STAT 212-501 (Fall 2022) Name ______________________

Instructor: Dr. Sharmistha Guha

Solutions to the Practice Final Exam

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

- **Do not open** the exam unless instructed to do so. The exam **duration is 2 hours.**
- There are **18 questions worth 5 points each** for a **total of 90 points.** In addition, **everyone gets 10 bonus points!** Your final score will be out of 100.
- There is **NO partial credit. Only ONE answer is correct** for each of the questions. Advice: Do **not** be hasty! Read **all** options *carefully* before choosing your final answer!
- **Attachments:** Attached with the exam are **distribution tables** (5 pages, one each for Normal, t, χ^2, F and Q tables) and some **R output** (2 pages) that you need for answering some of the questions.
- You may use **four** paper formula sheets (with writing on front and back), **four** paper blank sheets for scratch work, a standard scientific calculator (with **no** internet access).
- **Mark your answers clearly** (and fully) with a number 2 pencil 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.

  Good luck!

1. The method we discussed for identifying influential observations in regression is:

(a) The Sign test.

(b) The Kruskal-Wallis test.

(c) The Shapiro-Wilk test.

(d) The Cook's D statistic.

(e) The flop shot.

See pg. 58-61 of Chapter 1C notes.

2. A disadvantage of using the Sign Test to test a hypothesis about a population median is:

(a) It requires more assumptions than does the Wilcoxon signed-rank test.

(b) It is less powerful than a t-test when the population is in fact Normally distributed.

(c) It requires more assumptions than does the t-test.

(d) It is more computationally tedious than other methods.

(e) It can only be used when Dr.Guha says so.

See pg. 227 of Chapter 4 notes.

3. In a regression setting, when the variance of a positive response variable Y tends to increase as x increases, a good way to try to rectify this problem is to:

(a) Multiply each Y-value by the same large, positive number.

(b) Take the natural logarithm of each Y-value.

(c) Use $1/x$ as the independent variable.

(d) Divide each Y-value by the same large, positive number.

(e) Use both x and x^2 in the regression model.

See pg. 74-76 of Chapter 1B notes. We discussed the potential scenarios and the benefits of applying log transformations in details several times in class.

4. A sports reporter has studied data on professional golfers and determined that the following model reasonably approximates the relationship between average driving distance (x, in yards) and average eighteen hole score (Y):

$$Y = 182.9 - 0.7394x + 0.001203x^2 + \epsilon,$$

where ϵ is Normally distributed with mean 0 and standard deviation 0.25. Suppose a pro golfer's average driving distance is 285 yards. What is the probability that his average eighteen hole score is more than 70.25?

(a) 0.072.

(b) 0.360.

(c) 0.640.

(d) 0.928.

(e) 0.500.

$Y|x$ has a Normal distribution with mean $182.9 - 0.7394x + 0.001203x^2$ and standard deviation 0.25 under the assumed model. Hence, given $x = 285$, and letting $Z \sim N(0, 1)$, we have:

$$\begin{aligned} P(Y > 70.25) &= P\left(Z > \frac{70.25 - 182.9 + 0.7394(285) - 0.001203(285)^2}{0.25}\right) \\ &= P(Z > 1.463) \\ &= 0.072. \end{aligned}$$

5. 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), ϵ 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) 7.1.

(b) 6.8.

(c) 8.6.

(d) 10.

(e) Cannot be determined from the information given.

6. Four hypothesis tests are conducted independently of each other. If the level of significance of each test is 0.05, then the experimentwise error rate is:

(a) 0.10.

(b) 0.1921.

(c) 0.2916.

(d) 0.1855.

(e) 0.6561.

The experimentwise error rate (EWER) is the probability of rejecting one or more null hypotheses when in fact all the null hypotheses are true. Let R_i denote the event that the ith hypothesis is rejected, $i = 1, 2, 3, 4$. Then, using pg. 83 of Chapter 2A notes,

$$\begin{aligned} EWER = P(R_1 \cup R_2 \cup R_3 \cup R_4) &= 1 - P(R_1^c \cap R_2^c \cap R_3^c \cap R_4^c) \\ &= 1 - P(R_1^c)P(R_2^c)P(R_3^c)P(R_4^c) \\ &= 1 - (0.95)^4 \\ &= 0.1855. \end{aligned}$$

7. An experiment was conducted to ascertain gas mileage obtained with three grades of gasoline: regular, extra and premium. Four cars were used in the experiment, with all three gasolines being used (at different times) in each of the four cars. The order in which a given car received the three gasolines was determined randomly. The response variable was miles per gallon of gasoline. The following (partially completed) ANOVA table was obtained from this experiment's data.

Source of variation	Degrees of freedom	Sum of squares	Mean square	F
Gasolines		47.17		
Cars	3	390.25		
Error				
Total	11	448.92		

Use this information to answer the following **2 questions** (the current one and the next).

The value of the F-statistic when testing whether mean miles per gallon differ significantly between gasolines is given by:

(a) F-statistic = 8.2.

(b) F-statistic = 9.2.

(c) F-statistic = 14.3.

(d) F-statistic = 10.1.

(e) F-statistic = 12.3.

This is a Randomized Block Design (RBD) setting with $k = 3$ treatments (gasolines) and $n = 4$ blocks (cars). Using pg. 166-167 of Chapter 2B notes and the partial ANOVA table given, the SSE is given by:

$$SSE = SST - SSTr - SSB = 448.92 - 390.25 - 47.17 = 11.5,$$

The degrees of freedom for the treatments and error are given by $k-1 = 2$ and $(k-1)(n-1) = 6$ respectively. Hence, the F-statistic for testing equality of the treatment means equals:

$$F \equiv F_{Tr} = \frac{MSTr}{MSE} = \frac{SSTr/(k-1)}{SSE/\{(k-1)(n-1)\}} = \frac{47.17/2}{11.5/6} = 12.3.$$

8. In the F-test in Question (7), i.e., $H_0 =$ mean miles per gallon are same between gasolines; consider the F statistic you obtained. Assume $\alpha = 0.05$. Which of the following is true?

(a) From the F-table, $F_{6,2;0.05} = 19.33$, so we reject the null.

(b) From the F-table, $F_{6,2;0.05} = 19.33$, so we fail to reject the null.

(c) From the F-table, $F_{2,6;0.05} = 5.14$, so we reject the null.

(d) From the F-table, $F_{2,6;0.05} = 5.14$, so we fail to reject the null.

(e) Not enough information is given to draw a conclusion.

Under H_0, F_{Tr} follows a $F_{k-1,(k-1)(n-1)} \equiv F_{2,6}$ distribution and we reject H_0 for large values of F_{Tr}. From the F-table, we have: $F_{2,6;0.05} = 5.14 < 12.3$, so we reject the null.

9. Regression data consisting of $n = 200$ cases were collected. The response variable is y and there were four independent variables: x_1, x_2, x_3 and x_4. The accompanying output (available as a 2-page pdf document attached with this exam) shows various information obtained from these data in an "R" session. Use the information and the attachment above to answer the following **5 questions** (i.e. the current one and the next 4 questions).

Let M_1 be the linear model containing only the variables x_1 and x_4, and M_2 the linear model containing all four independent variables. The value of the F-statistic for testing:

$$H_0 : \text{Correct model is } M_1 \quad \text{vs.} \quad H_a : \text{Correct model is } M_2$$

(a) Is 72.9.

(b) Is 97.4.

(c) Is 150.5.

(d) Is 294.8.

(e) Cannot be determined from the information given.

You should use the reduction method of testing here. From the "R" output, the relevant SSEs for the full and reduced models are $SSE_f = 83,792$ and $SSE_r = 146,443$, respectively. Hence, the required F-statistic is:

$$F = \frac{(SSE_r - SSE_f)/(4-2)}{SSE_f/(n-4-1)} = \frac{(146,443 - 83,792)/2}{83,792/195} = 72.9.$$

10. The proportion of variance in y explained by the model containing x_2 and x_3 is:

(a) 0.855.

(b) 0.940.

(c) 0.364.

(d) 0.636.

(e) 0.750.

This is simply the multiple R^2 given in the summary for the model with only x_2 and x_3.

11. Consider the following information obtained using "R":

Model number	Variables used	R^2
1	x_1	0.7452
2	x_2	0.0493
3	x_1, x_2	0.8212
4	x_1, x_2, x_3	0.8212
5	x_1, x_2, x_3, x_4	0.8545

Based on this information, the most parsimonious model that still provides a near optimum fit is given by model number:

(a) 1.

(b) 2.

(c) 3.

(d) 4.

(e) 5.

Model 3 has an R^2 that is barely smaller than those of models 4 and 5, while being substantially larger than those of 1 and 2. Therefore, model 3 fits the bill.

12. Consider the output for the model containing all four of the independent variables. Which of the following is the best conclusion to draw from the model utility test?

(a) All four of the independent variables are useful.

(b) The variable x_1 should definitely be included in the model.

(c) None of the independent variables is useful.

(d) At least one of the 4 independent variables is useful.

(e) The variable x_2 should definitely be included in the model.

The P-value for the model utility test is extraordinarily small, $2 \cdot 10^{-16}$, and therefore there is strong evidence that at least one of the four independent variables is useful. However, as is always the case with a model utility test, this is all that can be said from the result of the test (and nothing can be said about the importance of any specific predictor in the model).

13. The estimate of the error variance in the model containing x_1 and x_2 is:

(a) 199.

(b) 84,908.

(c) 0.8548.

(d) 20.76.

(e) 431.

From the given `R` output, the estimate of σ^2 is $SSE/(n-k-1) = 84,908/(200-2-1) = 431.0$.

14. I graded a quiz for 12 students in my class. Their scores (out of 10 points) are as follows:

$$8, \ 6, \ 6, \ 8, \ 8, \ 10, \ 9, \ 7, \ 10, \ 9, \ 7, \ 8.$$

Use this information to answer the following **2 questions** (the current one and the next).

Using this data, suppose I now use the Sign test to test the following hypotheses for the median (M) of the population of my students' scores: $H_0 : M = 7.5$ vs. $H_a : M \neq 7.5$. Then, the Z-statistic is:

(e) Z-statistic = 1.1547.

(a) Z-statistic = -1.1547.

(b) Z-statistic = -1.1747.

(c) Z-statistic = 1.3547.

(d) Z-statistic = -1.8547.

The number of observations in this data that are greater than the null value 7.5 of the median is $Y = 8$. Using pg. 222-224 of Chapter 4 notes, the Z-statistic for the Sign test then equals:

$$Z = \frac{Y - n/2}{\sqrt{n}/2} = \frac{8 - 12/2}{\sqrt{12}/2} = \frac{2}{\sqrt{12}/2} = 1.1547.$$

15. The approximate P-value for this Sign test described in Question 14 above is closest to:

(a) P-value = 0.125.

(b) P-value = 0.75.

(c) P-value = 0.95.

(d) P-value = 0.165.

(e) P-value = 0.25.

Since the alternative H_a is two-sided, and $Z \sim N(0, 1)$ under H_0, the P-value for this test is $2P(Z > |Z_{obs}|) = 2P(Z > 1.1547)$, where $Z \sim N(0, 1)$ and $Z_{obs} = 1.1547$ is the observed value of the test statistic Z. Using the Normal table, this roughly equals $2(1-0.8749) = 0.25$.

16. A biomedical researcher wants to test the effectiveness of a synthetic antitoxin. Twelve randomly selected subjects are tested for resistance to a particular poison. They are retested after receiving the antitoxin. Below are the differences between before and after readings for the subjects. A +ve value indicates that the subject's resistance to the poison has increased.

Note: For your benefit, these are already sorted in increasing order of their absolute values.

Subject	1	2	3	4	5	6	7	8	9	10	11	12
Difference	-0.1	-0.2	-0.5	0.6	0.7	0.9	0.9	-1.2	1.3	-1.6	2.0	2.7

Use this information to answer the following **2 questions** (the current one and the next).

Let M be the median difference between before and after resistance readings. Then, the Z-statistic for the Signed-Rank test of $H_0 : M = 0$ vs. $H_a : M > 0$ is closest to:

(a) Z-statistic = 0.377.

(b) Z-statistic = 0.577.

(c) Z-statistic = 1.177.

(d) Z-statistic = 2.177.

(e) Z-statistic = 1.269.

Since, the null value of the median difference is 0, to conduct the Signed-Rank test we simply need to look at the ranks of the absolute values of the observations (of differences) and take the sum of those ranks for which the observations were positive. Since the data was already sorted, and accounting for one pair of ties (at 0.9), the ranks and the signs are given by:

Subject	1	2	3	4	5	6	7	8	9	10	11	12
Difference	-0.1	-0.2	-0.5	0.6	0.7	0.9	0.9	-1.2	1.3	-1.6	2.0	2.7
Absolute Values	0.1	0.2	0.5	0.6	0.7	0.9	0.9	1.2	1.3	1.6	2.0	2.7
Ranks	1	2	3	4	5	6.5	6.5	8	9	10	11	12
Signs	-	-	-	+	+	+	+	-	+	-	+	+

Using pg. 228-230 of Chapter 4 notes, the sum of the ranks for all the positive differences is: $S_+ = 4 + 5 + 6.5 + 6.5 + 9 + 11 + 12 = 54$, and the Z-statistic for the Signed-Rank test is:

$$Z = \frac{S_+ - n(n+1)/4}{\sqrt{n(n+1)(2n+1)/24}} = \frac{54 - 12(13)/4}{\sqrt{12(13)(25)/24}} = \frac{54 - 39}{\sqrt{162.5}} = 1.177.$$

17. The P-value for the Signed-Rank test of $H_0 : M = 0$ vs. $H_a : M > 0$ is closest to:

(a) P-value = 0.282.

(b) P-value = 0.564.

(c) P-value = 0.564.

(d) P-value = 0.242.

(e) P-value = 0.121.

Since the alternative H_a is one-sided and 'right-sided', we reject for large values of Z. Under H_0, $Z \sim N(0, 1)$. Hence, the P-value for this test is $P(Z > Z_{obs}) = P(Z > 1.177)$ where $Z \sim N(0, 1)$. Using the Normal table, this P-value roughly equals $1 - 0.8790 = 0.1210$.

18. Over the past five years an insurance company has had a mix of 40% Whole Life policies, 20% Universal Life policies, 25% Annual Renewable-Term (ART) policies, and 15% other types of policies. A change in this mix over the long haul could require a change in the commission structure, reserves, and possibly investments. A sample of 1000 policies issued over the last few months gave the following results.

Category	Number of Policies
Whole Life	320
Universal Life	280
ART	240
Other	160

Use this information to answer the following **2 questions** (the current and the next one).

It turns out that at a level of significance 0.05, one can reject the hypothesis that the policies issued in the last few months match the historical percentages. Which of the following is true of the χ^2 test used to reach this conclusion?

(a) The test statistic is 10.71.

(b) The test statistic is 41.90.

(c) The test statistic is 40.90.

(d) The test statistic is 69.07.

(e) The test statistic is 49.07.

This corresponds to a Multinomial experiment with $k = 4$ categories and the χ^2 goodness-of-fit test with category probabilities fully specified. Under H_0, i.e. assuming no change in the historical percentages, the expected numbers of policies in the four categories are as follows:

$$0.40(1000) = 400, \quad 0.20(1000) = 200, \quad 0.25(1000) = 250 \quad \text{and} \quad 0.15(1000) = 150.$$

Using pg. 179-181 of Chapter 3 notes, the χ^2 statistic is thus

$$\chi^2 = \frac{(320-400)^2}{400} + \frac{(280-200)^2}{200} + \frac{(240-250)^2}{250} + \frac{(160-150)^2}{150} = 49.07.$$

19. Which of the following is the best conclusion?

(a) $\chi^2_{3;0.05} = 7.815$, so we reject the null.

(b) $\chi^2_{3;0.05} = 7.815$, so we fail to reject the null.

(c) The percentages of Whole Life and Universal Life policies have both increased substantially.

(d) $\chi^2_{4;0.05} = 9.488$, so we reject the null.

(e) The insurance salesman who collected the data fudged the numbers.

Under H_0, $\chi^2 \sim \chi^2_{k-1} \equiv \chi^2_3$, and we reject H_0 for large values of χ^2. Thus, the P-value is $P(\chi^2_3 > \chi^2_{obs}) = P(\chi^2_3 > 49.07)$, where $\chi^2_{obs} = 49.07$ is the observed value of the test statistic. Comparing this value with the percentiles of the χ^2_3 distribution, we observe that $\chi^2_{3;0.005} = 12.838$ and $49.07 > 12.838$, and so, P-value $= P(\chi^2_3 > 49.07) < P(\chi^2_3 > 12.838) = 0.005$.

```
> fit=lm(y~x1+x2+x3+x4)
> fit12=lm(y~x1+x2)
> fit23=lm(y~x2+x3)
> fit14=lm(y~x1+x4)

> summary(fit)

Call:
lm(formula = y ~ x1 + x2 + x3 + x4)

Residuals:
    Min      1Q  Median      3Q     Max 
-49.396 -13.384  -1.188  11.823  54.457 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  8.71778   19.08414   0.457    0.648    
x1           5.23047    0.30209  17.314   <2e-16 ***
x2          -3.44166    0.37161  -9.262   <2e-16 ***
x3          -0.01945    0.25754  -0.076    0.940    
x4           2.31569    1.43913   1.609    0.109    
---
Residual standard error: 20.73 on 195 degrees of freedom
Multiple R-squared:  0.8567,    Adjusted R-squared:  0.8538 
F-statistic: 291.4 on 4 and 195 DF,  p-value: < 2.2e-16

> anova(fit)
Analysis of Variance Table

Response: y
           Df Sum Sq Mean Sq   F value Pr(>F)    
x1          1 435690  435690 1013.9356 <2e-16 ***
x2          1  64091   64091  149.1531 <2e-16 ***
x3          1      4       4    0.0085 0.9267    
x4          1   1113    1113    2.5892 0.1092    
Residuals 195  83792     430                     
---

> summary(fit12)

Call:
lm(formula = y ~ x1 + x2)

Residuals:
    Min      1Q  Median      3Q     Max 
-53.090 -11.848  -1.162  12.081  56.250 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  11.9296    18.9933   0.628    0.531    
x1            5.1922     0.1571  33.056   <2e-16 ***
x2           -3.4910     0.2863 -12.194   <2e-16 ***
---
Residual standard error: 20.76 on 197 degrees of freedom
Multiple R-squared:  0.8548,    Adjusted R-squared:  0.8533 
F-statistic: 579.8 on 2 and 197 DF,  p-value: < 2.2e-16

> anova(fit12)
Analysis of Variance Table

Response: y
           Df Sum Sq Mean Sq F value    Pr(>F)    
x1          1 435690  435690  1010.9 < 2.2e-16 ***
x2          1  64091   64091   148.7 < 2.2e-16 ***
Residuals 197  84908     431                      
---
```

```
> summary(fit23)

Call:
lm(formula = y ~ x2 + x3)

Residuals:
    Min      1Q  Median      3Q     Max
-83.683 -20.668  -1.358  22.768  81.655

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 128.3294    28.1482   4.559 9.01e-06 ***
x2           -6.6487     0.5115 -12.998  < 2e-16 ***
x3            3.7776     0.2122  17.799  < 2e-16 ***
---
Residual standard error: 32.89 on 197 degrees of freedom
Multiple R-squared:  0.6355,    Adjusted R-squared:  0.6318
F-statistic: 171.7 on 2 and 197 DF,  p-value: < 2.2e-16

> anova(fit23)
Analysis of Variance Table

Response: y
           Df Sum Sq Mean Sq F value    Pr(>F)
x2          1  28825   28825  26.645 5.952e-07 ***
x3          1 342745  342745 316.820 < 2.2e-16 ***
Residuals 197 213120    1082
---

> summary(fit14)

Call:
lm(formula = y ~ x1 + x4)

Residuals:
    Min      1Q  Median      3Q     Max
-83.880 -17.343  -0.782  17.766  76.185

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -123.7013    20.5277  -6.026 8.12e-09 ***
x1             4.9875     0.2054  24.280  < 2e-16 ***
x4             3.5022     1.8883   1.855   0.0651 .
---
Residual standard error: 27.26 on 197 degrees of freedom
Multiple R-squared:  0.7495,    Adjusted R-squared:  0.747
F-statistic: 294.8 on 2 and 197 DF,  p-value: < 2.2e-16

> anova(fit14)
Analysis of Variance Table

Response: y
           Df Sum Sq Mean Sq  F value  Pr(>F)
x1          1 435690  435690 586.1064 < 2e-16 ***
x4          1   2557    2557   3.4396 0.06515 .
Residuals 197 146443     743
---
```

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 T Distribution

df	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.5 Percentiles of the Chi-square Distribution

df	0.5%	1%	2.5%	5%	10%	90%	95%	97.5%	99%	99.5%
1	0.000	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	14.041	30.813	33.924	36.781	40.289	42.796
23	9.260	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.559
25	10.520	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314	46.928
26	11.160	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290
27	11.808	12.879	14.573	16.151	18.114	36.741	40.113	43.195	46.963	49.645
28	12.461	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.993
29	13.121	14.256	16.047	17.708	19.768	39.087	42.557	45.722	49.588	52.336
30	13.787	14.953	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672
40	20.707	22.164	24.433	26.509	29.051	51.805	55.758	59.342	63.691	66.766
60	35.534	37.485	40.482	43.188	46.459	74.397	79.082	83.298	88.379	91.952
80	51.172	53.540	57.153	60.391	64.278	96.578	101.879	106.629	112.329	116.321

Table A.6 Percentiles of the F Distribution (ν_1 = Numerator df; ν_2 = Denominator df)

		ν_1										
ν_2	α	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									
ν	α	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
∞	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