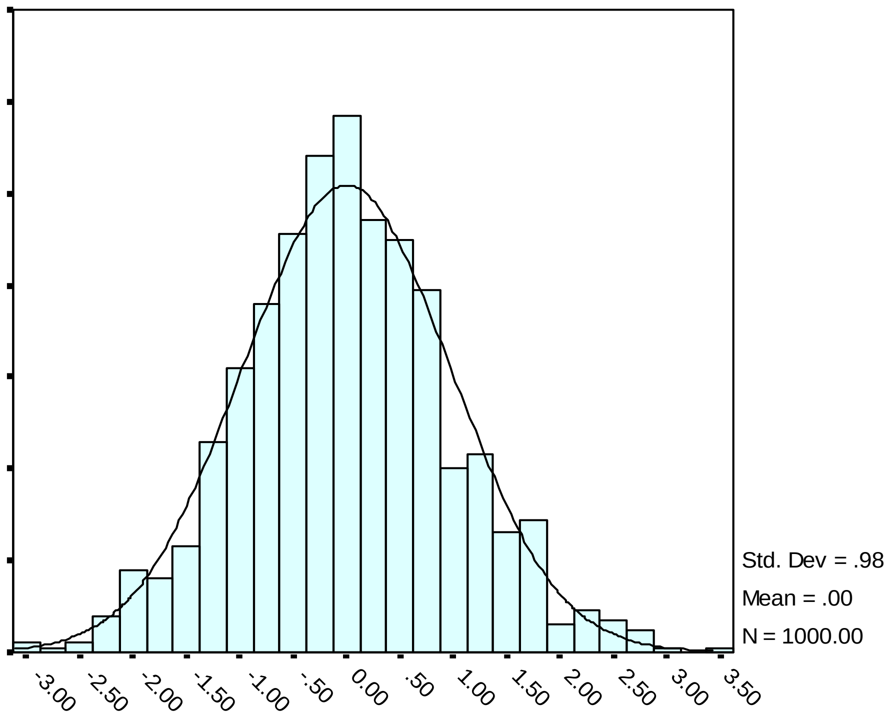
Lecture 2: The Normal Distribution

• 2. 1 Normal Distributions

- It is a theoretical function, typically represented by a graph of the functional relation generated by that function

- The x-axis represents all the specific values x that a random variable X might take

- The y-axis represents the density function $f(x)$ for each value x of X

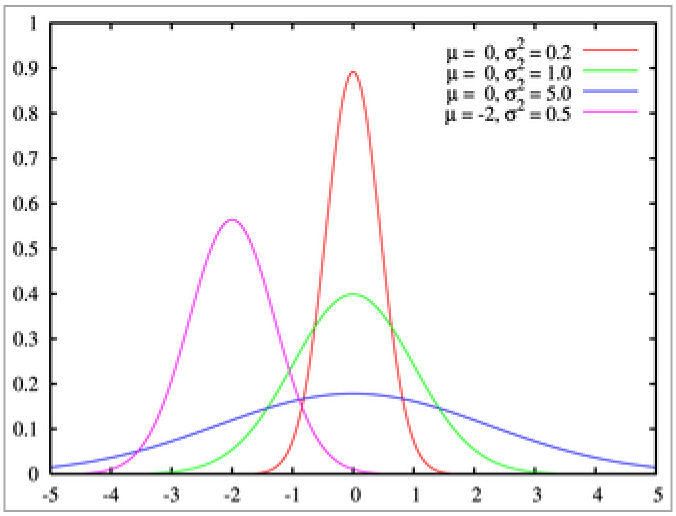


- The mathematical rule for a normal density function is as follows:

$$f(x;\mu,\sigma^2)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\left(\frac{(x-\mu)^2}{2\sigma^2}\right)}$$

where the precise density value assigned to any value x by this rule cannot be found unless μ and σ^2 are specified

- The **normal distribution**, also called **Gaussian distribution** (named after Carl Friedrich Gauss, a German mathematician, although Gauss was not the first to work with it), is an extremely important probability distribution in many fields. It is a family of distributions of the same general form, differing in their *location* and *scale* parameters: the mean ("average") and standard deviation ("variability"), respectively. The **standard normal distribution** is the normal distribution with a mean of zero and a standard deviation of one (the green curves in the plots). It is often called the **bell curve** because the graph of its probability density resembles a bell.



- **2. 2 Characterization of the normal distribution**

- There are various ways to characterize a probability distribution. The most visual is the probability density function (plot at the top), which represents how likely each value of the random variable is. The cumulative distribution function (below) is a conceptually cleaner way to specify the same information, but to the untrained eye its plot is much less informative (see below). Equivalent ways to specify the normal distribution are: the moments, the cumulants, the characteristic function, the moment-generating function, the cumulant-generating function, and Maxwell's theorem. Some of these are very useful for theoretical work, but not intuitive. See probability distribution for a discussion.

- ## 2. 2.1 Probability density function

- The probability density function of the **normal distribution** with mean μ and variance σ^2 (equivalently, standard deviation σ) is an example of a Gaussian function,

$$f(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) = \frac{1}{\sigma}\varphi\left(\frac{x-\mu}{\sigma}\right),$$

where

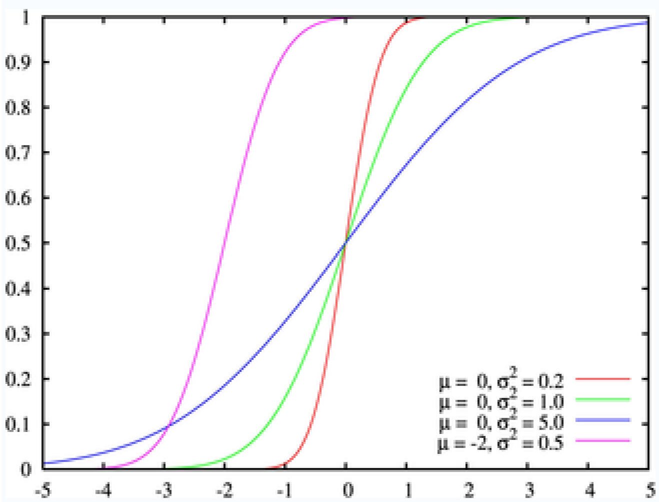
$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

is the density function of the "standard" normal distribution, i.e., the normal distribution with μ = 0 and σ = 1.

- To indicate that a random variable *X* has a distribution with density function *f*(· ; μ, σ), we write

$$X \sim N(\mu, \sigma^2).$$

- probability density function of the normal distribution for various parameter values.



- Some notable qualities of the normal distribution:

  - The density function is symmetric about its mean value.

  - The mean is also its mode and median.

  - 68.26894921371% of the area under the curve is within one standard deviation of the mean.

  - 95.44997361036% of the area is within two standard deviations.

  - 99.73002039367% of the area is within three standard deviations.

  - The inflection points of the curve occur at one standard deviation away from the mean.

• 2. 3 Standardized Scores

- A standardized score shows the relative status of a score in a distribution

- A standardized score, or ***z score***, expresses the deviation of a given score from the mean in standard deviation units as

$$z = \frac{x - \bar{x}}{S}$$

- When all the raw scores in a distribution are converted to z scores the new distribution has a mean of 0

$$\bar{z} = \sum_{i=1}^{N} \frac{z_i}{N} = \sum_{i=1}^{N} \frac{(x_i - \bar{x})}{NS}$$, because N and S are constant over the summation, then

$$\bar{z} = \frac{1}{NS} \sum_{i=1}^{N} (x_i - \bar{x}) = 0$$ (the sum of the deviations about the mean is always 0)

- When all the raw scores in a distribution are converted to z scores the new distribution has a standard deviation of 1.00

$$S_z{}^2 = \sum_{i=1}^{N} \frac{(x_i - \bar{x})^2}{NS^2} = \frac{S^2}{S^2} = 1$$, and S = 1

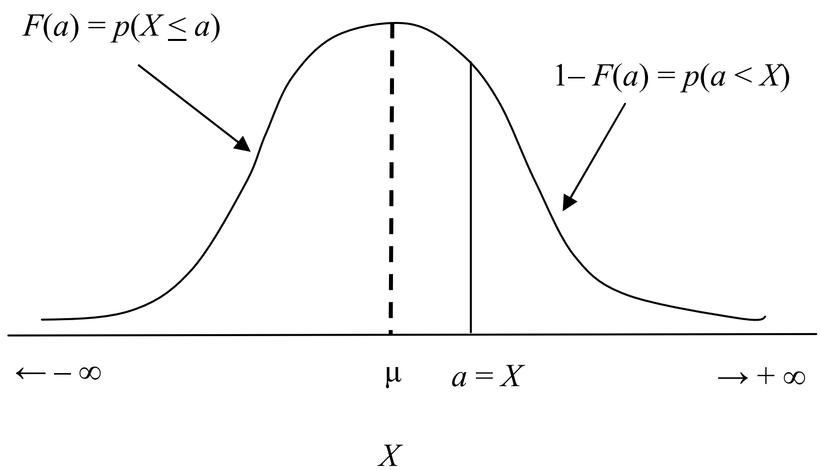
- Although creating standardized scores changes the $\bar{x}$ and S, it does not alter the form of the distribution and the frequencies of the z scores are the same as the frequencies of the corresponding raw scores

• 2. 4 Cumulative Probabilities and Areas for the Normal Distribution

- The cumulative probability at some value of $X = a$ can be expressed as $F(a)$, or

$$F(a) = p(X \leq a)$$

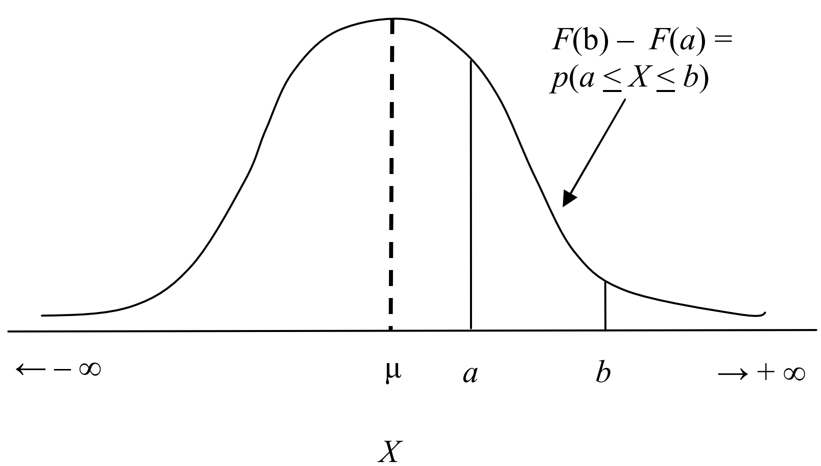
which represents the area under the normal curve between $-\infty$ and a (a included)



A cumulative probability

- Cumulative normal probabilities can be used to find the probability of any interval. For example, the probability that a particular value of X lies between a and b (if $a < b$), given some μ and σ^2, can be expressed as

$p(a \leq X \leq b) = F(b) - F(a)$, and this can be represented graphically as



The probability of an interval in a normal distribution

- The cumulative probabilities for different standardized values can be found in most tables of the normal distribution

  - for a given *z-score*, the table of a standardized normal distribution gives the cumulative probability up to and including that standardized score in a normal distribution

    $F(a) = p(z \leq a)$.

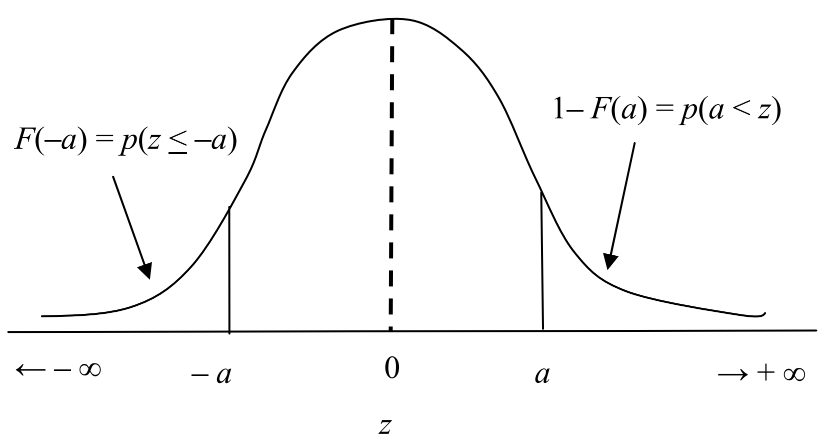
  - Example 1: what is the normal probability of a score $x = 57.5$ given a normal distribution of mean = 50 and standard deviation = 5?

$$z = \frac{57.5 - 50}{5} = 1.5$$

This represents a cumulative probability = .933, approximately. This is the probability of observing a z value less than or equal to 1.5 in a normal distribution

  - If the z score is negative, $F(-z) = 1 - F(z)$

    $1 - .933 = .067$



Probabilities of the extreme tails in a normal distribution

- Example 2: What is the cumulative probability of a score $x = 100$, in a normal distribution with mean = 107 and standard deviation = 70

$$z = \frac{100 - 107}{70} = -.1$$

The cumulative probability of z equal .1 is .5398. Hence, the cumulative frequency of z equal – .1 must be

$F(-.1) = .1 - .5398 = .4602$

- Example 3: What proportion of cases in a normal distribution lies within one standard deviation of the mean?

$p(-1 \le z \le 1) = F(1) - F(-1) =$

$.8413 - .1587 = .6826$

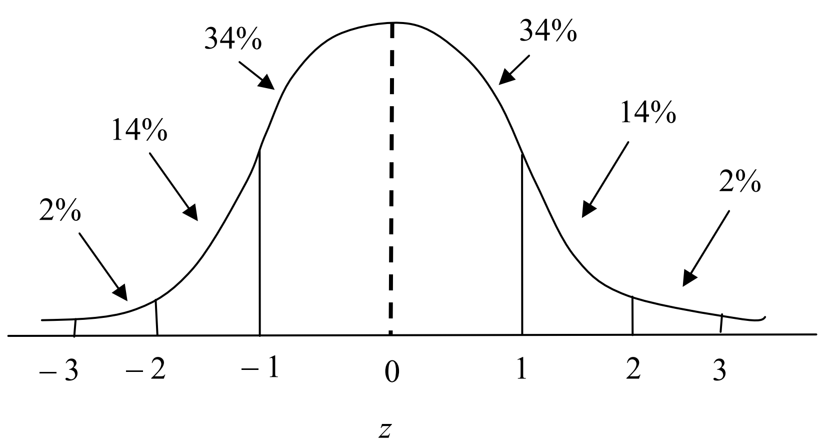
and this indicates that about 68% of all cases in a normal distribution must be contained within one standard deviation of the mean

- Similarly, if we want to find the proportion of cases to be found within two standard deviation from the mean

$p(-2 \le z \le 2) = F(2) - F(-2) =$

$.9772 - .0228 = .9544$,

and this indicates that about 95% of all cases in a normal distribution must be contained within two standard deviations of the mean



• 2. 5 Important Properties of Normal Population Distributions

- Normal distributions have very desirable mathematical properties that allow solutions to numerous statistical problems

- Given random samples of N independent observations, each drawn from a normal population, the distribution of the sample means x is normal, irrespective of the size of N, ($N > 0$)

  - The sampling distribution of means will be normal with expectation equal to μ and standard error equal to $\frac{\sigma}{\sqrt{N}}$

- Given random and independent observations, the sample mean $\bar{x}$ and the sample variance S^2 are independent *if and only if* the population distribution is normal

  - When a normal distribution is sampled, the information contained in the sample mean does not affect the value of the sample variance, and vice versa

  - Unless the distribution sampled is normal, the sample mean and sample variance are not independent across samples

• 2. 6 The Central Limit Theorem

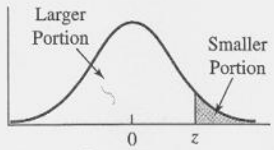
- If a population has a finite variance σ^2 and a finite mean μ, the distribution of sample means from samples of N independent observations approaches the form of a normal distribution with variance $\frac{\sigma^2}{N}$ and mean μ as the sample size N increases. When N is very large, the sampling distribution of the mean is approximately normal regardless of the distribution of the parent population

- The closer to a normal distribution the original population distribution, the closer the sampling distribution of the mean will be to a normal distribution for any sample size

- If the distribution of the original population is normal, the sampling distribution of the mean will also be normal, regardless of the sample size

- **2. 7 Tables of the Normal Distribution**

Table E.10
The Normal Distribution (z)



z	Mean to z	Larger Portion	Smaller Portion	z	Mean to z	Larger Portion	Smaller Portion
.00	0.0000	0.5000	0.5000	.40	0.1554	0.6554	0.3446
.01	0.0040	0.5040	0.4960	.41	0.1591	0.6591	0.3409
.02	0.0080	0.5080	0.4920	.42	0.1628	0.6628	0.3372
.03	0.0120	0.5120	0.4880	.43	0.1664	0.6664	0.3336
.04	0.0160	0.5160	0.4840	.44	0.1700	0.6700	0.3300
.05	0.0199	0.5199	0.4801	.45	0.1736	0.6736	0.3264
.06	0.0239	0.5239	0.4761	.46	0.1772	0.6772	0.3228
.07	0.0279	0.5279	0.4721	.47	0.1808	0.6808	0.3192
.08	0.0319	0.5319	0.4681	.48	0.1844	0.6844	0.3156
.09	0.0359	0.5359	0.4641	.49	0.1879	0.6879	0.3121
.10	0.0398	0.5398	0.4602	.50	0.1915	0.6915	0.3085
.11	0.0438	0.5438	0.4562	.51	0.1950	0.6950	0.3050
.12	0.0478	0.5478	0.4522	.52	0.1985	0.6985	0.3015
.13	0.0517	0.5517	0.4483	.53	0.2019	0.7019	0.2981
.14	0.0557	0.5557	0.4443	.54	0.2054	0.7054	0.2946
.15	0.0596	0.5596	0.4404	.55	0.2088	0.7088	0.2912
.16	0.0636	0.5636	0.4364	.56	0.2123	0.7123	0.2877
.17	0.0675	0.5675	0.4325	.57	0.2157	0.7157	0.2843
.18	0.0714	0.5714	0.4286	.58	0.2190	0.7190	0.2810
.19	0.0753	0.5753	0.4247	.59	0.2224	0.7224	0.2776
.20	0.0793	0.5793	0.4207	.60	0.2257	0.7257	0.2743
.21	0.0832	0.5832	0.4168	.61	0.2291	0.7291	0.2709
.22	0.0871	0.5871	0.4129	.62	0.2324	0.7324	0.2676
.23	0.0910	0.5910	0.4090	.63	0.2357	0.7357	0.2643
.24	0.0948	0.5948	0.4052	.64	0.2389	0.7389	0.2611
.25	0.0987	0.5987	0.4013	.65	0.2422	0.7422	0.2578
.26	0.1026	0.6026	0.3974	.66	0.2454	0.7454	0.2546
.27	0.1064	0.6064	0.3936	.67	0.2486	0.7486	0.2514
.28	0.1103	0.6103	0.3897	.68	0.2517	0.7517	0.2483
.29	0.1141	0.6141	0.3859	.69	0.2549	0.7549	0.2451
.30	0.1179	0.6179	0.3821	.70	0.2580	0.7580	0.2420
.31	0.1217	0.6217	0.3783	.71	0.2611	0.7611	0.2389
.32	0.1255	0.6255	0.3745	.72	0.2642	0.7642	0.2358
.33	0.1293	0.6293	0.3707	.73	0.2673	0.7673	0.2327
.34	0.1331	0.6331	0.3669	.74	0.2704	0.7704	0.2296
.35	0.1368	0.6368	0.3632	.75	0.2734	0.7734	0.2266
.36	0.1406	0.6406	0.3594	.76	0.2764	0.7764	0.2236
.37	0.1443	0.6443	0.3557	.77	0.2794	0.7794	0.2206
.38	0.1480	0.6480	0.3520	.78	0.2823	0.7823	0.2177
.39	0.1517	0.6517	0.3483	.79	0.2852	0.7852	0.2148

continued

(*Source:* The entries in this table were computed by the author.)

Table E.10
Continued

z	Mean to z	Larger Portion	Smaller Portion	z	Mean to z	Larger Portion	Smaller Portion
.80	0.2881	0.7881	0.2119	1.29	0.4015	0.9015	0.0985
.81	0.2910	0.7910	0.2090	1.30	0.4032	0.9032	0.0968
.82	0.2939	0.7939	0.2061	1.31	0.4049	0.9049	0.0951
.83	0.2967	0.7967	0.2033	1.32	0.4066	0.9066	0.0934
.84	0.2995	0.7995	0.2005	1.33	0.4082	0.9082	0.0918
.85	0.3023	0.8023	0.1977	1.34	0.4099	0.9099	0.0901
.86	0.3051	0.8051	0.1949	1.35	0.4115	0.9115	0.0885
.87	0.3078	0.8078	0.1922	1.36	0.4131	0.9131	0.0869
.88	0.3106	0.8106	0.1894	1.37	0.4147	0.9147	0.0853
.89	0.3133	0.8133	0.1867	1.38	0.4162	0.9162	0.0838
.90	0.3159	0.8159	0.1841	1.39	0.4177	0.9177	0.0823
.91	0.3186	0.8186	0.1814	1.40	0.4192	0.9192	0.0808
.92	0.3212	0.8212	0.1788	1.41	0.4207	0.9207	0.0793
.93	0.3238	0.8238	0.1762	1.42	0.4222	0.9222	0.0778
.94	0.3264	0.8264	0.1736	1.43	0.4236	0.9236	0.0764
.95	0.3289	0.8289	0.1711	1.44	0.4251	0.9251	0.0749
.96	0.3315	0.8315	0.1685	1.45	0.4265	0.9265	0.0735
.97	0.3340	0.8340	0.1660	1.46	0.4279	0.9279	0.0721
.98	0.3365	0.8365	0.1635	1.47	0.4292	0.9292	0.0708
.99	0.3389	0.8389	0.1611	1.48	0.4306	0.9306	0.0694
1.00	0.3413	0.8413	0.1587	1.49	0.4319	0.9319	0.0681
1.01	0.3438	0.8438	0.1562	1.50	0.4332	0.9332	0.0668
1.02	0.3461	0.8461	0.1539	1.51	0.4345	0.9345	0.0655
1.03	0.3485	0.8485	0.1515	1.52	0.4357	0.9357	0.0643
1.04	0.3508	0.8508	0.1492	1.53	0.4370	0.9370	0.0630
1.05	0.3531	0.8531	0.1469	1.54	0.4382	0.9382	0.0618
1.06	0.3554	0.8554	0.1446	1.55	0.4394	0.9394	0.0606
1.07	0.3577	0.8577	0.1423	1.56	0.4406	0.9406	0.0594
1.08	0.3599	0.8599	0.1401	1.57	0.4418	0.9418	0.0582
1.09	0.3621	0.8621	0.1379	1.58	0.4429	0.9429	0.0571
1.10	0.3643	0.8643	0.1357	1.59	0.4441	0.9441	0.0559
1.11	0.3665	0.8665	0.1335	1.60	0.4452	0.9452	0.0548
1.12	0.3686	0.8686	0.1314	1.61	0.4463	0.9463	0.0537
1.13	0.3708	0.8708	0.1292	1.62	0.4474	0.9474	0.0526
1.14	0.3729	0.8729	0.1271	1.63	0.4484	0.9484	0.0516
1.15	0.3749	0.8749	0.1251	1.64	0.4495	0.9495	0.0505
1.16	0.3770	0.8770	0.1230	1.65	0.4505	0.9505	0.0495
1.17	0.3790	0.8790	0.1210	1.66	0.4515	0.9515	0.0485
1.18	0.3810	0.8810	0.1190	1.67	0.4525	0.9525	0.0475
1.19	0.3830	0.8830	0.1170	1.68	0.4535	0.9535	0.0465
1.20	0.3849	0.8849	0.1151	1.69	0.4545	0.9545	0.0455
1.21	0.3869	0.8869	0.1131	1.70	0.4554	0.9554	0.0446
1.22	0.3888	0.8888	0.1112	1.71	0.4564	0.9564	0.0436
1.23	0.3907	0.8907	0.1093	1.72	0.4573	0.9573	0.0427
1.24	0.3925	0.8925	0.1075	1.73	0.4582	0.9582	0.0418
1.25	0.3944	0.8944	0.1056	1.74	0.4591	0.9591	0.0409
1.26	0.3962	0.8962	0.1038	1.75	0.4599	0.9599	0.0401
1.27	0.3980	0.8980	0.1020	1.76	0.4608	0.9608	0.0392
1.28	0.3997	0.8997	0.1003	1.77	0.4616	0.9616	0.0384

Table E.10
Continued

z	Mean to z	Larger Portion	Smaller Portion	z	Mean to z	Larger Portion	Smaller Portion
1.78	0.4625	0.9625	0.0375	2.28	0.4887	0.9887	0.0113
1.79	0.4633	0.9633	0.0367	2.29	0.4890	0.9890	0.0110
1.80	0.4641	0.9641	0.0359	2.30	0.4893	0.9893	0.0107
1.81	0.4649	0.9649	0.0351	2.31	0.4896	0.9896	0.0104
1.82	0.4656	0.9656	0.0344	2.32	0.4898	0.9898	0.0102
1.83	0.4664	0.9664	0.0336	2.33	0.4901	0.9901	0.0099
1.84	0.4671	0.9671	0.0329	2.34	0.4904	0.9904	0.0096
1.85	0.4678	0.9678	0.0322	2.35	0.4906	0.9906	0.0094
1.86	0.4686	0.9686	0.0314	2.36	0.4909	0.9909	0.0091
1.87	0.4693	0.9693	0.0307	2.37	0.4911	0.9911	0.0089
1.88	0.4699	0.9699	0.0301	2.38	0.4913	0.9913	0.0087
1.89	0.4706	0.9706	0.0294	2.39	0.4916	0.9916	0.0084
1.90	0.4713	0.9713	0.0287	2.40	0.4918	0.9918	0.0082
1.91	0.4719	0.9719	0.0281	2.41	0.4920	0.9920	0.0080
1.92	0.4726	0.9726	0.0274	2.42	0.4922	0.9922	0.0078
1.93	0.4732	0.9732	0.0268	2.43	0.4925	0.9925	0.0075
1.94	0.4738	0.9738	0.0262	2.44	0.4927	0.9927	0.0073
1.95	0.4744	0.9744	0.0256	2.45	0.4929	0.9929	0.0071
1.96	0.4750	0.9750	0.0250	2.46	0.4931	0.9931	0.0069
1.97	0.4756	0.9756	0.0244	2.47	0.4932	0.9932	0.0068
1.98	0.4761	0.9761	0.0239	2.48	0.4934	0.9934	0.0066
1.99	0.4767	0.9767	0.0233	2.49	0.4936	0.9936	0.0064
2.00	0.4772	0.9772	0.0228	2.50	0.4938	0.9938	0.0062
2.01	0.4778	0.9778	0.0222	2.51	0.4940	0.9940	0.0060
2.02	0.4783	0.9783	0.0217	2.52	0.4941	0.9941	0.0059
2.03	0.4788	0.9788	0.0212	2.53	0.4943	0.9943	0.0057
2.04	0.4793	0.9793	0.0207	2.54	0.4945	0.9945	0.0055
2.05	0.4798	0.9798	0.0202	2.55	0.4946	0.9946	0.0054
2.06	0.4803	0.9803	0.0197	2.56	0.4948	0.9948	0.0052
2.07	0.4808	0.9808	0.0192	2.57	0.4949	0.9949	0.0051
2.08	0.4812	0.9812	0.0188	2.58	0.4951	0.9951	0.0049
2.09	0.4817	0.9817	0.0183	2.59	0.4952	0.9952	0.0048
2.10	0.4821	0.9821	0.0179	2.60	0.4953	0.9953	0.0047
2.11	0.4826	0.9826	0.0174	2.61	0.4955	0.9955	0.0045
2.12	0.4830	0.9830	0.0170	2.62	0.4956	0.9956	0.0044
2.13	0.4834	0.9834	0.0166	2.63	0.4957	0.9957	0.0043
2.14	0.4838	0.9838	0.0162	2.64	0.4959	0.9959	0.0041
2.15	0.4842	0.9842	0.0158	2.65	0.4960	0.9960	0.0040
2.16	0.4846	0.9846	0.0154	2.66	0.4961	0.9961	0.0039
2.17	0.4850	0.9850	0.0150	2.67	0.4962	0.9962	0.0038
2.18	0.4854	0.9854	0.0146	2.68	0.4963	0.9963	0.0037
2.19	0.4857	0.9857	0.0143	2.69	0.4964	0.9964	0.0036
2.20	0.4861	0.9861	0.0139	2.70	0.4965	0.9965	0.0035
2.21	0.4864	0.9864	0.0136	2.71	0.4966	0.9966	0.0034
2.22	0.4868	0.9868	0.0132	2.72	0.4967	0.9967	0.0033
2.23	0.4871	0.9871	0.0129	2.73	0.4968	0.9968	0.0032
2.24	0.4875	0.9875	0.0125	2.74	0.4969	0.9969	0.0031
2.25	0.4878	0.9878	0.0122	2.75	0.4970	0.9970	0.0030
2.26	0.4881	0.9881	0.0119	2.76	0.4971	0.9971	0.0029
2.27	0.4884	0.9884	0.0116	2.77	0.4972	0.9972	0.0028

continued