

## Topic 10 - The Normal Distribution

### Statistics for Managers

June 3, 1999

The normal random variable appears in many important applications.

- Errors in astronomical observations
- Errors in physical measurements
- Deviation from benchmark measures quality control
- Errors in educational measurement, SAT, GRE, IOWA etc.
- Errors in survey studies: census data, market data, economic data
- Risk in finance: stock market data, portfolio theory, stock returns, option prices

The normal random variable is the first example of a continuous random variable. For a discrete random variable, probabilities of events are obtained by summing the probability mass function. For a continuous random variable, however, probabilities of events are obtained by integrating the probability mass function over a range of values of  $y$ , a method from calculus that computes areas under curves. Thus, the area under the probability mass function plays a big role. The graph of a probability mass function of a continuous random variable is also called a histogram. A histogram is a curve where an area under the curve over a range of  $y$  values represents the probability of the range of  $y$ -values occurring.

## Continuous Random Variable

A random variable  $Y$  that can be any real number in an interval is continuous. Its probability mass function  $f(y)$  determines the probability of the event  $\{b \leq Y \leq a\}$  by the integral formula:

$$P(b \leq Y \leq a) = \int_b^a f(y) dy$$

Clearly, the total integral is one:

$$P(-\infty < Y < +\infty) = \int_{-\infty}^{+\infty} f(y) dy = 1$$

The expect value and variance are given by integration:

$$E(Y) = \int_b^a y f(y) dy = \mu$$

$$VAR(Y) = \int_b^a (y - \mu)^2 f(y) dy = \sigma^2$$

Calculation formula for variance:

$$VAR(Y) = \int_b^a y^2 f(y) dy - \mu^2 = \sigma^2$$

Note that for continuos random variables the probability of the open and closed intervals are the same, that is,

$$P(b \leq Y \leq a) = P(b < Y \leq a) = P(b \leq Y < a) = \int_b^a f(y) dy$$

Note that the term, probability density function, is used for a continuos random variable, whereas, the term, probability mass function, is used for discrete random variables.

## Normal Probability Density Function

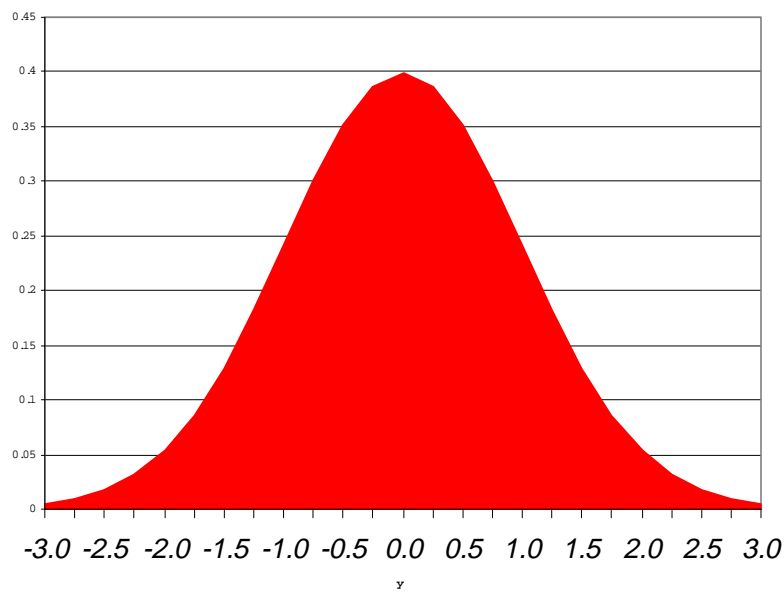
The probability density function  $f(y)$  of the normal distribution with mean  $\mu$  and standard deviation  $\sigma$ , denoted by  $N(\mu, \sigma)$ .

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

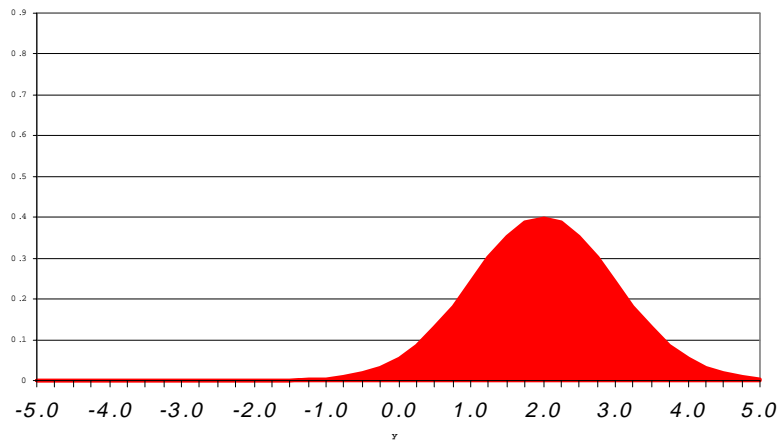
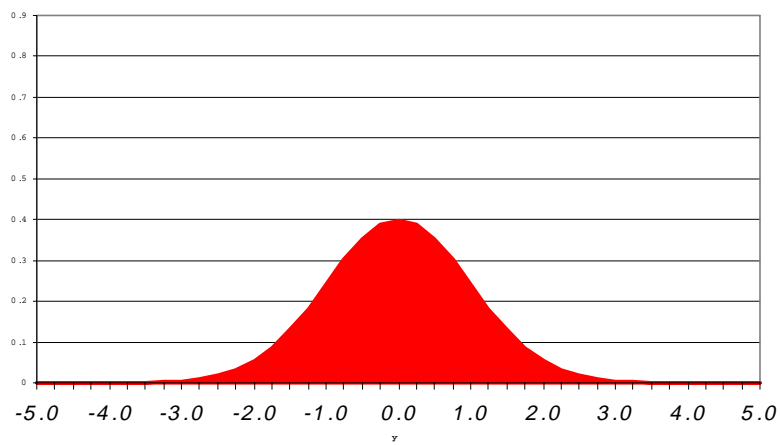
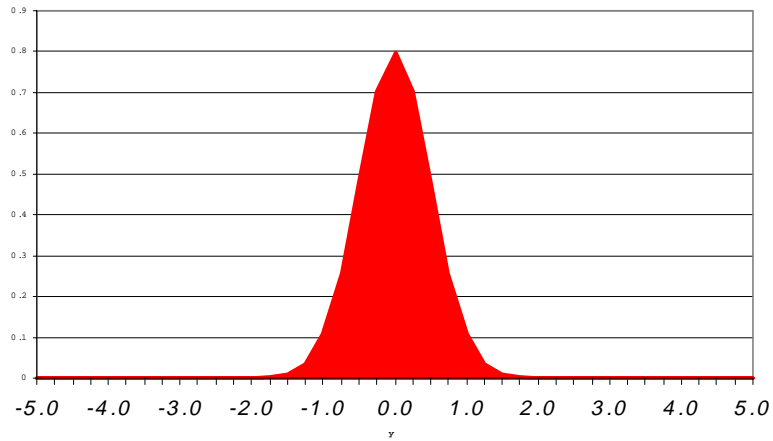
Expected value and variance

$$E(Y) = \mu$$

$$VAR(Y) = \sigma^2$$



• Figure 1 Histogram of  $N(0,1)$



• Figure 2 Histograms of  $N(0, 0.5)$ ,  $N(0, 1)$ , and  $N(2, 1)$ , respectively

## Standard Normal Distribution

The normal distribution with mean 0 and standard deviation 1,  $N(0,1)$ , is called the standard normal distribution.

## The Z-Score

Let  $Y$  be a random variable with expected value  $\mu$  and standard deviation  $\sigma$ , the Z score random variable is defined as follows:

$$Z = \frac{Y - \mu}{\sigma}$$

If  $Y$  is a normal random variable then the Z score is a standard normal random variable.

Let  $y$  be an arbitrary value, the z score value of  $y$  is, using expected value  $\mu$  and standard deviation  $\sigma$  is

$$z = \frac{y - \mu}{\sigma}$$

### Example 1.

Let  $Y$  be  $N(1, 1.5)$ . The z-scores of -0.5 and 2.5 are:

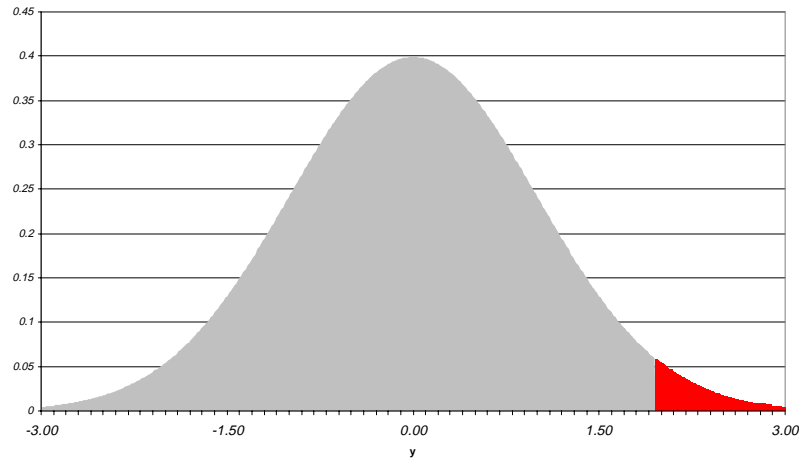
$$\frac{-0.5 - 1}{1.5} = -1 \quad \frac{2.5 - 1}{1.5} = 1$$

## Examples for the Standard Normal Distribution

### Example 2.

Find the probability that a standard normal random variable is greater than 1.96. (Use the complementary event rule of probability.)

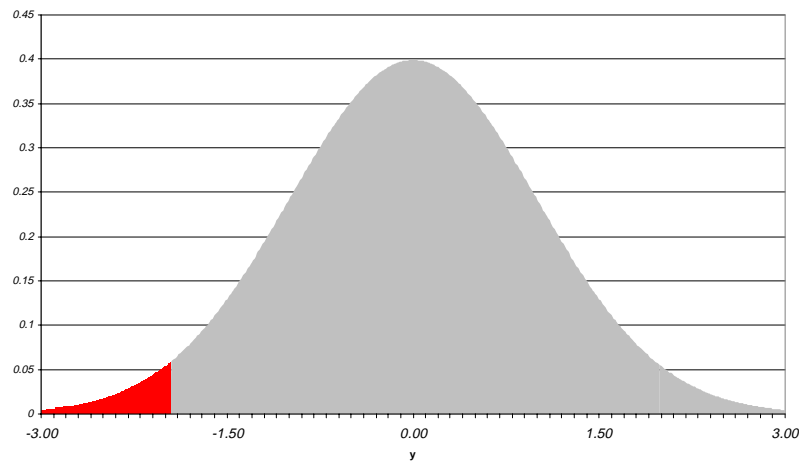
$$P(Z > 1.96) = 1 - P(Z \leq 1.96) = 1 - 0.975 = 0.025$$



### Example 3.

Find the probability that a standard normal random variable is less than -1.96. (Use the symmetry of the normal histogram.)

$$P(Z \leq -1.96) = P(Z > 1.96) = 1 - P(Z \leq 1.96) = 1 - 0.975 = 0.025$$



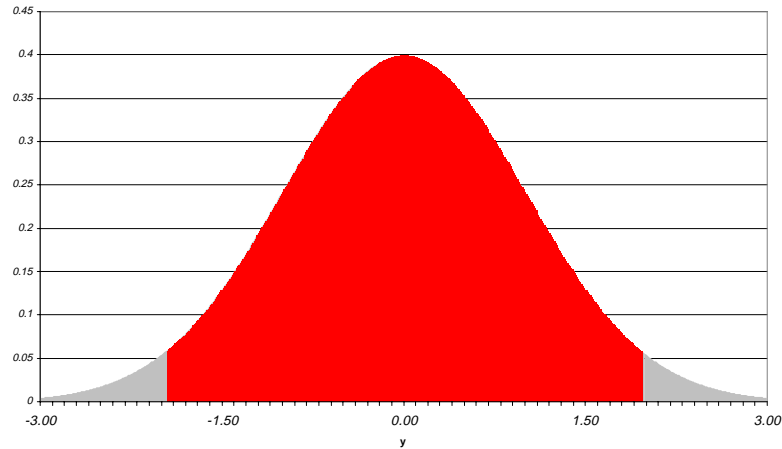
**Example 4.**

Find the probability that a standard normal random variable is between  $-1.96$  and  $1.96$ .

$$P(-1.96 < Z \leq 1.96) = P(Z \leq 1.96) - P(Z \leq -1.96) = 0.975 - 0.025 = 0.95$$

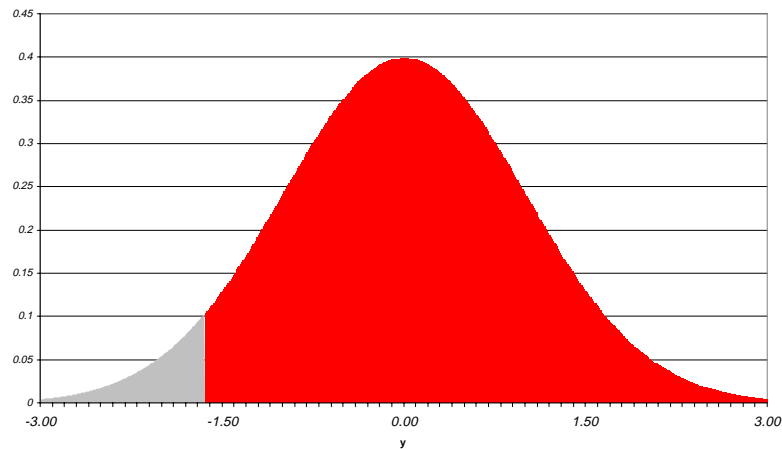
OR

$$P(-1.96 < Z \leq 1.96) = 1 - 2P(Z > 1.96) = 1 - 2(0.025) = 0.95$$



Find the probability that a standard normal random variable is greater than and  $-1.64$ .

$$P(Z > -1.64) = 1 - P(Z \leq -1.64) = 1 - P(Z > 1.64) = P(Z \leq 1.64) = 0.9495$$



Note the important relationship that for any positive  $z$ , we have

$$P(Z > -z) = P(Z \leq z)$$

## Finding Probabilities for an Arbitrary Normal Distribution

Let  $Y$  be a normal random variable with expected value  $\mu$  and standard deviation  $\sigma$ ,  $N(\mu, \sigma)$ . To find the probability that  $Y$  is less than  $y$ , convert  $y$  to the  $Z$  score using  $\mu$  and  $\sigma$ :

$$z = \frac{y - \mu}{\sigma}$$

The probability that  $Y$  is less than  $y$  is equal to the probability that the standard normal random variable  $Z$  is less than  $z$ . Similarly, the probability that  $Y$  is greater than  $y$  is equal to the probability that the standard normal random variable  $Z$  is greater than  $z$ .

### Example 5.

Let  $Y$  be  $N(1.0, 1.5)$ . Find the probability  $Y$  is less than 2.5.

$$z = \frac{2.5 - 1.0}{1.5} = 1.0$$

$$P(Y \leq 2.5) = P(Z \leq 1.0) = .8413$$

### Example 6.

Let  $Y$  be  $N(1.0, 1.5)$ . Find the probability that  $Y$  is larger than zero.

$$z = \frac{0 - 1.0}{1.5} \cong -0.67$$

$$P(Y > 0) = P(Z > -0.67) = P(Z \leq 0.67) = 0.7476$$

### Example 7.

Let  $Y$  be  $N(\mu, \sigma)$ , where the values of  $\mu$  and  $\sigma$  will be specified later. Find the probability that  $Y$  falls within two standard deviations of the expected value. This is, find the probability that  $Y$  falls between  $\mu - 2\sigma$  and  $\mu + 2\sigma$ . The  $Z$  scores of each of these limits are  $-2$  and  $+2$ , respectively.

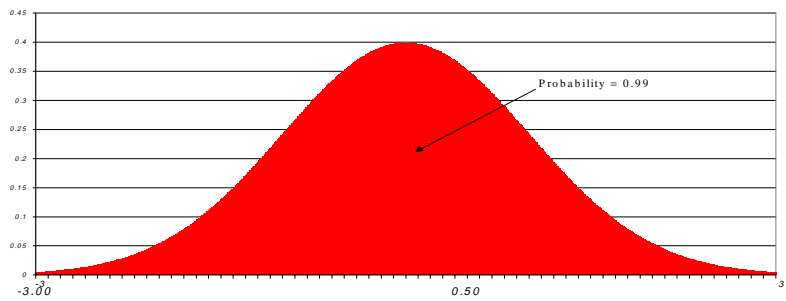
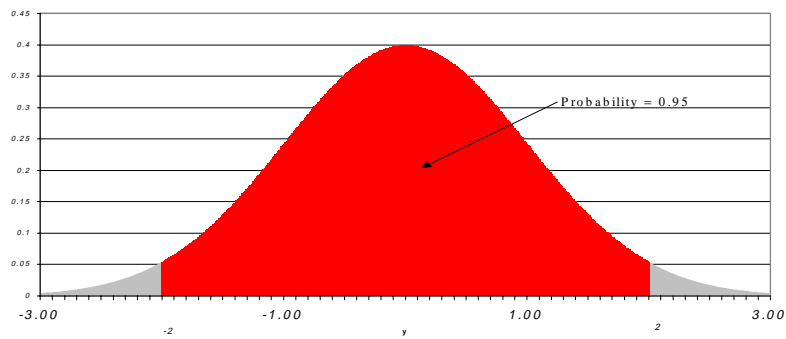
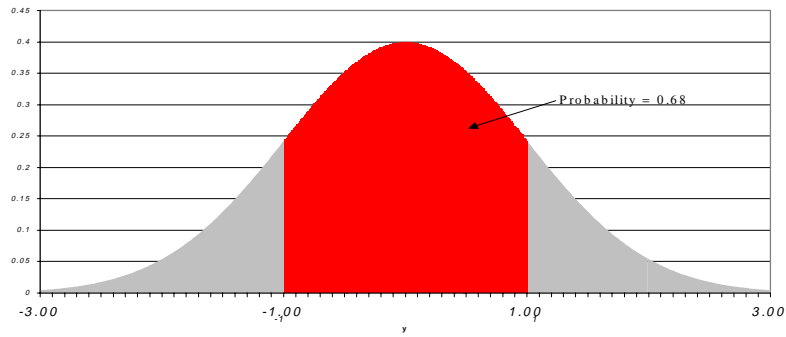
$$\frac{(\mu - 2\sigma) - \mu}{\sigma} = -2$$

$$\frac{(\mu + 2\sigma) - \mu}{\sigma} = +2$$

$$P(\mu - 2\sigma < Y \leq \mu + 2\sigma) = P(-2 < Z \leq +2) \cong P(-1.96 < Z \leq +1.96) \cong 0.95$$

## Probable Errors

k	Probability of Y falling within $k\sigma$ units of the expected value
1	0.68
2	0.95
3	0.99

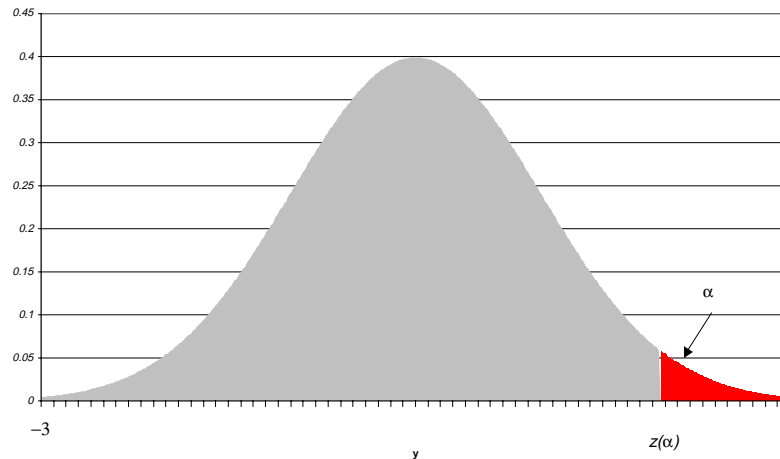


• Figure 3 Probable Errors

## Standard Normal Tail Probabilities and Critical Values

The number  $\alpha$  is a **tail probability** if it is less than 0.5. The value  $z_\alpha$  is called the **critical value** corresponding to  $\alpha$  if  $Z$  is greater than  $z_\alpha$  with probability  $\alpha$ .

$$P(Z > z_\alpha) = \alpha$$



• Figure 4 Tail Probability

### Note

$$P(Z \leq -z_\alpha) = P(Z \geq z_\alpha) = \alpha$$

### Example 8.

The probability that a standard normal random variable is greater than 1.96 is 0.025; thus,  $z_{0.025} = 1.96$ .

### Example 9.

The probability that the absolute value of a standard normal random variable is greater than  $z_{0.025}$  is 0.05; where,  $z_{0.025} = 1.96$ .

$$P(|Z| > z_{0.025}) = P(Z \leq -z_{0.025} \text{ OR } Z > z_{0.025}) =$$

$$P(Z \leq -z_{0.025}) + P(Z > z_{0.025}) = 0.025 + 0.025 = 0.05$$

**Example 10.**

The probability that the absolute value of a standard normal random variable is less than  $z_{0.025}$  is 0.95, where,  $z_{0.025} = 1.96$ .

$$P(|Z| \leq z_{0.025}) = 1 - P(|Z| > z_{0.025}) = 1 - 0.05 = 0.95$$

**Example 11.**

The probability that the absolute value of a standard normal random variable is greater than  $z_{\alpha/2}$  is  $\alpha$ .

$$P(|Z| > z_{\alpha/2}) = \alpha$$

**Example 12.**

The probability that the absolute value of a standard normal random variable is less than  $z_{\alpha/2}$  is  $1 - \alpha$ .

$$P(|Z| \leq z_{\alpha/2}) = 1 - P(|Z| > z_{\alpha/2}) = 1 - \alpha$$

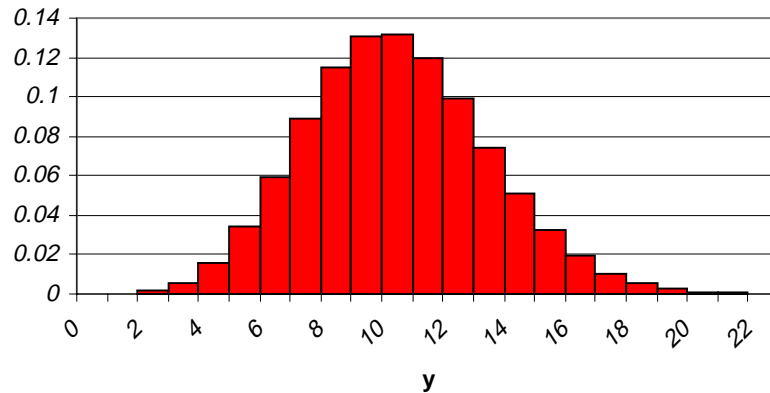
**Critical Values of the Standard Normal Distribution**

$\alpha$	$z_{\alpha}$
0.1000	1.28
0.0500	1.64
0.0250	1.96
0.0100	2.33
0.0050	2.58
0.0025	2.81
0.0010	3.09

## Normal Approximation to the Binomial Distribution

For large  $n$  such that  $np \geq 5$ , the binomial distribution  $BIN(n, p)$  is approximately a normal distribution with mean  $\mu = np$  and standard deviation  $\sigma = [np(1-p)]^{1/2}$

### Example 13. Binomial distributions for increasing number of trials



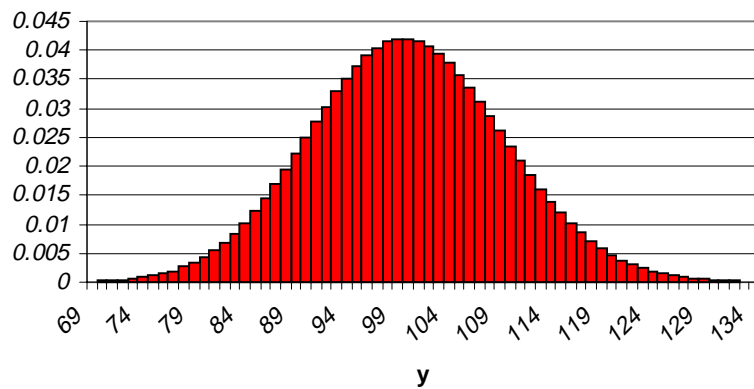
• Figure 5 Histogram of  $BIN(100, 0.1)$ ;  $\mu = 10$ ,  $\sigma = 3$

Using the normal approximation, let's find the probability that the number of successes is less than or equal to 14. Since the rectangle representing  $P(Y = 14)$  extends  $\frac{1}{2}$  of a unit beyond 14, a good approximation to the Binomial area is  $P(Y \leq 14.5)$ . Calculating the z-score of 14.5:

$$\frac{14.5 - 10}{3} = 1.5$$

We see that the approximate probability is  $P(Z \leq 1.5) = 0.9332$ .

Had we not used the  $\frac{1}{2}$  unit adjustment, we would use a z-score of 1.33 to get  $P(Z \leq 1.33) = 0.9082$ . The  $\frac{1}{2}$  unit adjustment is called the "correction for continuity."



• Figure 6 Histogram of BIN (1000, 0.1);  $\mu = 100$ ,  $\sigma = 9.486833$

Again, using the normal approximation, let's find the probability that the number of successes is less than or equal to 104. Since the rectangle representing  $P(Y = 104)$  extends  $\frac{1}{2}$  of a unit beyond 104, a good approximation to Binomial area is  $P(Y \leq 104.5)$ . Calculating the z-score of 104.5:

$$\frac{104.5 - 100}{9.486833} \cong 0.47$$

We see that the approximate probability is  $P(Z \leq 0.47) = 0.6808$ .

Had we not used the  $\frac{1}{2}$  unit adjustment, we would use a z-score of 0.42 to get  $P(Z \leq 0.42) = 0.6628$ .

## Homework 1

A soda-dispensing machine fills 12-oz can of soda using a normal distribution with mean 12.1 oz and standard deviation of 0.02 oz.

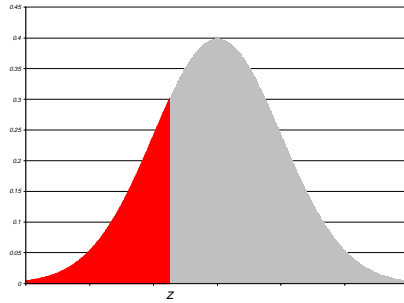
- a. What is the probability that a randomly selected soda contains less than 12 ounces? (Hint: An exact answer may be found using the Excel function NORMDIST.)
- b. What proportion of the soda cans contain between 12.068 ounces and 12.125 ounces of soda?
- c. The plant manager wants to know the amount at which 98.5% of the sodas contain more soda. Help him find this value.

## Homework 2

Suppose  $Y$  has a binomial distribution with  $n = 100$  and  $p = 0.10$ .

- a. Approximate  $P(Y < 15)$  using a normal distribution.
- b. Approximate  $P(Y \leq 15)$  using a normal distribution.

**Table 1. Standard Normal Distribution Function**

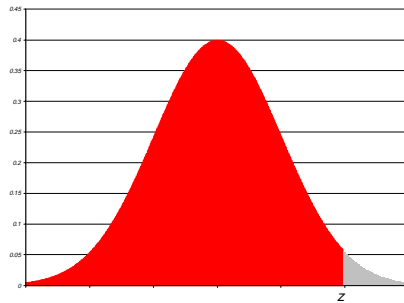


Entries are  $F(z) = P(Z \leq z)$ ,  $z \leq 0$ . Example: To find  $F(-0.75)$ , find the cell with row  $-0.7$  and column  $-0.05$ , which is  $0.2266$ .

-0.09	-0.08	-0.07	-0.06	-0.05	-0.04	-0.03	-0.02	-0.01	0	
0.0010	0.0010	0.0011	0.0011	0.0011	0.0012	0.0012	0.0013	0.0013	0.0013	-3.0
0.0014	0.0014	0.0015	0.0015	0.0016	0.0016	0.0017	0.0018	0.0018	0.0019	-2.9
0.0019	0.0020	0.0021	0.0021	0.0022	0.0022	0.0023	0.0024	0.0025	0.0026	-2.8
0.0026	0.0027	0.0028	0.0029	0.0030	0.0031	0.0032	0.0033	0.0034	0.0035	-2.7
0.0036	0.0037	0.0038	0.0039	0.0040	0.0041	0.0043	0.0044	0.0045	0.0047	-2.6
0.0048	0.0049	0.0051	0.0052	0.0054	0.0055	0.0057	0.0059	0.0060	0.0062	-2.5
0.0064	0.0066	0.0068	0.0069	0.0071	0.0073	0.0075	0.0078	0.0080	0.0082	-2.4
0.0084	0.0087	0.0089	0.0091	0.0094	0.0096	0.0099	0.0102	0.0104	0.0107	-2.3
0.0110	0.0113	0.0116	0.0119	0.0122	0.0125	0.0129	0.0132	0.0136	0.0139	-2.2
0.0143	0.0146	0.0150	0.0154	0.0158	0.0162	0.0166	0.0170	0.0174	0.0179	-2.1
0.0183	0.0188	0.0192	0.0197	0.0202	0.0207	0.0212	0.0217	0.0222	0.0228	-2.0
0.0233	0.0239	0.0244	0.0250	0.0256	0.0262	0.0268	0.0274	0.0281	0.0287	-1.9
0.0294	0.0301	0.0307	0.0314	0.0322	0.0329	0.0336	0.0344	0.0351	0.0359	-1.8
0.0367	0.0375	0.0384	0.0392	0.0401	0.0409	0.0418	0.0427	0.0436	0.0446	-1.7
0.0455	0.0465	0.0475	0.0485	0.0495	0.0505	0.0516	0.0526	0.0537	0.0548	-1.6
0.0559	0.0571	0.0582	0.0594	0.0606	0.0618	0.0630	0.0643	0.0655	0.0668	-1.5
0.0681	0.0694	0.0708	0.0721	0.0735	0.0749	0.0764	0.0778	0.0793	0.0808	-1.4
0.0823	0.0838	0.0853	0.0869	0.0885	0.0901	0.0918	0.0934	0.0951	0.0968	-1.3
0.0985	0.1003	0.1020	0.1038	0.1056	0.1075	0.1093	0.1112	0.1131	0.1151	-1.2
0.1170	0.1190	0.1210	0.1230	0.1251	0.1271	0.1292	0.1314	0.1335	0.1357	-1.1
0.1379	0.1401	0.1423	0.1446	0.1469	0.1492	0.1515	0.1539	0.1562	0.1587	-1.0
0.1611	0.1635	0.1660	0.1685	0.1711	0.1736	0.1762	0.1788	0.1814	0.1841	-0.9
0.1867	0.1894	0.1922	0.1949	0.1977	0.2005	0.2033	0.2061	0.2090	0.2119	-0.8
0.2148	0.2177	0.2206	0.2236	0.2266	0.2296	0.2327	0.2358	0.2389	0.2420	-0.7
0.2451	0.2483	0.2514	0.2546	0.2578	0.2611	0.2643	0.2676	0.2709	0.2743	-0.6
0.2776	0.2810	0.2843	0.2877	0.2912	0.2946	0.2981	0.3015	0.3050	0.3085	-0.5
0.3121	0.3156	0.3192	0.3228	0.3264	0.3300	0.3336	0.3372	0.3409	0.3446	-0.4
0.3483	0.3520	0.3557	0.3594	0.3632	0.3669	0.3707	0.3745	0.3783	0.3821	-0.3
0.3859	0.3897	0.3936	0.3974	0.4013	0.4052	0.4090	0.4129	0.4168	0.4207	-0.2
0.4247	0.4286	0.4325	0.4364	0.4404	0.4443	0.4483	0.4522	0.4562	0.4602	-0.1
0.4641	0.4681	0.4721	0.4761	0.4801	0.4840	0.4880	0.4920	0.4960	0.5000	0.0

Note:  $P(Z \leq -4) = 0.0000317$ ;  $P(Z \leq -5) = 0.000000287$ ;  $P(Z \leq -6) = 0.0000000099$ ;

**Table 2. Standard Normal Distribution Function**



Entries are  $F(z) = P(Z \leq z)$ ,  $z \geq 0$ . Example: To find  $F(1.96)$ , find the cell with row 1.9 and column 0.06, which is 0.9750.

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990