THE UNIVERSITY OF TEXAS AT AUSTIN

**Department of Information, Risk,
and Operations Management**

BA 386T **Tom Shively**

PROBABILITY CONCEPTS AND NORMAL DISTRIBUTIONS

The fundamental idea underlying any statistical analysis is the idea that probability distributions represent uncertainty. This concept will be used over and over throughout the semester. The goal of a statistical analysis is to determine what the probability distribution is in a specific situation (for example, by estimating the mean and variance of the distribution) and how to use the resulting distribution to make appropriate inferences (for example, to forecast future observations of quarterly sales for a company and to obtain a realistic measure of how accurate the forecast is likely to be).

The goal of this topic summary note is to develop several important probability concepts in very simple contexts. The ideas will be applied to more realistic situations and "real-world" problems throughout the semester but the basic probability concepts and intuition developed here will never change.

One of the important philosophies of the course is to develop intuition about specific concepts in very simple contexts so you get a good understanding of the basic concept. The advantage of using simple contexts initially is that the concept is more easily explained without a complicated context to confuse the issue. Once the concept is understood it is then much easier to apply to the real-world problems that we are actually interested in. If you try to learn the basic concepts in complicated contexts it makes things much more difficult.

Probability Concept #1: Random Variables

This section begins with two definitions. These are the only two formal definitions that will be given in the course. However, it is very helpful to have them as the semester goes along. Specific examples illustrating the definitions follow immediately.

Definition #1: A random variable is a variable that take takes on numerical values determined by the outcome of a random experiment. A random variable is typically denoted by a capital letter such as X or Y.

Note that there are two parts to the definition of a random variable. First, the outcome must be a numerical value. Second, the outcome must be determined by a random experiment (as we

will see, a random experiment is defined very broadly – for example, collecting a random sample of observations is a random experiment).

Definition #2: The probability function expresses the probability that the random variable X takes on the specific value x. The notation that is often used is $pr(X = x)$.

Examples of random variables and probability functions

This section provides three examples to illustrate the concepts of random variables and probability functions. The ideas captured by these examples apply to a wide variety of more complex real-world problems that will be discussed during the semester.

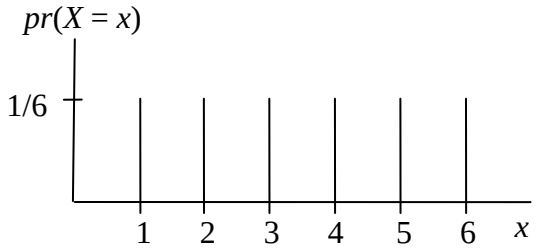
Example #1: Outcome of the roll of a six-faced die:

The outcome of the roll of a die is a random variable. The outcome is a numerical value (in particular, the possible outcomes are 1, 2, 3, 4, 5 and 6). The random experiment that determines the outcome is physically rolling the die.

Very important idea: There is uncertainty about the outcome of the roll of a die. This uncertainty is represented by a probability distribution. The important and fundamental idea to understand is that anytime there is uncertainty about the outcome of a random experiment, this uncertainty can be represented by a probability distribution. This is true in very simple cases such as rolling a die as well as in more complicated situations that will arise later in the semester in real-world problems.

Assuming the die is fair, there are six possible outcomes with equal probability 1/6 associated with each outcome. This information is summarized by the probability function written in numerical form on the left and shown in graphical form on the right:

x	$pr(X = x)$
1	1/6
2	1/6
3	1/6
4	1/6
5	1/6
6	1/6



In the graphical representation of the probability function, the height of the spike at each possible outcome is its associated probability.

It is always useful to graph a probability distribution. A great deal can be learned by just observing the graph, particularly in more complicated situations.

The probability distribution summarizes all available information regarding the uncertainty associated with the outcome of a roll of a die. The information contained in the distribution above (and available at a glance) is that there are six possible outcomes and each outcome is equally likely. Both these pieces of information will be important, for example, if you are planning to gamble on the outcome of the roll of a die. This is a fairly straightforward example but it illustrates in a simple context the basic idea that a probability distribution concisely summarizes all available information about the uncertainty of the outcome of a future event (in this case the outcome of a roll of a die).

There are two equivalent ways to think about probability in this context. First, if the die is rolled many, many times, then on one-sixth of all rolls the number one will be face up. Equivalently, for a given roll of the die there is a one in six chance that a one will be face up (i.e. a probability of 1/6 that a one will be face up). Similar probability statements can be made about the other five possible outcomes (i.e. 2, 3, 4, 5 and 6).

Thinking about probability in this way (which is hopefully fairly intuitive) provides a useful way to interpret probability statements in a wide variety of contexts – see, for example, the discussion of interpreting probability statements in the context of the MBA salary/normal distribution example later in these notes.

Aside

The outcome of a roll of a die is a *discrete* random variable because the outcome can only take on the discrete values: 1, 2, 3, 4, 5 or 6. The definition of a discrete random variable is not particularly important for this class but it is included here for completeness. Most of the random variables we will discuss this semester are continuous random variables such as those in examples #2 and #3 below.

End Aside

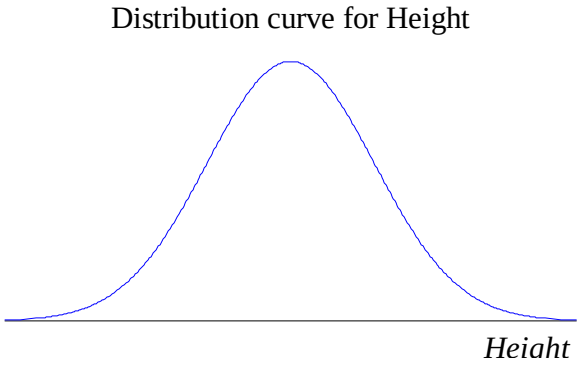
Example #2: Height of a man selected at random from the people walking past the business school

The height of a man selected at random from the people walking past the business school is a random variable. A person's height is a numerical value and the random experiment that determines the value is "selecting a man *at random* from the people walking past the business school."

This is an example of a continuous random variable because a man's height can take on any value in a continuum of values. For example, the man's height might be 70 inches, 71 inches or

any value in the continuum between 70 and 71 inches. The same statement holds for any two reasonable heights.

Since there is uncertainty about the height of the man who is selected, there is a probability distribution that represents this uncertainty. A possible distribution curve is drawn below.



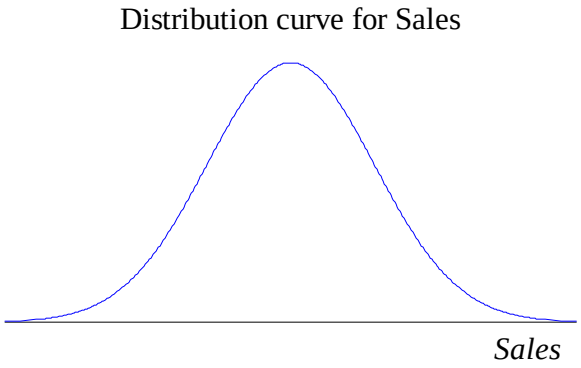
The distribution curve is a continuous curve (not spikes at discrete values as for the discrete random variable in example #1) because the height of a man selected at random is a continuous random variable. The interpretation of continuous distributions and how to compute probabilities associated with various outcomes (such as "what is the probability that the height of the man selected will be between 70 and 71 inches") will be discussed in the section "Normal Distribution" later in these notes.

The important point to take away from this example is that there is uncertainty about what the height of the man selected will be and this uncertainty is represented by a probability distribution.

Example #3: Sales next quarter for a specific company

Sales next quarter for a specific company is a random variable. Quarterly sales is a numerical value and the random experiment that determines the value is "letting the economy run for the next quarter." There are all kinds of forces that affect future sales whose magnitude and impact are uncertain. As mentioned following the definition of a random variable, a random experiment is defined very broadly. Anytime there is uncertainty about the effect of future events on a value of interest (examples of a value of interest are future sales, profits, stock returns, exchange rates, interest rates, etc.) we will interpret the value of interest to have been generated by the random experiment of "letting the economy run."

Since there is uncertainty about next quarter's sales there is a probability distribution that represents this uncertainty. A possible distribution is drawn below.



As discussed later in the semester this distribution can be used to make a prediction of next quarter sales and, just as importantly, to give a realistic and meaningful measure of how accurate the prediction is likely to be. Intuitively, if there is a large amount of uncertainty about what sales will be next quarter (as discussed in the next section "Probability Concept #2: Mean and Standard Deviation" a large amount of uncertainty corresponds to a very "spread out" distribution) then there is likely to be a significant prediction error. Conversely, if there is only a small amount of uncertainty about what sales will be next quarter then the prediction error is likely to be small. The concepts of prediction and confidence in the prediction are discussed in detail later in the semester.

As with the first two examples, the important point to take away from this example is that there is uncertainty about company sales next quarter and this uncertainty is represented by a probability distribution.

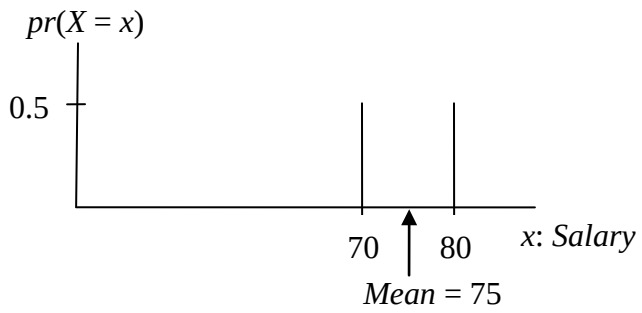
Probability Concept #2: Mean and Standard Deviation

The mean of a probability distribution is a measure of the center of the distribution while the standard deviation is a measure of its dispersion (how spread out the distribution is). The best way to understand the concepts of mean and standard deviation is graphically. We will first consider the mean and then the standard deviation.

Mean

Consider an MBA program at a specific business school. There is uncertainty regarding how much a student graduating from this school will make in their first job so there is a probability distribution that represents this uncertainty. For the sake of this example, suppose there are only two possible salaries a graduating student might make: $70,000 or $80,000. Further, suppose there is a 50/50 chance (i.e. a probability of 0.50) a student will make $70,000 and a 50/50 chance a student will make $80,000. This is an unrealistically simple situation but it is useful for explaining the intuition related to the mean of a distribution.

The probability function representing the uncertainty in salary (in thousands of dollars) for an MBA student graduating from this school is:



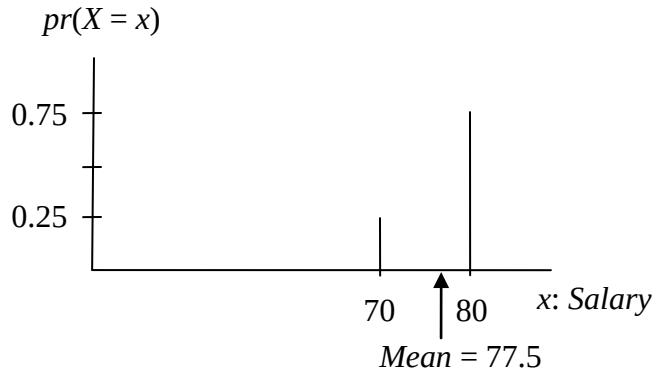
Intuitively, the center of this distribution is 75. One way to interpret the mean graphically is that it is the spot where a teeter-totter will exactly balance, with the horizontal axis of the probability function representing the board of the teeter-totter and the spikes in the probability function representing the weight of the people sitting on it.

The notation we will use for the mean is the Greek letter μ ("mu"). When you see "μ" you should think immediately of the mean (center) of the distribution, as in the figure above. The Greek letter μ is essentially short-hand for writing "the mean of the distribution."

Mathematically, the mean of a distribution is the weighted average of its values where the weights are the probabilities associated with each value. The mean of the above distribution is

$$\mu = Mean = (70)(0.5) + (80)(0.5) = 75.$$

Now consider an MBA program at a second business school. Suppose for this school there is a 25% chance (i.e. a probability of 0.25) that a graduating student will make \$70,000 in their first job and a 75% chance a student will make \$80,000. Note that this is a slightly different distribution than the one for the first school. The probability function representing possible salaries at the second school is



The mean of this distribution has shifted to the right because some of the "weight" associated with *Salary* has shifted from 70 to 80. This is reflected in the new mean

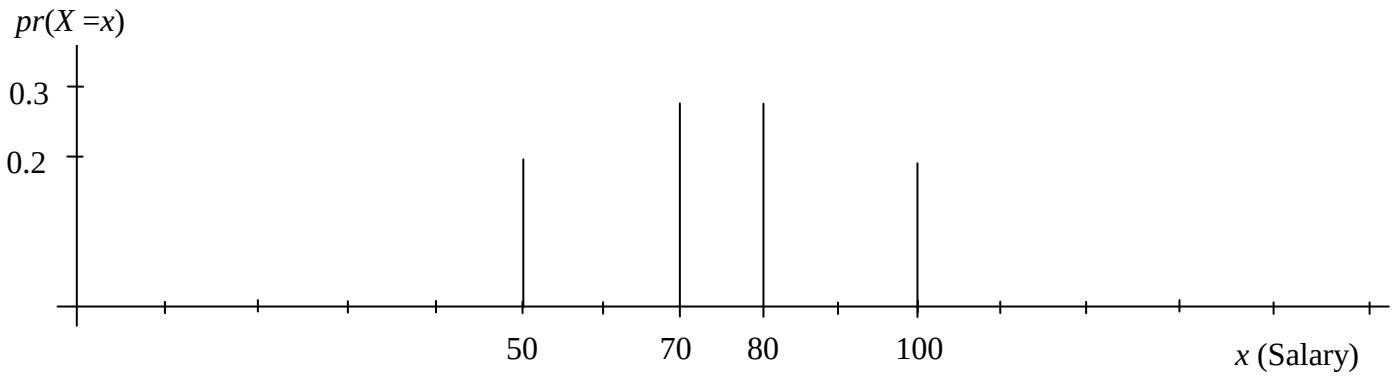
$$\mu = Mean = (70)(0.25) + (80)(0.75) = 77.5.$$

The teeter-totter analogy still works because the center of the teeter-totter (which is analogous to the mean) will be further to the right if the person sitting on the right side is heavier than the person sitting on the left side.

Standard deviation

The standard deviation is a measure of the spread of a distribution. The easiest way to understand intuitively what the standard deviation represents is by considering the concepts of standard deviation and spread graphically.

Consider an MBA program at a specific business school (School #1). For the sake of this example, suppose there are four possible salaries a graduating student might make in their first job: \$50,000 with probability 0.2, \$70,000 with probability 0.3, \$80,000 with probability 0.3 and \$100,000 with probability 0.2. The probability function representing the uncertainty in the salary of an MBA student graduating from this school is

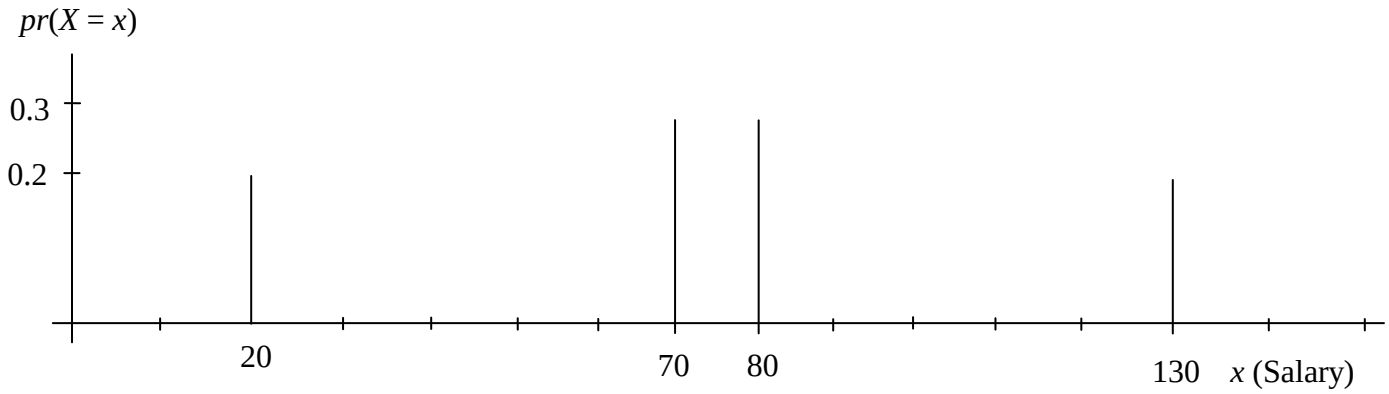


The mean of this distribution is

$$\mu = Mean = (50)(0.2) + (70)(0.3) + (80)(0.3) + (100)(0.2) = 75.$$

The mean can still be thought of as the spot where a teeter-totter will exactly balance even when the random variable (*Salary* in this case) can take on more than two values.

Now consider an MBA program at a second school (School #2). Suppose the probability distribution representing the possible salaries a student graduating from this school might make in their first job is



The mean of this distribution is also 75 since

$$\mu = Mean = (20)(0.2) + (70)(0.3) + (80)(0.3) + (130)(0.2) = 75.$$

However, the salary distributions for Schools #1 and #2 are considerably different even though the means are the same. The distribution for School #2 is more "spread out." There is more uncertainty associated with the second distribution than the first (i.e. there is more uncertainty associated with the salary an MBA student graduating from School #2 will make as compared to the salary a student from School #1 will make).

Important Aside: Summation notation

Summation notation is a very useful notation that represents a short-hand way of writing down the sum of several (possibly many) numbers. This notation will be used in this class as well as other classes such as finance. In particular, it will be used in the formula for standard deviation.

To explain the notation, consider the probability function for salaries from School #1. Let x_1 represent the first possible salary so $x_1 = 50$. Also, let $pr(X = x_1)$ represent the probability associated with $x_1 = 50$ so $pr(X = x_1) = 0.2$. Similarly, let x_2, x_3 and x_4 represent the other possible values (so $x_2 = 70$, $x_3 = 80$, and $x_4 = 100$) and $pr(X = x_2)$, $pr(X = x_3)$ and $pr(X = x_4)$ represent the associated probabilities (so $pr(X = x_2) = 0.3$, $pr(X = x_3) = 0.3$ and $pr(X = x_4) = 0.2$).

Using this notation, the mean can now be written

$$\mu = Mean = \quad (50)(0.2) \quad + \quad (70)(0.3) \quad + \quad (80)(0.3) \quad + \quad (100)(0.2)$$

$$= x_1 pr(X = x_1) + x_2 pr(X = x_2) + \ x_3 pr(X = x_3) + \ x_4 pr(X = x_4) \ .$$

This can be written more concisely in summation notation as

$$\mu = Mean = \sum_{i=1}^{4} x_i \, pr(X = x_i)$$

where $x_i \, pr(X = x_i)$ represents the value and probability associated with the i^{th} point. Thus, for i = 1, we have $x_1 \, pr(X = x_1)$; for i = 2, we have $x_2 \, pr(X = x_2)$; etc. The notation $\sum_{i=1}^{4}$ means to sum the terms $x_i \, pr(X = x_i)$ for i = 1, 2, 3 and 4.

When you see the summation notation $\sum_{i=1}^{4} x_i \, pr(X = x_i)$ you should think of

$$(50)(0.2) + (70)(0.3) + (80)(0.3) + (100)(0.2) = 75$$

i.e. you should think of the sum of each x-value times the probability associated with the x-value.

End Aside

The standard deviation is a measure of the spread, or dispersion, of a distribution. A natural way to measure the spread of a distribution is to look at the average distance each point is from the center of the distribution (i.e. the average distance each point is from the mean of the distribution). The greater this average distance is the more spread out the distribution is.

For School #1, the center of the distribution is μ = 75 so the distance the i^{th} point is from the center is $(x_i - 75)$. For example, for the first point (i = 1), the distance is $x_1 - 75 = 50 - 75 = -25$; for the second point, the distance is $x_2 - 75 = 70 - 75 = -5$; etc.

We want the average distance so we take each distance $(x_i - 75)$ and multiply it by its associated probability to give

$$(x_i - 75) \, pr(X = x_i)$$

and then add these values across all four points (i = 1, 2, 3 and 4)

$$\sum_{i=1}^{4} (x_i - 75) \, pr(X = x_i)$$

to give the average distance. While this is an intuitively appealing measure of spread at first glance, it unfortunately will always be zero. The reason is that the positive and negative distances will always cancel out.

To see this, note that the negative distance $(x_1 - 75) = 50 - 75 = -25$ associated with the first point cancels in the summation with the positive distance $(x_4 - 75) = 100 - 75 = 25$ associated

with the fourth point. Similarly, the negative distance $(x_2 - 75) = 70 - 75 = -5$ associated with the second point cancels with the positive distance $(x_3 - 75) = 80 - 75 = 5$ associated with the third point.

The result is that

$$\sum_{i=1}^{4}(x_i - 75)\,pr(X = x_i)$$

$$= (50 - 75)(0.2) + (70 - 75)(0.3) + (80 - 75)(0.3) + (100 - 75)(0.2)$$

$$= (-25)(0.2) + (-5)(0.3) + (5)(0.3) + (25)(0.2)$$

$$= 0$$

which isn't very useful as a measure of spread.

Aside

It is straightforward, though tedious, to show that $\sum_{i=1}^{n}(x_i - \mu)\,pr(X = x_i) = 0$ for every distribution no matter how complicated the distribution is.

End Aside

A natural way to avoid the problem of positive and negative distances cancelling out is to use the absolute value of the distance each point is from the center of the distribution rather than the actual distance itself, i.e. to use

$$\sum_{i=1}^{4}|\,x_i - 75\,|\,pr(X = x_i)$$

as a measure of spread. While this is a perfectly reasonable and intuitive measure there are some technical problems with using it (which are not important and do not add anything to an intuitive understanding of the measure of spread we will actually use).

Instead of using the average of the absolute values of the distances to avoid the problem of negative and positive terms cancelling out in the summation we use the average of the squares of the distances. Therefore,

$$\sum_{i=1}^{4}(x_i - 75)^2\,pr(X = x_i)$$

is the measure of spread we will use. It is interpreted as the average distance squared each point is from the center of the distribution. The greater the average distance squared is the more spread out the distribution is.

This measure is called the variance of the distribution. The term "variance" comes from the fact that it measures the "variability" in the possible outcomes that *Salary* can take on. The notation used for the variance is σ^2 (σ is the Greek letter "Sigma"). When you see "σ^2" you should think immediately of the spread (dispersion) of the distribution. The notation σ^2 is essentially short-hand for writing "the spread of the distribution."

For School #1, the variance is

$$\sigma^2 = \sum_{i=1}^{4} (x_i - 75)^2 \, pr(X = x_i)$$

$$= (50 - 75)^2 (0.2) + (70 - 75)^2 (0.3) + (80 - 75)^2 (0.3) + (100 - 75)^2 (0.2)$$

$$= 265$$

This means the average distance squared each point is from the center of the distribution is 265.

For School #2, the variance is

$$\sigma^2 = \sum_{i=1}^{4} (x_i - 75)^2 \, pr(X = x_i)$$

$$= (20 - 75)^2 (0.2) + (70 - 75)^2 (0.3) + (80 - 75)^2 (0.3) + (130 - 75)^2 (0.2)$$

$$= 1225$$

The distribution for School #2 is clearly more spread out than the distribution for School #1 and this is reflected in its larger variance.

The variance σ^2 is an intuitively reasonable measure of the spread of a distribution. The disadvantage of the variance is that its units can be difficult to interpret in a meaningful way. For example, in the salary example discussed here, the units are dollars squared (because the units on the distances from each point to the center of the distribution are in dollars and these values are squared to obtain the variance). To put the units back into the original scale, in this case dollars, we often work with $\sigma = \sqrt{\sigma^2}$. The units for σ are dollars and are easily interpretable.

σ is called the standard deviation of the distribution. The term "standard deviation" comes from the fact that its value represents, roughly speaking, the "typical" (or "standard") distance (or "deviation") that each point is from the center of the distribution. When you see "σ" you should think immediately of the spread of the distribution.

The variance of the probability distribution for School #1 is $\sigma^2 = 265$. This means the standard deviation of the distribution is $\sigma = \sqrt{\sigma^2} = \sqrt{265} = 16.28$. Similarly, the variance of the probability distribution for School #2 is $\sigma^2 = 1225$. This means its standard deviation is $\sigma = \sqrt{\sigma^2} = \sqrt{1225} = 35$.

The standard deviation of the distribution for School #2 is approximately twice the standard deviation of the distribution for School #1. Comparing the two distributions graphically shows that, roughly speaking, the second distribution is about twice as spread out as the first distribution. This is what is reflected in the two standard deviations.

A final point to understand is that the standard deviation σ represents the same information as the variance σ^2, although the units are different (dollars and dollars squared). The reason the information content is the same is that if you know the standard deviation you can compute the variance, and vice versa.

Interpretating the spread of the distribution as a measure of risk

The spread, or dispersion, of a distribution can be interpreted as risk in many contexts. In the context of MBA salaries, there is more risk associated with School #2 than School #1. It is possible that a student from School #2 will make a very high salary ($130,000) but there is an equal probability they will make a very low salary ($20,000). If you were making a decision to attend a business school based solely on your anticipated salary at graduation, you would need to determine if the possibility of making $130,000 is worth the risk that you might make only $20,000. School #1 has a smaller standard deviation (i.e. salaries that are less spread out) so there is less uncertainty, or risk, associated with attending this school than School #2.

Probability Concept #3: Adding a Constant to a Random Variable and Multiplying a Random Variable by a Constant

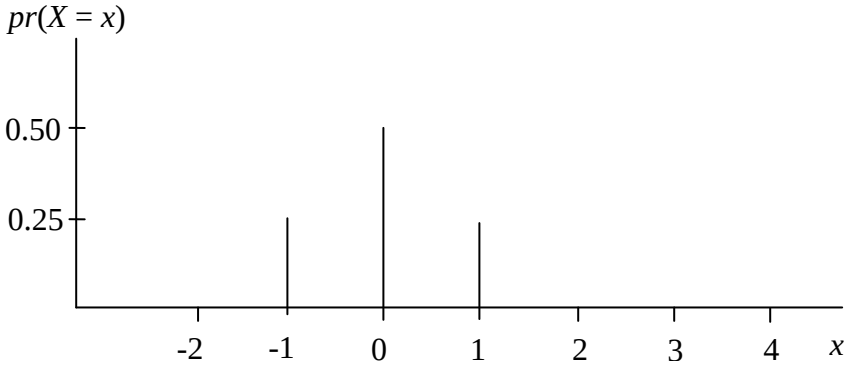
The concept of adding a constant to a random variable is an important one that is used in explaining and understanding regression. It is also used in computing probabilities under the normal distribution curve (see the next section "Normal Distribution"). The concept of multiplying a random variable by a constant is used in computing probabilities under the normal distribution curve.

The two concepts will be discussed in this section in a very simple context. The ideas will then be applied in contexts that are much more important. As discussed earlier, it is helpful to learn the basic concepts in as simple a context as possible so the ideas do not get lost because the context is complicated.

Adding a constant to a random variable

Suppose we have a random variable X that has a probability 0.25 of taking on the value -1, a probability 0.50 of taking on the value 0, and a probability 0.25 of taking on the value 1. The probability distribution is given below on the left in numerical form and on the right in graphical form.

x	$pr(X = x)$
-1	0.25
0	0.50
1	0.25



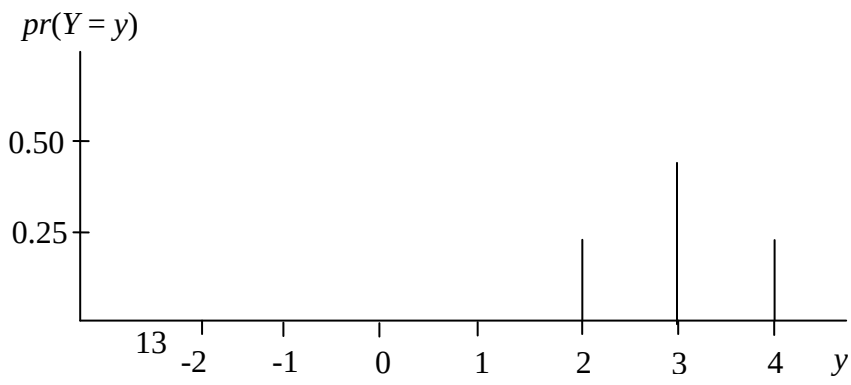
Now consider the random variable $Y = X + 3$. We want to find the probability distribution for Y. To obtain this distribution, note that Y takes on the value 2 if X takes on the value -1, and X takes on the value -1 with probability 0.25. Therefore, Y takes on the value 2 with probability 0.25. Graphically, the probability spike associated with $x = -1$ is shifted three units to the right and is now associated with $y = 2$.

Similarly, Y takes on the value 3 if X takes on the value 0, and X takes on the value 0 with probability 0.50. Therefore, Y takes on the value 3 with probability 0.50.

Finally, Y takes on the value 4 if X takes on the value 1, and X takes on the value 1 with probability 0.25. Therefore, Y takes on the value 4 with probability 0.25.

The distribution for Y is given below in numerical and graphical form.

x	$pr(X = x)$ and $pr(Y = y)$	$y = x + 3$
-1	0.25	2
0	0.50	3
1	0.25	4



The important point to take away from this example is that adding a positive constant to a random variable shifts the distribution to the right by the amount of the constant but does not change the spread of the distribution. In this example, the probabilities of 0.25, 0.50 and 0.25 associated with the *x*-values of -1, 0 and 1 are shifted three units to the right to the *y*-values of 2, 3 and 4. The mean (center) of the *X* distribution is also shifted three units to the right. Since the mean of *X* is 0 the mean of *Y* is 3.

Subtracting a positive constant from a random variable will shift the distribution to the left by the amount of the constant (be sure to understand why).

Multiplying a random variable by a constant

Consider the same distribution for the random variable *X* used in the previous section.

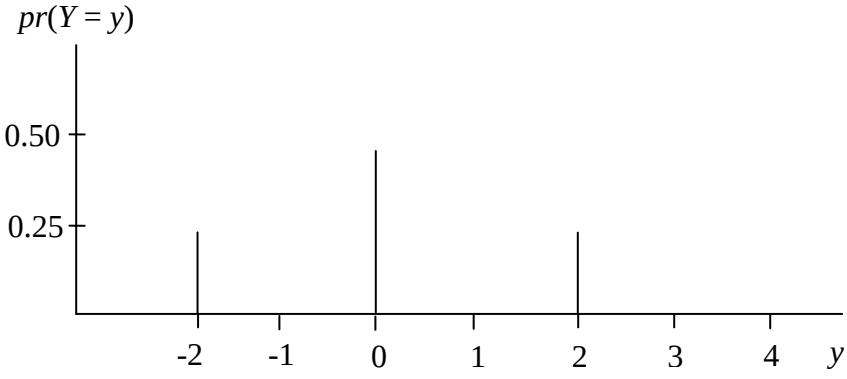
Now consider the new random variable $Y = 2X$. We want to find the probability distribution for *Y*. To obtain this distribution, note that *Y* takes on the value -2 if *X* takes on the value -1, and *X* takes on the value -1 with probability 0.25. Therefore, *Y* takes on the value -2 with probability 0.25. Graphically, the probability spike associated with $x = -1$ is now associated with $y = -2$.

Similarly, *Y* takes on the value 0 if *X* takes on the value 0, and *X* takes on the value 0 with probability 0.50. Therefore, *Y* takes on the value 0 with probability 0.50.

Finally, *Y* takes on the value 2 if *X* takes on the value 1, and *X* takes on the value 1 with probability 0.25. Therefore, *Y* takes on the value 2 with probability 0.25.

The distribution for *Y* is given below in numerical and graphical form.

x	$pr(X = x)$ and $pr(Y = y)$	$y = 2x$
-1	0.25	-2
0	0.50	0
1	0.25	2



The important point to take away from this example is that multiplying a random variable with a mean of zero by a constant greater than one increases the spread of the distribution but

does not change its center. In this example, the probabilities of 0.25, 0.50 and 0.25 associated with the *x*-values of -1, 0 and 1 are shifted to the (more spread out) *y*-values of -2, 0 and 2.

Multiplying a random variable with a mean of zero by a constant between 0 and 1 decreases the spread of the distribution but does not change its center (be sure to understand why). In this course, we will not be concerned with what happens when a random variable *X* is multiplied by a negative constant or when the random variable *X* has a mean different from zero.

Normal Distribution

The normal distribution is one of the most important distributions in statistics. It is the distribution we will use extensively throughout the semester.

The normal distribution is completely characterized by two parameters:

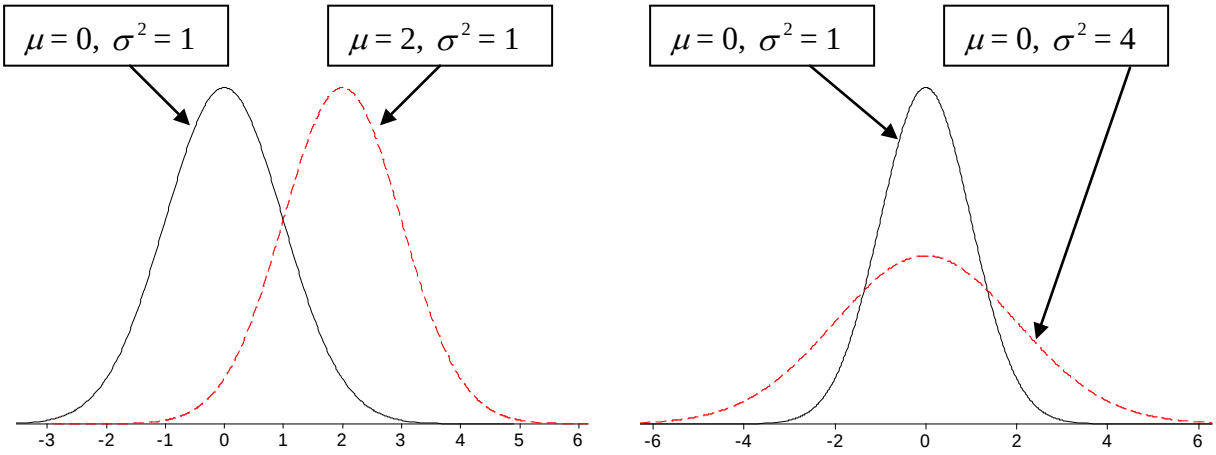
(1) The mean (μ), which is a measure of the center of the distribution; and

(2) The variance (σ^2), which is a measure of the spread of the distribution.

Graphical interpretation of the normal distribution parameters μ and σ^2

The normal distribution is a bell-shaped distribution. To get an intuitive feel for the meaning of its two parameters μ and σ^2, it is useful to look at the distribution graphically using the figures below. The figure on the left graphs two normal distributions with the same variance (i.e. same spread) but different means (i.e. different centers). The distribution graphed with a solid line is a normal distribution with mean $\mu = 0$ and variance $\sigma^2 = 1$. The distribution graphed with a dashed line is a normal distribution with mean $\mu = 2$ and variance $\sigma^2 = 1$. The important point to understand in comparing these distributions is that changing the mean from 0 to 2 shifts the distribution by two units to the right but does not change the spread of the distribution.

The figure on the right graphs two normal distributions with the same mean (i.e. same center) but different variances (i.e. different spread). The distribution graphed with a solid line is a normal distribution with mean $\mu = 0$ and variance $\sigma^2 = 1$. The distribution graphed with a dashed line is a normal distribution with mean $\mu = 0$ and variance $\sigma^2 = 4$. The important point to understand in comparing these two distributions is that changing the variance changes the spread of the distribution but does not affect the center (mean) of the distribution.



Notation for the normal distribution

The notation for a random variable X that has a normal distribution with mean μ and variance σ^2 is

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

The "~" means "is distributed as" and the "N" means "normal distribution". The first term in parentheses is always the mean and the second term is always the variance. Therefore, the notation "$X \sim N(\mu, \sigma^2)$" is short-hand for writing "the random variable X is normally distributed with mean μ and variance σ^2".

It is important to note that the convention we will always use in this class is that the second parameter is the variance, not the standard deviation. Since the variance and standard deviation represent the same information (i.e. providing one allows the other one to be determined) it doesn't matter which is used in the normal distribution notation but it is important to be consistent to remove any chance of misunderstanding.

Example

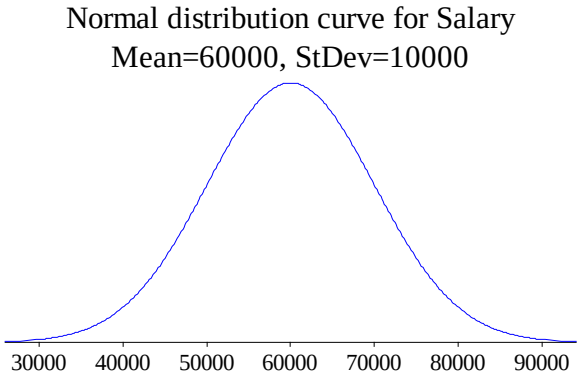
Consider the population of all MBA students in the U.S. who graduated last spring. This is a very large population. We will assume salaries in this population are normally distributed with mean \$60,000 and variance $\sigma^2 = (\$10{,}000)^2$, i.e. $Salary \sim N(60{,}000, (10{,}000)^2)$.

In practice, it is very important to check the assumption that salaries are normally distributed using a sample of salaries collected from the population (i.e. using a sample of data). We will

discuss this in detail later in the semester. In the current example, we will assume salaries are normally distributed to remove a level of complexity.

Also, the mean μ and variance σ^2 of a population will not typically be known. In practice, both parameters will have to be estimated using a sample of MBA salaries. This will be discussed in detail later in the semester (see the topic summary note "Estimation and Sampling Distributions"). Again, to remove complexity from the current example, we will assume the population mean μ = \$60,000 and population variance σ^2 = (\$10,000)^2 are known.

The graphical representation of the $N(60{,}000, (10{,}000)^2)$ distribution for the population of MBA salaries is given below.

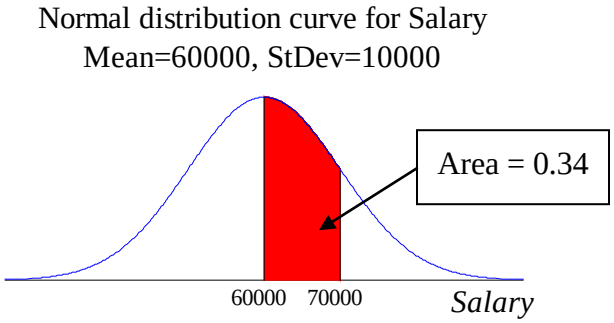


The interpretation of the population mean μ = \$60,000 is that it is the average (mean) salary of all spring MBA graduates. This value can be computed in principle by asking every spring MBA graduate in the U.S. what their current salary is and dividing the sum of their salaries by the number of graduates. In practice, this is not feasible and we will have to estimate the population mean from a sample of salary data.

The interpretation of the population variance σ^2 = (\$10,000)^2 in the context of a normal distribution is slightly more complicated and will be discussed later in this note after probability calculations for the normal distribution are discussed.

One of the most important concepts related to the normal distribution is that the area under the curve between two values represents probability. Calculating this area will be discussed in detail but first it is important to understand the interpretation of probability in the context of the MBA salary example. (Similar interpretations apply to other normal distributions but it is easiest to understand the concept in the context of a specific example.)

The area under the normal curve between 60,000 and 70,000 is 0.34. This area is straightforward to compute as shown later in this note but for the time being accept that the area is 0.34.



There are two equivalent interpretations of the area under the curve between 60,000 and 70,000. The first interpretation is that 34% of all MBA graduates in the population make between $60,000 and $70,000. The second interpretation is that if a single MBA graduate is selected at random from the population then there is a 34% chance (i.e. a probability of 0.34) that the person will make between $60,000 and $70,000. The equivalency makes sense intuitively because if 34 out of every 100 spring MBA graduates make between $60,000 and $70,000, there is a 34% chance that one of the people in this group will be selected.

Computing the area under the normal distribution curve

The area under a normal distribution curve is determined using a table of normal distribution probabilities. The process for using the normal distribution table is a purely mechanical process and is outlined below. This process will be used frequently throughout the semester in a variety of contexts. While the context will change the basic process of computing probabilities will always be the same.

Aside

You are <u>not</u> responsible for this aside. It is only included in case it helps motivate why tables are used to compute probabilities for the normal distribution.

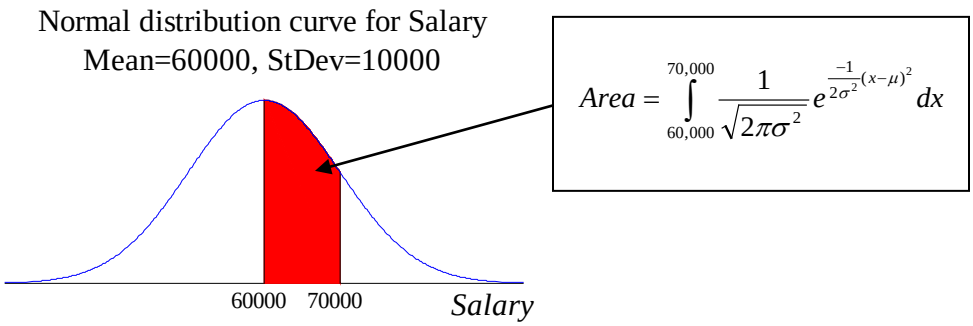
Mathematically, the normal curve is a function and from calculus (which you do not need to know for this course) the area under the curve is an integral. The function for the normal curve is

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

and the area under the curve between 60,000 and 70,000 is

$$Area = \int_{60,000}^{70,000} \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-1}{2\sigma^2}(x-\mu)^2} dx .$$

Graphically, the integral represents the shaded area below.



Unfortunately, this integral cannot be done analytically. Therefore, we have to resort to using a table to give (approximate) values. The probability values in the table are obtained using numerical integration. A numerical integration can be made extremely precise but it is still only an approximation.

In any case, using the table removes the need to deal with the integral above.

End Aside

Computing a probability for a random variable $X \sim N(\mu, \sigma^2)$, or equivalently, computing an area under the normal distribution curve, is a two-step process. Several examples are given below to illustrate the process.

Step #1: Convert the distribution from $X \sim N(\mu, \sigma^2)$ to $Z \sim N(0, 1)$. The distribution for Z is called the standard normal distribution because it has a mean of zero and a variance of one.

Step #2: Compute the probability for Z using the standard normal distribution table. As shown in the examples below, this will give you the probability of interest for the initial random variable X. The standard normal table is given at the end of this topic summary note.

The reason for converting from $X \sim N(\mu, \sigma^2)$ to $Z \sim N(0, 1)$ in step #1 is so only one table is required. Without the conversion step, tables would be needed for every possible combination of μ and σ^2 and this is not feasible.

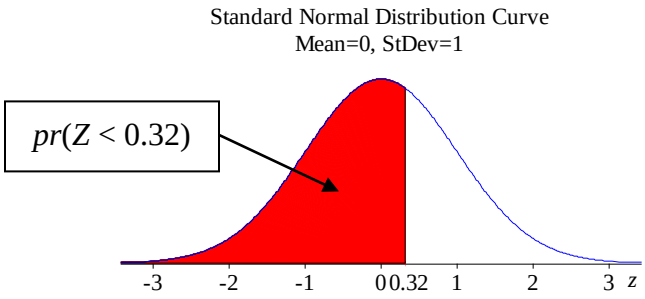
We will consider examples of step #2 first and then step #1.

Examples for step #2 (Using the standard normal distribution table)

Example #1: Compute $pr(Z < 0.32)$ where $Z \sim N(0, 1)$

My strong suggestion is to always draw a graph to represent the probability being calculated. It takes very little time to do and greatly reduces the chance of making a silly mistake.

Graphically, $pr(Z < 0.32)$ is represented by the area under the curve to the left of 0.32.



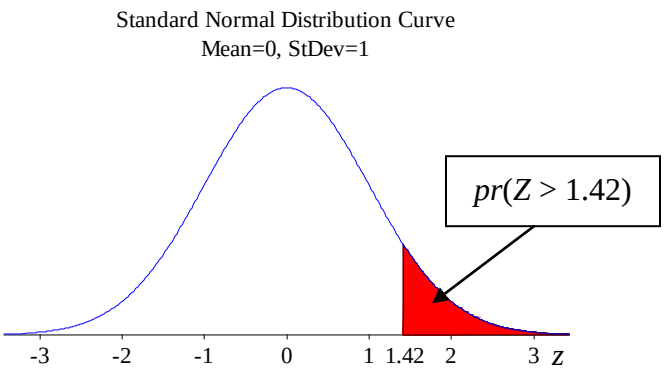
The standard normal distribution table gives the probability that the standard normal random variable Z is less than "little z" (i.e. $pr(Z < z)$) where "little z" is a specific number (for this problem, $z = 0.32$).

To find $pr(Z < 0.32)$, look down the left hand column in the standard normal table for 0.3 and then across the top row for .02. The intersection of this row and column gives $pr(Z < 0.32) = 0.6255$. There is an arrow pointing to this number in the table (at the end of this note).

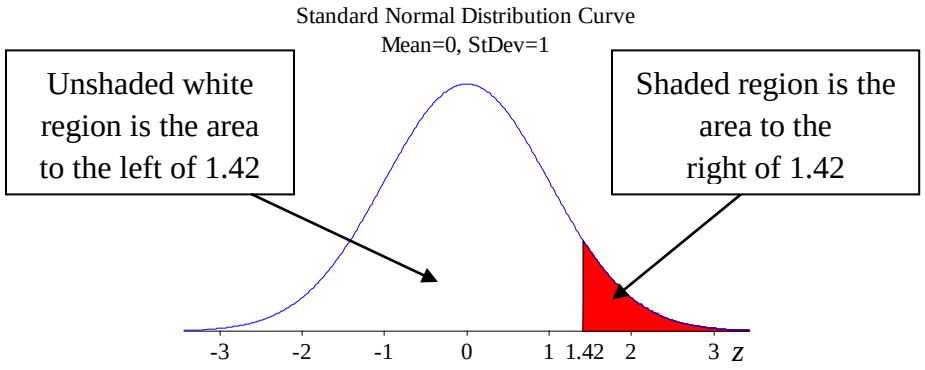
The convention used in the table is that the digits before and immediately after the decimal place (for this problem, 0.3) are given in the left hand column and the second digit following the decimal is given in the top row (for this problem, .02). Positive values of z are given on the first page of the table and negative values of z are given on the second page.

Example #2: Compute $pr(Z > 1.42)$ where $Z \sim N(0, 1)$

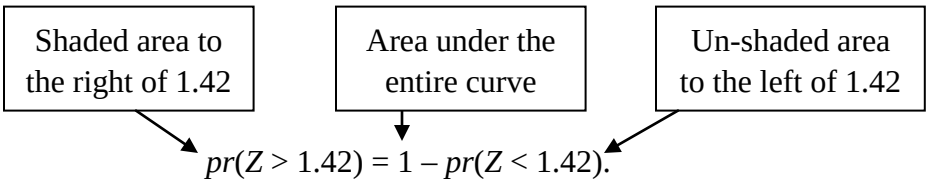
Graphically, $pr(Z > 1.42)$ is represented by the shaded area under the curve to the right of 1.42.



The standard normal table only gives areas to the left of a value. Graphically, to compute the area to the right of 1.42, we take the area under the entire curve (which is one) and subtract off the area we don't want (the area to the left of 1.42, which is represented by the white un-shaded area) to leave the area that we want (the shaded area to the right of 1.42).



Therefore,



We can look 1.42 up in the standard normal table to give $pr(Z < 1.42) = 0.9222$ (there is an arrow pointing to this number in the table). Using this value gives

$$pr(Z > 1.42) = 1 - pr(Z < 1.42)$$

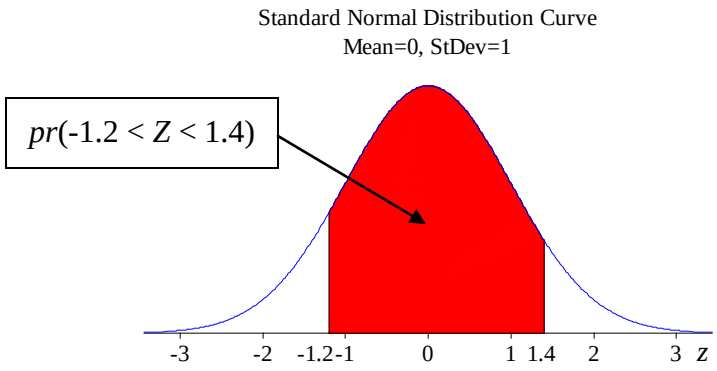
$$= 1 - 0.9222$$

$$= 0.0778.$$

Example #3: Compute $pr(-1.2 < Z < 1.4)$ where $Z \sim N(0, 1)$

We will often need to compute interval probabilities such as $pr(-1.2 < Z < 1.4)$ during the semester. For example, if we are interested in predicting next quarter's sales for a company, interval probabilities will be used to quantify how accurate the prediction is likely to be.

Graphically, $pr(-1.2 < Z < 1.4)$ is represented by the shaded area under the curve between -1.2 and 1.4.



To compute the shaded area, we take the total area to the left of 1.4 (this is the shaded area we are interested in plus the un-shaded white area to the left of -1.2) and subtract off the un-shaded white area to the left of -1.2. Therefore,

Shaded area between -1.2 and 1.4

Total area to the left of 1.4

Un-shaded area to the left of -1.2

$$pr(-1.2 < Z < 1.40) = pr(Z < 1.4) - pr(Z < -1.2)$$

$$= 0.9192 - 0.1151$$

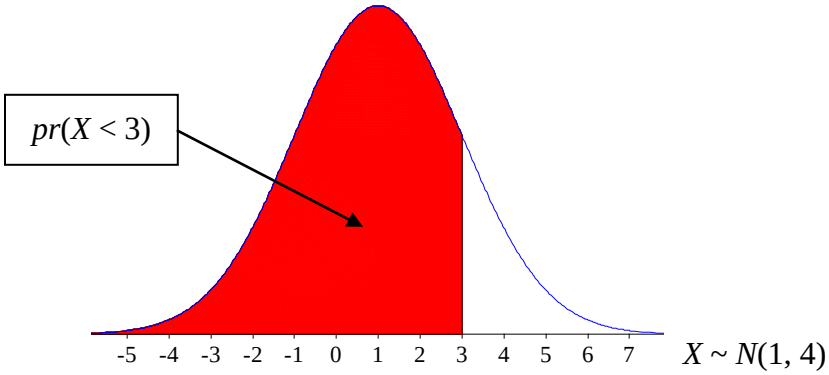
$$= 0.8041.$$

$pr(Z < 1.4) = 0.9192$ and $pr(Z < -1.2) = 0.1151$ are read off the standard normal table – there are arrows pointing to these numbers in the table (note that $pr(Z < -1.2)$ is given on the second page of the table).

<u>Examples for step #1</u> (Converting from $X \sim N(\mu, \sigma^2)$ to $Z \sim N(0, 1)$)

Example #4: Compute $pr(X < 3)$ where $X \sim N(\mu = 1, \sigma^2 = 4 = (2)^2)$

Graphically, $pr(X < 3)$ is represented by the shaded area to the left of 3 under the $N(1, 4)$ curve.



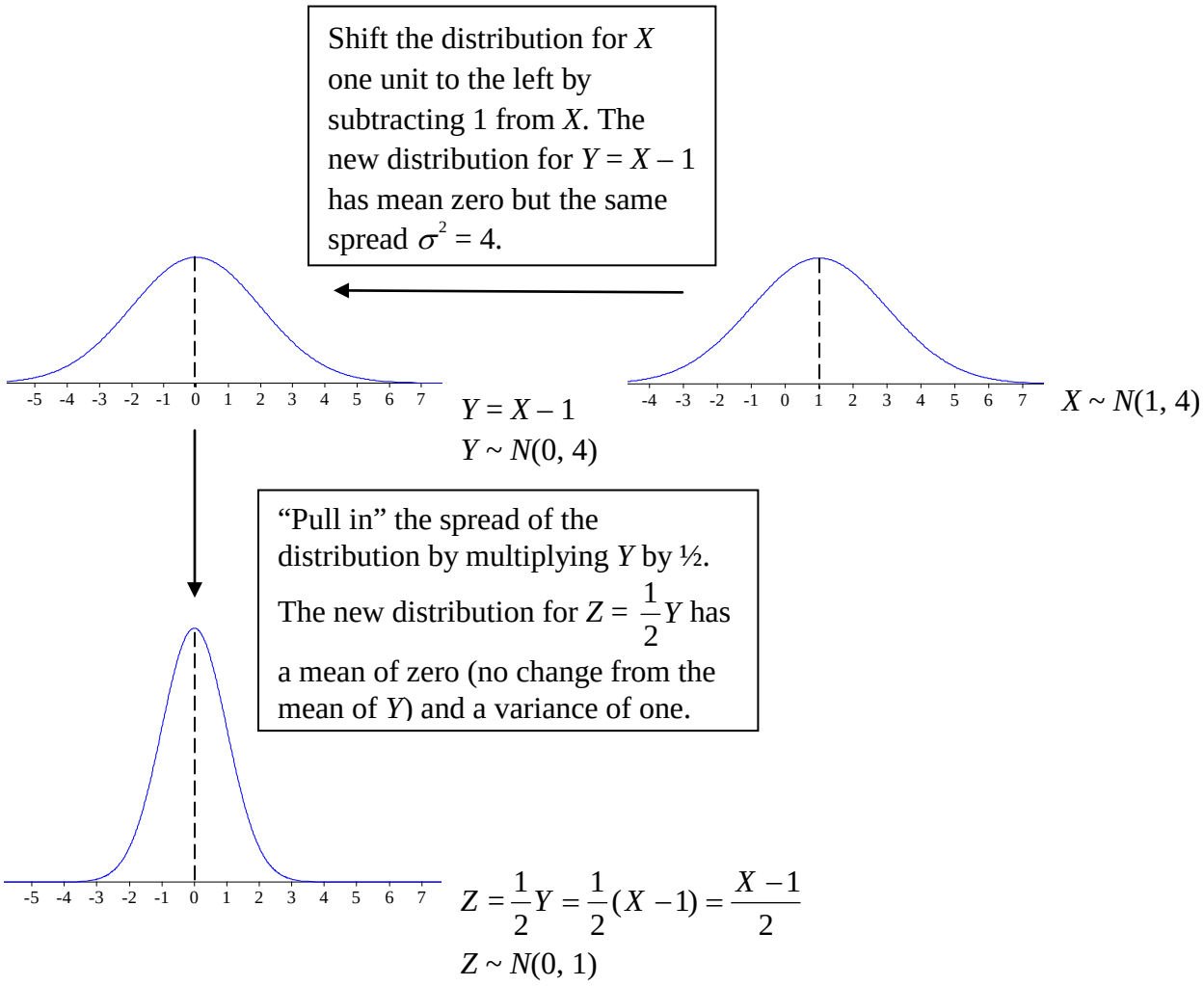
To compute this area (probability) we need to "standardize" the random variable X so that the resulting random variable Z has a $N(0, 1)$ distribution, and then look up the appropriate probability for Z in the standard normal table.

The standardization step is done in two stages. We first create the random variable $Y = X - 1$. Subtracting one shifts the distribution for X by one unit to the left but does not change the spread of the distribution (recall the discussion of "Adding a constant to a random variable" earlier in these notes). The resulting distribution for Y has a mean of zero because the mean, or center, of the distribution for X is one and the entire distribution has been shifted one unit to the left. This is shown graphically in the figure on the next page.

The variance for Y (see the distribution on the top left in the figure on the next page) is four so the distribution is "too spread out" (we want a variance of one). Recall from the discussion "Multiplying a random variable by a constant" that if we multiple a random variable with a mean of zero by a constant less than one that the distribution becomes "less spread out." We will create the new random variable Z by multiplying Y by ½, i.e.

$$Z = \frac{1}{2}Y = \frac{1}{2}(X - 1) = \frac{X - 1}{2}.$$

This decreases the spread by the appropriate amount so the resulting random variable Z has a variance of one (see the distribution in the bottom left in the figure on the next page). It is not intuitively obvious why we multiply by the specific number ½. However, it should be clear intuitively (based on the discussion in "Multiplying a random variable by a constant") why we multiply by a number less than one to reduce the spread of the distribution. You can just accept that the specifiec value required to obtain a variance of one for Z in this problem is $\frac{1}{2}$ (in general, it is $\frac{1}{\sigma}$).



The new random variable $Z = \frac{X-1}{2}$ has a $N(0, 1)$ distribution and we can look up probabilities involving Z in the standard normal table.

The specific steps to compute $pr(X < 3)$ are:

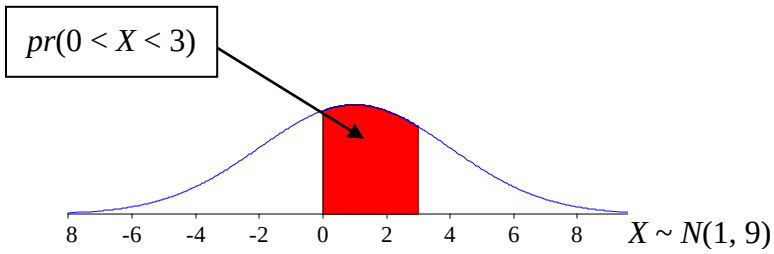
$pr(X < 3)$	Original probability to compute
$= pr\left(\frac{X-1}{2} < \frac{3-1}{2}\right)$	Step #1: Convert from $X \sim N(1, 4)$ to $Z \sim N(0, 1)$
$= pr(Z < 1)$	Rewrite $\frac{X-1}{2}$ as Z and $\frac{3-1}{2}$ as 1
$= 0.8413$	Step #2: Use the standard normal table to look up the appropriate probability for Z

The conversion step (step #1) requires that we subtract 1 from X and divide by 2 on the left of the inequality sign to give $Z = \frac{X-1}{2} \sim N(0, 1)$. Whatever is done on the left side of the inequality sign must also be done on the right side. Subtracting 1 and dividing by 2 on the right gives $\frac{3-1}{2} = 1$.

The important point to take away from this example is that to compute a probability such as $pr(X < 3)$ we first convert $X \sim N(1, 4)$ to the standard normal random variable $Z \sim N(0, 1)$, and then look up the appropriate probability for Z in the standard normal table. Because we keep equality at each line of the calculation, the probability $pr(Z < 1)$ computed for Z is equal to the probability $pr(X < 3)$ required for X.

Example #5: Compute $pr(0 < X < 3)$ where $X \sim N(\mu = 1, \sigma^2 = 9 = (3)^2)$

Graphically, $pr(0 < X < 3)$ is represented by the shaded area between 0 and 3.



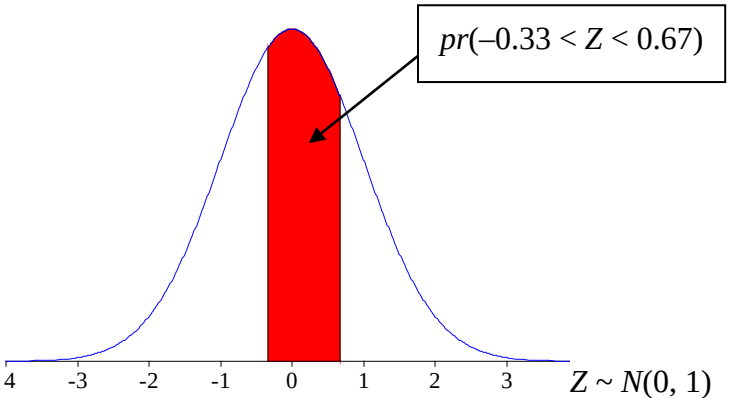
To compute this probability, $X \sim N(1, 9 = (3)^2)$ must be converted to $Z \sim N(0, 1)$ and then the standard normal tables are used to compute the probability.

In terms of the probability calculation,

$pr(0 < X < 3)$	Original probability to compute
$= pr\left(\frac{0-1}{3} < \frac{X-1}{3} < \frac{3-1}{3}\right)$	Step #1: Convert from $X \sim N(1, 9 = (3)^2)$ to $Z \sim N(0, 1)$
$= pr(-0.33 < Z < 0.67)$	Rewrite $\frac{0-1}{3}$ as -0.33, $\frac{X-1}{3}$ as Z, and $\frac{3-1}{3}$ as 0.67.

The conversion step requires that we subtract $\mu = 1$ from X and divide by $\sigma = 3$ in the center of the two inequality signs to give $Z = \frac{X - 1}{3} \sim N(0, 1)$. Whatever is done in the center must also be done on both ends of the inequality. Subtracting $\mu = 1$ and dividing by $\sigma = 3$ on both ends of the inequality gives $\frac{0 - 1}{3} = -0.33$ and $\frac{3 - 1}{3} = 0.67$.

Graphically, the shaded area in the figure below represents $pr(-0.33 < Z < 0.67)$. To compute the shaded area, we take the total area to the left of 0.67 (this is the shaded area we are interested in plus the un-shaded white area to the left of –0.33) and subtract off the un-shaded white area to the left of –0.33.



Therefore,

Shaded area between -0.33 and 0.67

Total area to the left of 0.67

Un-shaded area to the left of -0.33

$pr(-0.33 < Z < 0.67) = pr(Z < 0.67) - pr(Z < -0.33)$

$= 0.7486 - 0.3707$

$= 0.3779.$

Step #2: Use the standard normal table to look up appropriate probabilities for Z

<u>General formula for converting from $X \sim N(\mu, \sigma^2)$ to $Z \sim N(0, 1)$</u>

If $X \sim N(\mu, \sigma^2)$, where μ and σ^2 are specific numbers, we convert X to Z by first subtracting off the mean μ to give $Y = X - \mu$. In the above problem, this corresponds to subtracting off $\mu = 1$. The random variable Y has a mean of zero because the distribution for X is shifted by μ units.

We then divide Y by σ (or equivalently, multiply by $\frac{1}{\sigma}$) to give

$$Z = \frac{1}{\sigma}Y = \frac{1}{\sigma}(X - \mu) = \frac{X - \mu}{\sigma}.$$

In the above problem, this corresponds to dividing by $\sigma = 3$ (or equivalently, multiply by $\frac{1}{\sigma} = \frac{1}{3}$.)

Note that if $\sigma > 1$ (i.e. the variance for X is greater than one), then the spread is too large. Dividing by a σ value greater than one is equivalent to multiplying by a number less than one, with the result that the variance of the new random variable Z has a smaller variance than X (in particular, the variance of Z is one).

Similarly, if $\sigma < 1$ (i.e. the variance for X is less than one), then the spread is too small. Dividing by a σ value less than one is equivalent to multiplying by a number greater than one, with the result that the variance of the new random variable Z has a larger variance than X (in particular, the variance is one).

It can be a bit confusing to think about the conversion step in general terms. To get a good understanding of the conversion, try a couple of specific combinations of μ and σ (for example, try standardizing a $N(3, 4)$ distribution and a $N(2, 0.09)$ distribution).

To summarize, the formula

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

is the general formula for converting from $X \sim N(\mu, \sigma^2)$ to $Z \sim N(0, 1)$. It will work for any combination of μ and σ.

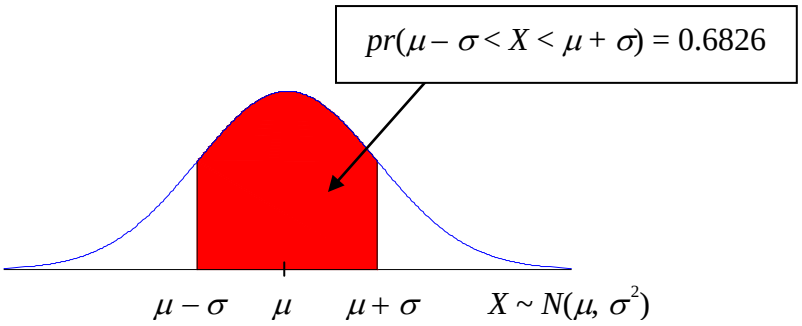
Interpreting σ in the context of a normal distribution

Suppose the random variable X has a $N(\mu, \sigma^2)$ distribution. Then (as shown below), there is a 68% chance X will fall within one standard deviation σ of the mean μ. Similarly, there is a 95% chance X will fall within two standard deviations of the mean.

For example, in the MBA salary example where X represents *Salary* and $X \sim N(60{,}000, (10{,}000)^2)$, 68% of all spring MBA graduates in the population make between $\mu - \sigma$ = \$60,000 – \$10,000 = \$50,000 and $\mu + \sigma$ = \$60,000 + \$10,000 = \$70,000. Or equivalently, if we pick an MBA graduate at random from the population there is a 68% chance this person will make between \$50,000 and \$70,000.

Similarly, 95% of all spring MBA graduates make between $\mu - 2\sigma$ = \$60,000 – 2(\$10,000) = \$40,000 and $\mu + 2\sigma$ = \$60,000 + 2(\$10,000) = \$80,000. Or equivalently, if we pick an MBA graduate at random from the population there is a 95% chance this person will make between \$40,000 and \$80,000.

Graphically, 68% of the area under the normal distribution curve falls between $\mu - \sigma$ and $\mu + \sigma$. (To be more precise, 68.26% of the area under the normal distribution curve falls between $\mu - \sigma$ and $\mu + \sigma$.)



In terms of the probability calculation,

$pr(\mu - \sigma < X < \mu + \sigma)$	Original probability to compute
$= pr\left(\frac{(\mu - \sigma) - \mu}{\sigma} < \frac{X - \mu}{\sigma} < \frac{(\mu + \sigma) - \mu}{\sigma}\right)$	Step #1: Convert from $X \sim N(\mu, \sigma^2)$ to $Z \sim N(0,1)$
$= pr(-1 < Z < 1)$	Rewrite $\frac{(\mu - \sigma) - \mu}{\sigma}$ as –1, $\frac{X - \mu}{\sigma}$ as Z, and $\frac{(\mu + \sigma) - \mu}{\sigma}$ as 1

Note that this probability calculation holds whatever the values of μ and σ are in a given situation. (The calculation of $pr(-1 < Z < 1)$ is completed below.)

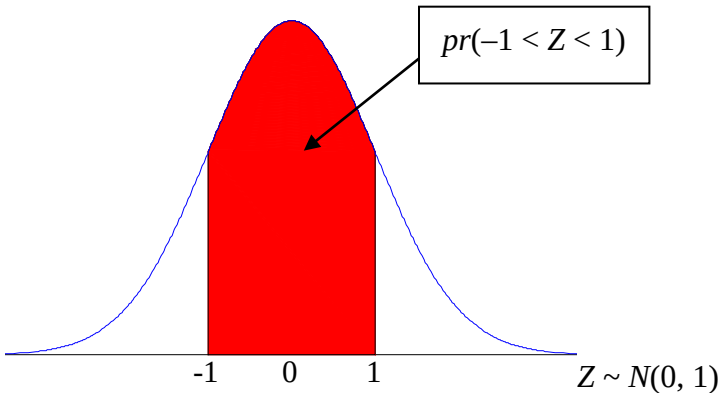
The conversion step requires that we subtract μ from X and divide by σ in the center of the two inequality signs to give $Z = \frac{X - \mu}{\sigma} \sim N(0, 1)$. Whatever is done in the center must also be done on both ends of the inequality. When we subtract μ and divide by σ on both ends of the inequality they cancel out to give $\frac{(\mu - \sigma) - \mu}{\sigma} = \frac{-\sigma}{\sigma} = -1$ and $\frac{(\mu + \sigma) - \mu}{\sigma} = \frac{\sigma}{\sigma} = 1$.

Aside

Anytime you are asked to compute a probability where μ and σ are not assigned specific numerical values (as in this situation) both μ and σ must cancel out of the calculation. If they don't this means a mistake was made somewhere in the calculation.

End Aside

To complete the probability calculation above we need to compute $pr(-1 < Z < 1)$. Graphically, the shaded area in the figure below represents this probability To compute the shaded area, we take the total area to the left of 1.0 (this is the shaded area we are interested in plus the un-shaded white area to the left of -1.0) and subtract off the un-shaded white area to the left of –1.0.



Therefore,

Shaded area between -1.0 and 1.0

Total area to the left of 1.0

Un-shaded area to the left of -1.0

$$pr(-1.0 < Z < 1.0) = pr(Z < 1.0) - pr(Z < -1.0)$$

$$= 0.8413 - 0.1587$$

$$= 0.6826.$$

Step #2: Use the standard normal table to look up appropriate probabilities for Z

The 68% and 95% intervals of $(\mu - \sigma, \mu + \sigma)$ and $(\mu - 2\sigma, \mu + 2\sigma)$, respectively, are worth memorizing as they will come up regularly during the semester.

Be sure to do the calculation that shows 95% of the area under the $N(\mu, \sigma^2)$ curve falls within two standard deviations of the mean. (To be more precise, 95% of the area under the $N(\mu, \sigma^2)$

curve falls within 1.96 standard deviations of the mean. However, in order to use round numbers we will use the statement "95% of the area under the $N(\mu, \sigma^2)$ curve falls within two standard deviations of the mean.")

Aside

For any probability calculation involving the normal distribution when the mean and variance are specific numerical values it is possible to compute areas under the normal distribution curve using a formula in Excel. For example, Excel can be used to compute $pr(X < 3)$ in Example #4 where $X \sim N(\mu = 1, \sigma^2 = (2)^2 = 4)$. If you type the formula "=NORMDIST(3,1,2,TRUE)" into an Excel cell the resulting value will be $pr(X < 3) = 0.8413$.

You can use Excel to compute normal distribution probabilities in homework problems if you want to but I recommend against it for three reasons. First, since laptops will not be used on the exam (for reasons discussed in class) you will need to use the standard normal table on the exam to compute probabilities.

Second, you cannot compute probabilities of the type

$$pr(\mu - 1.5\sigma < X < \mu + 1.5\sigma)$$

where $X \sim N(\mu, \sigma^2)$ and μ and σ are not specified numerical values. The reason is that Excel requires specific numbers to be input into the NORMDIST formula. These types of problems will occur during the semester and arise in real-world problems (for example, in computing confidence intervals other than 68% and 95% intervals).

Finally, in my opinion, it is easier to use the table than to start up Excel and type in the NORMDIST formula, although this is a matter of personal preference that may vary across people.

End Aside

Standard Normal Distribution Table

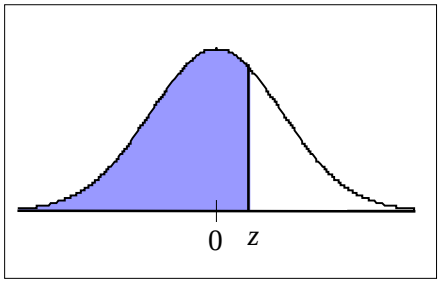


Table entry is the probability at or below *z*

Pr(*Z* < 0.32) for Example #1

Pr(*Z* < 1.42) for Example #2

Pr(*Z* < 1.4) for Example #3

Standard normal distribution probabilities for positive values of *z*

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.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
	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
	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
	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
	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
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998

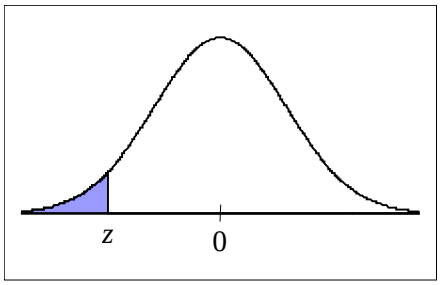


Table entry is the probability at or below *z*

Standard normal distribution probabilities for negative values of *z*

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
-3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
-3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
-3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
-3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
-2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
-2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
-2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
-2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
-2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
-2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
-1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
-1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
-1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
-1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
-1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
-1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
-1.2	**0.1151**	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
-1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
-1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
[illegible]	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
[illegible]	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
[illegible]	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
-0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
-0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
-0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
-0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
-0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
-0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
-0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641

Pr(*Z* < -1.2) for Example #3