McGill University

Faculty of Science

MATH 204

PRINCIPLES OF STATISTICS II

Final Examination

Date: 23rd April 2007 Time: 2pm-5pm

Examiner: Dr. David A. Stephens
Associate Examiner: Dr. Russell Steele

Please write your answers in the answer booklets provided.

This paper contains six questions. Each question carries 25 marks. Credit will be given for all questions attempted. The total mark available is 150 but rescaling of the final mark may occur

Candidates may take one double-sided sheet of Letter-sized (216 $\times$ 279 mm) or A4-sized (210 $\times$ 297mm) paper with handwritten notes into the examination room.

Calculators may be used. Relevant statistical tables are provided.

Dictionaries and Translation dictionaries are permitted.

1. In an experimental study of nutrition, laboratory animals were allocated at random to one of four different diets, A,B,C and D. The response measurement was the weight gain (in grammes) of each animal over one week.

The data recorded are tabulated below; entries in the rows of the table are the weight gains for animals allocated to each diet.

Diet							
A	0.54	1.98	0.65	0.52	1.92	1.48	0.97
B	1.24	1.82	1.39	1.25	1.29		
C	2.05	2.18	1.94	2.50	1.98	2.17	1.83
D	1.88	6.23	3.51	3.77	1.25	0.72	

(a) What is the fundamental difference between an *experimental* and an *observational* scientific study ?

4 MARKS

(b) In the experimental study described above, what kind of design is being used ?

2 MARKS

(c) Using the data, an ANOVA analysis is to be carried out. The ANOVA table below contains some missing entries marked by the notation $\star$.

SOURCE	DF	SS	MS	F
TREATMENTS	$\star$	11.355	$\star$	$\star$
ERROR	21	$\star$	$\star$	
TOTAL	$\star$	35.053		

Write out the ANOVA table in full, filling in the missing values **using the information already given in the table**.

10 MARKS

(d) What is the conclusion of the ANOVA analysis ? State clearly the null and alternative hypothesis, the test statistic, the null distribution, and the conclusion.

A table of the Fisher-F distribution is provided on page 13. Entries in the table are the 0.05 tail quantile of the Fisher-F(ν_1, ν_2) *distribution, for different values of* ν_1 *and* ν_2.

5 MARKS

(e) What other information is required to confirm that the conclusion of the test is valid ? Justify your answer in terms of the assumptions underlying the ANOVA analysis.

4 MARKS

2. (a) Explain the structure of a *balanced complete randomized block design with replication* for one treatment factor and one blocking factor. Explain why this kind of design differs from a *factorial design* with two treatment factors.

6 MARKS

(b) In a study of the heart drug enalaprilat, nine patients with congestive heart disease were treated with four different dose levels of the drug (labelled 0 to 3), and their active heart rate (in beats per minute) recorded. The objective of the study was to find out whether there are any differences between the heart rates at different dose levels. In the analysis, the patients are to be treated as levels of a (fixed) blocking factor.

Reference: CS Maskin, S Ocken, B Chadwick, and TH LeJemtel (1985), Comparative systemic and renal effects of dopamine and angiotensin - converting enzyme inhibition with enalaprilat in patients with heart failure *Circulation*, **72**, pp 846-852 (adapted)

The data from the study are displayed in the table below.

	Dose Level			
Patient	0	1	2	3
1	96	92	86	92
2	110	106	108	114
3	89	86	85	83
4	95	78	78	83
5	128	124	118	118
6	100	98	199	94
7	72	68	67	71
8	79	75	74	74
9	100	106	104	102

The data were analyzed in SPSS; the output is included on page 8. The variable **id** denotes the different patients, whereas **dose** denotes dose level.

(i) Explain the results labelled Analysis 1, the model that was fitted, the hypotheses tested, and the conclusions that are indicated.

9 MARKS

(ii) Explain the results labelled Analysis 2; in particular, explain what has gone wrong and why.

6 MARKS

(c) If the assumptions underlying the ANOVA tests were not believed to hold, briefly describe one alternative approach that could be used to test for differences between responses at different dose levels, taking into account the blocking structure.

4 MARKS

3. Measurements of bacteria counts following the culturing of the bacterium *Staphylococcus aureus* at three temperatures (27, 35 or 43 Celsius, labelled T1 to T3 respectively), for five different concentrations of the nutrient tryptone (0.6, 0.8, 1.0, 1.2 or 1.4, labelled C1 to C5 respectively) were recorded. Two independent replicate experiments were made at each of the 3×5 factor level combinations, with the objective of finding out which combination promoted the highest level of bacteria growth.

(a) What is the name of this kind of design ?

3 MARKS

(b) List the five (main effects and interactions) models that can be fitted to these data if the two variables are treated as factors. Use the standard model notation, where, for two factors A and B, the main effects plus interaction model is denoted

$$A + B + A.B$$

For this example, use the notation

C	for the concentration factor
T	for the temperature factor

5 MARKS

(c) Output from the SPSS analysis of the response data (on the log scale) is given on page 9. In the output, the variable **conc** denotes concentration, and **temp** denotes temperature.

Summarize in detail what the output illustrates. Comment specifically on

(i) The statistical evidence for including or excluding an interaction term.

(ii) The most appropriate model in ANOVA-F terms, using **backward selection**. Recall that the ANOVA-F test for comparing nested models is based on the statistic

$$F = \frac{(SSE_R - SSE_C)/(k-g)}{SSE_C/(n-k-1)}$$

where

- SSE_R is the error sum of squares for the **Reduced Model**, specified using $g+1$ parameters including the intercept.
- SSE_C is the error sum of squares for the **Complete Model**, specified using $k+1$ parameters including the intercept.

If the reduced model is an adequate simplification of the complete model, then

$$F \sim \text{Fisher-F}(k-g, n-k-1)$$

(iii) Overall model fit.

12 MARKS

(d) In the analysis described in (c), concentration is treated as a **factor predictor** taking five levels. However, it is possible instead to treat concentration as a **continuous covariate**.

How many **fewer** parameters does the **main effects only** model

$$C + T$$

contain if concentration is treated as a continuous covariate rather than a factor predictor ? Justify your answer.

5 MARKS

4. For lung transplantation it is desirable for the donor's lungs to have similar total capacity to those of the recipient. Total lung capacity (TLC) is difficult to measure, so a predictive model for TLC in terms of Age, Sex and Height of the donor is sought.

Measurements of TLC were collected from 32 patients pre-transplant, and their age, sex and height recorded, and analyzed using a general linear regression model.

Reference: Otulana BA, Higenbottam TW, Scott JP, Clelland C, Hutter, J and Wallwork J. (1989) Pulmonary function monitoring allows diagnosis of rejection in heart-lung transplant recipients. *Transplant Proceedings*, 1989;21:25834.

Output from an SPSS analysis is recorded on pages 11 to 12. Analyses of different models are presented under different sub-headings. In the analysis, sex is a factor predictor, and is coded as 1 for female, 2 for male.

(a) Given the results of all the analyses presented, decide on the best model to explain the observed variation in the response, total lung capacity. Justify your answer.

12 MARKS

(b) Explain the apparently contradictory results of Analyses 2 and 4.

4 MARKS

(c) Using analyses available, predict the total lung capacity for a 30-year-old female who is 165 cm tall. Recall that to predict in a linear regression model of the form

$$y = \beta_0 + \beta_1 x_1 + \cdots \beta_k x_k + \epsilon$$

the prediction is

$$\widehat{y} = \widehat{\beta}_0 + \widehat{\beta}_1 x_1 + \cdots + \widehat{\beta}_k x_k$$

where $\widehat{\beta}_0, \widehat{\beta}_1, \ldots, \widehat{\beta}_k$ are the estimated coefficients.

6 MARKS

(d) In Analysis 2 and Analysis 4, some of the parameter estimates are given the note

This parameter is set to zero because it is redundant.

Briefly explain this note.

3 MARKS

5. The following data relate to a study of dental enamel erosion by swimming pool chlorine, in a case-control observational study of swimmers; 49 cases of dental enamel erosion, and 245 controls, were recruited, and asked if they swam for greater than or equal to six hours, or less than six hours per week.

Reference: Centerwall BS *et al.* (1986), Erosion of dental enamel among competitive swimmers at a gas-chlorinated swimming pool, *American Journal of Epidemiology*, **123**(4) pp 641-647.

Swim ≥ 6 hrs.	Erosion of Enamel Yes	No	Total
Yes	32	118	150
No	17	127	144
Total	49	245	294

For this $r \times c$ table $(r = 2, c = 2)$, we wish to test the null hypothesis of **independence** using a Chi-squared test, and the odds-ratio statistic.

(a) Form the table of **expected values** under the null hypothesis with entries $\widehat{n}_{ij}$ given by the formula

$$\widehat{n}_{ij} = \frac{n_{i.} n_{.j}}{n} \qquad i = 1, 2, \; j = 1, 2.$$

where

$n_{i.}$ is the row total for row i $\qquad$ $n_{.j}$ is the column total for column j.

6 MARKS

(b) Compute the Chi-squared statistic

$$X^2 = \sum_{i=1}^{2} \sum_{j=1}^{2} \frac{(n_{ij} - \widehat{n}_{ij})^2}{\widehat{n}_{ij}}$$

4 MARKS

(c) Complete the test at the $\alpha = 0.05$ significance level of the null hypothesis, recalling that if the independence hypothesis is true, $X^2 \stackrel{\cdot}{\sim}$ Chisquared$((r-1)(c-1))$.

The table on page 14 contains the 0.05 and 0.01 tail quantiles of the Chisquared(ν) distribution, for $\nu = 1$ to 20.

4 MARKS

(d) Are the required conditions for the Chi-squared test met for this analysis ? Justify your answer.

3 MARKS

(e) Using the log odds ratio, its standard error and the test statistic Z

$$\log \widehat{\psi} = \log \left(\frac{n_{11} \, n_{22}}{n_{12} \, n_{21}} \right) \qquad \text{s.e.}(\log \widehat{\psi}) = \sqrt{\frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}}$$

$$Z = \frac{\log \widehat{\psi}}{\text{s.e.}(\log \widehat{\psi})}$$

test for an association between the factor and disease status. Use the result that under the hypothesis of no association, $Z \stackrel{\cdot}{\sim} N(0, 1)$, so that the critical values in a two-tailed test are ± 1.96.

8 MARKS

6. (a) Explain briefly what each of the following named tests is used for:

(i) Levene's Test

(ii) Friedman's Test

(iii) Fisher's Exact Test

9 MARKS

(b) Using the Kruskal-Wallis Test, assess the statistical significance of the following data. The data correspond to the comparison of three treatments (1-Control, 2-Relaxation, 3-Dietary) for chronic headache sufferers; data in the table are the improvement measures (y) and corresponding ranks (r) for the 18 patients in this randomized experimental study.

Reference: Fentress, DW, Masek, BJ, Mehegan, JE, and Benson, H, (1986). Biofeedback and relaxation-response training in the treatment of pediatric migraine. *Developmental Medicine and Child Neurology*, **28**, 138146.

Group	1	1	1	1	1	1	2	2	2	2	2	2	3	3	3	3	3	3
y	62	74	86	74	91	37	69	43	100	94	100	98	50	-120	100	-288	4	-76
r	11	8.5	7	8.5	6	14	10	13	2	5	2	4	12	17	2	18	15	16

Recall that the Kruskal-Wallis test statistic is

$$H = \frac{12}{n(n+1)} \sum_{j=1}^{k} \frac{R_j^2}{n_j} - 3(n+1)$$

where R_j is the rank sum for group $j = 1, 2, 3$, and k is the number of groups being compared. If the relevant null hypothesis, H_0, is **true**, then for large n,

$$H \approx \text{Chisquared}(k-1).$$

The table on page 14 contains the 0.05 and 0.01 tail quantiles of the Chisquared(ν) distribution, for $\nu = 1$ to 20.

12 MARKS

(c) If the data y were assumed to be Normally distributed, what kind of test might be used to analyze the data ? List any extra assumptions that need to be made.

4 MARKS

Output for Question 2

Analysis 1:

Tests of Between-Subjects Effects

Dependent Variable: y

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	12673.278[a]	11	1152.116	3.625	.004
Intercept	331008.444	1	331008.444	1041.415	.000
id	12107.056	8	1513.382	4.761	.001
dose	566.222	3	188.741	.594	.625
Error	7628.278	24	317.845		
Total	351310.000	36			
Corrected Total	20301.556	35			

a. R Squared = .624 (Adjusted R Squared = .452)

Parameter Estimates

Dependent Variable: y

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Intercept	99.444	10.293	9.661	.000	78.200	120.688
[id=1]	-11.500	12.606	-.912	.371	-37.518	14.518
[id=2]	6.500	12.606	.516	.611	-19.518	32.518
[id=3]	-17.250	12.606	-1.368	.184	-43.268	8.768
[id=4]	-19.500	12.606	-1.547	.135	-45.518	6.518
[id=5]	19.000	12.606	1.507	.145	-7.018	45.018
[id=6]	19.750	12.606	1.567	.130	-6.268	45.768
[id=7]	-33.500	12.606	-2.657	.014	-59.518	-7.482
[id=8]	-27.500	12.606	-2.181	.039	-53.518	-1.482
[id=9]	0[a]	.	.	.	.	.
[dose=1]	4.222	8.404	.502	.620	-13.123	21.568
[dose=2]	.222	8.404	.026	.979	-17.123	17.568
[dose=3]	9.778	8.404	1.163	.256	-7.568	27.123
[dose=4]	0[a]	.	.	.	.	.

a. This parameter is set to zero because it is redundant.

Analysis 2:

Tests of Between-Subjects Effects

Dependent Variable: y

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	20301.556[a]	35	580.044	.	.
Intercept	331008.444	1	331008.444	.	.
id	12107.056	8	1513.382	.	.
dose	566.222	3	188.741	.	.
id * dose	7628.278	24	317.845	.	.
Error	.000	0	.		
Total	351310.000	36			
Corrected Total	20301.556	35			

a. R Squared = 1.000 (Adjusted R Squared = .)

Output for Question 3

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	9.356	14	.668	.917	.562
Intercept	636.336	1	636.336	873.407	.000
conc	2.658	4	.665	.912	.482
temp	6.378	2	3.189	4.377	.032
conc * temp	.320	8	.040	.055	1.000
Error	10.929	15	.729		
Total	656.621	30			
Corrected Total	20.284	29			

R Squared = .461 (Adjusted R Squared = -.042)

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	9.036	6	1.506	3.080	.023
Intercept	636.336	1	636.336	1301.179	.000
conc	2.658	4	.665	1.359	.279
temp	6.378	2	3.189	6.521	.006
Error	11.248	23	.489		
Total	656.621	30			
Corrected Total	20.284	29			

R Squared = .445 (Adjusted R Squared = .301)

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	6.378	2	3.189	6.191	.006
Intercept	636.336	1	636.336	1235.470	.000
temp	6.378	2	3.189	6.191	.006
Error	13.907	27	.515		
Total	656.621	30			
Corrected Total	20.284	29			

R Squared = .314 (Adjusted R Squared = .264)

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	2.658	4	.665	.943	.456
Intercept	636.336	1	636.336	902.557	.000
conc	2.658	4	.665	.943	.456
Error	17.626	25	.705		
Total	656.621	30			
Corrected Total	20.284	29			

R Squared = .131 (Adjusted R Squared = -.008)

Output for Question 4

Analysis 1:

Tests of Between-Subjects Effects

Dependent Variable: Total Lung Capacity

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	4.411[a]	1	4.411	1.712	.201
Intercept	96.398	1	96.398	37.412	.000
Age	4.411	1	4.411	1.712	.201
Error	77.301	30	2.577		
Total	1267.557	32			
Corrected Total	81.712	31			

a. R Squared = .054 (Adjusted R Squared = .022)

Analysis 2:

Tests of Between-Subjects Effects

Dependent Variable: Total Lung Capacity

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	25.312[a]	1	25.312	13.464	.001
Intercept	1185.845	1	1185.845	630.764	.000
sex	25.312	1	25.312	13.464	.001
Error	56.400	30	1.880		
Total	1267.557	32			
Corrected Total	81.712	31			

a. R Squared = .310 (Adjusted R Squared = .287)

Parameter Estimates

Dependent Variable: Total Lung Capacity

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Intercept	6.977	.343	20.354	.000	6.277	7.677
[sex=1]	-1.779	.485	-3.669	.001	-2.769	-.789
[sex=2]	0[a]	.	.	.	.	.

a. This parameter is set to zero because it is redundant.

Output for Question 4

Analysis 3:

Tests of Between-Subjects Effects

Dependent Variable: Total Lung Capacity

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	39.549[a]	1	39.549	28.140	.000
Intercept	14.904	1	14.904	10.604	.003
Height	39.549	1	39.549	28.140	.000
Error	42.163	30	1.405		
Total	1267.557	32			
Corrected Total	81.712	31			

a. R Squared = .484 (Adjusted R Squared = .467)

Parameter Estimates

Dependent Variable: Total Lung Capacity

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	-9.740	2.991	-3.256	.003	-15.849	-3.632
Height	.095	.018	5.305	.000	.058	.131

Analysis 4:

Tests of Between-Subjects Effects

Dependent Variable: Total Lung Capacity

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	42.818[a]	3	14.273	10.275	.000
Intercept	4.771	1	4.771	3.435	.074
sex	.080	1	.080	.057	.813
Height	17.361	1	17.361	12.498	.001
sex * Height	.025	1	.025	.018	.895
Error	38.894	28	1.389		
Total	1267.557	32			
Corrected Total	81.712	31			

a. R Squared = .524 (Adjusted R Squared = .473)

Parameter Estimates

Dependent Variable: Total Lung Capacity

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	-5.828	4.977	-1.171	.251	-16.023	4.367
[sex=1]	-1.728	7.221	-.239	.813	-16.520	13.064
[sex=2]	0[a]	.	.	.	.	.
Height	.074	.029	2.577	.016	.015	.132
[sex=1] * Height	.006	.043	.133	.895	-.083	.094
[sex=2] * Height	0[a]	.	.	.	.	.

a. This parameter is set to zero because it is redundant.

Output for Question 4

Analysis 5:

Tests of Between-Subjects Effects

Dependent Variable: Total Lung Capacity

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	47.718[a]	6	7.953	5.849	.001
Intercept	7.568	1	7.568	5.566	.026
sex	.348	1	.348	.256	.617
Age	3.105	1	3.105	2.284	.143
Height	12.936	1	12.936	9.514	.005
sex * Age	1.150	1	1.150	.846	.367
sex * Height	.359	1	.359	.264	.612
Age * Height	3.389	1	3.389	2.492	.127
Error	33.994	25	1.360		
Total	1267.557	32			
Corrected Total	81.712	31			

a. R Squared = .584 (Adjusted R Squared = .484)

Analysis 6:

Tests of Between-Subjects Effects

Dependent Variable: Total Lung Capacity

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	50.492[a]	7	7.213	5.545	.001
Intercept	8.781	1	8.781	6.750	.016
sex	1.476	1	1.476	1.135	.297
Age	4.933	1	4.933	3.792	.063
Height	13.967	1	13.967	10.737	.003
sex * Age	3.033	1	3.033	2.331	.140
sex * Height	1.321	1	1.321	1.016	.324
Age * Height	5.050	1	5.050	3.882	.060
sex * Age * Height	2.774	1	2.774	2.132	.157
Error	31.220	24	1.301		
Total	1267.557	32			
Corrected Total	81.712	31			

a. R Squared = .618 (Adjusted R Squared = .506)

Table of the Fisher-F distribution

Entries in table are the $\alpha = 0.05$ tail quantile of Fisher-F(ν_1, ν_2) distribution
ν_1 given in columns, ν_2 given in rows.

$\nu_2 \backslash \nu_1$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88	242.98	243.91	244.69	245.36	245.95	246.46	246.92	247.32	247.69	248.01	248.31	248.58	248.83	249.05	249.26
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.40	19.41	19.42	19.42	19.43	19.43	19.44	19.44	19.44	19.45	19.45	19.45	19.45	19.45	19.46
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.76	8.74	8.73	8.71	8.70	8.69	8.68	8.67	8.67	8.66	8.65	8.65	8.64	8.64	8.63
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.94	5.91	5.89	5.87	5.86	5.84	5.83	5.82	5.81	5.80	5.79	5.79	5.78	5.77	5.77
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.70	4.68	4.66	4.64	4.62	4.60	4.59	4.58	4.57	4.56	4.55	4.54	4.53	4.53	4.52
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.03	4.00	3.98	3.96	3.94	3.92	3.91	3.90	3.88	3.87	3.86	3.86	3.85	3.84	3.83
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.60	3.57	3.55	3.53	3.51	3.49	3.48	3.47	3.46	3.44	3.43	3.43	3.42	3.41	3.40
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.31	3.28	3.26	3.24	3.22	3.20	3.19	3.17	3.16	3.15	3.14	3.13	3.12	3.12	3.11
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.10	3.07	3.05	3.03	3.01	2.99	2.97	2.96	2.95	2.94	2.93	2.92	2.91	2.90	2.89
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.94	2.91	2.89	2.86	2.85	2.83	2.81	2.80	2.79	2.77	2.76	2.75	2.75	2.74	2.73
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.82	2.79	2.76	2.74	2.72	2.70	2.69	2.67	2.66	2.65	2.64	2.63	2.62	2.61	2.60
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.72	2.69	2.66	2.64	2.62	2.60	2.58	2.57	2.56	2.54	2.53	2.52	2.51	2.51	2.50
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.63	2.60	2.58	2.55	2.53	2.51	2.50	2.48	2.47	2.46	2.45	2.44	2.43	2.42	2.41
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.57	2.53	2.51	2.48	2.46	2.44	2.43	2.41	2.40	2.39	2.38	2.37	2.36	2.35	2.34
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.51	2.48	2.45	2.42	2.40	2.38	2.37	2.35	2.34	2.33	2.32	2.31	2.30	2.29	2.28
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.46	2.42	2.40	2.37	2.35	2.33	2.32	2.30	2.29	2.28	2.26	2.25	2.24	2.24	2.23
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.41	2.38	2.35	2.33	2.31	2.29	2.27	2.26	2.24	2.23	2.22	2.21	2.20	2.19	2.18
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.37	2.34	2.31	2.29	2.27	2.25	2.23	2.22	2.20	2.19	2.18	2.17	2.16	2.15	2.14
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.34	2.31	2.28	2.26	2.23	2.21	2.20	2.18	2.17	2.16	2.14	2.13	2.12	2.11	2.11
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.31	2.28	2.25	2.22	2.20	2.18	2.17	2.15	2.14	2.12	2.11	2.10	2.09	2.08	2.07
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.28	2.25	2.22	2.20	2.18	2.16	2.14	2.12	2.11	2.10	2.08	2.07	2.06	2.05	2.05
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.26	2.23	2.20	2.17	2.15	2.13	2.11	2.10	2.08	2.07	2.06	2.05	2.04	2.03	2.02
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.24	2.20	2.18	2.15	2.13	2.11	2.09	2.08	2.06	2.05	2.04	2.02	2.01	2.01	2.00
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.22	2.18	2.15	2.13	2.11	2.09	2.07	2.05	2.04	2.03	2.01	2.00	1.99	1.98	1.97
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.20	2.16	2.14	2.11	2.09	2.07	2.05	2.04	2.02	2.01	2.00	1.98	1.97	1.96	1.96
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.18	2.15	2.12	2.09	2.07	2.05	2.03	2.02	2.00	1.99	1.98	1.97	1.96	1.95	1.94
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.17	2.13	2.10	2.08	2.06	2.04	2.02	2.00	1.99	1.97	1.96	1.95	1.94	1.93	1.92
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.15	2.12	2.09	2.06	2.04	2.02	2.00	1.99	1.97	1.96	1.95	1.93	1.92	1.91	1.91
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.14	2.10	2.08	2.05	2.03	2.01	1.99	1.97	1.96	1.94	1.93	1.92	1.91	1.90	1.89
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.13	2.09	2.06	2.04	2.01	1.99	1.98	1.96	1.95	1.93	1.92	1.91	1.90	1.89	1.88
31	4.16	3.30	2.91	2.68	2.52	2.41	2.32	2.25	2.20	2.15	2.11	2.08	2.05	2.03	2.00	1.98	1.96	1.95	1.93	1.92	1.91	1.90	1.88	1.88	1.87
32	4.15	3.29	2.90	2.67	2.51	2.40	2.31	2.24	2.19	2.14	2.10	2.07	2.04	2.01	1.99	1.97	1.95	1.94	1.92	1.91	1.90	1.88	1.87	1.86	1.85

Table of the **Chisquared**(ν) distribution

Entries in table are the $\alpha = 0.05$ and $\alpha = 0.01$ tail quantiles

ν	1	2	3	4	5	6	7	8	9	10
$\alpha = 0.05$	3.841	5.991	7.815	9.488	11.070	12.592	14.067	15.507	16.919	18.307
$\alpha = 0.01$	6.635	9.210	11.345	13.277	15.086	16.812	18.475	20.090	21.666	23.209

ν	11	12	13	14	15	16	17	18	19	20
$\alpha = 0.05$	19.675	21.026	22.362	23.685	24.996	26.296	27.587	28.869	30.144	31.410
$\alpha = 0.01$	24.725	26.217	27.688	29.141	30.578	32.000	33.409	34.805	36.191	37.566