NAME:____

VERSION A

Choose the best answer. Do not write letters in the margin or communicate with other students in any way; if you do you will receive a 0. If you have a question note it on your exam and ask for clarification when your exam is returned. In the meantime choose the best answer. Neither the proctors nor Dr. Cox will answer statistics questions during the exam.

Please check each question and possible answers thoroughly as questions at the bottom of a page sometimes run onto the next page. Please verify that your test version and scantron version are the same.

This exam has 25 questions. Cost for turning in exam late: 1st minute is 10 points; 2nd minute is 20 additional points (30 total); 3rd minute is 30 points (60 total) ; 4th minute is 40 points (100 total); 5th minute is 50 points (150 total); 6th minute is 80 points (230 total) and no exams are graded past that point.

As noted in the syllabus entering your student ID incorrectly on your scantron results in a 10 point deduction.

Use the output below to answer the appropriate questions in the exam. Be sure to check which table the question references.

ANOVA Table 1

	SS	DF	MS	F	p
Between	1023.914	3	341.305	14.567	0.0000023
Within	843.461	36	23.429		

Here is regression output from a model where screen size (in diagonal centimeters) is the independent variable and battery life in hours is the dependent variable. The data are from a random sample of laptops (source, Donnelly 2015).

Regression Table 2

Regression Statistics	
R Square	0.2478041
Adjusted R Square	0.1224381
Standard Error	0.3988209
Observations	8

ANOVA

	SS	DF	MS	F	p
Regression	0.314	1	0.314	1.977	0.2093591
Residual	0.954	6	0.159		

	coefficients	standard error	t stat	p-value
intercept	5.4715834	1.1355027	4.81864	0.00294
screen size (in diagonal centimeters)	??	0.031827	-1.40593	0.20936

1. You are going to test the following hypothesis, $H_0 : \mu = 25$ at the .05 level of significance. The report from the satisticians shows that the p-value for this test is .121.

   (a) You fail to reject the null hypothesis.
   (b) You fail to reject the null hypothesis for the one tailed version of the test only.
   (c) You reject the null hypothesis.
   (d) None of the above.

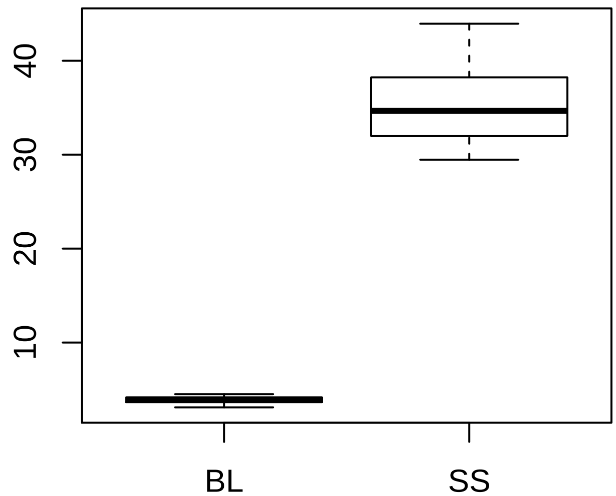
2. Suppose you see that all values of your independent variable are the same. You use this independent variable in your simple regression. Choose the most correct statement below.

   (a) Your intercept term will be zero and the coefficient on your independent variable will be the average of your dependent variable.
   (b) The R square value will be 0.
   (c) This cannot happen unless you have fewer observations for the independent variable than for the dependent variable.
   (d) The sum of squares due to regression will be 0.
   (e) B and C are correct.

3. What is the median for the following data set? 1, 1.5, 2, 2.5, 3, 3.5, 4, 5, 6.

   (a) 4.5.
   (b) 4.
   (c) 3.5.
   (d) 3.
   (e) none of the above.

4. Consider the graph below. Choose the best answer below



(a) The mean for BL is greater than the mean for SS.

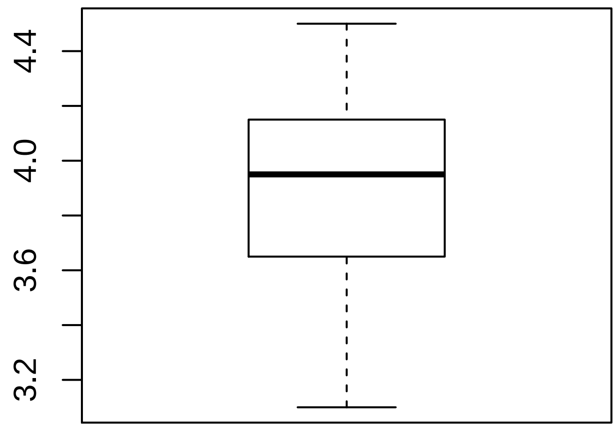
(b) The graph clearly shows that the coefficient of variation for BL is smaller than the coefficient of variation for SS.

(c) The interquartile range for BL is larger than the interquartile range of SS.

(d) The graph clearly shows that the mean of SS is larger than the median of SS.

(e) None of the above.

5. The graph(s) shown here is/are



(a) scatter plots

(b) histograms

(c) stem and leaf displays

(d) box and whiskers plots

6. The probability that a z value is less than 0.46 is

(a) 0.6772419

(b) 0.6483984

(c) 0.8314724

(d) 0.4540693

(e) 0.4840466

7. Consult Table 1. From the table you can conclude that the total sum of squares is

(a) -180.4527778

(b) 933.6875

(c) 1867.375
(d) 843.4611111
(e) 1333.8392857

8. Consult Table 1. From the table you can conclude that the total number of observations used in this analysis/experiment was?
   (a) 40
   (b) 38
   (c) 3
   (d) 5
   (e) 36

9. Consult Table 1. From the table what can you conclude concerning the null hypothesis?
   (a) fail to reject
   (b) reject the null
   (c) cannot be determined
   (d) depends on the number of observations.

10. Consult Table 1. What is the test statistic?
    (a) 7.2836593.
    (b) 14.5673185.
    (c) 23.4294753.
    (d) 0.0000023.
    (e) 341.3046296.

11. Consult Table 1. What is the critical value for a test at the .05 level?
    (a) 3.16
    (b) 3.86
    (c) 4.38
    (d) 1.88
    (e) 2.87

12. Suppose that you are conducting an ANOVA, with $\alpha = .05$, and you find that the overall variance is 100 and the groupwise variances are 0, 0, .5 and .1 and all groups have 10 of observations. (Note the groupwise variance is the variance of a specific group so that group 3 has a variance of .5.)

    (a) This will lead to a small F statistic and you will fail to reject the null hypothesis.
    (b) This will lead to a large F statistic and you will reject the null hypothesis.
    (c) It is impossible to tell anything from this since we do not know the number of overall observations.
    (d) We can find the number of overall observations but we cannot find the F statistic given this information.
    (e) None of the above.

13. Suppose that the number of homework assigments a professor gives follows a Poisson distribution with a mean of 13. What is the probability of drawing a professor that gives exactly 12 assignments?

    (a) 0.0549718
    (b) 0.0828592
    (c) 0.1143679
    (d) 0.0984185
    (e) 0.1099398

14. Suppose that you collect data on apartment prices in Tempe. You look at 50 different apartments and find a mean of 766 and a standard deviation of 190.9. Construct a 95% confidence interval for the mean apartment price. The interval is

    (a) [715.7468203, 816.2531797]
    (b) [705.7468203, 826.2531797]
    (c) [711.7468203, 820.2531797]
    (d) [640.5721383, 738.2278617]

15. Suppose that you collect data on apartment prices in Tempe. You look at 50 different apartments and find a mean of 766 and a standard deviation of 190.9. Test the hypothesis, $H_0 : \mu = 739$ at the .05 level of significance.

    (a) the test statistic is 1.0000986 so we rejcet the null
    (b) the test statistic is 0.7200888 so we fail to reject

(c) the test statistic is 1.0000986 so we fail to reject

(d) the test statistic is 1.2101085 so we rejcet the null

16. Suppose you have a random variable that is exponentially distributed with a mean of 42. What is the probability of observing a random variable drawn from this distribution with a value of less than 20?

    (a) 0.3167896

    (b) 0.6141787

    (c) 0.4352819

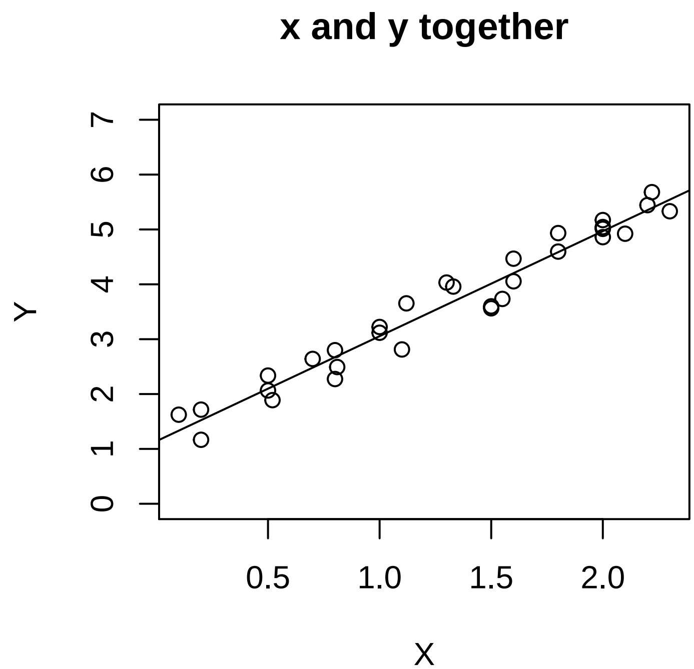
    (d) 0.3788548

17. If you are conducting ANOVA and you find an F statistic that is less than 1 then this proves that the null hypothesis is true.

    (a) True.

    (b) False.

18. The following graph shows various pairs of x and y values. Suppose y is the dependent variable in a regression and x is the independent variable. Choose the best option below.



(a) The regression results will show an intercept of about 1 and an estimated slope of about 2.

(b) The regression results will show an intercept of about 0 and an estimated slope of about 2.

(c) The regression results will show an intercept of about 5 and an estimated slope of about -2.

(d) The regression results will show an intercept of about 1 and an estimated slope of about 1.

(e) The regression results cannot be determined but because X varies so much in the graph we know that R^2 will be relatively small.

19. Suppose that the total sum of squares is 400 and SSR=300. The R^2 is

(a) .25.

(b) .75.

(c) 1.33

(d) 100.

20. Consider the regression output in Table 2. What is the estimated change in battery life in hours associated with a 1 unit increase in screen size (in diagonal centimeters) ?

    (a) 1.977
    (b) -0.0639236
    (c) -0.0447465
    (d) -0.0344204
    (e) 0.3988209

21. Consider the regression output in Table 2. What is the estimated variance of the error term?

    (a) 0.1590581
    (b) 0.1994104
    (c) 0.3988209
    (d) 1.977

22. Consider the regression output in Table 2. What is the percentage of variation in battery life in hours that can be explained by the variation in screen size (in diagonal centimeters)?

    (a) 24.7804124%
    (b) 12.2438144%
    (c) -23.7804124%
    (d) 29.7364948%

23. Consider the regression output in Table 2. What is the predicted or estimated battery life in hours when screen size (in diagonal centimeters) is 33 centimeters?

    (a) 3.9949476
    (b) 3.9457264
    (c) 4.0396942
    (d) 4.0844407
    (e) 3.9188785

24. Consider the regression output in Table 2. Suppose you want to test the hypothesis that screen size (in diagonal centimeters) reduces battery life by 30 minutes per additional inch, i.e. $H_0 : \beta_1 = -.5$ (30 minutes is .5 of an hour). What is the test statistic for this hypothesis?

(a) -18.5952313
(b) 14.3040241
(c) 10.0128168
(d) -1.4059323
(e) -5.7216096

25. Consider the regression output in Table 2. Suppose you want to test the hypothesis that screen size (in diagonal centimeters) has no impact on battery life in hours, i.e. $H_0 : \beta_1 = 0$. What is your conclusion for this hypothesis test? (Use $\alpha = .05$.)

(a) This is inconclusive unless we know whether it is a right tail or a left tail test.
(b) This cannot be determined without the appropriate df.
(c) rejcet the null
(d) fail to reject

t Distribution

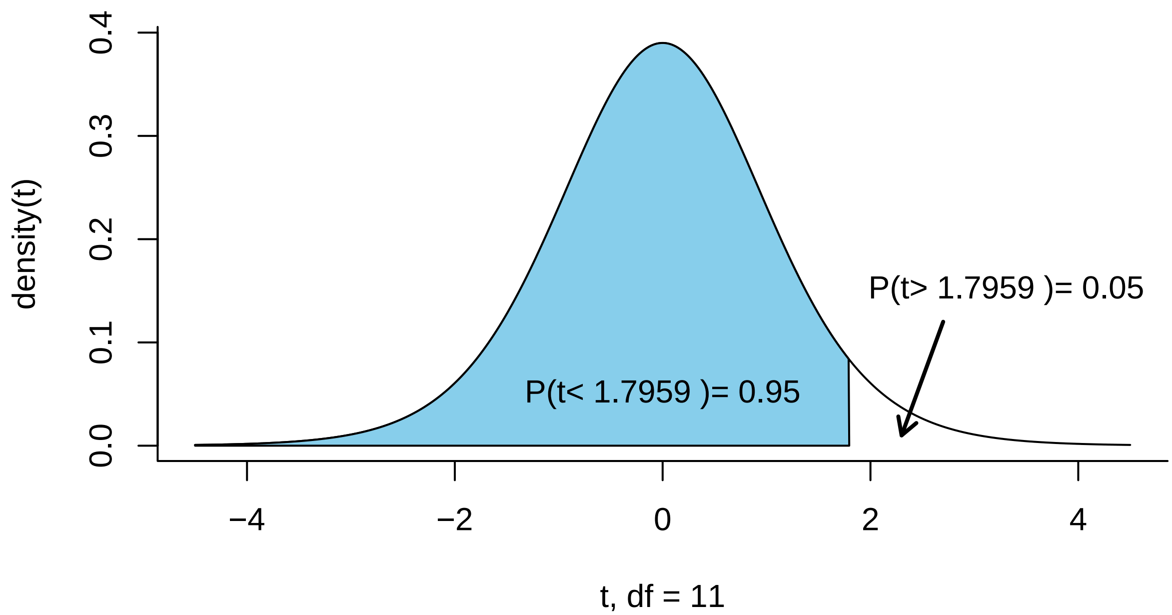


Table of t values. Degrees of Freedom are in the "df" column and *right tail* probabilities are in the top row. Example: for df=11 and $P(t > 2.7181) = .01$.					
	Right tail probability				
df	0.1	0.05	0.025	0.01	0.005
1	3.0777	6.3138	12.7062	31.8205	63.6567
2	1.8856	2.92	4.3027	6.9646	9.9248
3	1.6377	2.3534	3.1824	4.5407	5.8409
4	1.5332	2.1318	2.7764	3.7469	4.6041
5	1.4759	2.015	2.5706	3.3649	4.0321
6	1.4398	1.9432	2.4469	3.1427	3.7074
7	1.4149	1.8946	2.3646	2.998	3.4995
8	1.3968	1.8595	2.306	2.8965	3.3554
9	1.383	1.8331	2.2622	2.8214	3.2498
10	1.3722	1.8125	2.2281	2.7638	3.1693
11	1.3634	1.7959	2.201	**2.7181**	3.1058
12	1.3562	1.7823	2.1788	2.681	3.0545
13	1.3502	1.7709	2.1604	2.6503	3.0123
14	1.345	1.7613	2.1448	2.6245	2.9768
15	1.3406	1.7531	2.1314	2.6025	2.9467
16	1.3368	1.7459	2.1199	2.5835	2.9208
17	1.3334	1.7396	2.1098	2.5669	2.8982
18	1.3304	1.7341	2.1009	2.5524	2.8784
19	1.3277	1.7291	2.093	2.5395	2.8609
20	1.3253	1.7247	2.086	2.528	2.8453

Table of t values **continued**. For values not on table approximate based on available values.					
	Right tail probability				
df	0.1	0.05	0.025	0.01	0.005
21	1.3232	1.7207	2.0796	2.5176	2.8314
22	1.3212	1.7171	2.0739	2.5083	2.8188
23	1.3195	1.7139	2.0687	2.4999	2.8073
24	1.3178	1.7109	2.0639	2.4922	2.7969
25	1.3163	1.7081	2.0595	2.4851	2.7874
26	1.315	1.7056	2.0555	2.4786	2.7787
27	1.3137	1.7033	2.0518	2.4727	2.7707
28	1.3125	1.7011	2.0484	2.4671	2.7633
29	1.3114	1.6991	2.0452	2.462	2.7564
30	1.3104	1.6973	2.0423	2.4573	2.75
31	1.3095	1.6955	2.0395	2.4528	2.744
32	1.3086	1.6939	2.0369	2.4487	2.7385
33	1.3077	1.6924	2.0345	2.4448	2.7333
34	1.307	1.6909	2.0322	2.4411	2.7284
35	1.3062	1.6896	2.0301	2.4377	2.7238
36	1.3055	1.6883	2.0281	2.4345	2.7195
37	1.3049	1.6871	2.0262	2.4314	2.7154
38	1.3042	1.686	2.0244	2.4286	2.7116
39	1.3036	1.6849	2.0227	2.4258	2.7079
40	1.3031	1.6839	2.0211	2.4233	2.7045
41	1.3025	1.6829	2.0195	2.4208	2.7012
42	1.302	1.682	2.0181	2.4185	2.6981
43	1.3016	1.6811	2.0167	2.4163	2.6951
44	1.3011	1.6802	2.0154	2.4141	2.6923
45	1.3006	1.6794	2.0141	2.4121	2.6896
46	1.3002	1.6787	2.0129	2.4102	2.687
47	1.2998	1.6779	2.0117	2.4083	2.6846
48	1.2994	1.6772	2.0106	2.4066	2.6822
49	1.2991	1.6766	2.0096	2.4049	2.68
50	1.2987	1.6759	2.0086	2.4033	2.6778
51	1.2984	1.6753	2.0076	2.4017	2.6757
52	1.298	1.6747	2.0066	2.4002	2.6737
53	1.2977	1.6741	2.0057	2.3988	2.6718
54	1.2974	1.6736	2.0049	2.3974	2.67
55	1.2971	1.673	2.004	2.3961	2.6682
56	1.2969	1.6725	2.0032	2.3948	2.6665
57	1.2966	1.672	2.0025	2.3936	2.6649
58	1.2963	1.6716	2.0017	2.3924	2.6633
59	1.2961	1.6711	2.001	2.3912	2.6618
60	1.2958	1.6706	2.0003	2.3901	2.6603
61	1.2956	1.6702	1.9996	2.389	2.6589
62	1.2954	1.6698	1.999	2.388	2.6575
63	1.2951	1.6694	1.9983	2.387	2.6561
64	1.2949	1.669	1.9977	2.386	2.6549
65	1.2947	1.6686	1.9971	2.3851	2.6536
66	1.2945	1.6683	1.9966	2.3842	2.6524
67	1.2943	1.6679	1.996	2.3833	2.6512
68	1.2941	1.6676	1.9955	2.3824	2.6501
69	1.2939	1.6672	1.9949	2.3816	2.649
70	1.2938	1.6669	1.9944	2.3808	2.6479

Table of t values **continued**. For values not on table approximate based on available values.					
	Right tail probability				
df	0.1	0.05	0.025	0.01	0.005
71	1.2936	1.6666	1.9939	2.38	2.6469
72	1.2934	1.6663	1.9935	2.3793	2.6459
73	1.2933	1.666	1.993	2.3785	2.6449
74	1.2931	1.6657	1.9925	2.3778	2.6439
75	1.2929	1.6654	1.9921	2.3771	2.643
76	1.2928	1.6652	1.9917	2.3764	2.6421
77	1.2926	1.6649	1.9913	2.3758	2.6412
78	1.2925	1.6646	1.9908	2.3751	2.6403
79	1.2924	1.6644	1.9905	2.3745	2.6395
80	1.2922	1.6641	1.9901	2.3739	2.6387
81	1.2921	1.6639	1.9897	2.3733	2.6379
82	1.292	1.6636	1.9893	2.3727	2.6371
83	1.2918	1.6634	1.989	2.3721	2.6364
84	1.2917	1.6632	1.9886	2.3716	2.6356
85	1.2916	1.663	1.9883	2.371	2.6349
86	1.2915	1.6628	1.9879	2.3705	2.6342
87	1.2914	1.6626	1.9876	2.37	2.6335
88	1.2912	1.6624	1.9873	2.3695	2.6329
89	1.2911	1.6622	1.987	2.369	2.6322
90	1.291	1.662	1.9867	2.3685	2.6316
91	1.2909	1.6618	1.9864	2.368	2.6309
92	1.2908	1.6616	1.9861	2.3676	2.6303
93	1.2907	1.6614	1.9858	2.3671	2.6297
94	1.2906	1.6612	1.9855	2.3667	2.6291
95	1.2905	1.6611	1.9853	2.3662	2.6286
96	1.2904	1.6609	1.985	2.3658	2.628
97	1.2903	1.6607	1.9847	2.3654	2.6275
98	1.2902	1.6606	1.9845	2.365	2.6269
99	1.2902	1.6604	1.9842	2.3646	2.6264
100	1.2901	1.6602	1.984	2.3642	2.6259
120	1.2886	1.6577	1.9799	2.3578	2.6174
150	1.2872	1.6551	1.9759	2.3515	2.609
200	1.2858	1.6525	1.9719	2.3451	2.6006
300	1.2844	1.6499	1.9679	2.3388	2.5923
500	1.2832	1.6479	1.9647	2.3338	2.5857
1000	1.2824	1.6464	1.9623	2.3301	2.5808
10000+	1.2816	1.645	1.9602	2.3267	2.5763

Notes: for values not on the table, e.g. where df are 101 you can either approximate or for homework and test taking in ECN221 you can simply use the previous available value on the table. For example, you have a problem where df=125 and you are looking for a right tail probability of .05 then you can simply go to the row where df=120 and find a value of 1.6577 and use that value. For *left tail* probabilities the t values are the negative of the values in the table since the t distribution is symmetric around 0. For example, for 50 df and a left tail probability of .1 the t value is -1.2987.

Standard Normal, aka z Distribution

Table of z values and probabilities for the standard normal distribution. z is in the first column plus the top row. Each cell shows $P(X \leq z)$. For example $P(Z \leq 1.04) = .8508$. For $z < 0$ subtract the value from 1, e.g., $P(X \leq -1.04) = 1 - .8508 = .1492$.

z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0	0.5	0.504	0.508	0.512	0.516	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	0.591	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	0.648	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.67	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.695	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.719	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	0.758	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.791	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	0.834	0.8365	0.8389
1	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	0.877	0.879	0.881	0.883
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.898	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	0.937	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	0.975	0.9756	0.9761	0.9767
2	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	0.983	0.9834	0.9838	0.9842	0.9846	0.985	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.989
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	0.992	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.994	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	0.996	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.997	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	0.998	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	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.999	0.999
3.1	0.999	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999

Table of F values for right tail probabilities of .05. Degrees of freedom for the denominator are in the first column and degrees of freedom for the numerator are in the top row.

denom. df	numerator df									
	1	2	3	4	5	7	10	15	50	100
1	161.448	199.5	215.707	224.583	230.162	236.768	241.882	245.95	251.774	253.041
2	18.513	19	19.164	19.247	19.296	19.353	19.396	19.429	19.476	19.486
3	10.128	9.552	9.277	9.117	9.013	8.887	8.786	8.703	8.581	8.554
4	7.709	6.944	6.591	6.388	6.256	6.094	5.964	5.858	5.699	5.664
5	6.608	5.786	5.409	5.192	5.05	4.876	4.735	4.619	4.444	4.405
6	5.987	5.143	4.757	4.534	4.387	4.207	4.06	3.938	3.754	3.712
7	5.591	4.737	4.347	4.12	3.972	3.787	3.637	3.511	3.319	3.275
8	5.318	4.459	4.066	3.838	3.687	3.5	3.347	3.218	3.02	2.975
9	5.117	4.256	3.863	3.633	3.482	3.293	3.137	3.006	2.803	2.756
10	4.965	4.103	3.708	3.478	3.326	3.135	2.978	2.845	2.637	2.588
11	4.844	3.982	3.587	3.357	3.204	3.012	2.854	2.719	2.507	2.457
12	4.747	3.885	3.49	3.259	3.106	2.913	2.753	2.617	2.401	2.35
13	4.667	3.806	3.411	3.179	3.025	2.832	2.671	2.533	2.314	2.261
14	4.6	3.739	3.344	3.112	2.958	2.764	2.602	2.463	2.241	2.187
15	4.543	3.682	3.287	3.056	2.901	2.707	2.544	2.403	2.178	2.123
16	4.494	3.634	3.239	3.007	2.852	2.657	2.494	2.352	2.124	2.068
17	4.451	3.592	3.197	2.965	2.81	2.614	2.45	2.308	2.077	2.02
18	4.414	3.555	3.16	2.928	2.773	2.577	2.412	2.269	2.035	1.978
19	4.381	3.522	3.127	2.895	2.74	2.544	2.378	2.234	1.999	1.94
20	4.351	3.493	3.098	2.866	2.711	2.514	2.348	2.203	1.966	1.907
21	4.325	3.467	3.072	2.84	2.685	2.488	2.321	2.176	1.936	1.876
22	4.301	3.443	3.049	2.817	2.661	2.464	2.297	2.151	1.909	1.849
23	4.279	3.422	3.028	2.796	2.64	2.442	2.275	2.128	1.885	1.823
24	4.26	3.403	3.009	2.776	2.621	2.423	2.255	2.108	1.863	1.8
25	4.242	3.385	2.991	2.759	2.603	2.405	2.236	2.089	1.842	1.779
26	4.225	3.369	2.975	2.743	2.587	2.388	2.22	2.072	1.823	1.76
27	4.21	3.354	2.96	2.728	2.572	2.373	2.204	2.056	1.806	1.742
28	4.196	3.34	2.947	2.714	2.558	2.359	2.19	2.041	1.79	1.725
29	4.183	3.328	2.934	2.701	2.545	2.346	2.177	2.027	1.775	1.71
30	4.171	3.316	2.922	2.69	2.534	2.334	2.165	2.015	1.761	1.695
35	4.121	3.267	2.874	2.641	2.485	2.285	2.114	1.963	1.703	1.635
40	4.085	3.232	2.839	2.606	2.449	2.249	2.077	1.924	1.66	1.589
50	4.034	3.183	2.79	2.557	2.4	2.199	2.026	1.871	1.599	1.525
60	4.001	3.15	2.758	2.525	2.368	2.167	1.993	1.836	1.559	1.481
80	3.96	3.111	2.719	2.486	2.329	2.126	1.951	1.793	1.508	1.426
100	3.936	3.087	2.696	2.463	2.305	2.103	1.927	1.768	1.477	1.392
1000	3.851	3.005	2.614	2.381	2.223	2.019	1.84	1.676	1.363	1.26

Chapter 3:

sample mean: $\bar{x} = \sum_{i=1}^{n} x_i/n$

sample variance: $s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$.

sample standard deviation: $s = \sqrt{s^2}$.

Coefficient of Variation: $CV = \frac{s}{\bar{x}}\left(100\%\right)$

sample z-Score: $z = \frac{x_i - \bar{x}}{s}$

Interquartile Range: IQR $= Q_3 - Q_1$.

Sample Covariance: $s_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{n-1}$

Sample Correlation Coefficient: $r_{xy} = s_{xy}/(s_x s_y)$

Chapter 4:

The complement rule: $P(A) + P(A') = 1$

addition rule: $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$

conditional probability: $P(A|B) = \frac{P(A \text{ and } B)}{P(B)}$

Bayes' Theorem

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{P(A_1)P(B|A_1) + P(A_2)P(B|A_2) + \cdots + P(A_n)P(B|A_n)}$$

Combinations: ${}_nC_x = \frac{n!}{(n-x)!x!}$

Chapter 5:

Expected Value and mean of a Discrete Probability Distribution:

$$E(x) = \mu = \sum_{i=1}^{n} x_i P(x_i)$$

Variance of a Discrete Probability Distribution:

$$\sigma^2 = \sum_{i=1}^{n}(x_i - \mu)^2 P(x_i)$$

Chapter 5 continued:

Binomial Probability Dist.: $P(x,n) = \frac{n!}{(n-x)!x!}p^x(q)^{(n-x)}$

Mean of a Binomial Distribution: $\mu = np$

Standard Dev. of a Binomial Distribution: $\sigma = \sqrt{npq}$

Poisson Probability Distribution: $P(x) = \frac{\lambda^x e^{-\lambda}}{x!}$

Chapter 6:

Normal PDF: $f(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-(1/2)[(x-\mu)/\sigma]^2}$

the z-score: $z = \frac{x-\mu}{\sigma}$

Exponential PDF: $f(x) = \lambda e^{-\lambda x}$

Exponential CDF: $P(x \le a) = 1 - e^{-a\lambda}$

Standard Dev. of Exponential Dist.: $\sigma = \mu = \frac{1}{\lambda}$

Continuous Uniform PDF

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{if } a \le x \le b \\ 0 & \text{otherwise} \end{cases}$$

Uniform CDF: $P(x_1 \le x \le x_2) = \frac{x_2 - x_1}{b-a}$

mean of the continuous uniform dist.: $\mu = \frac{a+b}{2}$

standard dev. of the continuous uniform dist.: $\sigma = \frac{b-a}{\sqrt{12}}$

Chapter 7:

standard error of the mean: $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$.

z-score for the mean: $z_{\bar{x}} = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}}$

Chapter 7 continued:

sample proportion: $\bar{p} = \frac{x}{n}$

standard error of the proportion: $\sigma_{\bar{p}} = \sqrt{\frac{p(1-p)}{n}}$

Chapter 8:

Confidence Interval for the mean (σ known):

$$\bar{x} \pm z_{\alpha/2}\sigma_{\bar{x}}$$

margin of error for a CI for the mean: $ME_{\bar{x}} = z_{\alpha/2}\sigma_{\bar{x}}$

approximate standard error of the mean: $\hat{\sigma}_{\bar{x}} = \frac{s}{\sqrt{n}}$.

Confidence Interval for the mean (σ unknown):

$$\bar{x} \pm t_{\alpha/2}\hat{\sigma}_{\bar{x}}$$

Sample Size needed to Estimate a population mean

$$n = \frac{(z_{\alpha/2})^2\sigma^2}{(ME_{\bar{x}})^2}$$

Sample Size needed to Estimate the population proportion

$$n = \frac{(z_{\alpha/2})^2\bar{p}(1-\bar{p})}{(ME_p)^2}$$

Chapter 9:

the z-test statistic for a hypothesis test for the population mean (when σ is known)

$$z_{\bar{x}} = \frac{\bar{x} - \mu_{H_0}}{\sigma/\sqrt{n}}$$

the t-test statistic for a hypothesis test for the population mean (when σ is unknown)

$$t_{\bar{x}} = \frac{\bar{x} - \mu_{H_0}}{s/\sqrt{n}}$$

Chapter 10:

the mean of the sampling distribution for the difference in means:

$$\mu_{\bar{x}_1-\bar{x}_2} = \mu_{\bar{x}_1} - \mu_{\bar{x}_2}$$

the standard error of the difference between two means:

$$\sigma_{\bar{x}_1-\bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

the z-test statistic for a hypothesis test for the difference between two means (σ_1 and σ_2 known)

$$z_{\bar{x}} = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)_{H_0}}{\sigma_{\bar{x}_1-\bar{x}_2}}$$

the t-test statistic for a hypothesis test for the difference between two means (σ_1 and σ_2 unknown but equal)

$$t_{\bar{x}} = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)_{H_0}}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

pooled variance: $s_p^2 = \dfrac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{(n_1-1)+(n_2-1)}$

the t-test statistic for a hypothesis test for the difference between two means (σ_1 and σ_2 unknown and unequal)

$$t_{\bar{x}} = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)_{H_0}}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}}$$

Confidence Interval for the difference between the means of two independent populations (σ_1 and σ_2 unknown but equal)

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2}\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}$$

the matched-pair difference: $d = x_1 - x_2$

the mean of matched-pair difference: $\bar{d} = \dfrac{\sum_{i=1}^n d_i}{n}$

the standard deviation of the matched-pair differences

$$s_d = \sqrt{\frac{\sum_{i=1}^n (d_i - \bar{d})^2}{n-1}}$$

the t-Test Statistic for a Matched-Pair hypothesis test for the mean

$$t_{\bar{x}} = \frac{\bar{d} - (\mu_d)_{H_0}}{s_d/\sqrt{n}}$$

Chapter 11:

the total sum of squares (SST): $SST = \sum_{j=1}^{k}\sum_{i=1}^{n_j}(x_{ij} - \bar{\bar{x}})^2$

the mean square total (MST): $MST = \dfrac{SST}{n_T - 1}$

the partitioning of the Total Sum of Squares (SST) for a One-Way ANOVA: SST=SSB+SSW.

sum of squares between (SSB): $SSB = \sum_{j=1}^{k} n - j(\bar{x}_j - \bar{\bar{x}})^2$

the mean square between (MSB): $MSB = \dfrac{SSB}{k-1}$

sum of squares within (SSW): $SSW = \sum_{j=1}^{k}\sum_{i=1}^{n_j}(x_{ij} - \bar{x}_j)^2$

the mean square within (MSW): $MSW = \dfrac{SSW}{n_T - k}$

the F-test statistic for One-Way ANOVA: $F_{\bar{x}} = \dfrac{MSB}{MSW}$

Tukey-Kramer critical range:

$$CR_{ij} = Q_\alpha\sqrt{\frac{MSW}{2}\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}$$

Chapter 14: simple linear regression model for a population $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$

$$\hat{y} = b_0 + b_1 x \qquad e_i = y_i - \hat{y}_i$$

sum of squares error: $SSE = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$

total sum of squares (SST): $SST = \sum(y - \bar{y})^2$

$$SST = SSR + SSE$$

sum of squares regression (SSR): $SSR = \sum(\hat{y} - \bar{y})^2$

Chapter 14 continued:

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

F-statistic for the coef. of determination: $F = \dfrac{SSR}{SSE/(n-2)}$

Standard Error of the Estimate, $s_e = \sqrt{SSE/(n-2)}$.
Confidence Interval (CI) for an average value of Y:

$$CI = \hat{y}^* \pm t_{\alpha/2} s_e \sqrt{\frac{1}{n} + \frac{(x-\bar{x})^2}{\sum x^2 - ((\sum x)^2/n)}}$$

Prediction Interval (PI) fpr a specific value of y:

$$PI = \hat{y}^* \pm t_{\alpha/2} s_e \sqrt{1 + \frac{1}{n} + \frac{(x-\bar{x})^2}{\sum x^2 - ((\sum x)^2/n)}}$$

t-test statistic for the regression slope: $t = \dfrac{b_1 - \beta_1}{s_b}$

the standard error of a slope: $s_b = \dfrac{s_e}{\sqrt{\sum x^2 - n(\bar{x})^2}}$

confid. interval for the pop. slope: $CI = b_1 \pm t_{\alpha/2} s_b$

Chapter 15:

mean square regression (MSR): $MSR = SSR/k$

mean square error (MSE): $MSE = SSE/(n-k-1)$

F-test stat. for the overall regression model: $F = \dfrac{MSR}{MSE}$

adjusted multiple coef. of det.: $R_A^2 = 1 - (1 - R^2)\dfrac{n-1}{n-k-1}$

variance inflation factor: $VIF_j = \dfrac{1}{1 - R_j^2}$

Other Math Rule Reminders:

$$e^x = \exp(x) \quad \text{and} \quad \ln 1 = 0 \quad \text{and} \quad \ln e = 1$$

$$x! = (x)(x-1)(x-2)\cdots(2)(1) \quad \text{and} \quad 0! = 1 \quad \text{and} \quad x^0 = 1$$

Key

1. a
2. e
3. d
4. e
5. d
6. a
7. c
8. a
9. b
10. b
11. e
12. b, note that the groupwise variances imply $SSW = 0 + 0 + 9 * .5 + 9 * .1 = 5.4$ we know this because $s_j^2 = SSE_j/(n_j - 1)$ and all groups have 10 observations. Then we also know that SST=39*100 since there are 40 observations and the overall variance is 100 and the sample variance is just $SST/(n_T - 1)$. So this will give us a MSW of 5.4/36 and MSB of $(3900 - 5.4)/3 = 1298.2$ so that your F statistic is over 7000.
13. e
14. c, use $\bar{x} \pm t_{\{.025,n-1\}}(s/\sqrt{n})$
15. c
16. d
17. b, note that failing to reject (aka accepting) the null hypothesis does not mean that the null is true, it is possible that we commit a Type II error when we fail to reject the null.
18. a
19. b
20. c

21. a
22. a
23. a
24. b
25. d