Math 2200 Spring 2016, Final Exam

You may use *any* calculator. You may use a 4×6 inch notecard as a cheat sheet.

1. If A and B are events that satisfy the following three properties
   P(A) = 0.725,
   P(A ∪ B) = 0.780, and
   A and B are independent,
   then what is P(B)?

   A) 0.080 B) 0.095 C) 0.110 D) 0.125 E) 0.140
   F) 0.155 G) 0.170 H) 0.185 I) 0.200 J) 0.215

Solution. *Because A and B are independent events, we have $P(A \cap B) = P(A)\,P(B) = 0.725\,P(B)$. It follows that*

$$\begin{aligned} 0.780 &= P(A \cup B) \\ &= P(A) + P(B) - P(A \cap B) \\ &= 0.725 + P(B) - 0.725\,P(B) \\ &= 0.725 + (1 - 0.725)\,P(B) \\ &= 0.725 + 0.275\,P(B). \end{aligned}$$

Thus,

$$P(B) = \frac{0.780 - 0.725}{0.275} = 0.200.$$

Answer: ***I***

2. At Midwestern University, 1% of the female students are taller than 1.825m, and 4% of the male students are taller than 1.825m. There are 40,000 students at Midwestern University, of which 18,000 are male. From the entire student body, one student is selected at random. That selected student is taller than 1.825m. What is the probability that the selected student is female?

   A) 0.1275 B) 0.1488 C) 0.1701 D) 0.1914 E) 0.2127
   F) 0.2340 G) 0.2553 H) 0.2766 I) 0.2979 J) 0.3192

Solution. *Let F, M, and T be the events that a randomly selected student is, respectively, female, male, and taller than 1.825m. Then P(F) = (40000 - 18000)/40000 = 0.55, P(M) = 18000/40000 = 0.45, $P(T\,|\,F) = 0.1$, and $P(T\,|\,M) = 0.4$. By Bayes's Law*

$$P(F\,|\,T) = \frac{P(T\,|\,F)\,P(F)}{P(T\,|\,F)\,P(F) + P(T\,|\,M)\,P(M)} = \frac{(0.1)(0.55)}{(0.1)(0.55) + (0.4)(0.45} = 0.2340426.$$

Answer: ***F***

3. There was a young person of Bantry,
   Who frequently slept in the pantry,
   When disturbed by the mice,
   She appeased them with rice,
   That judicious young person from Bantry.

   There were three types of mice that would check out the pantry: field mice, dormice, and agouti mice. When a mouse entered the pantry looking for a tasty morsel, it was three times as likely to be a dormouse as an agouti mouse, and two times as likely to be a field mouse as a dormouse. That young person of Bantry did not want to encourage the field mice, so only 1 grain was given to a field mouse who showed up. Dormice were somehat more welcome, relatively speaking, and received 2 grains. The young person of Bantry was most favorably inclined towards agouti mice, which were fed 5 grains. Let X be the number of grains that were dispensed when a mouse disturbed the pantry sleeper. What was E(X)? (The next problem will ask for the variance of X.)

A) 1.0 B) 1.1 C) 1.2 D) 1.3 E) 1.4
F) 1.5 G) 1.6 H) 1.7 I) 1.8 J) 1.9

Solution. *Thank you Edward Lear. When a mouse came knocking, let p be the probability that it was an agouti mouse, q the probability that it was a dormouse, and r the probability that it was a field mouse. Then $p+q+r=1$. We are given that $q=3p$ and $r=2q=6p$. Therefore $p+3p+6p=1$, or $10p=1$, or $p=0.1$. It follows that $q=0.3$ and $r=0.6$. We have E(X) = 1×P(X = 1) + 2×P(X = 2) + 5×P(X = 5) $= 1\times 0.6+2\times 0.3+5\times 0.1$, or 1.7.*
Answer: ***H***

4. Let X be the random variable of the preceding problem. What is the variance of X?

A) 1.35 B) 1.37 C) 1.39 D) 1.41 E) 1.43
F) 1.45 G) 1.47 H) 1.49 I) 1.51 J) 1.53

Solution. *We have*

$$\begin{aligned}
\mathit{Var(X)} &= (1-1.7)^2\times \mathit{P(X=1)}+(2-1.7)^2\times \mathit{P(X=2)}+(5-1.7)^2\times \mathit{P(X=5)}\\
&= (1-1.7)^2\times 0.6+(2-1.7)^2\times 0.3+(5-1.7)^2\times 0.1\\
&= 1.410.
\end{aligned}$$

An alternative calculation is

$$\begin{aligned}
\mathit{Var(X)} &= E\left(X^2\right)-E(X)^2\\
&= \left((1)^2\times \mathit{P(X=1)}+(2)^2\times \mathit{P(X=2)}+(5)^2\times \mathit{P(X=5)}\ \right)-(1.7)^2\\
&= \left((1)^2\times 0.6+(2)^2\times 0.3+(5)^2\times 0.1\right)-(1.7)^2\\
&= 1.410.
\end{aligned}$$

Answer: ***D***

5. Ripped (more like ripped-off) from a Penn State exam! History suggests that scores on the Math portion of the Standard Achievement Test (SAT) are normally distributed with a mean of 529 and a variance of 5732. History also suggests that scores on the Verbal portion of the SAT are normally distributed with a mean of 474 and a variance of 6368. Select two students at random. Let X denote the first student's Math score, and let Y denote the second student's Verbal score. What is P(X > Y)?

A) 0.5037 B) 0.5350 C) 0.5663 D) 0.5976 E) 0.6289
F) 0.6602 G) 0.6915 H) 0.7228 I) 0.7541 J) 0.7854

Solution. *The first thing to notice is that "watermelon" problems are found on exams across the nation. Second: the Penn State examiners (assuming that they did not also rip off this problem from someplace else) were imaginative, albeit deceptive. The data in this problem is entirely fabricated. The SAT has been called the Scholastic Assessment Test and the Scholastic Aptitude Test, but never the Standard Achievement Test. Nevertheless, it is an excellently contrived exam question. To solve it, we first observe that $X-Y$ is a normal* *random variable with mean* $529 - 474$, *or 55, and standard deviation* $\sqrt{5732 + 6368}$, *or 110. It follows that*

$$\begin{aligned} P(X > Y) &= P(X - Y > 0) \\ &= P\left(\frac{X - Y - 55}{110} > \frac{-55}{110}\right) \\ &= P\left(Z > \frac{-1}{2}\right) \\ &= P\left(Z < \frac{1}{2}\right) \\ &= \Phi\,(0.5) \\ &= 0.6914625. \end{aligned}$$

Answer: ***G***

6. The probability that a woman will be alive one year after she has survived her first heart attack is 0.62. Suppose that, in a study of women who have just had a first heart attack, 1600 participants are randomly selected. Calculate the *approximate* probability P that more than 1000 are alive one year after they have had their first heart attacks. Use the normal approximation with correction for continuity. *If you use the Phi table, then you must interpolate quite accurately.* Whatever the method you used to obtain the value of P, round P to 4 decimal places: you will have a number of the form $0.d_1d_2d_3d_4$. Answer this problem with
$$10^{(1/0.d_1d_2d_3d_4)}.$$
(The purpose of this problem is to calculate P in the specified way. The purpose of the requested answer is to bully you into calculating P in the specified way.)

A) 1027.943 B) 1032.340 C) 1036.737 D) 1041.134 E) 1045.531
F) 1049.928 G) 1054.325 H) 1058.722 I) 1063.119 J) 1067.516

Solution. *Let* X_1, X_2, … , X_{1600} *be the i.i.d. Bernoulli trials, each with* $p = 0.62$ *and* $q = 0.38$. *Then*

$$X_1, +X_2, + \cdots + X_{1600} \approx N\left(1600\,p, \sqrt{1600\,p\,q}\right).$$

It follows that

$$\begin{aligned} P &= P(X_1, +X_2, + \cdots + X_{1600} > 1000) \\ &= P(X_1, +X_2, + \cdots + X_{1600} \geq 1001) \\ &\approx P\left(N\left(1600\,p, \sqrt{1600\,p\,q}\right) \geq 1000.5\right) \\ &= P\left(\frac{N\left(1600\,p, \sqrt{1600\,p\,q}\right) - 1600\,p}{\sqrt{1600\,p\,q}} \geq \frac{1000.5 - 1600\,p}{\sqrt{1600\,p\,q}}\right) \\ &= P(Z \geq 0.4377955) \\ &= 1 - P(Z < 0.4377955) \\ &= 1 - \Phi(0.4377955) \\ &= 1 - \left(0.6664 + \frac{77955}{100000} \times (0.6700 - 0.6664)\right) \\ &= 1 - \left(0.6664 + \frac{77955}{100000} \times (0.6700 - 0.6664)\right) \\ &= 1 - 0.6692064 \\ &= 0.3307936. \end{aligned}$$

Rounded to 4 decimal places, we have $P = 0.3308$, *and* $10^{1/P} = 1054.325$.
Answer: ***G***

7. Statistics concerning coronary heart disease tend to be gender-linked. Accordingly, statistical investigations have often focused on a single sex. Together, the Nurses' Health Study of 121,700 female nurses and the Health Professional Follow-up Study of 51,529 males have provided data that have been useful for comparison. In a 2014 report that appeared in the British Medical Journal (BMJ), researchers limited their attention to participants in these two studies who experienced a non-fatal first heart attack while participating in one of the two studies. Of the 2258 women in that group, 1576 were alive at the time of the BMJ study. For the men, the comparable numbers were 1389 out of 1840. If p_M and p_F are the population proportions of male and female survivors respectively, a hypothesis test of the null hypothesis $p_M = p_F$ against the alternative $p_M > p_F$ would conclusively reject the null hypothesis. To see why, calculate and answer with the z-score of $\widehat{p_M} - \widehat{p_F}$ under the null hypothesis.

A) 2.8974 B) 3.0297 C) 3.1620 D) 3.2943 E) 3.4266
F) 3.5589 G) 3.6912 H) 3.8235 I) 3.9558 J) 4.0881

Solution. *Let* n *denote the sample size, 1840, of males and* m *the sample size, 2258, of females. The mean of* $\widehat{p_M} - \widehat{p_F}$ *is* $p_M - p_F$, *which is 0 under the null hypothesis. The standard deviation of* $\widehat{p_M} - \widehat{p_F}$ *is*

$$\begin{aligned}\sqrt{\frac{p_M\,(1-p_M)}{n} + \frac{p_F\,(1-p_F)}{m}} &\approx \sqrt{\frac{\widehat{p_M}\,(1-\widehat{p_M})}{n} + \frac{\widehat{p_F}\,(1-\widehat{p_F})}{m}} \\ &= \sqrt{\frac{1389\,(1840-1389)}{1840^3} + \frac{1576\,(2258-1576)}{2258^3}} \\ &= 0.01392558.\end{aligned}$$

The z-score of the observed value of the test statistic $\widehat{p_M} - \widehat{p_F}$ *under the null hypothesis is*

$$\frac{\widehat{p_M} - \widehat{p_F} - 0}{\sqrt{p_M\,(1-p_M)/n + p_F\,(1-p_F)/m}} \approx \frac{1389/1840 - 1576/2258}{0.01392558} = 4.088053.$$

Answer: ***J***

8. Reading and 'riting and 'rithmetic, taught to the tune of the hick'ry stick. That was in 1907. Fast forward to 2008. Here is a cross-tabulation of the SAT statistics by exam and by gender.

	Test-Takers	Critical Reading		Writing		Mathematics	
	Number	Mean	SD	Mean	SD	Mean	SD
Male	704,226	504	114	533	116	488	111
Female	812,764	500	110	500	111	501	109

Let's consider the Critical Reading exam. The cited means, $\mu_M = 504$ and $\mu_F = 500$ are true population means because the College Board is able to conduct a census. Likewise, the standard deviations $\sigma_M = 114$ and $\sigma_F = 110$ are population standard deviations. Let us now proceed on a hypothetical path. Assume that the population standard deviations are known and have the given values. Assume, however, that the given means are not population means but sample means based on random samples of 2800 male test-takers and 3200 female test-takers. Let us test the null hypothesis that $\mu_M = \mu_F$ against the alternative that $\mu_M > \mu_F$. What is the p-value?

A) 0.0389 B) 0.0502 C) 0.0615 D) 0.0728 E) 0.0841
F) 0.0954 G) 0.1067 H) 0.1180 I) 0.1293 J) 0.1406

Solution. *Let X_M be the score of a randomly selected male. X_F be the score of a randomly selected female. Let σ_M and σ_F be the standard deviations of X_M and X_F respectively. We are given that $\sigma_M = 114$ and $\sigma_F = 110$. The observed value of $\overline{X_M}$ based on a random sample of size $n = 2800$ is 504. The observed value of $\overline{X_F}$ based on a random sample of size $m = 3200$ is 500. The observed value of $\overline{X_M} - \overline{X_F}$ is $504 - 500$, or 4. The standard deviation σ_D of $\overline{X_M} - \overline{X_F}$ is given by*

$$\sigma_D = \sqrt{\frac{\sigma_M^2}{n} + \frac{\sigma_F^2}{m}} = \sqrt{\frac{(114)^2}{2800} + \frac{(110)^2}{3200}} = 2.902185.$$

We can now calculate the requested p-value:

$$\begin{aligned}
\textit{p-value} &= P\left(\overline{X_M} - \overline{X_F} \geq 4 \,|\, \mu_M = \mu_F\right) \\
&= P\left(\frac{\overline{X_M} - \overline{X_F} - (\mu_M - \mu_F)}{2.902185} \geq \frac{4 - (\mu_M - \mu_F)}{2.902185} \,\middle|\, \mu_M = \mu_F\right) \\
&= P\left(Z \geq \frac{4}{2.902185}\right) \\
&= P(Z \geq 1.378272) \\
&= 1 - P(Z \leq 1.378272) \\
&= 1 - \Phi(1.3782720) \\
&= 1 - 0.9159403 \\
&= 0.08405966.
\end{aligned}$$

Answer: ***E***

9. In a laboratory study, 9 mice were infected with amyloid plaques characteristic of Alzheimer's disease. A potential treatment was administered to 5 of the mice, and 4 of the infected mice were left untreated as a control group. Let X denote the number of years that a treated mouse lives after infection and treatment. Let Y be the number of years that an untreated mouse lives after infection. Assume that X and Y are normally distributed. The observed values of X and Y in the experiment were

Treated (X)	2.1	5.3	1.4	4.6	0.9
Untreated (Y)	1.9	0.5	2.8	3.1	

The sample standard deviations are $S_X = 1.9705$ and $S_Y = 1.1673$. Using $\overline{X} - \overline{Y}$ as the test statistic, test the null hypothesis that the population means satisfy $\mu_X = \mu_Y$ against the alternative that $\mu_X \neq \mu_Y$. Do not assume equal variances (so don't pool). Assign the degrees of freedom for the test statistic conservatively. What is the positive endpoint of the critical region at significance level 0.05?

A) 2.2502 B) 2.3739 C) 2.4976 D) 2.6213 E) 2.7450
F) 2.8687 G) 2.9924 H) 3.1161 I) 3.2398 J) 3.3635

Solution. *We assign the degrees of freedom to be 3: one less than the minimum of the two sample sizes. From the tables, we find that $t_{0.025,3} = 3.1824$. The positive endpoint of the critical region is*

$$t_{0.025,3} \times \sqrt{\frac{S_X^2}{5} + \frac{S_Y^2}{4}}, \qquad \textit{or} \qquad 3.1824 \times \sqrt{\frac{(1.9705)^2}{5} + \frac{(1.167)^2}{4}}, \qquad \textit{or} \qquad 3.363492.$$

Answer: ***J***

10. An experiment involved a random sample of 6 individuals who were somewhat depressed to very depressed. At the beginning of the experiment, each subject filled out a Life Satisfaction questionnaire. Scores on this well-being scale are normally distributed and range from 5 to 35: the higher the score, the greater is one's satisfaction with life. After a course of treatment with the antidepressant fluoxetine, each subject completed the questionnaire again. The scores are tabulated.

	Subject 1	Subject 2	Subject 3	Subject 4	Subject 5	Subject 6
Before Treatment (X)	18	6	9	13	16	14
After Treatment (Y)	16	19	22	20	16	21

In a traditional one-sided hypothesis test of $\mu_{\mathrm{Y}} = \mu_{\mathrm{X}}$ against the alternative that $\mu_{\mathrm{Y}} > \mu_{\mathrm{X}}$, by what amount does the test statistic exceed the critical value? Use significance level 0.05 and the course test statistic, not the standardized test statistic.

A) 0.8564 B) 0.9507 C) 1.0450 D) 1.1393 E) 1.2336
F) 1.3279 G) 1.4222 H) 1.5165 I) 1.6108 J) 1.7051

Solution. *Clearly this is a small sample—t-scores will be used instead of z-scores. Because values of X and Y are paired and not independent, we analyze this as a one-sample test of the difference of questionnaire scores. Let U denote the differences of the scores: -2, 13, 13, 7, 0, 7. We calculate* $\overline{U} = 6.333333$ *and* $S_U = 6.314006$. *The sample mean 6.3333 is the observed value of the test statistic. We look up* $t_{0.05\,,\,6-1} = 2.0150$. *The critical value cv is given by*

$$cv = t_{0.05\,,\,6-1}\,\frac{S_U}{\sqrt{6}} = 2.0150 \times \frac{6.314006}{\sqrt{6}} = 5.1940.$$

The requested value $\overline{U} - cv$ *is* $6.3333 - 5.1940$, *or 1.1393.*
Answer: ***D***

11. Breaking news! The NY primary election was held 19 April 2016. In an exit poll, Republican primary voters were asked, "Which best describes your feeling about a Donald Trump presidency?". Four choices were offered, but, for the purposes of exam questioning, we will merge the two positives and the two negatives. The percentages are tabulated. Also, 1391 Democratic Party voters were asked, "Which best describes your feeling about a Hillary Clinton presidency?". Four choices were offered, but, for the purposes of exam questioning, we will merge the two positives and the two negatives. The numbers for the two choices (after the megers) are tabulated.

	Republicans on Trump	Democrats on Clinton
Scared or concerned	37%	460
Optimistic or excited	63%	931
Total	100%	1391

A classical hypothesis test at significance level 0.01 rejects the hypothesis that the distribution of feelings Democrats have for their likely nominee fits the distribution of feelings Republicans have for their likely nominee. By how much does the observed test statistic exceed the critical value?

A) 1.6667 B) 1.7685 C) 1.8703 D) 1.9721 E) 2.0739
F) 2.1757 G) 2.2775 H) 2.3793 I) 2.4811 J) 2.5829

Solution. *A column of expected values is calculated by multiplying 1391 by 0.37 and 0.63.*

	Republicans on Trump	*Observed Count*	*Expected Count*
Scared or concerned	*37%*	*460*	*514.67*
Optimistic or excited	*63%*	*931*	*876.33*
Total	*100%*	*1391*	*1391.00*

The observed value if the χ^2_{2-1} *test statistic is*

$$\chi^2_{obs} = \frac{(460 - 514.67)^2}{514.67} + \frac{(931 - 876.33)^2}{876.33} = 9.2178.$$

The critical value is $\chi^2_{0.01,1}$*, or 6.6349. The answer is* $\chi^2_{obs} - \chi^2_{0.01,1}$*, or* $9.2178 - 6.6349$*, or 2.5829.*
Answer: ***J***

12. Benford's Law is an observation about the frequency distribution of leading digits in many real-life sets of numerical data. The law states that in many naturally occurring collections of numbers, about 30.1% of the listed numbers begin with the digit 1, about 17.6% with the digit 2, 12.5% with the digit 3, 9.7% with the digit 4, 7.9% with the digit 5, and 22.2% with one of the digits 6, 7, 8, or 9. It is often asserted that the IRS believes that naturally-occuring numbers on tax returns follow Benford's Law, and, accordingly, counts leading digits to flag potential tax cheats. (Just one of the many ways sneaky statisticians monitor sneaky number manipulators.) Do the counts of the 60 tallest structures by category fit Benford's Law? The counts are given in the table below.

Leading digit	1	2	3	4	5	6, 7, 8, or 9
Benford's Law	30.1%	17.6%	12.5%	9.7%	7.9%	22.2%
Actual Counts	18	8	8	6	10	10

To decide if the observed distribution fits the distribution predicted by Benford's Law, report the p-value (using a fit of the distributions as the null hypothesis).

A) 0.0091 B) 0.0562 C) 0.1033 D) 0.1504 E) 0.1976
F) 0.2446 G) 0.2917 H) 0.3388 I) 0.3859 J) 0.4330

Solution. *The expected counts are obtained by multiplying 60 by each of the given percentages:*

Leading digit	*1*	*2*	*3*	*4*	*5*	*6, 7, 8, or 9*
Observed	*18*	*8*	*8*	*6*	*10*	*10*
Expected	*18.06*	*10.56*	*7.50*	*5.82*	*4.74*	*13.32*

Expected Counts under Null Hypothesis of Fit

The observed value v *of the* χ^2_5 *test statistic is given by*

$$v = \frac{(18 - 18.06)^2}{18.06} + \frac{(8 - 10.56)^2}{10.56} + \frac{(8 - 7.50)^2}{7.50} + \frac{(6 - 5.82)^2}{5.82} + \frac{(10 - 4.74)^2}{4.74} + \frac{(10 - 13.32)^2}{13.32} = 7.32426.$$

The p-value is

$$p\text{-}value = P\left(\chi^2_5 > v\right) = P\left(\chi^2_5 > 7.32426\right) = 0.1976.$$

The p-value just recorded was obtained using R. Using the provided χ^2*-table the calculation would begin by scanning the line for df = 5 and locating the consecutive bracketing numbers 9.2364, corresponding to* $\alpha = 0.100$*, and 7.2893, corresponding to* $\alpha = 0.200$*. By "bracketing" we mean that the observed test statistic 7.32426 lies between 9.2364 and 7.2893. At this point we can be quite confident of the correct answer choice. The p-value we seek lies between the two corresponding values of* α*, namely 0.100 and 0.200. There are three answer choices in this range, but 7.32426 is much closer to the bracketing value 7.2893 than to the bracketing value 9.2364, an we therefore expect the p-value to be much closer to 0.200 than to 0.100. That reasoning leads us to answer choice 0.1976. But let's do the interpolation anyway.*

$$p\text{-}value = \frac{(0.100 - 0.200)}{(9.2364 - 7.2893)}(7.32426 - 7.2893) + 0.200 = 0.1982045.$$

Answer: ***E***

13. In a 1993 survey of 1074 Californian men in the 35–44 age group, the employment and marital statuses of the surveyees were cross-tabulated.

	MNS	WDS	NM
Employed	679	103	114
Unemployed	63	10	20
Not in labor force	42	18	25

Employment Status by Marital Status
(MNS = married, not separated; WDS = widowed, divorced, or separated; NM = never married)

Are the three distributions of employment statuses homogeneous across marital categories? The test statistic is a sum of 9 components—one component for each cell. If we start with the component from the last entry of the bottom row, then we need not continue with calculating the test statistic. The contribution from the cited cell by itself exceeds the critical value at significance level 0.05. By how much?

A) 0.6484 B) 0.9505 C) 1.2526 D) 1.5547 E) 1.8568
F) 2.1589 G) 2.4610 H) 2.7631 I) 3.0652 J) 3.3673

Solution. *In the solution, we will complete a table of expected values under the null hypothesis of homogeneity. To answer this problem, we actually only need to calculate the last of the 9 cell entries, however. The number of MNS in the survey is 784. The number of WDS in the survey is 131. The number of NM in the survey is 159. The proportion of MNS is* $p_1 = 784/1074 = 0.7299814$*, the proportion of WDS is* $p_2 = 131/1074 = 0.1219739$*, and the proportion of NM is* $p_3 = 159/1074 = 0.1480447$*. If these proportions held in each of the employment statuses, there there would be* 896×0.7299814*, or 654.0633, employed MNS,* 896×0.1219739*, or 109.2886, employed WDS,* 896×0.1480447*, or 132.6481, employed NM,* 93×0.7299814*, or 67.88827, unemployed MNS,* 93×0.1219739*, or 11.34357, unemployed WDS,* 93×0.1480447*, or 13.76816, unemployed NM,* 85×0.7299814*, or 62.04842, MNS not in labor force,* 85×0.1219739*, or 10.36778, WDS not in labor force, and* 85×0.1480447*, or 12.58380, NM not in labor force. These values are tabulated as follows:*

	MNS	*WDS*	*NM*
Employed	*654.0633*	*1109.2886*	*132.6481*
Unemployed	*67.88827*	*11.34357*	*13.76816*
Not in labor force	*62.04842*	*10.36778*	*12.58380*

Expected under True Null Hypothesis of Homogeneity

The component of the chi square test statistic arising from the last cell is $(25 - 12.58380)^2/12.58380$*, or 12.25083. The critical value is* $\chi^2_{0.05,\,(3-1)\times(3-1)}$*, or 9.4877. The answer is* $12.25083 - 9.4877$*, or 2.76313.*

Answer: ***H***

14. The Physicians Health Study, reported in 1988 in the New England Journal of Medicine, conclusively established a dependence between suffering heart attacks (myocardial infarction, or MI) and following an aspirin regimen (namely, the latter reduces the former). Consider the following data from the study:

	MI	No MI
Placebo	189	10845
Aspirin	104	10933

What is the observed value of the test statistic for deciding whether MI and aspirin therapy are independent? (Don't bother with the p-value. It is so small that the trial was stopped early: the conclusion was so overwhelming that it would have been unethical to continue the experiment and allow the participants in the control group to suffer preventable heart attacks.)

A) 22.5648 B) 23.1771 C) 23.7894 D) 24.4017 E) 25.0140
F) 25.6263 G) 26.2386 H) 26.8509 I) 27.4632 J) 28.0755

Solution. *Here is the table with marginal totals:*

	MI	*No MI*	*Total*
Placebo	*189*	*10845*	*11034*
Aspirin	*104*	*10933*	*11037*
Total	*293*	*21778*	*22071*

What is the If MI and aspirin therapy were not related, then the expected number in each cell would equal the product of the row total and the column total divided by the table total. This would result in the following table of expected counts under the null hypothesis of independence:

	MI	*No MI*
Placebo	*146.48*	*10887.52*
Aspirin	*146.52*	*10890.48*

The test statistic is

$$\frac{(189-146.48)^2}{146.48}+\frac{(10845-10887.52)^2}{10887.52}+\frac{(104-146.52)^2}{146.52}+\frac{(10933-10890.48)^2}{10890.48},$$

which evaluates to 25.0140.
Answer: ***E***

15. The table below records bivariate data: the opercular breathing rate Y of a goldfish (in counts per minute) and temperature X (in degrees centigrade) of the water in which the goldfish is swimming.

X	9	12	15	18	21	24	27
Y	49	46	60	86	87	110	124

In terms of unknown coefficients β_0, β_1, and ρ, the true linear model is $y = \beta_0 + \beta_1 x$ with linear correlation ρ. The regression line calculated from the 7 tabulated observations is $y = b_0 + b_1 x$ with sample Pearson correlation r. Sample statistics are: $\overline{\mathrm{X}} = 18$, $\overline{\mathrm{Y}} = 80.2857$, $\mathrm{S_X} = 6.4807$, $\mathrm{S_Y} = 30.1038$, $r = 0.9739$, $b_0 = -1.1429$, and $b_1 = 4.5238$. In a classical one-sided hypothesis test of $\mathrm{H_0} : \beta_1 = 0$ versus the alternative $\mathrm{H_a} : \beta_1 > 0$, what is the endpoint of the critical region if b_1 is the test statistic and 0.05 is the significance level? (The next two problems continue with this data.)

A) 0.581 B) 0.704 C) 0.827 D) 0.950 E) 1.073
F) 1.196 G) 1.319 H) 1.442 I) 1.565 J) 1.688

Solution. *With $n = 7$, we calculate $SST = (n-1)\,S_Y^2 = 6\cdot(30.1038)^2 = 5437.4286$, $SSR = r^2\,SST = (0.9739)^2(5437.4286) = 5157.2989$, $SSE = SST - SSR = 280.2857$, $S_e = \sqrt{SSE/(n-2)} = \sqrt{280.2857/5} = 7.4871$, and $SE(b_1) = \left(1/\sqrt{n-1}\right)\, S_e/S_X = 0.471645$. The endpoint of the critical region is $t_{0.05,n-2}\times SE(b_1)$, or 2.0150×0.471645, or 0.95039.*
Answer: ***D***

16. What is the lower endpoint of the 60% confidence interval (centered at $b_1 = 4.5238$) for the slope of the regression line?

    A) 3.6581 B) 3.7198 C) 3.7815 D) 3.8432 E) 3.9049
    F) 3.9666 G) 4.0283 H) 4.0900 I) 4.1517 J) 4.2134

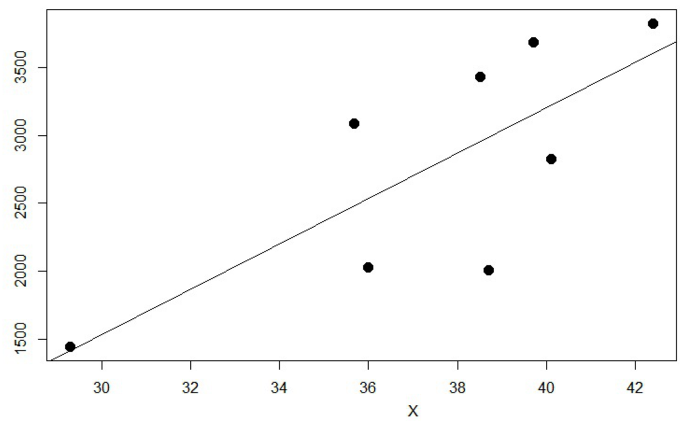
Solution. *We calculated* $SE(b_1) = 0.471645$*. Now we look up* $t_{0.200,7-2} = 0.9195$*. The lower endpoint of the 60% confidence interval for* β_1 *is* $b_1 - t_{0.200,7-1} \times SE(b_1)$*, or or* $4.5238 - 0.9195 \times 0.471645$*, or 4.0901.*
Answer: ***H***

17. Suppose that Sushi the goldfish is swimming in water that is 25.5 degrees centigrade. Let $\widehat{y_*}$ be the breathing rate for Sushi predicted by the regression line. What is the smallest integer U that is greater than the upper endpoint of a 95% confidence interval (centered at $\widehat{y_*}$) for Sushi's breathing rate?

    A) 129 B) 130 C) 131 D) 132 E) 133
    F) 134 G) 135 H) 136 I) 137 J) 138

Solution. *Let* $\widehat{y}$ *be the value of* Y *that is predicted by the regression line for* $x_* = 25.5$*. Then* $SE(\widehat{y}) = \sqrt{SE(b_1) \cdot (25.5 - \overline{x}) + (8/7) \cdot S_e^2}$*. From the preceding work, we have* $SE(b_1) = 0.4716449725$ *and* $S_e = 7.487131822$*. Therefore,* $SE(\widehat{y}) = \sqrt{(0.4716449725)^2 \cdot (25.5 - 18)^2 + (8/7)) \cdot (7.487131822)^2} = 8.750889$*. Because* $t_{0.025,5} = 2.5706$*, it follows that the lower bound of a 95% confidence interval for the regression line prediction corresponding to* $x = 25.5$ *is* $-1.143 + 4.524 \cdot 25.5 + 2.5706 \cdot 8.750889$*, or 136.8.*
Answer: ***I***

18. Let X be human birth weight (in grams) and let Y be gestational age at birth (in weeks). To study a linear model $y = \beta_0 + \beta_1 x$ with linear correlation ρ for these variables, a Boston hospital recorded several pairs of observations. We will use only 8 of the data points in the problem. The scatter plot below shows the eight points we have selected, as well as the regression line $y = b_0 + b_1 x$ that is based on the 8 plotted points.



Sample statistics are $S_X = 3.9778$, SST $= 5394283.495$, and SSR $= 3101449.708$. That is all you need! In a one-sided hypothesis test of $H_0 : \beta_1 = 0$ versus the alternative $H_a : \beta_1 > 0$, what is the p-value?

    A) 0.0002 B) 0.0072 C) 0.0146 D) 0.0218 E) 0.0290
    F) 0.0362 G) 0.0434 H) 0.0506 I) 0.0578 b J) 0.0650

Solution. *First we calculate the sample linear correlation:*

$$r^2 = \frac{SSR}{SST} = \frac{3101449.708}{5394283.495} = 0.5749511888.$$

It follows that $r = \sqrt{0.5749511888} = 0.758255$. *(Note: we are using the positive square root because the scatter plot shows that* $r > 0$.*)*

Next, we calculate S_Y*:*

$$S_Y = \sqrt{\frac{1}{n-1}\, SST} = \sqrt{\frac{1}{7}\,(5394283.495)} = 877.8450.$$

The next sample statistic we calculate is b_1*:*

$$b_1 = r\,\frac{S_Y}{S_X} = (0.758255)\,\frac{877.8450}{3.9778} = 167.3363.$$

Three more sample statistics to go. We have $SSE = SST - SSR = 5394283.495 - 3101449.708 = 2292833.787$. *Therefore,*

$$S_e = \sqrt{\frac{1}{n-2}\, SSE} = \sqrt{\frac{1}{6}\,(2292833.787)} = 618.174.$$

It follows that

$$SE(b_1) = \frac{1}{\sqrt{n-1}}\,\frac{S_e}{S_X} = \frac{1}{\sqrt{7}}\,\frac{618.174}{3.9778} = 58.7379.$$

At last, the p-value can be calculated:

$$\begin{aligned}
p\text{-}value &= P\left(t_{n-2} > \frac{b_1}{SE(b_1)}\right) \\
&= P\left(t_6 > \frac{167.3363}{58.7379}\right) \\
&= P(t_6 > 2.8489) \\
&= 0.0146.
\end{aligned}$$

Answer: ***C***

19. The preceding problem concerned a hypothesis test to decide whether the slope of a linear model is zero (the null hypothesis) or positive (the alternative). Because the sign of the correlation coefficient is the same as the sign of the slope of the regression line, another approach is to test whether $H_0 : \rho = 0$ or $H_a : \rho > 0$. Using significance level $\alpha = 0.025$, perform a classical test of the aforementioned one-sided hypothesis test. Calculate the observed value T of the course test statistic, find the critical value v, and answer with $T - v$.

    A) 0.3807 B) 0.4020 C) 0.4233 D) 0.4446 E) 0.4659
    F) 0.4872 G) 0.5085 H) 0.5298 I) 0.5511 J) 0.5724

Solution. *The test statistic is* $r\sqrt{n-2}/\sqrt{1-r^2}$, *which has observed value* $T = 0.758255\,\sqrt{8-2}/\sqrt{1-0.5749511888}$, *or* $T = 2.848863$. *The critical value of a one-sided test with significance level 0.025 is* $v = t_{0.025,\,8-2} = 2.4469$. *The answer is* $T - v = 2.848863 - 2.4469 = 0.4020$.

Answer: ***B***

20. Where's the beef? In 1999, food scientists evaluated several treatments of bologna with the goal of assessing the quality of the bologna-type food product obtained by replacing different percentages of meat with various blends of vegetable protein (*Utilization of soy protein isolate and konjac blends in a low-fat bologna (model system)*, Meat Science, Volume 53, Issue 1, September 1999, 4557). We will consider the texture measurements of three of the treatments consisting of a fixed percentage of a konjac blend and different percentages, namely 1.1%, 2.2%, and 4.4%, of soy protein).

					$\overline{X_{i\cdot}}$	S_i^2
Treatment 1	94.1	97.9	96.2	95.8	96.0	2.4333
Treatment 2	95.6	97.6	99.8	97.8	97.7	2.9467
Treatment 3	90.2	92.1	93.8	92.3	92.1	2.1800

As the headers indicate, the marginal columns on the right report the group treatment sample means and the group treatment sample variances. What is the value of the sample statistic that is compared with the number 4 to justify the ANOVA assumption of equal variances? (The next two questions continue with this dataset.)
A) 1.1204 B) 1.3517 C) 1.5830 D) 1.8143 E) 2.0456
F) 2.2769 G) 2.5082 H) 2.7395 I) 2.9708 J) 3.2021

Solution. *The ratio of the largest treatment group sample variance to the smallest, 2.9467/2.1800, or 1.3517, is compared to 4. Because this ratio is smaller than 4, the rule of thumb is that the variances of the treatment groups are equal.*
Answer: ***B***

21. For the data introduced in the preceding problem, what is the test statistic used to test the null hypothesis that the three treatment means are all equal against the alternative hypothesis that at least two different values are found among the means?
A) 7.550 B) 8.657 C) 9.764 D) 10.871 E) 11.978
F) 13.085 G) 14.192 H) 15.299 I) 16.406 J) 17.513

Solution. *The number of treatments is $m = 3$ and the number of samples is $n = 4$ for each treatment. The overall mean is $\overline{X_{\cdot\cdot}} = (96.0+97.7+92.1)/3 = 95.2667$. The null hypothesis estimate of the population variance is*

$$MSB = S_0^2 = \frac{n}{m-1}\sum_{i=1}^{3}\left(\overline{X_{i\cdot}} - \overline{X_{\cdot\cdot}}\right)^2 = \frac{4}{2}\left((96.0-95.2667)^2 + (96.2-95.2667)^2 + (92.1-95.2667)^2\right) = 32.9733.$$

The average treatment group sample variance is

$$MSW = \overline{S^2} = \frac{1}{3}\left(S_1^2 + S_2^2 + S_3^2\right) = \frac{1}{3}\left(2.4333 + 2.9467 + 2.1800\right) = 2.52.$$

The test statistic is

$$\frac{MSB}{MSW} = \frac{S_0^2}{\overline{S^2}} = \frac{32.9733}{2.52} = 13.0846.$$

Answer: ***F***

22. For the hypothesis test described in the preceding problem, what is the p-value?
A) 0.0022 B) 0.0042 C) 0.0061 D) 0.0082 E) 0.0102
F) 0.0122 G) 0.0142 H) 0.0161 I) 0.0181 J) 0.0201

Solution. *The p-value is given by*

$$p - value = P\left(F_{m-1\,,\,m(n-1)} > \frac{MSB}{MSW}\right) = P\left(F_{2,9} > 13.0846\right) = 0.002169.$$

The p-value just given was obtained using the software package Maple—specifically, the command

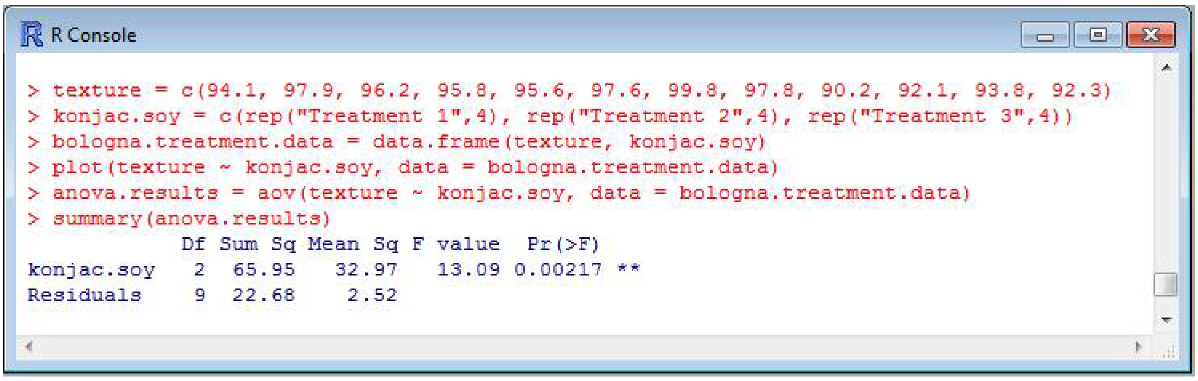
```
1-stats[statevalf,cdf,fratio[2,9]](13.0846);
```

If instead we use the given F-table, we find the points $(p, x) = (10.11, 0.005)$ *and* $(p, x) = (13.41, 0.002)$, *and we pass a line through them:*

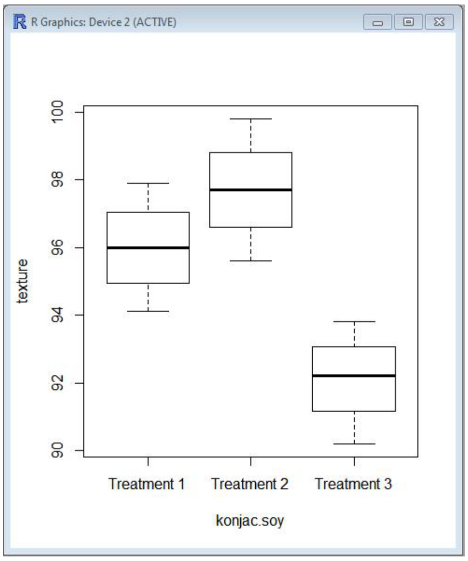
$$p = \frac{(0.002 - 0.005)}{(13.41 - 10.11)}\,(x - 13.41) + 0.002,$$

or $p = -0.00090909\,x + 0.014190909$. *For* $x = 13.0846$, *we solve* $p = (-0.00090909)(13.0846) + 0.014190909$, *or p = 0.002296. This approximation differs from correct answer choice (A) by only 0.0001. In fact, the interpolation was not even necessary for making a choice. Because the observed value 13.0846 of the test statistic is between the tabulated values* $f = 10.11$, *which corresponds to* $p = 0.005$, *and* $f = 13.41$, *which corresponds to* $p = 0.001$, *we infer that the requested p-value must lie between 0.001 and 0.005. Of the answer choices, only (A) and (B) fit the bill. Notice that choice (A), 0.0022, is about 0.001 from* $p = 0.001$ *and answer choice (B), 0.0042, is about 0.001 from* $p = 0.005$. *Because the observed value 13.0846 of the test statistic is much closer to the value 13.41, which corresponds to the p of choice (A), than to 10.11, which corresponds to the p of choice (B), there can be little doubt as to the correct choice.*

Finally, we will present a screen capture from an R session. Along with the boxplot that is shown, the code results in the displayed ANOVA table. The entry in the last column is the requested p-value.



```
R Console
> texture = c(94.1, 97.9, 96.2, 95.8, 95.6, 97.6, 99.8, 97.8, 90.2, 92.1, 93.8, 92.3)
> konjac.soy = c(rep("Treatment 1",4), rep("Treatment 2",4), rep("Treatment 3",4))
> bologna.treatment.data = data.frame(texture, konjac.soy)
> plot(texture ~ konjac.soy, data = bologna.treatment.data)
> anova.results = aov(texture ~ konjac.soy, data = bologna.treatment.data)
> summary(anova.results)
            Df Sum Sq Mean Sq F value  Pr(>F)
konjac.soy   2  65.95   32.97   13.09 0.00217 **
Residuals    9  22.68    2.52
```



Answer: ***A***

23. Beef. It's what's for dinner. Once upon a time, long ago and far away, Barbara Ryan and Brian Joiner, developers of MINITAB, a statistical software package that some people apparently still use, put on their aprons, turned on the oven, mixed up several batches of meat loaf mixtures, formed several meat bricks, recorded the weights of said bricks, and, one at a time, positioned each meat loaf in the oven and baked it. You will notice that each raw loaf was treated by the heat of the oven at the position in which the loaf was placed. In non-technical language, this treatment is often called "baking." When each culinary masterpiece was removed from the oven, the liquid drippings were weighed and each weight was divided by the weight of the raw loaf that produced the drippings. The table below shows the percentages of liquid drippings for 12 loaves: four different oven positions, and, for each position, one loaf from each of three batches of meat loaf mixture. The statistician-cooks were interested in the effects of oven position. Their experimental design controlled for the effects that different mixtures would have on drippings.

	Batch 1	Batch 2	Batch 3	$X_{i\cdot}$
Oven Position 1	5.04	5.12	7.84	18.00
Oven Position 2	3.87	6.48	7.70	18.05
Oven Position 3	6.43	5.81	8.59	20.83
Oven Position 4	4.46	4.28	6.34	15.08
$X_{\cdot j}$	19.80	21.69	30.47	71.96

Note: The marginal entries are row and column sums, not means. You may (or may not) want to know that the sum of the squares of all non-marginal table entries is 456.6336. What is the observed value of the test statistic of the null hypothesis that there is no difference in the means for the treatments against the alternative that at least one pair of different treatment means? (There are two more questions concerning this dataset. In case you are wondering where this is going, you will ultimately be asked if oven position has an effect on drippings, if the meatloaf mixture has an effect on drippings, and if so, whether or not the ANOVA results are significant or highly significant.)

A) 1.5679 B) 1.7782 C) 1.9885 D) 2.1988 E) 2.4091
F) 2.6194 G) 2.8297 H) 3.0400 I) 3.2503 J) 3.4606

Solution. *Here $m = 4$ is the number of treatments and $n = 3$ is the number of observations per treatment. We calculate* $\overline{X_{\cdot\cdot}} = \frac{1}{m\,n}\,X_{\cdot\cdot} = \frac{1}{12}\,71.96 = 5.996667$. *Therefore,*

$$
\begin{aligned}
SST &= \sum_{i=1}^{m}\sum_{j=1}^{n} X_{ij}^2 - m\,n\,\overline{X_{\cdot\cdot}}^2 \\
&= 456.6336 - 12\,(5.996667)^2 \\
&= 25.11342,
\end{aligned}
$$

$$
\begin{aligned}
SS(Tr) &= n\sum_{i=1}^{m} \overline{X_{i\cdot}}^2 - m\,n\,\overline{X_{\cdot\cdot}}^2 \\
&= 3\left(\left(\frac{18}{3}\right)^2 + \left(\frac{18.05}{3}\right)^2 + \left(\frac{20.83}{3}\right)^2 + \left(\frac{15.08}{3}\right)^2\right) - 12\,(5.996667)^2 \\
&= \frac{1}{3}\,\big((18)^2 + (18.05)^2 + (20.83)^2 + (15.08)^2\big) - 12\,(5.996667)^2 \\
&= 5.512419,
\end{aligned}
$$

$$
\begin{aligned}
SS(Bl) &= m\sum_{j=1}^{n} \overline{X_{\cdot j}}^2 - m\,n\,\overline{X_{\cdot\cdot}}^2 \\
&= 4\left(\left(\frac{19.80}{4}\right)^2 + \left(\frac{21.69}{4}\right)^2 + \left(\frac{30.47}{4}\right)^2\right) - 12\,(5.996667)^2 \\
&= \frac{1}{4}\,\big((19.80)^2 + (21.69)^2 + (30.47)^2\big) - 12\,(5.996667)^2 \\
&= 16.20907,
\end{aligned}
$$

$$
\begin{aligned}
SSE &= SST - \big(SS(Tr) + SS(Bl)\big) \\
&= 25.11342 - (5.512419 + 16.20907) \\
&= 3.391931.
\end{aligned}
$$

We next obtain the mean squares by dividing by the appropriate degrees of freedom:

$$\begin{aligned}
\mathit{MS(Tr)} &= \frac{1}{m-1}\,\mathit{SS(Tr)} \\
&= \frac{1}{3}(5.512419) \\
&= 1.837473,
\end{aligned}$$

$$\begin{aligned}
\mathit{MS(Bl)} &= \frac{1}{n-1}\,\mathit{SS(Bl)} \\
&= \frac{1}{2}(16.20907) \\
&= 8.104535,
\end{aligned}$$

$$\begin{aligned}
\mathit{MSE} &= \frac{1}{(m-1)(n-1)}\,\mathit{SSE} \\
&= \frac{1}{6}(3.391931) \\
&= 0.5653218.
\end{aligned}$$

To test whether oven positions have different effects on the mean drippings, we use the observed value for the test statistic MS(Tr)/MSE, namely 1.837473/0.5653218, or 3.250313.

This particular problem is solved, but, while we are here, let us estimate the p-value. We have, p-value = $P(F_{3,6} \geq 3.250313) > P(F_{3,6} \geq 3.29) = 0.100$, *obtained directly from the given F table. This tells us that the null hypothesis is* not *rejected.*
Answer: ***I***

24. For the test of the null hypothesis that the meat loaf mixture does not affect mean drippings against the alternative hypothesis that the mixture does have an effect, what is the p-value?

A) 0.0010	B) 0.0031	C) 0.0052	D) 0.0073	E) 0.0144
F) 0.0236	G) 0.0328	H) 0.0420	I) 0.0512	J) 0.0604

Solution. *Most of the work has been done. The observed value of the test statistic is MS(Bl)/MSE, or 8.104535/0.5653218, or 14.33614. The p-value is* $P(F_{2,6} \geq 14.33614) = 0.005182109$. *The value just obtained was via the R command* `pf(14.33614,df1=2,df2=6,lower.tail=FALSE)`. *Using th provided F table leads to the answer even more easily. The table shows that* $P(F_{2,6} \geq 14.54) = 0.005$, *and, given the answer choices, answer C is the only reasonable candidate. The p-value is less than 0.01, which means that the rejection of the null hypothesis is highly significant.*
Answer: ***C***

25. The ANOVA calculations of the preceding two problems allow us to test two hypotheses. For one test the null hypothesis is that oven position does not affect drippings and, for the other test, the null hypothesis is that the meat loaf mixture does not affect drippings. For each test, there are three possible outcomes: a highly significant rejection of the null hypothesis, coded by $**$, a significant rejection of the null hypothesis, coded by $*$, and retention of the null hypothesis, coded by a blank space. Answer with the column letter that reports the results of our hypotheses tests.

	A	B	C	D	E	F	G	H	I
Oven Position Test				$*$	$*$	$*$	$**$	$**$	$**$
Mixture Test		$*$	$**$		$*$	$**$		$*$	$**$

Solution. *This question has been answered by our work in the previous two problems. We use the space saved to insert an ANOVA table.*

Variance Source	*Sum of Squares*	*Degrees of Freedom*	*Mean Square*	*Observed F-statistic*	*Tail Area Above F*
Treatments	*5.512419*	3	*1.837473*	*3.250313*	*0.0999*
Blocks	*16.20907*	2	*8.104535*	*14.33614*	*0.0052*
Error	*3.391931*	6	*0.5653218*		
Total	*25.11342*	11			

ANOVA Table for Beef Drippings (With Mixture Blocks)

As a check of the answers to Problems 23–25, the analysis of variance command of R was used. The start was this table

Observation	*OvenPosition*	*Batch*	*Drippings*
1	*ov1*	*b1*	*5.04*
2	*ov1*	*b2*	*5.12*
3	*ov1*	*b3*	*7.84*
4	*ov2*	*b1*	*3.87*
5	*ov2*	*b2*	*6.48*
6	*ov2*	*b3*	*7.70*
7	*ov3*	*b1*	*6.43*
8	*ov3*	*b2*	*5.81*
9	*ov3*	*b3*	*8.59*
10	*ov4*	*b1*	*4.46*
11	*ov4*	*b2*	*4.28*
12	*ov4*	*b3*	*6.34*

saved as the file meatLoafData.txt in the `data` *subfolder of a subfolder* `stats` *of a* `C` *directory. The following screen capture shows how the filre was read into an R session, and the commands used to produce an ANOVA table reflecting the mixture blocks.*

```
> meat.loaf.data="C:/stats/data/meatLoafData.txt"
> meatLoaf = read.table(meat.loaf.data,header=TRUE)
> meatLoaf
   Observation OvenPosition Batch Drippings
1            1          ov1    b1      5.04
2            2          ov1    b2      5.12
3            3          ov1    b3      7.84
4            4          ov2    b1      3.87
5            5          ov2    b2      6.48
6            6          ov2    b3      7.70
7            7          ov3    b1      6.43
8            8          ov3    b2      5.81
9            9          ov3    b3      8.59
10          10          ov4    b1      4.46
11          11          ov4    b2      4.28
12          12          ov4    b3      6.34
> meatLoaf.aov = aov(Drippings ~ OvenPosition+Batch,data = meatLoaf)
> summary(meatLoaf.aov)
             Df Sum Sq Mean Sq F value  Pr(>F)
OvenPosition  3  5.512   1.837    3.25 0.10206
Batch         2 16.209   8.105   14.34 0.00518 **
Residuals     6  3.392   0.565
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
>
```

Meat Loaf ANOVA Drippings Table

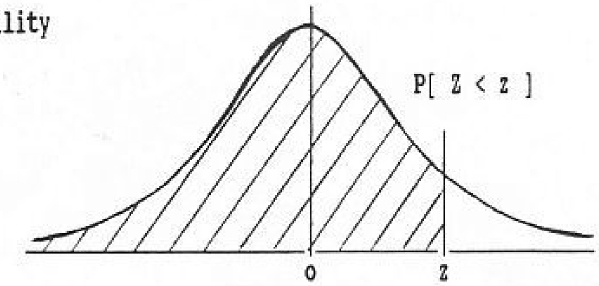
Answer: ***C***

STANDARD STATISTICAL TABLES

1. Areas under the Normal Distribution

The table gives the cumulative probability up to the standardised normal value z
i.e.

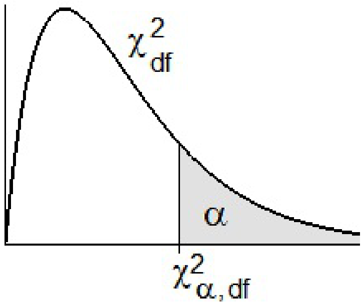
$$P[\ Z < z\ ] = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \exp(-\tfrac{1}{2}Z^2)\ dZ$$



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	0.7020	0.7054	0.7089	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	0.8804	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	0.9865	0.9868	0.9871	0.9874	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	0.9924	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	0.9980	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

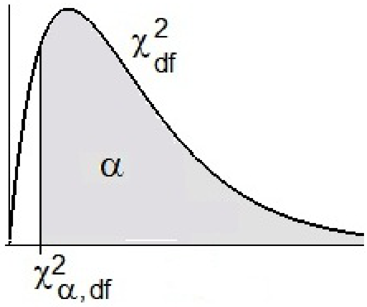
z	3.00	3.10	3.20	3.30	3.40	3.50	3.60	3.70	3.80	3.90
P	0.9986	0.9990	0.9993	0.9995	0.9997	0.9998	0.9998	0.9999	0.9999	1.0000

Values of $\chi^2_{\alpha,\mathrm{df}}$ $\quad P(\chi^2_{\mathrm{df}} \geqslant \chi^2_{\alpha,\mathrm{df}}) = \alpha$



df \ α	0.005	0.010	0.025	0.050	0.100	0.200	0.250	0.300	0.400	0.500
1	7.8794	6.6349	5.0239	3.8415	2.7055	1.6424	1.3233	1.0742	0.7083	0.4549
2	10.5966	9.2103	7.3778	5.9915	4.6052	3.2189	2.7726	2.4079	1.8326	1.3863
3	12.8382	11.3449	9.3484	7.8147	6.2514	4.6416	4.1083	3.6649	2.9462	2.3660
4	14.8603	13.2767	11.1433	9.4877	7.7794	5.9886	5.3853	4.8784	4.0446	3.3567
5	16.7496	15.0863	12.8325	11.0705	9.2364	7.2893	6.6257	6.0644	5.1319	4.3515
6	18.5476	16.8119	14.4494	12.5916	10.6446	8.5581	7.8408	7.2311	6.2108	5.3481
7	20.2777	18.4753	16.0128	14.0671	12.0170	9.8032	9.0371	8.3834	7.2832	6.3458
8	21.9550	20.0902	17.5345	15.5073	13.3616	11.0301	10.2189	9.5245	8.3505	7.3441
9	23.5894	21.6660	19.0228	16.9190	14.6837	12.2421	11.3888	10.6564	9.4136	8.3428
10	25.1882	23.2093	20.4832	18.3070	15.9872	13.4420	12.5489	11.7807	10.4732	9.3418
11	26.7568	24.7250	21.9200	19.6751	17.2750	14.6314	13.7007	12.8987	11.5298	10.3410
12	28.2995	26.2170	23.3367	21.0261	18.5493	15.8120	14.8454	14.0111	12.5838	11.3403
13	29.8195	27.6882	24.7356	22.3620	19.8119	16.9848	15.9839	15.1187	13.6356	12.3398
14	31.3193	29.1412	26.1189	23.6848	21.0641	18.1508	17.1169	16.2221	14.6853	13.3393
15	32.8013	30.5779	27.4884	24.9958	22.3071	19.3107	18.2451	17.3217	15.7332	14.3389
16	34.2672	31.9999	28.8454	26.2962	23.5418	20.4651	19.3689	18.4179	16.7795	15.3385
17	35.7185	33.4087	30.1910	27.5871	24.7690	21.6146	20.4887	19.5110	17.8244	16.3382
18	37.1565	34.8053	31.5264	28.8693	25.9894	22.7595	21.6049	20.6014	18.8679	17.3379
19	38.5823	36.1909	32.8523	30.1435	27.2036	23.9004	22.7178	21.6891	19.9102	18.3377
20	39.9968	37.5662	34.1696	31.4104	28.4120	25.0375	23.8277	22.7745	20.9514	19.3374
21	41.4011	38.9322	35.4789	32.6706	29.6151	26.1711	24.9348	23.8578	21.9915	20.3372
22	42.7957	40.2894	36.7807	33.9244	30.8133	27.3015	26.0393	24.9390	23.0307	21.3370
23	44.1813	41.6384	38.0756	35.1725	32.0069	28.4288	27.1413	26.0184	24.0689	22.3369
24	45.5585	42.9798	39.3641	36.4150	33.1962	29.5533	28.2412	27.0960	25.1063	23.3367
25	46.9279	44.3141	40.6465	37.6525	34.3816	30.6752	29.3389	28.1719	26.1430	24.3366
30	53.6720	50.8922	46.9792	43.7730	40.2560	36.2502	34.7997	33.5302	31.3159	29.3360
40	66.7660	63.6907	59.3417	55.7585	51.8051	47.2685	45.6160	44.1649	41.6222	39.3353
50	79.4900	76.1539	71.4202	67.5048	63.1671	58.1638	56.3336	54.7228	51.8916	49.3349
60	91.9517	88.3794	83.2977	79.0819	74.3970	68.9721	66.9815	65.2265	62.1348	59.3347
70	104.2149	100.4252	95.0232	90.5312	85.5270	79.7146	77.5767	75.6893	72.3583	69.3345
80	116.3211	112.3288	106.6286	101.8795	96.5782	90.4053	88.1303	86.1197	82.5663	79.3343
90	128.2989	124.1163	118.1359	113.1453	107.5650	101.0537	98.6499	96.5238	92.7614	89.3342
100	140.1695	135.8067	129.5612	124.3421	118.4980	111.6667	109.1412	106.9058	102.9459	99.3341

Chi-Squared Values—Right Tails.



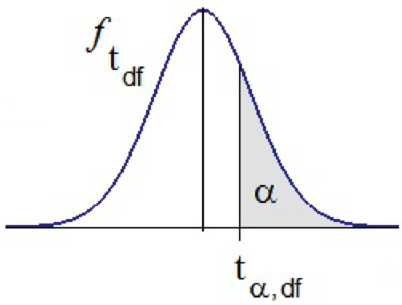
Values of $\chi^2_{\alpha,df}$ $\quad P(\chi^2_{df} \geq \chi^2_{\alpha,df}) = \alpha$

df \ α	0.600	0.700	0.750	0.800	0.900	0.950	0.975	0.990	0.995
1	0.2750	0.1485	0.1015	0.0642	0.0158	0.0039	0.0010	0.0002	0.0000
2	1.0217	0.7133	0.5754	0.4463	0.2107	0.1026	0.0506	0.0201	0.0100
3	1.8692	1.4237	1.2125	1.0052	0.5844	0.3518	0.2158	0.1148	0.0717
4	2.7528	2.1947	1.9226	1.6488	1.0636	0.7107	0.4844	0.2971	0.2070
5	3.6555	2.9999	2.6746	2.3425	1.6103	1.1455	0.8312	0.5543	0.4117
6	4.5702	3.8276	3.4546	3.0701	2.2041	1.6354	1.2373	0.8721	0.6757
7	5.4932	4.6713	4.2549	3.8223	2.8331	2.1673	1.6899	1.2390	0.9893
8	6.4226	5.5274	5.0706	4.5936	3.4895	2.7326	2.1797	1.6465	1.3444
9	7.3570	6.3933	5.8988	5.3801	4.1682	3.3251	2.7004	2.0879	1.7349
10	8.2955	7.2672	6.7372	6.1791	4.8652	3.9403	3.2470	2.5582	2.1559
11	9.2373	8.1479	7.5841	6.9887	5.5778	4.5748	3.8157	3.0535	2.6032
12	10.1820	9.0343	8.4384	7.8073	6.3038	5.2260	4.4038	3.5706	3.0738
13	11.1291	9.9257	9.2991	8.6339	7.0415	5.8919	5.0088	4.1069	3.5650
14	12.0785	10.8215	10.1653	9.4673	7.7895	6.5706	5.6287	4.6604	4.0747
15	13.0297	11.7212	11.0365	10.3070	8.5468	7.2609	6.2621	5.2293	4.6009
16	13.9827	12.6243	11.9122	11.1521	9.3122	7.9616	6.9077	5.8122	5.1422
17	14.9373	13.5307	12.7919	12.0023	10.0852	8.6718	7.5642	6.4078	5.6972
18	15.8932	14.4399	13.6753	12.8570	10.8649	9.3905	8.2307	7.0149	6.2648
19	16.8504	15.3517	14.5620	13.7158	11.6509	10.1170	8.9065	7.6327	6.8440
20	17.8088	16.2659	15.4518	14.5784	12.4426	10.8508	9.5908	8.2604	7.4338
21	18.7683	17.1823	16.3444	15.4446	13.2396	11.5913	10.2829	8.8972	8.0337
22	19.7288	18.1007	17.2396	16.3140	14.0415	12.3380	10.9823	9.5425	8.6427
23	20.6902	19.0211	18.1373	17.1865	14.8480	13.0905	11.6886	10.1957	9.2604
24	21.6525	19.9432	19.0373	18.0618	15.6587	13.8484	12.4012	10.8564	9.8862
25	22.6156	20.8670	19.9393	18.9398	16.4734	14.6114	13.1197	11.5240	10.5197
30	27.4416	25.5078	24.4776	23.3641	20.5992	18.4927	16.7908	14.9535	13.7867
40	37.1340	34.8719	33.6603	32.3450	29.0505	26.5093	24.4330	22.1643	20.7065
50	46.8638	44.3133	42.9421	41.4492	37.6886	34.7643	32.3574	29.7067	27.9907
60	56.6200	53.8091	52.2938	50.6406	46.4589	43.1880	40.4817	37.4849	35.5345
70	66.3961	63.3460	61.6983	59.8978	55.3289	51.7393	48.7576	45.4417	43.2752
80	76.1879	72.9153	71.1445	69.2069	64.2778	60.3915	57.1532	53.5401	51.1719
90	85.9925	82.5111	80.6247	78.5584	73.2911	69.1260	65.6466	61.7541	59.1963
100	95.8078	92.1289	90.1332	87.9453	82.3581	77.9295	74.2219	70.0649	67.3276

Chi-Squared Values—Central Hump + Right Tails.

Values of $t_{\alpha,df}$

$P(t_{df} \geq t_{\alpha,df}) = \alpha$

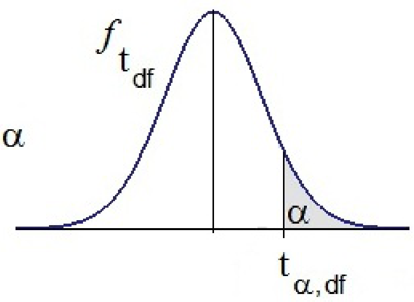


df \ α	.450	.400	.350	.300	.250	.200
1	.1584	.3249	.5095	.7265	1.0000	1.3764
2	.1421	.2887	.4447	.6172	.8165	1.0607
3	.1366	.2767	.4242	.5844	.7649	.9785
4	.1338	.2707	.4142	.5686	.7407	.9410
5	.1322	.2672	.4082	.5594	.7267	.9195
6	.1311	.2648	.4043	.5534	.7176	.9057
7	.1303	.2632	.4015	.5491	.7111	.8960
8	.1297	.2619	.3995	.5459	.7064	.8889
9	.1293	.2610	.3979	.5435	.7027	.8834
10	.1289	.2602	.3966	.5415	.6998	.8791
11	.1286	.2596	.3956	.5399	.6974	.8755
12	.1283	.2590	.3947	.5386	.6955	.8726
13	.1281	.2586	.3940	.5375	.6938	.8702
14	.1280	.2582	.3933	.5366	.6924	.8681
15	.1278	.2579	.3928	.5357	.6912	.8662
16	.1277	.2576	.3923	.5350	.6901	.8647
17	.1276	.2573	.3919	.5344	.6892	.8633
18	.1274	.2571	.3915	.5338	.6884	.8620
19	.1274	.2569	.3912	.5333	.6876	.8610
20	.1273	.2567	.3909	.5329	.6870	.8600
21	.1272	.2566	.3906	.5325	.6864	.8591
22	.1271	.2564	.3904	.5321	.6858	.8583
23	.1271	.2563	.3902	.5317	.6853	.8575
24	.1270	.2562	.3900	.5314	.6848	.8569
25	.1269	.2561	.3898	.5312	.6844	.8562
26	.1269	.2560	.3896	.5309	.6840	.8557
27	.1268	.2559	.3894	.5306	.6837	.8551
28	.1268	.2558	.3893	.5304	.6834	.8546
29	.1268	.2557	.3892	.5302	.6830	.8542
30	.1267	.2556	.3890	.5300	.6828	.8538
40	.1265	.2550	.3881	.5286	.6807	.8507
50	.1263	.2547	.3875	.5278	.6794	.8489
60	.1262	.2545	.3872	.5272	.6786	.8477
70	.1261	.2543	.3869	.5268	.6780	.8468
80	.1261	.2542	.3867	.5265	.6776	.8461
90	.1260	.2541	.3866	.5263	.6772	.8456
100	.1260	.2540	.3864	.5261	.6770	.8452

Student-t Values—Right Tails $\alpha = 0.45, 0.40, 0.35, 0.30, 0.25, 0.20.$

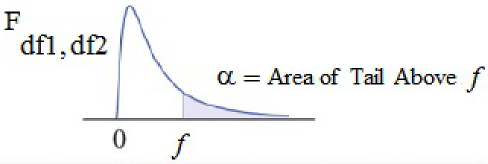
Values of $t_{\alpha,df}$

$P(t_{df} \geqslant t_{\alpha,df}) = \alpha$



df \ α	.150	.100	.050	.025	.010	.005
1	1.9626	3.0777	6.3138	12.7062	31.8205	63.6567
2	1.3862	1.8856	2.9200	4.3027	6.9646	9.9248
3	1.2498	1.6377	2.3534	3.1824	4.5407	5.8409
4	1.1896	1.5332	2.1318	2.7764	3.7469	4.6041
5	1.1558	1.4759	2.0150	2.5706	3.3649	4.0321
6	1.1342	1.4398	1.9432	2.4469	3.1427	3.7074
7	1.1192	1.4149	1.8946	2.3646	2.9980	3.4995
8	1.1081	1.3968	1.8595	2.3060	2.8965	3.3554
9	1.0997	1.3830	1.8331	2.2622	2.8214	3.2498
10	1.0931	1.3722	1.8125	2.2281	2.7638	3.1693
11	1.0877	1.3634	1.7959	2.2010	2.7181	3.1058
12	1.0832	1.3562	1.7823	2.1788	2.6810	3.0545
13	1.0795	1.3502	1.7709	2.1604	2.6503	3.0123
14	1.0763	1.3450	1.7613	2.1448	2.6245	2.9768
15	1.0735	1.3406	1.7531	2.1314	2.6025	2.9467
16	1.0711	1.3368	1.7459	2.1199	2.5835	2.9208
17	1.0690	1.3334	1.7396	2.1098	2.5669	2.8982
18	1.0672	1.3304	1.7341	2.1009	2.5524	2.8784
19	1.0655	1.3277	1.7291	2.0930	2.5395	2.8609
20	1.0640	1.3253	1.7247	2.0860	2.5280	2.8453
21	1.0627	1.3232	1.7207	2.0796	2.5176	2.8314
22	1.0614	1.3212	1.7171	2.0739	2.5083	2.8188
23	1.0603	1.3195	1.7139	2.0687	2.4999	2.8073
24	1.0593	1.3178	1.7109	2.0639	2.4922	2.7969
25	1.0584	1.3163	1.7081	2.0595	2.4851	2.7874
26	1.0575	1.3150	1.7056	2.0555	2.4786	2.7787
27	1.0567	1.3137	1.7033	2.0518	2.4727	2.7707
28	1.0560	1.3125	1.7011	2.0484	2.4671	2.7633
29	1.0553	1.3114	1.6991	2.0452	2.4620	2.7564
30	1.0547	1.3104	1.6973	2.0423	2.4573	2.7500
40	1.0500	1.3031	1.6839	2.0211	2.4233	2.7045
50	1.0473	1.2987	1.6759	2.0086	2.4033	2.6778
60	1.0455	1.2958	1.6706	2.0003	2.3901	2.6603
70	1.0442	1.2938	1.6669	1.9944	2.3808	2.6479
80	1.0432	1.2922	1.6641	1.9901	2.3739	2.6387
90	1.0424	1.2910	1.6620	1.9867	2.3685	2.6316
100	1.0418	1.2901	1.6602	1.9840	2.3642	2.6259

Student-t Values—Right Tails $\alpha = 0.15, 0.10, 0.05, 0.025, 0.010, 0.005$.



Values of *F*-Distributions

α	df1	df2 2	3	4	6	8	9	10	12	15	16	18	20
.100	2	9.00	5.46	4.32	3.46	3.11	3.01	2.92	2.81	2.70	2.67	2.62	2.59
.050	2	19.00	9.55	6.94	5.14	4.46	4.26	4.10	3.89	3.68	3.63	3.55	3.49
.025	2	39.00	16.04	10.65	7.26	6.06	5.71	5.46	5.10	4.77	4.69	4.56	4.46
.010	2	99.00	30.82	18.00	10.92	8.65	8.02	7.56	6.93	6.36	6.23	6.01	5.85
.005	2	199.00	49.80	26.28	14.54	11.04	10.11	9.43	8.51	7.70	7.51	7.21	6.99
.002	2	499.00	92.99	42.72	20.81	14.91	13.41	12.33	10.90	9.68	9.40	8.95	8.62
.001	2	999.00	148.50	61.25	27.00	18.49	16.39	14.91	12.97	11.34	10.97	10.39	9.95
.100	3	9.16	5.39	4.19	3.29	2.92	2.81	2.73	2.61	2.49	2.46	2.42	2.38
.050	3	19.16	9.28	6.59	4.76	4.07	3.86	3.71	3.49	3.29	3.24	3.16	3.10
.025	3	39.17	15.44	9.98	6.60	5.42	5.08	4.83	4.47	4.15	4.08	3.95	3.86
.010	3	99.17	29.46	16.69	9.78	7.59	6.99	6.55	5.95	5.42	5.29	5.09	4.94
.005	3	199.17	47.47	24.26	12.92	9.60	8.72	8.08	7.23	6.48	6.30	6.03	5.82
.002	3	499.17	88.45	39.27	18.34	12.84	11.44	10.45	9.15	8.03	7.78	7.38	7.07
.001	3	999.17	141.11	56.18	23.70	15.83	13.90	12.55	10.80	9.34	9.01	8.49	8.10
.100	4	9.24	5.34	4.11	3.18	2.81	2.69	2.61	2.48	2.36	2.33	2.29	2.25
.050	4	19.25	9.12	6.39	4.53	3.84	3.63	3.48	3.26	3.06	3.01	2.93	2.87
.025	4	39.25	15.10	9.60	6.23	5.05	4.72	4.47	4.12	3.80	3.73	3.61	3.51
.010	4	99.25	28.71	15.98	9.15	7.01	6.42	5.99	5.41	4.89	4.77	4.58	4.43
.005	4	199.25	46.19	23.15	12.03	8.81	7.96	7.34	6.52	5.80	5.64	5.37	5.17
.002	4	499.25	85.98	37.39	17.01	11.71	10.38	9.43	8.19	7.14	6.90	6.52	6.23
.001	4	999.25	137.10	53.44	21.92	14.39	12.56	11.28	9.63	8.25	7.94	7.46	7.10
.100	5	9.29	5.31	4.05	3.11	2.73	2.61	2.52	2.39	2.27	2.24	2.20	2.16
.050	5	19.30	9.01	6.26	4.39	3.69	3.48	3.33	3.11	2.90	2.85	2.77	2.71
.025	5	39.30	14.88	9.36	5.99	4.82	4.48	4.24	3.89	3.58	3.50	3.38	3.29
.010	5	99.30	28.24	15.52	8.75	6.63	6.06	5.64	5.06	4.56	4.44	4.25	4.10
.005	5	199.30	45.39	22.46	11.46	8.30	7.47	6.87	6.07	5.37	5.21	4.96	4.76
.002	5	499.30	84.42	36.21	16.16	11.00	9.70	8.79	7.59	6.57	6.34	5.97	5.70
.001	5	999.30	134.58	51.71	20.80	13.48	11.71	10.48	8.89	7.57	7.27	6.81	6.46

F Values—Right Tails $\alpha = 0.100, 0.0.50, 0.025, 0.010, 0.005, 0.002, 0.001$ and Selected Degrees of Freedom