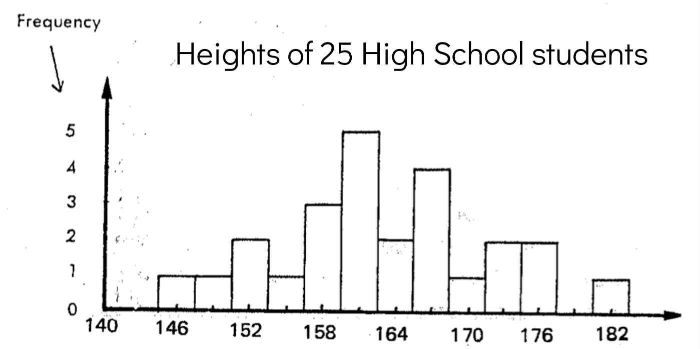
Normal Distribution

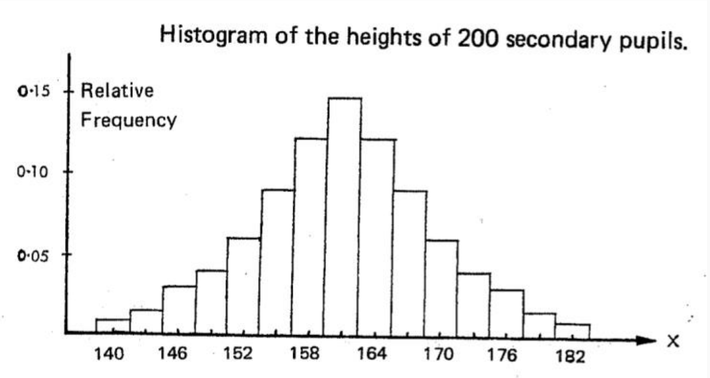
Probability Density Graph

Know this

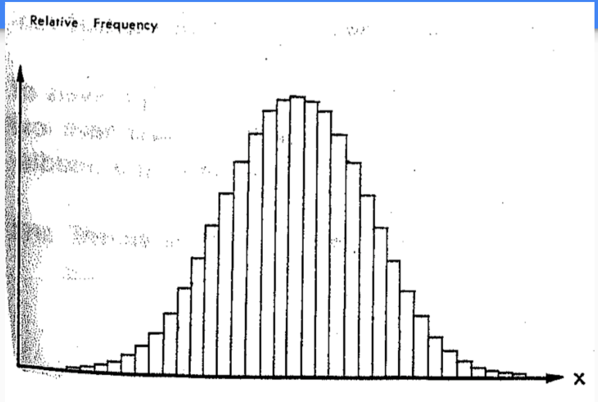
When a continuous population is sampled, the distribution of the sample can be represented on a relative frequency histogram.



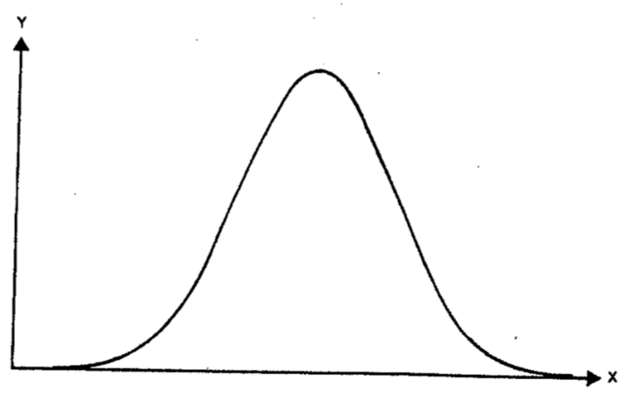
When larger samples of the population are taken, the width of the intervals on the x-axis becomes smaller. The resulting histogram represents the distribution of the population more accurately.



In the long run, as the sample size increases, the resulting graph is called a probability density graph.



A smooth curve through the tops of the rectangles would look like this:



Discuss:
What are the characteristics of this curve?
What does the curve represent?
Is it OK to use a smooth curve to represent this data?
Will all probability density graphs be this shape? Examples?

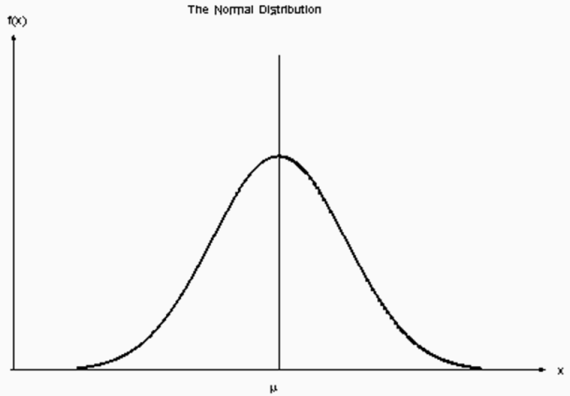
Normal Distribution

Know this

If X is normally distributed $X \sim N(\mu, \sigma^2)$

then its probability density function is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad \text{for} \quad -\infty < x < \infty.$$



The Z-distribution or Standard Normal distribution such that $Z \sim N(0,1)$ is:

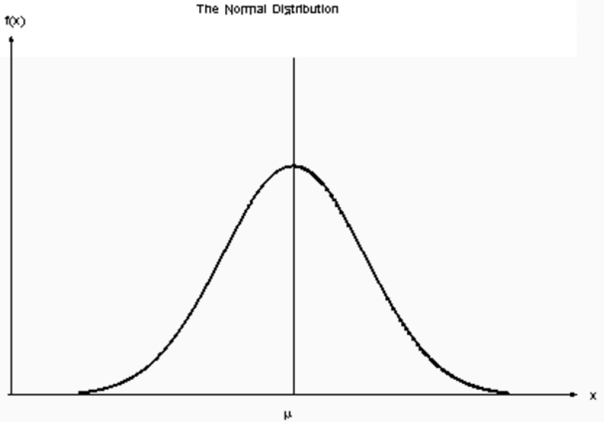
$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, \quad -\infty < z < \infty \text{ where } \mu = 0, \quad \sigma = 1$$

Notes: Probability Density Function

When a continuous population is sampled, the distribution of the sample can be represented on a relative frequency histogram.

When larger samples of the population are taken, the width of the intervals on the x-axis becomes smaller. The resulting histogram represents the distribution of the population more accurately.

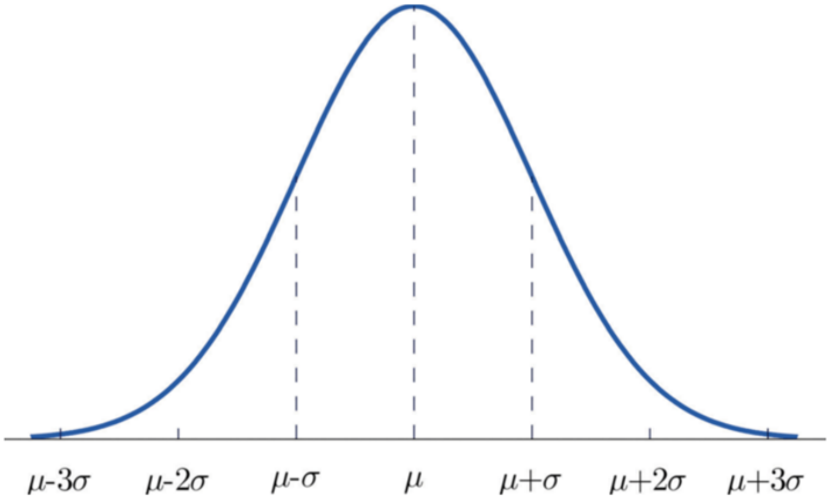
In the long run, as the sample size increases, the resulting graph is called a probability density graph.

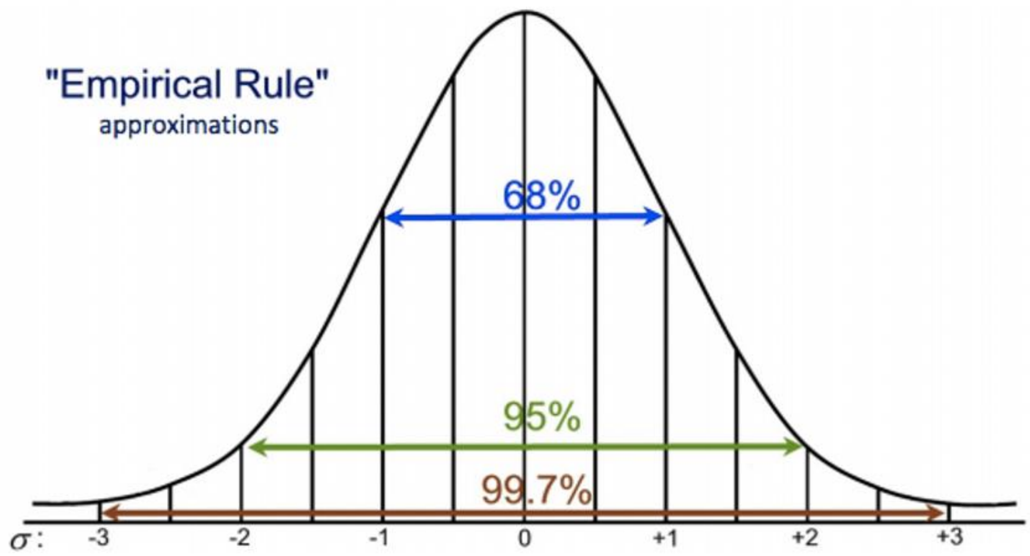


http://itl.nist.gov/div898/handbook/pmc/section5/pmc51.htm

Notes: Normal Distribution

Probabilities correspond to areas underneath the normal curve.





- 68% of the distribution lies within **one** standard deviation of the mean.
- 95% of the distribution lies within **two** standard deviations of the mean.
- 99.7% of the distribution lies within **three** standard deviations of the mean.

Approximately 68% of observations lie in the region $\mu - \sigma \le x \le \mu + \sigma$

Approximately 95% of observations lie in the region $\mu - 2\sigma \le x \le \mu + 2\sigma$

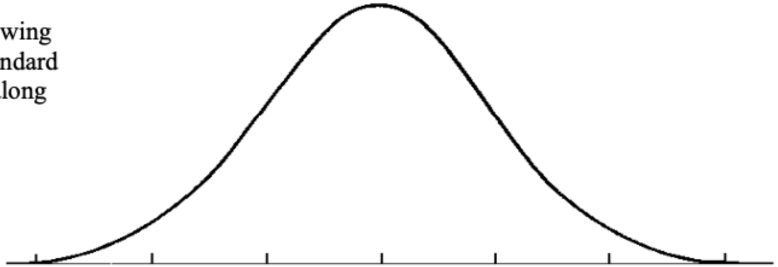
Approximately 99% of observations lie in the region $\mu - 3\sigma \le x \le \mu + 3\sigma$

Example: Using Empirical Rule

In the United States, the average height of an adult male is 5’ 9” (69 inches). Male height is distributed according to a normal distribution curve with a standard deviation of approximately 2.5 inches.

1. Label the graph at the right showing the mean and the respective standard deviation values (±1, ±2, ±3) along the bottom.

2. Indicate on the graph, the locations of 68%, 95% and 99.7% of the population, referred to as the "empirical rule".



3. What percent of heights falls below 66.5” ?

4. What percent of the population falls below 69” ?

5. What percent of the heights lies within 2 standard deviations of the mean?

6. What percent of the heights lies within 1 standard deviation above the mean?

7. A male with a height of 71.5” is taller than what percent of the population?

8. According to the empirical rule, 95% of heights fall between what two values?

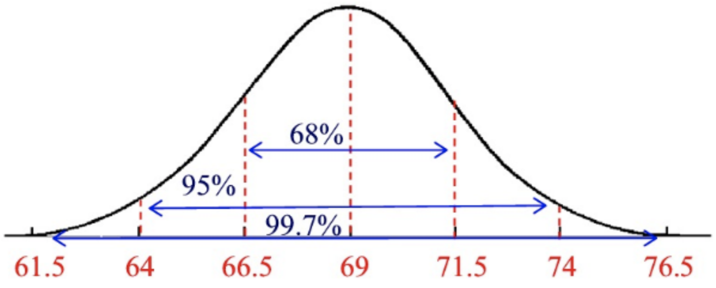
9. Of 200 males, how many men can be expected to have a height between 69” and 71.5”?

Answers

In the United States, the average height of an adult male is 5’ 9” (69 inches). Male height is distributed according to a normal distribution curve with a standard deviation of approximately 2.5 inches.

1. Label the graph at the right showing the mean and the respective standard deviation values (±1, ±2, ±3) along the bottom.

2. Indicate on the graph, the locations of 68%, 95% and 99.7% of the population, referred to as the "empirical rule".



3. What percent of heights falls below 66.5” ? 16%

4. What percent of the population falls below 69” ? 50%

5. What percent of the heights lies within 2 standard deviations of the mean? 95%

6. What percent of the heights lies within 1 standard deviation above the mean? 34%

7. A male with a height of 71.5” is taller than what percent of the population? 84%

8. According to the empirical rule, 95% of heights fall between what two values? 64” and 74”

9. Of 200 males, how many men can be expected to have a height between 69” and 71.5”? 68 men

Example: GDC

Example 1 - If $X \sim N(20,16)$, draw a diagram for each question and then find:

a) $P(X < 23)$

b) $P(X > 14)$

c) $P(16 < X < 24)$

d) Hence, without a GDC, find $P(20 < X < 24)$

Example 2 - A study showed that the mean duration of a certain strain of flu virus was 12 days with a standard deviation of 3 days. If you caught this type of flu that it would last:

a) longer than 17 days?

b) between 13 and 15 days?

Example 3 - Weights of a certain breed of cat follow a normal distribution with mean 16kg and variance 16kg^2. In a sample of 2000 such cats, estimate the number that will have a weight above 13kg.

Example 4 - When Ali participates in long-jump competitions, the lengths of her jumps are normally distributed with mean 5.2m and standard deviation 0.7m. She needs to jump 6m to qualify for the school team.

a) What is the probability she will qualify with a single jump?

b) If she is allowed three jumps, what is the probability that she will qualify for the school team?

Notes: Standard Normal Distribution

Every normal X-distribution can be transformed into the standard normal distribution, or ____________________,

using the transformation ___________________.

No matter what the parameters μ and σ of the X-distribution are, we always end up with _______________

Example: Sketch a diagram. Calculate the probabilities.

$P(Z < -1.2) =$

$P(Z < 1.2) =$

$P(Z > -1.2) =$

$P(Z > 1.2) =$

$P(-1.2 < Z < 1.2)$

Examples:

Example 4 – The scores of Chesney's most recent tests in Biology and Math are shown below (as well as the cohort mean and standard deviation). Assume the score on both tests were normally distributed, which test did Chesney do better in comparison to her classmates?

	Biology	Math
Chesney's Score	82	80
Mean	78	77
Standard Deviation	6.1	3.2

Standard Normal Distribution

Standard normal distribution is a normal curve that has a mean of zero and a standard deviation of one.

A normal distribution is transformed to the standard normal distribution using $Z = \frac{X - \mu}{\sigma}$

Inverse Normal Calculations

Example:

Scores on the SAT verbal test in recent years follow the $X \sim N(505, 110^2)$ distribution.

How high must a student score in order to place in the top 10% of all students taking the SAT?

Example:

Given that $Z \sim N(0,1)$ find *a* such that

$a)\ \ p(Z < a) = 0.5478$

$b)\ \ p(Z > a) = 0.6$

$c)\ \ p(-k \le Z \le k) = 0.95$

$d)\ \ p(-k \le Z \le k) = 0.99$