Exam Empirical Methods

VU University Amsterdam, Faculty of Exact Sciences

15.15 – 18.00h, December 16, 2014

- Always motivate your answers.
- Write your answers in English.
- Only the use of a simple, non-graphical calculator is allowed.
- Programmable/graphical calculators, laptops, mobile phones, smart watches, books, own formula sheets, etc. are not allowed.
- On the last four pages of the exam, some formulas and tables that you may want to use can be found.
- The total number of points you can receive is 90: Grade $= 1 + \dfrac{\text{points}}{10}$.
- The division of points per question and subparts is as follows:

Question	1	2	3	4	5	6	7
Part a)	2	4	3	2	10	2	2
Part b)	2	3	3	8	2	6	1
Part c)	2	3	5	3	-	4	6
Part d)	2	3	2	-	-	2	4
Part e)	-	-	2	-	-	-	2
Total	8	13	15	13	12	14	15

- If you are asked to perform a test, do not only give the conclusion of your test, but report:
  1. the hypotheses in terms of the population parameter of interest;
  2. the significance level;
  3. the test statistic and its distribution under the null hypothesis;
  4. the observed value of the test statistic;
  5. the P-value or the critical value(s);
  6. whether or not the null hypothesis is rejected and why;
  7. finally, phrase your conclusion in terms of the context of the problem.

1. Are the following statements sensible/correct? Briefly motivate your answer.

   a) For data at the nominal level of measurement a Pareto bar chart is a better visualisation type than a pie chart.

   b) In a stratified sample, subjects are divided into sections, then some of these sections are randomly selected and *all* subjects in these selected sections are chosen.

   c) Outside temperatures (in °C) are at the ratio level of measurement.

   d) The median of a dataset of size $n \geq 3$ is always larger than the mean.

2. The weather on a particular day is classified as cold, mild or warm. There is a probability of 0.30 that it is cold and a probability of 0.45 that it is mild. In addition, on each day it may either rain or not rain. On cold days there is a probability of 0.30 that it will rain, on mild days there is a probability of 0.10 that it will rain and on warm days there is a probability of 0.05 that it will rain.
   For the questions below: show how your answer was obtained and name the rules or properties you use.

   a) Show that the probability that it rains on a particular day equals 0.1475.

   b) Are the events $A = \{\text{mild weather}\}$ and $B = \{\text{no rain}\}$ independent events?

   c) Compute the probability that it is either cold or it rains on a particular day.

   d) If it is raining on a particular day, what is the probability that it is warm?

3. Assume that the amount of sugar contained in 1-kg packs is normally distributed with a mean of $\mu = 1.01$ kg and a standard deviation of $\sigma = 0.012$.

   a) What is the probability that a single pack of sugar contains less than 1.00 kg of sugar?

   b) What is the probability that the mean weight of a random sample of $n = 16$ sugar packs is more than 1.00 kg?

   Now assume that the amount of sugar contained in 1-kg packs from company A is normally distributed with unknown mean μ and unknown standard deviation σ. The weight of $n = 25$ randomly selected sugar packets from company A is measured and the sample mean equals $\overline{x} = 1.005$ and the sample standard deviation $s = 0.008$.

   c) Construct a 90% confidence interval for μ.

   d) What is the interpretation of the confidence interval obtained in part c)?

   e) Based on your answer of part c), could company A argue that the population mean of the sugar packs they produce equals 1.01?

4. A random-number generator is supposed to produce a sequence of 0s and 1s with each value being equally likely to be a 0 or a 1 and all values being independent. In an examination of the random-number generator, a sequence of 50,000 values is obtained of which 25,264 are 0s.

   a) What is your population parameter of interest if you want to test whether the random-number generator produces 0s and 1s with equal probability? Also give a point estimate for it.

b) Use the P-value method to test the claim that the random-number generator produces 0s and 1s with equal probability at significance level $\alpha = 0.01$.

c) If a 99% confidence interval with margin of error $E = 0.002$ were required for your chosen population parameter of part a), how many values should be investigated?

5. A new teaching method for a statistics course is being evaluated. A set of 182 students is randomly split up in two groups, 1 and 2, each consisting of 91 students. In Group 1 the standard teaching method is used, while in group 2 the new teaching method is tried. At the end of the course all students take the same exam (grade between 1 and 10, 1:worst, 10:best). Some sample statistics regarding the exam scores which you may or may not use in your analysis are shown below:
$\overline{x}_1 = 6.13, s_1 = 1.12, \overline{x}_2 = 6.66, s_2 = 1.91, s_p = 1.57.$

   a) Test with a suitable hypothesis test (motivate your choice) the claim that the new teaching method is better than the standard teaching method. Take significance level 0.05.

   b) The test you performed in part a) should only be used if certain requirements are met. What are these requirements and are they met in this case?

6. Three drugs are compared with respect to whether or not they cause an allergic reaction to patients. A group of $n = 300$ patients is randomly split into into three groups of 100 patients, each of which is given one of the three drugs. The results of this experiment are in the table below:

	Allergic reaction	No allergic reaction	Total
Drug A	77	23	100
Drug B	64	36	100
Drug C	69	31	100
Total	210	90	300

   a) In order to investigate whether the three drugs can be considered equivalent in terms of allergic reactions they cause, should you use a test of independence or a test of homogeneity?

   b) Using a chi-square test, investigate the claim that the three drugs can be considered equivalent in terms of allergic reactions they cause. Take significance level $\alpha = 0.01$. *The observed value of the test-statistic is 4.10, so you do not have to compute this value!*

   c) The test in part b) should only be used under certain conditions. What are these conditions and are they satisfied in this case?

   d) Could the Fisher exact test be used in this case to test the claim that drug A causes less allergic reactions than drug B?

7. The download times (in milliseconds) of 9 randomly selected files were measured. Furthermore, the file sizes are measured in MB. The file sizes and the download times are stored in respective datasets x and y. A linear regression analysis was carried out with

explanatory variable 'file size' and response variable 'download time'. Some sample statistics of the data that you may or may not use are:

$$\overline{x} = 5.56, \overline{y} = 128.44, s_x = 2.26, s_y = 27.47, r = 0.93, \sqrt{\frac{1-r^2}{n-2}} = 0.135,$$
$$b_0 = 65.34, s_{b_0} = 9.76, b_1 = 11.34, s_{b_1} = 1.64.$$

Furthermore, a scatterplot of the downloading times against the file sizes is shown in the left graph of Figure 1 below. The middle graph shows a normal Q-Q plot of the residuals of the regression analysis and the right graph shows a residual plot of the residuals against the values of the x variable, i.e. the file sizes.

a) Give the regression equation. What is the predicted download time for a file of size 5.0 MB?

b) What proportion of the variation in the y variable can be accounted for by the regression equation?

c) Test the claim that $\beta_1 = 0$, i.e. that there is no linear relationship between the explanatory variable 'file size' and response variable 'download time'. Take significance level $\alpha = 0.05$.

d) For the test in part c) certain requirements about the errors have to be met. For instance, the errors should be indepedent, which may be assumed. What are the remaining requirements and is it reasonable to assume that they are indeed met?

e) In view of the scatterplot and your answers of parts b), c) and d), do you judge that the linear regression model is an appropriate model for these data?

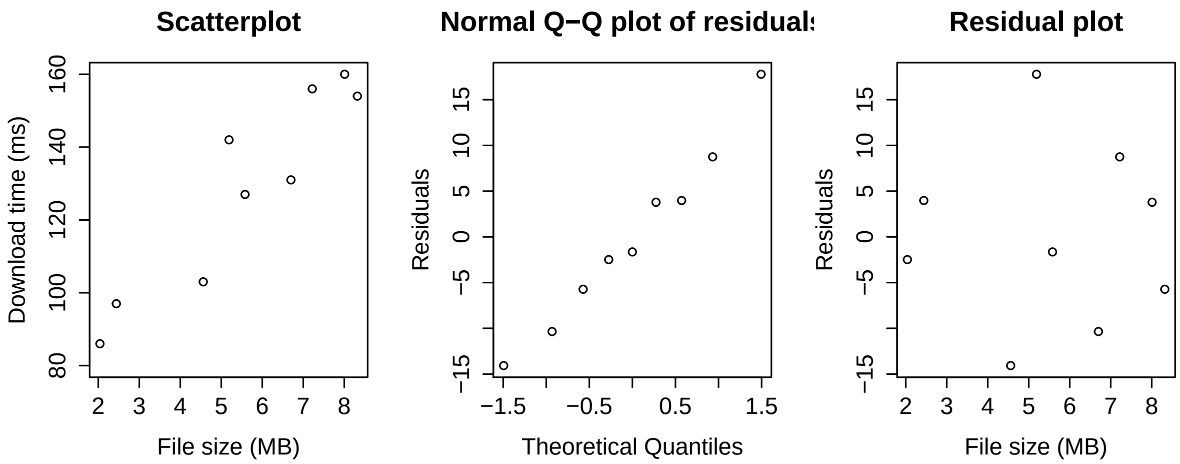


Figure 1: Scatterplot, normal Q-Q plot of residuals and residual plot.

Formulas and Tables for Exam Empirical Methods

Probability

We use the following notation:
Ω sample space, P probability measure.
$B, A_1, A_2, \dots, A_m$ events,
$A_1, A_2, \dots, A_m$ a partition of Ω with $P(A_i) > 0$ for all $i \in \{1, 2, \dots, m\}$.

Law of Total Probability:

$$P(B) = \sum_{i=1}^{m} P(B \cap A_i) = \sum_{i=1}^{m} P(B|A_i)P(A_i).$$

Bayes' Theorem:

$$P(A_r|B) = \frac{P(A_r \cap B)}{\sum_{i=1}^{m} P(B|A_i)P(A_i)} = \frac{P(B|A_r)P(A_r)}{\sum_{i=1}^{m} P(B|A_i)P(A_i)}.$$

Two *independent* samples

(The statements below hold if certain requirements are met.)

For two *independent* samples,
(i) if σ_1 and σ_2 are unknown and $\sigma_1 \neq \sigma_2$, the test statistic

$$T_2 = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}$$

has a t-distribution with approximately $\tilde{n}$ degrees of freedom under the null hypothesis. We use the conservative estimate $\tilde{n} = \min\{n_1 - 1, n_2 - 1\}$.
(ii) if σ_1 and σ_2 are unknown and $\sigma_1 = \sigma_2$, then the test statistic

$$T_2^{\text{eq}} = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s_p^2/n_1 + s_p^2/n_2}}$$

has a t-distribution with $n_1 + n_2 - 2$ degrees of freedom under the null hypothesis. Here s_p is the square root of the pooled sample variance s_p^2 given by

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

(iii) if σ_1 and σ_2 are known, then the test statistic

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$$

has a standard normal distribution under the null hypothesis.
(iv) if $p_1 = p_2$, the test statistic

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\bar{p}(1-\bar{p})/n_1 + \bar{p}(1-\bar{p})/n_2}}$$

approximately has a standard normal distribution. Here $\bar{p} = (x_1 + x_2)/(n_1 + n_2)$ is the pooled sample proportion.
(v) the margin of error for a $1 - \alpha$ confidence interval for $p_1 - p_2$ is given by

$$E = z_{\alpha/2}\sqrt{\hat{p}_1(1-\hat{p}_1)/n_1 + \hat{p}_2(1-\hat{p}_2)/n_2}.$$

Correlation

Under certain conditions the test statistic

$$T_{cor} = \frac{r - \rho}{\sqrt{(1-r^2)/(n-2)}}$$

has a t-distribution with $n-2$ degrees of freedom. Here ρ is the population linear correlation coefficient and r is the sample linear correlation coefficient given by

$$r = \frac{1}{n-1}\sum_{i=1}^{n}\Big[\frac{(x_i - \bar{x})(y_i - \bar{y})}{s_x s_y}\Big].$$

Linear regression

Let β_0 be the unknown intercept and β_1 the unknown slope of a linear regression model with one explanatory variable, and let b_0 and b_1 be the corresponding estimators, i.e. the intercept and slope of the regression line (the 'best' line). Then b_0 and b_1 are given by

$$b_1 = r\frac{s_y}{s_x}$$

and

$$b_0 = \bar{y} - b_1\bar{x}.$$

If certain requirements are met, then the test statistic

$$T_1 = \frac{b_1 - \beta_1}{s_{b_1}}$$

has a t-distribution with $n-2$ degrees of freedom. Here s_{b_1} is the standard error (i.e. estimated standard deviation) of the estimator b_1.

NEGATIVE *z* Scores

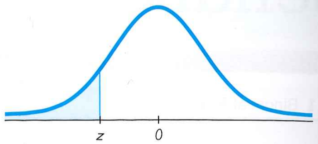


Table 2 Standard Normal (*z*) Distribution: Cumulative Area from the LEFT

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.50 and lower	.0001									
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051 *	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505 *	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
−1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
−1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
−0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
−0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
−0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
−0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
−0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
−0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
−0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
−0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
−0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
−0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

NOTE: For values of *z* below 3.49, use 0.0001 for the area.
*Use these common values that result from interpolation:

z Score	Area
−1.645	0.0500
−2.575	0.0050

(*continued*)

POSITIVE *z* Scores

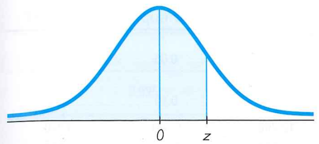


Table 2 (*continued*) Cumulative Area from the LEFT

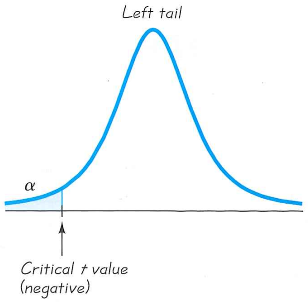
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.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
3.50 and up	.9999									

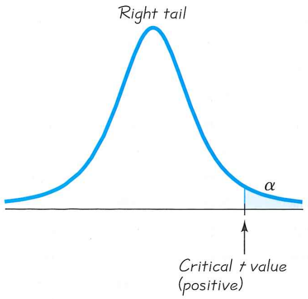
NOTE: For values of *z* above 3.49, use 0.9999 for the area.
*Use these common values that result from interpolation:

z Score	Area
1.645	0.9500
2.575	0.9950

Common Critical Values

Confidence Level	Critical Value
0.90	1.645
0.95	1.96
0.99	2.575





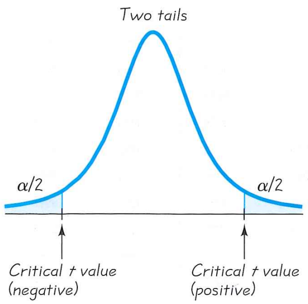


Table 3 *t* Distribution: Critical *t* Values

	Area in One Tail 0.005	0.01	0.025	0.05	0.10
Degrees of Freedom	Area in Two Tails 0.01	0.02	0.05	0.10	0.20
1	**63.657**	31.821	**12.706**	6.314	3.078
2	**9.925**	6.965	**4.303**	2.920	1.886
3	**5.841**	4.541	**3.182**	2.353	1.638
4	**4.604**	3.747	**2.776**	2.132	1.533
5	**4.032**	3.365	**2.571**	2.015	1.476
6	**3.707**	3.143	**2.447**	1.943	1.440
7	**3.499**	2.998	**2.365**	1.895	1.415
8	**3.355**	2.896	**2.306**	1.860	1.397
9	**3.250**	2.821	**2.262**	1.833	1.383
10	**3.169**	2.764	**2.228**	1.812	1.372
11	**3.106**	2.718	**2.201**	1.796	1.363
12	**3.055**	2.681	**2.179**	1.782	1.356
13	**3.012**	2.650	**2.160**	1.771	1.350
14	**2.977**	2.624	**2.145**	1.761	1.345
15	**2.947**	2.602	**2.131**	1.753	1.341
16	**2.921**	2.583	**2.120**	1.746	1.337
17	**2.898**	2.567	**2.110**	1.740	1.333
18	**2.878**	2.552	**2.101**	1.734	1.330
19	**2.861**	2.539	**2.093**	1.729	1.328
20	**2.845**	2.528	**2.086**	1.725	1.325
21	**2.831**	2.518	**2.080**	1.721	1.323
22	**2.819**	2.508	**2.074**	1.717	1.321
23	**2.807**	2.500	**2.069**	1.714	1.319
24	**2.797**	2.492	**2.064**	1.711	1.318
25	**2.787**	2.485	**2.060**	1.708	1.316
26	**2.779**	2.479	**2.056**	1.706	1.315
27	**2.771**	2.473	**2.052**	1.703	1.314
28	**2.763**	2.467	**2.048**	1.701	1.313
29	**2.756**	2.462	**2.045**	1.699	1.311
30	**2.750**	2.457	**2.042**	1.697	1.310
31	**2.744**	2.453	**2.040**	1.696	1.309
32	**2.738**	2.449	**2.037**	1.694	1.309
33	**2.733**	2.445	**2.035**	1.692	1.308
34	**2.728**	2.441	**2.032**	1.691	1.307
35	**2.724**	2.438	**2.030**	1.690	1.306
36	**2.719**	2.434	**2.028**	1.688	1.306
37	**2.715**	2.431	**2.026**	1.687	1.305
38	**2.712**	2.429	**2.024**	1.686	1.304
39	**2.708**	2.426	**2.023**	1.685	1.304
40	**2.704**	2.423	**2.021**	1.684	1.303
45	**2.690**	2.412	**2.014**	1.679	1.301
50	**2.678**	2.403	**2.009**	1.676	1.299
60	**2.660**	2.390	**2.000**	1.671	1.296
70	**2.648**	2.381	**1.994**	1.667	1.294
80	**2.639**	2.374	**1.990**	1.664	1.292
90	**2.632**	2.368	**1.987**	1.662	1.291
100	**2.626**	2.364	**1.984**	1.660	1.290
200	**2.601**	2.345	**1.972**	1.653	1.286
300	**2.592**	2.339	**1.968**	1.650	1.284
400	**2.588**	2.336	**1.966**	1.649	1.284
500	**2.586**	2.334	**1.965**	1.648	1.283
1000	**2.581**	2.330	**1.962**	1.646	1.282
2000	**2.578**	2.328	**1.961**	1.646	1.282
Large	**2.576**	2.326	**1.960**	1.645	1.282

Table 4 Chi-Square (χ^2) Distribution

Degrees of Freedom	Area to the Right of the Critical Value 0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
1	—	—	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.071	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.299
13	3.565	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	14.042	30.813	33.924	36.781	40.289	42.796
23	9.260	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.559
25	10.520	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314	46.928
26	11.160	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290
27	11.808	12.879	14.573	16.151	18.114	36.741	40.113	43.194	46.963	49.645
28	12.461	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.993
29	13.121	14.257	16.047	17.708	19.768	39.087	42.557	45.722	49.588	52.336
30	13.787	14.954	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672
40	20.707	22.164	24.433	26.509	29.051	51.805	55.758	59.342	63.691	66.766
50	27.991	29.707	32.357	34.764	37.689	63.167	67.505	71.420	76.154	79.490
60	35.534	37.485	40.482	43.188	46.459	74.397	79.082	83.298	88.379	91.952
70	43.275	45.442	48.758	51.739	55.329	85.527	90.531	95.023	100.425	104.215
80	51.172	53.540	57.153	60.391	64.278	96.578	101.879	106.629	112.329	116.321
90	59.196	61.754	65.647	69.126	73.291	107.565	113.145	118.136	124.116	128.299
100	67.328	70.065	74.222	77.929	82.358	118.498	124.342	129.561	135.807	140.169

Source: Donald B. Owen, *Handbook of Statistical Tables.*

Degrees of Freedom

$n - 1$	**Confidence interval or hypothesis test** for a standard deviation σ or variance σ^2
$k - 1$	**Goodness-of-fit** with *k* categories
$(r - 1)(c - 1)$	**Contingency table** with *r* rows and *c* columns
$k - 1$	**Kruskal-Wallis test** with *k* samples