THE UNIVERSITY OF QUEENSLAND
AUSTRALIA

This exam paper must not be removed from the venue

Venue ____________

Seat Number ______

Student Number |__|__|__|__|__|__|__|__|

Family Name ____________

First Name ____________

Solutions

School of Mathematics & Physics

EXAMINATION

Semester One Final Examinations, 2019

STAT1201 Analysis of Scientific Data

This paper is for St Lucia Campus students.

Examination Duration: 120 minutes

Reading Time: 10 minutes

Exam Conditions:

This is a Central Examination

This is a Closed Book Examination - specified materials permitted

During reading time - write only on the rough paper provided

This examination paper will be released to the Library

Materials Permitted In The Exam Venue:

(No electronic aids are permitted e.g. laptops, phones)

Calculators - Casio FX82 series or UQ approved (labelled)

An annotated copy of A Portable Introduction to Data Analysis (any edition) is also permitted.

Materials To Be Supplied To Students:

None

Instructions To Students:

Additional exam materials (eg. answer booklets, rough paper) will be provided upon request.

There are **45** marks available on this exam from **4** questions.

Write your answers in the spaces provided on pages 2 – 12 of this examination paper. Show your working and state conclusions where appropriate. Pages 13 – 20 give formulas and statistical tables. Those pages will not be marked.

The textbook can have any amount of annotation on its pages. Loose sheets of paper or post-it notes are not permitted. Page tabs are allowed.

For Examiner Use Only

Question	Mark

Total ______

Question 1

10 marks

A study evaluated the effect of 8-hour nightly continuous positive airway pressure (CPAP) treatment on fasting insulin levels in patients with prediabetes and suffering from obstructive sleep apnoea. In the study 39 patients were randomly assigned to either the 8-hour nightly CPAP treatment group or an oral placebo group. The fasting insulin level was measured for each patient at the beginning of the study and after two weeks treatment.

(a) At the beginning of the study each patient's percentage of body fat was measured. The 26 patients in the CPAP treatment group had a mean body fat percentage of 36.8% and a standard deviation of 7.8% while the 13 patients in the oral placebo group had mean body fat percentage of 32.7% and a standard deviation of 4.3%. Is there any evidence of a difference in the mean body fat percentage between the two groups? State the null and alternative hypotheses, and use an appropriate test statistic to determine the *P*-value. What do you conclude? [5 marks]

let μ_1 = mean body fat % for CPAP group

and μ_2 = mean body fat % for placebo group.

Test $H_0: \mu_1 = \mu_2$ vs $H_1: \mu_1 \neq \mu_2$

test statistic $t = \dfrac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}$

$= \dfrac{(0.368 - 0.327) - 0}{\sqrt{0.078^2/26 + 0.043^2/13}}$

$= \dfrac{0.041}{0.0194} = 2.11$

We compare this to a T_{12} distribution $(df = \min(n_1-1, n_2-1))$.

one sided p-value $= P(T_{12} \geq 2.11)$ is between 0.025 and 0.05.

The p-value for this two-sided test is between 0.05 and 0.1.

There is weak evidence of a difference in mean body fat % between the two groups.

(b) After the two week trial, the 26 subjects in the CPAP group experienced an average decrease in fasting insulin level of 5.7 (pmol/L) with a standard deviation of 26.7 (pmol/L). Does this give any evidence of a decrease in mean fasting insulin level for patients receiving the CPAP treatment? State the null and alternative hypotheses, and use an appropriate test statistic to determine the *P*-value. What do you conclude? [4 marks]

let μ = decrease in mean fasting insulin level after two weeks of CPAP treatment.

Test $H_0: \mu = 0$ vs $H_1: \mu > 0$

Test statistic $t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{5.7 - 0}{26.7/\sqrt{26}} = \frac{5.7}{5.23} = 1.089$

we compare this to a T_{25} distribution

p-value = $P(T_{25} \geq 1.089)$ is between 0.1 and 0.25

There is inconclusive evidence of a decrease in mean fasting insulin level.

(c) At the start of the study the researchers determined that a change (positive or negative) of 5 (pmol/L) in mean fasting insulin was clinically significant. From related research they thought that the standard deviation of the changes in fasting insulin would be about 25 (pmol/L). The researchers made the following calculation in R:

```
power.t.test(n=26,delta=5,sd=25,sig.level=0.05,type='one.sample')

     One-sample t test power calculation

              n = 26
          delta = 5
             sd = 25
      sig.level = 0.05
          power = 0.1638018
    alternative = two.sided
```

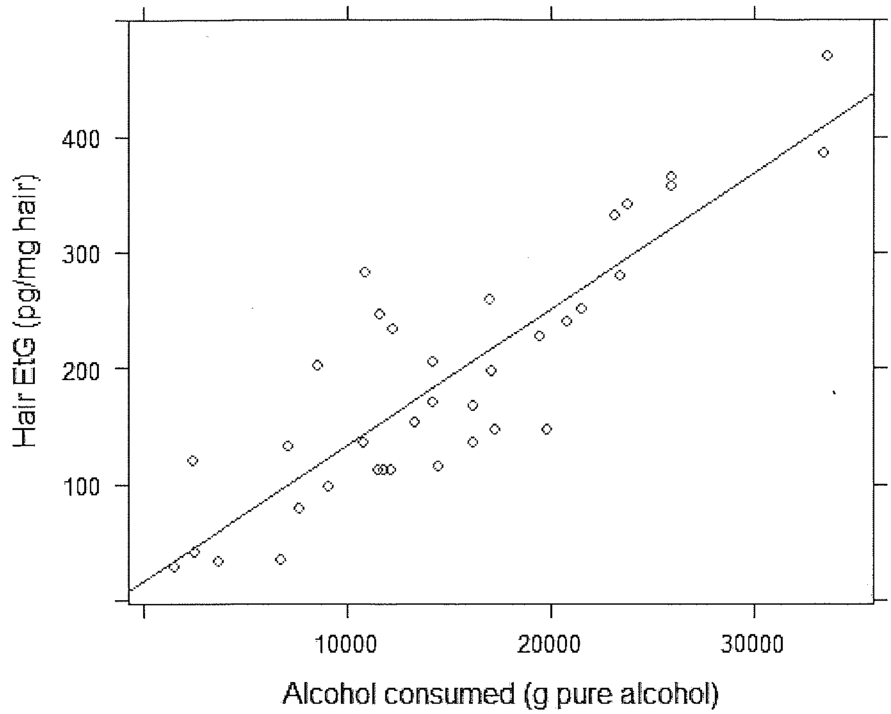
Approximately, what is the probability that the researchers would have been able to detect a change of 5 (pmol/L) in the mean fasting insulin levels with this study, assuming such a change was present? [1 mark]

The probability of detecting a change is given by the power. From the output, this is 0.164.

Question 2

11 marks

Ethyl glucuronide (EtG), an alcohol metabolite, is known to accumulate in hair. A study was undertaken to determine its suitability as a marker of chronic and excessive alcohol consumption. The study involved 36 patients entering a detoxification program Hair samples (first 3 cm segments) were taken from each patient and the EtG level measured (units pg/mg hair) by gas chromatography coupled to mass spectrometry. Interviews were conducted with patients to determine alcohol consumption (units: g pure alcohol) over the past three months. The data are displayed in the figure below together with the fitted least squares line.



The output on the next page show the results of a linear regression fit in R for the relationship between EtG levels in hair and past alcohol consumption.

```
lm(formula = hair ~ alcohol)

Residuals:
    Min      1Q  Median      3Q     Max
-101.97  -30.04  -15.74   43.16  138.28

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) 15.872784  19.611596   0.809    0.424
alcohol      0.011776   0.001155  10.195 7.09e-12 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 54.69 on 34 degrees of freedom
Multiple R-squared:  0.7535,   Adjusted R-squared:  0.7463
F-statistic: 103.9 on 1 and 34 DF,  p-value: 7.092e-12
```

(a) Briefly interpret the value 15.872784 in the regression output. [1 mark]

A person with no alcohol consumption in the past 3 months would have ~ 15.87 pg/mg hair of EtG in their hair on average.

(b) Briefly interpret the value 0.011776 in the regression output. [1 mark]

A 1g increase in alcohol consumption in the past 3 months results in a mean increase of ~0.0118 pg/mg hair of EtG.

(c) Give a 95% confidence interval for the underlying slope of the linear relationship between the level of EtG in hair and alcohol consumption. [2 marks]

CI for β_1 (slope) is $b_1 \pm t^* \, s.e.(b_1)$

The output says 34 degrees of freedom – not given in T-table so we use 30 degrees of freedom.

$0.011776 \pm 2.042 \times 0.001155$

$= 0.011776 \pm 0.002358$

(d) Does the regression analysis provide evidence that mean EtG concentration in hair will be different to zero when a patient's past alcohol consumption is zero? State the null and alternative hypotheses, and report the appropriate test statistic and *P*-value from the output. What do you conclude? [3 marks]

Let β_0 = mean concentration of EtG in hair for a person with zero past alcohol consumption.

Test $H_0 : \beta_0 = 0$ vs $H_1 : \beta_0 \neq 0$

reported test statistic is $t = 0.809$

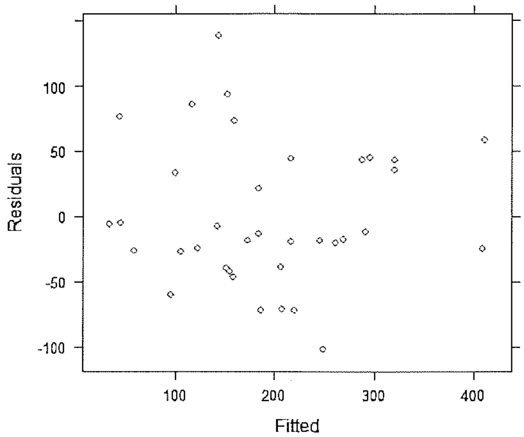
reported p-value is 0.424

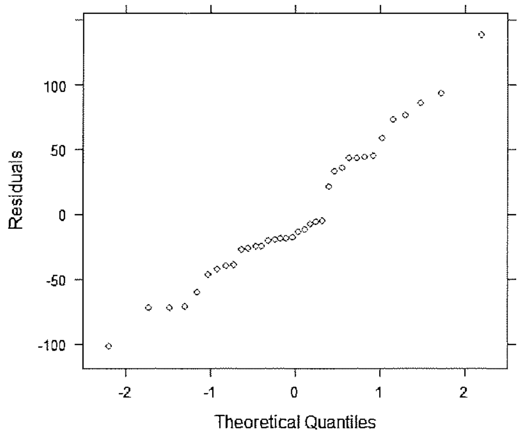
There is inconclusive evidence that the mean concentration of EtG is not zero for a person with zero past alcohol consumption.

(e) A patient whose past alcohol consumption was 19800 g pure alcohol had EtG level of 145 pg/mg hair in their hair sample. What is the residual associated with this patient in the regression model. [1 mark]

$$\text{Residual} = \text{observed} - \text{predicted}$$
$$= 145 - (0.011776 \times 19800 + 15.8727)$$
$$\simeq 145 - 249 \approx -104 \text{ pg/mg hair}$$

(f) The following figures were generated by R to help check the assumptions underlying the linear regression. The theoretical quantiles on the horizontal axis of the plot on the left are those of the standard normal distribution.





Comment on the validity of the assumptions underlying linear regression for this data with reference to these figures and the figure on page 4. [3 marks]

- The even spread in the plot of residuals vs fitted values is consistent with the assumption of constant variance.

- The quantile plot shows a relatively straight line. This is in agreement with the normality assumption

- The plot on pg 4 shows a linear relationship between hair EtG and alcohol consumed. Again, this is in agreement with the model assumption that the mean is a linear function.

Question 3

11 marks

A study was made to identify characteristics of children presenting at two hospitals (Casey and Clayton) suffering accidental poisoning. Patients aged under 13 and admitted to the emergency department of one of the hospitals during the period 1 July 2009 and 30 June 2012 were included in the study.

Of the 1248 patients admitted to Clayton hospital the source of the poison in 496 cases was household cleaning products while of the 1057 patients admitted to Casey hospital, household cleaning products was the source in 406 cases.

(a) Overall, what proportion of accidental poisoning cases did household cleaning products cause? [1 mark]

$$\frac{496 + 406}{1248 + 1057} \simeq 0.391$$

(b) What is the difference between the two hospitals in the proportions of poisonings caused by household cleaning products? [1 mark]

$$\frac{496}{1248} - \frac{406}{1057} \simeq 0.0133$$

(c) Give a 95% confidence interval for the true difference in the proportions of children accidentally poisoned with household cleaning products between the two hospitals. What does this interval say about accidental poisoning rates of children due to household cleaning products in the catchment areas of the two hospitals? [4 marks]

95% CI for the difference of proportions

$$\hat{p}_1 - \hat{p}_2 \pm 1.96 \times \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

$$= 0.0133 \pm 1.96 \times \sqrt{\frac{0.397(1-0.397)}{1248} + \frac{0.384(1-0.384)}{1057}}$$

$$= 0.0133 \pm 0.03996$$

As the CI covers 0, there does not appear to be a difference in the accidental poisoning rates due to household cleaning products between the two hospitals.

(e) The following table gives the summary of the major emergency department discharge categories for children presenting at the two hospitals in the study with accidental poisoning.

	Discharged to usual residence directly	Admitted to a ward of the same hospital	Transferred to a different hospital
Casey	361	754	133
Clayton	317	596	144

Based on this table, is there evidence of an association between hospital and discharge category of children? [5 marks]

Test H_0: no association between hospital & discharge category.
H_1: some association between hospital & discharge category.

test statistic $X^2 = \sum_i \frac{(E_i - O_i)^2}{E_i}$

$E_{11} = \frac{678 \times 1248}{2305} \simeq 367.1$

$E_{21} = 678 - 367.1 \simeq 310.9$

$E_{12} = \frac{1350 \times 1248}{2305} \simeq 730.9$

$E_{22} = 1350 - 730.9 \simeq 619.1$

$E_{31} = \frac{133 \times 1248}{2305} = 150$

$E_{23} = 277 - 150 = 127$

$X^2 = \frac{(367.1-361)^2}{367.1} + \frac{(310.9-317)^2}{310.9} + \frac{(730.9-754)^2}{730.9} + \frac{(619.1-596)^2}{619.1} + \frac{(150-133)^2}{150} + \frac{(127-144)^2}{127}$

$\simeq 5.99$

Under H_0 X^2 has approximately χ^2_{df} distribution

where $df = (\#\text{rows} - 1) \times (\#\text{columns} - 1) = (2-1)\times(3-1) = 2$

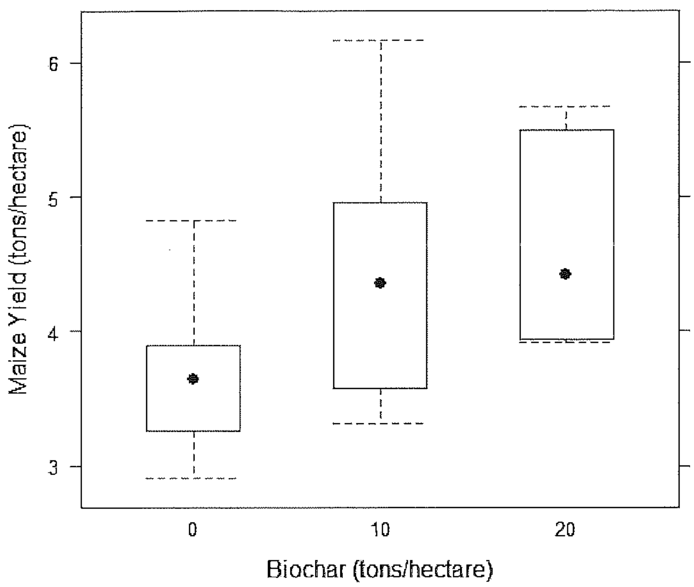
p-value $= P(\chi^2_2 \geq 5.99) \simeq 0.05$

There is mild\ weak evidence of an association between admitting hospital and discharge category.

Question 4

13 marks

Biochar is the carbonaceous product from the burning of biomass. When added to the soil, biochar has been shown to improve the soil's chemical properties. A study was conducted in Nepal to see if biochar could also be used to improve the grain yield from maize crops. An invasive forest shrub was used as feedstock to produce the biochar used in the study. The field trial was conducted on private rain-fed agricultural land which was divided into twenty-four plots of $10m^2$. Treatments of 0, 10 and 20 tons/hectare of biochar with eight replications each were applied in a completely randomised design to the plots. In addition to biochar, equal mineral fertiliser and farmyard manure was applied to each plot according to standard farming practice. The following plot shows a summary of the results:



The following table shows the corresponding summary statistics:

Treatment	n	$\bar{x}$	s
0 t/ha (control)	8	3.67	0.59
10 t/ha	8	4.40	0.97
20 t/ha	8	4.66	0.77

(a) Briefly describe the distribution of the maize yield from plots treated with 10 tons/hectare of biochar. [1 mark]

The distribution appears slightly skewed to the right.
There are no outliers.

(b) Researchers were interested in testing if biochar affects the mean yield from maize crops. State the null hypothesis in words for this study. [1 mark]

Let μ_i = mean maize yield from treatment of i ton/hectare biochar
with $i = 0, 10, 20$

$H_0 : \mu_0 = \mu_{10} = \mu_{20}$

(c) An analysis of variance was conducted to compare the mean maize yield for the three treatment levels of biochar. Ignoring the treatment groups, the sample standard deviation for the yields was 0.87. Based on this and the values given in the previous table, complete the following ANOVA table. [5 marks]

Source	DF	SS	MS	F
Treatment Groups	2	4.235	2.117	3.375
Residual	21	13.173	0.627	
Total	23	17.408		

Residual SS = $(s_1^2 + s_2^2 + s_3^2) \times (8-1) = 13.173$

Total SS = $s^2 \times 23 = 17.408$

(d) Is there significant evidence of a difference in mean maize yield across the three treatment levels of biochar at the 10% level? Justify your conclusion. [2 marks]

Comparing the f statistic with the $F_{2,21}$ distribution gives the p-value between 0.05 and 0.1. There is evidence of a difference in mean maize yield at the 10% level.

(e) Tukey's Honest Significant Differences were calculated in R. The output is given below.

```
  Tukey multiple comparisons of means
    90% family-wise confidence level

Fit: aov(formula = yield ~ treatment, data = farm)

$treatment
           diff        lwr      upr     p adj
10-0  0.7343754 -0.1279105 1.596661 0.1787670
20-0  0.9919266  0.1296407 1.854213 0.0525616
20-10 0.2575512 -0.6047348 1.119837 0.7954369
```

Based on this output, which pairwise comparisons, if any, are significantly different at the 10% level? Why do we use Tukey's Honest Significant Difference to do pairwise comparisons of treatments instead of multiple two sample t-tests at the 10% level? [2 marks]

- There is a significant difference between the 0 and 20 ton/hectare treatments at the 10% level.
- Doing multiple two sample t-tests will inflate the type I error rate.

(f) How might the assumptions underlying the one-way analysis of variance be violated in this study. [2 marks]

The assumptions of one-way anova are independence, constant variance and normality. There is no obvious departure from constant variance & normality in the plots given the sample size. (Maybe a lack of independence between neighbouring plots.)

END OF EXAMINATION

Formulas and Statistical Tables

BASICS

$$\overline{x} = \frac{\sum x_j}{n} \qquad s = \sqrt{\frac{\sum (x_j - \overline{x})^2}{n-1}}$$

STANDARDISING

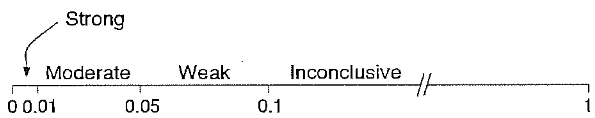
If $X \sim \text{Normal}(\mu, \sigma)$ then $Z = \frac{X-\mu}{\sigma} \sim \text{Normal}(0,1)$

BINOMIAL RANDOM VARIABLES

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$$ (but is usually available in tables)

$$\text{E}(X) = np \quad \text{sd}(X) = \sqrt{np(1-p)} \quad \hat{P} = \frac{X}{n} \quad \text{E}(\hat{P}) = p \quad \text{sd}(\hat{P}) = \sqrt{\frac{p(1-p)}{n}}$$

P-VALUES AND ERRORS



	Decision	
	Retain	Reject
H_0 is true	Correct $(1-\alpha)$	Type I Error (α)
H_0 is false	Type II Error (β)	Correct $(1-\beta)$

TESTS AND CONFIDENCE INTERVALS BASED ON STANDARD ERRORS

$$t = \frac{\text{estimate} - \text{hypothesised}}{\text{se(estimate)}} \qquad \text{estimate} \pm t^* \text{se(estimate)}$$

$$\text{se}(\overline{x}) = \frac{s}{\sqrt{n}} \qquad \text{se}(\overline{x}_1 - \overline{x}_2) = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \qquad \text{se}(r) = \sqrt{\frac{1-r^2}{n-2}}$$

$$\text{se}(\hat{p}) = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \qquad \text{se}(\hat{p}_1 - \hat{p}_2) = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

Use t for means, correlation and regression. Use z for proportions.

Regression

$$y = b_0 + b_1 x \qquad y = b_0 + b_1 x + b_2 x_1 \qquad x_1 = \begin{cases} 1, & \text{if Group B} \\ 0, & \text{if Group A} \end{cases}$$

Pooled Variance

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

ANOVA Tables

$$\text{DFT} = n - 1 \qquad \text{DFG} = k - 1$$

$$\text{MS} = \frac{\text{SS}}{\text{DF}} \qquad R^2 = \frac{\text{SSG}}{\text{SST}} \qquad s_p = \sqrt{MSR} \qquad F = \frac{\text{MSG}}{\text{MSR}}$$

Bonferroni Correction for k Comparisons

$$\alpha = \frac{0.05}{k}$$

Odds and Odds Ratios

$$\text{Odds} = \frac{p}{1-p} \qquad \text{OR} = \frac{\text{Odds for group B}}{\text{Odds for group A}} \qquad \text{se}(\ln(\text{OR})) = \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}$$

Chi-Squared Tests

$$\text{expected} = \frac{(\text{row total}) \times (\text{column total})}{\text{overall total}} \qquad \chi^2 = \sum \frac{(\text{observed} - \text{expected})^2}{\text{expected}}$$

$$\text{df} = (\#\text{ rows} - 1) \times (\#\text{ columns} - 1)$$

Sign Test

Count of positive values is $X \sim \text{Binomial}(n, 0.5)$

Signed-Rank Test

$S =$ sum of ranks corresponding to positive differences

Table 12.1: Standard Normal distribution

	Second decimal place of z									
z	0	1	2	3	4	5	6	7	8	9
0.0	0.500	0.496	0.492	0.488	0.484	0.480	0.476	0.472	0.468	0.464
0.1	0.460	0.456	0.452	0.448	0.444	0.440	0.436	0.433	0.429	0.425
0.2	0.421	0.417	0.413	0.409	0.405	0.401	0.397	0.394	0.390	0.386
0.3	0.382	0.378	0.374	0.371	0.367	0.363	0.359	0.356	0.352	0.348
0.4	0.345	0.341	0.337	0.334	0.330	0.326	0.323	0.319	0.316	0.312
0.5	0.309	0.305	0.302	0.298	0.295	0.291	0.288	0.284	0.281	0.278
0.6	0.274	0.271	0.268	0.264	0.261	0.258	0.255	0.251	0.248	0.245
0.7	0.242	0.239	0.236	0.233	0.230	0.227	0.224	0.221	0.218	0.215
0.8	0.212	0.209	0.206	0.203	0.200	0.198	0.195	0.192	0.189	0.187
0.9	0.184	0.181	0.179	0.176	0.174	0.171	0.169	0.166	0.164	0.161
1.0	0.159	0.156	0.154	0.152	0.149	0.147	0.145	0.142	0.140	0.138
1.1	0.136	0.133	0.131	0.129	0.127	0.125	0.123	0.121	0.119	0.117
1.2	0.115	0.113	0.111	0.109	0.107	0.106	0.104	0.102	0.100	0.099
1.3	0.097	0.095	0.093	0.092	0.090	0.089	0.087	0.085	0.084	0.082
1.4	0.081	0.079	0.078	0.076	0.075	0.074	0.072	0.071	0.069	0.068
1.5	0.067	0.066	0.064	0.063	0.062	0.061	0.059	0.058	0.057	0.056
1.6	0.055	0.054	0.053	0.052	0.051	0.049	0.048	0.047	0.046	0.046
1.7	0.045	0.044	0.043	0.042	0.041	0.040	0.039	0.038	0.038	0.037
1.8	0.036	0.035	0.034	0.034	0.033	0.032	0.031	0.031	0.030	0.029
1.9	0.029	0.028	0.027	0.027	0.026	0.026	0.025	0.024	0.024	0.023
2.0	0.023	0.022	0.022	0.021	0.021	0.020	0.020	0.019	0.019	0.018
2.1	0.018	0.017	0.017	0.017	0.016	0.016	0.015	0.015	0.015	0.014
2.2	0.014	0.014	0.013	0.013	0.013	0.012	0.012	0.012	0.011	0.011
2.3	0.011	0.010	0.010	0.010	0.010	0.009	0.009	0.009	0.009	0.008
2.4	0.008	0.008	0.008	0.008	0.007	0.007	0.007	0.007	0.007	0.006
2.5	0.006	0.006	0.006	0.006	0.006	0.005	0.005	0.005	0.005	0.005
2.6	0.005	0.005	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004
2.7	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003
2.8	0.003	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002
2.9	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.001	0.001	0.001
3.0	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
3.1	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
3.2	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
3.3										

This table gives $P(Z \geq z)$ for $Z \sim \text{Normal}(0,1)$. Critical values of the Normal distribution, the z^* values such that $P(Z \geq z^*) = p$ for a particular p, can be found from the ∞ row of Table 14.2.

Table 14.2: Critical values of Student's T distribution

		Probability p								
	df	0.25	0.10	0.05	0.025	0.01	0.005	0.001	0.0005	0.0001
	1	1.000	3.078	6.314	12.71	31.82	63.66	318.3	636.6	3183.1
	2	0.816	1.886	2.920	4.303	6.965	9.925	22.33	31.60	70.70
	3	0.765	1.638	2.353	3.182	4.541	5.841	10.21	12.92	22.20
	4	0.741	1.533	2.132	2.776	3.747	4.604	7.173	8.610	13.03
	5	0.727	1.476	2.015	2.571	3.365	4.032	5.893	6.869	9.678
	6	0.718	1.440	1.943	2.447	3.143	3.707	5.208	5.959	8.025
	7	0.711	1.415	1.895	2.365	2.998	3.499	4.785	5.408	7.063
	8	0.706	1.397	1.860	2.306	2.896	3.355	4.501	5.041	6.442
	9	0.703	1.383	1.833	2.262	2.821	3.250	4.297	4.781	6.010
	10	0.700	1.372	1.812	2.228	2.764	3.169	4.144	4.587	5.694
	11	0.697	1.363	1.796	2.201	2.718	3.106	4.025	4.437	5.453
Q1 (a) →	12	0.695	1.356	1.782	2.179	2.681	3.055	3.930	4.318	5.263
	13	0.694	1.350	1.771	2.160	2.650	3.012	3.852	4.221	5.111
	14	0.692	1.345	1.761	2.145	2.624	2.977	3.787	4.140	4.985
	15	0.691	1.341	1.753	2.131	2.602	2.947	3.733	4.073	4.880
	16	0.690	1.337	1.746	2.120	2.583	2.921	3.686	4.015	4.791
	17	0.689	1.333	1.740	2.110	2.567	2.898	3.646	3.965	4.714
	18	0.688	1.330	1.734	2.101	2.552	2.878	3.610	3.922	4.648
	19	0.688	1.328	1.729	2.093	2.539	2.861	3.579	3.883	4.590
	20	0.687	1.325	1.725	2.086	2.528	2.845	3.552	3.850	4.539
	21	0.686	1.323	1.721	2.080	2.518	2.831	3.527	3.819	4.493
	22	0.686	1.321	1.717	2.074	2.508	2.819	3.505	3.792	4.452
	23	0.685	1.319	1.714	2.069	2.500	2.807	3.485	3.768	4.415
	24	0.685	1.318	1.711	2.064	2.492	2.797	3.467	3.745	4.382
Q1 (b) →	25	0.684	1.316	1.708	2.060	2.485	2.787	3.450	3.725	4.352
	26	0.684	1.315	1.706	2.056	2.479	2.779	3.435	3.707	4.324
	27	0.684	1.314	1.703	2.052	2.473	2.771	3.421	3.690	4.299
	28	0.683	1.313	1.701	2.048	2.467	2.763	3.408	3.674	4.275
	29	0.683	1.311	1.699	2.045	2.462	2.756	3.396	3.659	4.254
Q2 (c) →	30	0.683	1.310	1.697	2.042	2.457	2.750	3.385	3.646	4.234
	40	0.681	1.303	1.684	2.021	2.423	2.704	3.307	3.551	4.094
	50	0.679	1.299	1.676	2.009	2.403	2.678	3.261	3.496	4.014
	60	0.679	1.296	1.671	2.000	2.390	2.660	3.232	3.460	3.962
	70	0.678	1.294	1.667	1.994	2.381	2.648	3.211	3.435	3.926
	80	0.678	1.292	1.664	1.990	2.374	2.639	3.195	3.416	3.899
	90	0.677	1.291	1.662	1.987	2.368	2.632	3.183	3.402	3.878
	100	0.677	1.290	1.660	1.984	2.364	2.626	3.174	3.390	3.862
	∞	0.674	1.282	1.645	1.960	2.326	2.576	3.090	3.291	3.719

This table gives t^* such that $P(T \geq t^*) = p$, where $T \sim$ Student(df).

Table 19.3: F distribution

d	p	n = 1	2	3	4	5	6	7	8	9
2	0.100	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38
	0.050	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4
	0.010	98.5	99.0	99.2	99.2	99.3	99.3	99.4	99.4	99.4
	0.001	999	999	999	999	999	999	999	999	999
3	0.100	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24
	0.050	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
	0.010	34.1	30.8	29.5	28.7	28.2	27.9	27.7	27.5	27.3
	0.001	167	148	141	137	135	133	132	131	130
4	0.100	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94
	0.050	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00
	0.010	21.2	18.0	16.7	16.0	15.5	15.2	15.0	14.8	14.7
	0.001	74.1	61.2	56.2	53.4	51.7	50.5	49.7	49.0	48.5
5	0.100	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32
	0.050	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77
	0.010	16.3	13.3	12.1	11.4	11.0	10.7	10.5	10.3	10.2
	0.001	47.2	37.1	33.2	31.1	29.8	28.8	28.2	27.6	27.2
6	0.100	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96
	0.050	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10
	0.010	13.7	10.9	9.78	9.15	8.75	8.47	8.26	8.10	7.98
	0.001	35.5	27.0	23.7	21.9	20.8	20.0	19.5	19.0	18.7
7	0.100	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72
	0.050	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
	0.010	12.2	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72
	0.001	29.2	21.7	18.8	17.2	16.2	15.5	15.0	14.6	14.3
8	0.100	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56
	0.050	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39
	0.010	11.3	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91
	0.001	25.4	18.5	15.8	14.4	13.5	12.9	12.4	12.0	11.8
9	0.100	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44
	0.050	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18
	0.010	10.6	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35
	0.001	22.9	16.4	13.9	12.6	11.7	11.1	10.7	10.4	10.1
10	0.100	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35
	0.050	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02
	0.010	10.0	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94
	0.001	21.0	14.9	12.6	11.3	10.5	9.93	9.52	9.20	8.96

This table gives f^* such that $P(F_{n,d} \geq f^*) = p$.

Table 19.3: F distribution (continued)

		n								
d	p	1	2	3	4	5	6	7	8	9
11	0.100	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27
	0.050	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90
	0.010	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63
	0.001	19.7	13.8	11.6	10.3	9.58	9.05	8.66	8.35	8.12
12	0.100	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21
	0.050	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80
	0.010	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39
	0.001	18.6	13.0	10.8	9.63	8.89	8.38	8.00	7.71	7.48
13	0.100	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.20	2.16
	0.050	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71
	0.010	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19
	0.001	17.8	12.3	10.2	9.07	8.35	7.86	7.49	7.21	6.98
14	0.100	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12
	0.050	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65
	0.010	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03
	0.001	17.1	11.8	9.73	8.62	7.92	7.44	7.08	6.80	6.58
15	0.100	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09
	0.050	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59
	0.010	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89
	0.001	16.6	11.3	9.34	8.25	7.57	7.09	6.74	6.47	6.26
16	0.100	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06
	0.050	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54
	0.010	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78
	0.001	16.1	11.0	9.01	7.94	7.27	6.80	6.46	6.19	5.98
17	0.100	3.03	2.64	2.44	2.31	2.22	2.15	2.10	2.06	2.03
	0.050	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49
	0.010	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68
	0.001	15.7	10.7	8.73	7.68	7.02	6.56	6.22	5.96	5.75
18	0.100	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	2.00
	0.050	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46
	0.010	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60
	0.001	15.4	10.4	8.49	7.46	6.81	6.35	6.02	5.76	5.56
19	0.100	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.98
	0.050	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42
	0.010	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52
	0.001	15.1	10.2	8.28	7.27	6.62	6.18	5.85	5.59	5.39
20	0.100	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96
	0.050	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39
	0.010	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46
	0.001	14.8	9.95	8.10	7.10	6.46	6.02	5.69	5.44	5.24

Table 19.3: F distribution (continued)

		n								
d	p	1	2	3	4	5	6	7	8	9
21	0.100	2.96	2.57	2.36	2.23	2.14	2.08	2.02	1.98	1.95
Q4(d) →	0.050	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37
	0.010	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40
	0.001	14.6	9.77	7.94	6.95	6.32	5.88	5.56	5.31	5.11
22	0.100	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93
	0.050	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34
	0.010	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35
	0.001	14.4	9.61	7.80	6.81	6.19	5.76	5.44	5.19	4.99
23	0.100	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.95	1.92
	0.050	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32
	0.010	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30
	0.001	14.2	9.47	7.67	6.70	6.08	5.65	5.33	5.09	4.89
24	0.100	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91
	0.050	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30
	0.010	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26
	0.001	14.0	9.34	7.55	6.59	5.98	5.55	5.23	4.99	4.80
25	0.100	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89
	0.050	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28
	0.010	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22
	0.001	13.9	9.22	7.45	6.49	5.89	5.46	5.15	4.91	4.71
26	0.100	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88
	0.050	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27
	0.010	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18
	0.001	13.7	9.12	7.36	6.41	5.80	5.38	5.07	4.83	4.64
27	0.100	2.90	2.51	2.30	2.17	2.07	2.00	1.95	1.91	1.87
	0.050	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25
	0.010	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15
	0.001	13.6	9.02	7.27	6.33	5.73	5.31	5.00	4.76	4.57
28	0.100	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87
	0.050	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24
	0.010	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12
	0.001	13.5	8.93	7.19	6.25	5.66	5.24	4.93	4.69	4.50
29	0.100	2.89	2.50	2.28	2.15	2.06	1.99	1.93	1.89	1.86
	0.050	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22
	0.010	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09
	0.001	13.4	8.85	7.12	6.19	5.59	5.18	4.87	4.64	4.45
30	0.100	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85
	0.050	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21
	0.010	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07
	0.001	13.3	8.77	7.05	6.12	5.53	5.12	4.82	4.58	4.39
∞	0.100	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63
	0.050	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88
	0.010	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41
	0.001	10.8	6.91	5.42	4.62	4.10	3.74	3.47	3.27	3.10

Table 22.4: χ^2 distribution

df	Probability p								
	0.975	0.95	0.25	0.10	0.05	0.025	0.01	0.005	0.001
1	0.001	0.004	1.323	2.706	3.841	5.024	6.635	7.879	10.83
Q3(e) 2	0.051	0.103	2.773	4.605	5.991	7.378	9.210	10.60	13.82
3	0.216	0.352	4.108	6.251	7.815	9.348	11.34	12.84	16.27
4	0.484	0.711	5.385	7.779	9.488	11.14	13.28	14.86	18.47
5	0.831	1.145	6.626	9.236	11.07	12.83	15.09	16.75	20.52
6	1.237	1.635	7.841	10.64	12.59	14.45	16.81	18.55	22.46
7	1.690	2.167	9.037	12.02	14.07	16.01	18.48	20.28	24.32
8	2.180	2.733	10.22	13.36	15.51	17.53	20.09	21.95	26.12
9	2.700	3.325	11.39	14.68	16.92	19.02	21.67	23.59	27.88
10	3.247	3.940	12.55	15.99	18.31	20.48	23.21	25.19	29.59
11	3.816	4.575	13.70	17.28	19.68	21.92	24.72	26.76	31.26
12	4.404	5.226	14.85	18.55	21.03	23.34	26.22	28.30	32.91
13	5.009	5.892	15.98	19.81	22.36	24.74	27.69	29.82	34.53
14	5.629	6.571	17.12	21.06	23.68	26.12	29.14	31.32	36.12
15	6.262	7.261	18.25	22.31	25.00	27.49	30.58	32.80	37.70
16	6.908	7.962	19.37	23.54	26.30	28.85	32.00	34.27	39.25
17	7.564	8.672	20.49	24.77	27.59	30.19	33.41	35.72	40.79
18	8.231	9.390	21.60	25.99	28.87	31.53	34.81	37.16	42.31
19	8.907	10.12	22.72	27.20	30.14	32.85	36.19	38.58	43.82
20	9.591	10.85	23.83	28.41	31.41	34.17	37.57	40.00	45.31
21	10.28	11.59	24.93	29.62	32.67	35.48	38.93	41.40	46.80
22	10.98	12.34	26.04	30.81	33.92	36.78	40.29	42.80	48.27
23	11.69	13.09	27.14	32.01	35.17	38.08	41.64	44.18	49.73
24	12.40	13.85	28.24	33.20	36.42	39.36	42.98	45.56	51.18
25	13.12	14.61	29.34	34.38	37.65	40.65	44.31	46.93	52.62
26	13.84	15.38	30.43	35.56	38.89	41.92	45.64	48.29	54.05
27	14.57	16.15	31.53	36.74	40.11	43.19	46.96	49.64	55.48
28	15.31	16.93	32.62	37.92	41.34	44.46	48.28	50.99	56.89
29	16.05	17.71	33.71	39.09	42.56	45.72	49.59	52.34	58.30
30	16.79	18.49	34.80	40.26	43.77	46.98	50.89	53.67	59.70
40	24.43	26.51	45.62	51.81	55.76	59.34	63.69	66.77	73.40
50	32.36	34.76	56.33	63.17	67.50	71.42	76.15	79.49	86.66
60	40.48	43.19	66.98	74.40	79.08	83.30	88.38	91.95	99.61
70	48.76	51.74	77.58	85.53	90.53	95.02	100.4	104.2	112.3
80	57.15	60.39	88.13	96.58	101.9	106.6	112.3	116.3	124.8
90	65.65	69.13	98.65	107.6	113.1	118.1	124.1	128.3	137.2
100	74.22	77.93	109.1	118.5	124.3	129.6	135.8	140.2	149.4

This table gives x^* such that $P(X^2 \geq x^*) = p$, where $X^2 \sim \chi^2(\text{df})$.