Resit Exam Statistical Methods

Vrije Universiteit Amsterdam, Faculty of Science

18.45 – 21.30h, February 22, 2023

- In general: **always motivate your answers and write down your calculations precisely.**
- Write your answers in English.
- Only the use of a simple, non-graphical, non-programmable calculator is allowed.
- On the last five pages of the exam, you can find some helpful formulas and tables.
- The total number of points you can receive is 70: $\text{Grade} = 1 + \frac{\text{points}}{7.77}$.
- The division of points per question and subparts is as follows:

Question	1	2	3	4	5	6
Part a)	3	3	2	2	1	3
Part b)	3	3	2	2	8	8
Part c)	3	2	2	2	3	3
Part d)	–	3	2	6	–	2
Part e)	–	–	2	–	–	–
Totals	9	11	10	12	12	16

- 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 (for which it is also important to check whether the conditions are met!);
  4. the observed value of the test statistic;
  5. the P-value or the critical value(s) (if a question states which method you should use, you are not free to choose);
  6. whether or not the null hypothesis is rejected and why;
  7. finally, phrase your conclusion in terms of the context of the problem.

Good luck!

1. Ava and Eve are identical twin sisters, but Ava tells the truth with a probability of 0.75 and Eve tells the truth with a probability of 0.25 (independently of each other). Suppose you meet one of them on the street. Without any additional information, it is Ava or Eve, each with a probability of 50%. However, when you ask her: "Are you Ava?", she replies with "Yes". Solve the following problems.

   a) For an encounter with a randomly selected sister, find the probability that the person you have met will reply with "Yes" to the question "Are you Ava?". *Note: you are not required to model the probability space.*

   b) Find the conditional probability that you have truly encountered Ava taking into account that the encountered sister had replied with "Yes" to the question "Are you Ava?". *Note: if you were not able to solve a), you may use for part b) the wrong probability $P(A) = 0.61$.*

   c) Find the probability to encounter Eve and that the encountered sister replies with "No" to the question "Are you Ava?".

   Note: always state which theorems/laws you use, and explicitly write down your calculations. Also write down the formulas you are using, so that your usage of them is comprehensible.

---

2. Suppose that a (very) small telecom company hires you to analyze their call center – in particular the waiting times for the customers who call. The call center consists of only four employees with initials A., B., C., and D.; for simplicity, we assume that they are working around the clock, without taking breaks, tiring, taking holidays etc. A customer who calls will be connected to A., B., C., and D., with the probabilities 10%, 25%, 30%, and 35%, respectively. In half of the cases, each employee needs exactly 10 minutes to solve a case. In the other half of the cases, A., B., C., and D. need 40, 20, 15, and 5 minutes, respectively.

   a) Determine the probability space that models the waiting time of a customer who calls.

      Tip: you could, for example, fill an appropriate table for the probabilities related to the different waiting times. But still provide explicit calculations that show how you derived them in detail.

   b) Calculate the expected waiting time for a customer.

      Note: only if you were not able to find the probabilities in a), make up some probabilities to proceed with part b), but do not choose too trivial probabilities.

   c) Suppose that 10,000 customers call per year. Find an approximation for the sample average of exactly those 10,000 customer waiting times.

   d) For the 10,000 customers mentioned in part c), find an approximation for the sampling distribution of the random variable that models the average of those 10,000 customer waiting times. Use that the standard deviation of the waiting time of a single customer is $\sigma \approx 7.66$.

   Note: only if you were not able to find the expected value in b), you could use the wrong *number 22.5 as the expected waiting time of one customer in c) and d).*

   Note: always state which theorems/laws you use, and explicitly write down your calculations. Also write down the formulas you are using, so that your usage of them is comprehensible.

---

3. This exam question touches upon miscellaneous topics of the course.

   a) Name and describe two sampling techniques for collecting data for a study.

      Note: we will not judge here how useful or appropriate these techniques are.

   b) Describe what a histogram is and what kind of data it can illustrate.

   c) For a normally distributed random variable X with mean $\mu = -3$ and standard deviation $\sigma = 2$, find the probability that X is bigger than -2.5.

   d) Describe in your own words the main differences between normal distributions and chi-squared distributions.

   e) Describe in your own words what p-values are and what they are used for.

   Note: always state which theorems/laws you use, and explicitly write down your calculations. Also write down the formulas you are using, so that your usage of them is comprehensible.

---

4. In many parts of the world, including the UK, tea is a big deal, and in particular, it matters whether one pours the milk first in the cup and after that the tea, or the tea first and then add the milk. A British lady claims to be able to tell whether the milk or the tea was added first to her teacup. To test this, she was given 30 cups of tea, of which 15 had milk added first and 15 had tea added first. The lady had to choose which were which. She did not know the number of cups that had milk or tea put in first, and she was not told of success along the way. Her guesses are in the contingency table below.

first	poured	actual milk	actual tea	
guessed	milk	10	3	13
	tea	5	12	17
		15	15	30

   a) Which test should be used to test her claim that she can tell whether milk or tea was poured first in her teacup?

   Suppose that instead we want to investigate the null hypothesis: "The *actual* and *guessed* first pours of milk and tea are independent." against the undirected alternative hypothesis of dependence.

   b) Compute the table of expected frequencies in case the null hypothesis is true.

   c) To test the above-stated hypotheses, certain requirements have to be met. Describe the requirements and whether they are met.

   d) Perform a suitable test to test the above-stated hypotheses. Take significance level $\alpha = 0.01$. You don't have to restate the hypotheses.
      (See the first page of this exam for detailed instructions about testing.)

---

5. A couple of years ago, a company started making smartphones which are as sustainable and fair (e.g. without child labour) as possible. Let's call the product NicePhone. The company grew and extended to many countries in Europe, and now wants to investigate how many people who own a smartphone are aware of the existence of the brand NicePhone. For this analysis, the researcher asked 150 German and 200 Dutch citizens which were randomly chosen and own at least one smartphone. In the German sample, 29 people knew the brand NicePhone; in the Dutch sample, 64 people knew the brand NicePhone. Let p_1 denote the proportion of smart phone users in Germany who know the brand NicePhone, and let p_2 denote the proportion of smart phone users in the Netherlands who know the brand NicePhone.

   a) Give, based on the data, a point estimate for the difference between the proportions of smart phone users who know the brand NicePhone in Germany and in the Netherlands.

b) Perform a suitable test to investigate the claim that there is a difference in the proportions of smart phone users who know the brand NicePhone between Germans and Dutch people. Take significance level $\alpha = 0.01$.
(See the first page of this exam for detailed instructions about testing.)

c) In general, what are the differences in testing population proportions and population means (both in two samples), i.e., how are the statistical tests different from one another? Explain at least three differences.

---

6. During the pandemic, lecture videos were made about the different topics of a course. The lecturer wants to investigate whether the number of questions asked by students on Canvas about a video depends on the length of the video. The number of questions and the length of a video were measured for 30 videos and stored in the respective data set y and x. A linear regression analysis was carried out with explanatory variable 'length' and response variable 'number of questions'. Some characteristics of the data that you may or may not use are:

$$\overline{x} = 136.701, \quad \overline{y} = 39.325, \quad s_x = 26.047, \quad s_y = 11.700,$$

$$s_{b_0} = 10.859, \quad s_{b_1} = 0.078, \quad r = 0.392, \quad \sqrt{\frac{1-r^2}{n-2}} = 0.174.$$

Furthermore, a scatter plot of the number of questions against the length of a video and the 'best-fit' line, as well as a normal QQ plot of the residuals of a linear regression of 'number of questions' on 'length' are shown in Figure 1.

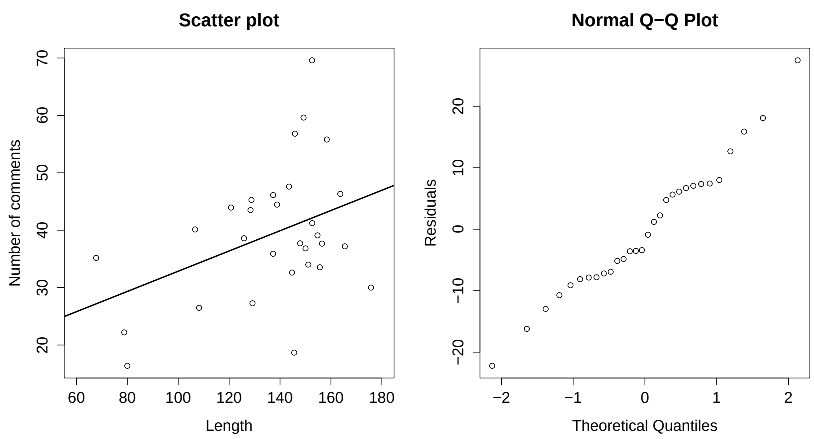


Figure 1: Scatter plot and normal QQ plot of residuals.

a) Based on these data, give the least-squares estimates of the slope and the intercept of the regression line.

b) Perform a suitable test to investigate the claim that the slope parameter is $\beta_1 = 0$. Take significance level $\alpha = 0.05$. *(See the first page of this exam for detailed instructions about testing.)*

c) Explain in general: What are normal QQ plots, i.e., what do they display and what kind of conclusions can be drawn from them?

d) What does the QQ plot in Figure 1 tell us?

Formulas and Tables for Exam Statistical Methods

Probability

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

$$\textit{Law of Total Probability}: \quad P(B) = \sum_{i=1}^{m} P(B \cap A_i) = \sum_{i=1}^{m} P(B|A_i)P(A_i).$$

$$\textit{Bayes' Theorem}: \quad P(A_r|B) = \frac{P(A_r \cap B)}{P(B)} = \frac{P(B|A_r)P(A_r)}{\sum_{i=1}^{m} P(B|A_i)P(A_i)}.$$

One sample testing

(i) If σ is known and certain requirements are met, the test statistic $Z = \dfrac{\overline{X}_n - \mu_0}{\sigma/\sqrt{n}}$ has a standard normal distribution under the null hypothesis. The margin of error for the $(1-\alpha)$ % confidence interval for the true population mean μ is $E = z_{\alpha/2}\sigma/\sqrt{n}$.

(ii) If σ is unknown and certain requirements are met, the test statistic $T = \dfrac{\overline{X}_n - \mu_0}{S_n/\sqrt{n}}$ has a t-distribution with $n-1$ degrees of freedom under the null hypothesis. The margin of error for the $(1-\alpha)$ % confidence interval for the true population mean μ is $E = t_{n-1,\alpha/2}s_n/\sqrt{n}$.

(iii) If certain requirements are met, the test statistic $Z = \dfrac{\hat{P}_n - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$ has a standard normal distribution under the null hypothesis. The margin of error for the $(1-\alpha)$ % confidence interval for the true population proportion p is $E = z_{\alpha/2}\sqrt{\hat{p}_n(1-\hat{p}_n)}/\sqrt{n}$.

Two independent samples

(i) If σ_1 and σ_2 are unknown, $\sigma_1 \neq \sigma_2$, and certain requirements are met, the test statistic

$$T_2 = \frac{(\overline{X}_1 - \overline{X}_2) - d_0}{\sqrt{S_1^2/n_1 + S_2^2/n_2}},$$

where d_0 is the hypothetical difference of μ_1 and μ_2 under H_0, has a t-distribution with approximately $\tilde{n}$ degrees of freedom under the null hypothesis.
We use the conservative estimate $\tilde{n} = \min\{n_1 - 1, n_2 - 1\}$.
The margin of error for the $(1-\alpha)$ % confidence interval for the true difference in population means, $\mu_1 - \mu_2$, is $E = t_{\tilde{n},\alpha/2}\sqrt{s_1^2/n_1 + s_2^2/n_2}$.
(ii) If σ_1 and σ_2 are unknown, $\sigma_1 = \sigma_2$, and certain requirements are met, the test statistic

$$T_2^{\text{eq}} = \frac{(\overline{X}_1 - \overline{X}_2) - d_0}{\sqrt{S_p^2/n_1 + S_p^2/n_2}}$$

where d_0 is the hypothetical difference of μ_1 and μ_2 under H_0, has a t-distribution with $n_1 + n_2 - 2$ degrees of freedom under the null hypothesis.
Here S_p^2 is the pooled sample variance given by

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}.$$

The margin of error for the $(1-\alpha)$ % confidence interval for the true difference in population means, $\mu_1 - \mu_2$, is $E = t_{n_1+n_2-2,\alpha/2}\sqrt{s_p^2/n_1 + s_p^2/n_2}$.
(iii) If certain requirements are met, the test statistic

$$Z_p = \frac{(\hat{P}_1 - \hat{P}_2) - d_0}{\sqrt{\overline{P}(1-\overline{P})/n_1 + \overline{P}(1-\overline{P})/n_2}},$$

where d_0 is the hypothetical difference of p_1 and p_2 under H_0, approximately has a standard normal distribution under the null hypothesis.
Here $\overline{P} = (X_1 + X_2)/(n_1 + n_2)$ is the pooled sample proportion.
The margin of error for the $(1-\alpha)$ % confidence interval for the true difference in population proportions, $p_1 - p_2$, is $E = z_{\alpha/2}\sqrt{\hat{p}_1(1-\hat{p}_1)/n_1 + \hat{p}_2(1-\hat{p}_2)/n_2}$.

Two dependent samples

If certain requirements are met, the test statistic

$$T_d = \frac{\overline{D} - d_0}{S_d/\sqrt{n}},$$

where d_0 is the hypothetical difference of μ_1 and μ_2 under H_0, has a t-distribution with $n-1$ degrees of freedom under the null hypothesis.
The margin of error for the $(1-\alpha)$ % confidence interval for the true difference in population means, $\mu_1 - \mu_2$, is $E = t_{n-1,\alpha/2}\sqrt{s_d^2/n}$.

Correlation

If certain requirements are me, the test statistic $T_\rho = \dfrac{R}{\sqrt{(1-R^2)/(n-2)}}$ has a t-distribution with $n-2$ degrees of freedom under the null hypothesis that the population correlation is equal to zero.

Linear regression

Let b_0 and b_1 respectively be the corresponding estimators of the unknown intercept and slope of a linear regression model with one explanatory variable. Then b_0 and b_1 are given by

$$b_1 = r\frac{s_y}{s_x} \quad \text{and} \quad b_0 = \bar{y} - b_1\bar{x}.$$

If certain requirements are met, then the test statistic $T_\beta = \dfrac{b_1}{s_{b_1}}$ has a t-distribution with $n-2$ degrees of freedom under the null hypothesis that the population slope is equal to zero.

Goodness-of-Fit or Test of Independence/Homogeneity

If there are k different categories, or the contingency table has r rows and c columns, and certain requirements are met, then the test statistic $X^2 = \sum_i \dfrac{(O_i - E_i)^2}{E_i}$ approximately has a chi-square distribution with $k-1$ or $(r-1)(c-1)$ degrees of freedom, respectively, under the null hypothesis.

NEGATIVE *z* Scores

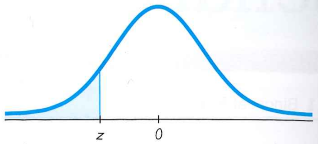


Table 2 Standard Normal (*z*) Distribution: Cumulative Area from the LEFT

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

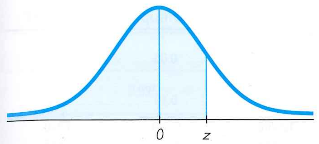


Table 2 *(continued)* Cumulative Area from the LEFT

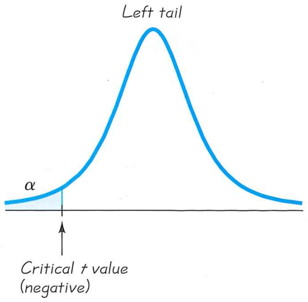
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4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

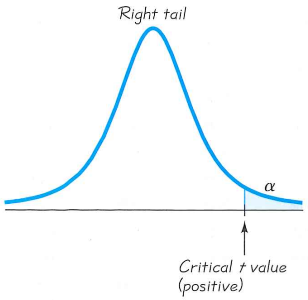
NOTE: For values of *z* above 3.49, use 0.9999 for the area.
*Use these common values that result from interpolation:

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





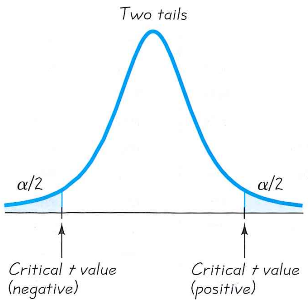


Table 3 *t* Distribution: Critical *t* Values

	Area in One Tail				
	0.005	0.01	0.025	0.05	0.10
	Area in Two Tails				
Degrees of Freedom	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

Appendix Tables

Table 4 Chi-Square (χ^2) Distribution

	Area to the Right of the Critical Value									
Degrees of Freedom	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