MATH577201

This question paper consists of 11 printed pages, each of which is identified by the reference **MATH5772**.

Statistical tables are provided at the end of this paper. Only approved basic scientific calculators may be used.

Examination for the Module MATH5772
(May / June 2005)

Multivariate and Cluster Analysis

Time allowed: **3 hours**

Attempt not more than FOUR questions.
All questions carry equal marks.

1. (a) Let $\boldsymbol{x}$ be a p-dimensional random vector with $E(\boldsymbol{x}) = \boldsymbol{\mu}$ and $\text{var}(\boldsymbol{x}) = \Sigma = (\sigma_{ij})$. Let $\mathcal{P} = (\rho_{ij})$ be the correlation matrix of $\boldsymbol{x}$. Define the correlations ρ_{ij} in terms of the covariances σ_{ij} and also write down a matrix equation defining $\mathcal{P}$ in terms of Σ and a diagonal matrix Δ.

Use your matrix equation to show that $\mathcal{P}$ is positive semi-definite. You should prove any results about Σ that you use.

(b) Let $\boldsymbol{x} \sim N_p(\boldsymbol{\mu}, \Sigma)$. Write down expressions for the squared Euclidean, Pearson, and Mahalanobis distances between $\boldsymbol{x}$ and $\boldsymbol{\mu}$. Briefly explain the similarities and differences between these distance measures.

Show that the squared Mahalanobis distance between $\boldsymbol{x}$ and $\boldsymbol{\mu}$ has a χ^2_p distribution.

(c) Consider two random vectors $\boldsymbol{x} \sim N_2(\boldsymbol{\mu}_x, \Sigma)$ and $\boldsymbol{y} \sim N(\boldsymbol{\mu}_y, \Sigma)$ with densities $f_{\boldsymbol{x}}(\boldsymbol{x})$ and $f_{\boldsymbol{y}}(\boldsymbol{y})$ respectively. Given that $\boldsymbol{\mu}_x = [1, 5]^T$ and $\boldsymbol{\mu}_y = [5, 1]^T$, sketch the contours of the densities of $\boldsymbol{x}$ and $\boldsymbol{y}$ on one graph.

Let $\boldsymbol{z}$ be a finite mixture distribution with density

$$f_{\boldsymbol{z}}(\boldsymbol{z}) = \frac{1}{2} f_{\boldsymbol{x}}(\boldsymbol{x}) + \frac{1}{2} f_{\boldsymbol{y}}(\boldsymbol{y}).$$

Find $E(\boldsymbol{z})$ and give the structure of $\text{var}(\boldsymbol{z})$ (note that you do not have to find an explicit form for $\text{var}(\boldsymbol{z})$). Use the structure of $\text{var}(\boldsymbol{z})$ to explain why the Mahalanobis distance is generally a poor choice for use in hierarchical cluster analysis.

(d) Find the change in the squared Euclidean, Pearson, and Mahalanobis distances between $\boldsymbol{x}$ and $\boldsymbol{\mu}$ if the random vector $\boldsymbol{x}$ has distribution $N_p(\boldsymbol{\mu}, k\Sigma)$ instead of $N_p(\boldsymbol{\mu}, \Sigma)$ for some positive real constant k.

(e) Briefly explain how the result from part (b), that the Mahalanobis distance between $\boldsymbol{x}$ and $\boldsymbol{\mu}$ has a χ^2_p distribution, can be used to detect outliers in a sample of data.

The table below presents data $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{30}$ on the thickness of cork bark measured on the north, east, south, and west sides of 30 trees, along with transformed data data $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_{30}$ where $\boldsymbol{y}_r = A\boldsymbol{x}_r$ with

$$A = \begin{bmatrix} 1 & -1 & 1 & -1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix}.$$

Also tabulated are the Mahalanobis distances $d^2_M(\boldsymbol{x}_r, \bar{\boldsymbol{x}})$ between $\boldsymbol{x}_r$ and $\bar{\boldsymbol{x}}$, and $d^2_M(\boldsymbol{y}_r, \bar{\boldsymbol{y}})$ between $\boldsymbol{y}_r$ and $\bar{\boldsymbol{y}}$, calculated using the sample variance matrices. Use these Mahalanobis distances to comment on any outliers in the original and transformed data sets. What effect has the transformation had?

North (x_1)	East (x_2)	South (x_3)	West (x_4)	$d^2_M(\boldsymbol{x}_r, \bar{\boldsymbol{x}})$	y_1	y_2	y_3	$d^2_M(\boldsymbol{y}_r, \bar{\boldsymbol{y}})$
59.5	55.3	73.5	59.7	18.3	18.0	-13.9	-4.3	9.2
70.4	55.2	72.4	54.1	17.0	33.5	-2.1	1.1	9.9
59.0	54.5	49.0	49.1	8.4	4.5	10.1	5.4	1.8
48.9	40.4	44.2	43.3	3.6	9.5	4.7	-2.9	2.8
42.2	41.8	35.4	32.4	6.4	3.4	6.8	9.5	0.9
70.4	65.4	64.9	53.6	12.7	16.3	5.4	11.8	3.8
37.6	33.0	35.7	31.5	7.7	8.9	1.9	1.5	0.4
45.3	44.4	40.1	37.7	3.1	3.3	5.1	6.7	0.2
33.0	37.8	32.1	33.0	10.7	-5.7	0.8	4.8	2.5
59.1	52.9	59.3	59.1	8.4	6.4	-0.2	-6.2	2.6
40.1	35.8	38.0	42.1	5.1	0.2	2.1	-6.3	3.0
45.4	45.2	47.5	51.3	12.1	-3.5	-2.1	-6.1	3.6
49.5	40.9	52.9	48.5	4.4	13.0	-3.4	-7.6	3.8
46.1	45.6	43.0	35.5	7.7	8.1	3.1	10.1	1.6
31.7	34.7	19.5	22.5	15.1	-6.0	12.2	12.2	2.9
45.3	38.7	54.4	47.0	10.1	14.0	-9.1	-8.2	5.6
47.1	41.4	37.6	43.2	8.0	0.1	9.5	-1.7	3.8
59.7	53.8	60.9	50.3	2.8	16.5	-1.1	3.5	1.9
53.3	51.0	54.4	51.8	4.1	4.9	-1.1	-0.8	0.8
49.8	41.6	45.3	41.8	3.0	11.7	4.4	-0.2	2.0
48.3	46.6	42.4	43.4	3.0	0.6	5.9	3.2	0.3
49.8	38.0	44.5	42.0	10.0	14.3	5.3	-4.0	5.5
65.3	58.2	64.4	56.1	3.3	15.4	0.9	2.1	1.5
60.3	52.4	66.8	58.0	5.2	16.8	-6.5	-5.6	4.1
68.1	57.5	70.2	55.9	6.5	24.9	-2.1	1.7	4.9
70.0	57.3	70.7	61.0	6.7	22.4	-0.7	-3.7	5.4
34.0	25.2	30.4	31.1	15.9	8.1	3.5	-5.9	3.8
55.8	45.1	54.6	56.1	11.0	9.2	1.2	-11.0	7.2
29.3	29.0	25.8	35.3	11.6	-9.2	3.5	-6.3	4.8
73.8	62.6	61.4	58.7	20.5	13.9	12.4	4.0	7.4

CONTINUED...

2. (a) Let $\{\boldsymbol{x}_r; r = 1, \ldots, n\}$ be an independent sample from a $N_p(\boldsymbol{\mu}, \Sigma)$ distribution. Show that the log-likelihood of the parameters $\boldsymbol{\mu}$ and Σ given the data matrix X is

$$l(\boldsymbol{\mu}, \Sigma | X) = -\frac{n}{2} \log |\Sigma| - \frac{1}{2} \sum_{r=1}^{n} (\boldsymbol{x}_r - \boldsymbol{\mu})^T \Sigma^{-1} (\boldsymbol{x}_r - \boldsymbol{\mu}) + C,$$

giving an explicit expression for for the constant C. By writing $(\boldsymbol{x}_r - \boldsymbol{\mu})$ as $(\boldsymbol{x}_r - \bar{\boldsymbol{x}} + \bar{\boldsymbol{x}} - \boldsymbol{\mu})$, show that the maximum likelihood estimate of $\boldsymbol{\mu}$ is $\widehat{\boldsymbol{\mu}} = \bar{\boldsymbol{x}}$.

(b) Consider testing $H_0 : \boldsymbol{\mu} = \mathbf{0}$ against $H_1 : \boldsymbol{\mu} \neq \mathbf{0}$ with Σ known. Show that the likelihood ratio test statistic λ is equal to the the Mahalanobis distance between $\bar{\boldsymbol{x}}$ and $\mathbf{0}$ with respect to Σ/n, i.e.

$$\lambda = n\bar{\boldsymbol{x}}^T \Sigma^{-1} \bar{\boldsymbol{x}},$$

and state its distribution under H_0.

(c) Now consider a two sample problem with independent samples $\{\boldsymbol{x}_r; r = 1, \ldots, n_x\}$ and $\{\boldsymbol{y}_s; s = 1, \ldots, n_y\}$ from $N_p(\boldsymbol{\mu}_x, \Sigma)$ and $N_p(\boldsymbol{\mu}_y, \Sigma)$ distributions respectively, where Σ is a known variance matrix common to the two distributions.

Let $\boldsymbol{d} = \bar{\boldsymbol{x}} - \bar{\boldsymbol{y}}$. Write down the distribution of $\boldsymbol{d}$. Construct a test statistic based on a quadratic form involving $\boldsymbol{d}$ and Σ to test $H_0 : \boldsymbol{\mu}_x = \boldsymbol{\mu}_y$ against $H_1 : \boldsymbol{\mu}_x \neq \boldsymbol{\mu}_y$. State the distribution of this test statistic when H_0 is true.

Explain briefly what changes are needed to the test statistic and its distribution under H_0 when Σ is not known and must be estimated from the data.

(d) Measurements of cranial length and breadth in millimetres were made on $n_x = 14$ male and $n_y = 35$ female frogs. Sample means and sums of squares matrices were

$$\bar{\boldsymbol{x}} = \begin{bmatrix} 21.8 \\ 22.8 \end{bmatrix}, \qquad \bar{\boldsymbol{y}} = \begin{bmatrix} 22.9 \\ 24.4 \end{bmatrix},$$

$$(n_x - 1)S_x = \begin{bmatrix} 240 & 248 \\ 248 & 270 \end{bmatrix}, \qquad (n_y - 1)S_y = \begin{bmatrix} 601 & 690 \\ 690 & 830 \end{bmatrix}.$$

Carry out a hypothesis test to determine whether male and female frogs have significantly different mean cranial measurements.

Regardless of whether you find a significant difference between the two samples, construct 95% simultaneous confidence intervals for the unit co-ordinate vectors and interpret your findings.

Hint You may find it useful to recall that the T^2 and F distributions are related by the equation

$$T^2(p, m) = \frac{mp}{m - p + 1} F(p, m - p + 1).$$

CONTINUED...

3. (a) Suppose A is a $p \times p$ symmetric matrix with ordered distinct eigenvalues $\lambda_1 > \lambda_2 > \cdots > \lambda_p$. Let $\boldsymbol{x}_i$ be the unit eigenvector corresponding to λ_i for $i = 1, \ldots, n$. Show that $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ are orthogonal if $i \neq j$.

Prove that A can be written as $A = \Gamma\Lambda\Gamma^T$ where Γ is an orthogonal matrix and Λ is a diagonal matrix. Show how Γ and Λ can be constructed from the eigenvalues and eigenvectors of A.

(b) Let $\boldsymbol{x}$ be a random vector with mean vector $\boldsymbol{\mu}$ and variance matrix Σ, where Σ has spectral decomposition $\Sigma = \Gamma\Lambda\Gamma^T$. Write down an equation defining the principal component transformation which transforms $\boldsymbol{x}$ into a new random vector $\boldsymbol{y}$.

Find $E(\boldsymbol{y})$ and $\mathrm{var}(\boldsymbol{y})$, noting any special structure of the diagonal elements of $\mathrm{var}(\boldsymbol{y})$.

Given a data matrix X, explain how a sample principal component analysis can be carried out. Why is this sample principal component analysis useful in reducing the effective dimension of the data?

(In your answer, you can use without proof the fact that $\boldsymbol{a}^T\Sigma\boldsymbol{a}$ is maximised over all possible unit vectors $\boldsymbol{a}$ when $\boldsymbol{a}$ is the eigenvector of Σ corresponding to the largest eigenvalue of Σ. Any further results that you assume about variance maximisation should be clearly stated but need not be proved.)

(c) Show that the eigenvalues of X^TX are also eigenvalues of XX^T, where X is an $n \times p$ data matrix. When does this imply that principal component analysis can be used in place of multi-dimensional scaling, and why might we prefer to use principal components?

(d) In a study of beef cattle, the length, height, girth, and percentage body fat of 150 Normandy cows were measured. A principal components analysis of these data using the correlation matrix in R gave the following output, where "`pc.bdy.ft`" is the variable "percentage of body fat".

```
> summary(prc)
Importance of components:
                       Comp.1  Comp.2  Comp.3  Comp.4
Standard deviation      1.388   1.006   0.920   0.462
Proportion of Variance 0.482   0.253   0.212   0.053
Cumulative Proportion  0.482   0.735   0.947   1.000

> loadings(prc)

Loadings:
            Comp.1 Comp.2 Comp.3 Comp.4
length      0.537  0.223  0.631 -0.513
girth       0.677 -0.037  0.034  0.743
height      0.498 -0.052 -0.753 -0.427
pc.bdy.fat -0.071  0.973 -0.183  0.123
```

Explain what information is given by this output. Suggest an interpretation of the principal components analysis. How might you use the above output in any further analysis of the data?

4. (a) Suppose p-dimensional measurements, represented as a random vector $\boldsymbol{x}$, can be made on individuals from two populations Π_1 and Π_2 with probability density functions $f_1(\boldsymbol{x})$ and $f_2(\boldsymbol{x})$ respectively. Briefly distinguish between the situations when cluster analysis or discriminant analysis might be applied to data obtained from the two populations.

(b) In discriminant analysis, the maximum likelihood allocation rule defines two regions, R_1 and R_2 in $\mathbb{R}^p$. State how the pdfs $f_1(\boldsymbol{x})$ and $f_2(\boldsymbol{x})$ are used to define R_1 and R_2, and explain how these regions are used to construct the allocation rule.

Now assume that individuals from Π_i follow a $N_p(\boldsymbol{\mu}_i, \Sigma)$ distribution for $i = 1, 2$. Show that in this case the region R_1 is

$$R_1 = \left\{ \boldsymbol{x} : (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^T \Sigma^{-1} \left[\boldsymbol{x} - \frac{1}{2}(\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2) \right] \geqslant 0 \right\}.$$

For the bivariate case, sketch a diagram showing how the means $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ and the contours of the density functions $f_1(\boldsymbol{x})$ and $f_2(\boldsymbol{x})$ relate to the line separating R_1 and R_2.

(c) The classification probabilities of the allocation rule in (b) are defined to be $p_{ij} = P(\boldsymbol{x} \text{ allocated to } \Pi_i | \boldsymbol{x} \text{ from } \Pi_j)$, $i, j = 1, 2$. Show that the misclassification probabilities are given by $p_{12} = p_{21} = \Phi(-\Delta/2)$, where Φ is the cumulative distribution function of a standard univariate normal and $\Delta^2 = (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^T \Sigma^{-1} (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)$.

(d) Suppose we have measurements on length and breath in millimetres of 50 human skulls from pre-dynastic Egypt (population 1) and 50 skulls from Roman-era Egypt (population 2). The mean vectors and pooled covariance matrix are

$$\bar{\boldsymbol{x}}_1 = \begin{bmatrix} 134 \\ 132 \end{bmatrix} \qquad \bar{\boldsymbol{x}}_2 = \begin{bmatrix} 128 \\ 136 \end{bmatrix} \qquad S = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}.$$

Find the equation of the line dividing R_1 and R_2, and use parametric substitution to evaluate the misclassification probabilities p_{12} and p_{21}. In a large sample of skulls, what proportion would you expect to be misclassified?

If you were only able to use one of the variables (either length or breadth) to classify skulls, which would you choose to minimise the number of misclassifications? Why is your chosen measurement likely to result in fewer misclassified skulls?

To which population would you assign a skull with height 135mm, breadth 133mm? Why can you not use your misclassification probabilities above to say how confident you are in your classification of this skull?

5. Given an $n \times n$ symmetric matrix of dissimilarities, $\Delta = (\delta_{rs})$, classical multidimensional scaling is used to find an $n \times p$ matrix X, where the Euclidean distances between the points $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ specified by the rows of X match the dissimilarities as closely as possible. The classical multidimensional scaling solution is produced by the following algorithm.

(i) Construct $A = (a_{rs})$ where $a_{rs} = -\frac{1}{2}\delta_{rs}^2$.

(ii) Set $B = HAH$, where $H = I_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T$ is the centring matrix.

(iii) Take the spectral decomposition $B = \Gamma\Lambda\Gamma^T$ of B, where Γ and Λ are constructed from the eigenvectors and eigenvalues of B in the usual way.

(iv) Set $\Gamma_1 = \left[\boldsymbol{\gamma}_{(1)}, \ldots, \boldsymbol{\gamma}_{(p)}, \mathbf{0}, \ldots, \mathbf{0}\right]$ and $\Lambda_1 = \mathrm{diag}(\lambda_1, \ldots, \lambda_p, 0, \ldots, 0)$.

(v) Let $X = \Gamma_1\Lambda_1^{1/2}$.

(a) Show that Δ is a Euclidean distance matrix if and only if B is positive semi-definite. (Recall that B is positive semi-definite if $\boldsymbol{c}^T B\boldsymbol{c} \geqslant 0$ for all $\boldsymbol{c} \in \mathbb{R}^n$.)

(b) A group of political science lecturers was asked to assess the dissimilarities on a scale of 1 (very similar) to 9 (very dissimilar) between certain political leaders prominent at the time of the second world war. The dissimilarity matrix obtained for (in order) Hitler, Mussolini, Churchill, Roosevelt, Stalin, and Attlee is given by

$$\Delta = \begin{bmatrix} 0 & 2 & 7 & 8 & 5 & 9 \\ 2 & 0 & 8 & 8 & 8 & 9 \\ 7 & 8 & 0 & 3 & 5 & 8 \\ 8 & 8 & 3 & 0 & 8 & 7 \\ 5 & 8 & 5 & 8 & 0 & 7 \\ 9 & 9 & 8 & 7 & 7 & 0 \end{bmatrix}.$$

Is it possible to construct a set of points which has Euclidean distance matrix Δ? How many dimensions do you think would be best to represent this dissimilarity matrix? You may find the following edited *R* output helpful (where Δ is represented as `Delta`).

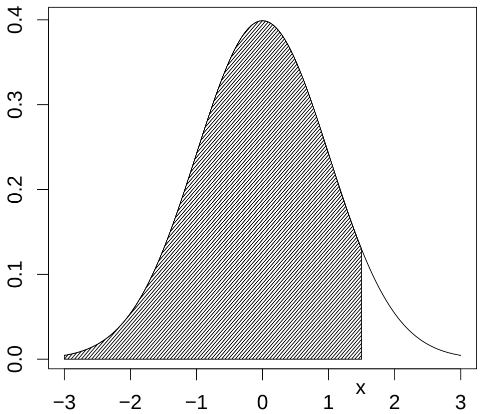
```
> A = -0.5 * Delta^2
> H = diag(rep(1, 6)) - 1/6 * rep(1, 6) %*% t(rep(1, 6))
> B = H %*% A %*% H
> tmp = eigen(B)
> round(tmp$values, 1)
[1] 63.8 37.4 31.5  0.0 -0.9 -5.8
> round(tmp$vectors, 2)
           [,1]  [,2]  [,3]  [,4]  [,5]  [,6]
Hitler     0.53  0.02  0.10 -0.41  0.52  0.53
Mussolini  0.56 -0.03 -0.44 -0.41 -0.46 -0.34
Churchill -0.26 -0.47  0.30 -0.41 -0.53  0.42
Roosevelt -0.39 -0.48 -0.33 -0.41  0.49 -0.33
Stalin     0.00  0.28  0.70 -0.41  0.06 -0.51
Attlee    -0.44  0.69 -0.33 -0.41 -0.08  0.23
```

Construct a set of points in two-dimensional space representing these political leaders, sketch a plot of these points, and interpret your plot.

Normal Distribution Function Tables

The first table gives

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2} dt$$



and this corresponds to the shaded area in the figure to the right. $\Phi(x)$ is the probability that a random variable, normally distributed with zero mean and unit variance, will be less than or equal to x. When $x < 0$ use $\Phi(x) = 1 - \Phi(-x)$, as the normal distribution with mean zero is symmetric about zero. To interpolate, use the formula

$$\Phi(x) \approx \Phi(x_1) + \frac{x - x_1}{x_2 - x_1} \big(\Phi(x_2) - \Phi(x_1)\big)$$

Table 1

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0.00	0.5000	**0.50**	0.6915	**1.00**	0.8413	**1.50**	0.9332	**2.00**	0.9772	**2.50**	0.9938
0.05	0.5199	**0.55**	0.7088	**1.05**	0.8531	**1.55**	0.9394	**2.05**	0.9798	**2.55**	0.9946
0.10	0.5398	**0.60**	0.7257	**1.10**	0.8643	**1.60**	0.9452	**2.10**	0.9821	**2.60**	0.9953
0.15	0.5596	**0.65**	0.7422	**1.15**	0.8749	**1.65**	0.9505	**2.15**	0.9842	**2.65**	0.9960
0.20	0.5793	**0.70**	0.7580	**1.20**	0.8849	**1.70**	0.9554	**2.20**	0.9861	**2.70**	0.9965
0.25	0.5987	**0.75**	0.7734	**1.25**	0.8944	**1.75**	0.9599	**2.25**	0.9878	**2.75**	0.9970
0.30	0.6179	**0.80**	0.7881	**1.30**	0.9032	**1.80**	0.9641	**2.30**	0.9893	**2.80**	0.9974
0.35	0.6368	**0.85**	0.8023	**1.35**	0.9115	**1.85**	0.9678	**2.35**	0.9906	**2.85**	0.9978
0.40	0.6554	**0.90**	0.8159	**1.40**	0.9192	**1.90**	0.9713	**2.40**	0.9918	**2.90**	0.9981
0.45	0.6736	**0.95**	0.8289	**1.45**	0.9265	**1.95**	0.9744	**2.45**	0.9929	**2.95**	0.9984
0.50	0.6915	**1.00**	0.8413	**1.50**	0.9332	**2.00**	0.9772	**2.50**	0.9938	**3.00**	0.9987

The inverse function $\Phi^{-1}(p)$ is tabulated below for various values of p.

Table 2

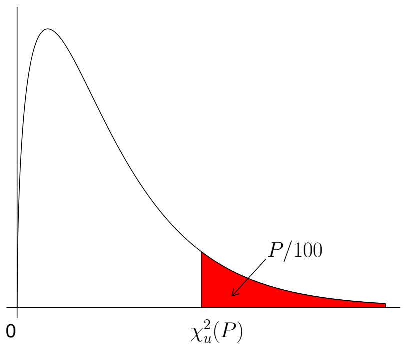
p	**0.900**	**0.950**	**0.975**	**0.990**	**0.995**	**0.999**	**0.9995**
$\Phi^{-1}(p)$	1.2816	1.6449	1.9600	2.3263	2.5758	3.0902	3.2905

Percentage Points of the χ^2-Distribution

This table gives the percentage points $\chi^2_\nu(P)$ for various values of P and degrees of freedom ν, as indicated by the figure to the right.

If X is a variable distributed as χ^2 with ν degrees of freedom, $P/100$ is the probability that $X \geq \chi^2_\nu(P)$.

For $\nu > 100$, $\sqrt{2X}$ is approximately normally distributed with mean $\sqrt{2\nu - 1}$ and unit variance.



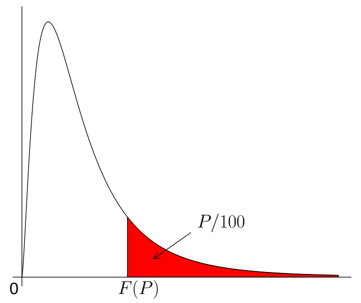
	Percentage points P						
ν	**10**	**5**	**2.5**	**1**	**0.5**	**0.1**	**0.05**
1	2.706	3.841	5.024	6.635	7.879	10.828	12.116
2	4.605	5.991	7.378	9.210	10.597	13.816	15.202
3	6.251	7.815	9.348	11.345	12.838	16.266	17.730
4	7.779	9.488	11.143	13.277	14.860	18.467	19.997
5	9.236	11.070	12.833	15.086	16.750	20.515	22.105
6	10.645	12.592	14.449	16.812	18.548	22.458	24.103
7	12.017	14.067	16.013	18.475	20.278	24.322	26.018
8	13.362	15.507	17.535	20.090	21.955	26.124	27.868
9	14.684	16.919	19.023	21.666	23.589	27.877	29.666
10	15.987	18.307	20.483	23.209	25.188	29.588	31.420
11	17.275	19.675	21.920	24.725	26.757	31.264	33.137
12	18.549	21.026	23.337	26.217	28.300	32.909	34.821
13	19.812	22.362	24.736	27.688	29.819	34.528	36.478
14	21.064	23.685	26.119	29.141	31.319	36.123	38.109
15	22.307	24.996	27.488	30.578	32.801	37.697	39.719
16	23.542	26.296	28.845	32.000	34.267	39.252	41.308
17	24.769	27.587	30.191	33.409	35.718	40.790	42.879
18	25.989	28.869	31.526	34.805	37.156	42.312	44.434
19	27.204	30.144	32.852	36.191	38.582	43.820	45.973
20	28.412	31.410	34.170	37.566	39.997	45.315	47.498
25	34.382	37.652	40.646	44.314	46.928	52.620	54.947
30	40.256	43.773	46.979	50.892	53.672	59.703	62.162
40	51.805	55.758	59.342	63.691	66.766	73.402	76.095
50	63.167	67.505	71.420	76.154	79.490	86.661	89.561
80	96.578	101.879	106.629	112.329	116.321	124.839	128.261

CONTINUED...

5 Percent Points of the F-Distribution

This table gives the percentage points $F_{\nu_1,\nu_2}(P)$ for $P = 0.05$ and degrees of freedom ν_1, ν_2, as indicated by the figure to the right.
The lower percentage points, that is the values $F'_{\nu_1,\nu_2}(P)$ such that the probability that $F \le F'_{\nu_1,\nu_2}(P)$ is equal to $P/100$, may be found using the formula

$$F'_{\nu_1,\nu_2}(P) = 1/F_{\nu_1,\nu_2}(P)$$



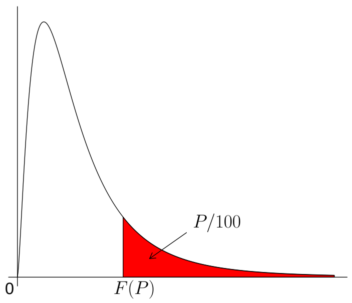
	ν_1								
ν_2	**1**	**2**	**3**	**4**	**5**	**6**	**12**	**24**	∞
2	18.513	19.000	19.164	19.247	19.296	19.330	19.413	19.454	19.496
3	10.128	9.552	9.277	9.117	9.013	8.941	8.745	8.639	8.526
4	7.709	6.944	6.591	6.388	6.256	6.163	5.912	5.774	5.628
5	6.608	5.786	5.409	5.192	5.050	4.950	4.678	4.527	4.365
6	5.987	5.143	4.757	4.534	4.387	4.284	4.000	3.841	3.669
7	5.591	4.737	4.347	4.120	3.972	3.866	3.575	3.410	3.230
8	5.318	4.459	4.066	3.838	3.687	3.581	3.284	3.115	2.928
9	5.117	4.256	3.863	3.633	3.482	3.374	3.073	2.900	2.707
10	4.965	4.103	3.708	3.478	3.326	3.217	2.913	2.737	2.538
11	4.844	3.982	3.587	3.357	3.204	3.095	2.788	2.609	2.404
12	4.747	3.885	3.490	3.259	3.106	2.996	2.687	2.505	2.296
13	4.667	3.806	3.411	3.179	3.025	2.915	2.604	2.420	2.206
14	4.600	3.739	3.344	3.112	2.958	2.848	2.534	2.349	2.131
15	4.543	3.682	3.287	3.056	2.901	2.790	2.475	2.288	2.066
16	4.494	3.634	3.239	3.007	2.852	2.741	2.425	2.235	2.010
17	4.451	3.592	3.197	2.965	2.810	2.699	2.381	2.190	1.960
18	4.414	3.555	3.160	2.928	2.773	2.661	2.342	2.150	1.917
19	4.381	3.522	3.127	2.895	2.740	2.628	2.308	2.114	1.878
20	4.351	3.493	3.098	2.866	2.711	2.599	2.278	2.082	1.843
25	4.242	3.385	2.991	2.759	2.603	2.490	2.165	1.964	1.711
30	4.171	3.316	2.922	2.690	2.534	2.421	2.092	1.887	1.622
40	4.085	3.232	2.839	2.606	2.449	2.336	2.003	1.793	1.509
50	4.034	3.183	2.790	2.557	2.400	2.286	1.952	1.737	1.438
100	3.936	3.087	2.696	2.463	2.305	2.191	1.850	1.627	1.283
∞	3.841	2.996	2.605	2.372	2.214	2.099	1.752	1.517	1.002

CONTINUED...

1 Percent Points of the F-Distribution

This table gives the percentage points $F_{\nu_1,\nu_2}(P)$ for $P = 0.01$ and degrees of freedom ν_1, ν_2, as indicated by the figure to the right.
The lower percentage points, that is the values $F'_{\nu_1,\nu_2}(P)$ such that the probability that $F \leq F'_{\nu_1,\nu_2}(P)$ is equal to $P/100$, may be found using the formula

$$F'_{\nu_1,\nu_2}(P) = 1/F_{\nu_1,\nu_2}(P)$$



	ν_1								
ν_2	**1**	**2**	**3**	**4**	**5**	**6**	**12**	**24**	**∞**
2	98.503	99.000	99.166	99.249	99.299	99.333	99.416	99.458	99.499
3	34.116	30.817	29.457	28.710	28.237	27.911	27.052	26.598	26.125
4	21.198	18.000	16.694	15.977	15.522	15.207	14.374	13.929	13.463
5	16.258	13.274	12.060	11.392	10.967	10.672	9.888	9.466	9.020
6	13.745	10.925	9.780	9.148	8.746	8.466	7.718	7.313	6.880
7	12.246	9.547	8.451	7.847	7.460	7.191	6.469	6.074	5.650
8	11.259	8.649	7.591	7.006	6.632	6.371	5.667	5.279	4.859
9	10.561	8.022	6.992	6.422	6.057	5.802	5.111	4.729	4.311
10	10.044	7.559	6.552	5.994	5.636	5.386	4.706	4.327	3.909
11	9.646	7.206	6.217	5.668	5.316	5.069	4.397	4.021	3.602
12	9.330	6.927	5.953	5.412	5.064	4.821	4.155	3.780	3.361
13	9.074	6.701	5.739	5.205	4.862	4.620	3.960	3.587	3.165
14	8.862	6.515	5.564	5.035	4.695	4.456	3.800	3.427	3.004
15	8.683	6.359	5.417	4.893	4.556	4.318	3.666	3.294	2.868
16	8.531	6.226	5.292	4.773	4.437	4.202	3.553	3.181	2.753
17	8.400	6.112	5.185	4.669	4.336	4.102	3.455	3.084	2.653
18	8.285	6.013	5.092	4.579	4.248	4.015	3.371	2.999	2.566
19	8.185	5.926	5.010	4.500	4.171	3.939	3.297	2.925	2.489
20	8.096	5.849	4.938	4.431	4.103	3.871	3.231	2.859	2.421
25	7.770	5.568	4.675	4.177	3.855	3.627	2.993	2.620	2.169
30	7.562	5.390	4.510	4.018	3.699	3.473	2.843	2.469	2.006
40	7.314	5.179	4.313	3.828	3.514	3.291	2.665	2.288	1.805
50	7.171	5.057	4.199	3.720	3.408	3.186	2.562	2.183	1.683
100	6.895	4.824	3.984	3.513	3.206	2.988	2.368	1.983	1.427
∞	6.635	4.605	3.782	3.319	3.017	2.802	2.185	1.791	1.003

END