BRASE • BRASE

TENTH EDITION

Understandable Statistics

Concepts and Methods

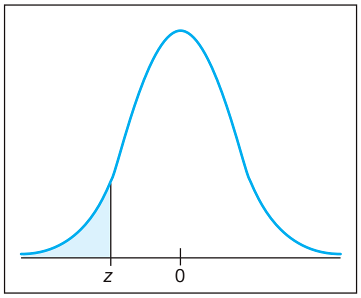


Table entry for *z* is the area to the left of *z*.

Areas of a Standard Normal Distribution

(a) Table of Areas to the Left of *z*

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−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

For values of *z* less than −3.49, use 0.000 to approximate the area.

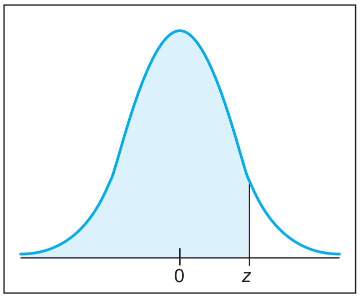


Table entry for z is the area to the left of z.

Areas of a Standard Normal Distribution *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

For z values greater than 3.49, use 1.000 to approximate the area.

Areas of a Standard Normal Distribution *continued*

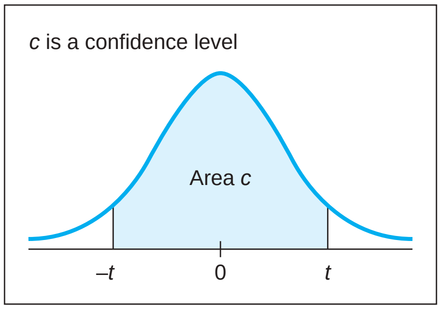
(b) Confidence Interval Critical Values z_c

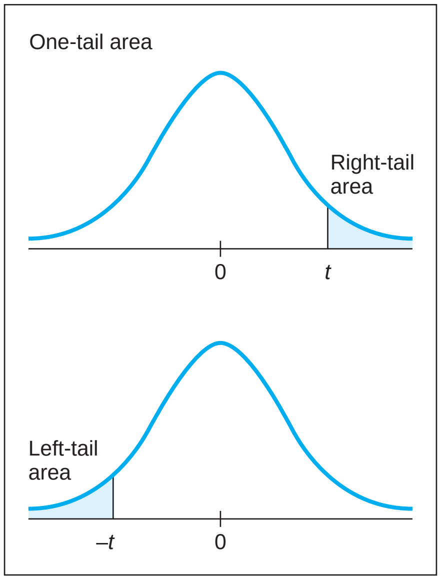
Level of Confidence c	Critical Value z_c
0.70, or 70%	1.04
0.75, or 75%	1.15
0.80, or 80%	1.28
0.85, or 85%	1.44
0.90, or 90%	1.645
0.95, or 95%	1.96
0.98, or 98%	2.33
0.99, or 99%	2.58

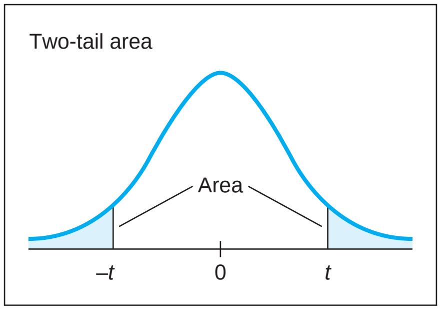
Areas of a Standard Normal Distribution *continued*

(c) Hypothesis Testing, Critical Values z_0

Level of Significance	$\alpha = 0.05$	$\alpha = 0.01$
Critical value z_0 for a left-tailed test	-1.645	-2.33
Critical value z_0 for a right-tailed test	1.645	2.33
Critical values $\pm z_0$ for a two-tailed test	± 1.96	± 2.58



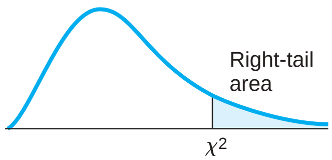




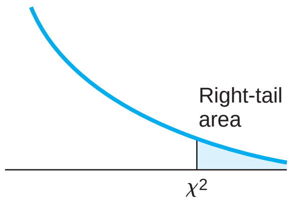
Critical Values for Student's *t* Distribution

one-tail area	0.250	0.125	0.100	0.075	0.050	0.025	0.010	0.005	0.0005
two-tail area	0.500	0.250	0.200	0.150	0.100	0.050	0.020	0.010	0.0010
d.f. \ *c*	0.500	0.750	0.800	0.850	0.900	0.950	0.980	0.990	0.999
1	1.000	2.414	3.078	4.165	6.314	12.706	31.821	63.657	636.619
2	0.816	1.604	1.886	2.282	2.920	4.303	6.965	9.925	31.599
3	0.765	1.423	1.638	1.924	2.353	3.182	4.541	5.841	12.924
4	0.741	1.344	1.533	1.778	2.132	2.776	3.747	4.604	8.610
5	0.727	1.301	1.476	1.699	2.015	2.571	3.365	4.032	6.869
6	0.718	1.273	1.440	1.650	1.943	2.447	3.143	3.707	5.959
7	0.711	1.254	1.415	1.617	1.895	2.365	2.998	3.499	5.408
8	0.706	1.240	1.397	1.592	1.860	2.306	2.896	3.355	5.041
9	0.703	1.230	1.383	1.574	1.833	2.262	2.821	3.250	4.781
10	0.700	1.221	1.372	1.559	1.812	2.228	2.764	3.169	4.587
11	0.697	1.214	1.363	1.548	1.796	2.201	2.718	3.106	4.437
12	0.695	1.209	1.356	1.538	1.782	2.179	2.681	3.055	4.318
13	0.694	1.204	1.350	1.530	1.771	2.160	2.650	3.012	4.221
14	0.692	1.200	1.345	1.523	1.761	2.145	2.624	2.977	4.140
15	0.691	1.197	1.341	1.517	1.753	2.131	2.602	2.947	4.073
16	0.690	1.194	1.337	1.512	1.746	2.120	2.583	2.921	4.015
17	0.689	1.191	1.333	1.508	1.740	2.110	2.567	2.898	3.965
18	0.688	1.189	1.330	1.504	1.734	2.101	2.552	2.878	3.922
19	0.688	1.187	1.328	1.500	1.729	2.093	2.539	2.861	3.883
20	0.687	1.185	1.325	1.497	1.725	2.086	2.528	2.845	3.850
21	0.686	1.183	1.323	1.494	1.721	2.080	2.518	2.831	3.819
22	0.686	1.182	1.321	1.492	1.717	2.074	2.508	2.819	3.792
23	0.685	1.180	1.319	1.489	1.714	2.069	2.500	2.807	3.768
24	0.685	1.179	1.318	1.487	1.711	2.064	2.492	2.797	3.745
25	0.684	1.198	1.316	1.485	1.708	2.060	2.485	2.787	3.725
26	0.684	1.177	1.315	1.483	1.706	2.056	2.479	2.779	3.707
27	0.684	1.176	1.314	1.482	1.703	2.052	2.473	2.771	3.690
28	0.683	1.175	1.313	1.480	1.701	2.048	2.467	2.763	3.674
29	0.683	1.174	1.311	1.479	1.699	2.045	2.462	2.756	3.659
30	0.683	1.173	1.310	1.477	1.697	2.042	2.457	2.750	3.646
35	0.682	1.170	1.306	1.472	1.690	2.030	2.438	2.724	3.591
40	0.681	1.167	1.303	1.468	1.684	2.021	2.423	2.704	3.551
45	0.680	1.165	1.301	1.465	1.679	2.014	2.412	2.690	3.520
50	0.679	1.164	1.299	1.462	1.676	2.009	2.403	2.678	3.496
60	0.679	1.162	1.296	1.458	1.671	2.000	2.390	2.660	3.460
70	0.678	1.160	1.294	1.456	1.667	1.994	2.381	2.648	3.435
80	0.678	1.159	1.292	1.453	1.664	1.990	2.374	2.639	3.416
100	0.677	1.157	1.290	1.451	1.660	1.984	2.364	2.626	3.390
500	0.675	1.152	1.283	1.442	1.648	1.965	2.334	2.586	3.310
1000	0.675	1.151	1.282	1.441	1.646	1.962	2.330	2.581	3.300
∞	0.674	1.150	1.282	1.440	1.645	1.960	2.326	2.576	3.291

For degrees of freedom *d.f.* not in the table, use the closest *d.f.* that is *smaller*.



For *d.f.* ≥ 3



For *d.f.* = 1 or 2

The χ^2 Distribution

	Right-tail Area									
d.f.	.995	.990	.975	.950	.900	.100	.050	.025	.010	.005
1	0.0^4393	0.0^3157	0.0^3982	0.0^2393	0.0158	2.71	3.84	5.02	6.63	7.88
2	0.0100	0.0201	0.0506	0.103	0.211	4.61	5.99	7.38	9.21	10.60
3	0.072	0.115	0.216	0.352	0.584	6.25	7.81	9.35	11.34	12.84
4	0.207	0.297	0.484	0.711	1.064	7.78	9.49	11.14	13.28	14.86
5	0.412	0.554	0.831	1.145	1.61	9.24	11.07	12.83	15.09	16.75
6	0.676	0.872	1.24	1.64	2.20	10.64	12.59	14.45	16.81	18.55
7	0.989	1.24	1.69	2.17	2.83	12.02	14.07	16.01	18.48	20.28
8	1.34	1.65	2.18	2.73	3.49	13.36	15.51	17.53	20.09	21.96
9	1.73	2.09	2.70	3.33	4.17	14.68	16.92	19.02	21.67	23.59
10	2.16	2.56	3.25	3.94	4.87	15.99	18.31	20.48	23.21	25.19
11	2.60	3.05	3.82	4.57	5.58	17.28	19.68	21.92	24.72	26.76
12	3.07	3.57	4.40	5.23	6.30	18.55	21.03	23.34	26.22	28.30
13	3.57	4.11	5.01	5.89	7.04	19.81	22.36	24.74	27.69	29.82
14	4.07	4.66	5.63	6.57	7.79	21.06	23.68	26.12	29.14	31.32
15	4.60	5.23	6.26	7.26	8.55	22.31	25.00	27.49	30.58	32.80
16	5.14	5.81	6.91	7.96	9.31	23.54	26.30	28.85	32.00	34.27
17	5.70	6.41	7.56	8.67	10.09	24.77	27.59	30.19	33.41	35.72
18	6.26	7.01	8.23	9.39	10.86	25.99	28.87	31.53	34.81	37.16
19	6.84	7.63	8.91	10.12	11.65	27.20	30.14	32.85	36.19	38.58
20	7.43	8.26	8.59	10.85	12.44	28.41	31.41	34.17	37.57	40.00
21	8.03	8.90	10.28	11.59	13.24	29.62	32.67	35.48	38.93	41.40
22	8.64	9.54	10.98	12.34	14.04	30.81	33.92	36.78	40.29	42.80
23	9.26	10.20	11.69	13.09	14.85	32.01	35.17	38.08	41.64	44.18
24	9.89	10.86	12.40	13.85	15.66	33.20	36.42	39.36	42.98	45.56
25	10.52	11.52	13.12	14.61	16.47	34.38	37.65	40.65	44.31	46.93
26	11.16	12.20	13.84	15.38	17.29	35.56	38.89	41.92	45.64	48.29
27	11.81	12.88	14.57	16.15	18.11	36.74	40.11	43.19	46.96	49.64
28	12.46	13.56	15.31	16.93	18.94	37.92	41.34	44.46	48.28	50.99
29	13.21	14.26	16.05	17.71	19.77	39.09	42.56	45.72	49.59	52.34
30	13.79	14.95	16.79	18.49	20.60	40.26	43.77	46.98	50.89	53.67
40	20.71	22.16	24.43	26.51	29.05	51.80	55.76	59.34	63.69	66.77
50	27.99	29.71	32.36	34.76	37.69	63.17	67.50	71.42	76.15	79.49
60	35.53	37.48	40.48	43.19	46.46	74.40	79.08	83.30	88.38	91.95
70	43.28	45.44	48.76	51.74	55.33	85.53	90.53	95.02	100.4	104.2
80	51.17	53.54	57.15	60.39	64.28	96.58	101.9	106.6	112.3	116.3
90	59.20	61.75	65.65	69.13	73.29	107.6	113.1	118.1	124.1	128.3
100	67.33	70.06	74.22	77.93	82.36	118.5	124.3	129.6	135.8	140.2

Source: Biometricka, June 1964, The χ^2 Distribution, H. L. Herter (Table 7). Used by permission of Oxford University Press.

INSTUCTOR'S ANNOTATED EDITION

TENTH EDITION

Understandable Statistics

Concepts and Methods

Charles Henry Brase
Regis University

Corrinne Pellillo Brase
Arapahoe Community College

BROOKS/COLE
CENGAGE Learning™

Australia • Brazil • Japan • Korea • Mexico • Singapore • Spain • United Kingdom • United States

This book is dedicated to the memory of
a great teacher, mathematician, and friend

Burton W. Jones
Professor Emeritus, University of Colorado

BROOKS/COLE
CENGAGE Learning™

Understandable Statistics: Concepts and Methods, Tenth Edition

Charles Henry Brase, Corrinne Pellillo Brase

Editor in Chief: Michelle Julet

Publisher: Richard Stratton

Senior Sponsoring Editor: Molly Taylor

Senior Editorial Assistant: Shaylin Walsh

Media Editor: Andrew Coppola

Marketing Manager: Ashley Pickering

Marketing Communications Manager: Mary Anne Payumo

Content Project Manager: Jill Clark

Art Director: Linda Helcher

Senior Manufacturing Buyer: Diane Gibbons

Senior Rights Acquisition Specialist, Text: Katie Huha

Rights Acquisition Specialist, Images: Mandy Groszko

Text Permissions Editor: Sue Howard

Production Service: Elm Street Publishing Services

Cover Designer: RHDG

Cover Image: © Anup Shah

Compositor: Integra Software Services, Ltd. Pvt.

Library of Congress Control Number: 2009942998

Student Edition:
ISBN-13: 978-0-8400-4838-7
ISBN-10: 0-8400-4838-6

Annotated Instructor's Edition:
ISBN-13: 978-0-8400-5456-2
ISBN-10: 0-8400-5456-4

Brooks/Cole
20 Channel Center Street
Boston, MA 02210
USA

Cengage Learning is a leading provider of customized learning solutions with office locations around the globe, including Singapore, the United Kingdom, Australia, Mexico, Brazil and Japan. Locate your local office at international.cengage.com/region

Cengage Learning products are represented in Canada by Nelson Education, Ltd.

For your course and learning solutions, visit **www.cengage.com.**
Purchase any of our products at your local college store or at our preferred online store **www.cengagebrain.com**.

Printed in the United States of America
1 2 3 4 5 6 7 14 13 12 11 10

Contents

4 Elementary Probability Theory 130

5 The Binomial Probability Distribution and Related Topics 180

6 Normal Curves and Sampling Distributions 248

Critical Thinking

Students need to develop critical thinking skills in order to understand and evaluate the limitations of statistical methods. *Understandable Statistics: Concepts and Methods* makes students aware of method appropriateness, assumptions, biases, and justifiable conclusions.

CRITICAL THINKING

Bias and Variability

Whenever we use a sample statistic as an estimate of a population parameter, we need to consider both *bias* and *variability* of the statistic.

A sample statistic is **unbiased** if the mean of its sampling distribution equals the value of the parameter being estimated.

The spread of the sampling distribution indicates the **variability of the statistic.** The spread is affected by the sampling method and the sample size. Statistics from larger random samples have spreads that are smaller.

We see from the central limit theorem that the sample mean $\bar{x}$ is an unbiased estimator of the mean μ when $n \geq 30$. The variability of $\bar{x}$ decreases as the sample size increases.

In Section 6.6, we will see that the sample proportion $\hat{p}$ is an unbiased estimator of the population proportion of successes p in binomial experiments with sufficiently large numbers of trials n. Again, we will see that the variability of $\hat{p}$ decreases with increasing numbers of trials.

The sample variance s^2 is an unbiased estimator for the population variance σ^2.

◄ Critical Thinking

Critical thinking is an important skill for students to develop in order to avoid reaching misleading conclusions. The Critical Thinking feature provides additional clarification on specific concepts as a safeguard against incorrect evaluation of information.

Interpretation ►

Increasingly, calculators and computers are used to generate the numeric results of a statistical process. However, the student still needs to correctly interpret those results in the context of a particular application. The Interpretation feature calls attention to this important step. Interpretation is stressed in examples, guided exercises, and in the problem sets.

The probability is 0.9988 that $\bar{x}$ is between 2350 and 2650.

(c) *Interpretation* At the end of each day, the inspector must decide to accept or reject the accumulated milk that has been held in cold storage awaiting shipment. Suppose the 42 samples taken by the inspector have a mean bacteria count $\bar{x}$ that is *not* between 2350 and 2650. If you were the inspector, what would be your comment on this situation?

SOLUTION: The probability that $\bar{x}$ is between 2350 and 2650 for milk that is not contaminated is very high. If the inspector finds that the average bacteria count for the 42 samples is not between 2350 and 2650, then it is reasonable to conclude that there is something wrong with the milk. If $\bar{x}$ is less than 2350, you might suspect someone added chemicals to the milk to artificially reduce the bacteria count. If $\bar{x}$ is above 2650, you might suspect some other kind of biologic contamination.

1, 22, and 23, we'll apply
it theorem to solve
lving a *sum* of random

11. *Interpretation* A job-performance evaluation form has these categories:

    1 = excellent; 2 = good; 3 = satisfactory; 4 = poor; 5 = unacceptable

    Based on 15 client reviews, one employee had

    median rating of 4; mode rating of 1

    The employee was pleased that most clients had rated her as excellent. The supervisor said improvement was needed because at least half the clients had rated the employee at the poor or unacceptable level. Comment on the different perspectives.

12. *Critical Thinking: Data Transformation* In this problem, we explore the effect on the mean, median, and mode of adding the same number to each data value. Consider the data set 2, 2, 3, 6, 10.
    (a) Compute the mode, median, and mean.
    (b) Add 5 to each of the data values. Compute the mode, median, and mean.
    (c) Compare the results of parts (a) and (b). In general, how do you think the mode, median, and mean are affected when the same constant is added to each data value in a set?

◄ NEW! Critical Thinking and Interpretation Exercises

In every section and chapter problem set, Critical Thinking problems provide students with the opportunity to test their understanding of the application of statistical methods and their interpretation of their results. Interpretation problems ask students to apply statistical results to the particular application.

Statistical Literacy

No language can be spoken without learning the vocabulary, including statistics. *Understandable Statistics: Concepts and Methods* introduces statistical terms with deliberate care.

SECTION 6.1 PROBLEMS

1. *Statistical Literacy* Which, if any, of the curves in Figure 6-10 look(s) like a normal curve? If a curve is not a normal curve, tell why.
2. *Statistical Literacy* Look at the normal curve in Figure 6-11, and find μ, $\mu + \sigma$, and σ.

FIGURE 6-10

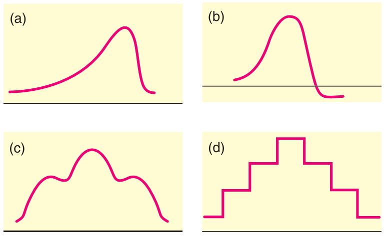
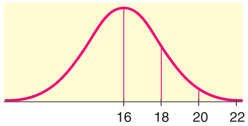


FIGURE 6-11



◄ Statistical Literacy Problems

In every section and chapter problem set, Statistical Literacy problems test student understanding of terminology, statistical methods, and the appropriate conditions for use of the different processes.

Definition Boxes ►

Whenever important terms are introduced in text, yellow definition boxes appear within the discussions. These boxes make it easy to reference or review terms as they are used further.

Box-and-Whisker Plots

Five-number summary

The quartiles together with the low and high data values give us a very useful *five-number summary* of the data and their spread.

Five-number summary

Lowest value, Q_1, median, Q_3, highest value

Box-and-whisker plot

We will use these five numbers to create a graphic sketch of the data called a *box-and-whisker plot*. Box-and-whisker plots provide another useful technique from exploratory data analysis (EDA) for describing data.

IMPORTANT WORDS & SYMBOLS

▲ REVISED! Important Words & Symbols

The Important Words & Symbols within the Chapter Review feature at the end of each chapter summarizes the terms introduced in the Definition Boxes for student review at a glance. Page numbers for first occurrance of term are given for easy reference.

Statistical Literacy

Linking Concepts: Writing Projects ▶

Much of statistical literacy is the ability to communicate concepts effectively. The Linking Concepts: Writing Projects feature at the end of each chapter tests both statistical literacy and critical thinking by asking the student to express their understanding in words.

LINKING CONCEPTS: WRITING PROJECTS

Discuss each of the following topics in class or review the topics on your own. Then write a brief but complete essay in which you summarize the main points. Please include formulas and graphs as appropriate.

1. An average is an attempt to summarize a collection of data into just *one* number. Discuss how the mean, median, and mode all represent averages in this context. Also discuss the differences among these averages. Why is the mean a balance point? Why is the median a midway point? Why is the mode the most common data point? List three areas of daily life in which you think one of the mean, median, or mode would be the best choice to describe an "average."
2. Why do we need to study the variation of a collection of data? Why isn't the average by itself adequate? We have studied three ways to measure variation. The range, the standard deviation, and, to a large extent, a box-and-whisker plot all indicate the variation within a data collection. Discuss similarities and differences among these ways to measure data variation. Why would it seem reasonable to pair the median with a box-and-whisker plot and to pair the mean with the standard deviation? What are the advantages and disadvantages of each method of describing data spread? Comment on statements such as the following: (a) The range is easy to compute, but it doesn't give much information; (b) although the standard deviation is more complicated to compute, it has some significant applications; (c) the box-and-whisker plot is fairly easy to construct, and it gives a lot of information at a glance.

9. *Basic Computation: Testing μ, σ Unknown* A random sample of 25 values is drawn from a mound-shaped and symmetric distribution. The sample mean is 10 and the sample standard deviation is 2. Use a level of significance of 0.05 to conduct a two-tailed test of the claim that the population mean is 9.5.
   (a) *Check Requirements* Is it appropriate to use a Student's t distribution? Explain. How many degrees of freedom do we use?
   (b) What are the hypotheses?
   (c) Compute the sample test statistic t.
   (d) Estimate the P-value for the test.
   (e) Do we reject or fail to reject H_0?
   (f) *Interpret* the results.

◀ NEW! Basic Computation Problems

These problems focus student attention on relevant formulas, requirements, and computational procedures. After practicing these skills, students are more confident as they approach real-world applications.

Expand Your Knowledge Problems ▶

Expand Your Knowledge problems present optional enrichment topics that go beyond the material introduced in a section. Vocabulary and concepts needed to solve the problems are included at point-of-use, expanding students' statistical literacy.

(b) Suppose the EPA has established an average chlorine compound concentration target of no more than 58 mg/l. Comment on whether this wetlands system meets the target standard for chlorine compound concentration.

17. *Expand Your Knowledge: Harmonic Mean* When data consist of rates of change, such as speeds, the *harmonic mean* is an appropriate measure of central tendency. for n data values,

$$\textbf{Harmonic mean} = \frac{n}{\sum \frac{1}{x}}, \quad \text{assuming no data value is 0}$$

Suppose you drive 60 miles per hour for 100 miles, then 75 miles per hour for 100 miles. Use the harmonic mean to find your average speed.

18. *Expand Your Knowledge: Geometric Mean* When data consist of percentages, ratios, growth rates, or other rates of change, the *geometric mean* is a useful measure of central tendency. For n data values,

$$\textbf{Geometric mean} = \sqrt[n]{\text{product of the } n \text{ data values}}, \quad \text{assuming all data values are positive}$$

To find the *average growth factor* over 5 years of an investment in a mutual fund with growth rates of 10% the first year, 12% the second year, 14.8% the third year, 3.8% the fourth year, and 6% the fifth year, take the geometric mean of 1.10, 1.12, 1.148, 1.038, and 1.16. Find the average growth factor of this investment.

Note that for the same data, the relationships among the harmonic, geometric, and arithmetic means are harmonic mean ≤ geometric mean ≤ arithmetic mean (*Source: Oxford Dictionary of Statistics*).

Direction and Purpose

Real knowledge is delivered through direction, not just facts. *Understandable Statistics: Concepts and Methods* ensures the student knows what is being covered and why at every step along the way to statistical literacy.

Chapter Preview ► Questions

Preview Questions at the beginning of each chapter give the student a taste of what types of questions can be answered with an understanding of the knowledge to come.

Normal Curves and Sampling Distributions


Dana White/PhotoEdit

PREVIEW QUESTIONS

What are some characteristics of a normal distribution? What does the empirical rule tell you about data spread around the mean? How can this information be used in quality control? (Section 6.1)

Can you compare apples and oranges, or maybe elephants and butterflies? In most cases, the answer is no—unless you first standardize your measurements. What are a standard normal distribution and a standard z score? (Section 6.2)

How do you convert any normal distribution to a standard normal distribution? How do you find probabilities of "standardized events"? (Section 6.3)

As humans, our experiences are finite and limited. Consequently, most of the important decisions in our lives are based on sample (incomplete) information. What is a probability sampling distribution? How will sampling distributions help us make good decisions based on incomplete information? (Section 6.4)

There is an old saying: All roads lead to Rome. In statistics, we could recast this saying: All probability distributions average out to be normal distributions (as the sample size increases). How can we take advantage of this in our study of sampling distributions? (Section 6.5)

The binomial and normal distributions are two of the most important probability distributions in statistics. Under certain limiting conditions, the binomial can be thought to evolve (or envelope)

s concept in the real world? (Section 6.6)

st cases, we will not be successful all the time,

FOCUS PROBLEM

Benford's Law: The Importance of Being Number 1

Benford's Law states that in a wide variety of circumstances, numbers have "1" as their first nonzero digit disproportionately often. Benford's Law applies to such diverse topics as the drainage areas of rivers; properties of chemicals; populations of towns; figures in newspapers, magazines, and government reports; and the half-lives of radioactive atoms!

Specifically, such diverse measurements begin with "1" about 30% of the time, with "2" about 18% of time, and with "3" about 12.5% of the time. Larger digits occur less often. For example, less than 5% of the numbers in circumstances such as these begin with the digit 9. This is in dramatic contrast to a random sampling situation, in which each of the digits 1 through 9 has an equal chance of appearing.

The first nonzero digits of numbers taken from large bodies of numerical records such as tax returns, population studies, government records, and so forth, show the probabilities of occurrence as displayed in the table on the next page.

Corbis

8. *Focus Problem: Benford's Law* Again, suppose you are the auditor for a very large corporation. The revenue file contains millions of numbers in a large computer data bank (see Problem 7). You draw a random sample of $n = 228$ numbers from this file and $r = 92$ have a first nonzero digit of 1. Let p represent the population proportion of all numbers in the computer file that have a leading digit of 1.
   i. Test the claim that p is more than 0.301. Use $\alpha = 0.01$.
   ii. If p is in fact larger than 0.301, it would seem there are too many numbers in the file with leading 1's. Could this indicate that the books have been "cooked" by artificially lowering numbers in the file? Comment from the point of view of the Internal Revenue Service. Comment from the perspective of the Federal Bureau of Investigation as it looks for "profit skimming" by unscrupulous employees.
   iii. Comment on the following statement: "If we reject the null hypothesis at level of significance α, we have not *proved* H_0 to be false. We can say that the probability is α that we made a mistake in rejecting H_0." Based on the outcome of the test, would you recommend further investigation before accusing the company of fraud?

▲ Chapter Focus Problems

The Preview Questions in each chapter are followed by Focus Problems, which serve as more specific examples of what questions the student will soon be able to answer. The Focus Problems are set within appropriate applications and are incorporated into the end-of-section exercises, giving students the opportunity to test their understanding.

Direction and Purpose

Focus Points ►

Each section opens with bulleted Focus Points describing the primary learning objectives of the section.

SECTION 3.1

Measures of Central Tendency: Mode, Median, and Mean

FOCUS POINTS

- Compute mean, median, and mode from raw data.
- Interpret what mean, median, and mode tell you.
- Explain how mean, median, and mode can be affected by extreme data values.
- What is a trimmed mean? How do you compute it?
- Compute a weighted average.

The average price of an ounce of gold is $1200. The Zippy car averages 39 miles per gallon on the highway. A survey showed the average shoe size for women is size 9.

Average

In each of the preceding statements, *one* number is used to describe the entire sample or population. Such a number is called an *average*. There are many ways to compute averages, but we will study only three of the major ones.

The easiest average to compute is the *mode*.

Mode

The **mode** of a data set is the value that occurs most frequently.

EXAMPLE 1 MODE

Count the letters in each word of this sentence and give the mode. The numbers of letters in the words of the sentence are

5 3 7 2 4 4 2 4 8 3 4 3 4

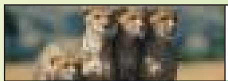
Scanning the data, we see that 4 is the mode because more words have 4 letters than any other number. For larger data sets, it is useful to order—or sort—the data before scanning them for the mode.

LOOKING FORWARD

In our future work with inferential statistics, we will use the mean $\bar{x}$ from a random sample to estimate the population parameter μ (Chapter 7) or to make decisions regarding the value of μ (Chapter 8).

▲ NEW! Looking Forward

This feature shows students where the presented material will be used later. It helps motivate students to pay a little extra attention to key topics.



Chapter Review

SUMMARY

Organizing and presenting data are the main purposes of the branch of statistics called descriptive statistics. Graphs provide an important way to show how the data are distributed.

- Frequency tables show how the data are distributed within set classes. The classes are chosen so that they cover all data values and so that each data value falls within only one class. The number of classes and the class width determine the class limits and class boundaries. The number of data values falling within a class is the class frequency.
- A histogram is a graphical display of the information in a frequency table. Classes are shown on the horizontal axis, with corresponding frequencies on the vertical axis. Relative-frequency histograms show relative frequencies on the vertical axis. Ogives show cumulative frequencies on the vertical axis. Dotplots are like histograms, except that the classes are individual data values.
- Bar graphs, Pareto charts, and pie charts are useful to show how quantitative or qualitative data are distributed over chosen categories.
- Time-series graphs show how data change over set intervals of time.
- Stem-and-leaf displays are effective means of ordering data and showing important features of the distribution.

Graphs aren't just pretty pictures. They help reveal important properties of the data distribution, including the shape and whether or not there are any outliers.

◄ Chapter Summaries

The Summary within each Chapter Review feature now also appears in bulleted form, so students can see what they need to know at a glance.

Real-World Skills

Statistics is not done in a vacuum. *Understandable Statistics: Concepts and Methods* gives students valuable skills for the real world with technology instruction, genuine applications, actual data, and group projects.

REVISED! Tech Notes ▶

Tech Notes appearing throughout the text give students helpful hints on using TI-4 Plus and TI-*n*spire (with 84 Plus keypad) and TI-83 calculators, Microsoft Excel 2007, and Minitab to solve a problem. They include display screens to help students visualize and better understand the solution.

TECH NOTES *Stem-and-leaf display*

TI-84Plus/TI-83Plus/TI-*n*spire Does not support stem-and-leaf displays. You can sort the data by using keys **Stat ➤ Edit ➤ 2:SortA.**

Excel 2007 Enter your data and select the data you want to sort. On the **Home** ribbon, click the **Sort and Filter** button in the **Editing** group of the ribbon and select the desired sorting option.

Minitab Use the menu selections **Graph ➤ Stem-and-Leaf** and fill in the dialogue box.

Minitab Stem-and-Leaf Display (for Data in Guided Exercise 4)

```
Stem-and-Leaf of Scores     N=35
Leaf Unit=1.0

   1        8    3
   5        9    2789
  14       10    123455669
 (11)      11    01222267789
  10       12    045568
   4       13    125
   1       14    3
```

The values shown in the left column represent depth. Numbers above the value in parentheses show the cumulative number of values from the top to the stem of the middle value. Numbers below the value in parentheses show the cumulative number of values from the bottom to the stem of the middle value. The number in parentheses shows how many values are on the same line as the middle value.

Using Technology

Binomial Distributions

Although tables of binomial probabilities can be found in most libraries, such tables are often inadequate. Either the value of p (the probability of success on a trial) you are looking for is not in the table, or the value of n (the number of trials) you are looking for is too large for the table. In Chapter 6, we will study the normal approximation to the binomial. This approximation is a great help in many practical applications. Even so, we sometimes use the formula for the binomial probability distribution on a computer or graphing calculator to compute the probability we want.

Applications

The following percentages were obtained over many years of observation by the U.S. Weather Bureau. All data listed are for the month of December.

Location	Long-Term Mean % of Clear Days in Dec.
Juneau, Alaska	18%
Seattle, Washington	24%
Hilo, Hawaii	36%
Honolulu, Hawaii	60%
Las Vegas, Nevada	75%
Phoenix, Arizona	77%

Adapted from *Local Climatological Data,* U.S. Weather Bureau publication, "Normals, Means, and Extremes" Table.

In the locations listed, the month of December is a relatively stable month with respect to weather. Since weather patterns from one day to the next are more or less the same, it is reasonable to use a binomial probability model.

1. Let r be the number of clear days in December. Since December has 31 days, $0 \le r \le 31$. Using appropriate computer software or calculators available to you, find the probability $P(r)$ for each of the listed locations when $r = 0, 1, 2, \ldots, 31$.
2. For each location, what is the expected value of the probability distribution? What is the standard deviation?

You may find that using cumulative probabilities and appropriate subtraction of probabilities, rather than addition of probabilities, will make finding the solutions to Applications 3 to 7 easier.

3. Estimate the probability that Juneau will have at most 7 clear days in December.
4. Estimate the probability that Seattle will have from 5 to 10 (including 5 and 10) clear days in December.
5. Estimate the probability that Hilo will have at least 12 clear days in December.
6. Estimate the probability that Phoenix will have 20 or more clear days in December.
7. Estimate the probability that Las Vegas will have from 20 to 25 (including 20 and 25) clear days in December.

Technology Hints

TI-84Plus/TI-83Plus/TI-*n*spire (with TI-84 Plus keypad), Excel 2007, Minitab

The Tech Note in Section 5.2 gives specific instructions for binomial distribution functions on the TI-84Plus/TI-83Plus/TI-*n*spire (with TI-84Plus keypad) calculators, Excel 2007, and Minitab.

SPSS

In SPSS, the function **PDF.BINOM(q,n,p)** gives the probability of q successes out of n trials, where p is the probability of success on a single trial. In the data editor, name a variable r and enter values 0 through n. Name another variable Prob_r. Then use the menu choices **Transform ➤ Compute.** In the dialogue box, use Prob_r for the target variable. In the function group, select **PDF and Noncentral PDF**. In the function box, select **PDF.BINOM(q,n,p)**. Use the variable r for q and appropriate values for n and p. Note that the function **CDF.BINOM(q,n,p)**, from the **CDF and Noncentral CDF** group, gives the cumulative probability of 0 through q successes.

◀ REVISED! Using Technology

Further technology instruction is available at the end of each chapter in the Using Technology section. Problems are presented with real-world data from a variety of disciplines that can be solved by using TI-84 Plus, TI-*n*spire (with 84 Plus keypad) and TI-83 Plus calculators, Microsoft Excel 2007, and Minitab.

Real-World Skills

EXAMPLE 13 CENTRAL LIMIT THEOREM

A certain strain of bacteria occurs in all raw milk. Let x be the bacteria count per milliliter of milk. The health department has found that if the milk is not contaminated, then x has a distribution that is more or less mound-shaped and symmetrical. The mean of the x distribution is $\mu = 2500$, and the standard deviation is $\sigma = 300$. In a large commercial dairy, the health inspector takes 42 random samples of the milk produced each day. At the end of the day, the bacteria count in each of the 42 samples is averaged to obtain the sample mean bacteria count $\bar{x}$.

(a) Assuming the milk is not contaminated, what is the distribution of $\bar{x}$?

SOLUTION: The sample size is $n = 42$. Since this value exceeds 30, the central limit theorem applies, and we know that $\bar{x}$ will be approximately normal, with mean and standard deviation

◄ UPDATED! **Applications**

Real-world applications are used from the beginning to introduce each statistical process. Rather than just crunching numbers, students come to appreciate the value of statistics through relevant examples.

entered about the mean in which almost all the
found.

10. *Vending Machine: Soft Drinks* A vending machine automatically pours soft drinks into cups. The amount of soft drink dispensed into a cup is normally distributed with a mean of 7.6 ounces and standard deviation of 0.4 ounce. Examine Figure 6-3 and answer the following questions.
    (a) Estimate the probability that the machine will overflow an 8-ounce cup.
    (b) Estimate the probability that the machine will not overflow an 8-ounce cup.
    (c) The machine has just been loaded with 850 cups. How many of these do you expect will overflow when served?

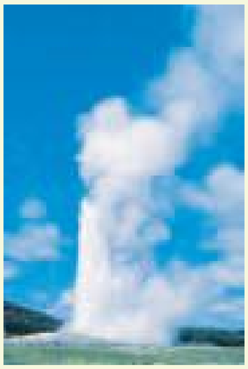
11. *Pain Management: Laser Therapy* "Effect of Helium-Neon Laser Auriculotherapy on Experimental Pain Threshold" is the title of an article in the journal *Physical Therapy* (Vol. 70, No. 1, pp. 24–30). In this article, laser therapy was discussed as a useful alternative to drugs in pain management of chronically ill patients. To measure pain threshold, a machine was used that delivered low-voltage direct current to different parts of the body (wrist, neck, and back). The machine measured current in milliamperes (mA). The pretreatment experimental group in the study had an average threshold of pain (pain was first detectable) at $\mu = 3.15$ mA with

Most exercises in each section ► are applications problems.

DATA HIGHLIGHTS: GROUP PROJECTS

Break into small groups and discuss the following topics. Organize a brief outline in which you summarize the main points of your group discussion.

1. *The Story of Old Faithful* is a short book written by George Marler and published by the Yellowstone Association. Chapter 7 of this interesting book talks about the effect of the 1959 earthquake on eruption intervals for Old Faithful Geyser. Dr. John Rinehart (a senior research scientist with the National Oceanic and Atmospheric Administration) has done extensive studies of the eruption intervals before and after the 1959 earthquake. Examine Figure 3-11. Notice the general shape. Is the graph more or less symmetrical? Does it have a single mode frequency? The mean interval between eruptions has remained steady at about 65 minutes for the past 100 years. Therefore, the 1959 earthquake did not significantly change the mean, but it did change the distribution of eruption intervals. Examine Figure 3-12. Would you say there are really two frequency modes, one shorter and the other longer? Explain. The overall mean is about the same for both graphs, but one graph has a much larger standard deviation (for eruption intervals) than the other. Do no calculations, just look at both graphs, and then explain which graph has the smaller and which has the larger standard deviation. Which distribution will have the larger coefficient of variation? In everyday terms, what would this mean if you were actually at Yellowstone waiting to see the next eruption of Old Faithful? Explain your answer.



Old Faithful Geyser, Yellowstone National Park

FIGURE 3-11

Typical Behavior of Old Faithful Geyser Before 1959 Quake

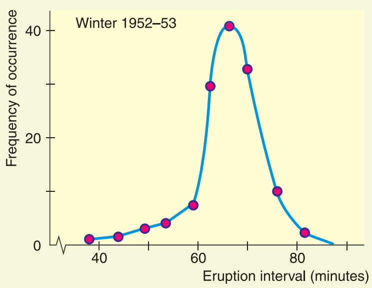
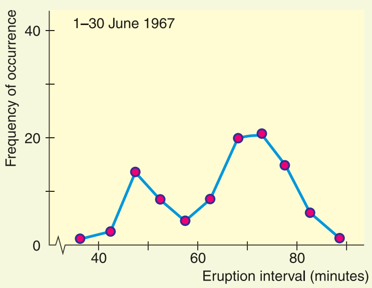


FIGURE 3-12

Typical Behavior of Old Faithful Geyser After 1959 Quake



◄ **Data Highlights: Group Projects**

Using Group Projects, students gain experience working with others by discussing a topic, analyzing data, and collaborating to formulate their response to the questions posed in the exercise.

Making the Jump

Get to the "Aha!" moment faster. *Understandable Statistics: Concepts and Methods* provides the push students need to get there through guidance and example.

PROCEDURE

How to test μ when σ is unknown

Requirements

Let x be a random variable appropriate to your application. Obtain a simple random sample (of size n) of x values from which you compute the sample mean $\bar{x}$ and the sample standard deviation s. If you can assume that x has a normal distribution or simply a mound-shaped and symmetric distribution, then any sample size n will work. If you cannot assume this, use a sample size $n \geq 30$.

Procedure

1. In the context of the application, state the *null and alternate hypotheses* and set the *level of significance* α.
2. Use $\bar{x}$, s, and n from the sample, with μ from H_0, to compute the *sample test statistic*.

d.f. for testing μ when σ unknown

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} \quad \text{with degrees of freedom } d.f. = n - 1$$

3. Use the Student's t distribution and the type of test, one-tailed or two-tailed, to find (or estimate) the *P-value* corresponding to the test statistic.
4. *Conclude* the test. If P-value $\leq \alpha$, then reject H_0. If P-value $> \alpha$, then do not reject H_0.
5. *Interpret your conclusion* in the context of the application.

◄ REVISED! **Procedures and Requirements**

Procedure display boxes summarize simple step-by-step strategies for carrying out statistical procedures and methods as they are introduced. Requirements for using the procedures are also stated. Students can refer back to these boxes as they practice using the procedures.

Guided Exercises ►

Students gain experience with new procedures and methods through Guided Exercises. Beside each problem in a Guided Exercise, a completely worked-out solution appears for immediate reinforcement.

GUIDED EXERCISE 11 Probability regarding $\bar{x}$

In mountain country, major highways sometimes use tunnels instead of long, winding roads over high passes. However, too many vehicles in a tunnel at the same time can cause a hazardous situation. Traffic engineers are studying a long tunnel in Colorado. If x represents the time for a vehicle to go through the tunnel, it is known that the x distribution has mean $\mu = 12.1$ minutes and standard deviation $\sigma = 3.8$ minutes under ordinary traffic conditions. From a histogram of x values, it was found that the x distribution is mound-shaped with some symmetry about the mean.

Engineers have calculated that, *on average,* vehicles should spend from 11 to 13 minutes in the tunnel. If the time is less than 11 minutes, traffic is moving too fast for safe travel in the tunnel. If the time is more than 13 minutes, there is a problem of bad air quality (too much carbon monoxide and other pollutants).

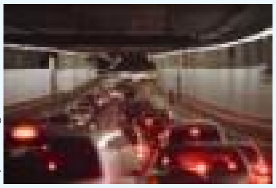
Under ordinary conditions, there are about 50 vehicles in the tunnel at one time. What is the probability that the mean time for 50 vehicles in the tunnel will be from 11 to 13 minutes? We will answer this question in steps.

(a) Let $\bar{x}$ represent the sample mean based on samples of size 50. Describe the $\bar{x}$ distribution.

→ From the central limit theorem, we expect the $\bar{x}$ distribution to be approximately normal, with mean and standard deviation

$$\mu_{\bar{x}} = \mu = 12.1 \qquad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{3.8}{\sqrt{50}} \approx 0.54$$

(b) Find $P(11 < \bar{x} < 13)$.



Jupiter Images

→ We convert the interval

$11 < \bar{x} < 13$

to a standard z interval and use the standard normal probability table to find our answer. Since

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \approx \frac{\bar{x} - 12.1}{0.54}$$

$\bar{x} = 11$ converts to $z \approx \frac{11 - 12.1}{0.54} = -2.04$

and $\bar{x} = 13$ converts to $z \approx \frac{13 - 12.1}{0.54} = 1.67$

Therefore,

$$P(11 < \bar{x} < 13) = P(-2.04 < z < 1.67) = 0.9525 - 0.0207 = 0.9318$$

(c) *Interpret* your answer to part (b).

→ It seems that about 93% of the time, there should be no safety hazard for average traffic flow.

Preface

Welcome to the exciting world of statistics! We have written this text to make statistics accessible to everyone, including those with a limited mathematics background. Statistics affects all aspects of our lives. Whether we are testing new medical devices or determining what will entertain us, applications of statistics are so numerous that, in a sense, we are limited only by our own imagination in discovering new uses for statistics.

Overview

The tenth edition of *Understandable Statistics: Concepts and Methods* continues to emphasize concepts of statistics. Statistical methods are carefully presented with a focus on understanding both the *suitability of the method* and the *meaning of the result*. Statistical methods and measurements are developed in the context of applications.

Critical thinking and interpretation are essential in understanding and evaluting information. Statistical literacy is fundamental for applying and comprehending statistical results. In this edition we have expanded and highlighted the treatment of statistical literacy, critical thinking, and interpretation.

We have retained and expanded features that made the first nine editions of the text very readable. Definition boxes highlight important terms. Procedure displays summarize steps for analyzing data. Examples, exercises, and problems touch on applications appropriate to a broad range of interests.

New with the tenth edition is CourseMate, encompassing all interactive online products and services with this text. Online homework powered by a choice of Enhanced WebAssign or Aplia is now available through CengageBrain.com Also available through CourseMate are over 100 data sets (in Microsoft Excel, Minitab, SPSS, and TI-84Plus/TI-83Plus/TI-*n*spire with 84plus keypad ASCII file formats), lecture aids, a glossary, statistical tables, intructional video (also available on DVDs), an Online Multimedia eBook, and interactive tutorials.

Major Changes in the Tenth Edition

With each new edition, the authors reevaluate the scope, appropriateness, and effectiveness of the text's presentation and reflect on extensive user feedback. Revisions have been made throughout the text to clarify explanations of important concepts and to update problems.

Critical Thinking, Interpretation, and Statistical Literacy

The tenth edition of this text continues and expands the emphasis on critical thinking, interpretation, and statistical literacy. Calculators and computers are very good at providing numerical results of statistical processes. However, numbers from a computer or calcultor display are meaningless unless the user knows how to interpret the results and if the statistical process is appropriate. This text helps students determine whether or not a statistical method or process is appropriate. It helps students understand what a statistic measures. It helps students interpret the results of a confidence interval, hypothesis test, or liner regression model.

New Problems Featuring Basic Computation

Calculators and computer software automatically calculate designated statistical measurements. However, students gain an appreciation and understanding of what the measurements mean by studying and using the basic formulas. Basic computation problems focus attention on using formulas with small data sets. Students see not only how the formulas work, but also how the resulting measurements relate to the displayed data set.

There are more than 200 new and revised problems that feature basic computations, interpretation, and statistical literacy. Studets are asked to cheek that the use of specific probability distributions and inferential methods are appropriate.

Normal Distributions and Sampling Distributions in Same Chapter

Chapter 6 of the tenth edition includes both an introduction to normal distributions as well as an introduction to sampling distributions. Putting both topics in the same chapter streamlines the course, and gives an immediate, important application of normal distributions. The chapter also includes the normal approximation to the bionomial distribution.

New Content

The uniform probability distribution and the exponential probability distribution are introduced in Expand Your Knowledge problems in Chapter 6. Polynomial regression (also known as curvilinear regression) is discussed in Expand Your Knowledge problems of Section 9.4 Multiple Regression.

Excel 2007 and TI-*n*spire Calculator (with 84plus keypad)

Excel 2007 instructions are included in in the *Tech Notes* and *Using Technology* features.

The TI-*n*spire calculator with the 84Plus keypad is also addressed. Instructions for using the TI-nspire calculator with the *n*spire keypad are included in the separate Technology Guide for the TI-84Plus graphing calculator.

Looking Forward

A new *Looking Forward* feature briefly points out how current subject matter will be used in later chapters.

Other Changes

Examples in probability (Chapter 4) have been revised to incoporate sample spaces that are small and easy for students to visualize.

Throughout the tenth edition, examples, guided exercies, and problem sets feature interpretation. They also address the suitability of using a specified statistical method or process for analysis.

Chapter 6 Normal Curves and Sampling Distribution of the tenth edition combines material from Chapters 6 and 7 of the ninth edition. Chapter 7 Estimation, Chapter 8 Hypothesis Testing, Chapter 9 Correlation and Regression, Chapter 10 Chi-Square and F Distribution, and Chapter 11 Nonparametric Statistics of the tenth edition correspond respectively to Chapters 8–12 of the ninth edition.

Continuing Content

Introduction of Hypothesis Testing Using *P*-Values

In keeping with the use of computer technology and standard practice in research, hypothesis testing is introduced using *P*-values. The critical region method is still supported, but not given primary emphasis.

Use of Student's *t* Distribution in Confidence Intervals and Testing of Means

If the normal distribution is used in confidence intervals and testing of means, then the *population standard deviation must be known*. If the population standard deviation is not known, then under conditions described in the text, the Student's *t* distribution is used. This is the most commonly used procedure in statistical research. It is also used in statistical software packages such as Microsoft Excel, Minitab, SPSS, and TI-84Plus/TI-83Plus/TI-*n*spire calculators.

Confidence Intervals and Hypothesis Tests of Difference of Means

If the normal distribution is used, then both population standard deviations must be known. When this is not the case, the Student's *t* distribution incorporates an approximation for *t*, with a commonly used conservative choice for the degrees of freedom. Satterthwaite's approximation for the degrees of freedom as used in computer software is also discussed. The pooled standard deviation is presented for appropriate applications ($\sigma_1 \approx \sigma_2$).

Features in the Tenth Edition

Chapter and Section Lead-ins

- *Preview Questions* at the beginning of each chapter are keyed to the sections.
- *Focus Problems* at the beginning of each chapter demonstrate types of questions students can answer once they master the concepts and skills presented in the chapter.
- *Focus Points* at the beginning of each section describe the primary learning objectives of the section.

Carefully Developed Pedagogy

- *Examples* show students how to select and use appropriate procedures.
- *Guided Exercises* within the sections give students an opportunity to work with a new concept. Completely worked-out solutions appear beside each exercise to give immediate reinforcement.
- *Definition boxes* highlight important definitions throughout the text.
- *Procedure displays* summarize key strategies for carrying out statistical procedures and methods. Conditions required for using the procedure are also stated.
- NEW! *Looking Forward* features give a brief preview of how a current topic is used later.
- *Labels* for each example or guided exercise highlight the technique, concept, or process illustrated by the example or guided exercise. In addition, labels for

section and chapter problems describe the field of application and show the wide variety of subjects in which statistics is used.

- *Section and chapter problems* require the student to use all the new concepts mastered in the section or chapter. Problem sets include a variety of real-world applications with data or settings from identifiable sources. Key steps and solutions to odd-numbered problems appear at the end of the book.
- NEW! *Basic Computation problems* ask students to practice using formulas and statistical methods on very small data sets. Such practice helps students understand what a statistic measures.
- *Statistical Literacy problems* ask students to focus on correct terminology and processes of appropriate statistical methods. Such problems occur in every section and chapter problem set.
- NEW! *Interpretation problems* ask students to explain the meaning of the statistical results in the context of the application.
- *Critical Thinking problems* ask students to analyze and comment on various issues that arise in the application of statistical methods and in the interpretation of results. These problems occur in every section and chapter problem set.
- *Expand Your Knowledge problems* present enrichment topics such as negative binomial distribution; conditional probability utilizing binomial, Poisson, and normal distributions; estimation of standard deviation from a range of data values; and more.
- *Cumulative review problem sets* occur after every third chapter and include key topics from previous chapters. Answers to *all* cumulative review problems are given at the end of the book.
- *Data Highlights and Linking Concepts* provide group projects and writing projects.
- *Viewpoints* are brief essays presenting diverse situations in which statistics is used.
- *Design and photos* are appealing and enhance readability.

Technology within the Text

- *Tech Notes* within sections provide brief point-of-use instructions for the TI-84Plus, TI-83Plus and TI-*n*spire (with 84plus keypad) calculators, Microsoft Excel 2007 and Minitab.
- *Using Technology* sections have been revised to show the use of SPSS as well as the TI-84Plus, TI-83Plus and TI-*n*spire (with 84plus keypad) calculators, Microsoft Excel, and Minitab.

Alternate Routes Through the Text

Understandable Statistics: Concepts and Methods, Tenth Edition, is designed to be flexible. It offers the professor a choice of teaching possibilities. In most one-semester courses, it is not practical to cover all the material in depth. However, depending on the emphasis of the course, the professor may choose to cover various topics. For help in topic selection, refer to the Table of Prerequisite Material on page 1.

- *Introducing linear regression early.* For courses requiring an early presentation of linear regression, the descriptive components of linear regression (Sections 9.1 and 9.2) can be presented any time after Chapter 3. However, inference topics involving predictions, the correlation coefficient ρ, and the slope of the least-squares line β require an introduction to confidence intervals (Sections 7.1 and 7.2) and hypothesis testing (Sections 8.1 and 8.2).
- *Probability.* For courses requiring minimal probability, Section 4.1 (What Is Probability?) and the first part of Section 4.2 (Some Probability Rules—Compound Events) will be sufficient.

Acknowledgments

It is our pleasure to acknowledge the prepublication reviewers of this text. All of their insights and comments have been very valuable to us. Reviewers of this text include:

Reza Abbasian, Texas Lutheran University
Paul Ache, Kutztown University
Kathleen Almy, Rock Valley College
Polly Amstutz, University of Nebraska at Kearney
Delores Anderson, Truett-McConnell College
Robert J. Astalos, Feather River College
Lynda L. Ballou, Kansas State University
Mary Benson, Pensacola Junior College
Larry Bernett, Benedictine University
Kiran Bhutani, The Catholic University of America
Kristy E. Bland, Valdosta State University
John Bray, Broward Community College
Bill Burgin, Gaston College
Toni Carroll, Siena Heights University
Pinyuen Chen, Syracuse University
Emmanuel des-Bordes, James A. Rhodes State College
Jennifer M. Dollar, Grand Rapids Community College
Larry E. Dunham, Wor-Wic Community College
Andrew Ellett, Indiana University
Ruby Evans, Keiser University
Mary Fine, Moberly Area Community College
Rebecca Fouguette, Santa Rosa Junior College
Rene Garcia, Miami-Dade Community College
Larry Green, Lake Tahoe Community College
Shari Harris, John Wood Community College
Janice Hector, DeAnza College
Jane Keller, Metropolitan Community College
Raja Khoury, Collin County Community College
Diane Koenig, Rock Valley College
Charles G. Laws, Cleveland State Community College
Michael R. Lloyd, Henderson State University
Beth Long, Pellissippi State Technical and Community College
Lewis Lum, University of Portland
Darcy P. Mays, Virginia Commonwealth University
Charles C. Okeke, College of Southern Nevada, Las Vegas
Peg Pankowski, Community College of Allegheny County
Ram Polepeddi, Westwood college, Denver North Campus
Ron Spicer, Colorado Technical University
Azar Raiszadeh, Chattanooga State Technical Community College
Traei Reed, St.Johns River Community College
Michael L. Russo, Suffolk County Community College
Janel Schultz, Saint Mary's University of Minnesota
Sankara Sethuraman, Augusta State University
Stephen Soltys, West Chester university of Pennsylvania
Winson Taam, Oakland University
Jennifer L. Taggart, Rockford College
William Truman, University of North Carolina at Pembroke
Bill White, University of South Carolina Upstate
Jim Wienckowski, State University of New York at Buffalo

Stephen M. Wilkerson, Susquehanna University
Hongkai Zhang, East Central University
Shunpu Zhang, University of Alaska, Fairbanks
Cathy Zucco-Teveloff, Trinity College

We would especially like to thank George Pasles for his careful accuracy review of this text. We are especially appreciative of the excellent work by the editorial and production professionals at Brooks/Cole Cengage. In particular we thank Molly Taylor, Shaylin Walsh, Jill Clark, and Heather Johnson.

Without their creative insight and attention to detail, a project of this quality and magnitude would not be possible. Finally, we acknowledge the cooperation of Minitab, Inc., SPSS, Texas Instruments, and Microsoft.

Charles Henry Brase

Corrinne Pellillo Brase

Additional Resources—Get More from Your Textbook!

Instructor Resources

Annotated Instructors's Edition (**AIE**) Answers to all exercises, teaching comments, and pedagogical suggestions appear in the margin, or at the end of the text in the case of large graphs.

Solution Builder Contains complete solutions to all exercises in the text, including those in the Chapter Review and Cumulative Review Problems in online format. Solution Builder allows instructors to create customized, secure PDF printouts of solutions matched exactly to the exercises assigned for class. Available to adoptions by signing up at www.cengage.com/solutionbuilder

ExamView® Allows instructors to create, deliver, and customize tests for class in print and online formats and features automatic grading. This electronic test bank features more than 450 questions based on the text. All test items are also provided in PDF and Microsoft® Word formats for instructors who opt not to use the software component

PowerLecture Provides the instructor with dynamic media tools for teaching. Create, deliver, and customize tests (both print and online) in minutes with ExamView® Computerized Testing Featuring Algorithmic Equations. Easily build solution sets for homework or exams using Solution Builder's online solutions manual. Microsoft® PowerPoint® lecture slides and figures from the book are also included on this CD-ROM.

Student Resources

Student Solutions Manual Provides solutions to the odd-numbered section and chapter exercises and to all the Cumulative Review exercises in the student textbook.

Instructional DVDs Hosted by Dana Mosely, these text-specific DVDs cover all sections of the text and provide explanations of key concepts, examples, exercises, and applications in a lecture-based format. DVDs are close-captioned for the hearing-impaired.

Aplia Is an online interactive learning solution that helps students improve comprehension—and their grade—by integrating a variety of mediums and tools such as video, tutorials, practice tests, and an interactive eBook. Created by a professor to enhance his own courses, Aplia provides automatically graded assignments with detailed, immediate feedback on every question, and innovative teaching materials. More than 1,000,000 students have used Aplia at over 1,800 institutions.

MINITAB® and IBM SPSS Statistics CD-ROMs These statistical software packages manipulate and interpret data to produce textual, graphical, and tabular results. MINITAB and/or SPSS may be packaged with the textbook. Student versions are available.

CourseMate Brings course concepts to life with interactive learning, study, and exam preparation tools that support the printed textbook. Watch student comprehension soar as your class works with the printed textbook and the textbook-specific website. Statistics CourseMate goes beyond the book to deliver what you need. Find the following and more at www.cengage.com/statistics/brase.

- **Engagement Tracker** a first-of-its-kind tool that monitors student engagement in the course. Online student quizzes
- **Technology Guides** Separate guides exist with information and examples for each of four technology tools. Guides are available for the TI-84Plus, TI-S3Plus, and TI-*n*spire graphing calculators, Minitab software (version 14) Microsoft Excel (2008/2007), and SPSS Statistics software.
- **Interactive Teaching and Learning Tools** include glossary flashcards, online datasets (in Microsoft Excel, Minitab, SPSS, and Tl-84Plus/TI-83Plus/TI-*n*spire with 84Plus keypad ASCII file formats), statistical tables and formulae, and more.
- **Multimedia eBook** Integrates numerous assets such as video explanations and tutorials to expand upon and reinforce concepts as they appear in die text.

Enhanced WebAssign Offers an extensive online program for Statistics to encourage the practice that's so critical for concept mastery. The meticulously crafted pedagogy and exercises in Brase and Brase's text become even more effective in Enhanced WebAssign.

CengageBrain.com Provides the freedom to purchase online homework and other materials a-la-carte exactly what you need, when you need it.

For more information, visit **http://www.cengage.com/statistics/brase** or contact your local Cengage Learning sales representative.

Table of Prerequisite Material

Chapter		Prerequisite Sections
1	Getting Started	None
2	Organizing Data	1.1, 1.2
3	Averages and Variation	1.1, 1.2, 2.1
4	Elementary Probability Theory	1.1, 1.2, 2.1, 3.1, 3.2
5	The Binomial Probability Distribution and Related Topics	1.1, 1.2, 2.1, 3.1, 3.2, 4.1, 4.2 4.3 useful but not essential
6	Normal Curves and Sampling Distributions	
	(omit 6.6)	1.1, 1.2, 2.1, 3.1, 3.2, 4.1, 4.2, 5.1
	(include 6.6)	also 5.2, 5.3
7	Estimation	
	(omit 7.3 and parts of 7.4)	1.1, 1.2, 2.1, 3.1, 3.2, 4.1, 4.2, 5.1, 6.1, 6.2, 6.3, 6.4, 6.5
	(include 7.3 and all of 7.4)	also 5.2, 5.3, 6.6
8	Hypothesis Testing	
	(omit 8.3 and part of 8.5)	1.1, 1.2, 2.1, 3.1, 3.2, 4.1, 4.2, 5.1, 6.1, 6.2, 6.3, 6.4, 6.5
	(include 8.3 and all of 8.5)	also 5.2, 5.3, 6.6
9	Correlation and Regression	
	(9.1 and 9.2)	1.1, 1.2, 3.1, 3.2
	(9.3 and 9.4)	also 4.1, 4.2, 5.1, 6.1, 6.2, 6.3, 6.4, 6.5, 7.1, 7.2, 8.1, 8.2
10	Chi-Square and F Distributions	
	(omit 10.3)	1.1, 1.2, 2.1, 3.1, 3.2, 4.1, 4.2, 5.1, 6.1, 6.2, 6.3, 6.4,
	(include 10.3)	6.5, 8.1 also 7.1
11	Nonparametric Statistics	1.1, 1.2, 2.1, 3.1, 3.2, 4.1, 4.2, 5.1, 6.1, 6.2, 6.3, 6.4, 6.5, 8.1, 8.3