3,137;0.05}$ .
- (b) The  $F$ -statistic is 4.245 and we would conclude that  $x_2$ ,  $x_6$  and  $x_7$  are *not* needed in the model if  $4.245 > F_{3,137;0.05}$ .
- (c) The  $F$ -statistic is 308.64 and we would conclude that  $x_2$ ,  $x_6$  and  $x_7$  are needed in the model if  $308.64 > F_{3,137;0.05}$ .
- (d) The  $F$ -statistic is 308.64 and we would conclude that  $x_2$ ,  $x_6$  and  $x_7$  are *not* needed in the model if  $308.64 > F_{3,137;0.05}$ .
- (e) The  $F$ -statistic is 308.64 and we would conclude that  $x_2$ ,  $x_6$  and  $x_7$  are needed in the model if  $308.64 > F_{5,137;0.05}$ .

You should use the reduction method here. The full model has an SSE of 7.573 and the reduced model, the one with  $x_1$ ,  $x_3$ ,  $x_4$ ,  $x_5$  and  $x_8$ , has an SSE of 8.277. Therefore the  $F$ -statistic is:

$$F = \frac{(8.277 - 7.573)/3}{7.573/137} = 4.245.$$

We would conclude that the variables  $x_2$ ,  $x_6$  and  $x_7$  are needed in the model if 4.245 is larger than  $F_{3,137;0.05}$ .

5. A plot of the residuals for the full model (i.e., the one containing all 8 independent variables) is provided for you in the last page of the R output attachment. Which of the following is the best conclusion based on this plot?

- (a) An assumption of Normally distributed error terms seems reasonable.

- (b) There is a clear decrease in the variance of residuals as the predicted value increases.
- (c) There appears to be no outliers that have a large influence on the fitted model.
- ☒ (d) These residuals do not give us any reason to believe that the model assumptions have been violated.
- (e) Both options (b) and (c) are true.

The residuals appear to be randomly scattered, as they should be when the model assumptions are met.

6. Given that  $\sum_{i=1}^{146} (y_i - \bar{y})^2 = 189.4666$ , the estimate of the error variance using the full model is:

- (a)  $189.4666/145$ .
- (b)  $189.4666/137$ .
- ☒ (c)  $189.4666(1 - 0.960028)/137$ .
- (d)  $189.4666(0.960028)/137$ .
- (e)  $0.960028$ .

We know that:

$$\begin{aligned}
 \hat{\sigma}^2 &= SSE/(n - k - 1) \\
 &= (SST - SSR)/(n - k - 1) \\
 &= (SST - SST \cdot R^2)/(n - k - 1) \\
 &= SST(1 - R^2)/(n - k - 1) \\
 &= 189.4666(1 - 0.960028)/137.
 \end{aligned}$$

7. A plot of Cook's D for the full model (i.e., the one containing all 8 independent variables) is provided for you in the last page of the R output attachment. Which of the following is the best conclusion based on this plot?

- (a) An assumption of Normally distributed error terms seems reasonable.
- (b) There are no extremely large residuals.
- ☒ (c) There appear to be no outliers that have a large influence on the fitted model.
- (d) An assumption of constant variance for the error terms seems reasonable.
- (e) Both options (b) and (c) are true.

The Cook's  $D$  values are not the same thing as residuals. They indicate whether leaving out a data value changes the estimated regression coefficients very much. Since no Cook's  $D$  value is larger than 1.5 (in fact, all are less than 0.15), there are no influential observations.

8. A group of entomologists were studying data involving spruce moths. The numbers of moths caught in 60 different traps were recorded. The traps varied according to where they were located on a tree: top, middle, lower or ground. Fifteen traps were used for each location, and an analysis of variance was conducted to determine what effect, if any, the locations had on the number of moths caught. The following partial ANOVA table was determined from the data. (The lower case italicized letters in the table denote entries that are not given to you.)

Source of variation	Degrees of freedom	Sum of squares	Mean square	$F$
Location	$a$	$b$	$c$	$d$
Error	56	3261.6	$e$	
Total	$f$	$g$		

Use the information and the table above to answer the following 4 questions (i.e. the current one and the next 3 questions).

The number of moths caught in different traps varies even when all the traps are in the same location. We term this standard deviation  $\sigma$ . The best estimate of  $\sigma$  from the ANOVA table is:

- (a)  $(g - 3261.6)/56$ .
- (b)  $g/f$ .
- (c)  $\sqrt{g/f}$ .
- (d)  $3261.6/56$ .
- ☒ (e)  $\sqrt{3261.6/56}$ .

We estimate  $\sigma^2$  using the MSE, in this case  $3261.6/56$ . So, the estimate of  $\sigma$  is  $\sqrt{3261.6/56}$ .

9. Let  $\mu_t$ ,  $\mu_m$ ,  $\mu_\ell$  and  $\mu_g$  denote the average number of moths caught in the top, middle, lower and ground parts of a tree, respectively. The value of the  $F$ -statistic for testing  $H_0 : \mu_t = \mu_m = \mu_\ell = \mu_g$  is:

- (a)  $(a/b)/58.24$ .
- (b)  $b/3261.6$ .
- ☒ (c)  $(b/3)/58.24$ .

(d)  $(b/a)/3261.6$ .

(e)  $(a/b)/(g/f)$ .

The value of the  $F$ -statistic is:  $MSTr/MSE = (b/a)/(3261.6/56) = (b/3)/58.24$ .

**10.** It turns out that the value of the  $F$ -statistic is 11.34. If we test the null hypothesis  $H_0 : \mu_t = \mu_m = \mu_\ell = \mu_g$  using  $\alpha = 0.05$ , then which of the following is correct? (Remember, if you can't find a certain degree of freedom on the  $F$ -table, choose the next smaller one on the table.)

(a) Since  $F$  is smaller than the appropriate table value, we cannot reject  $H_0$ .

(b) Since  $F$  is larger than the appropriate table value, we may conclude that  $\mu_t \neq \mu_m$ ,  $\mu_m \neq \mu_\ell$  and  $\mu_\ell \neq \mu_g$ .

(c) We cannot reject equality of the four means since  $F > F_{3,50;0.05} = 2.79$ .

☒ (d) It is reasonable to conclude that the four means are not all the same since  $F > F_{3,50;0.05} = 2.79$ .

(e) Since we do not know the  $P$ -value it is impossible to draw a conclusion.

The correct degrees of freedom are 3 and 56, but 56 is not in the table. So we use the next smaller degrees of freedom, 50. We would reject the hypothesis of equal means if  $F \geq F_{3,50;0.05} = 2.79$ , and since  $11.34 > 2.79$ , it is reasonable to conclude that the four means are not all the same.

**11.** A multiple linear regression model relating  $Y$  with  $x_1$  and  $x_2$  has the form:

$$Y = 10 + x_1 - 3x_2 + 0.9x_1x_2 + \epsilon,$$

where, for every choice of  $(x_1, x_2)$ ,  $\epsilon$  has a Normal distribution with mean 0 and standard deviation 1. The expected value of  $Y$  when  $x_1 = 1$  and  $x_2 = 2$  is:

☒ (a) 6.8.

(b) 7.1.

(c) 8.6.

(d) 10.

(e) Cannot be determined from the information given.

The expected value of  $Y$  given  $x_1 = 1$  and  $x_2 = 2$  is:

$$E(Y) = 10 + x_1 - 3x_2 + 0.9x_1x_2 = 10 + 1 - 6 + 0.9(2) = 6.8.$$

**12.** Suppose we want to test 5 null hypotheses:  $H_0^{(1)}, \dots, H_0^{(5)}$ , and for simplicity, assume that they are tested using 5 independent datasets. Suppose each hypothesis is now tested using a procedure that controls the respective Type I error rate at a level 0.1. Then, the experimentwise error rate for testing all these hypotheses simultaneously is given by: (**Note:** this question carries **4 points**.)

- (a) 0.1.
- (b)  $(0.1)^5$ .
- (c) 0.5905.
- (d) 0.9999.
- ☒ (e) 0.4095.

Using the formula and calculations on pg. 135-136 of Chapter 2A notes, the experimentwise error rate (EWER) is given by:  $\alpha_E = 1 - (1 - \alpha)^m$ . Here, there are  $m = 5$  hypotheses (tested independently using 5 different datasets) and for each test, the Type I error rate is controlled at a level  $\alpha = 0.1$ . Hence, the  $\text{EWER} = 1 - (1 - 0.1)^5 = 1 - 0.9^5 = 1 - 0.5905 = 0.4095$ .

```

1 X=cbind(x1,x2,x3,x4,x5,x6,x7,x8)
2
3 leaps(X,y,method='r2',nbest=2)$which
4
5      1      2      3      4      5      6      7      8      AIC      BIC
6 1 FALSE FALSE FALSE FALSE FALSE FALSE FALSE TRUE    269.69    278.64
7 1 FALSE FALSE  TRUE FALSE FALSE FALSE FALSE FALSE    302.86    311.81
8 2 FALSE FALSE  TRUE  TRUE FALSE FALSE FALSE FALSE     48.05     59.98
9 2 FALSE FALSE  TRUE FALSE FALSE FALSE  TRUE FALSE    121.99    133.92
10 3 FALSE FALSE  TRUE  TRUE FALSE FALSE FALSE  TRUE     20.04     34.96
11 3 FALSE FALSE  TRUE  TRUE FALSE  TRUE FALSE FALSE     39.80     54.72
12 4 FALSE FALSE  TRUE  TRUE  TRUE FALSE FALSE  TRUE     15.29     33.19
13 4  TRUE FALSE  TRUE  TRUE FALSE FALSE FALSE  TRUE     16.29     34.19
14 5  TRUE FALSE  TRUE  TRUE  TRUE FALSE FALSE  TRUE      9.29     30.18
15 5 FALSE FALSE  TRUE  TRUE  TRUE FALSE  TRUE  TRUE     10.83     31.72
16 6  TRUE FALSE  TRUE  TRUE  TRUE  TRUE FALSE  TRUE      5.83     29.70
17 6  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE FALSE  TRUE      7.05     30.92
18 7  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE FALSE  TRUE      4.06     30.91
19 7  TRUE FALSE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE      5.11     31.98
20 8  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE      2.32     32.16
21
22 $label
23 [1] "(Intercept)" "1"          "2"          "3"          "4"
24 [6] "5"            "6"            "7"            "8"
25
26 $size
27 [1] 2 2 3 3 4 4 5 5 6 6 7 7 8 8 9
28
29 $r2
30 [1] 0.7253822 0.6553326 0.9406433 0.9015029 0.9516711 0.9446654 0.9538540
31 [8] 0.9535363 0.9563137 0.9558510 0.9579181 0.9575638 0.9589913 0.9586936
32 [15] 0.9600280
33
34
35 > fit=lm(y~x1+x2+x3+x4+x5+x6+x7+x8)
36 > anova(fit)
37
38 Analysis of Variance Table
39
40 Response: y
41      Df Sum Sq Mean Sq F value    Pr(>F)
42 x1      1 41.861  41.861 757.2479 < 2.2e-16 ***
43 x2      1 37.150  37.150 672.0372 < 2.2e-16 ***
44 x3      1 46.334  46.334 838.1606 < 2.2e-16 ***
45 x4      1 53.633  53.633 970.2110 < 2.2e-16 ***
46 x5      1  0.912   0.912  16.5013 8.144e-05 ***
47 x6      1  0.444   0.444   8.0349 0.005283 **
48 x7      1  0.333   0.333   6.0215 0.015387 *
49 x8      1  1.226   1.226  22.1824 6.008e-06 ***
50 Residuals 137  7.573   0.055
51 ---
52
53
54
55 > fit1=lm(y~x2+x6+x7)
56 > anova(fit1)
57
58 Analysis of Variance Table
59
60 Response: y
61      Df Sum Sq Mean Sq F value    Pr(>F)
62 x2      1 10.159  10.159  24.553 2.028e-06 ***
63 x6      1 63.442  63.442 153.329 < 2.2e-16 ***
64 x7      1 57.110  57.110 138.025 < 2.2e-16 ***
65 Residuals 142 58.755   0.414
66 ---

```

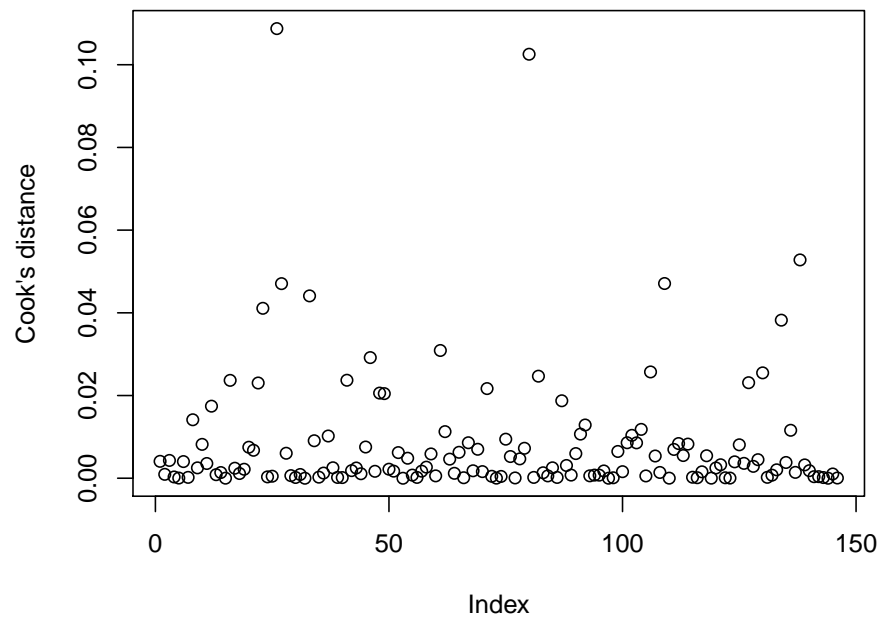


```

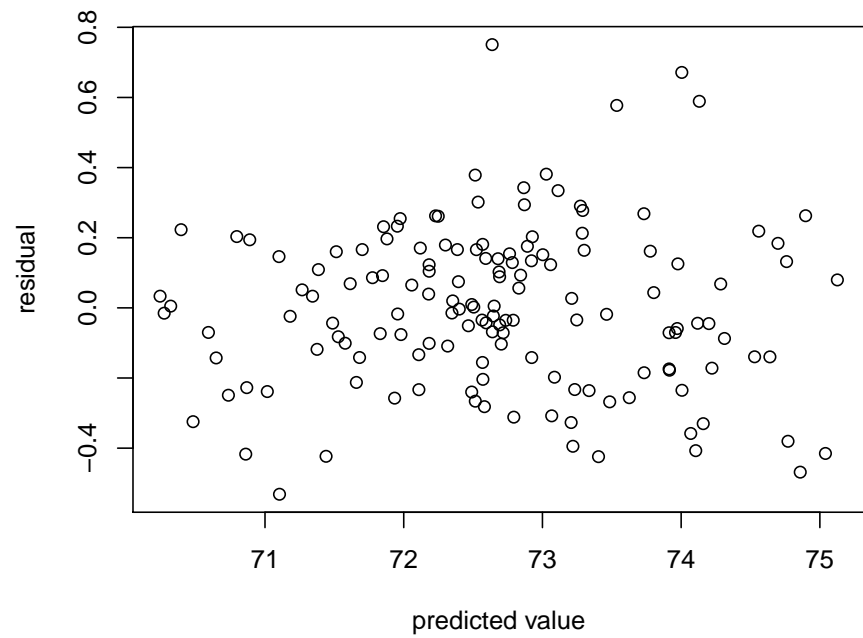
67
68
69
70 > fit2=lm(y~x1+x3+x4+x5+x8)
71 > anova(fit2)
72
73 Analysis of Variance Table
74
75 Response: y
76      Df Sum Sq Mean Sq  F value    Pr(>F)
77 x1      1 41.861  41.861   708.037 < 2.2e-16 ***
78 x3      1 82.765  82.765 1399.901 < 2.2e-16 ***
79 x4      1 54.243  54.243  917.479 < 2.2e-16 ***
80 x5      1  0.830   0.830   14.041 0.0002607 ***
81 x8      1  1.490   1.490   25.205 1.542e-06 ***
82 Residuals 140  8.277   0.059
83 ---
84
85
86
87 > fit3=lm(y~x3+x4+x8)
88 > summary(fit3)
89
90 Call:
91 lm(formula = y ~ x3 + x4 + x8)
92
93 Residuals:
94      Min       1Q   Median       3Q      Max
95 -0.66244 -0.17566  0.02268  0.16736  0.84612
96
97 Coefficients:
98             Estimate Std. Error t value Pr(>|t|)
99 (Intercept) 61.240140   1.071359  57.161 < 2e-16 ***
100 x3          -0.195544   0.007981 -24.500 < 2e-16 ***
101 x4           0.820476   0.041074  19.976 < 2e-16 ***
102 x8          -0.032671   0.005740  -5.692 6.93e-08 ***
103 ---
104
105 Residual standard error: 0.2539 on 142 degrees of freedom
106 Multiple R-squared:  0.9517,    Adjusted R-squared:  0.9507
107 F-statistic: 932.1 on 3 and 142 DF,  p-value: < 2.2e-16
108
109 > anova(fit3)
110
111 Analysis of Variance Table
112
113 Response: y
114      Df Sum Sq Mean Sq  F value    Pr(>F)
115 x3      1 124.164 124.164 1925.500 < 2.2e-16 ***
116 x4      1  54.057  54.057  838.301 < 2.2e-16 ***
117 x8      1   2.089   2.089   32.402 6.926e-08 ***
118 Residuals 142   9.157   0.064
119
120

```

**Cook's distance for full model**



**Residuals from full model**



**Table A.3** The Cumulative Distribution Function for the Standard Normal Distribution: Values of  $\Phi(z)$  for Nonnegative  $z$ 

$z$	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

Table A.4 Percentiles of the <i>T</i> Distribution						
<i>df</i>	90%	95%	97.5%	99%	99.5%	99.9%
1	3.078	6.314	12.706	31.821	63.657	318.309
2	1.886	2.920	4.303	6.965	9.925	22.327
3	1.638	2.353	3.183	4.541	5.841	10.215
4	1.533	2.132	2.777	3.747	4.604	7.173
5	1.476	2.015	2.571	3.365	4.032	5.893
6	1.440	1.943	2.447	3.143	3.708	5.208
7	1.415	1.895	2.365	2.998	3.500	4.785
8	1.397	1.860	2.306	2.897	3.355	4.501
9	1.383	1.833	2.262	2.822	3.250	4.297
10	1.372	1.812	2.228	2.764	3.169	4.144
11	1.363	1.796	2.201	2.718	3.106	4.025
12	1.356	1.782	2.179	2.681	3.055	3.930
13	1.350	1.771	2.160	2.650	3.012	3.852
14	1.345	1.761	2.145	2.625	2.977	3.787
15	1.341	1.753	2.132	2.603	2.947	3.733
16	1.337	1.746	2.120	2.584	2.921	3.686
17	1.333	1.740	2.110	2.567	2.898	3.646
18	1.330	1.734	2.101	2.552	2.879	3.611
19	1.328	1.729	2.093	2.540	2.861	3.580
20	1.325	1.725	2.086	2.528	2.845	3.552
21	1.323	1.721	2.080	2.518	2.831	3.527
22	1.321	1.717	2.074	2.508	2.819	3.505
23	1.319	1.714	2.069	2.500	2.807	3.485
24	1.318	1.711	2.064	2.492	2.797	3.467
25	1.316	1.708	2.060	2.485	2.788	3.450
26	1.315	1.706	2.056	2.479	2.779	3.435
27	1.314	1.703	2.052	2.473	2.771	3.421
28	1.313	1.701	2.048	2.467	2.763	3.408
29	1.311	1.699	2.045	2.462	2.756	3.396
30	1.310	1.697	2.042	2.457	2.750	3.385
40	1.303	1.684	2.021	2.423	2.705	3.307
80	1.292	1.664	1.990	2.374	2.639	3.195
∞	1.282	1.645	1.960	2.326	2.576	3.090

**Table A.6** Percentiles of the  $F$  Distribution ( $\nu_1$  = Numerator df;  $\nu_2$  = Denominator df)

		$\nu_1$										
$\nu_2$	$\alpha$	1	2	3	4	5	6	7	8	12	24	1,000
1	0.10	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	60.71	62.00	63.30
	0.05	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	243.9	249.1	254.2
2	0.10	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.41	9.45	9.49
	0.05	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.41	19.45	19.49
3	0.10	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.22	5.18	5.13
	0.05	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.74	8.64	8.53
4	0.10	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.90	3.83	3.76
	0.05	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	5.91	5.77	5.63
5	0.10	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.27	3.19	3.11
	0.05	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.68	4.53	4.37
6	0.10	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.90	2.82	2.72
	0.05	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.00	3.84	3.67
7	0.10	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.67	2.58	2.47
	0.05	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.57	3.41	3.23
8	0.10	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.50	2.40	2.30
	0.05	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.28	3.12	2.93
10	0.10	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.28	2.18	2.06
	0.05	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	2.91	2.74	2.54
12	0.10	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.15	2.04	1.91
	0.05	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.69	2.51	2.30
14	0.10	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.05	1.94	1.80
	0.05	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.53	2.35	2.14
16	0.10	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	1.99	1.87	1.72
	0.05	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.42	2.24	2.02
20	0.10	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.89	1.77	1.61
	0.05	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.28	2.08	1.85
30	0.10	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.77	1.64	1.46
	0.05	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.09	1.89	1.63
50	0.10	2.81	2.41	2.20	2.06	1.97	1.90	1.84	1.80	1.68	1.54	1.33
	0.05	4.03	3.18	2.79	2.56	2.40	2.29	2.20	2.13	1.95	1.74	1.45
100	0.10	2.76	2.36	2.14	2.00	1.91	1.83	1.78	1.73	1.61	1.46	1.22
	0.05	3.94	3.09	2.70	2.46	2.31	2.19	2.10	2.03	1.85	1.63	1.30
1,000	0.10	2.71	2.31	2.09	1.95	1.85	1.78	1.72	1.68	1.55	1.39	1.08
	0.05	3.85	3.00	2.61	2.38	2.22	2.11	2.02	1.95	1.76	1.53	1.11

**Table A.7** Percentiles of the Studentized Range Distribution ( $Q_{\alpha, k, \nu}$  for  $\nu = 0.10$  and  $\alpha = 0.05$ )

		$k$									
$\nu$	$\alpha$	2	3	4	5	6	7	8	9	10	11
5	0.10	2.85	3.72	4.26	4.66	4.98	5.24	5.46	5.65	5.82	5.96
	0.05	3.63	4.60	5.22	5.67	6.03	6.33	6.58	6.80	6.99	7.17
6	0.10	2.75	3.56	4.06	4.43	4.73	4.97	5.17	5.34	5.50	5.64
	0.05	3.46	4.34	4.90	5.30	5.63	5.89	6.12	6.32	6.49	6.65
7	0.10	2.68	3.45	3.93	4.28	4.55	4.78	4.97	5.14	5.28	5.41
	0.05	3.34	4.16	4.68	5.06	5.36	5.61	5.81	6.00	6.16	6.30
8	0.10	2.63	3.37	3.83	4.17	4.43	4.65	4.83	4.99	5.13	5.25
	0.05	3.26	4.04	4.53	4.89	5.17	5.40	5.60	5.77	5.92	6.05
10	0.10	2.56	3.27	3.70	4.02	4.26	4.46	4.64	4.78	4.91	5.03
	0.05	3.15	3.88	4.33	4.65	4.91	5.12	5.30	5.46	5.60	5.72
12	0.10	2.52	3.20	3.62	3.92	4.16	4.35	4.51	4.65	4.78	4.89
	0.05	3.08	3.77	4.20	4.51	4.75	4.95	5.12	5.26	5.39	5.51
13	0.10	2.50	3.18	3.59	3.88	4.12	4.30	4.46	4.60	4.72	4.83
	0.05	3.05	3.73	4.15	4.45	4.69	4.88	5.05	5.19	5.32	5.43
14	0.10	2.49	3.16	3.56	3.85	4.08	4.27	4.42	4.56	4.68	4.79
	0.05	3.03	3.70	4.11	4.41	4.64	4.83	4.99	5.13	5.25	5.36
16	0.10	2.47	3.12	3.52	3.80	4.03	4.21	4.36	4.49	4.61	4.71
	0.05	3.00	3.65	4.05	4.33	4.56	4.74	4.90	5.03	5.15	5.26
18	0.10	2.45	3.10	3.49	3.77	3.98	4.16	4.31	4.44	4.55	4.65
	0.05	2.97	3.61	4.00	4.28	4.49	4.67	4.82	4.95	5.07	5.17
20	0.10	2.44	3.08	3.46	3.74	3.95	4.12	4.27	4.40	4.51	4.61
	0.05	2.95	3.58	3.96	4.23	4.44	4.62	4.77	4.89	5.01	5.11
25	0.10	2.42	3.04	3.42	3.68	3.89	4.06	4.20	4.32	4.43	4.53
	0.05	2.91	3.52	3.89	4.15	4.36	4.53	4.67	4.79	4.90	4.99
30	0.10	2.40	3.02	3.39	3.65	3.85	4.02	4.15	4.27	4.38	4.47
	0.05	2.89	3.49	3.84	4.10	4.30	4.46	4.60	4.72	4.82	4.92
40	0.10	2.38	2.99	3.35	3.60	3.80	3.96	4.10	4.21	4.32	4.41
	0.05	2.86	3.44	3.79	4.04	4.23	4.39	4.52	4.63	4.73	4.82
60	0.10	2.36	2.96	3.31	3.56	3.75	3.91	4.04	4.15	4.25	4.34
	0.05	2.83	3.40	3.74	3.98	4.16	4.31	4.44	4.55	4.65	4.73
80	0.10	2.35	2.94	3.29	3.54	3.73	3.88	4.01	4.12	4.22	4.31
	0.05	2.81	3.38	3.71	3.95	4.13	4.28	4.40	4.51	4.60	4.69
$\infty$	0.10	2.33	2.90	3.24	3.48	3.66	3.81	3.93	4.04	4.13	4.21
	0.05	2.77	3.31	3.63	3.86	4.03	4.17	4.29	4.39	4.47	4.55