

5.1 Standardizing with z-scores

Bob got an 80 on a Math 135 test and Betty got a 92 on Math 117 test. Who did better? You can look at the raw scores or determine a relative score based on the each did compared to the rest of their classmates.

The Math 135 test had a mean of 72 with a standard deviation of 4.
The Math 117 test had a mean of 82 with a standard deviation of 6.

Scores are standardized by using the **z-score**. The z-score measures the distance of a score to the mean in standard deviations. Therefore z-scores have no units. The z-score is computed using the following formula:

$$z = \frac{x - \bar{x}}{s} = \frac{\text{Score} - \text{mean}}{\text{Standard Deviation}}$$

Bob

$$z = \frac{80 - 72}{4} = \frac{8}{4} = 2$$

Betty

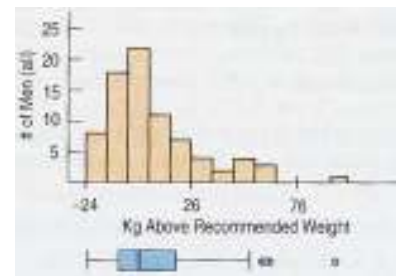
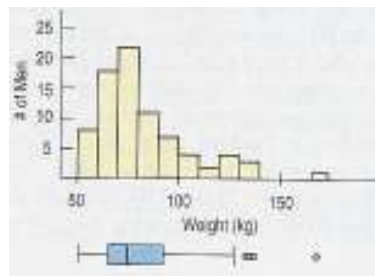
$$z = \frac{92 - 82}{6} = \frac{10}{6} \approx 1.67$$

Bob has a higher relative score since his z-score is higher.

5.2 Shifting and Scaling

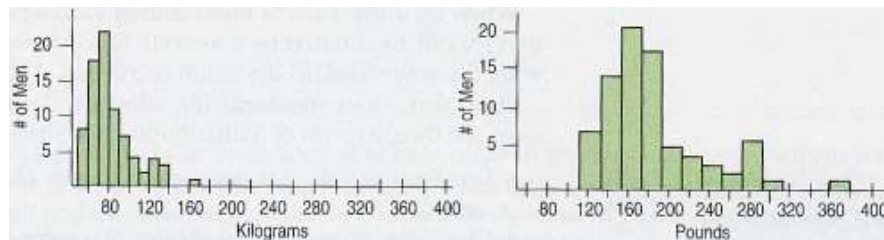
Shift data by adding or subtracting a constant to each data value. The measures of positions (center, quartiles, minimum, and maximum) change by the same constant value, but the spread will remain unchanged.

Fig 5.2 and 5.3
p. 113.



Rescale data by multiplying or dividing a constant to each data value. All measures of position and spread will be multiplied or divided by that same value.

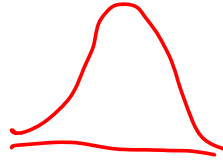
Fig 5.4
p. 114



Standardizing data into z-scores is just shifting the values by the mean and rescaling by the standard deviation.

- Standardizing into z-scores does not change the SHAPE of the distribution of a variable.
- Standardizing into z-scores changes the CENTER by making the mean 0.
- Standardizing into z-scores changes the SPREAD by making the standard deviation 1.

5.3 Normal Models



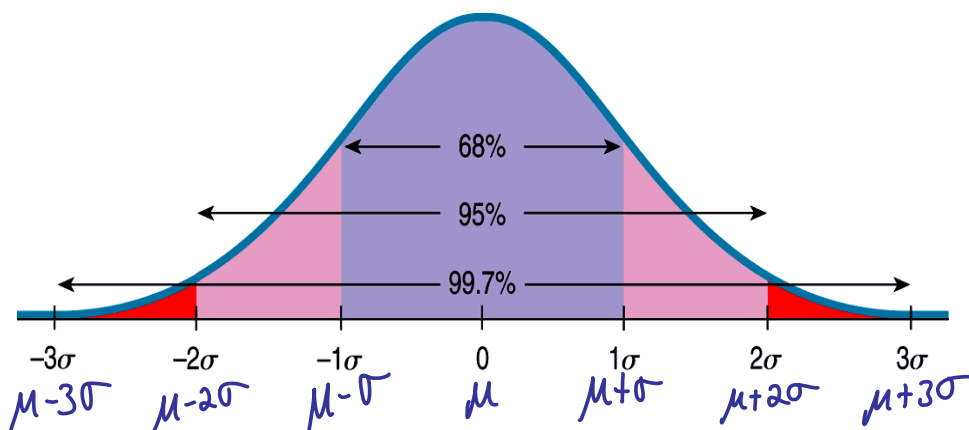
Normal Models (p. 117) are statistical models of distributions whose shapes are bell-shaped. These models will be unimodal and roughly symmetric. These distributions can be used to determine how extreme a z-score is.

Notation: $N(\mu, \sigma)$ a Normal distribution with mean of μ and a standard deviation of σ . This mean and standard deviation are NOT \bar{x} and s . They do not come from data. They are used to help specify a model and are called **parameters** of the model. In statistic, Greek letters are used to represent parameters. \bar{x} and s are summaries of data and are called **statistics**.

The z-score for a Normal model is $z = \frac{x - \mu}{\sigma}$. Then we will only need the model $N(0,1)$ which is called the standard Normal model or standard Normal distribution.

Since standardizing does not change the shape of the distribution the standard Normal distribution should only be used for data that is unimodal and symmetric. This can be checked by making a histogram. (Nearly Normal Condition)

68-95-99.7 Rule (p. 118) - In a Normal model, APPROXIMATELY 68% of the values fall within 1 standard deviation of the mean, 95% of the values fall within 2 standard deviation of the mean, and 99.7% of the values fall within 3 standard deviation of the mean.

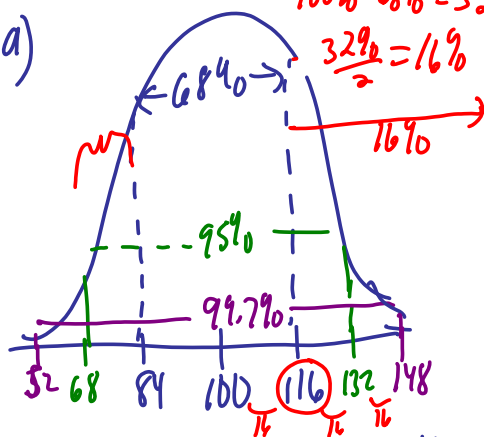


Example: p. 133, #8

$$N(100, 16)$$

Normal model with $\mu=100$ & $\sigma=16$

a)



$$68\%: \mu - \sigma = 100 - 16 = 84$$

$$\mu + \sigma = 100 + 16 = 116$$

$$95\%: \mu - 2\sigma = 100 - 2(16) = 68$$

$$\mu + 2\sigma = 100 + 2(16) = 132$$

$$99.7\%: \mu - 3\sigma = 100 - 3(16) = 52$$

$$\mu + 3\sigma = 100 + 3(16) = 148$$

b) Middle 95% are between 68 and 132

c) About 16% are above 116

d) $95\% - 68\% = 27\%$ $\frac{27\%}{2} = 13.5\%$ About 13.5% are between 68 and 84

e) $100\% - 95\% = 5\%$
 $\frac{5\%}{2} = 2.5\%$ About 2.5% are above 132.

5.4 Finding Normal Percentiles

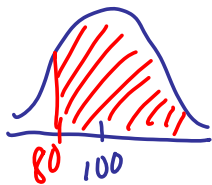
Distr.

Normalcdf(Left Endpt, Right Endpt, μ, σ) InvNorm(Percent to Left, μ, σ)

p. 133, # 10

$$N(100, 16) \quad \mu=100 \quad \sigma=16$$

a) over 80

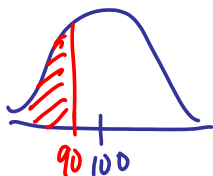


$$\text{Normalcdf}(80, 100000, 100, 16)$$

.8944

89.44% have IQ scores over 80.

b) under 90



$$\text{Normalcdf}(-100000, 90, 100, 16)$$

.2660

26.6% have IQ scores under 90

c) between 112 & 132



$$\text{Normalcdf}(112, 132, 100, 16)$$

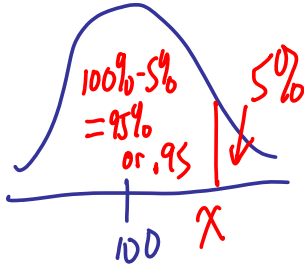
20.39% have IQ scores between 112 and 132.

p. 137, #44

$$N(100, 16)$$

$$\mu = 100 \quad \sigma = 16$$

a) Bound for Highest 5%

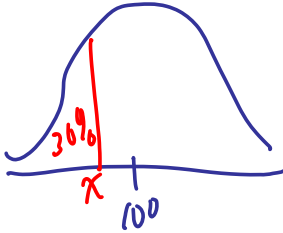


$$\text{INVNorm}(.95, 100, 16)$$

$$x = 126.3$$

The upper 5% are above 126.3

b) Bound for Lowest 30%

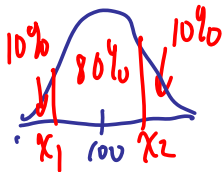


$$\text{INVNorm}(.3, 100, 16)$$

$$x = 91.6$$

The lower 30% are below 91.6

c) Bounds for middle 80%



$$\text{INVNorm}(.1, 100, 16) + \text{INVNorm}(.9, 100, 16)$$

$$x_1 = 79.5 \quad x_2 = 120.5$$

Middle 80% are between 79.5 & 120.5.

5.5 Normal Probability Plots

How to tell if something is Normal? Can draw a Histogram of the data, but it still might not be clear.

Normal Probability Plot (p. 127): A plot with data values on the vertical axis and z-scores on the horizontal axis. If the distribution is roughly Normal, the plot is roughly a diagonal straight line. Deviations from a straight line indicate the distribution if not Normal.

Fig 5.8, p.127

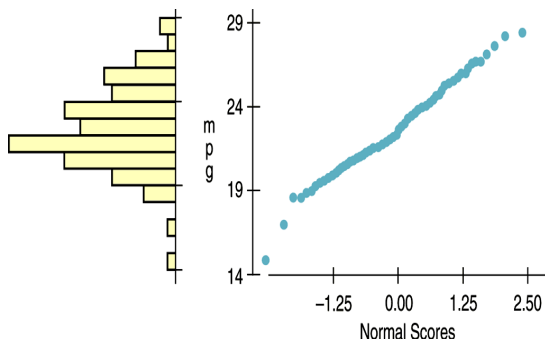


Fig 5.9, p.127

