Chapter 2 § 2

Standard Normal Calculations

Definitions:

Standardizing – Changing the units in a normal distribution of size σ and a mean μ as center.

Z – Scores – Standardized observation from symmetric distributions to be expressed in a common scale

Standardized observation – describes how many standard deviations the original observation falls away from the mean and in which direction

Standard normal distribution – a normal distribution N (0, 1) with a mean 0 and a standard deviation of 1.

Normal probability plot – A method for assessing normality.

Formula:

Z-Score

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

Example:

The height of young women in Collierville High School is N (64.5, 2.5)

$$Z = \frac{\text{height} - 64.5}{2.5}$$

A woman's standardized height is the number of standard deviations by which her height is differs from the mean height of all young women. A woman 68 inches tall, for example, has standardized height

$$Z = \frac{68 - 64.5}{2.5} = 1.4$$

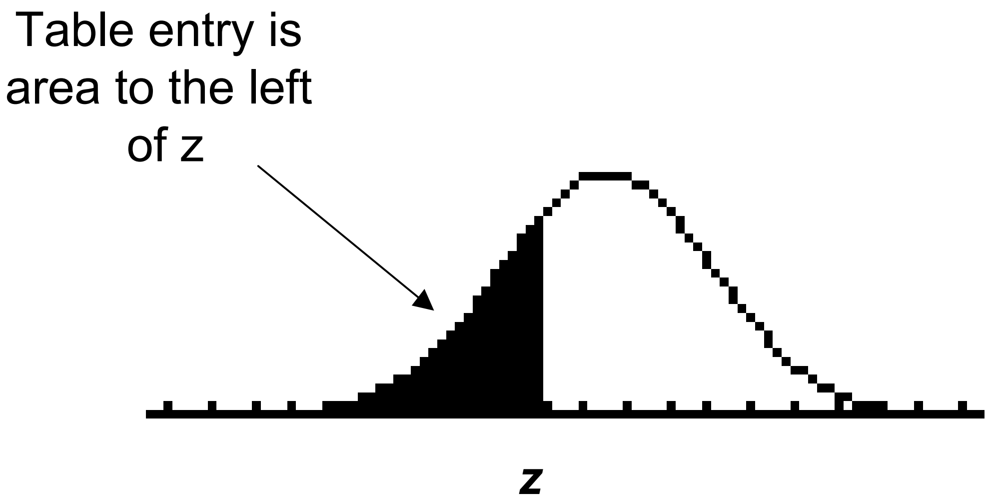
1.4 standard deviations above the mean.

Similarly, a woman 60 inches tall has a standard deviation of

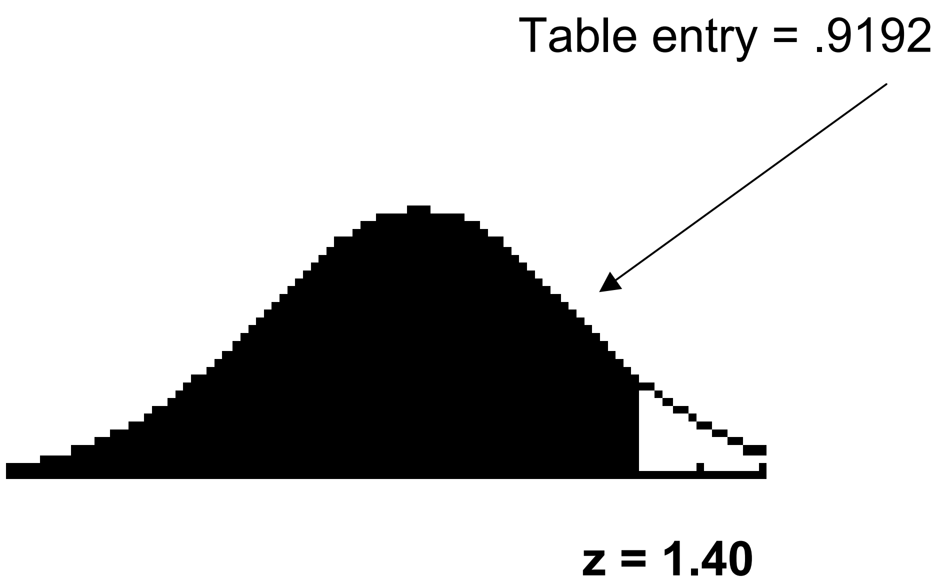
$$Z = \frac{60 - 64.5}{2.5} = -1.8$$

1.8 standard deviations below the mean.

Table A (page 834) is a table of areas under the standard normal curve. The table entry for each value **z** is the area under the curve to the left of ***z***.

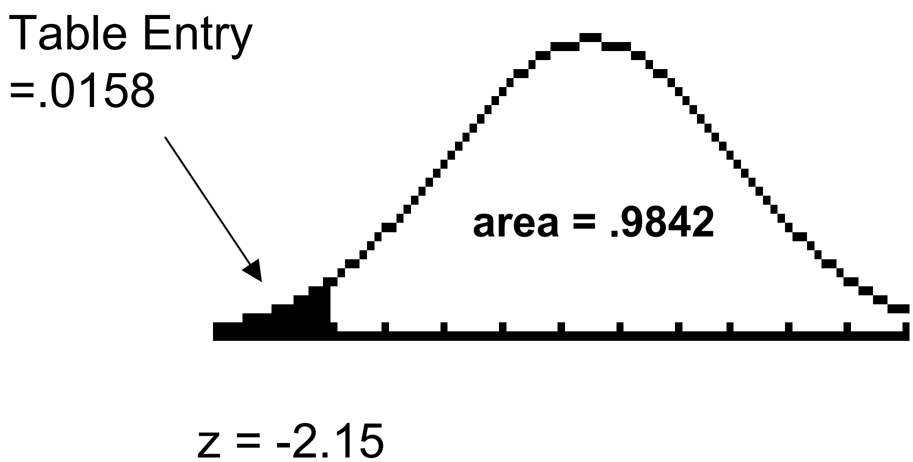


Find the proportion of observations from the standard normal distribution that are les than 1.4.



≈ 92%

Find the proportion of observations from the standard normal distribution that are greater than -2.15.



Thus the area above the z – score of -2.15 is ≈ 98%.

Note:

Area under the density curve is a proportion of the observation in a distribution. Thus by finding an area under the curve for a range of values, the proportion of observations can be found.

4 Steps for Finding Normal Proportions

State the problem in terms of the observed variable x.

Standardize x to restate the problem in terms of a standard normal variable z.

Draw a picture to show the area under the standard normal curve.

Find the required area under the standard normal curve, using Table A and the fact that the total area under the curve is 1.

The level of cholesterol in the blood for 14-year-old boys has a mean of 170 milligrams of cholesterol per deciliter of blood and a standard deviation of 30 mg/dl. What percent of 14-year-old boys have more than 240mg/dl of cholesterol?

State the problem:

Call the level of cholesterol in the blood x. The variable x has the N (170,30) distribution. We want to find the proportion of boys with x > 240.

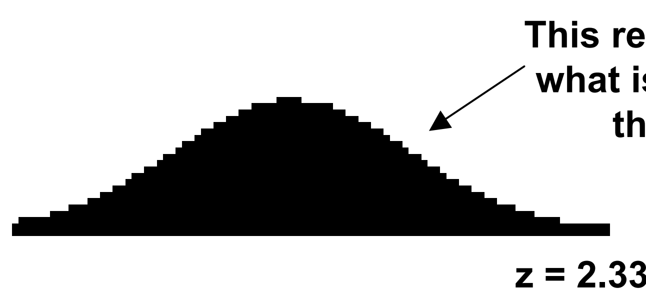
Standardize x:

Subtract the mean, then divide by the standard deviation, to turn into a standard normal z:

$$\frac{x - 170}{30} > \frac{240 - 170}{30}$$

$$z > 2.33$$

Draw a Picture:



Use the Table:

From table A, we see that the proportion of observations less than 2.33 is .9901. About 99% of the boys have cholesterol levels less than 240. The area to the right of 2.33 is 1 - .9901 = .0099. This is about 1%. Only 1% of the boys have high blood pressure.

What percent of 14-year-old boys have blood cholesterol between 150 and 210 mg/dl?

State the problem:

Call the level of cholesterol in the blood x. We want the proportion of boys with $150 \le x \le 210$.

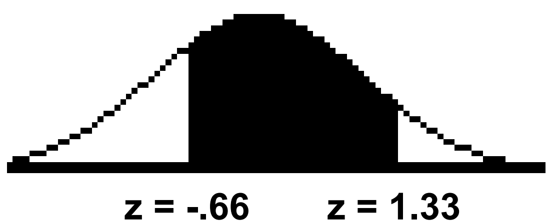
Standardize x:

$$150 \le x \le 210$$

$$\frac{150 - 170}{30} \qquad \frac{210 - 170}{30}$$

$$-.667 \le x \le 1.334$$

Draw a Picture:



Use the Table:

From table A, we see that the proportion of observations less than 1.33 is ≈ .9082 and the proportion of observations less than -.66 is ≈ .2546.

Area between -.667 and 1.334
= area below 1.334 – area below -.667
= .9082 - .2546 = .6536

About 65% of boys have a cholesterol levels between 150 to 210 mg/dl.

Locate the point on a normal curve that corresponds to the top 10% of cholesterol levels for 14-yearold boys.

State the problem:

We want to find the score x with area .1 to its right under the normal curve.

Use the Table:

Looking at Table A for the entry closest to .9, we find .8997. This entry corresponds to z = 1.28. So 1.28 is the standardized value with area .9 to its left.

Unstandardize x:

To transform the solution from the z back to the original x scale. We know that the standardized value of the unknown is z = 1.28. So x itself satisfies

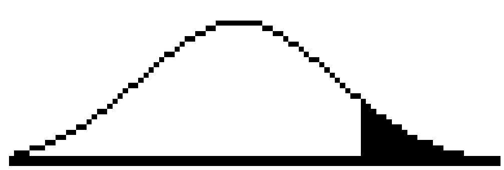
$$\frac{x - 170}{30} = 1.28$$

Solving this equation for x gives

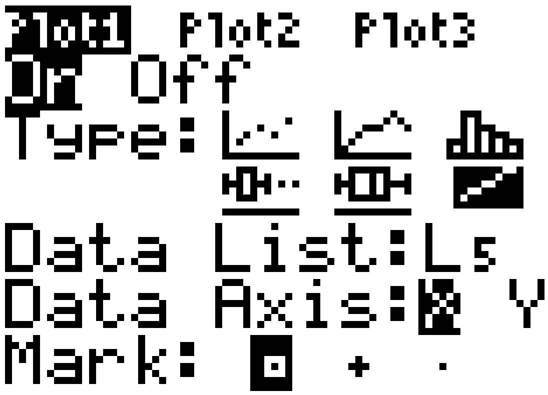
$$x = 170 + (30)(1.28) = 208.4$$

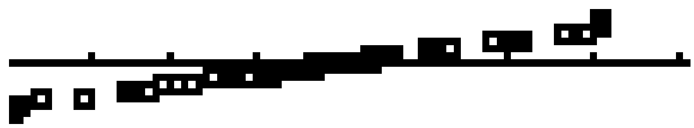
Thus the "unstandardized" meaning of $z \approx 1.28$. We see that a 14-year-old boy must have a cholesterol level of at least 208.4 to place in the highest 10%.

Draw a Picture:



Given a set a data we can run a **normal probability plot** to determine if the data distribution is close to a normal distribution.





If the data distribution is close to a normal distribution, the plotted points will lie close to a straight line. Conversely, nonnormal data will show a nonlinear trend. Outliers appear as points that are far away from the overall pattern of the plot. Since the above plot is quite linear, our conclusion is that it is reasonable to believe that the data are from a normal distribution.