**Normal distribution:**

Normal distributions are a family of distributions that have the same general shape. They are symmetric with scores more concentrated in the middle than in the tails. The general shape is a bell-shaped density curve described by its mean μ and standard deviation, σ. The density curve is symmetrical, centered about its mean, with its spread determined by its standard deviation.
The standard deviation is a statistic that tells you how tightly all the various examples are clustered around the mean in a set of data. If this curve is flatter and more spread out, the standard deviation is to be larger.
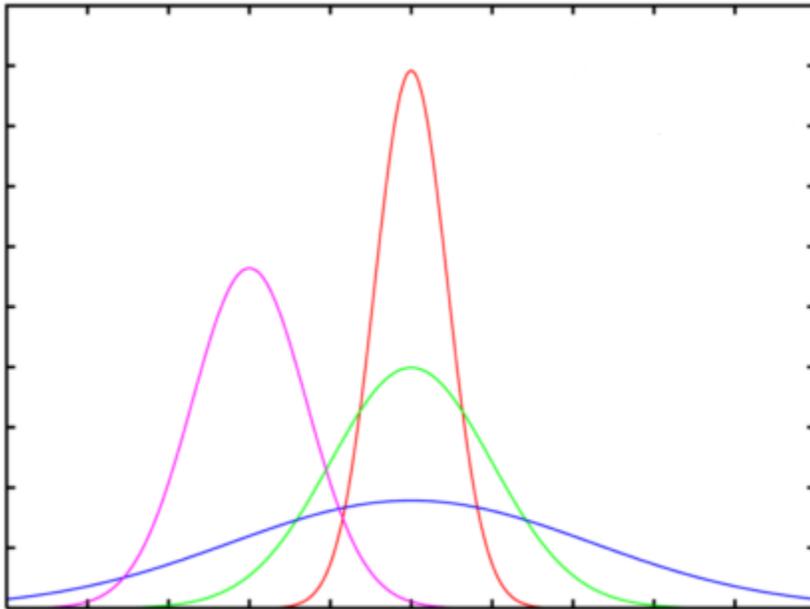
The mean is given by the x-value at the peak. In this example the red curve and the blue curve have the same mean; the lower mean corresponds to the light purple curve. The higher standard deviation corresponds to the most flatten curve, the blue one; it follows the purple cure; the smallest standard deviation corresponds to the red curve, the data appears clustered around the mean.

Normal distribution notation:

$$X \sim \mathcal{N}(\mu, \sigma)$$

It is read, "the variable X is normal distributed with mean, μ, and standard deviation, σ."

Finding a z-score, which is the number of standard deviations from the mean, use the following formula:

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

Example 1:
Jamie was told that her score on an aptitude test was 3 standard
deviations below the mean. If test scores were approximately normal
with μ=94 and σ=6, what was Jamie's score?
In symbolic form:

$$X \sim \mathcal{N}(94, 6)$$

Answer:
since the number of standard deviations from the mean is the z-score,
in this example $z = -3$ (whenever the value is below the mean it is
negative; above the mean it would be positive).
In this example we need to find x. After some algebra, solving the z-
score formula, we obtain:

$x = z \sigma + \mu$ Therefore, $x = (-3)6 + 94 = 76$.

Example 2:

Suppose $X \sim N(16.5, 0.5)$, and $x = 16$. Find and interpret the z-score.
Answer:

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

$$z = \frac{16 - 16.5}{0.5} = -1$$

The z-score tells us that $x = 16$ is one standard deviation below or to
the left of the mean.

The Empirical rule:

$https://online.stat.psu.edu/stat200/lesson/2/2.2/2.2.7$

This reading on the Empirical Rule is an extension of the previous
reading "Understanding the Normal Distribution." In the prior reading,
the goal was to develop an intuition of the interaction between
decreased probability and increased distance from the mean. In this
reading, we will practice applying the Empirical Rule to estimate
the specific probability of occurrence of a sample based on the range
of the sample, measured in standard deviations.

The graphic below is a representation of the Empirical Rule:
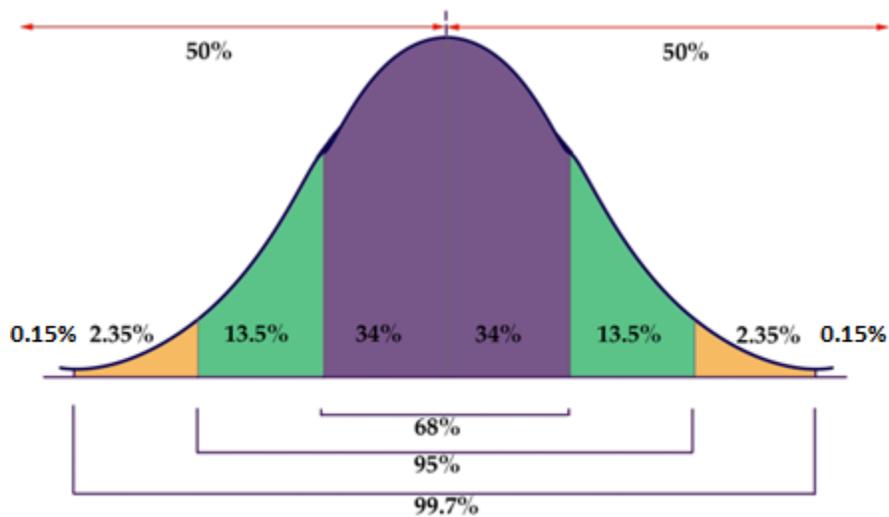
The graphic is a rather concise summary of the vital statistics of a Normal Distribution. Note how the graph resembles a bell? Now you know why the normal distribution is also called a *bell curve*.

- 50% of the data is above, and 50% below, the mean of the data
- Approximately 68% of the data occurs within 1 SD of the mean
- Approximately 95% occurs within 2 SD's of the mean
- Approximately 99.7% of the data occurs within 3 SDs of the mean

Example 3, questions based on the Empirical rule.

Suppose the IQ (intelligence test) scores are normal distributed with a mean equal to 100, and a standard deviation equal to 15; that is: $X \sim N(100, 15)$.

a) What is the probability that an individual IQ score is more than 130?

Find the z-score of 130, $z = \frac{130-100}{15} = 2$, that is, a score of 130 is 2 standard deviations above the mean. The answer is 2.5%

b) What is the probability that an individual IQ score is less than 85?

Find the z-score of 85, $z = \frac{85-100}{15} = -1$, that is, a score of 85 is 1 standard deviations below the mean. The answer is 16%

c) What is the probability that an individual IQ score lies in between 45 and 145?
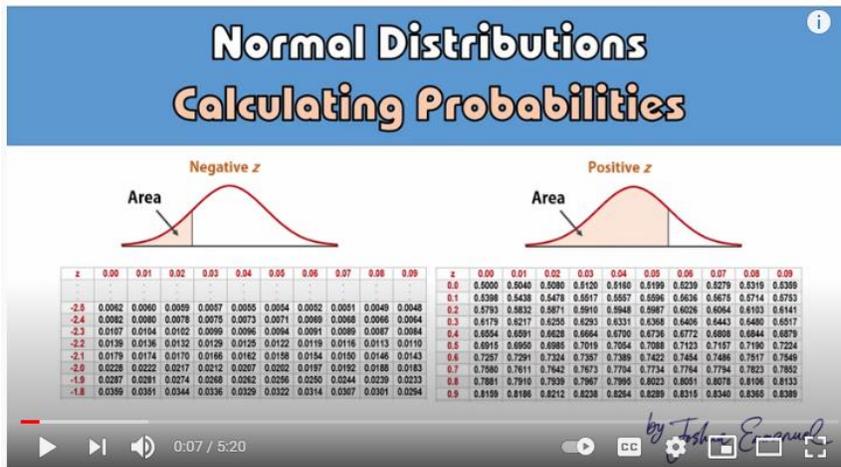
Find the corresponding z-scores:

$z = \frac{45-100}{15} = -3$ and $z = \frac{145-100}{15} = 3$

That is, three standard deviations below the mean and three standard deviations above the mean, the answer is 99.7% or almost the entire population of IQ scores.

Using tables to calculate probabilities based on the standard normal distribution:

https://youtu.be/p_KApjpyBHE

Central limit Theorem:

The central limit theorem states that if you have a population with mean μ and standard deviation σ and take sufficiently large random samples from the population then the distribution of the sample means will be approximately normally distributed.

For the random samples we take from the population, we can compute the mean of the sample means:

$$\mu_{\bar{x}} = \mu$$

and the standard deviation of the sample means:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

In plain language: the sample mean approaches (is equal to) the population mean; the standard deviation is given by dividing the population standard deviation by the square root of the sample size.

Example 4:

A manufacturing company wants to study the rate of production for lumber siding that is produced in one of their factories. Past data reveals that the production of ten boxes of lumber siding has a mean of 29 days, with standard deviation 9 days. The normal distribution for the population is shown by the dotted black line.

The company plans to take a random sample of 34 such batches of lumber siding and will calculate the mean production period of the sample to compare to the known production periods. Compute the mean and standard deviation of the sampling distribution of sample means for a sample of size 34. Round your answers to the nearest tenth.
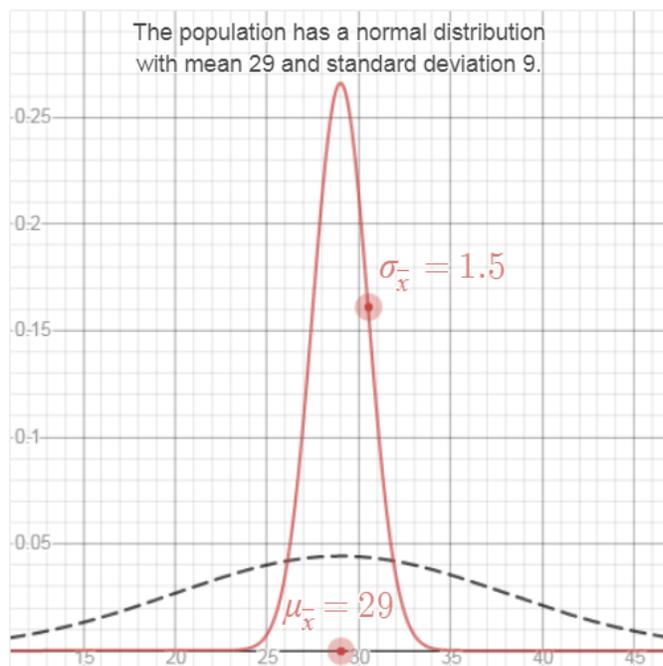
Answer:

The mean of the sampling distribution of sample means is equal to the population mean, or 29 days.

The standard deviation for a sample of size 34 is given by

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{9}{\sqrt{34}} = 1.5$$

The final graph on Knewton looks like this:
(next page)

Another question based on the previous example (4): What is the probability that the sample mean for a sample of size 34 will be more than 32?

Use tables, or on graphing calculators:

TI 84:

Press 2nd, distr (key VARS):

Select normalcdf:

```
  normalcdf
lower:
upper:
μ:
σ:
Paste
```

"More than 32" implies that we need the probability or area from 32 and beyond, up to infinity. Infinity on the calculator is indicated by E99 (E is reached by pressing 2nd and the comma key) or by entering four nines:

lower: 32

Upper: 9999

μ: 29

σ: $9/\sqrt{34}$

Or you may enter $\sigma = 1.5$ (see calculations above).
Hit enter twice, the calculator yields:
$0.0275 \approx 0.0228$ as percentage, $0.0228 \times 100 = 2.28\,\% = 2\%$ rounded to the nearest percent.

Now, **what if** the question read less instead of more, say:
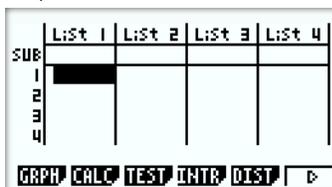What is the probability that the sample mean for a sample of size 34 will be less than 32?

In the calculator you must enter as lower, negative infinity, since less than a value corresponds to the area to the left of the given value (32 in this case), and the left extreme of the curve is negative infinity represented by -E99 or -9999; the upper is the value, 32; mean and standard deviations the same, 29 and 1.5 respectively. Another type of question is the one asking for probabilities in between values, like: What is the probability that the sample mean for a sample of size 34 will be between 31 and 32?; in that case lower is 21, upper is 31, mean 29, and the standard deviation 1.5.

On Casio graphing Calculator (Casio 9750GII):

F5, DIST:

```
    List 1 List 2 List 3 List 4
SUB
 1  ▮
 2
 3
 4
GRPH CALC TEST INTR DIST  ▷
```

F1, NORM:

Then choose F2, Ncd (We will never use Npd):

Ncd looks like this:

Update Lower, upper, Std Dev and mean as described for example 4, previous page and hit EXE.