INSTITUTE OF TECHNOLOGY, BLANCHARDSTOWN

Academic Year	2015-6
Year of Programme	2
Semester	Semester Two, First Sitting
Date of Examination	First exam
Time of Examination	First exam

Programme Code	Programme Title	Module Code
BN001	Higher Certificate in Engineering in Computer Engineering	ELTC H2021
BN009	Bachelor of Engineering in Mechatronic Engineering	MECH H2017
BN012	Bachelor of Engineering in Computer Engineering	ELTC H2021
BN117	Bachelor of Engineering (honours) in Computer Engineering in Mobile Systems	ELTC H2021
BN121	Bachelor of Engineering (honours) in Mechatronics	MECH H2017
BN903	Higher Certificate in Engineering in Mechatronic Engineering	MECH H2017

Module Title	Statistics and Probability

Internal Examiner: **Dr. Kan Tadd**

External Examiner(s): **Dr Muhammad ibn Musa Al-Khwarizmi**

Instructions to candidates:

1. **To ensure that you take the correct examination, please check that the module and programme that you are following are listed in the table above.**
2. **The information section, including statistical tables, is at the end of the paper.**
3. **Answer all five questions in this paper.**

DO NOT TURN OVER THIS PAGE UNTIL YOU ARE TOLD TO DO SO

Question 1 (20 marks)

Answer any two of the following three parts. **Each part carries 10 marks.**

a) Answer the following two questions. Each question carries 5 marks.

(i) Explain how a comparison of the mean and the median can give additional information about the distribution of numbers within a data set, illustrating your explanation with an example

(ii) For a list of data values x_1, x_2, … x_n, identify the two quantities defined in the equations below and explain what they measure:

$$\bar{x} = \frac{\sum_i x_i}{n}, \quad s^2 = \frac{\sum_i (x_i - \bar{x})^2}{n-1}.$$

(iii) Calculate appropriate estimates of the mean, the median and the standard deviation for the following frequency distribution:

Group	Frequency
50 to 52	4
52 to 54	6
54 to 56	9
56 to 58	21

b) For a frequency distribution, the frequency mean is given by the expression shown here:

$$\bar{x} = \frac{\sum_i m_i f_i}{n}$$, where the summation includes all groups.

Prove that this estimate is halfway between the minimum and maximum possible values of the mean for the distribution.

c) For a list of data values x_1, x_2, … x_n, the mean and standard deviation are defined in the equations below:

$$\bar{x} = \frac{\sum_i x_i}{n}, \quad s^2 = \frac{\sum_i (x_i - \bar{x})^2}{n-1}.$$

Show that the equation for s can be rewritten in the more practical form shown here:

$$s^2 = \frac{\sum_i x_i^{\,2} - n(\bar{x})^2}{n-1}.$$

d) The points listed here are paired sample values of the two variables x and y:

(X_1, Y_1), (X_2, Y_2), … (X_n, Y_n).

The sample correlation coefficient r is defined by the equation shown:

$$r = \frac{\sum \left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)}{\sqrt{\sum \left(X_i - \overline{X}\right)^2}\sqrt{\sum \left(Y_i - \overline{Y}\right)^2}}.$$

Show that this is equivalent to the more practical version given here:

$$r = \frac{\sum_i X_i Y_i - n\overline{X}\,\overline{Y}}{\sqrt{\sum_i X_i^{\,2} - n\overline{X}^2}\sqrt{\sum_i Y_i^{\,2} - n\overline{Y}^2}}.$$

e) The points listed here are paired sample values of the two variables x and y:

(X_1, Y_1), (X_2, Y_2), … (X_n, Y_n).

Prove that if the quantities x and y are linked by a linear equation $y = mx + c$, then the value of the sample correlation coefficient r is +1 or –1, depending on the sign of m. The sample correlation coefficient r is defined by the equation shown:

$$r = \frac{\sum \left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)}{\sqrt{\sum \left(X_i - \overline{X}\right)^2}\sqrt{\sum \left(Y_i - \overline{Y}\right)^2}}.$$

Question 2 (20 marks)

Answer two of the following three parts. **Each part carries 10 marks.**

a) Answer the following two questions. Each question carries 5 marks.

   i) Give a definition of an event and the probability of that event happening. Give an example of an event for which it is difficult to define a probability of it occurring, explaining why.

   ii) A program generates a random integer between 1 and 100, inclusive, with all numbers equally likely to come up. Identify the experiment and the event space in this example and so calculate the probability that the integer is a cube.

b) A box contains 100 components, of which 7 are defective. A set of 4 of the components are selected at random. Answer the following two questions. Each question carries 5 marks.

   i) Calculate the probability that of the 4 components chosen, 3 are defective. Identify the experiment and the event space in this case.

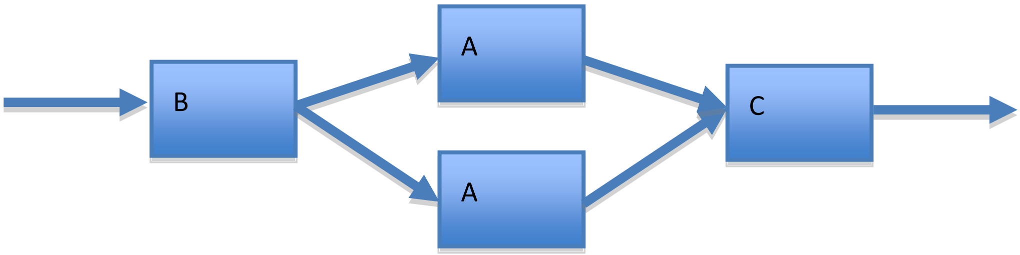
   ii) Show that ${}^{7}C_{3} = {}^{7}C_{4}$ using the definition of the quantity ${}^{n}C_{r}$. No marks will be awarded for simply calculating the two values and showing they are the same.

c) Answer the following two questions. Each question carries 5 marks.

   (i) Two fair dice are rolled and the numbers coming up are recorded. Let S be the sum of the two numbers produced. Create a distribution for this random variable.

   (ii) The Cumulative distribution function for the exponential distribution is given by the equation $F(x) = 1 - e^{-\lambda x}$. Find the probability density function and so find the expected value of variable X.

d) In the system shown in the diagram, the probability that a component of type A fails is 0.06, type B is 0.03 and type C is 0.02.



Answer the following two questions. Each question carries 5 marks.

i) Calculate the probability the overall system works, stating clearly the laws of probability used during the calculation.

ii) Justify the choice of which component is doubled up by calculating the probability of failure without doubled up components.

e) Answer the following two questions. Each question carries 5 marks.

(iii) For a continuous random variable X, give a definition of the cumulative distribution function and the probability density function. Show how the probability density function is used to calculate the probability of the event $a < X < b$, giving an example.

(iv) Two fair dice are rolled and the numbers coming up are recorded. Let D be the difference of the two numbers produced. Create a distribution for this random variable.

Question 3 (20 marks)

Answer two of the following four parts. **Each part carries 10 marks.**

a) An employee has established that during the working day, emails arrive at a rate of 5 in every hour. Answer the following two questions. Each question carries 5 marks.

   i) The employee has cleared their inbox. Calculate the probability that an email arrives within the next 10 minutes.

   ii) The employee attends a meeting for 30 minutes. Calculate the probability that 4 emails arrive within that time.

b) A factory is producing components, of which a certain proportion p are defective. They are packed in boxes of n components. Let X be the random variable of the number of defective components in a box. Answer the following two questions. Each question carries 5 marks.

   i) Show that the distribution for X is given by the equation shown here:

$$P[X = r] = {}^{n}\mathrm{C}_{r}\, p^{r}(1 - p)^{n - r}.$$

   ii) For the case where 1.2% of the components are defective and they are packed in boxes of 16, calculate the probability that a box has 3 or more defective components. The Poisson approximation may be used.

c) A faulty house alarm system is generating a false alarm twice in week. Calculate the probabilities of the following two events.

   i) The alarm goes off within the next 2 days.

   ii) The alarm goes off 4 times in 5 days.

d) A manufacturing facility is producing components which are required to have a capacitance of 250 pF. In fact it has been found that the capacitance C is a random variable following a normal distribution with mean 251.8 pF and standard deviation 6.1 pF. Answer the following two questions. Each question carries 5 marks.

i) Calculate a capacitance such that the probability a component chosen at random will have a lower capacitance is 0.99.

ii) Calculate a capacitance such that the probability a component chosen at random will have a lower capacitance is 0.1.

iii) Calculate a capacitance such that the probability a component chosen at random will have a greater capacitance is 0.98.

iv) Calculate a capacitance such that the probability a component chosen at random will have a higher capacitance is 0.025. Calculate the middle 95% range of capacitances.

e) A plant nursery uses a particular trace mineral in its production processes. Let S be the random variable of the quantity of the mineral used in a week at the facility. The manager has found that S is normally distributed with mean 2.55 g and standard deviation 0.245 g. Answer the following two questions. Each question carries 5 marks.

i) Calculate the probability that the nursery uses more than 2.6g of the trace mineral in a given week, and from this calculate the probability that the nursery uses between 2.5 g and 2.6 g in a given week.

ii) Calculate the middle range for 0.98, that is, the two values a and b such that

$$P[\ a < S < b\ ] = 0.98 \text{ and } P[\ S < a\ ] = P[\ S > b\ ].$$

f) Answer the following two questions. Each question carries 5 marks.

i) Let N be a random variable with event space $\{0, 1, 2, \ldots \infty\}$. The variable N follows the Poisson distribution, given by the equation shown here:

$$P[N = r] = e^{-\mu} \frac{\mu^r}{r!} .$$

Calculate $\sum_{r=0}^{\infty} P[N = r]$ and explain what this means for the random variable N.

ii) The Cumulative Distribution Function for a random variable is given by the integral equation shown here, where μ and σ are parameters:

$$P[X < a] = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{a} \exp\left[-\frac{1}{2}\left(\frac{u-\mu}{\sigma} \right)^2 \right] \mathrm{d}u .$$

Show that this leads to the standardisation rule shown below. Your answer should include an explanation of the variable Z:

$$P[X < a] = P\left[Z < \frac{a-\mu}{\sigma} \right].$$

Question 4 (20 marks)

In all parts to this question, the statement of the problem is followed by the information available for the required calculations.

Answer two of the following three parts. **Each part carries 10 marks.**

a) In an Institute of Technology, the number of students taking one of four academic programmes and reporting their stress as high (HS), low (LS) or none at all (NS) were counted in a survey. Carry out an appropriate statistical test to determine whether the data suggests there is link between level of stress and programme of study.

Available information

The counts are given in the table:

	Science	*Business*	*Computing*	*Engineering*
HS	53	51	41	25
LS	16	21	24	21
NS	6	12	13	17

Some values of the expression shown are available in the following table. The terms O_i and E_i are the observed and expected values, as usual.

$\frac{(O_i - E_i)^2}{E_i}$	*Science*	*Business*	*Computing*	*Engineering*
HS	2.59	0.24	0.23	3.21
LS	0.99	0.17	0.34	0.83
NS	?	0.15	0.02	?

b) A company is manufacturing components with 3 separate production processes. During the production process, 4 important faults in the components are identified. There are no instances of two faults or more on a given component. The management seek to identify whether there is a link between production method and identified fault; to investigate this question, in a given time-period, each incidence of a component with a fault was recorded with the production method. Carry out an appropriate statistical test to determine whether the data suggests there is link between the type of fault and the production method.

Available information

The counts are given in the table:

	Fault A	*Fault B*	*Fault C*	*Fault D*
Method 1	126	152	169	178
Method 2	154	151	157	153
Method 3	175	142	145	125

Some values of the expression shown are available in the following table. The terms O_i and E_i are the observed and expected values, as usual.

$\frac{(O_i - E_i)^2}{E_i}$	*Fault A*	*Fault B*	*Fault C*	*Fault D*
Method 1	5.649	0.000	0.385	3.105
Method 2	0.005	0.010	0.015	0.002
Method 3	?	0.007	0.265	?

c) A cement product manufacturer is making circular capstones. The manager has taken a random sample of 13 capstones and measured their diameters in centimetres. Carry out an appropriate statistical test on the two statements given here, using the data given below.

i) The mean diameter of the capstones is 108cm.

ii) The variance of the diameters is 5.2cm^2.

Available information

Let D_i, for i = 1 to 13, denote the diameter measured for each capstone. Then:

$$\sum_i D_i = 1410.8 \text{ and } \sum_i D_i^2 = 1{,}53123.25 .$$

d) A lecturer is seeking to prove that attendance is a strong predictor of final marks for a group of students. For each student, the attendance over the course of the semester and the final mark were recorded. Answer the following two questions. Each question carries 5 marks.

i) Determine, using an appropriate statistical test, whether or not there is a positive correlation between the two variables ‘attendance’ and ‘final mark’.

ii) Construct a simple linear equation that allows you to predict the final mark from the attendance. Comment on how useful this equation will be, given the result of the previous question .

Available information

The table of recorded values is given here.

Attendance	45	32	67	56	78	86	43
Final Mark	86	34	62	76	65	74	23

e) An engineer seeks to establish an empirical link between the temperature of a component and its resistance by measuring the resistance and a given temperature fifteen times. Answer the following two questions. Each question carries 5 marks.

i) Carry out an appropriate test to determine if the two variables are correlated.

ii) Calculate the coefficients of a simple equation to predict the resistance R from a given temperature T. Explain how this equation minimises errors in prediction and state whether you would consider it to be a good fit to the data.

Available information

Let T_i be the temperature recorded in Celsius and R_i the resistance in kilo-Ohms, for $i = 1$ to 15, for each of the components. The following sums are available:

$$\sum_i T_i = 907.0, \quad \sum_i R_i = 730.5, \quad \sum_i T_i R_i = 47{,}509.5, \quad \sum_i R_i^2 - n\overline{R}^2 = 2509.9,$$

$$\sum_i T_i^2 - n\overline{T}^2 = 11{,}625.73.$$

f) A lecturer delivering a module seeks to establish a link between the number of hours students spend working per week and the overall mark they achieve. The following data was collected for a group of 8 students, otherwise judged to be of equal ability, with X_i representing the working hours and Y_i the overall mark for each student.

X_i	27	32	10	8	12	4	33	15
Y_i	57	34	51	76	65	74	54	65

Answer the following three questions.

i) Calculate the value of the quantity $\sum_i X_i Y_i$.

ii) The sample correlation coefficient for this data is $r = -0.747$. Carry out an appropriate test to determine if the two variables are negatively correlated.

iii) Write down the least squares regression equation for this data, given that the sample value of one of the coefficients is $b = -0.896$. Comment on how close the equation fits the data and therefore how useful it is.

Question 5 (20 marks)

In all parts to this question, the statement of the problem is followed by the information available for the required calculations.

Answer two of the following three parts. **Each part carries 10 marks.**

a) The weights of the fruits of a random sample of 200 tomato plants have been measured during a given growing season. The data was organised as a frequency distribution, shown in the table below. Carry out an appropriate statistical test to determine if the heights of the plants follow a Normal distribution.

Available information

The terms O_i and E_i are the observed and expected values, as usual.

Lower bound	*Upper bound*	*Frequency*	*Values of* $\frac{(O_i - E_i)^2}{E_i}$
30.0	30.5	2	*0.654*
30.5	31.0	6	*0.001*
31.0	31.5	15	*1.018*
31.5	32.0	39	*0.061*
32.0	32.5	59	*0.529*
32.5	33.0	56	*2.663*
33.0	33.5	11	*7.027*
33.5	34.0	9	*?*
34.0	34.5	2	*?*
34.5	35.0	1	*2.506*

The following data is also available for the frequency distribution, where the symbols m_i and f_i have their usual meanings.

$$\sum_i m_i f_i = 6{,}461.50\,,\ \sum_i m_i^2 f_i = 208{,}860.0\,.$$

b) In an experiment, 250 values of the response variable were organised as a frequency distribution, shown in the table below. Carry out an appropriate statistical test to determine if the variable follows a Normal distribution.

Available information

The terms O_i and E_i are the observed and expected values, as usual.

Lower bound	*Upper bound*	*Frequency*	*Values of* $\frac{(O_i - E_i)^2}{E_i}$
130.0	130.5	6	2.249
130.5	131.0	15	0.961
131.0	131.5	44	2.804
131.5	132.0	109	?
132.0	132.5	56	0.382
132.5	133.0	11	?
133.0	133.5	9	5.942

The following data is available, where the symbols m_i and f_i have the usual meanings.

$$\sum_i m_i f_i = 39{,}531.5, \ \sum_i m_i^2 f_i = 5{,}209{,}216.0 .$$

c) A computer program is designed to produce a random integer between 1 and 5, with each integer equally likely to come up. The program is run a large number of times and the occurrence of each integer is counted, with the results shown in the table below. Carry out an appropriate statistical test to determine if the numbers follow a uniform distribution.

Available information

The terms O_i and E_i are the observed and expected values, as usual.

Integer	*1*	*2*	*3*	*4*	*5*
Frequency	45	62	51	39	59
Values of $\frac{(O_i - E_i)^2}{E_i}$	0.7508	2.2781	0.0008	?	?

d) A module is delivered in an Institute of Technology, with four different lecturers taking the workshop classes. The end-of-semester marks of the students in each class are shown in the table below. Carry out an appropriate test to determine if there is any difference in academic performance in the classes, as measured by the end-of-semester marks.

Available Information

The overall sum of squares is $\Sigma X_i^2 = 128{,}807$.

Supplier	*Number of fails:*	*Total:*
A	53, 45, 24, 66, 63, 65, 64, 61, 47, 55, 54, 53	650
B	53, 54, 55, 58, 57, 59, 56, 57	449
C	63, 62, 64, 61, 60, 69, 45, 54, 63, 74	615
D	47, 48, 49, 47, 46, 47, 48, 60, 68, 67	527

e) During its operations, an assembly plant receives deliveries of components from four suppliers. The deliveries come in containers of 500 from just one supplier. For 40 randomly selected containers, the components were tested and the number of fails noted. The results are shown in the table below. Carry out an appropriate test to determine if any supplier is producing more or less fails in its products.

Available Information

The overall sum of squares is $\Sigma X_i^2 = ?$.

Supplier	*Number of fails:*	*Total:*
A	23, 25, 24, 26, 23, 25, 24, 21, 27, 25, 24, 23	290
B	23, 24, 25, 28, 27, 29, 26, 27	209
C	23, 24, 22, 21, 20, 19, 22, 24, 23, 24	222
D	27, 28, 29, 27, 26, 27, 28, 30, 28, 27	277

f) An experiment was carried out with a random selection of 60 components produced by four distinct methods in three manufacturing plants. The 60 components were tested to destruction and these times-to-failure were recorded. Carry out an appropriate test to decide whether there is a difference in times-to-failure between the four plants and three methods.

Available information

The total sum of squares is 34.680. The sources of variation and sums of squares are shown here:

Source	*Sum of squares*
Production Method	2.542
Plant	6.609
Mixed effect	3.571

g) A two-factor experiment gave 50 values of the response variable. The first factor has 3 levels and the second has 4. Carry out a two-way analysis of variance to decide whether there is a difference in the response variable between the two factors.

Available information

The total sum of squares is 24.680. The sources of variation and sums of squares are shown here:

Source	*Sum of squares*
Production Method	3.242
Plant	5.109
Mixed effect	4.071

Information

Statistical tables

These are available in the last four pages.

Sample Statistics

For a list of n figures $X_1, X_2, \ldots, X_n$:

The sample mean is: $\overline{X} = \dfrac{\sum_i X_i}{n}$. The sample standard deviation is: $S^2 = \dfrac{\sum_i X_i^{\,2} - n\left(\overline{X}\right)^2}{n-1}$.

Tests on Means

For a set of n random data values with true mean μ, then:

If $n > 50$, the variable $Z = \dfrac{\overline{X} - \mu}{S/\sqrt{n}}$ follows the standard normal distribution.

If $n < 30$, the variable $T = \dfrac{\overline{X} - \mu}{S/\sqrt{n}}$ follows the t-distribution with $n - 1$ degrees of freedom.

Correlation

For a list of paired values (X_1, Y_1) to (X_n, Y_n):

The sample correlation coefficient r is: $r = \dfrac{\sum_i X_i Y_i - n\overline{X}\,\overline{Y}}{\sqrt{\sum_i X_i^{\,2} - n\overline{X}^2}\sqrt{\sum_i Y_i^{\,2} - n\overline{Y}^2}}$.

The statistic t given by $t = r\sqrt{\dfrac{n-2}{1-r^2}}$, follows the t-distribution with $n - 2$ degrees of freedom.

Regression

The equations for the parameter α and β in the least squares regression line are given by:

$$\beta = \frac{\sum_i X_i Y_i - n\overline{X}\,\overline{Y}}{\sum_i X_i^{\,2} - n\overline{X}^2} \text{ and } \overline{Y} = \alpha + \beta\overline{X} .$$

Categorical tests

Let O_i be the observed frequency values and let E_i be the expected values predicted from the Null Hypothesis.

- For a test for independence conducted on a table of values of size r by c, the quantity χ^2 has $(r-1)(c-1)$ degrees of freedom.
- For a goodness-of-fit test on a frequency distribution, the quantity χ^2 has $k - p - 1$ degrees of freedom, where k is the number of groups and p is the number of parameters calculated for the distribution under the Null Hypothesis.
- The sample value of χ^2 is calculated from the equation $\chi^2 = \sum_i \frac{(O_i - E_i)^2}{E_i}$,

One-way analysis of Variance

Using the symbol m for the overall mean, and m_j to be the mean for group j:

$$m = \overline{X} = \frac{\sum X_i}{N} \text{ and } m_j = \overline{X}_j = \frac{\sum_{\text{group } j} X_i}{n_j}.$$

The *total sum of squares*, denoted S_1, is: $S_1 = \sum \left(X_i - \overline{X}\right)^2 = \sum (X_i - m)^2 = \sum X_i^{\,2} - Nm^2$.

The sum of squares between the groups, denoted S_T, is $S_T = \sum_j n_j \left(m_j - m\right)^2$.

The sum of squares for the error is $S_E = \sum_{i,j} \left(X_i - m_j\right)^2 = \sum_j \sum_{\text{within } j} X_i - n_j m_j^{\,2}$.

Table of probabilities P[Z > a], where Z is the standard normal variable.

	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
3.5	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
3.6	0.0002	0.0002	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
3.7	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
3.8	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
3.9	0.0									
4.0	0.0									

One-Tailed Critical Values for the Student t distribution

Degrees of freedom	***Probability (level of significance)***			
	0.05	***0.025***	***0.01***	***0.005***
1	6.314	12.706	31.821	63.657
2	2.920	4.303	6.965	9.925
3	2.353	3.182	4.541	5.841
4	2.132	2.776	3.747	4.604
5	2.015	2.571	3.365	4.032
6	1.943	2.447	3.143	3.707
7	1.895	2.365	2.998	3.499
8	1.860	2.306	2.896	3.355
9	1.833	2.262	2.821	3.250
10	1.812	2.228	2.764	3.169
11	1.796	2.201	2.718	3.106
12	1.782	2.179	2.681	3.055
13	1.771	2.160	2.650	3.012
14	1.761	2.145	2.624	2.977
15	1.753	2.131	2.602	2.947
16	1.746	2.120	2.583	2.921
17	1.740	2.110	2.567	2.898
18	1.734	2.101	2.552	2.878
19	1.729	2.093	2.539	2.861
20	1.725	2.086	2.528	2.845
21	1.721	2.080	2.518	2.831
22	1.717	2.074	2.508	2.819
23	1.714	2.069	2.500	2.807
24	1.711	2.064	2.492	2.797
25	1.708	2.060	2.485	2.787
26	1.706	2.056	2.479	2.779

Critical Values for the χ^2 distribution; probabilities $P[\chi^2 > a]$.

Degrees of freedom	*Probability (level of significance)*							
	0.05	*0.025*	*0.01*	*0.005*	*0.995*	*0.99*	*0.975*	*0.95*
1	3.841	5.024	6.635	7.879	0.000	0.000	0.001	0.004
2	5.991	7.378	9.210	10.597	0.010	0.020	0.051	0.103
3	7.815	9.348	11.345	12.838	0.072	0.115	0.216	0.352
4	9.488	11.143	13.277	14.860	0.207	0.297	0.484	0.711
5	11.070	12.833	15.086	16.750	0.412	0.554	0.831	1.145
6	12.592	14.449	16.812	18.548	0.676	0.872	1.237	1.635
7	14.067	16.013	18.475	20.278	0.989	1.239	1.690	2.167
8	15.507	17.535	20.090	21.955	1.344	1.646	2.180	2.733
9	16.919	19.023	21.666	23.589	1.735	2.088	2.700	3.325
10	18.307	20.483	23.209	25.188	2.156	2.558	3.247	3.940
11	19.675	21.920	24.725	26.757	2.603	3.053	3.816	4.575
12	21.026	23.337	26.217	28.300	3.074	3.571	4.404	5.226
13	22.362	24.736	27.688	29.819	3.565	4.107	5.009	5.892
14	23.685	26.119	29.141	31.319	4.075	4.660	5.629	6.571
15	24.996	27.488	30.578	32.801	4.601	5.229	6.262	7.261
16	26.296	28.845	32.000	34.267	5.142	5.812	6.908	7.962
17	27.587	30.191	33.409	35.718	5.697	6.408	7.564	8.672
18	28.869	31.526	34.805	37.156	6.265	7.015	8.231	9.390
19	30.144	32.852	36.191	38.582	6.844	7.633	8.907	10.117
20	31.410	34.170	37.566	39.997	7.434	8.260	9.591	10.851
21	32.671	35.479	38.932	41.401	8.034	8.897	10.283	11.591
22	33.924	36.781	40.289	42.796	8.643	9.542	10.982	12.338
23	35.172	38.076	41.638	44.181	9.260	10.196	11.689	13.091
24	36.415	39.364	42.980	45.559	9.886	10.856	12.401	13.848

Critical Values for the F-distribution for level of significance 0.05

	Numerator degrees of freedom							
Denominator degrees of freedom.	1	2	3	4	5	6	7	8
1	161.448	199.500	215.707	224.583	230.162	233.986	236.768	238.883
2	18.513	19.000	19.164	19.247	19.296	19.330	19.353	19.371
3	10.128	9.552	9.277	9.117	9.013	8.941	8.887	8.845
4	7.709	6.944	6.591	6.388	6.256	6.163	6.094	6.041
5	6.608	5.786	5.409	5.192	5.050	4.950	4.876	4.818
6	5.987	5.143	4.757	4.534	4.387	4.284	4.207	4.147
7	5.591	4.737	4.347	4.120	3.972	3.866	3.787	3.726
8	5.318	4.459	4.066	3.838	3.687	3.581	3.500	3.438
9	5.117	4.256	3.863	3.633	3.482	3.374	3.293	3.230
10	4.965	4.103	3.708	3.478	3.326	3.217	3.135	3.072
11	4.844	3.982	3.587	3.357	3.204	3.095	3.012	2.948
12	4.747	3.885	3.490	3.259	3.106	2.996	2.913	2.849
13	4.667	3.806	3.411	3.179	3.025	2.915	2.832	2.767
14	4.600	3.739	3.344	3.112	2.958	2.848	2.764	2.699
15	4.543	3.682	3.287	3.056	2.901	2.790	2.707	2.641
16	4.494	3.634	3.239	3.007	2.852	2.741	2.657	2.591
17	4.451	3.592	3.197	2.965	2.810	2.699	2.614	2.548
18	4.414	3.555	3.160	2.928	2.773	2.661	2.577	2.510
19	4.381	3.522	3.127	2.895	2.740	2.628	2.544	2.477
20	4.351	3.493	3.098	2.866	2.711	2.599	2.514	2.447
21	4.325	3.467	3.072	2.840	2.685	2.573	2.488	2.420
22	4.301	3.443	3.049	2.817	2.661	2.549	2.464	2.397
23	4.279	3.422	3.028	2.796	2.640	2.528	2.442	2.375
24	4.260	3.403	3.009	2.776	2.621	2.508	2.423	2.355
25	4.242	3.385	2.991	2.759	2.603	2.490	2.405	2.337
26	4.225	3.369	2.975	2.743	2.587	2.474	2.388	2.321
27	4.210	3.354	2.960	2.728	2.572	2.459	2.373	2.305
28	4.196	3.340	2.947	2.714	2.558	2.445	2.359	2.291
29	4.183	3.328	2.934	2.701	2.545	2.432	2.346	2.278
30	4.171	3.316	2.922	2.690	2.534	2.421	2.334	2.266
35	4.121	3.267	2.874	2.641	2.485	2.372	2.285	2.217
40	4.085	3.232	2.839	2.606	2.449	2.336	2.249	2.180
45	4.057	3.204	2.812	2.579	2.422	2.308	2.221	2.152
50	4.034	3.183	2.790	2.557	2.400	2.286	2.199	2.130