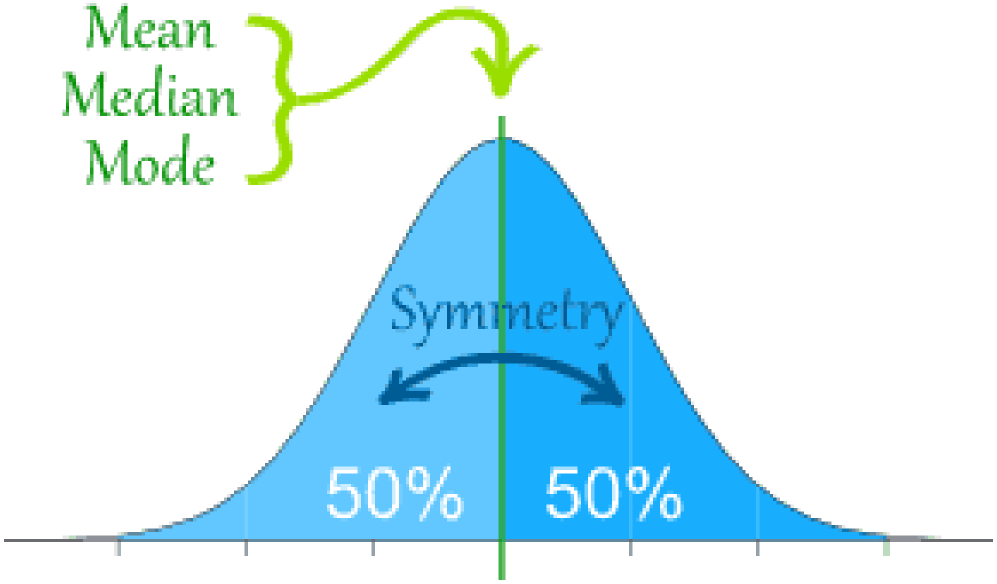
Probability Distribution

Normal Distribution

- The normal distribution is pattern for the distribution of a set of data which follows a ***bell shaped curve.***

- The bell shaped curve has several properties:
  - The curve ***concentrated in the center and decreases on either side***. This means that the data has less of a tendency to produce ***unusually extreme values***, compared to some other distributions.
  - The bell shaped curve is ***symmetric***.
  - The total ***area*** under the normal curve is ***equal to 1***.
  - The ***mean, median, mode*** for a normal distribution all have the ***same value***.

Normal Distribution

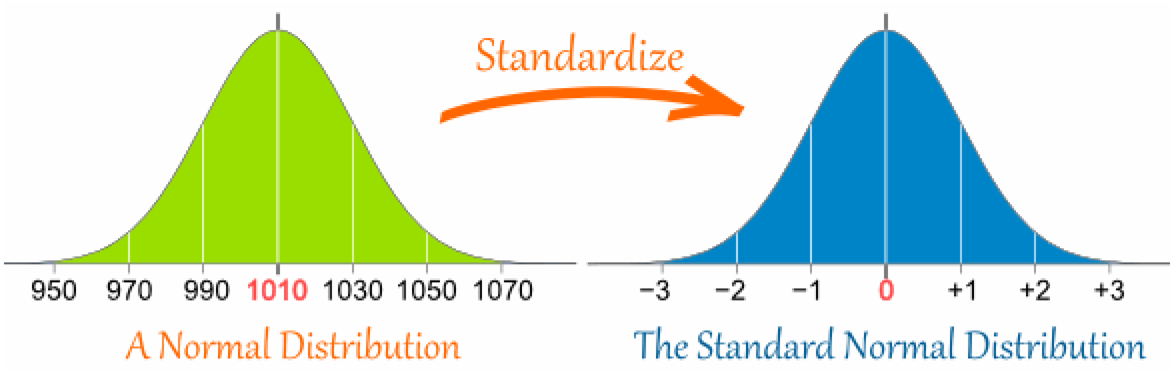


Normal Distribution

- The normal distribution is a ***continuous probability distribution***. This has several implications for probability.
- Find the area under the curve, means finding the probability of the variables.

Standard Normal Distribution

- ***The standard normal distribution*** is that normal variable, X transforms into standard normal distribution, Z with mean, $\mu = 0$ and variance, $\sigma^2 = 1$.
- $X \sim N(\mu , \sigma^2) \Rightarrow Z \sim N(0 , 1)$



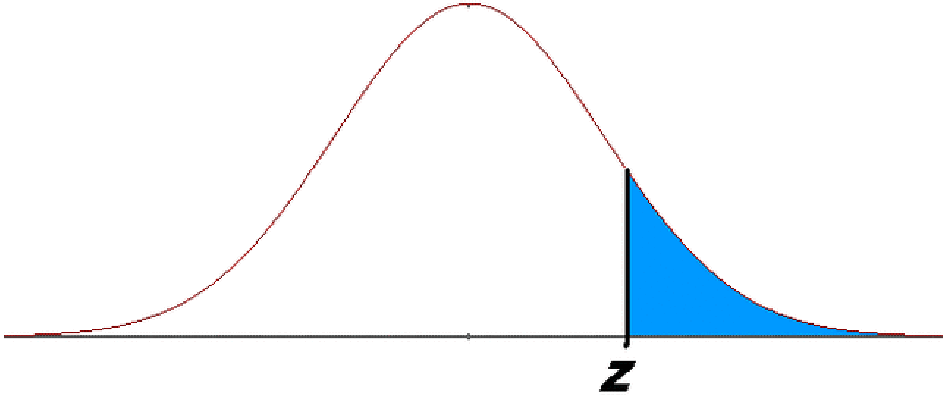
- ***We need standard normal table to look for probabilities.***

Standard Normal Distribution

- The standard score (Z score) is defined by the formula

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

- Its values are usually represented by the symbols z

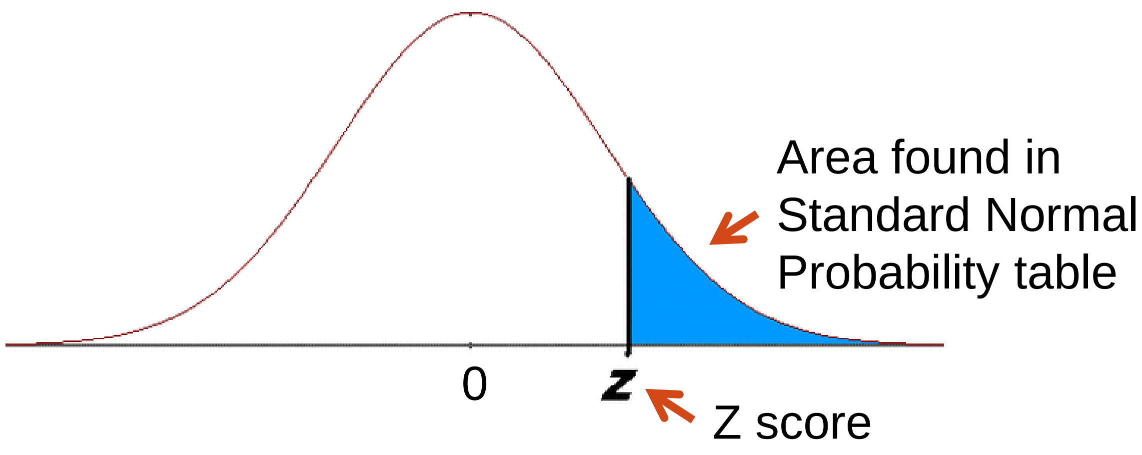


- ***We need standard normal table to look for probabilities.***

How to Use Standard Normal Probability Table

Z tables give the area in the tails of the normal distribution

- If Z = +, then the area is in the right hand tail
- If Z = - , then the area is in the left hand tail



Tips to Use Standard Normal Probability Table

1. $P(Z > a) = P(Z \geq a) = from\ table$

2. $P(Z < -a) = P(Z > a) = from\ table$

Tips to Use Standard Normal Probability Table (cont.)

3. $P(Z < a) = P(Z \leq a) = 1 - P(Z > a)$

4. $P(Z > -a) = 1 - P(Z > a) = from\ table$

Tips to Use Standard Normal Probability Table (cont.)

5. $P(a < Z < b) = P(Z > a) - P(Z > b)$

6. $P(-a < Z < -b) = P(Z > b) - P(Z > a)$

Tips to Use Standard Normal Probability Table (cont.)

7. $P(-a < Z < b) = 1 - P(Z > a) - P(Z > b)$

Exercise 1

1. If Z~N(0, 1), find

(a) $P(Z > 0.5)$

(b) $P(Z > 1.03)$

(c) $P(Z < 2.1)$

(d) $P(Z < 0.48)$

(e) $P(Z < -2.33)$

(f) $P(Z > -1.54)$

(g) $P(0.95 < Z < 1.42)$

(h) $P(-2.02 < Z < -0.55)$

(i) $P(-1.32 < Z < 2.20)$

Exercise 1 (cont.)

2. Given a standard normal distribution, find the value ok k such that:

(a) $P(Z > k) = 0.0268$

(b) $P(Z < k) = 0.8023$

Standard Normal Distribution

- **Z-score:** can use to establish the number of standard deviations from the mean and work out probabilities

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

where X_i = variable value,

μ = arithmetic mean, σ = standard deviation

Example 1

Given that *X* has the normal distribution *N(60,16)*, find:

a) $P(55<X<63)$

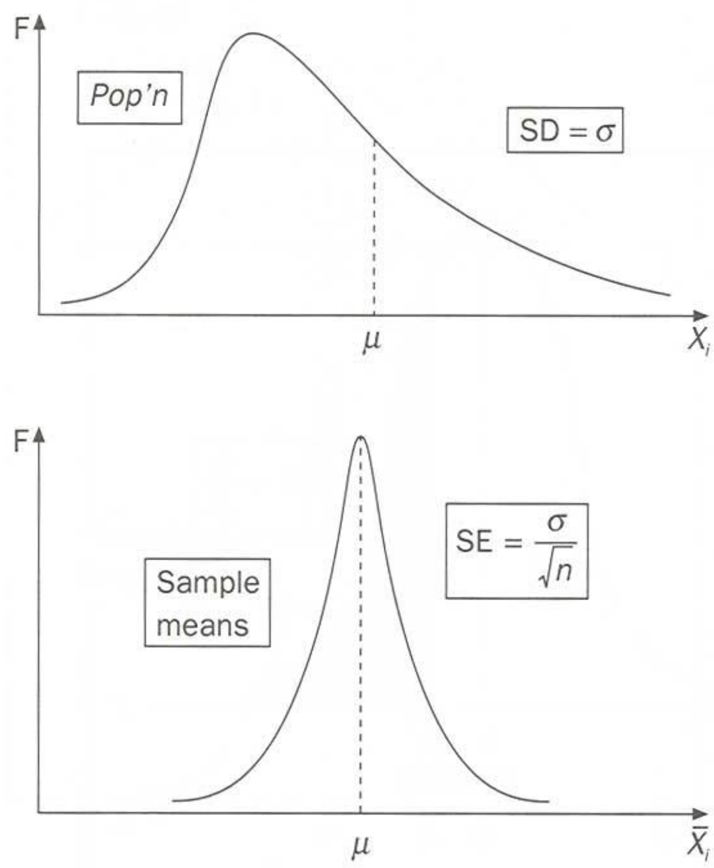
b) $P(X>49)$

Exercise 2

1. Suppose a variable has a normal distribution with a mean of 100 and a standard deviation of 10.Find the probability of an individual value (X_i) being:
   a) 115 or more
   b) 75 or less
   c) Between 75 and 115
2. The attendance at a night club is normally distributed, with a mean of 80 people and a standard deviation of 12 people. What is the probability that on any given night the attendance is:
   a) 74 people or more
   b) 83 people or less
   c) Between 70 and 85 people?

Central Limit Theorem

- Mostly, the population which the samples are selected is not normally distributed.



Central Limit Theorem

- Even if the ***population is not normal***, if sampling is ***random*** and the sample size (*n*) is **≥ *30***, then the distribution of sample *means* can be regarded as ***approximately normal*** with mean, $\mu_{\bar{x}} = \mu$ and standard deviation, $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$

$$\bar{X} \sim N(\mu_{\bar{x}}, {\sigma^2}_{\bar{x}})$$

$$\bar{X} \sim N(\mu, \frac{\sigma^2}{\sqrt{n}})$$

$$*** \ \boldsymbol{Standard\ error = \frac{\sigma}{\sqrt{n}}} \ ***$$

Central Limit Theorem

- The greater the sample size, the more confident we can be that any *sample statistic* represents the ‘population’ from which the sample is drawn
- We can then still use our Z score and Z tables for calculating probabilities for this distribution of sample means

Example 2

- The output of glass panels has a mean thickness of 4cm and a standard deviation of 1cm. If a random sample of 100 glass panels is taken, what is the probability of the sample mean having a thickness of between 3.9cm and 4.2cm?

Exercise 3

- A production line makes unit which are normally distributed with a mean weight of 200 grams and a standard deviation of 9 grams. What is the probability of a sample of 36 units having a (sample) mean weight of 203 grams or more?