MATH273001

This question paper consists of 11 printed pages, each of which is identified by the reference **MATH2730**.

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

Examination for the Module MATH2730
(January 2006)

ANALYSIS OF EXPERIMENTAL DATA

Time allowed: **2 hours**

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

Throughout this examination paper, replacing a subscript by a $\bullet$ denotes that the subscript has been summed over. A bar over a quantity indicates that averaging has taken place. For example, given data Y_{ij} for $i = 1, \ldots, t$ and $j = 1, \ldots, n$, use the notation

$$Y_{i\bullet} = \sum_{j=1}^{n} Y_{ij}, \qquad Y_{\bullet\bullet} = \sum_{i=1}^{t}\sum_{j=1}^{n} Y_{ij},$$

$$\overline{Y}_{i\bullet} = \frac{1}{n}\sum_{j=1}^{n} Y_{ij}, \quad \text{and} \quad \overline{Y}_{\bullet\bullet} = \frac{1}{nt}\sum_{i=1}^{t}\sum_{j=1}^{n} Y_{ij}.$$

1. Consider the one-way fixed and random effects ANOVA models

$$\text{FIXED} \qquad Y_{ij} = \mu + \alpha_i + \varepsilon_{ij} \qquad \sum_{i=1}^{t} \alpha_i = 0,$$

$$\text{RANDOM} \qquad Y_{ij} = \mu + A_i + \varepsilon_{ij} \qquad A_i \text{ i.i.d. } N(0, \sigma_A^2),$$

for $i = 1, \ldots, t$ and $j = 1, \ldots, n$. Let the total number of observations be $N = nt$. Here Y_{ij} is the jth observation in the ith group, the ε_{ij} are independently $N(0, \sigma^2)$ distributed (independent of the A_i) and represent random variation, and the α_i and A_i represent effects of the treatment groups.

(a) Briefly explain when the fixed and random effects models are appropriate. In particular, under what circumstances would you choose to use each model?

In a one-way ANOVA, we wish to test the null hypothesis that there are no treatment effects against the alternative that the treatment affects the mean response. Write down null and alternative hypotheses for the fixed and random effects models above. Your hypotheses should be in terms of the model parameters $\{\alpha_i\}$ and σ_A^2.

(b) Three experiments were carried out to study the yield of wheat in different treatment groups. For each of the following descriptions of the treatment groups, say whether a fixed or random effects model would be appropriate. Justify your choice in each case.

(i) To examine the effects of different fertilisers, each group was given a different brand of fertiliser.

(ii) All groups were given the same brand of fertiliser, but from different production batches, to check the consistency of the results from experiment (i).

(iii) In order to determine the best dosage of fertiliser to maximise cost-effectiveness, each group was given a different selected dose of fertiliser.

(c) Define the treatment and error sums of squares to be $SS_T = \sum_{i=1}^{t} n(\overline{Y}_{i\bullet} - \overline{Y}_{\bullet\bullet})^2$ and $SS_E = \sum_{i=1}^{t} \sum_{j=1}^{n} (Y_{ij} - \overline{Y}_{i\bullet})^2$ and the corresponding mean squares to be $MS_T = SS_T/(t-1)$ and $MS_E = SS_E/(N-t)$. We estimate μ by the least squares estimator $\widehat{\mu} = \overline{Y}_{\bullet\bullet}$ and σ^2 by the unbiased estimator $\widehat{\sigma}^2 = MS_E$.

(i) In the fixed effects model, the α_i are estimated by $\widehat{\alpha}_i = \overline{Y}_{i\bullet} - \overline{Y}_{\bullet\bullet}$ for $i = 1, \ldots, t$. Show that the $\widehat{\alpha}_i$ are unbiased estimators.

(ii) For the random effects model, $E(MS_T) = n\sigma_A^2 + \sigma^2$. Construct an unbiased estimator for σ_A^2 and show that your estimator is unbiased.

(d) A national supermarket chain wanted to determine whether typical sales were different in different stores. Five stores were chosen randomly from across the country. In each store, the values of eleven randomly chosen trolley loads of shopping were recorded.

Given that the sums of squares calculated from these data were $SS_T = 541.17$ and $SS_E = 933.14$, construct an ANOVA table to test the null hypothesis that there is no significant difference in the mean spend in the stores. Assuming that a random effects model is appropriate, estimate σ^2 and σ_A^2. Comment on your findings.

2. Consider the one-way fixed effects ANOVA model

$$Y_{ij} = \mu + \alpha_i + \varepsilon_{ij} \qquad i = 1, \ldots, t; \;\; j = 1, \ldots, n; \;\; \varepsilon_{ij} \text{ i.i.d. } N(0, \sigma^2).$$

If the ANOVA null hypothesis H_0:$\alpha_i = 0$ for all i is rejected, we know that some treatments are not equivalent. One way to investigate which treatments are significantly different is by using contrasts. Define a contrast to be a linear combination $C = \sum_i c_i Y_{i\bullet}$ with $\sum_i c_i = 0$. We use C to test the hypothesis H_0:$\sum_i c_i \mu_i = 0$ against H_A:$\sum_i c_i \mu_i \neq 0$, where $\mu_i = \mu + \alpha_i$.

(a) Let the sum of squares for the contrast C be

$$SS_C = \frac{C^2}{n \sum_i c_i^2}.$$

Show that, under H_0:$\sum_i c_i \mu_i = 0$, the distribution of SS_C is given by $SS_C/\sigma^2 \sim \chi_1^2$.
Hint: First write down the distribution of $Y_{i\bullet}$ and use this to find the distribution of C. Write down the expectation and variance of this distribution when H_0 is true. Then use the fact that if $Z \sim N(0,1)$, we have $Z^2 \sim \chi_1^2$ by definition.

(b) A manufacturer needs to outsource the production of a chemical. Before deciding on a supplier, the manufacturer asks four laboratories to manufacture five batches each. A numeric measurement of quality is assigned to each batch and these values are shown below. Given that $\sum_{ij} y_{ij}^2 = 319.9826$ and $\sum_i y_{i\bullet}^2 = 1599.53$, construct an ANOVA table and test the null hypothesis of no difference between laboratories. Comment on your results.

Quality measurements on batches of chemicals

Laboratory	Quality measurements					$y_{i\bullet}$
1	4.13	4.07	4.04	4.07	4.05	20.36
2	3.86	3.85	4.08	4.11	4.08	19.98
3	4.00	4.02	4.01	4.01	4.04	20.08
4	3.88	3.89	3.91	3.96	3.92	19.56

(c) Before gathering the data in part (b), prior experience suggested that laboratories one and three would be equivalent, as would laboratories two and four, although laboratories one and three would be different to laboratories two and four.

Construct a set of contrasts to test whether these suppositions are correct. Carry out your tests using a 5% significance level. Comment on your findings.

3. Consider the two-way fixed effects ANOVA model $Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk}$ for $i = 1, \ldots, a$, $j = 1, \ldots, b$ and $k = 1, \ldots, n$. Here, α_i represents level i of factor A, β_j represents level j of factor B, γ_{ij} represents the interaction between levels i and j of factors A and B respectively, and the ε_{ijk} are independently $N(0, \sigma^2)$ distributed representing random variation.

(a) List a set of constraints commonly applied to the parameters $\{\alpha_i\}$, $\{\beta_j\}$ and $\{\gamma_{ij}\}$. With reference to the number of cells or groups in the data, explain why these (or other) constraints must be applied.

(b) By minimising a suitable sum of squares, show that μ is estimated by $\widehat{\mu} = \overline{Y}_{\bullet\bullet\bullet}$ and derive least squares estimates $\{\widehat{\alpha_i}\}$ of the parameters $\{\alpha_i\}$ for $i = 1, \ldots, a$. Write down without proof least squares estimates of the parameters $\{\beta_j\}$ and $\{\gamma_{ij}\}$.

(c) A study was carried out to assess the performance of four different designs of air-conditioning unit in different regions of the U.S.A. Two air conditioning units of each design were installed in each region, thus obtaining a pair of observations for each combination of region and design. The time to failure of the units in months are given below. Cell sums $y_{ij\bullet}$, row sums $y_{i\bullet\bullet}$, column sums $y_{\bullet j\bullet}$ and the grand total $y_{\bullet\bullet\bullet}$ are also provided, all in italic type.

Time to failure of air-conditioning units

Region	Design												$y_{i\bullet\bullet}$
	A			B			C			D			
	Data		$y_{i1\bullet}$	Data		$y_{i2\bullet}$	Data		$y_{i3\bullet}$	Data		$y_{i4\bullet}$	
Northeast	58	49	*107*	35	24	*59*	72	60	*132*	61	64	*125*	*423*
Southeast	40	38	*78*	18	22	*40*	54	64	*118*	38	50	*88*	*324*
Northwest	63	59	*122*	44	16	*60*	81	60	*141*	52	48	*100*	*423*
Southwest	36	29	*65*	9	13	*22*	47	52	*99*	30	41	*71*	*257*
$y_{\bullet j\bullet}$			*372*			*181*			*490*			*384*	*1427*

Analysing the data in *R* produced the following edited output, where `ttf` refers to the time to failure.

```
> aircon.lm = lm(ttf ~ region * design)
> anova(aircon.lm)
              Df Sum Sq Mean Sq F value    Pr(>F)
region         a 2475.1       b 12.6502 0.0001714
design         3      c  2067.4 31.7002 5.768e-07
region:design  9  310.8    34.5       d 0.8325323
Residuals     16 1043.5    65.2
```

Fill in the missing values indicated by the letters `a` to `d` and interpret the resulting table.

For any terms which you believe should be included in the model, estimate the model parameters.

Given your parameter estimates, what are the predicted times to failure for each type of air conditioning unit when used in the northeastern U.S.A?

4. Consider the fixed effects nested design model

$$Y_{ijk} = \mu + \alpha_i + \beta_{j(i)} + \varepsilon_{ijk},$$

for $i = 1, \ldots, a$, $j = 1, \ldots, b$ and $k = 1, \ldots, n$. Here, α_i represents level i of factor A, $\beta_{j(i)}$ represents level j of factor B nested within level i of factor A, and the ε_{ijk} are independently $N(0, \sigma^2)$ distributed representing random variation. Assume that $\sum_i \alpha_i = 0$ and $\sum_j \beta_{j(i)} = 0$ for each $i = 1, \ldots, a$.

(a) Let the total sum of squares and sums of squares for main groups, sub groups and errors be defined by

$$SS_{TOT} = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(Y_{ijk} - \overline{Y}_{\bullet\bullet\bullet})^2, \qquad SS_{MG} = bn\sum_{i=1}^{a}(\overline{Y}_{i\bullet\bullet} - \overline{Y}_{\bullet\bullet\bullet})^2,$$

$$SS_{SG} = n\sum_{i=1}^{a}\sum_{j=1}^{b}(\overline{Y}_{ij\bullet} - \overline{Y}_{i\bullet\bullet})^2 \quad \text{and} \quad SS_E = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(Y_{ijk} - \overline{Y}_{ij\bullet})^2.$$

Show that $SS_{TOT} = SS_{MG} + SS_{SG} + SS_E$. Explain what each of the sums of squares quantities represents.

(b) Show that the error sum of squares can be written as

$$SS_E = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n} Y_{ijk}^2 - n\sum_{i=1}^{a}\sum_{j=1}^{b}\overline{Y}_{ij\bullet}^2. \tag{1}$$

Given that $E(Y_{ijk}^2) = (\mu + \alpha_i + \beta_{j(i)})^2 + \sigma^2$ and $E(\overline{Y}_{ij\bullet}^2) = (\mu + \alpha_i + \beta_{j(i)})^2 + \sigma^2/n$, use equation (1) to show that $E(SS_E) = ab(n-1)\sigma^2$.

(c) A process engineer was testing the yield of a product manufactured on three machines, each of which has three stations at which the product is manufactured. An experiment was conducted where three observations were taken from each station giving the yields below.

Process yields

	Machine 1			Machine 2			Machine 3		
Station:	1	2	3	1	2	3	1	2	3
	34.1	33.7	36.2	31.1	33.1	32.8	32.9	33.8	33.6
	30.3	34.9	36.8	33.5	34.7	35.1	33.0	33.4	32.8
	31.6	35.0	37.1	34.0	33.9	34.3	33.1	32.8	31.7
$y_{ij\bullet}$	96.0	103.6	110.1	98.6	101.7	102.2	99.0	100.0	98.1
$y_{i\bullet\bullet}$		309.7			302.5			297.1	

Explain why a nested model is appropriate to analyse these data.

Given that $\sum_{i,j,k} y_{ijk}^2 = 30688.51$, $\sum_{ij} y_{ij\bullet}^2 = 92005.27$ and $\sum_i y_{i\bullet\bullet}^2 = 275688.75$, construct an ANOVA table to determine whether the machines or stations have any significant effect on the yield of the process. Comment on your results.

CONTINUED...

5. A continuous response variable Y may be modelled in terms of a continuous predictor variable X by the centred simple linear regression model $Y = \alpha + \beta(X - \overline{X}) + \varepsilon$ where $\varepsilon \sim N(0, \sigma^2)$. Given data $\{(X_i, Y_i); i = 1, \ldots, n\}$, least squares estimates of α and β are $\widehat{\alpha} = \overline{Y}$ and $\widehat{\beta} = S_{XY}/S_{XX}$, where $S_{XY} = \sum_i (X_i - \overline{X})(Y_i - \overline{Y})$ and $S_{XX} = \sum_i (X_i - \overline{X})^2$. Define the residual mean square to be $MS_{RES} = (n-2)^{-1} \sum_i [Y_i - \widehat{\alpha} - \widehat{\beta}(X_i - \overline{X})]^2$.

(a) An alternative matrix form for the centred simple linear regression model is $\boldsymbol{y} = \boldsymbol{X}\,\boldsymbol{\theta} + \boldsymbol{\varepsilon}$. In this matrix form, $\boldsymbol{y}$ and $\boldsymbol{\varepsilon}$ are vectors of length n, $\boldsymbol{X}$ is an $n \times 2$ matrix and $\boldsymbol{\theta}$ is a vector of length 2. Explain how to construct $\boldsymbol{y}, \boldsymbol{X}, \boldsymbol{\theta}$ and $\boldsymbol{\varepsilon}$.

Show that S_{XY} can be written as $S_{XY} = \sum_i (X_i - \overline{X}) Y_i$.

The estimate of the parameter vector $\boldsymbol{\theta}$ is $\widehat{\boldsymbol{\theta}} = (\boldsymbol{X}^T \boldsymbol{X})^{-1} \boldsymbol{X}^T \boldsymbol{y}$. Show by substituting $\boldsymbol{X}$ and $\boldsymbol{y}$ that this gives the least squares estimates $\widehat{\alpha}$ and $\widehat{\beta}$.

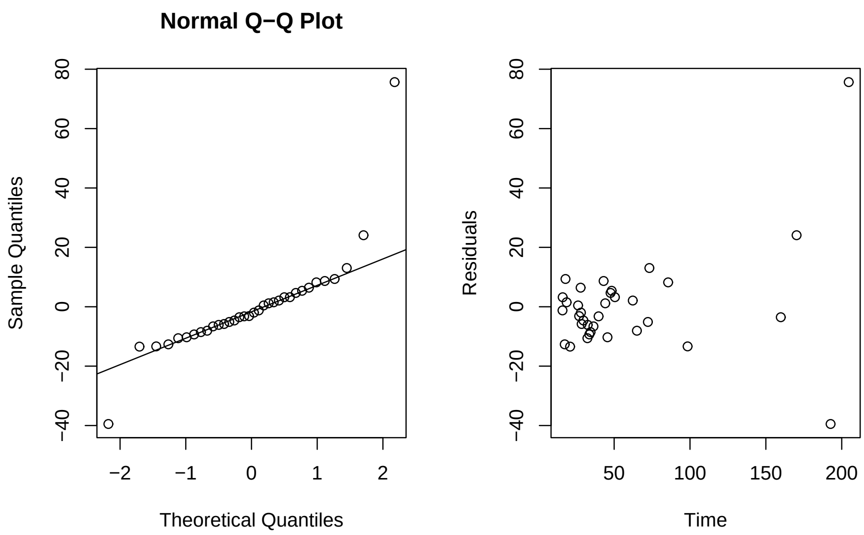
(b) Show that $\mathrm{var}(\widehat{\boldsymbol{\theta}}) = \begin{bmatrix} \sigma^2/n & 0 \\ 0 & S_{XX}/n \end{bmatrix}$.

Hint: You may use the facts that var$(\widehat{\boldsymbol{\theta}}) = E\{(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T\}$ *and* $E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^T) = \sigma^2 \boldsymbol{I}$, *where* $\boldsymbol{I}$ *is the identity matrix.*

(c) The winning time in minutes (y) taken to run 34 Scottish hill races in 1984 were recorded, along with the length of the races in miles (x). Given that $\overline{x} = 7.66$, $\overline{y} = 57.26$, $S_{xx} = 1016.360$, $S_{xy} = 8739.455$, and $MS_{RES} = 298$, fit a linear regression model predicting winning time from the length of the race. What would you expect the winning time to be for the Knock Hill race, which is three miles in length?

A normal QQ plot of the residuals and a plot of the residuals against the times are shown below. Comment on the implications of these plots. If they reveal any problems with the linear regression, indicate how you might remedy these problems.

Construct a 95% confidence interval for the expected time taken to complete a race of zero distance. Comment on any implications of your confidence interval.



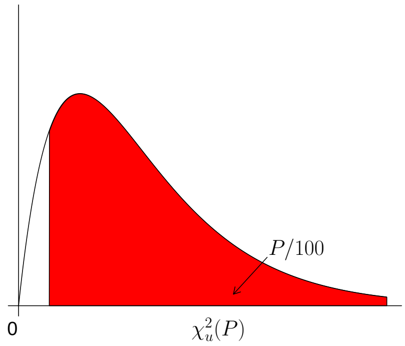
CONTINUED...

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	**99.95**	**99.9**	**99.5**	**99**	**97.5**	**95**	**90**	**80**
1	3.9e-07	1.6e-06	3.9e-05	1.6e-04	9.8e-04	3.9e-03	1.6e-02	6.4e-02
2	0.001	0.002	0.010	0.020	0.051	0.103	0.211	0.446
3	0.015	0.024	0.072	0.115	0.216	0.352	0.584	1.005
4	0.064	0.091	0.207	0.297	0.484	0.711	1.064	1.649
5	0.158	0.210	0.412	0.554	0.831	1.145	1.610	2.343
6	0.299	0.381	0.676	0.872	1.237	1.635	2.204	3.070
7	0.485	0.598	0.989	1.239	1.690	2.167	2.833	3.822
8	0.710	0.857	1.344	1.646	2.180	2.733	3.490	4.594
9	0.972	1.152	1.735	2.088	2.700	3.325	4.168	5.380
10	1.265	1.479	2.156	2.558	3.247	3.940	4.865	6.179
11	1.587	1.834	2.603	3.053	3.816	4.575	5.578	6.989
12	1.934	2.214	3.074	3.571	4.404	5.226	6.304	7.807
13	2.305	2.617	3.565	4.107	5.009	5.892	7.042	8.634
14	2.697	3.041	4.075	4.660	5.629	6.571	7.790	9.467
15	3.108	3.483	4.601	5.229	6.262	7.261	8.547	10.307
16	3.536	3.942	5.142	5.812	6.908	7.962	9.312	11.152
17	3.980	4.416	5.697	6.408	7.564	8.672	10.085	12.002
18	4.439	4.905	6.265	7.015	8.231	9.390	10.865	12.857
19	4.912	5.407	6.844	7.633	8.907	10.117	11.651	13.716
20	5.398	5.921	7.434	8.260	9.591	10.851	12.443	14.578
25	7.991	8.649	10.520	11.524	13.120	14.611	16.473	18.940
30	10.804	11.588	13.787	14.953	16.791	18.493	20.599	23.364
40	16.906	17.916	20.707	22.164	24.433	26.509	29.051	32.345
50	23.461	24.674	27.991	29.707	32.357	34.764	37.689	41.449
80	44.791	46.520	51.172	53.540	57.153	60.391	64.278	69.207

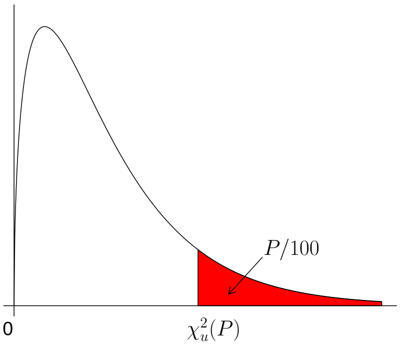
CONTINUED...

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.

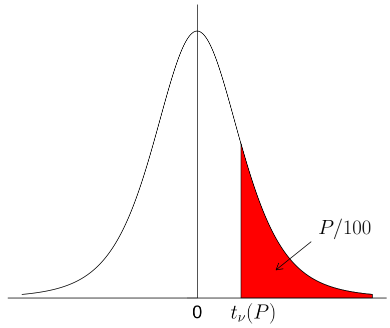


ν	Percentage points P						
	10	**5**	**2.5**	**1**	**0.5**	**0.1**	**0.05**
1	2.706	3.841	5.024	6.635	7.879	10.828	12.116
2	4.605	5.991	7.378	9.210	10.597	13.816	15.202
3	6.251	7.815	9.348	11.345	12.838	16.266	17.730
4	7.779	9.488	11.143	13.277	14.860	18.467	19.997
5	9.236	11.070	12.833	15.086	16.750	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...

Percentage Points of the t-Distribution

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



The lower percentage points are given by symmetry as $-t_u(P)$, and the probability that $|t| \geq t_u(P)$ is $2P/100$.

	Percentage points P						
ν	**10**	**5**	**2.5**	**1**	**0.5**	**0.1**	**0.05**
1	3.078	6.314	12.706	31.821	63.657	318.309	636.619
2	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922
21	1.323	1.721	2.080	2.518	2.831	3.527	3.819
25	1.316	1.708	2.060	2.485	2.787	3.450	3.725
30	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	1.303	1.684	2.021	2.423	2.704	3.307	3.551
50	1.299	1.676	2.009	2.403	2.678	3.261	3.496
70	1.294	1.667	1.994	2.381	2.648	3.211	3.435
100	1.290	1.660	1.984	2.364	2.626	3.174	3.390
∞	1.282	1.645	1.960	2.326	2.576	3.090	3.291

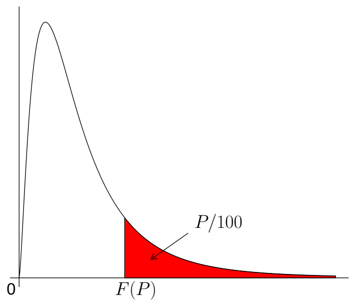
CONTINUED...

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 $F \leq F'_{\nu_1,\nu_2}(P)$ 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_2 \ ν_1	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

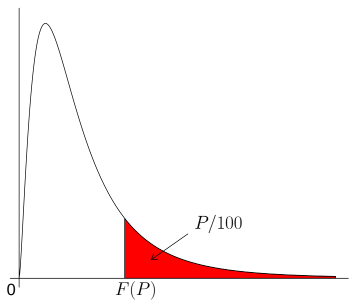
CONTINUED...

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 $F \le F'_{\nu_1,\nu_2}(P)$ 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	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