Math 2200 Spring 2017, Final Exam

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

1. Let X be the weight in mg of an adult female housefly (Musca domestica, L.) and let Y be the weight in mg of an adult male housefly. Let μ_{X} and μ_{Y} denote the unknown means of X and Y. Do not assume that X and Y are normally distributed. In a study of the effect of a sublethal dose of sodium arsenite on adult houseflies, 457 females and 402 males were weighed prior to treatment. The sample statistics were $\overline{\mathrm{X}} = 27.650$, $S_{\mathrm{X}} = 2.530$, $\overline{\mathrm{Y}} = 18.075$, and $S_{\mathrm{Y}} = 1.997$ (GAINES, J. C.,CLARE, S., RICHARDSON, C. H., Journal of Economic Entomology 1937 Vol.30 No.2 pp.363-366). Find a 99% confidence interval for $\mu_{\mathrm{X}} - \mu_{\mathrm{Y}}$. What is the lower limit of the interval?

   A) 8.7701 B) 8.8514 C) 8.9327 D) 9.0140 E) 9.0953
   F) 9.1766 G) 9.2579 H) 9.3392 I) 9.4205 J) 9.5018

Solution **F) 9.1766**

Because there is no assumption of normality, the student-t distribution cannot be used. However, the large sizes, n = 457 and m = 402, of both samples allow us to use normal approximations. Moreover, the large samples allow us to approximate the unknown population means with the sample means. Thus, the margin of error ME is

$$\mathrm{ME} = z_{0.005}\sqrt{\frac{S_{\mathrm{X}}^2}{n} + \frac{S_{\mathrm{Y}}^2}{m}} = 2.575829\sqrt{\frac{(2.530)^2}{457} + \frac{(1.997)^2}{402}} = 0.3984364.$$

The 99% confidence interval is

$$\overline{\mathrm{X}} - \overline{\mathrm{Y}} \pm \mathrm{ME} = 27.650 - 18.075 \pm 0.3984364 = 9.575 \pm 0.3984364,$$

or $[9.176564, 9.973436]$.

Here is the R code for the answer:

```
> n.X = 457; n.Y = 402
> m.X = 27.650; m.Y = 18.075
> S.X = 2.530; S.Y = 1.997
> SE = sqrt(S.X^2/n.X + S.Y^2/n.Y)
> SE
[1] 0.1546828
> z.99 = qnorm(0.995)
> z.99
[1] 2.575829
> ME = z.99*SE
> ME
[1] 0.3984365
> conf.int.99 = c(m.X - m.Y - ME , m.X - m.Y + ME)
> conf.int.99
[1] 9.176564 9.973436
```

2. As in the preceding problem, let X be the weight in mg of an adult female housefly and let Y be the weight in mg of an adult male housefly. Let μ_X and μ_Y denote the unknown means of X and Y. In this problem, *do* assume that X and Y are normally distributed. Suppose that a hypothetical random sample of 17 females and 12 males had resulted in the same sample statistics that were actually observed and reported in the preceding problem: $\overline{X} = 27.650$, $S_X = 2.530$, $\overline{Y} = 18.075$, and $S_Y = 1.997$. Using the hypothetical sample statistics, what would be the margin of error for a 99% confidence interval for $\mu_X - \mu_Y$? Follow all course conventions. Do not pool. In fact, the authors of the article noted that the variance of female weights could be expected to be larger than that of males because of differing egg developments.

   A) 0.4216 B) 0.6653 C) 0.9090 D) 1.1527 E) 1.3964
   F) 1.6401 G) 1.8838 H) 2.1275 I) 2.3712 J) 2.6149

Solution **J) 2.6149**

In this problem the sample sizes, $n = 17$ and $m = 12$, are too small to use the normal approximation. Moreover, the sizes are too small to expect that the sample variances are good estimates of the unknown population variances. In short, the method of the preceding problem is not appropriate. However, the assumption of normality of X and Y allows us to use a student-t distribution. According to course conventions we will use $\mathrm{df} = \min(n-1, m-1) = 11$ as a conservative choice for the degrees of freedom. Thus

$$\mathrm{ME} = t_{0.005,11}\sqrt{\frac{S_X^2}{n} + \frac{S_Y^2}{m}} = 3.105807\sqrt{\frac{(2.530)^2}{17} + \frac{(1.997)^2}{12}} = 2.614893.$$

Here is the R code for the answer:

```
> n.X = 17; n.Y = 12
> S.X = 2.530; S.Y = 1.997
> alpha = 0.01; d.f. = min(17,12)-1
> t.99 = qt(alpha/2, df = d.f., lower.tail=FALSE)
> t.99
[1] 3.105807
> SE = sqrt(S.X^2/n.X + S.Y^2/n.Y)
> SE
[1] 0.8419368
> ME = t.99*SE
> ME
[1] 2.614893
```

3. The political landscape was different when your examiner arrived at First President University. Let us travel back in time to January 1980, just before the inauguration of Ronald Reagan, and to April 1981, one year into the Reagan presidency. At both those historical moments, New York Times–CBS News Polls asked 1400 randomly sampled Americans if they identified themselves as Democrats or Republicans. (The size of each poll was 1400 and the two polls were independent.) In January 1980, only 462 of the 1400 surveyees identified themselves as either Republicans or Republican-leaning independents. In April 1981, the number of such surveyees had grown but was still only 574 of 1400. Let p_{1981} be the population proportion of Republicans or Republican-leaning independents in April 1981. Let p_{1980} be the corresponding proportion in January 1980. What is the upper limit of a 99% confidence interval for $p_{1981} - p_{1980}$?

   A) 0.1021 B) 0.1144 C) 0.1268 D) 0.1392 E) 0.1515
   F) 0.1639 G) 0.1763 H) 0.1887 I) 0.2010 J) 0.2134

Solution C) 0.1268

Let $n = 1400$ denote the sample size in 1981 and $m = 1400$ the sample size in 1980. The standard error of $\widehat{p_{1981}} - \widehat{p_{1980}}$ is

$$\begin{aligned} \text{SE} &= \sqrt{\frac{p_{1981}\,(1-p_{1981})}{n} + \frac{p_{1980}\,(1-p_{1980})}{m}} \\ &= \sqrt{\frac{574\,(1400-574)}{1400^3} + \frac{462\,(1400-462)}{1400^3}} \\ &= 0.01818555. \end{aligned}$$

The z-multiplier $z_{0.005}$ is 2.575829. The margin of error ME is ME $= z_{0.005}\,\text{SE} = 2.575829 \times 0.01818555 = 0.04684287$. The upper limit of a 99% confidence interval is

$$\frac{574}{1400} - \frac{462}{1400} + 0.04684287, \quad \text{or} \quad 0.1268429.$$

Here is the R code for the answer:

```
> n.1981 = 1400; n.1980 = 1400
> p.1981 = 574/1400; p.1980 = 462/1400
> SE = sqrt( p.1981*(1-p.1981)/n.1981 + p.1980*(1-p.1980)/n.1980)  # standard error
> SE
[1] 0.01818555
> z.99 = qnorm(0.995)                                               # The z-multiplier
> z.99
[1] 2.575829
> ME.99 = z.99*SE                                                   # margin of error
> conf.int.99 = c(p.1981-p.1980 - ME.99 , p.1981-p.1980 + ME.99)    # 99\% conf. int.
> conf.int.99
[1] 0.03315712 0.12684288
```

4. The amperage drawn by a circular saw is normally distributed with mean μ and standard deviation σ. Find the lower endpoint of a 95% confidence interval for the variance σ^2 based on the following observations of a random sample of amperage drawn: 5.6384, 6.2152, 5.5988, 6.3004, 5.6364, 5.4509. (The next problem will use some of the calculations needed for this problem.)

   A) 0.0497 B) 0.0600 C) 0.0703 D) 0.0806 E) 0.0909
   F) 0.1012 G) 0.1115 H) 0.1218 I) 0.1321 J) 0.1424

Solution A) 0.0497

If X is normally distributed with variance σ^2, and if S^2 is the sample variance of a random sample of size n drawn from X, then a $100\,(1-\alpha)\%$ confidence interval for σ^2 is

$$\left[\frac{(n-1)\,S^2}{\chi^2_{\alpha/2,n-1}}, \frac{(n-1)\,S^2}{\chi^2_{1-\alpha/2,n-1}}\right]$$

Here we calculate $S^2 = 0.127551$ and look up $\chi^2_{0.05/2,6-1} = \chi^2_{0.025,5} = 12.8325$ and $\chi^2_{1-0.05/2,6-1} = \chi^2_{0.975,5} = 0.8312116$. The requested confidence interval is

$$\left[\frac{5 \times 0.127551}{12.8325}, \frac{5 \times 0.127551}{0.8312116}\right] = [0.04969842, 0.76725950].$$

Here is the R code for the answer:

```
> X = c(5.6384, 6.2152, 5.5988, 6.3004, 5.6364, 5.4509)
> S = sd(X)
> S                                                     # the sample standard deviation
[1] 0.3571428
> S^2                                                   # the sample variance
[1] 0.127551
> chisq.divisor.left = qchisq(0.025, df = 5)
> chisq.divisor.left
[1] 0.8312116
> chisq.divisor.right = qchisq(0.975, df = 5)
> chisq.divisor.right
[1] 12.8325
> conf.int.95 = c( (6-1)*S^2/chisq.divisor.right , (6-1)*S^2/chisq.divisor.left)
> conf.int.95
[1] 0.04969841 0.76725945
```

5. What is the upper endpoint of the 95% confidence interval described in the preceding problem?

   A) 0.1954 B) 0.2771 C) 0.3588 D) 0.4405 E) 0.5222
   F) 0.6039 G) 0.6856 H) 0.7673 I) 0.8490 J) 0.9307

Solution **H) 0.7673**

See preceding solution.

6. Standing and supine systolic blood pressures were compared for 5 subjects. The blood pressure of each subject was measured in both positions.

Subject	1	2	3	4	5
Standing	162	128	139	145	151
Supine	160	137	138	149	158

   Test the null hypothesis that standing and supine systolic blood pressure have the same mean against the alternative that mean supine systolic blood pressure is greater than mean standing systolic blood pressure. Assume that the difference between supine and standing blood pressure is normally distributed. Report the p-value.

   A) 0.0193 B) 0.0446 C) 0.0699 D) 0.0952 E) 0.1205
   F) 0.1458 G) 0.1711 H) 0.1964 I) 0.2217 J) 0.2470

Solution **D) 0.0952**

Let X be supine systolic blood pressure and let Y be standing systolic blood pressure. The paired data results in the following differences $X_i - Y_i$ for $1 \le i \le 5$: $-2, 9, -1, 4, 7$. The sample mean of these differences is 3.4 and the sample standard deviation is 4.827007. The p-value is

$$P\left(t_{5-1} \ge \frac{3.4}{4.827007/\sqrt{5}}\right) = P\left(t_4 \ge 1.57502\right) = 0.09518603.$$

Here is the R code for the answer:

```
> X = c(160,137,138,149,158); Y = c(162,128,139,145,151); X.minus.Y = X - Y
> X.minus.Y
[1] -2  9 -1  4  7
> n = length(X.minus.Y)
> mu.obs = mean(X.minus.Y)
> mu.obs
[1] 3.4
> S.obs = sd(X.minus.Y)
> S.obs
[1] 4.827007
> mu.obs/(S.obs/sqrt(n))
[1] 1.57502
> pt(mu.obs/(S.obs/sqrt(n)), df = n-1, lower.tail = FALSE)
[1] 0.09518608
```

7. "In the movies and in certain kinds of romantic literature, we sometimes come across a deathbed scene in which a dying person holds onto life until some special event has occurred. For example, a mother might stave off death until her long-absent son returns from the wars. Do such feats of will occur in real life?"

   Those sentences were the introduction to a paper that statistician David Phillips published in 1972. He studied the question, Do some people postpone their deaths until after their birthdays? Phillips randomly selected 1251 deceased persons from a datbase of prominent Americans. For each such person, Phillips observed his or her birth and death months. He then assigned the person to one of three groups we will refer to as groups A, B, and C. If a person died in the month prior to his or her birth month, then Phillips placed that person in group A. If the person died in either his or her birth month or one of the next three months, then Phillips placed that person in group B. A person who was not placed in either group A or group B was placed in group C. The observed counts are tabulated below. Phillips asked, Does the observed distribution of the counts fit the distribution in which deaths are equally distributed by month (i.e., 1/12 of all deaths occur in any specified month, 2/12 of all deaths occur in any specified two months, 3/12 of all deaths occur in any specified 3 months, and so on)? We will test the null hypothesis that the observed distribution of deaths, as tabulated below, fits the proposed distribution against the alternative that the observed distribution does not fit.

Group	A	B	C	Total
Observed Count	86	472	693	1251

   If the null hypothesis were true, what would be the expected count for group C? (The next problem continues with this hypothesis test.)

   A) 681.75   B) 687.75   C) 693.75   D) 699.75   E) 705.75
   F) 711.75   G) 717.75   H) 723.75   I) 729.75   J) 735.75

Solution I) 729.75

Group A counts deaths in one month. That would be 1/12 of all deaths in the proposed distribution. Group B counts deaths in 4 months, or 4/12 of all deaths in the proposed distribution. Group C counts deaths in the remaining 7 months, or 7/12 of all deaths in the proposed distribution. The expected counts are obtained by multiplying 1251 by each of the proportions 1/12, 4/12, 7/12. The results of these multiplications are tabulated below.

Group	A	B	C	Total
Observed Count	86	472	693	1251
Expected	104.25	417.00	729.75	1251

8. To continue the death month hypothesis test of the preceding problem, calculate the test statistic and the critical value using 0.01 as the significance level. By how much does the test statistic exceed the critical value?

   A) 2.7273 B) 2.8480 C) 2.9687 D) 3.0894 E) 3.2101
   F) 3.3308 G) 3.4515 H) 3.5722 I) 3.6929 J) 3.8136

Solution D) 3.0894

From the preceding problem, we have the appropriate χ^2 goodness of fit table of observed counts and expected counts under null hypothesis:

Group	A	B	C	Total
Observed Count	86	472	693	1251
Expected	104.25	417.00	729.75	1251

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

$$v = \frac{(86 - 104.25)^2}{104.25} + \frac{(472 - 417)^2}{417} + \frac{(693 - 729.75)^2}{729.75} = 12.29976.$$

The crtical value is $\chi^2_{0.01,2} = 9.21034$. Thus, the test statistic exceeds the critical value by $12.29976 - 9.21034$, or 3.08942.

Here is the solution in R:

```
> E = 1251*c(1/12,4/12,7/12)
> E
[1] 104.25 417.00 729.75
> Obs = c(86,472,693)
> (Obs-E)^2/E
[1] 3.194844 7.254197 1.850719
> sum((Obs-E)^2/E)
[1] 12.29976
> pchisq(12.29976, df = 2, lower.tail=FALSE) # p-value (not required)
[1] 0.002133738
> qchisq(0.01, df = 2, lower.tail=FALSE) # classical hypothesis test critical value
[1] 9.21034
> 12.29976 - 9.21034 # answer
[1] 3.08942
```

9. A survey of drivers under 65 years old counted those drivers who had 0 or 1 accidents in the year prior to the survey. The accidents were divided into two types: major or minor. The categorical variable *Type of Accident* (with three values: None, Major, Minor) and the categorical variable *Age Group* are cross-tabulated below.

	None	Major	Minor	Total
$[16, 18)$	67	10	5	82
$[18, 26)$	42	6	5	53
$[26, 40)$	75	8	4	87
$[40, 65)$	56	4	6	66
Total	240	28	20	288

Are the conditional distributions of Type of Accident homogeneous across the age groups? In a classical chi-squared hypothesis test, what is the value of the test statistic? (A conclusion is not requested so a significance level need not be given. The next problem asks for the p-value of a contemporary hypothesis test.)

A) 3.0599 B) 3.1730 C) 3.2861 D) 3.3992 E) 3.5123
F) 3.6254 G) 3.7385 H) 3.8516 I) 3.9647 J) 4.0778

Solution **E) 3.5123**

The fraction of all table entries that are found in the first column is 240/288. Under the null hypothesis of homogeneity, this fraction would, for each row, be the fraction of the row entries found in the first column of the row. Thus, the first column entries would be (82)(240/288), (53)(240/288), (87)(240/288), and (66)(240/288) in the 1st, 2nd, 3rd, and 4th rows respectively. Generalizing this isdea, we see that, under the null hypothesis, the expected value in the ith row jth column is equal to (ith row total)(jth column total)/(table total). The table of expected counts is:

	None	Major	Minor	Total
$[16, 18)$	68.33333	7.972222	5.694444	82
$[18, 26)$	44.16667	5.152778	3.680556	53
$[26, 40)$	72.50000	8.458333	6.041667	87
$[40, 65)$	55.00000	6.416667	4.583333	66
Total	240	28	20	288

The test statistic is

$$\frac{(67-68.33333)^2}{68.33333} + \frac{(10-7.972222)^2}{7.972222} + \frac{(5-5.694444)^2}{5.694444}$$
$$+ \frac{(42-44.16667)^2}{44.16667} + \frac{(6-5.152778)^2}{5.152778} + \frac{(5-3.680556)^2}{3.680556}$$
$$+ \frac{(75-72.50000)^2}{72.50000} + \frac{(8-8.458333)^2}{8.458333} + \frac{(4-6.041667)^2}{6.041667}$$
$$+ \frac{(56-55.00000)^2}{55.00000} + \frac{(4-6.416667)^2}{6.416667} + \frac{(6-4.583333)^2}{4.583333},$$

which evaluates to 3.512298.

Here is the solution in R:

```
> Obs = matrix(c(67,10,5,42,6,5,75,8,4,56,4,6),ncol=3,byrow = TRUE)
> Obs
     [,1] [,2] [,3]
[1,]   67   10    5
[2,]   42    6    5
[3,]   75    8    4
[4,]   56    4    6
> E = matrix(rep(NA,12),ncol=3,byrow = TRUE)
> row.sums = rep(NA,4)
> for(i in 1:4){row.sums[i] = sum(Obs[i,])}
> col.sums = rep(NA,3)
> for(j in 1:3){col.sums[j] = sum(Obs[,j])}
> table.sum = sum(row.sums)
> table.sum == sum(col.sums)
[1] TRUE
> for(i in 1:4){for(j in 1:3){E[i,j]=row.sums[i]*col.sums[j]/table.sum}}
> E
         [,1]     [,2]     [,3]
[1,] 68.33333 7.972222 5.694444
[2,] 44.16667 5.152778 3.680556
[3,] 72.50000 8.458333 6.041667
[4,] 55.00000 6.416667 4.583333
> test.stat = sum( (Obs-E)^2/E )
> test.stat
[1] 3.512298
```

10. With regard to the test of homogeneity described in the preceding problem, what is the p-value of a contemporary test?

A) 0.1479	B) 0.2222	C) 0.2965	D) 0.3708	E) 0.4451
F) 0.5194	G) 0.5937	H) 0.6680	I) 0.7423	J) 0.8166

Solution **I) 0.7423**

There are 4 rows and 3 columns in the table. Therefore, the p-value is

$$P\left(\chi^2_{(4-1)(3-1)} \geq 3.512298\right) = 0.7423$$

The chi-squared table that is provided shows that $P\left(\chi^2_6 \geq 3.4546\right) = 0.75$ and $P\left(\chi^2_6 \geq 3.8276\right) = 0.70$. From these values you can deduce that the requested p-value is between 0.75 and 0.70 (and much closer to the former than to the latter). Because 0.7423 is the only answer choice bracketed by 0.70 and 0.75, an interpolation for a more precise evaluation of the p-value is unnecessary.

Here is the solution in R:

```
> pchisq(test.stat, df = (4-1)*(3-1), lower.tail=FALSE)
[1] 0.7423328
```

11. In an investigation of treatments that might prevent prostate cancer, 34,887 men were randomly assigned treatments of selenium, vitamin E, both selenium and vitamin E, or placebo. The men were followed and the incidences of prostate cancer in the four treatment groups were counted. The results are presented in the table below.

	No Cancer	Prostate Cancer	Total
Selenium	8177	575	8752
Vitamin E	8117	620	8737
Selenium and E	8147	555	8702
Placebo	8167	529	8696
Total	32608	2279	34887

In a test of the null hypothesis that Treatment and Development of Prostate Cancer are inependent, the null hypothesis is retained at significance level 0.05. By how much does the test statistic fall short of the critical value? (The next problem asks for a p-value.)

Source: Vitamin E and the risk of prostate cancer: the selenium and vitamin E cancer prevention trial (SELECT), Klein, E.A. and 20 co-authors, Journal of the American Medical Association 306 (2011), 1549-1556.

A) 0.0143 B) 0.0316 C) 0.0489 D) 0.0662 E) 0.0835
F) 0.1008 G) 0.1181 H) 0.1354 I) 0.1527 J) 0.1700

Solution **B) 0.0316**

Here is the table of expected counts under the null hypothesis:

	No Cancer	Prostate Cancer
Selenium	8180.274	571.7261
Vitamin E	8166.254	570.7462
Selenium and E	8133.540	568.4598
Placebo	8127.932	568.0679

Here is the solution in R:

```
> Obs = matrix(c(8177,575, 8117, 620, 8147, 555, 8167, 529), ncol = 2, byrow = TRUE)
> Obs
     [,1] [,2]
[1,] 8177  575
[2,] 8117  620
[3,] 8147  555
[4,] 8167  529
> E = matrix(rep(NA,8), ncol = 2)
> row.sums = rep(NA,4)
> for(i in 1:4){row.sums[i] = sum(Obs[i,])}
> row.sums
[1] 8752 8737 8702 8696
> col.sums = c(sum(Obs[,1]),sum(Obs[,2]))
> col.sums
[1] 32608  2279
> for(i in 1:4){for(j in 1:2){E[i,j]=row.sums[i]*col.sums[j]/34887}}
> E
```

```
         [,1]     [,2]
[1,] 8180.274 571.7261
[2,] 8166.254 570.7462
[3,] 8133.540 568.4598
[4,] 8127.932 568.0679
> test.stat = sum( (Obs-E)^2/E )
> test.stat
[1] 7.783171
> test.stat = 0   #  temporary value
>                 #  Start of alternative calculation, in case sum( (Obs-E)^2/E ) looks weird
> for(i in 1:4){for(j in 1:2){test.stat = test.stat + (Obs[i,j]-E[i,j])^2/E[i,j] }}
> test.stat
[1] 7.783171
> crit.val = qchisq(0.95, df = (4-1)*(2-1))
> crit.val
[1] 7.814728
> answer = crit.val - test.
> answer
[1] 0.031557
```

12. Like the preceding problem, this problem concerns a test of the null hypothesis that Treatment and Development of Prostate Cancer are independent. In a contemporary test of the null hypothesis, what is the p-value?

A) 0.0507 B) 0.0820 C) 0.1133 D) 0.1446 E) 0.1759
F) 0.2072 G) 0.2385 H) 0.2698 I) 0.3011 J) 0.3324

Solution **A) 0.0507**

There are 4 rows and two columns. The test statistic is 7.783171. The p-value is

$$P\left(\chi^2_{(4-1)(2-1)} \geq 7.783171\right) = P\left(\chi^2_3 \geq 7.783171\right) = 0.05071204.$$

Here is the solution in R:

```
> p.value = pchisq( test.stat, df = (4-1)*(2-1), lower.tail=FALSE)
> p.value
[1] 0.05071204
```

13. The table below records bivariate data for 8 women of Pima Indian heritage: the plasma glucose concentration Y and the diastolic blood pressure X.

X	102	78	58	74	85	64	56	76
Y	133	154	99	136	95	119	109	92

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 8 tabulated observations is $y = b_0 + b_1 x$ with sample Pearson correlation r. Sample statistics are: $\overline{\mathrm{X}} = 74.125$, $\overline{\mathrm{Y}} = 117.125$, $\mathrm{S_X} = 15.14159$, $\mathrm{S_Y} = 22.31871$, $r = 0.3026205$, $b_0 = 84.0606$, and $b_1 = 0.4460627$. In a classical one-sided hypothesis test of $\mathrm{H_0} : \beta_1 \leq 0$ 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 problem continues with this data.)

A) 0.7230 B) 0.8013 C) 0.8796 D) 0.9579 E) 1.0362
F) 1.1145 G) 1.1928 H) 1.2711 I) 1.3494 J) 1.4277

Solution F) 1.1145

With $n = 8$, we calculate $\mathrm{SST} = (n-1)\, S_Y^2 = 7 \cdot (22.31871)^2 = 3486.874$, $\mathrm{SSR} = r^2\, \mathrm{SST} = (0.3026205)^2 (3486.874) = 319.325$, $\mathrm{SSE} = \mathrm{SST} - \mathrm{SSR} = 3167.549$, $\mathrm{S}_e = \sqrt{\mathrm{SSE}/(n-2)} = \sqrt{280.2857/5} = 22.97661$, and $\mathrm{SE}\,(b_1) = \left(1/\sqrt{n-1}\right) \mathrm{S}_e/\mathrm{S}_\mathrm{X} = 0.5735423$. The endpoint of the critical region is $\mathrm{t}_{0.05,n-2} \times \mathrm{SE}\,(b_1)$, or 1.94318×0.5735423, or 1.114496. (Because $b_1 = 0.4460627 < 1.114496$, the null hypothesis would be retained.)

14. What is the lower endpoint of the 95% confidence interval for the slope of the regression line?

A) -1.2654 B) -1.1627 C) -1.0600 D) -0.9573 E) -0.8546
F) -0.7519 G) -0.6492 H) -0.5465 I) -0.4438 J) -0.3411

Solution D) -0.9573

We calculated $\mathrm{SE}\,(b_1) = 0.5735423$. We look up $t_{0.025,8-2} = 2.446912$. The lower endpoint of the 95% confidence interval for β_1 is $b_1 - t_{0.025,8-2} \times \mathrm{SE}\,(b_1)$, or $0.4460627 - 2.446912 \times 0.5735423$, or -0.9573448. (That the lower bound is negative is expected as a consequence of the preceding problem. In that problem a hypothesis test at significance level 0.05 retained the null hypothesis that $\beta_1 \leq 0$.)

15. The table below records bivariate data for 8 women of Pima Indian heritage: the body mass index Y and the diastolic blood pressure X. (The women are the same as in the preceding two problems, but from your point of view that is neither here nor there.)

X	102	78	58	74	85	64	56	76
Y	32.8	32.4	25.4	37.4	37.4	34.9	25.2	24.2

In terms of unknown coefficients β_0, β_1, and ρ, which are all different from the corresponding coefficients of the preceding two problems, the true linear model is $y = \beta_0 + \beta_1\, x$ with linear correlation ρ. The regression line calculated from the 8 tabulated observations is $y = b_0 + b_1\, x$ with sample Pearson correlation r. Sample statistics are $S_\mathrm{X} =$, $\mathrm{SST} = 213.2087$, and $\mathrm{SSR} = 44.61597$. Also, b_1 and r are positive. In the next problem you will be asked for the p-value of a one-sided hypothesis test of $\mathrm{H}_0 : \beta_1 = 0$ versus the alternative $\mathrm{H}_\mathrm{a} : \beta_1 > 0$. To answer that question, you will need, among other sample statistics, the value of b_1. Answer this question with the value of b_1. (Suggestion: start by finding the observed values of the sample statistics r and S_Y.)

A) 0.0215 B) 0.0578 C) 0.0941 D) 0.1304 E) 0.1667
F) 0.2030 G) 0.2393 H) 0.2756 I) 0.3119 J) 0.3482

Solution E) 0.1667

First we calculate the sample linear correlation:

$$r^2 = \frac{\mathrm{SSR}}{\mathrm{SST}} = \frac{44.61597}{213.2087} = 0.2092596.$$

It follows that $r = \sqrt{0.2092596} = 0.457449$. (Note: we are using the positive square root because we have been given $r > 0$.)

Next, we calculate S_Y:

$$S_\mathrm{Y} = \sqrt{\frac{1}{n-1}\,\mathrm{SST}} = \sqrt{\frac{1}{7}\,(213.2087)} = 5.518912.$$

The next sample statistic we calculate is b_1:

$$b_1 = r\,\frac{S_\mathrm{Y}}{S_\mathrm{X}} = (0.457449)\,\frac{5.518912}{15.14159} = 0.1667342.$$

16. Refer to the preceding problem in which X is diastolic blood pressure and Y is body mass index. 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 8 tabulated observations is $y = b_0 + b_1\, x$ with sample Pearson correlation r. As previously stated, given sample statistics are $S_{\mathrm{X}} = 15.14159$, $\mathrm{SST} = 213.2087$, and $\mathrm{SSR} = 44.61597$. Also, given were the inequalities $b_1 > 0$ and $r > 0$. In the previous problem you were asked for the value of the sample statistic b_1. What is the p-value of a one-sided hypothesis test of $\mathrm{H_0} : \beta_1 = 0$ versus the alternative $\mathrm{H_a} : \beta_1 > 0$?

A) 0.247 B) 0.277 C) 0.307 D) 0.337 E) 0.067
F) 0.097 G) 0.127 H) 0.157 I) 0.187 J) 0.217

Solution **G) 0.127**

First we calculate the sample linear correlation:

$$r^2 = \frac{\mathrm{SSR}}{\mathrm{SST}} = \frac{44.61597}{213.2087} = 0.2092596.$$

It follows that $r = \sqrt{0.2092596} = 0.457449$. (Note: we are using the positive square root because we have been given $r > 0$.)

Next, we calculate S_{Y}:

$$S_{\mathrm{Y}} = \sqrt{\frac{1}{n-1}\,\mathrm{SST}} = \sqrt{\frac{1}{7}\,(213.2087)} = 5.518912.$$

The next sample statistic we calculate is b_1:

$$b_1 = r\,\frac{S_{\mathrm{Y}}}{S_{\mathrm{X}}} = (0.457449)\,\frac{5.518912}{15.14159} = 0.1667342.$$

Three more sample statistics to go. We have $\mathrm{SSE} = \mathrm{SST} - \mathrm{SSR} = 213.2087 - 44.61597 = 168.5927$. Therefore,

$$S_e = \sqrt{\frac{1}{n-2}\,\mathrm{SSE}} = \sqrt{\frac{1}{6}\,(168.5927)} = 5.300829.$$

It follows that

$$\mathrm{SE}\,(b_1) = \frac{1}{\sqrt{n-1}}\,\frac{S_e}{S_{\mathrm{X}}} = \frac{1}{\sqrt{7}}\,\frac{5.300829}{15.14159} = 0.1323193.$$

At last, the p-value can be calculated:

$$\begin{aligned} p\text{-value} &= \mathrm{P}\left(\mathrm{t}_{n-2} > \frac{b_1}{\mathrm{SE}\,(b_1)}\right) \\ &= \mathrm{P}\left(\mathrm{t}_6 > \frac{0.1667342}{0.1323193}\right) \\ &= \mathrm{P}\,(\mathrm{t}_6 > 1.260089) \\ &= 0.1272142. \end{aligned}$$

17. The preceding two problems 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 $\mathrm{H}_0 : \rho = 0$ or $\mathrm{H}_a : \rho > 0$. Using significance level $\alpha = 0.05$, a classical one-sided hypothesis test of $\mathrm{H}_0 : \rho = 0$ against $\mathrm{H}_a : \rho > 0$ retains H_0. By how much does the test statistic fall short of the critical value?

A) 0.6252 B) 0.6831 C) 0.7410 D) 0.7989 E) 0.8568
F) 0.9147 G) 0.9726 H) 1.0305 I) 1.0884 J) 1.1463

Solution B) 0.6831

The test statistic is $r\sqrt{n-2}/\sqrt{1-r^2}$, or, using values for r and r^2 that have already been calculated, $(0.457449)\sqrt{8-2}/\sqrt{1-0.2092596}$, or 1.260089. The critical value of a one-sided test with significance level 0.05 is $t_{0.05,\,8-2} = 1.94318$. The answer is $1.94318 - 1.260089$, or 0.683091.

18. A random sample of 4 batteries was drawn from each of three battery types (the type being the "treatment"). The lifetimes of the batteries operating under low temperature conditions were recorded.

					$\overline{X_{i\cdot}}$	S_i^2
Treatment 1	130	155	74	180	134.75	2056.9173
Treatment 2	150	188	159	126	155.75	656.25
Treatment 3	138	110	168	160	144	674.6667

At significance level 0.05, a classical ANOVA test of the null hypothesis that the treatment means are all equal retains the null hypothesis. By how much does the critical value exceed the test statistic?
A) 3.3213 B) 3.4570 C) 3.5927 D) 3.7284 E) 3.8641
F) 3.9998 G) 4.1355 H) 4.2712 I) 4.4069 J) 4.5426

Solution E) 3.8641

The number of treatments is $m = 3$ and the number of samples is $n = 4$ for each treatment. The overall mean is $\overline{X_{\cdot\cdot}} = (134.75 + 155.75 + 144)/3 = 144.8333$. The null hypothesis estimate of the population variance is

$$\begin{aligned}
\text{MSB} &= S_0^2 \\
&= \frac{n}{m-1}\sum_{i=1}^{3}\left(\overline{X_{i\cdot}} - \overline{X_{\cdot\cdot}}\right)^2 \\
&= \frac{4}{2}\left((134.75 - 144.8333)^2 + (155.75 - 144.8333)^2 + (144 - 144.8333)^2\right) \\
&= 443.0833.
\end{aligned}$$

The average treatment group sample variance is

$$\begin{aligned}
\text{MSW} &= \overline{S^2} \\
&= \frac{1}{3}\left(S_1^2 + S_2^2 + S_3^2\right) \\
&= \frac{1}{3}\,(2056.9173 + 656.25 + 674.6667) \\
&= 1129.278.
\end{aligned}$$

The test statistic is

$$\frac{\text{MSB}}{\text{MSW}} = \frac{S_0^2}{\overline{S^2}} = \frac{443.0833}{1129.278} = 0.3923598.$$

The critical value is $f_{0.05,2,9} = 4.256495$. The critical value exceeds the test statistic by $4.256495 - 0.3923598$, or 3.864135.

Here is the solution using R. The R-session not only solves the problem, it also obtains the values presented in the two marginal columns of the data table given in the statement of the problem.

```
> B1 = c(130,  155,  74, 180)
> B1.obs.mean = mean(B1)
> B1.obs.mean
```

```
[1] 134.75
> B1.obs.var = sd(B1)^2
> B1.obs.var
[1] 2056.917
> B2 = c(150, 188, 159, 126)
> B2.obs.mean = mean(B2)
> B2.obs.mean
[1] 155.75
> B2.obs.var = sd(B2)^2
> B2.obs.var
[1] 656.25
> B3 = c(138,110, 168, 160)
> B3.obs.mean = mean(B3)
> B3.obs.mean
[1] 144
> B3.obs.var = sd(B3)^2
> B3.obs.var
[1] 674.6667
> battery.lifetimes = c(B1,B2,B3)
> treatment = c(rep("B1",4),rep("B2",4),rep("B3",4))
> batt.life.data.frame = data.frame(battery.lifetimes,treatment)
> batt.life.data.frame
   battery.lifetimes treatment
1                130        B1
2                155        B1
3                 74        B1
4                180        B1
5                150        B2
6                188        B2
7                159        B2
8                126        B2
9                138        B3
10               110        B3
11               168        B3
12               160        B3
> # The test statistic is found in the "F value" column of the next return.
> summary(aov(battery.lifetimes~treatment,data=batt.life.data.frame))
            Df Sum Sq Mean Sq F value Pr(>F)
treatment    2    886   443.1   0.392  0.686
Residuals    9  10163  1129.3
> crit.val = qf(0.05, df1 = 3-1, df2 = 3*(4-1), lower.tail=FALSE)
> crit.val
[1] 4.256495
> answer = crit.val - 0.392 #The 0.392 is from the "F val" column of the ANOVA table.
> answer                    #The accuracy of the answer is limited by the F value
[1] 3.864495
> # As an alternative, we can do the calculations ourselves
> m = 3  # number of treatments
> n = 4  # sample size per treatment
> grand.mean = mean(batt.life.data.frame[,1])
```

```
> grand.mean
[1] 144.8333
> MSB = (n/(m-1))*( (B1.obs.mean-grand.mean)^2 + (B2.obs.mean-grand.mean)^2 + (B3.obs.mean-grand.mean)^2 )
> MSB
[1] 443.0833
> MSW = (B1.obs.var + B2.obs.var + B3.obs.var)/3
> test.stat = MSB/MSW
> test.stat
[1] 0.3923599
> crit.val - test.stat
[1] 3.864135
```

19. The operating lifetimes of three types of batteries were tested by randomly selecting 3 batteries of each type, and, for each type, operating one battery at low temperature until failure, operating a second battery at medium temperature until failure, and operating the third battery at high temperature until failure. The observed lifetimes are recorded in the table below. Consider the type of battery a treatment and the temperature as a blocking factor.

	Low (-10C)	Medium (20C)	High (45C)	$X_{i\cdot}$
Battery 1	130	65	63	258
Battery 2	150	115	56	321
Battery 3	138	139	95	372
$X_{\cdot j}$	418	319	214	951

The marginal entries of this table are row and column *sums*, not means. The sum of the squares of all non-marginal table entries is 111,345.

There are two questions concerning this dataset. In this problem we are concerned with the treatment effects. In the next problem we will consider the block effects. At significance level 0.05, a classical hypothesis test retains the null hypothesis that the treatment effects are the same. By how much does the test statistic fall short of the critical value?

A) 0.8651 B) 1.3774 C) 1.8897 D) 2.4020 E) 2.9143
F) 3.4266 G) 3.9389 H) 4.4512 I) 4.9635 J) 5.4758

Solution **H) 4.4512**

Here $m = 3$ is the number of treatments and $n = 3$ is the number of observations per treatment. We calculate $\overline{\mathrm{X}_{\cdot\cdot}} = \frac{1}{m\,n}\,\mathrm{X}_{\cdot\cdot} = \frac{1}{9}\,951 = 105.6667$. Therefore,

$$
\begin{aligned}
\mathrm{SST} &= \sum_{i=1}^{m}\sum_{j=1}^{n}\mathrm{X}_{ij}^2 - m\,n\,\overline{\mathrm{X}_{\cdot\cdot}}^2 \\
&= 111345 - 9\,(105.6667)^2 \\
&= 10855.94,
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{SS(Tr)} &= n\sum_{i=1}^{m}\overline{\mathrm{X}_{i\cdot}}^2 - m\,n\,\overline{\mathrm{X}_{\cdot\cdot}}^2 \\
&= 3\left(\left(\frac{258}{3}\right)^2 + \left(\frac{321}{3}\right)^2 + \left(\frac{372}{3}\right)^2\right) - 9\,(105.6667)^2 \\
&= \frac{1}{3}\,\big((258)^2 + (321)^2 + (372)^2\big) - 9\,(105.6667)^2 \\
&= 2173.937,
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{SS(Bl)} &= m\sum_{j=1}^{n}\overline{\mathrm{X}_{\cdot j}}^2 - m\,n\,\overline{\mathrm{X}_{\cdot\cdot}}^2 \\
&= 3\left(\left(\frac{418}{3}\right)^2 + \left(\frac{319}{3}\right)^2 + \left(\frac{214}{3}\right)^2\right) - 9\,(105.6667)^2 \\
&= \frac{1}{3}\,\big((418)^2 + (319)^2 + (214)^2\big) - 9\,(105.6667)^2 \\
&= 6937.937,
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{SSE} &= \mathrm{SST} - \big(\mathrm{SS(Tr)} + \mathrm{SS(Bl)}\big) \\
&= 10855.94 - (2173.937 + 6937.937) \\
&= 1744.066.
\end{aligned}
$$

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

$$\begin{aligned}
\mathrm{MS(Tr)} &= \frac{1}{m-1}\,\mathrm{SS(Tr)} \\
&= \frac{1}{2}\,(2173.937) \\
&= 1086.968,
\end{aligned}$$

$$\begin{aligned}
\mathrm{MS(Bl)} &= \frac{1}{n-1}\,\mathrm{SS(Bl)} \\
&= \frac{1}{2}\,(6937.937) \\
&= 3468.968,
\end{aligned}$$

$$\begin{aligned}
\mathrm{MSE} &= \frac{1}{(m-1)(n-1)}\,\mathrm{SSE} \\
&= \frac{1}{4}\,(1744.066) \\
&= 436.0165.
\end{aligned}$$

To test whether there are different treatment effects on the mean lifetimea, we use the observed value for the test statistic MS(Tr)/MSE, namely 1086.968/436.0165, or 2.492952. The critical value is $f_{0.05,2,4} = 6.944272$. The amount by which the test statistic falls short of the critical value is $6.944272 - 2.492952$, or 4.45132.

The R-code for the solution is found on the next page. (Note: The ANOVA table gives the p-value, 0.19814, so the calculation continues beyond the table. Notice that the p-value indicates that, in a traditional hypothesis test, the null hypothesis is retained.)

```
> battery.lifetimes = c(130,65,63,150,115,56,138,139,95)
> battery.types = c(rep("B1",3),rep("B2",3),rep("B3",3))
> battery.blocks = c(rep(c("L","M","H"),3))
> battery.data = data.frame(battery.lifetimes,battery.types,battery.blocks)
> colnames(battery.data) = c("lifetimes","treatments","blocks")
> battery.data
  lifetimes treatments blocks
1       130         B1      L
2        65         B1      M
3        63         B1      H
4       150         B2      L
5       115         B2      M
6        56         B2      H
7       138         B3      L
8       139         B3      M
9        95         B3      H
> anova(lm(lifetimes~treatments+blocks,data=battery.data))
Analysis of Variance Table

Response: lifetimes
           Df Sum Sq Mean Sq F value  Pr(>F)
treatments  2   2174    1087  2.4931 0.19814
blocks      2   6938    3469  7.9564 0.04035 *
Residuals   4   1744     436
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
> test.stat = 2.4931                        # treatment row entry of F-val column
> crit.val = qf(0.05,2,4,lower.tail=FALSE)    # numerator Df = 2, denominator Df = 4
> crit.val
[1] 6.944272
> crit.val - test.stat
[1] 4.451172
```

20. With regard to the battery lifetimes discussed in the preceding problem, a classical hypothesis test at significance level 0.05 rejects the null hypothesis that the temperature effects are the same. By how much does the test statistic exceed the critical value?

A) 0.9120 B) 1.0121 C) 1.1122 D) 1.2123 E) 1.3124
F) 1.4125 G) 1.5126 H) 1.6127 I) 1.7128 J) 1.8129

Solution **B) 1.0121**

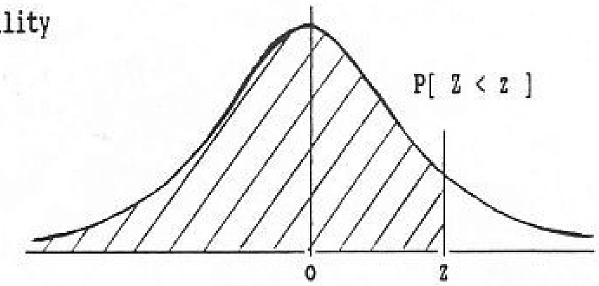
Most of the work has been done. The observed value of the test statistic is MS(Bl)/MSE, or 3468.968/436.0165, or 7.956048. The critical value is $f_{0.05,2,4} = 6.944272$. (This happens to be the same critical value for treatment effects. That is because, in this problem, $n = m$.) The test statistic exceeds the critical value by $7.956048 - 6.944272$, or 1.011776. (Note: The p-value in the ANOVA table, 0.04035, indicates a significant rejection of the null hypothesis that temperature effects on mean battery lifetimes are the same for the three tested temperatures.)

STANDARD STATISTICAL TABLES

1. Areas under the Normal Distribution

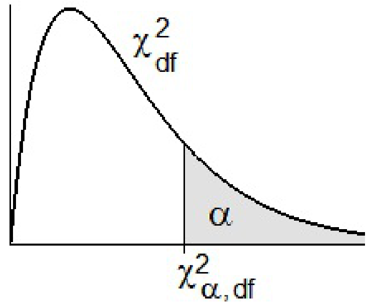
The table gives the cumulative probability up to the standardised normal value z
i.e.

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



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

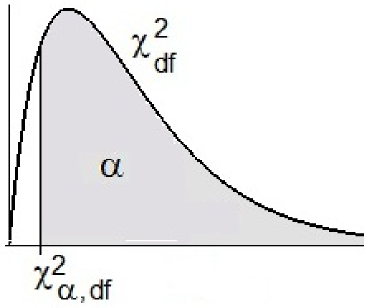
z	3.00	3.10	3.20	3.30	3.40	3.50	3.60	3.70	3.80	3.90
P	0.9986	0.9990	0.9993	0.9995	0.9997	0.9998	0.9998	0.9999	0.9999	1.0000



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

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

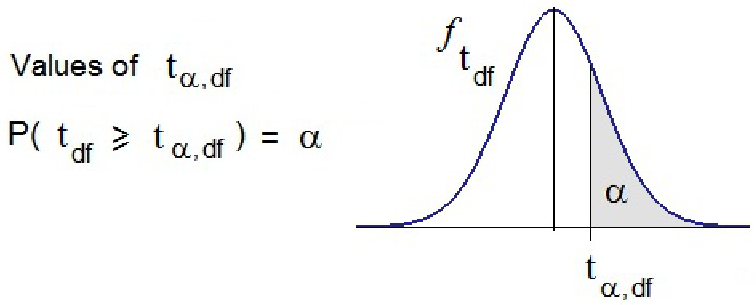
Chi-Squared Values—Right Tails.



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

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.

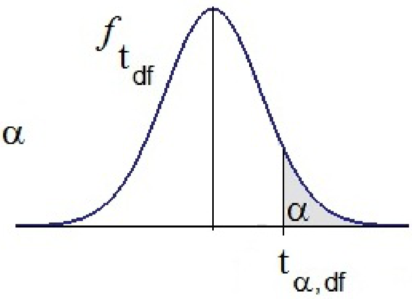


df \ α	.450	.400	.350	.300	.250	.200
1	.1584	.3249	.5095	.7265	1.0000	1.3764
2	.1421	.2887	.4447	.6172	.8165	1.0607
3	.1366	.2767	.4242	.5844	.7649	.9785
4	.1338	.2707	.4142	.5686	.7407	.9410
5	.1322	.2672	.4082	.5594	.7267	.9195
6	.1311	.2648	.4043	.5534	.7176	.9057
7	.1303	.2632	.4015	.5491	.7111	.8960
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 $\alpha = 0.45, 0.40, 0.35, 0.30, 0.25, 0.20.$

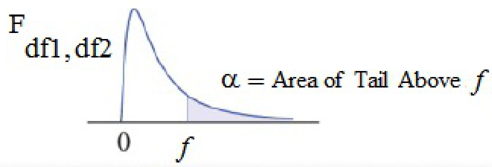
Values of $t_{\alpha,df}$

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



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