May 2015

Ph.D. COMPREHENSIVE EXAMINATIONS
DEPARTMENT OF STATISTICAL SCIENCES
UNIVERSITY OF TORONTO

APPLIED STATISTICS COMPREHENSIVE EXAMINATION

May 11, 2015, 12:30 p.m. – 4:30 p.m.

Sidney Smith Hall

(#questions = 6); (#pages = 19 including cover page and tables)

1. ATTEMPT ALL QUESTIONS.

2. Please work neatly and legibly.

3. **Start each question in a new book, with your name and the number of the question on the front cover.** If there is more than one book for a question, then also indicate which is the first book and which second, e.g., Jane Smith, Question 5, Book 1 of 2.

4. The questions are not in any special order, nor are they all of equal difficulty.

5. The problems may be improperly phrased or may contain a misprint. Should this happen, reflect it in your discussion. Faculty members are not available to answer questions during the exam.

6. You are NOT permitted any aids (e.g., books, notes, etc.) **aside from a single non programmable calculator.**

7. Good luck!

1. In a manufacturing process, it is acceptable for 2% of parts to be defective, but no more. When the process is running correctly, well under 1% of the parts are defective. Accordingly, the quality control engineers decide to take periodic random samples of parts and check how many are defective. They plan to test the null hypothesis that the probability of a defective part is 0.01 or less against the alternative that it is greater than 0.01, and if the null hypothesis is rejected at $\alpha = 0.05$, they will stop the production line and service the machines.

   (a) What is a reasonable model for this problem? Denote the observations by $Y_1, \ldots, Y_n$.

   (b) What is the parameter space?

   (c) State the null hypothesis in symbols.

   (d) The engineers decide to use the test statistic

   $$Z_n = \frac{\sqrt{n}(\overline{Y}_n - \theta_0)}{\sqrt{\theta_0(1-\theta_0)}}.$$

   What is the approximate distribution of the test statistic under the null hypothesis for large samples? Briefly justify your answer; you don't have to prove anything.

   (e) The engineers need to decide how many parts they will sample for testing. They want the smallest sample size that will ensure that if the true probability of a defective part is 0.02, the test will detect it with a probability of at least 0.90. What is the required sample size? The answer is a number. Show your work.

   (f) Why are the engineers unhappy when they see the answer?

2. An injection molding process that is producing an unacceptable percentage of burned parts is being studied by two engineers, Sally and Zhenhua. There are three two-level factors to be studied: injection pressure (P); screw RPM control (R); and injection speed (S). Sally conducts a factorial experiment and obtains the following results:

P (psi)	R (number of turns)	S	Percent Burned
1200	0.3	Slow	11
1200	0.3	Fast	17
1200	0.6	Slow	25
1200	0.6	Fast	29
1400	0.3	Slow	2
1400	0.3	Fast	9
1400	0.6	Slow	37
1400	0.6	Fast	40

Zhenhua claims that by designing a study to investigate one factor at a time it will require fewer observations to achieve the same precision (variance), compared to Sally's factorial experiment. Zhenhua proposes the following experimental design based on Sally's factorial experiment above.

Step 1. Factor P is thought to be the most important. By fixing the other factors at standard conditions (R=0.6, S=fast), two levels of P at 1200 and 1400 are compared. Here, P=1200 is chosen as it gives a smaller percent burned than at P=1400.

Step 2. The next most important factor is thought to be R. By fixing P=1200 from step 1 and S = fast (standard condition), the two levels of R at 0.3 and 0.6 are compared. Here, R=0.3 is chosen as it gives a smaller percent burned than at R=0.6.

Step 3. Vary the remaining factor S. Two levels of S are compared with P=1200 and R =0.3 based on steps 1 and 2. The S=slow level is chosen as it gives a smaller percent burned than at S=fast.

Suppose that the observations for each factorial run are independent and normally distributed with variance σ^2.

Answer the following questions.

(a) Show that the variance of any main effect in Sally's factorial design is $\sigma^2/2$.

(b) What is the smallest number of observations (experimental runs) that Zhenhua's design requires to obtain the same precision as Sally's factorial experiment in estimating all the main effects? Is the number of runs different than Sally's experiment? Explain your reasoning.

(c) Is it possible to estimate the interaction of pressure (P) and screw RPM control (R) using Zhenhua's design? If yes then estimate the interaction effect; otherwise estimate it using Sally's data.

(d) Did Zhenhua's approach miss any important factor settings that lead to the smallest percent burned? If he did miss an important factor settings then what are the factor settings and explain why his design might have missed these settings compared to Sally's approach. If he didn't miss any important factor settings then explain why the two designs lead to similar conclusions about which settings lead to the smallest percent burned.

3. The objective of *meta-analysis* is to estimate parameters or test hypotheses by combining the results of more than one statistical analysis. Typically the raw data are not available, and conclusions must be based on published statistics. Consider n independent tests of a null hypothesis or collection of null hypotheses. Some tests will provide stronger evidence against the null hypothesis than others. Let's draw an overall conclusion based on the p-values. Accordingly, let $P_1, \ldots, P_n$ denote the p-values from the n tests.

   (a) Assume that all the null hypotheses are true; call this the *combined* null hypothesis. Under the combined null hypothesis, what is the joint distribution of $P_1, \ldots, P_n$? Show your work. For simplicity, you may assume each null hypothesis would be rejected for large enough values of its respective test statistic, and that the cumulative distribution function of each test statistic is continuous and strictly increasing over its support.

   (b) Suppose we decide to base a test of the combined null hypothesis on the value of the lowest p-value, so that the combined null hypothesis will be rejected if the minimum p-value is small enough. What is the critical value assuming significance level α? Your answer is a formula. Show your work.

   (c) Give a formula for the p-value of the combined test based on the minimum p-value.

   (d) Here is a sample of ten p-values:

```
> sort(Pvalues)
 [1] 0.0233 0.0378 0.0523 0.0556 0.0641 0.0652 0.1494 0.1665 0.2721 0.3893
```

   For the test of the combined null hypothesis based on the minimum p-value,

      i. What is the critical value at $\alpha = 0.05$? The answer is a number.

      ii. Using this test, do you reject the combined null hypothesis at $\alpha = 0.05$? Answer Yes or No.

      iii. What is the p-value of the combined test? The answer is a number.

      iv. Is there strong evidence against the combined null hypothesis at this point?

   (e) In his *Statistical methods for research workers*, Fisher suggested a different test statistic: $Y = -2\sum_{i=1}^{n} \log(P_i)$. What is the distribution of Y under the combined null hypothesis? Show your work.

   (f) Compare the results of the calculation below to the appropriate table at the back of the exam:

```
> Y = -2*sum(log(Pvalues)); Y
[1] 48.58313
```

      i. What is the critical value for the combined test at $\alpha = 0.05$? The answer is a number.

      ii. What is the p-value of this combined test? You will need to give a range of numbers.

      iii. Does Fisher's test provide strong evidence against the combined null hypothesis?

(g) A serious problem in meta-analysis is that results that are not statistically significant tend to be under-reported. In fact, even when a few p-values over 0.05 are reported, they tend to be deliberately selected to make a point, and it is often safest to discard them. Accordingly, let $P_1, \ldots, P_n$ be a random sample of p-values, all less than α.

i. If the combined null hypothesis is true, what is the conditional distribution of P_j given $P_j < \alpha$? You need not show any work; just write down the answer.

ii. Modify Fisher's test to fit this situation. Write down a formula for the test statistic. What is its distribution under the the combined null hypothesis? Briefly explain in a sentence or two. You need not show all the work unless you really want to.

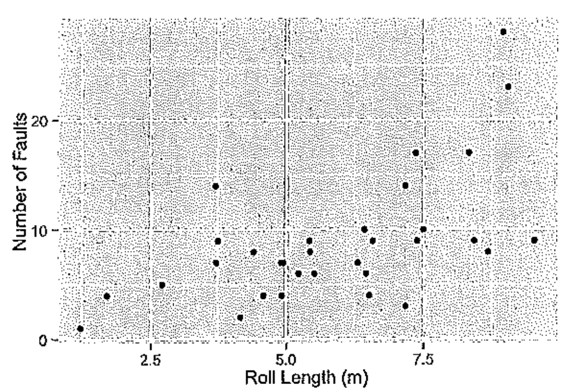
iii. A scientist mistakenly calculates Fisher's original test (not the modified version) on a sample of ten p-values that are all less than $\alpha = 0.05$ (see output that follows). It is a mistake because the journal only publishes results with $p < 0.05$. Correct the error; do it the easy way! Does the conclusion change when you use the modified test?

```
> sort(P)
 [1] 0.00068 0.00664 0.02199 0.02672 0.02976 0.03230 0.03256 0.03280
 [9] 0.03318 0.04658
> Y = -2*sum(log(P)); Y
[1] 80.01848
```

4. The data in Table 2 give the number of faults in rolls of textile fabric[1], followed by some R code. The questions follow the R code.

Table 2: Numbers of faults in rolls of textile fabric

Roll No.	Roll length (m)	No. of faults	Roll No.	Roll length (m)	No. of faults
1	5.51	6	17	5.43	8
2	6.51	4	18	8.42	9
3	8.32	17	19	9.05	23
4	3.75	9	20	5.42	9
5	7.15	14	21	5.22	6
6	8.68	8	22	1.22	1
7	2.71	5	23	6.57	9
8	6.30	7	24	1.70	4
9	4.91	7	25	7.38	9
10	3.72	7	26	3.71	14
11	6.45	6	27	7.35	17
12	4.41	8	28	7.49	10
13	8.95	28	29	4.95	7
14	4.58	4	30	7.16	3
15	6.42	10	31	9.52	9
16	4.92	4	32	4.17	2



[1]Set 2 of *Applied Statistics* by Cox & Snell, p.169

```
> data(cloth)
> head(cloth)
     x  y
1 1.22  1
2 1.70  4
3 2.71  5
4 3.71 14
5 3.72  7
6 3.75  9
> with(cloth, plot(x,y))

> model1 <- glm(y~x-1,family = poisson(link = identity), data = cloth)
> summary(model1)
...
Coefficients:
  Estimate Std. Error z value Pr(>|z|)
x  1.51024    0.08962   16.85   <2e-16 ***

...
(Dispersion parameter for poisson family taken to be 1)

    Null deviance:    Inf  on 32  degrees of freedom
Residual deviance: 64.537  on 31  degrees of freedom
AIC: 189.84

Number of Fisher Scoring iterations: 3

> model2 <- glm(y~x-1,family = quasipoisson(link = identity), data = cloth)
> summary(model2)

...
Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.8167  -1.1265  -0.2367   0.6274   3.4376

Coefficients:
  Estimate Std. Error t value Pr(>|t|)
x   1.5102     0.1328   11.38 1.35e-12 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

(Dispersion parameter for quasipoisson family taken to be 2.194371)

    Null deviance:    Inf  on 32  degrees of freedom
Residual deviance: 64.537  on 31  degrees of freedom
AIC: NA

Number of Fisher Scoring iterations: 3
```

(a) Write out in mathematical notation the model that is fit in `model1`, using y_i for the number of faults in the ith roll of cloth, and x_i for the length of the roll. Give an expression for the probability function for y_i, given x_i, the maximum likelihood estimator of the unknown parameter, and expression for the variance of the maximum likelihood estimator.

(b) Give an expression for the residual deviance from the fit of model 1. Can this be used as a measure of the fit of the Poisson model? Why or why not?

(c) In the output for `model2`, the dispersion parameter is estimated to be 2.19. Explain how this estimate was obtained, and what effect it has on the inference.

(d) In both `model1` and `model2` the systematic part of the model has no intercept. Does this seem appropriate in this context? Why or why not? What is the purpose of the command `link = identity`?

(e) Show that if it is assumed that the mean of the Poisson distribution follows a Gamma distribution with shape and scale parameters θ and ρ respectively, that the distribution of y_i is of the form

$$f(y_i \mid \theta, \rho) = \frac{\Gamma(y_i + \theta)}{\Gamma(y_i + 1)\Gamma(\theta)} \left(\frac{\rho}{\rho + x_i}\right)^{\theta} \left(\frac{x_i}{\rho + x_i}\right)^{y_i}, \quad y_i = 0, 1, 2, \dots. \qquad (1)$$

This is a negative binomial model with

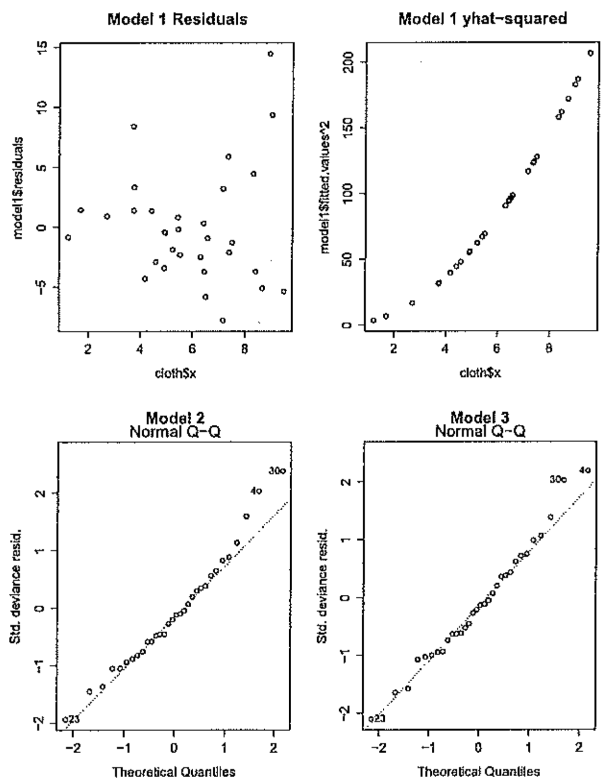
$$E(y_i) = \mu_i = \frac{\theta}{\rho} x_i, \quad \mathrm{Var}(y_i) = \mu_i + \frac{\mu_i^2}{\theta}.$$

(f) The summary of `model3` shows the fit of this negative binomial model.

```
> model3 = glm.nb(y ~ x - 1, link = identity, data = cloth)
> summary(model3)

...
Coefficients:
  Estimate Std. Error z value Pr(>|z|)
x   1.5105     0.1298   11.64   <2e-16 ***
---
...
    Null deviance:    Inf  on 32  degrees of freedom
Residual deviance: 30.795  on 31  degrees of freedom
AIC: 179.47

Number of Fisher Scoring iterations: 1


              Theta:  8.69
          Std. Err.:  4.21

 2 x log-likelihood:  -175.466
```

Which model seems more suitable to you for the analysis of this data, and why? Some plots which may be helpful are shown below.

Figure 1: Diagnostic Plots for Question 4



5. Research has shown that keyboard typing can be improved by typing the same words many times. This is sometimes referred to as the "learning effect". Suppose that two keyboards denoted by A and B are being compared in terms of typing efficiency (a keyboard is considered more "efficient" than another keyboard if the time to type a manuscript is shorter).

Six different manuscripts denoted by 1-6 are given to the same typist. The test is arranged in the following sequence:

1	2	3	4	5	6
A, B	A, B	A, B	A, B	A, B	A, B

Answer the following questions.

(a) How many times is this experiment replicated? Explain.

(b) Is blocking used in this experiment? If no, then explain why this is an unblocked experiment; if yes, what is the name of the blocking factor and how many levels does it have?

(c) Describe a serious design flaw in this experiment that will unfairly help the performance of keyboard B. How can the design flaw be eliminated from this experiment? Construct another experiment using six manuscripts and two keyboards that eliminates this flaw and would likely yield a better estimate of the difference in typing efficiency between the two keyboards. Explain your reasoning.

(d) A typist named Joe took the test described above. Each observation is the number of minutes it took Joe to type a manuscript on keyboards A and B.

Manuscript	A	B
1	19.00	13.84
2	20.26	16.43
3	19.84	13.35
4	21.77	14.28
5	20.23	15.18
6	20.64	15.19

The means (standard deviations) for keyboards A, B, and the difference between B and A are (respectively): 20.29(0.92); 14.71(1.11); −5.57(1.27).

Suppose that a typist similar to Joe is considering buying either keyboard A or B.

i. Which keyboard would you recommend? Explain your reasoning and perform any appropriate statistical tests.

ii. Suppose that you are the consulting statistician for this project. The project team asks you to recommend further testing to support which keyboard the team will recommend. Suggest at least one more experiment that would allow the team to make sound recommendations.

6. In a study reported in *Psychological Science* in 2014, the investigators described three experiments comparing students' retention of lecture material when they took notes on a laptop or in longhand (writing with a pen on paper). In one of the experiments, 65 students were randomly assigned to watch one of five TED talks, and assigned at random to the "longhand" or "laptop" instructions for taking notes. After watching the talks they were assigned some unrelated distractor tasks for about 15 minutes, and then asked to respond to a number of questions about the lecture they watched. The response of interest was the number of correct responses to these questions. The questions were divided into two groups: factual-recall questions and conceptual-application questions. The conceptual-application questions required more understanding of the material viewed.

   The authors state "mixed effects analysis of variance was used to test differences, with note-taking medium (laptop vs. longhand) as a fixed effect and lecture (which talk was viewed) as a random effect. ... On factual-recall questions, participants performed equally well across conditions; $F(1, 55) = 0.014$, $p = 0.91$. However, on conceptual-application questions, laptop participants performed significantly worse (mean $= -0.156$, SD $= 0.915$) than longhand participants (mean $= 0.154$, SD $=$ 1.08), $F(1, 55) = 9.99$, $p = 0.03$. Which lecture participants saw also affected performance on conceptual-application questions, $F(4, 55) = 12.52$, $p = 0.02$; however there was no significant interaction between lecture and note-taking medium, $F(4, 55) = 0.164$, $p = 0.96$." (The mean scores on the tests have been expressed in standardized units, so don't represent the actual number of correct responses.)

   (a) Assume 33 students used laptops and 32 students took notes in longhand, and that for each student one of the five lectures was randomly assigned, subject to balance with each lecture being used on 13 students. Complete the analysis of variance table below and **copy it to your answer book.**

Analysis of Variance

Source	df	F-statistic
	4	
Condition		
Interaction		
Error		
Total		——

(b) Write the mathematical form of the model, using the notation y_{ijk} for the score of the kth student on the jth test, under condition i ($1 =$ laptop; $2 =$ longhand). What are the assumptions on this model for the analysis in part (a)?

(c) Can you construct a two-sample t-test to compare the mean responses under conditions 1 (laptop) and 2 (longhand)? How does this differ from the F-test used by the authors?

(d) Why do you think the authors decided to model "talk" as a random effect, rather than a fixed effect? Do you agree with this decision?

(e) The general linear model for a mixed effects analysis has the form:

$$y = X\beta + Zb + \epsilon,$$

where y is an $n \times 1$ vector of responses, X is an $n \times p$ matrix of covariate values, β is a $p \times 1$ vector of fixed effect coefficients, Z is an $n \times q$ matrix of covariate values, b is a $q \times 1$ vector of random effects, and ϵ is an $n \times 1$ vector of random errors.

Assuming that $b \sim N(0, \Sigma_b)$ and $\epsilon \sim N(0, \sigma^2 I)$, show that the marginal density of y follows a normal distribution, and give an expression for its mean and covariance matrix.

(f) Mixed effects models are often used in longitudinal analysis, where subjects are repeatedly measured over time. Figure 2 shows the growth curves for 10 animals measured at five successive time points. The R code that follows gives two different analyses of the data. Compare and contrast the conclusions from the two analyses. Describe for a non-statistician the difference between the two analyses.

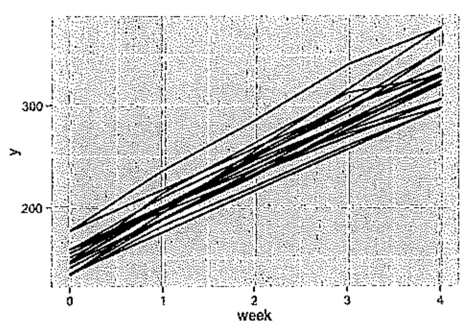


Figure 2: growth curves for 10 animals over 4 weeks

```
> growthcurves[1:10,]
  rat week   y
1   1    0 151
2   1    1 199
3   1    2 246
4   1    3 283
5   1    4 320
```

```
6     2     0 145
7     2     1 199
8     2     2 249
9     2     3 293
10    2     4 354

> is.factor(rat)
[1] TRUE
> is.factor(week)
[1] FALSE
> model1 <- lm(y ~ week + rat + week:rat, data = growthcurves)
> summary(model1)
...
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  155.400      5.455  28.486  < 2e-16 ***
week          42.200      2.227  18.948  < 2e-16 ***
rat2          -9.800      7.715  -1.270  0.21375
rat3           5.400      7.715   0.700  0.48936
rat4           5.600      7.715   0.726  0.47355
rat5         -17.800      7.715  -2.307  0.02812 *
rat6           8.200      7.715   1.063  0.29632
rat7         -10.000      7.715  -1.296  0.20479
rat8           2.400      7.715   0.311  0.75789
rat9          27.000      7.715   3.500  0.00148 **
rat10        -17.400      7.715  -2.255  0.03156 *
week:rat2      9.000      3.150   2.858  0.00769 **
week:rat3      3.800      3.150   1.207  0.23705
week:rat4     -6.600      3.150  -2.096  0.04467 *
week:rat5      4.600      3.150   1.461  0.15455
week:rat6      1.000      3.150   0.318  0.75306
week:rat7     -0.800      3.150  -0.254  0.80123
week:rat8      3.200      3.150   1.016  0.31775
week:rat9      8.000      3.150   2.540  0.01650 *
week:rat10    -2.000      3.150  -0.635  0.53024
---
...
> anova(model1)
Analysis of Variance Table

Response: y
          Df Sum Sq Mean Sq  F value    Pr(>F)
week       1 195541  195541 3942.3556 < 2.2e-16 ***
rat        9  14377    1597   32.2061 5.051e-13 ***
week:rat   9   1992     221    4.4632 0.0008918 ***
Residuals 30   1488      50
---

> model2 <- lmer(y ~ week + (week | rat), data = growthcurves)
> summary(model2)

Linear mixed model fit by REML ['lmerMod']
...
```

```
Random effects:
 Groups   Name        Variance Std.Dev. Corr
 rat      (Intercept) 156.25   12.500
          week         17.18    4.145   0.41
 Residual              49.60    7.043
Number of obs: 50, groups: rat, 10

Fixed effects:
            Estimate Std. Error t value
(Intercept)  154.760      4.313   35.88
week          44.220      1.488   29.72

Correlation of Fixed Effects:
     (Intr)
week 0.175

> anova(model2)
Analysis of Variance Table
     Df Sum Sq Mean Sq F value
week  1  43812   43812  883.31
```

TABLE A Table entry for z is under the standard normal curve to the left of z.

TABLE A Standard normal probabilities (*continued*)

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

CONT'D.

TABLE D *t* distribution critical values

	Upper tail probability p											
df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	0.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	0.765	0.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	0.741	0.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	0.727	0.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	0.718	0.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	0.711	0.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	0.706	0.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	0.703	0.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	0.700	0.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	0.697	0.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	0.695	0.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	0.694	0.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	0.692	0.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	0.691	0.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	0.690	0.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	0.689	0.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	0.688	0.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	0.688	0.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	0.687	0.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	0.686	0.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	0.686	0.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	0.685	0.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	0.685	0.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	0.684	0.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	0.684	0.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	0.684	0.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	0.683	0.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	0.683	0.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	0.683	0.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646
40	0.681	0.851	1.050	1.303	1.684	2.021	2.123	2.423	2.704	2.971	3.307	3.551
50	0.679	0.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678	2.937	3.261	3.496
60	0.679	0.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.915	3.232	3.460
80	0.678	0.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639	2.887	3.195	3.416
100	0.677	0.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.174	3.390
1000	0.675	0.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813	3.098	3.300
z^*	0.674	0.841	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.091	3.291
	50%	60%	70%	80%	90%	95%	96%	98%	99%	99.5%	99.8%	99.9%
	Confidence level C											

CONT'D

TABLE F Table entry for *p* is the critical value $(\chi^2)^*$ with probability *p* lying to its right

TABLE F χ^2 distribution critical values

	Tail probability *p*											
df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.32	1.64	2.07	2.71	3.84	5.02	5.41	6.63	7.88	9.14	10.83	12.12
2	2.77	3.22	3.79	4.61	5.99	7.38	7.82	9.21	10.60	11.98	13.82	15.20
3	4.11	4.64	5.32	6.25	7.81	9.35	9.84	11.34	12.84	14.32	16.27	17.73
4	5.39	5.99	6.74	7.78	9.49	11.14	11.67	13.28	14.86	16.42	18.47	20.00
5	6.63	7.29	8.12	9.24	11.07	12.83	13.39	15.09	16.75	18.39	20.51	22.11
6	7.84	8.56	9.45	10.64	12.59	14.45	15.03	16.81	18.55	20.25	22.46	24.10
7	9.04	9.80	10.75	12.02	14.07	16.01	16.62	18.48	20.28	22.04	24.32	26.02
8	10.22	11.03	12.03	13.36	15.51	17.53	18.17	20.09	21.95	23.77	26.12	27.87
9	11.39	12.24	13.29	14.68	16.92	19.02	19.68	21.67	23.59	25.46	27.88	29.67
10	12.55	13.44	14.53	15.99	18.31	20.48	21.16	23.21	25.19	27.11	29.59	31.42
11	13.70	14.63	15.77	17.28	19.68	21.92	22.62	24.72	26.76	28.73	31.26	33.14
12	14.85	15.81	16.99	18.55	21.03	23.34	24.05	26.22	28.30	30.32	32.91	34.82
13	15.98	16.98	18.20	19.81	22.36	24.74	25.47	27.69	29.82	31.88	34.53	36.48
14	17.12	18.15	19.41	21.06	23.68	26.12	26.87	29.14	31.32	33.43	36.12	38.11
15	18.25	19.31	20.60	22.31	25.00	27.49	28.26	30.58	32.80	34.95	37.70	39.72
16	19.37	20.47	21.79	23.54	26.30	28.85	29.63	32.00	34.27	36.46	39.25	41.31
17	20.49	21.61	22.98	24.77	27.59	30.19	31.00	33.41	35.72	37.95	40.79	42.88
18	21.60	22.76	24.16	25.99	28.87	31.53	32.35	34.81	37.16	39.42	42.31	44.43
19	22.72	23.90	25.33	27.20	30.14	32.85	33.69	36.19	38.58	40.88	43.82	45.97
20	23.83	25.04	26.50	28.41	31.41	34.17	35.02	37.57	40.00	42.34	45.31	47.50
21	24.93	26.17	27.66	29.62	32.67	35.48	36.34	38.93	41.40	43.78	46.80	49.01
22	26.04	27.30	28.82	30.81	33.92	36.78	37.66	40.29	42.80	45.20	48.27	50.51
23	27.14	28.43	29.98	32.01	35.17	38.08	38.97	41.64	44.18	46.62	49.73	52.00
24	28.24	29.55	31.13	33.20	36.42	39.36	40.27	42.98	45.56	48.03	51.18	53.48
25	29.34	30.68	32.28	34.38	37.65	40.65	41.57	44.31	46.93	49.44	52.62	54.95
26	30.43	31.79	33.43	35.56	38.89	41.92	42.86	45.64	48.29	50.83	54.05	56.41
27	31.53	32.91	34.57	36.74	40.11	43.19	44.14	46.96	49.64	52.22	55.48	57.86
28	32.62	34.03	35.71	37.92	41.34	44.46	45.42	48.28	50.99	53.59	56.89	59.30
29	33.71	35.14	36.85	39.09	42.56	45.72	46.69	49.59	52.34	54.97	58.30	60.73
30	34.80	36.25	37.99	40.26	43.77	46.98	47.96	50.89	53.67	56.33	59.70	62.16
40	45.62	47.27	49.24	51.81	55.76	59.34	60.44	63.69	66.77	69.70	73.40	76.09
50	56.33	58.16	60.35	63.17	67.50	71.42	72.61	76.15	79.49	82.66	86.66	89.56
60	66.98	68.97	71.34	74.40	79.08	83.30	84.58	88.38	91.95	95.34	99.61	102.7
80	88.13	90.41	93.11	96.58	101.9	106.6	108.1	112.3	116.3	120.1	124.8	128.3
100	109.1	111.7	114.7	118.5	124.3	129.6	131.1	135.8	140.2	144.3	149.4	153.2