# CS555B1 Data Analysis and Visualization

Lecture 2

Kia Teymourian

- All of class R examples are available on **Github**

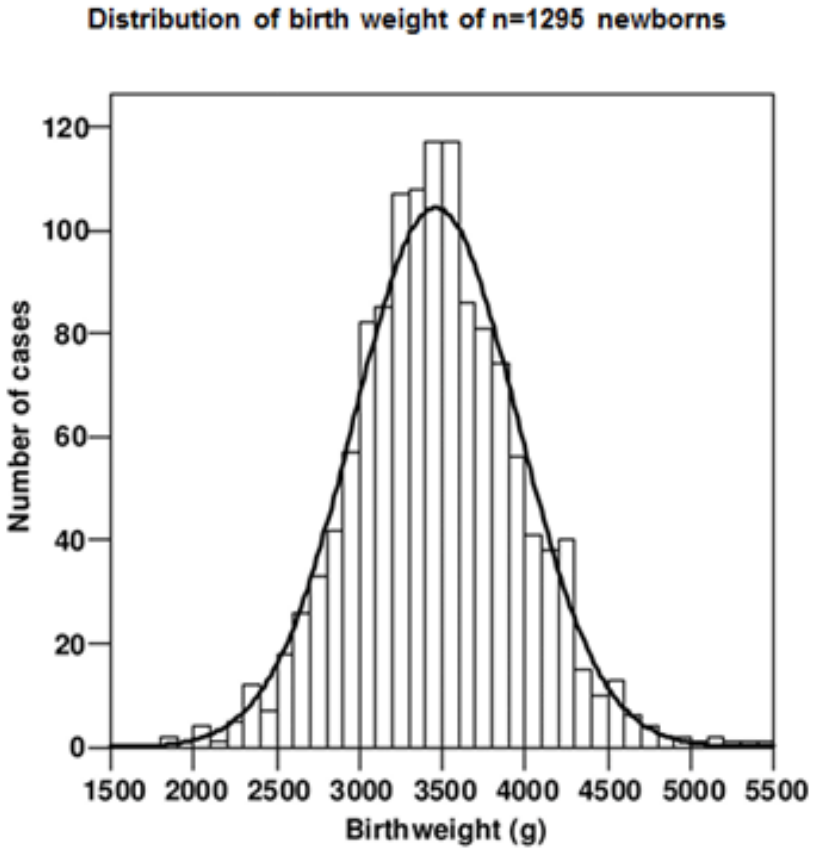
  **https://github.com/kiat/R-Examples**

- Github is a source code repository
- Git is a version control system
- You can download the code as zip file or use git to get the latest code.

# Normal Distribution

Many statistical inference procedures are based on the normal distribution.

The density curve (a mathematical model that represents the pattern of data) of a normal distribution has a very familiar, bell curve shape with a single peak. It is perfectly symmetrical.

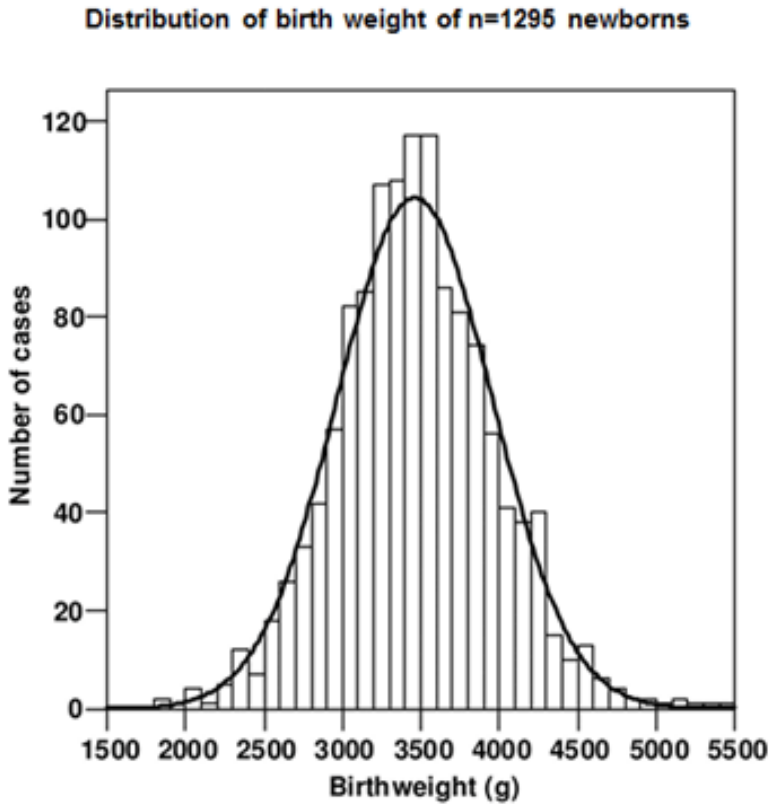



Pfab T et al. Circulation. 2006;114:1687-1692

# Normal Distribution

Many statistical inference procedures are based on the normal distribution.

The density curve (a mathematical model that represents the pattern of data) of a normal distribution has a very familiar, bell curve shape with a single peak. It is perfectly symmetrical.




Pfab T et al. Circulation. 2006;114:1687-1692

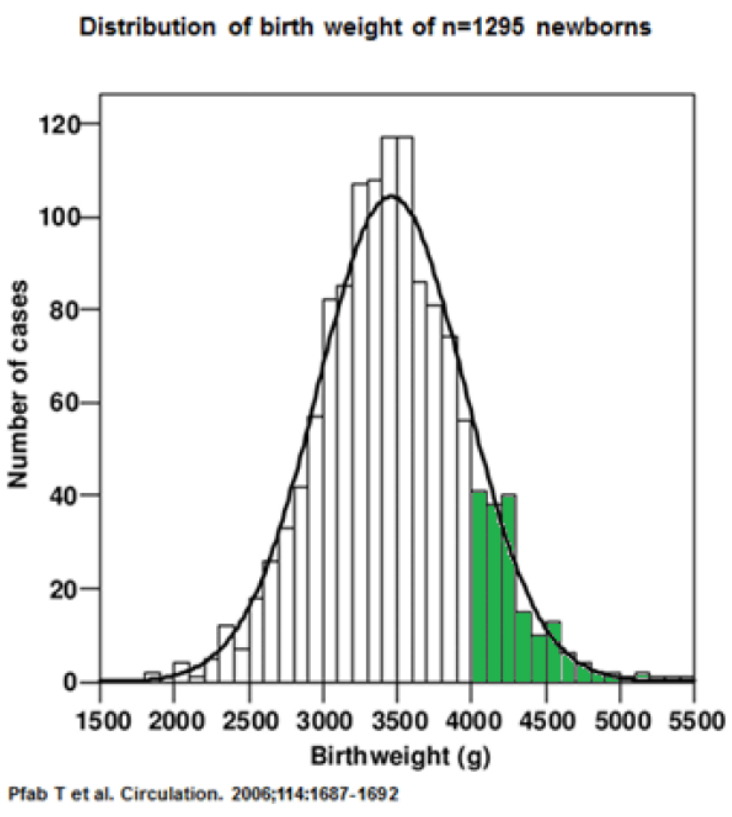
A 2006 paper (Low Birth Weight, a Risk Factor for Cardiovascular Diseases in Later Life, Is Already Associated with Elevated Fetal Glycosylated Hemoglobin at Birth) showed a histogram of the birth weights of newborns in their sample.

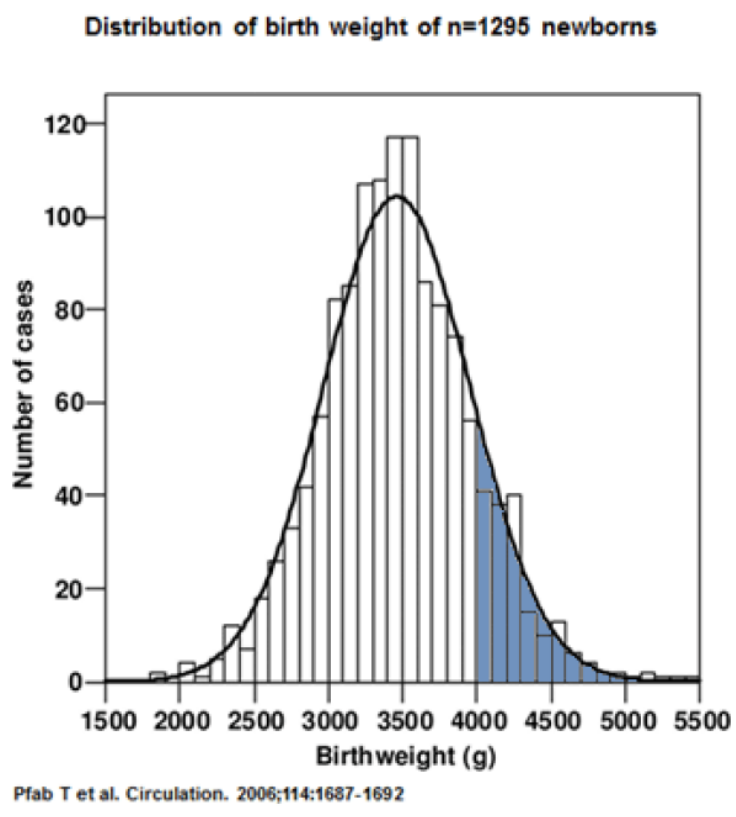
The normal density curve seems to fit the data well.
It is an “idealized” description of the data. It gives the general picture of the data, ignoring minor irregularities.
In situations like this, we can use properties of the normal distribution to make statements about the quantitative variable.

# the proportion of newborns who weighed > 4000g

Using the histogram, it is the proportion of the area of the bars of the histogram to the right of 4000g (where the area of the bars shaded in green).

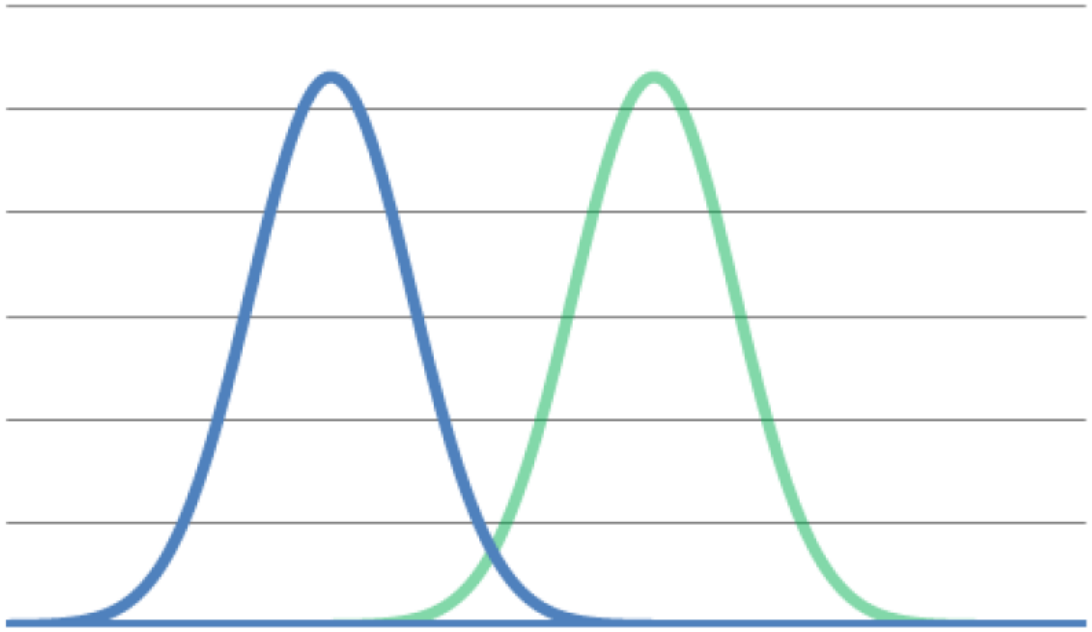
Alternatively, if we scale the curve such that the total area under the curve is 100%, then the area to the right of 4000g under the curve (shaded in blue) would be an approximation of the proportion of infants weighing more than 4000g.



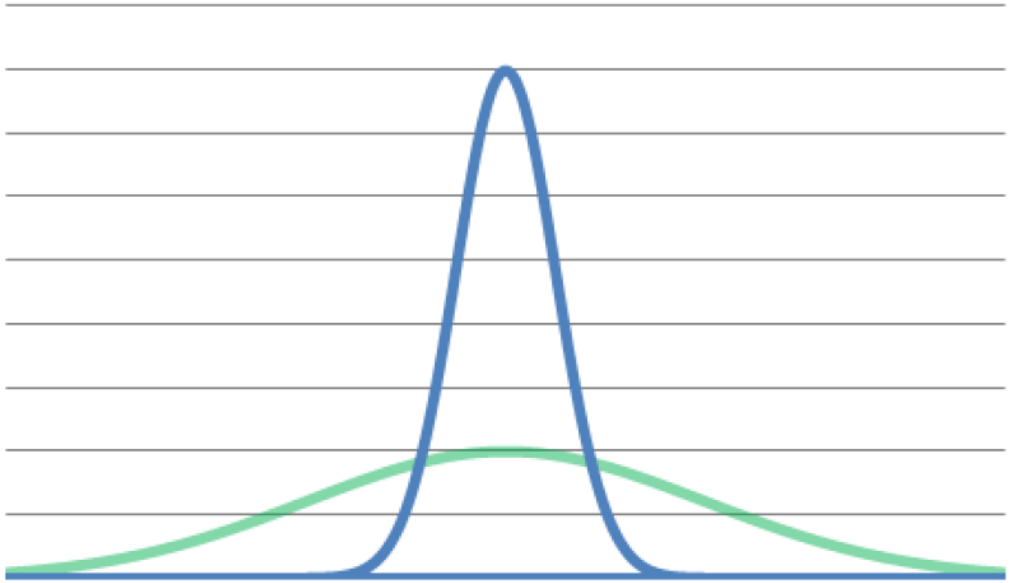



# Normal Distribution

For a normal distribution with mean, μ, and standard deviation, σ,
it is often denoted **N(μ, σ)**.
The mean is the center of the distribution and is the point that splits the area under the bell shaped curve in half.



Different Means

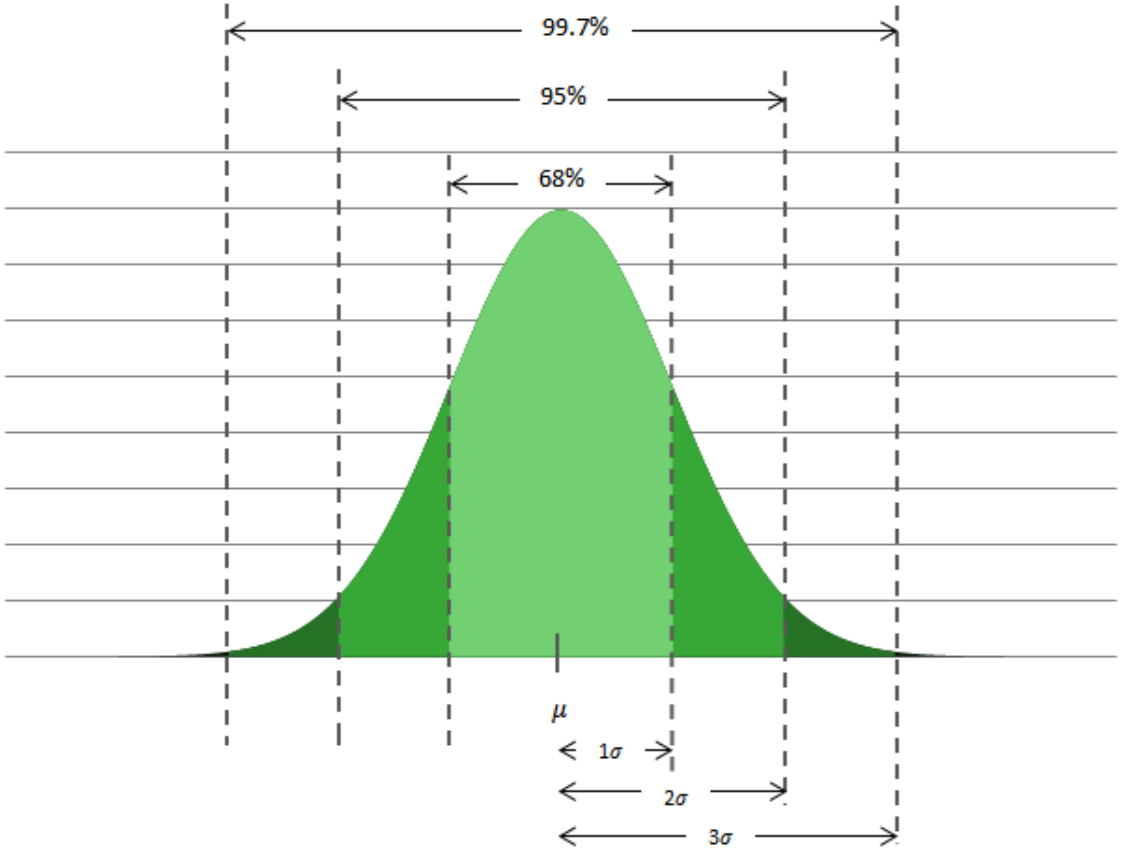


Different Standard Deviations

# 68-95-99.7 Rule

For a normal distribution with mean, μ, and standard deviation, σ,

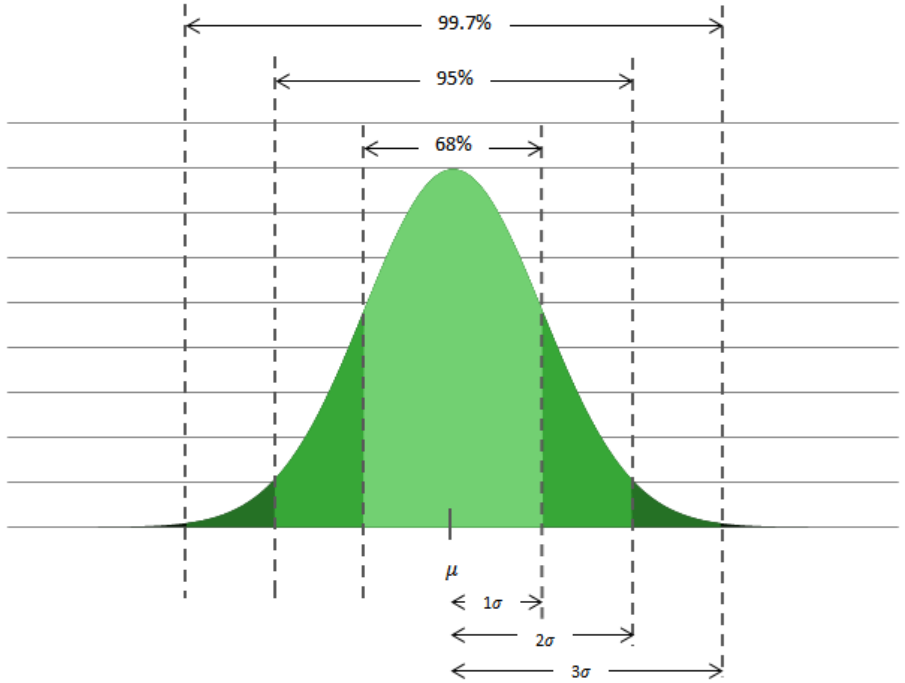
- 68% of the observations fall within one standard deviation of the mean
- 95% of the observations fall within two standard deviations of the mean
- 99.7% of the observations fall within three standard deviations of the mean



# Example – 68, 95, 99.7 Rules

- Let's assume that the distribution of birth weights is normally distributed with a mean of 3500 grams with a standard deviation of 500 grams.
  - What proportion of infants weigh between 3000 and 4000 grams at birth?
  - What proportion weigh between 2500 and 4500 grams?
  - What percentage weigh more than 4000 grams?

- 68 rule tells us that 68% of data is within (μ-σ) and (μ+σ)
- 95 rule tells us that 95% of data is within (μ-2σ) and (μ+2σ)

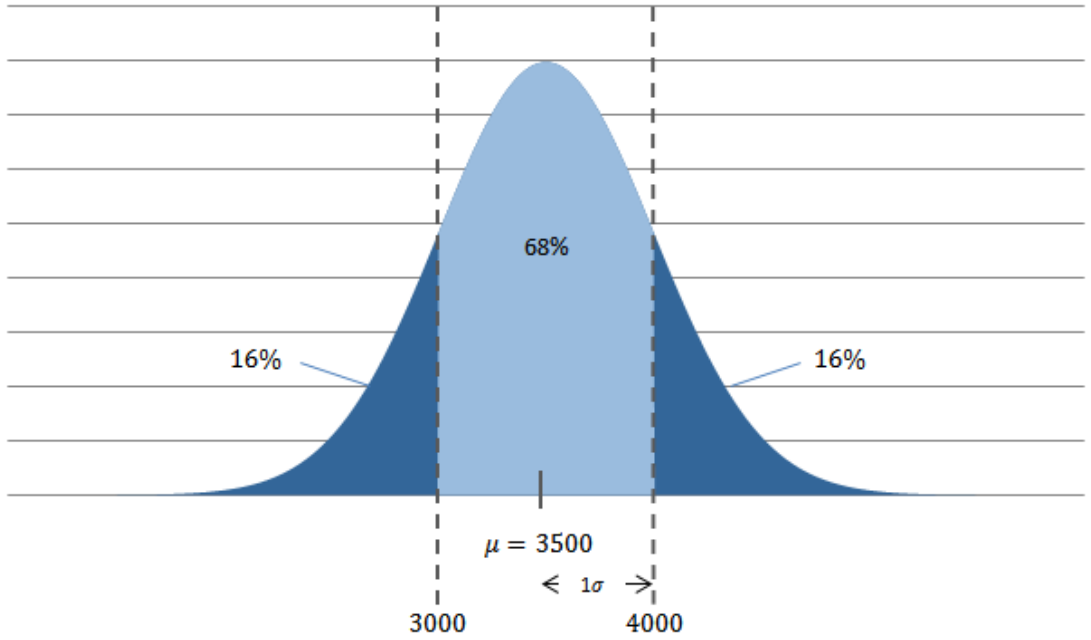


# Answer

- What proportion is between 3000 and 4000 grams at birth?
  Apply 68% rule, 3500 − 500 = 3000 grams and 3500+500=4000 grams

- What proportion weigh between 2500 and 4500 grams?
  95% rule, 3500-2*500=2500 and 3500+2*500=4500

- What percentage weigh more than 4000 grams?
  68% between 3000 grams and 4000 grams. 100%−68%=32% of the data are outside Considering the symmetry, (16%) is below 3000 grams and (16%) is above 4000 grams.



# Exercise

The distribution of SAT scores for the verbal section for high school seniors is approximately a normal distribution with a **mean of 504** and a **standard deviation of 111**.

**- What proportion of seniors score between 393 and 615?**

## Answer

$\mu = 504, \sigma = 111$

$393 - 504 = -111$
$615 - 504 = 111$

393 and 615 are one standard deviation away from the mean of 504
$393 = 504 - 111 = \mu - \sigma$
$615 = 504 + 111 = \mu + \sigma$

The 68-95-99.7 rule tells us that 68% of scores are between 393 and 615.

# Standardized Normal Distribution

- If we measure in standard deviation units, all normal distribution density curves are the same.

- If we define an area under a normal distribution in terms of standard deviation units, then regardless of the actual values of the mean and standard deviation, that area corresponds with the same proportion of observations.

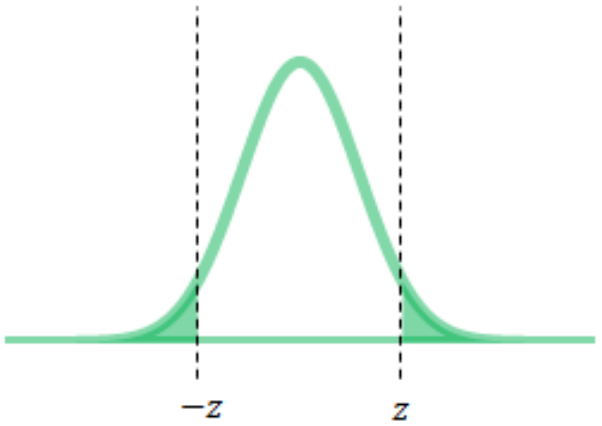
- Convert a value into standard deviation units by calculating its Z-score

- Z-score tells us how many standard deviation x is from the mean

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

# Standardized Normal Distribution

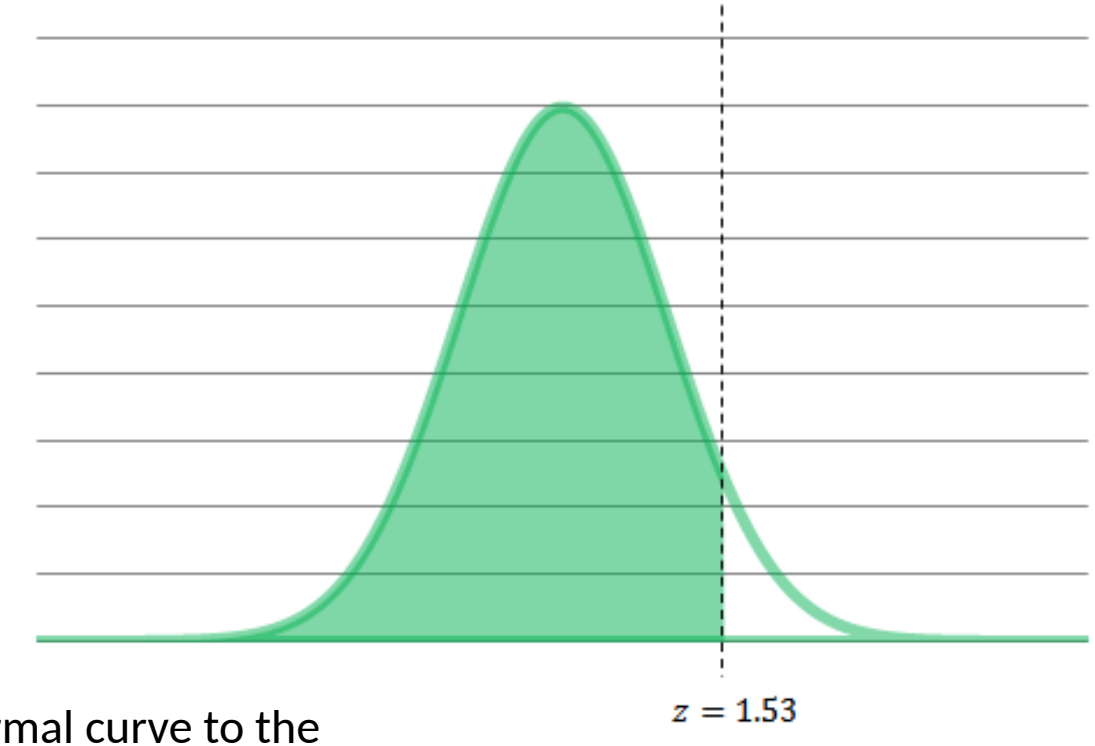
- If the original variable was normally distributed with a specific mean standard deviation, the distribution of standardized z values follows a normal distribution with a **mean of 0** and a **standard deviation of 1** (the standard Normal distribution).
- This means that regardless mean and standard deviation of a particular normal distribution, **we can "transform" it a standard Normal distribution.**
- We can use a single distribution (the standard Normal distribution) to compute areas under the curve (which translate into proportions of observations).

# Standardized Normal Distribution

- Convert a value into standard deviation units by calculating its Z-score
- Z-score tells us how many standard deviation x is from the mean

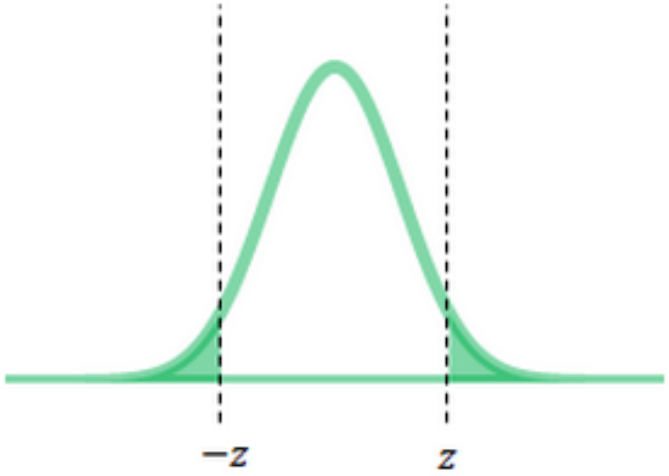
$$z = \frac{x - \mu}{\sigma}$$

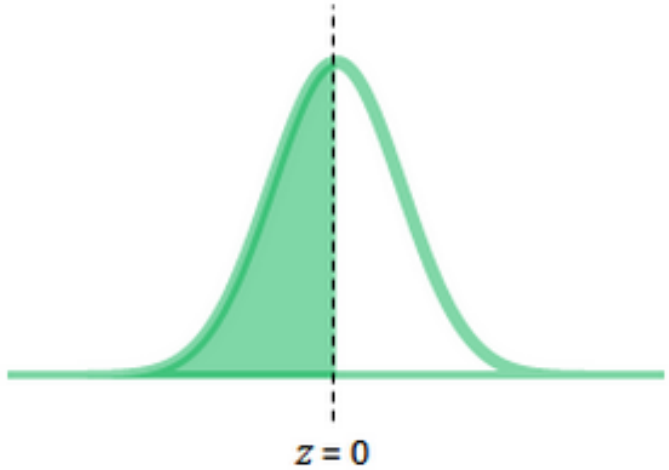



The area under the standard Normal curve to the left of z is 0.9370.

# Properties of Standardized Normal Distribution

- The total area under the standard normal curve is 1.00.
- The curve is perfectly symmetrical. If z is greater than 0, then the probability that a given observation is less than −z is equal to the probability that a given observation is greater than z.
- The standard normal curve is centered on 0.
- The probability that $z \geq a$ is equal to the probability that $z > a$. There is no area under the curve for a single point. Similarly, the probability that $z \leq a$ is equal to the probability that $z < a$.



## Table A. Standard Normal Probabilities (continued)

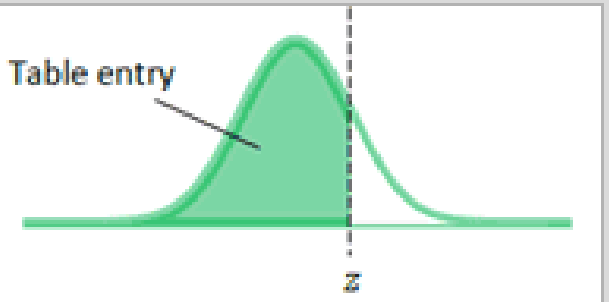


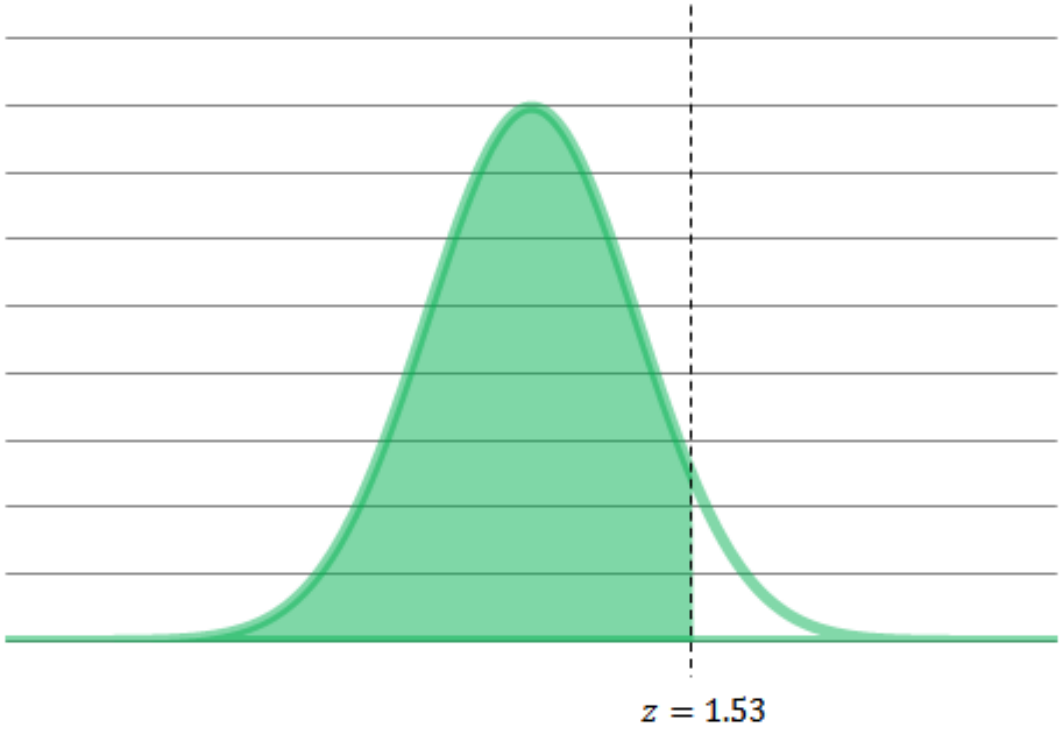

Table entry for $z$ is the area under the standard Normal curve to the left of $z$.

| z | .00 | .01 | .02 | .03 | .04 | .05 | .06 | .07 | .08 | .09 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0.00 | 0.5000 | 0.5040 | 0.5080 | 0.5120 | 0.5160 | 0.5199 | 0.5239 | 0.5279 | 0.5319 | 0.5359 |
| 0.10 | 0.5398 | 0.5438 | 0.5478 | 0.5517 | 0.5557 | 0.5596 | 0.5636 | 0.5675 | 0.5714 | 0.5753 |
| 0.20 | 0.5793 | 0.5832 | 0.5871 | 0.5910 | 0.5948 | 0.5987 | 0.6026 | 0.6064 | 0.6103 | 0.6141 |
| 0.30 | 0.6179 | 0.6217 | 0.6255 | 0.6293 | 0.6331 | 0.6368 | 0.6406 | 0.6443 | 0.6480 | 0.6517 |
| 0.40 | 0.6554 | 0.6591 | 0.6628 | 0.6664 | 0.6700 | 0.6736 | 0.6772 | 0.6808 | 0.6844 | 0.6879 |
| 0.50 | 0.6915 | 0.6950 | 0.6985 | 0.7019 | 0.7054 | 0.7088 | 0.7123 | 0.7157 | 0.7190 | 0.7224 |
| 0.60 | 0.7257 | 0.7291 | 0.7324 | 0.7357 | 0.7389 | 0.7422 | 0.7454 | 0.7486 | 0.7517 | 0.7549 |
| 0.70 | 0.7580 | 0.7611 | 0.7642 | 0.7673 | 0.7704 | 0.7734 | 0.7764 | 0.7794 | 0.7823 | 0.7852 |
| 0.80 | 0.7881 | 0.7910 | 0.7939 | 0.7967 | 0.7995 | 0.8023 | 0.8051 | 0.8078 | 0.8106 | 0.8133 |
| 0.90 | 0.8159 | 0.8186 | 0.8212 | 0.8238 | 0.8264 | 0.8289 | 0.8315 | 0.8340 | 0.8365 | 0.8389 |
| 1.00 | 0.8413 | 0.8438 | 0.8461 | 0.8485 | 0.8508 | 0.8531 | 0.8554 | 0.8577 | 0.8599 | 0.8621 |
| 1.10 | 0.8643 | 0.8665 | 0.8686 | 0.8708 | 0.8729 | 0.8749 | 0.8770 | 0.8790 | 0.8810 | 0.8830 |
| 1.20 | 0.8849 | 0.8869 | 0.8888 | 0.8907 | 0.8925 | 0.8944 | 0.8962 | 0.8980 | 0.8997 | 0.9015 |
| 1.30 | 0.9032 | 0.9049 | 0.9066 | 0.9082 | 0.9099 | 0.9115 | 0.9131 | 0.9147 | 0.9162 | 0.9177 |
| 1.40 | 0.9192 | 0.9207 | 0.9222 | 0.9236 | 0.9251 | 0.9265 | 0.9279 | 0.9292 | 0.9306 | 0.9319 |
| 1.50 | [illegible] | [illegible] | [illegible] | 0.9370 | 0.9382 | 0.9394 | 0.9406 | 0.9418 | 0.9429 | 0.9441 |
| 1.60 | 0.9452 | 0.9463 | 0.9474 | 0.9484 | 0.9495 | 0.9505 | 0.9515 | 0.9525 | 0.9535 | 0.9545 |
| 1.70 | 0.9554 | 0.9564 | 0.9573 | 0.9582 | 0.9591 | 0.9599 | 0.9608 | 0.9616 | 0.9625 | 0.9633 |
| 1.80 | 0.9641 | 0.9649 | 0.9656 | 0.9664 | 0.9671 | 0.9678 | 0.9686 | 0.9693 | 0.9699 | 0.9706 |
| 1.90 | 0.9713 | 0.9719 | 0.9726 | 0.9732 | 0.9738 | 0.9744 | 0.9750 | 0.9756 | 0.9761 | 0.9767 |

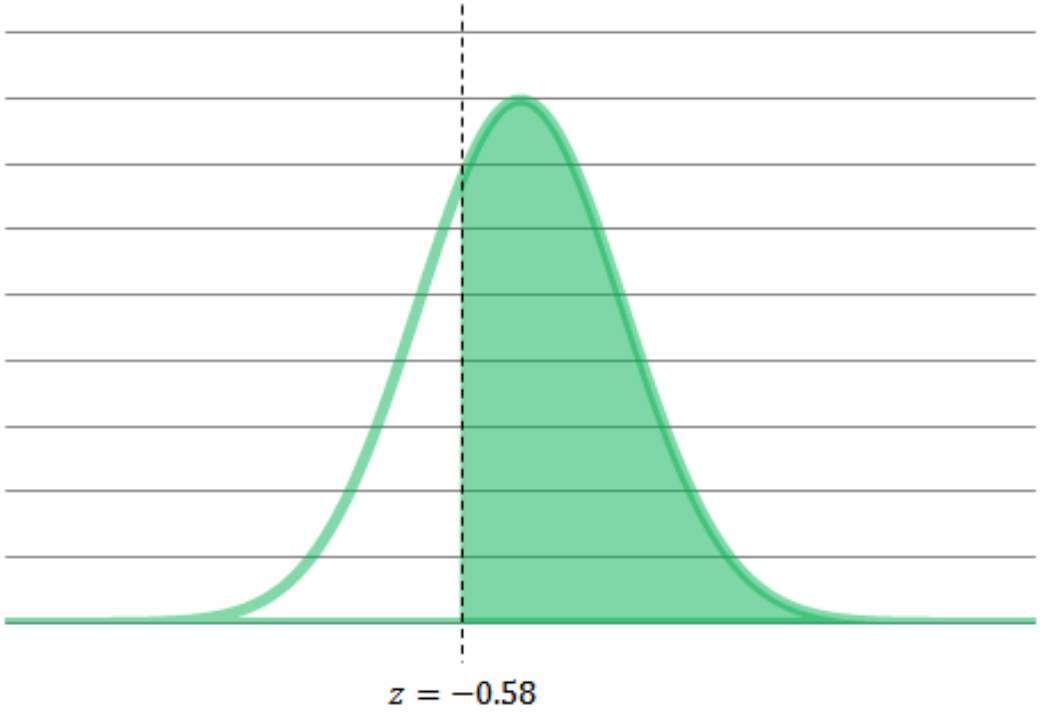
# Exercise

Using the tables,

- find the area under the standard normal curve to the left of z=1.53
- find the proportion of observations greater than z=−0.58.




The area under the standard Normal curve to the left of z is 0.937.




The area under the standard Normal curve to the right of z is 1 - 0.281 = 0.719 .

# Steps to solve the question

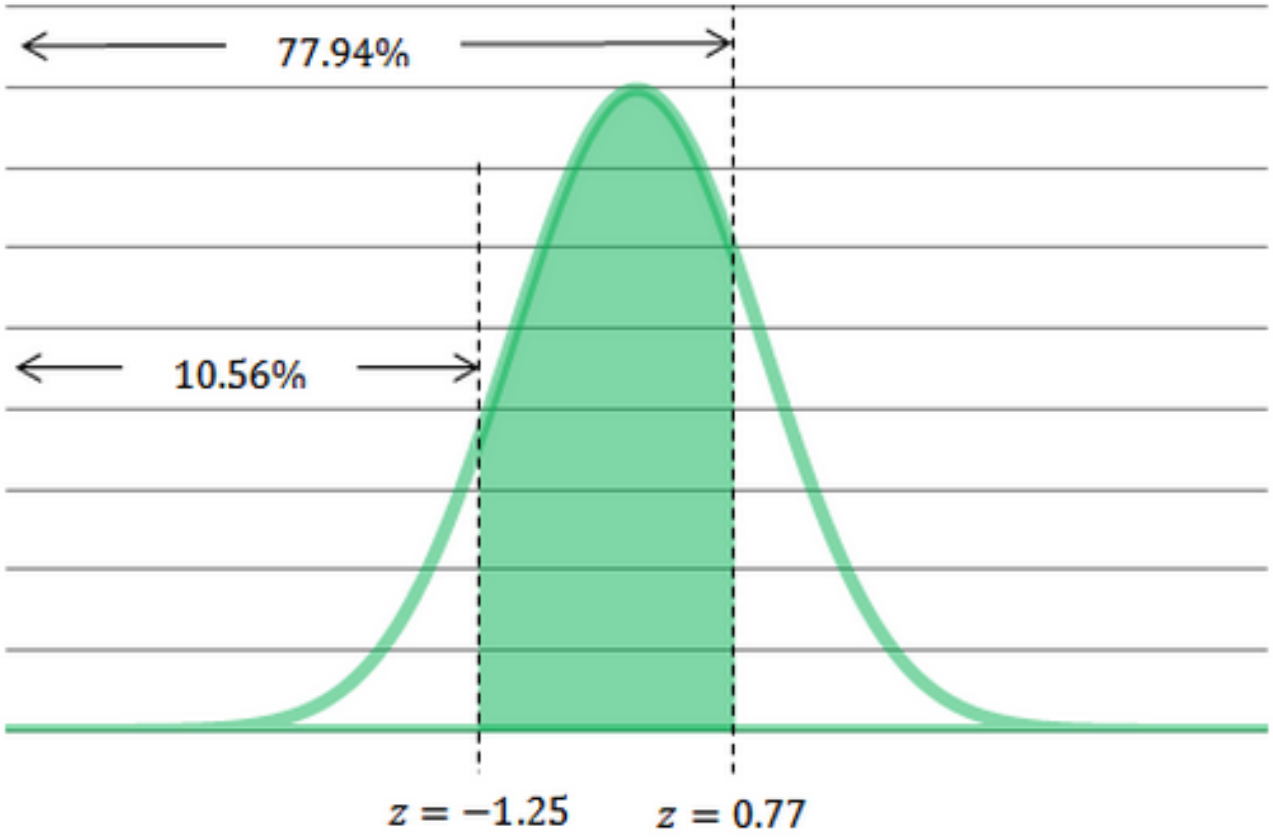
1. State the problem in terms of **x**.

2. Standardize **x** to restate the problem in terms of the standard normal variable **z**. Draw a picture to show the area under the standard normal curve.

3. Find the required area under the standard normal curve using Table A and keep in mind that the total area under the curve is 1.

# Exercise

Using the tables, find the area under the standard normal curve between z=−1.25 and z=0.77.




The area to the left of z=−1.25 is 0.1056.
The area to the left of z=0.77 is 0.7794.

The area between z=−1.25 and z=0.77 is equal to 0.7794−0.1056=0.6738.

# Normal Distribution – R commands

**pnorm() computes the probability** that a normally distributed random number will be less than the given number. This function is also called the "**Cumulative Distribution Function" (CDF)**.

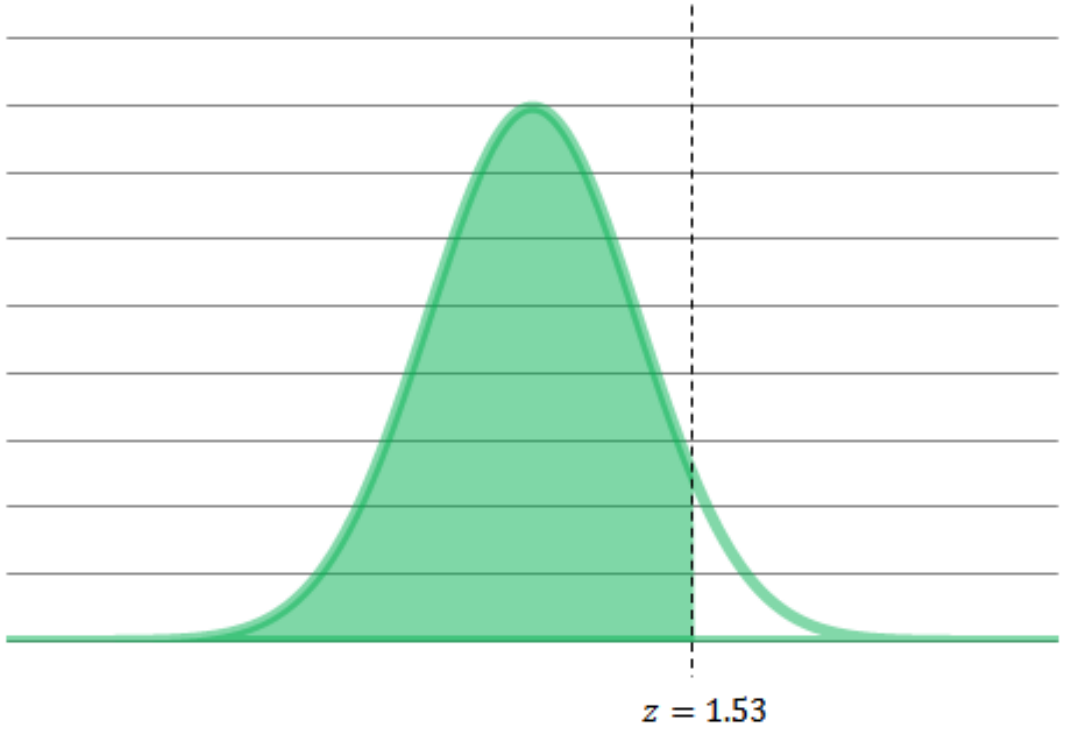
Calculate the area to the left of z: $P(Z \leq z)$
> pnorm(z)

For non-standardized normal distribution
> pnorm(x, mean=a, sd=b) # calculate the area to the left of x

# Find the area under the standard curve

z=1.53

R command
> pnorm(1.53)
0.9369916



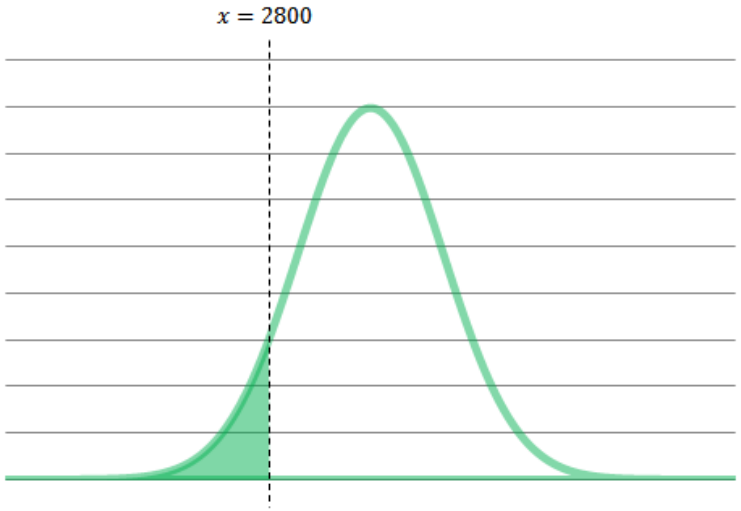

The area under the standard Normal curve to the left of z is 0.937.

# Standardized Normal Distribution – an example

Let's assume that the birth weights of newborns are normally distributed with a mean of 3500g with a standard deviation of 500g.

**What proportion of infants weigh less than 2800g?**

The variable x, the birth weight, has the N(3500,500) distribution




$$\frac{x - 3500}{500} < \frac{2800 - 3500}{500}$$

$$z < -1.40$$

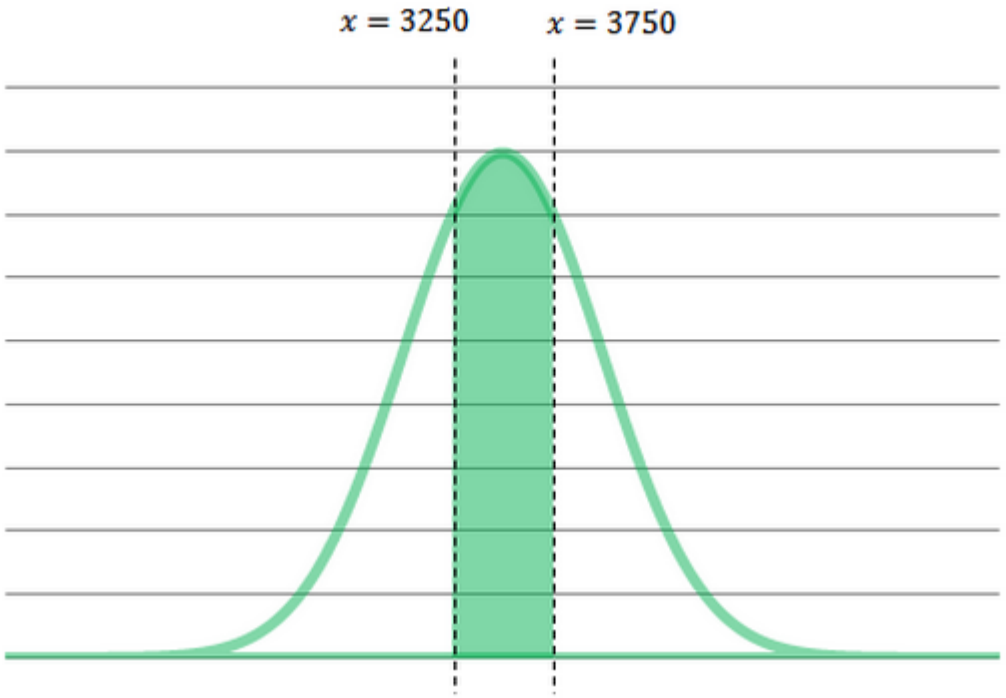
The value in Table A for z=−1.40 is 0.0808. As such, 8.08% of infants weigh less than 2800g.

R command
> pnorm(2800, mean=3500, sd=500)
0.08075666

# Standardized Normal Distribution – an example

Let’s assume that the birth weights of newborns are normally distributed with a mean of 3500g with a standard deviation of 500g.

What proportion of infants weigh between 3250g and 3750g?



# Standardized Normal Distribution – an example

1. State the problem in terms of x. Let x be the birth weight of infants. The variable x has the N(3500,500) distribution. We are interested in the proportion of infants with 3250<x<3750 grams.
2. Standardize x to restate the problem in terms of the standard normal variable z.

$$3250 < x < 3750$$

$$\frac{3250 - 3500}{500} < \frac{x - 3500}{500} < \frac{3750 - 3500}{500}$$

$$-0.50 < z < 0.50$$

3. Find the required area under the standard normal curve using the Table. The area between z=−0.50 and z=0.50 is 0.6915–0.3085=0.3830

R command

> pnorm(3750, mean=3500, sd=500) - pnorm(3250, mean=3500, sd=500)

0.3829249

# R functions on Normal Distribution

**dnorm()** - Given a set of values it returns the **height of the probability distribution at each point**.

If you only give the points it assumes you want to use a mean of 0 and standard deviation of 1.

```
density_standard_norm <- function(x){
1/sqrt(2*pi)*exp(-0.5*x^2)
}

> dnorm(0)
0.3989423

> dnorm(0, mean=4, sd=10)
0.03682701

> pnorm(0)
0.5
```

# R Function on Normal Distribution - qnorm

In statistics, quantiles are cut points dividing the range of a probability distribution **into contiguous intervals with equal probabilities.**

Given a probability p and a distribution, we want to calculate the corresponding quantile for p: the value x such that $P(X \leq x) = p$

```
> qnorm(x)
```

For non-standardized normal distribution

```
> qnorm(x, mean=a, sd=b)
```

# Plotting Normal Distribution

Define a vector

```
> x <- seq(from = -3, to = 3, length.out = 100)
```

Apply the distribution density function to the vector

```
> y <- dnorm(x)
```

Plot it

```
> plot(x, y, type="l")
```

Shade an area from -1 to 1: define the area by specifying points along the outer edges and then use polygon function to fill the shape

```
> xvalues <- x[x>=-1 & x<=1]
> yvalues <- y[x>=-1 & x<=1]
> region.x <- c(xvalues[1], xvalues, tail(xvalues, 1))
> region.y <- c(0, yvalues, 0)
> polygon(region.x, region.y, col="navy")
```

# Another way to shade an area

Define a vector
```
> curve(dnorm(x), xlim=c(-3,3), main='Normal Density')
```

To shade the region represented by P(-3 < X < -2).
The first vertex we want for our polygon is (-3,0).
```
> cord.x <- c(-3)
> cord.y <- c(0)
```

The 2nd vertex will be (-3, f(-3)), f(-3) is the normal density evaluated at -3.
```
> cord.x <- c(cord.x, -3)
> cord.y <- c(cord.y, dnorm(-3))
```

The 3rd and 4th vertices are (-2,f(-2)) and (-2,0)
```
> cord.x <- c(cord.x, -2, -2)
> cord.y <- c(cord.y, dnorm(-2), 0)

> polygon(cord.x,cord.y,col='skyblue')
```

http://www.r-bloggers.com/creating-shaded-areas-in-r/

# R Function on Normal Distribution

**Density, distribution function, quantile function and random generation** for the normal distribution with mean equal to mean and standard deviation equal to sd.

Usage

```
dnorm(x, mean = 0, sd = 1, log = FALSE)
pnorm(q, mean = 0, sd = 1, lower.tail = TRUE, log.p = FALSE)
qnorm(p, mean = 0, sd = 1, lower.tail = TRUE, log.p = FALSE)
rnorm(n, mean = 0, sd = 1)
```

To read the manual
> help(pnorm)

| | |
|---|---|
| x, q | vector of quantiles. |
| p | vector of probabilities. |
| n | number of observations. If length(n) > 1, the length is taken to be the number required. |
| mean | vector of means. |
| sd | vector of standard deviations. |
| log, log.p | logical; if TRUE, probabilities p are given as log(p). |
| lower.tail | logical; if TRUE (default), probabilities are P[X ≤ x] otherwise, P[X > x]. |

# Outliers

Outliers are data points that do not fit with the general pattern of the data. They can be detected using graphical summaries.

Quartiles divide a rank-ordered data set into four equal parts.
The values that divide each part are called the first, second, and third quartiles;
They are denoted by Q1, Q2, and Q3, respectively.

Q1 is the "middle" value in the first half of the rank-ordered data set.
Q2 is the median value in the set.
Q3 is the "middle" value in the second half of the rank-ordered data set.

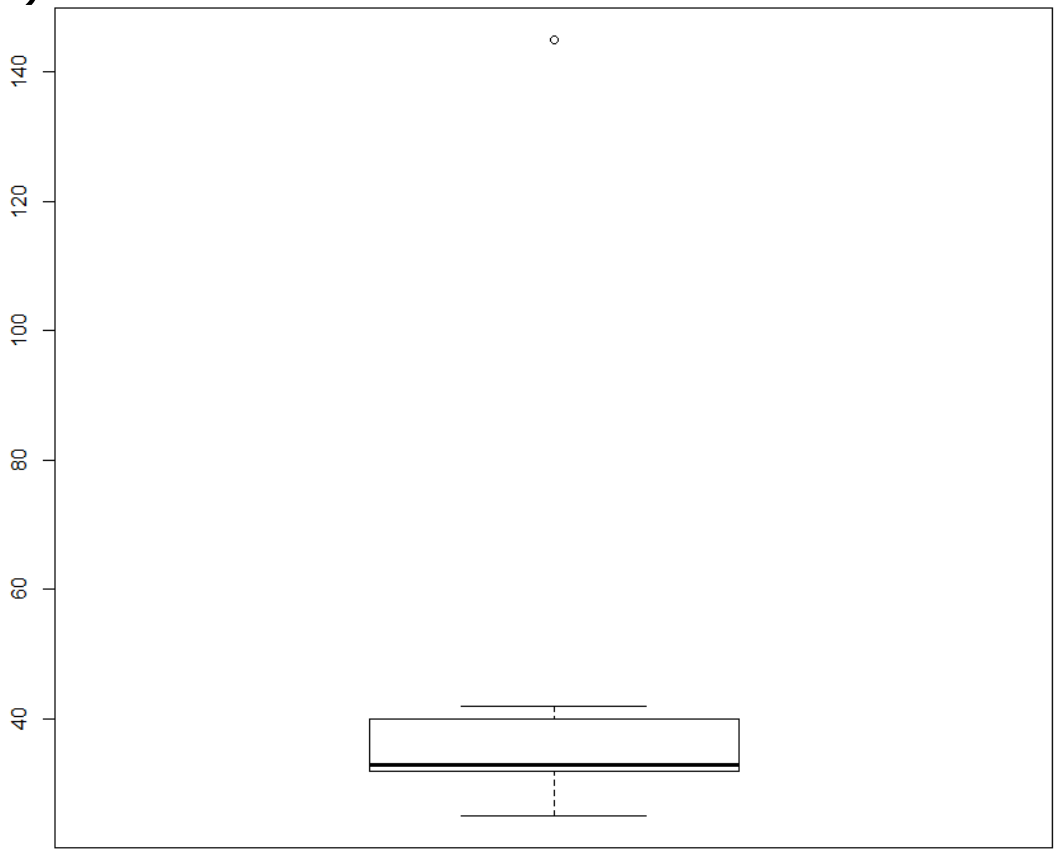
**IQR (inter-quartile range) = Q3 –Q1** It is a measure of variability.

Outliers are defined as any points that are
**<= Q1 – 1.5*IQR**
OR
**>= Q3 + 1.5*IQR**

# Outliers

The earnings (in thousands of dollars) for 9 people are:
35, 40 , 145 , 33 , 30 , 42 , 32 , 32 , 25

```
> earning <- c(35, 40, 145, 33, 30, 42, 32, 32, 25)
> summary(earning)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
     25      32      33      46      40     145
> boxplot(earning)
```



# Sampling

**The population: entire group of individuals or things that we want information about or to answer questions about.**

**A sample: a set of data collected and/or selected from a statistical population by a defined procedure.** The elements of a sample are known as sample points, sampling units or observations.

Often we gather information about only part of the population (sample) and use this information to make educated guesses about qualities of the larger group (population).

However, we need to be sure that our sample is selected in such a way that allows for inference to the population.

In many ways, all samples will have some type of responder bias since those who elect to participate in surveys tend to be different than those that who chose not to participate.

We must also be sensitive to biases that could be imposed by the questions that we ask (and the way that we ask them).

# Selection bias

**Biases** introduced by sampling are often termed selection biases.

In order to avoid them, we try to choose **samples by chance** (instead of letting a poll interviewer or responder choose who is selected for the sample).

Ensuring that each individual or element from the population has an **equal chance** of being selected is key to protecting against selection bias.

The methods used for the collection of the data that we analyze is critically important to our ability to answer questions and to make inferences and conclusions.

We should be mindful of sampling methods and potential for bias as we critically analyze our results.

# Inference about a population

**A parameter** is a number that describes a population.

**A statistic** is a number that is computed from the sample data.

Typical notation for population parameters and their corresponding sample statistic:

| | Population Parameter | Sample Statistic |
|---|---|---|
| Mean | $\mu$ ("mu") | $\bar{x}$ ("x bar") |
| Variance | $\sigma^2$ ("sigma squared") | $s^2$ |
| Standard Deviation | $\sigma$ ("sigma") | $s$ |
| Proportion | $p$ | $\hat{p}$ ("p hat") |

# A parameter vs. a statistic

A researcher was interested in estimating the mean income level for college graduates aged 25–30.

In order to do so, a random sample of 2000 college graduates was taken.

The mean income of the sample was =$45,455, which is a statistic.

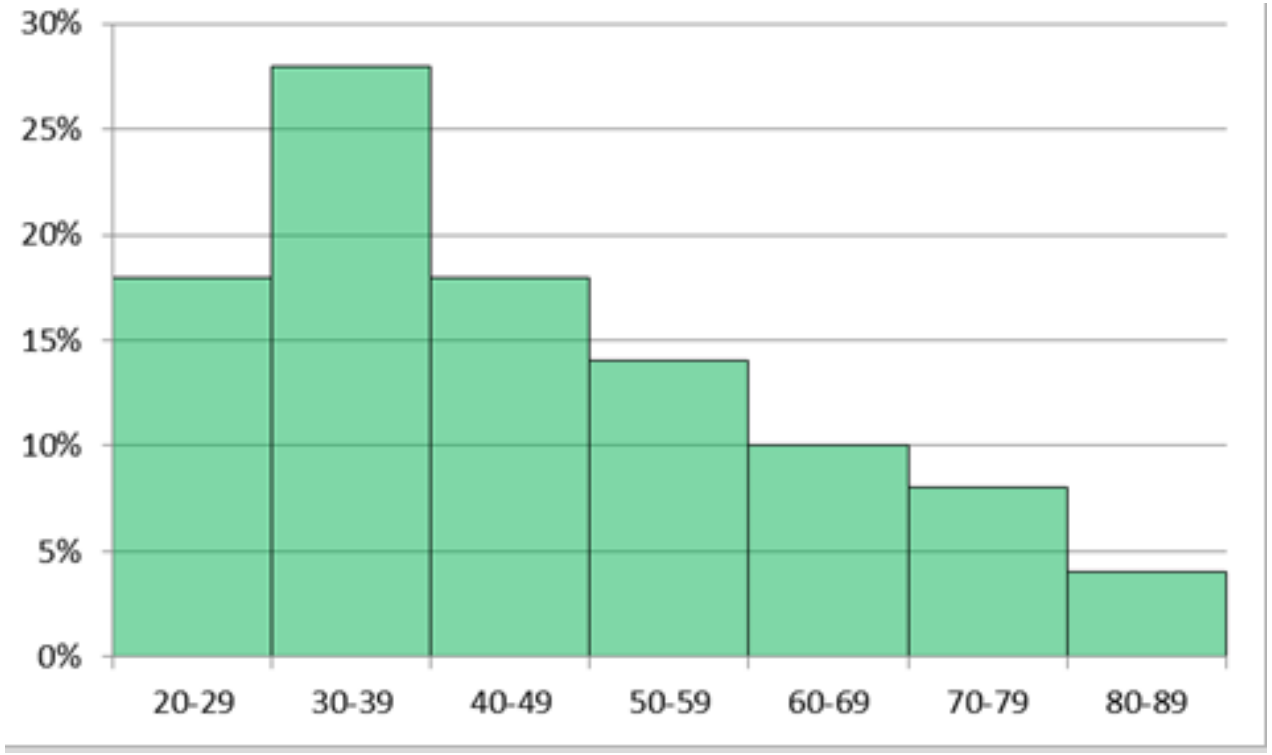
The unknown population parameter, μ, would be the mean of all college graduates aged 25–30.

It is unknown and is estimated by the sample mean.

The actual value of μ could only be obtained if you took the arithmetic average of all approximately 12 million college graduates aged 25–30.

# The Central Limit Theorem – an example

We would like to estimate the mean age of employees at company A. Let's assume that the population is made up of only 50 individuals with the following age distribution:




| | | | | |
|---|---|---|---|---|
| 42 | 20 | 32 | 47 | 31 |
| 66 | 25 | 64 | 25 | 46 |
| 76 | 56 | 32 | 20 | 50 |
| 60 | 58 | 31 | 83 | 51 |
| 22 | 32 | 64 | 49 | 75 |
| 40 | 43 | 54 | 44 | 62 |
| 46 | 27 | 32 | 49 | 37 |
| 38 | 59 | 33 | 59 | 73 |
| 26 | 26 | 83 | 71 | 39 |
| 35 | 33 | 35 | 28 | 35 |

The population mean of the ages of employees is 45.28 years

# The Central Limit Theorem – an example

Assume we take two random samples of size 5. The resulting ages in each of the two samples.

Sample 1: 20,44,46,20,44
Sample 2: 83,32,31,50,32

The first sample has a sample mean of 34.6 and a sample median of 44.
The second sample has a sample mean of 45.6 and a sample median of 32.

Neither the sample mean nor the sample median will always fall closer to the population mean in a given sample.

To evaluate each of these sample statistics and their ability to estimate the true population value, we must not rely on just one example (one sample).

We'd like to compare the distribution of the sample mean and sample median if we take hundreds of random samples of the same size.

# The Central Limit Theorem – an example

We want to evaluate how well the sample mean and sample median perform as estimators of the population mean by using a computer to randomly select 10,000 samples of size 5.

For each sample, we calculated the sample mean and the sample median. Here's what our results look like:

| Sample | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | $\bar{x}$ | $m$ |
|---|---|---|---|---|---|---|---|
| **1** | 20 | 44 | 46 | 20 | 44 | 34.6 | 43 |
| **2** | 83 | 32 | 31 | 50 | 32 | 45.6 | 32 |
| **3** | 58 | 49 | 31 | 32 | 50 | 44 | 49 |
| **4** | 49 | 38 | 31 | 71 | 32 | 44.2 | 38 |
| **5** | 40 | 27 | 76 | 47 | 62 | 50.4 | 47 |
| **6** | 59 | 27 | 32 | 66 | 46 | 46 | 46 |
| **7** | 31 | 64 | 38 | 42 | 62 | 47.4 | 42 |
| **8** | 83 | 49 | 50 | 39 | 22 | 48.6 | 49 |
| **9** | 40 | 26 | 22 | 49 | 54 | 38.2 | 40 |
| **10** | 27 | 47 | 64 | 39 | 54 | 46.2 | 47 |
| ... | | | | | | | |
| **10,000** | 25 | 26 | 33 | 60 | 71 | 43 | 33 |

# The Central Limit Theorem

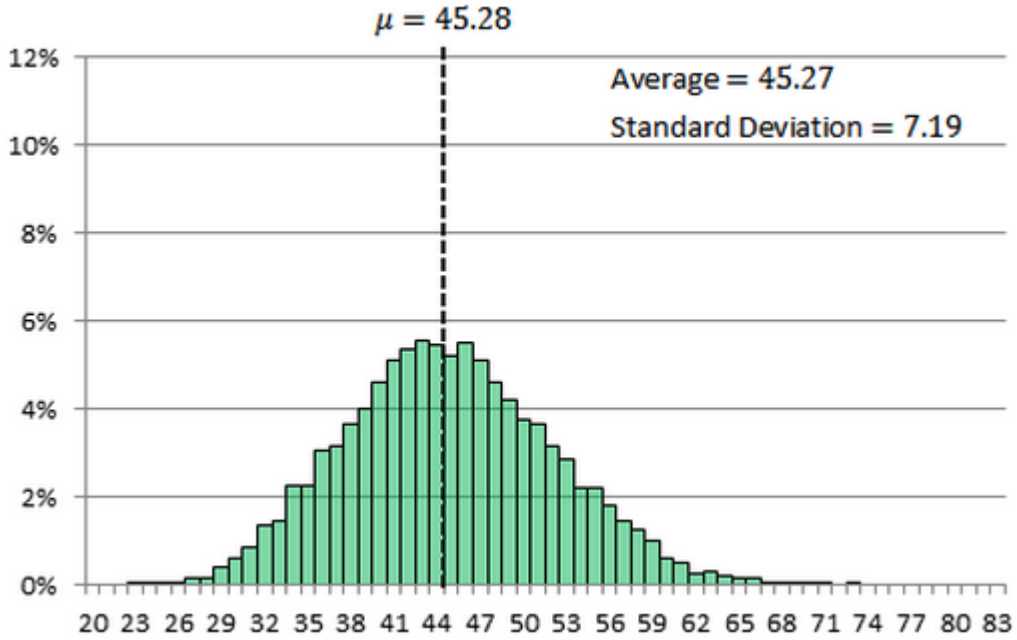
When the number of samples taken from a population is sufficiently large, the **sampling distribution of the sample mean**, will be approximately **normally distributed** with an expected value of μ and a standard deviation of σ (μ and σ are the mean and the standard deviation from the population).

Say you take a random sample of size n from a population with mean μ and standard deviation σ.

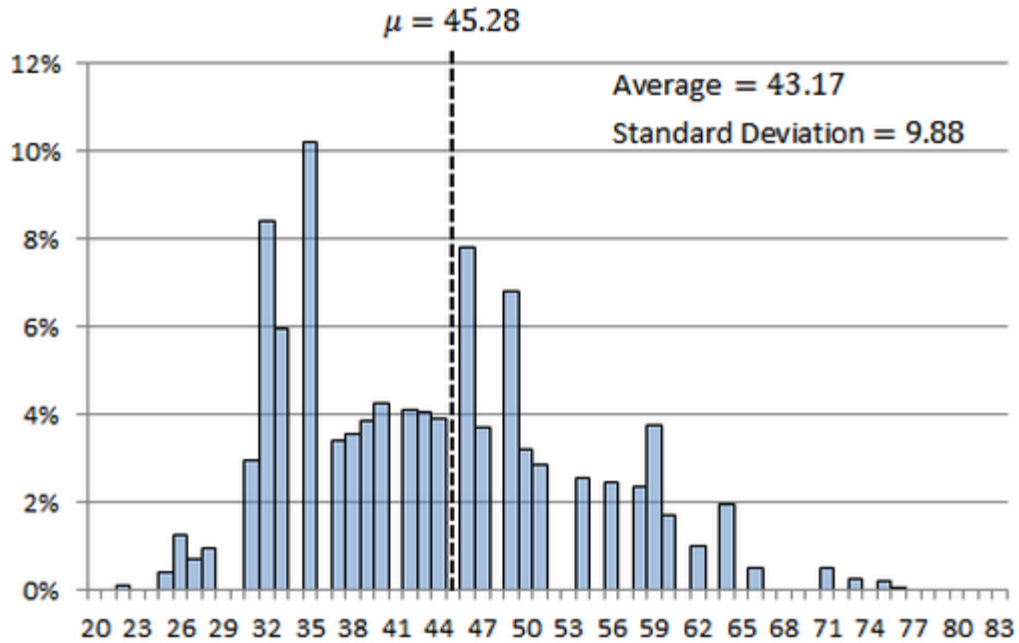
The larger the sample size, the closer the sampling distribution of the sample means will be to the normal distribution (and the smaller the variance will be of the sample mean).

# The Central Limit Theorem – an example

The distribution of the sample mean ($\bar{x}$) for the 10,000 samples of size 5 by using a histogram:



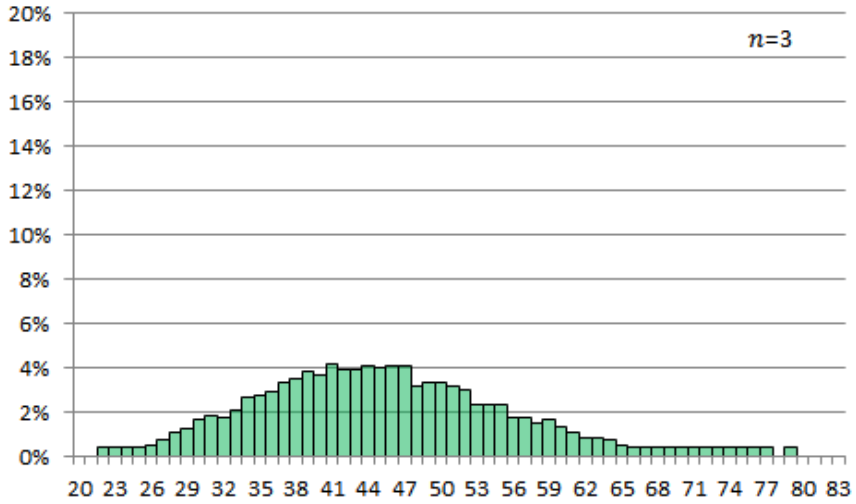

The distribution of the sample median (m) for the 10,000 samples of size 5:

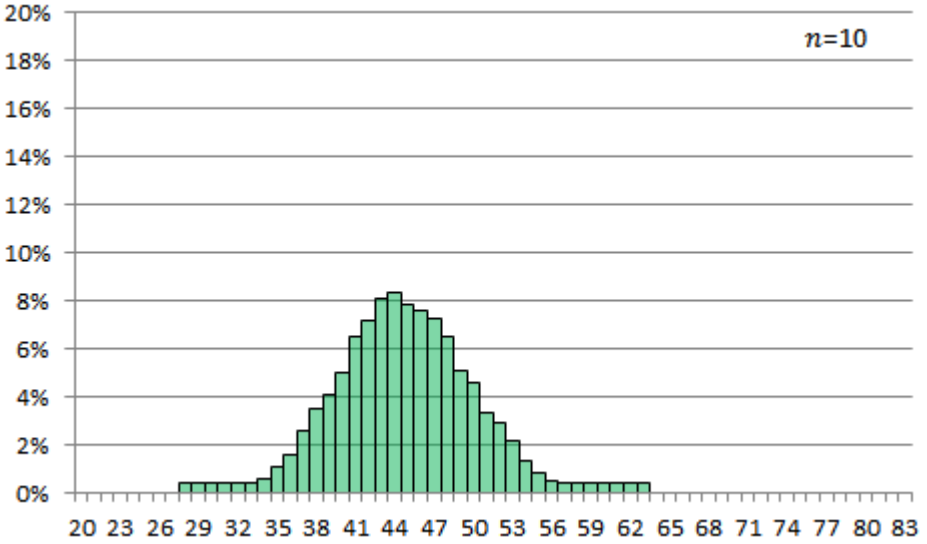


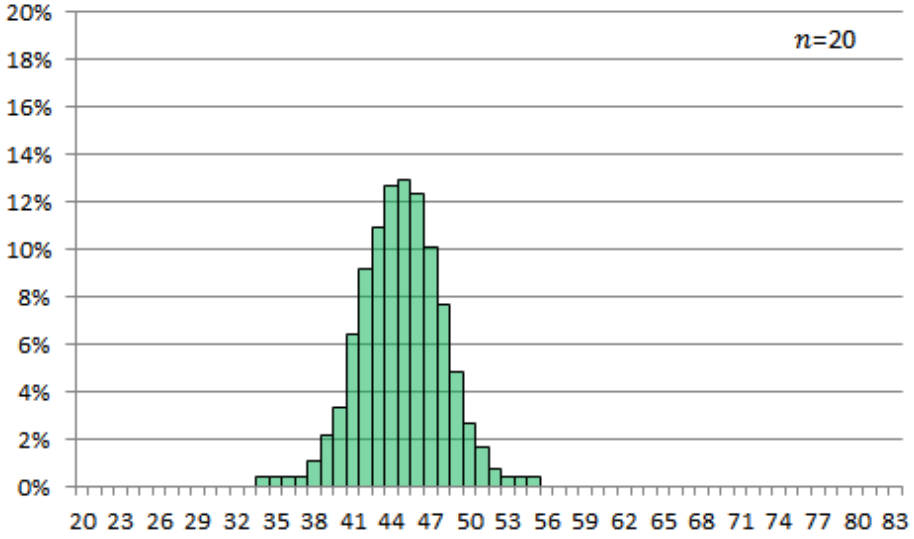

The distribution for the sample mean is centered over the true value μ and has less spread or variability than the distribution of the sample median.

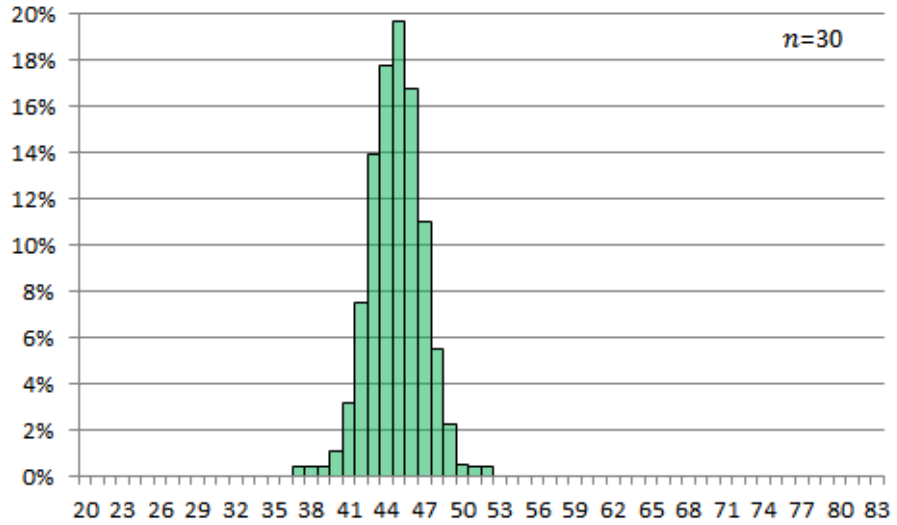
# The Central Limit Theorem – an example

The population has a mean of 45.28 and a standard deviation of 17.20




| $n$ | $\bar{x}_{\bar{x}}$ | $\sigma_{\bar{x}}$ |
|---|---|---|
| 3 | 45.25 | 9.57 |
| 5 | 45.27 | 7.19 |
| 10 | 45.20 | 4.80 |
| 15 | 45.28 | 3.72 |
| 20 | 45.25 | 2.97 |
| 25 | 45.28 | 2.43 |
| 30 | 45.28 | 1.97 |

# Try the Online demonstration

For better understanding try the following online demo

**https://gallery.shinyapps.io/CLT_mean/**

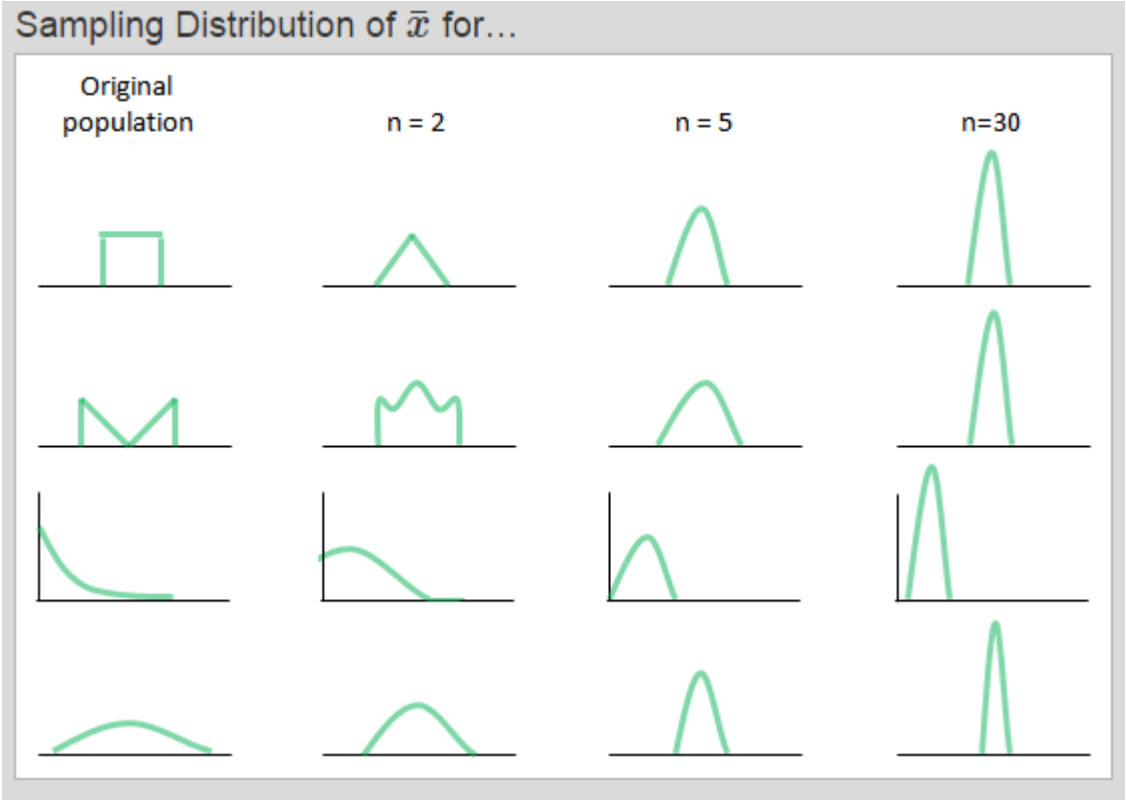
Change the parameters and scroll down to see the **“Sampling Distribution”**

# How large should n be?

If the underlying population is approximately normal to begin with, then even small values of n will give a sampling distribution of sample means that are normally distributed.

For **more skewed population distributions**, **n must be larger** before the sampling distribution is sufficiently normally distributed.

Generally, the rule of thumb is that n should be **n ≥30** for the distribution of the sample means to be reasonably normally distributed.

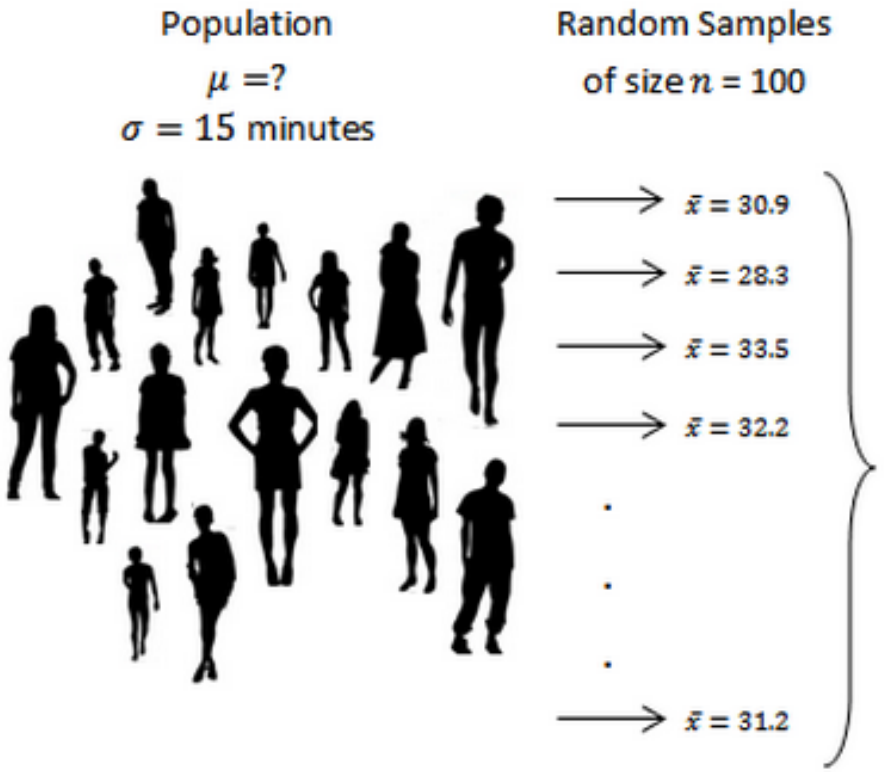


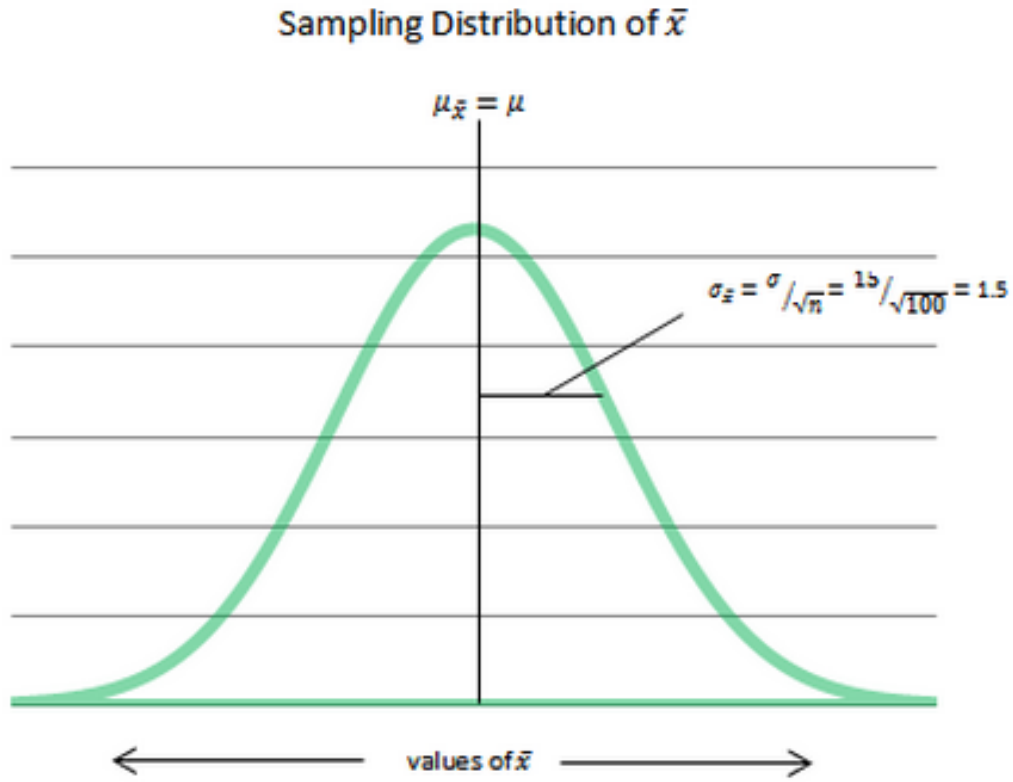
# Standard Deviation of Sample Mean

- If n >= 30 we know that the sample mean is approximately normally distributed with

$$\mu_{\bar{x}} = \mu \qquad SE = \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

Standard Error (SE)



# Exercise

Suppose we have selected a random sample of **n=36** from a population with a mean of **80** and a standard deviation of **6**.

Find the probability that the **sample mean** will be between **79** and **81**.

# Answer

The Central Limit Theorem tells us that regardless of the shape of the underlying population, the sampling distribution of $\bar{x}$ is approximately normal when n≥30.

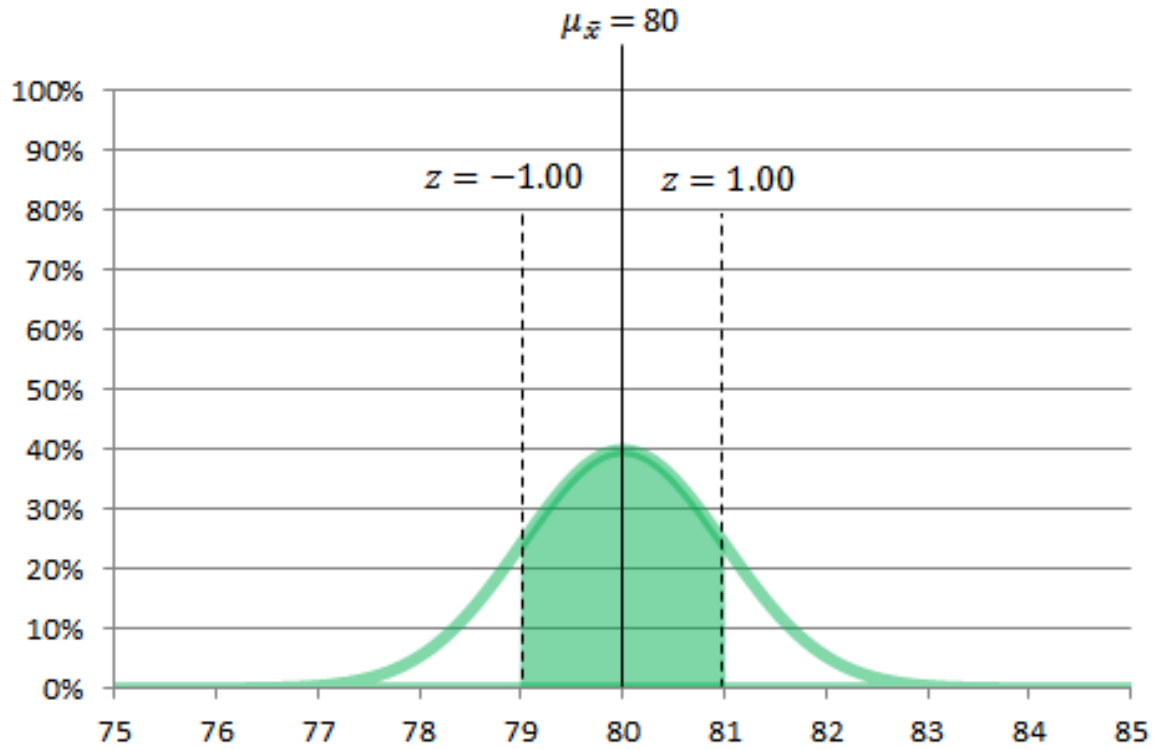
The sampling distribution will have a mean of $\mu_{\bar{x}} = 80$ and $\sigma_{\bar{x}} = \sigma/\sqrt{n} = 6/6 = 1$

We first standardize the distribution and then we use Table A to calculate the probabilities:

$$79 < x < 81$$

$$\frac{79-80}{1} < \frac{x-80}{1} < \frac{81-80}{1}$$

$$-1.00 < z < 1.00$$




0.8413–0.1587=0.68260.8413–0.1587=0.6826

68.26%