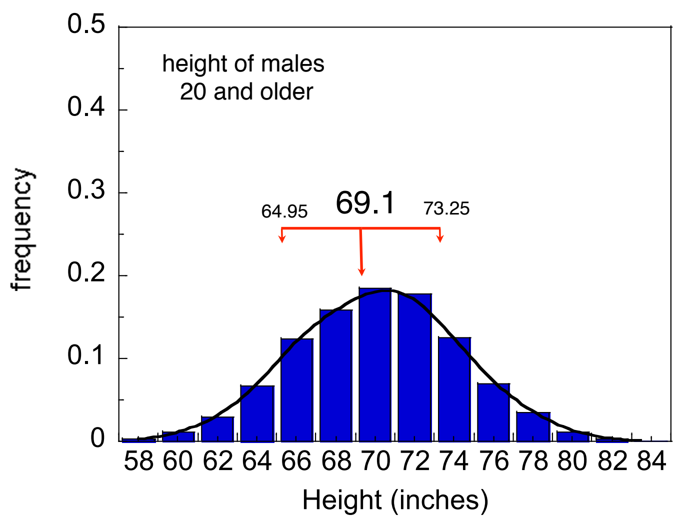
Statistics

- Let's consider a recent antropomorphic study of the US Population*: the height of males 20 and older.
- It was reported that their height was 69.1 ± 4.15 inches.
    - What does it mean?

* CDC

Statistics



Statistics

- Let' s consider a recent antropomorphic study of the US Population*: the height of males 20 and older.
- It was reported that their height was 69.1 ± 4.15 inches.

  - What does it mean?
  - Did they really measure the height of every member of the population?

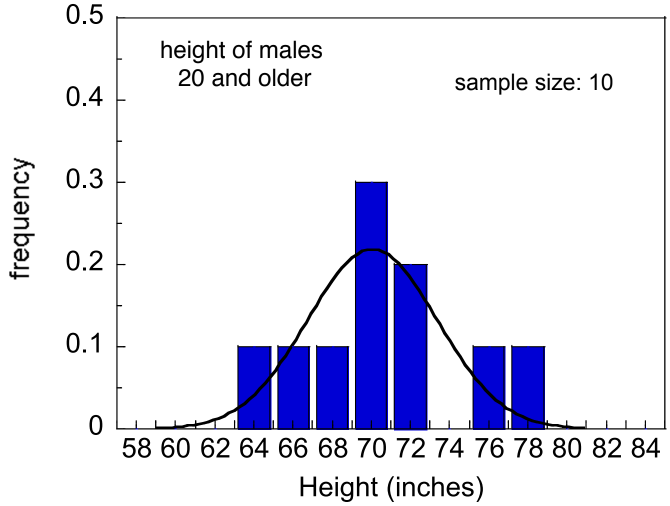
* CDC

Statistics

No!

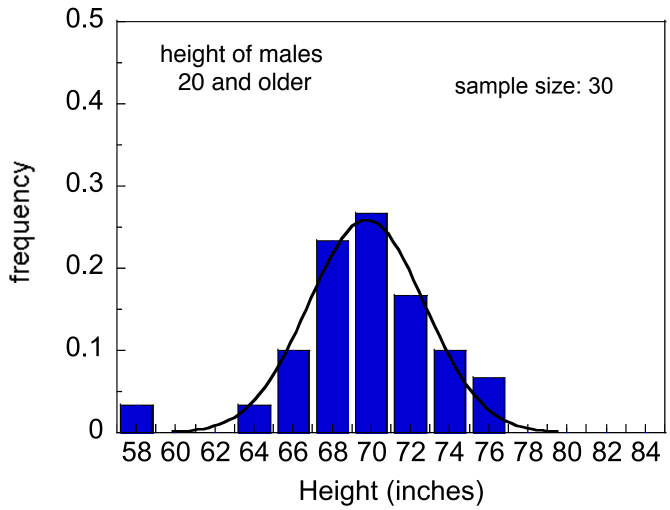
But it was necessary to sample a large segment of the population

Statistics



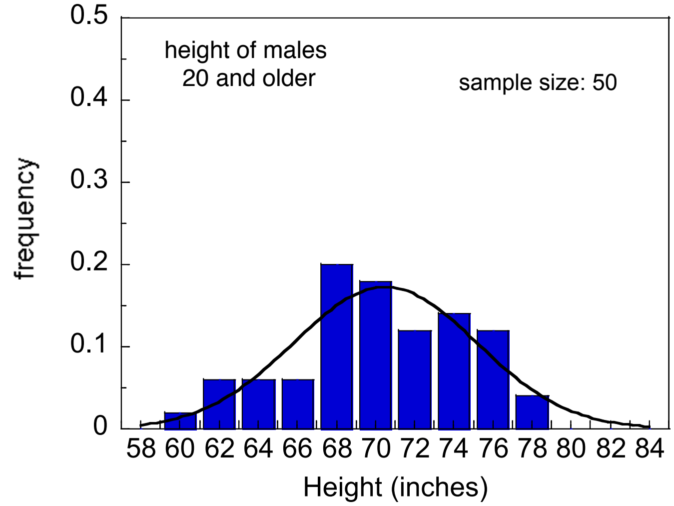
Lara-Curzio. Metodología de la Investigación. Escuela Politécnica Nacional. Quito, Ecuador 2015

Statistics

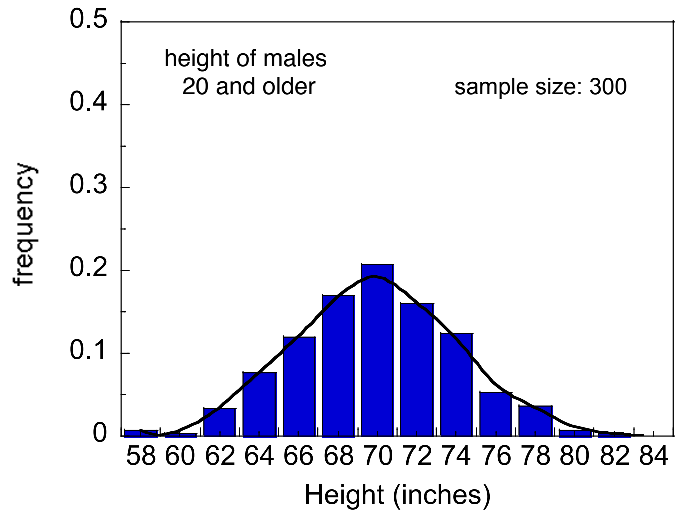


Lara-Curzio. Metodología de la Investigación. Escuela Politécnica Nacional. Quito, Ecuador 2015

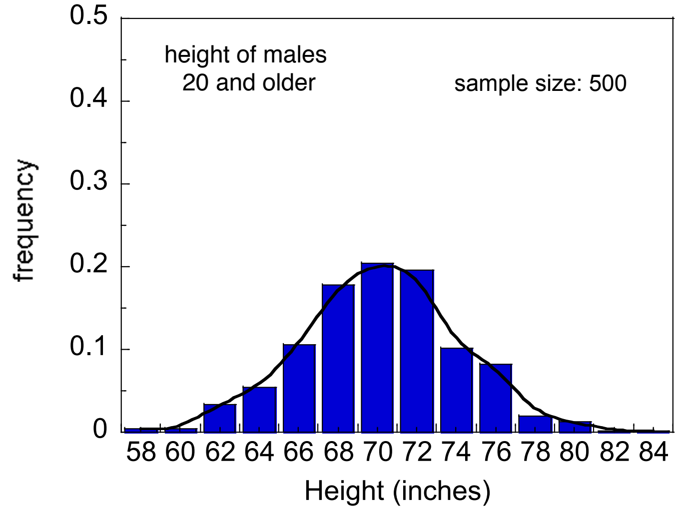
Statistics



Statistics

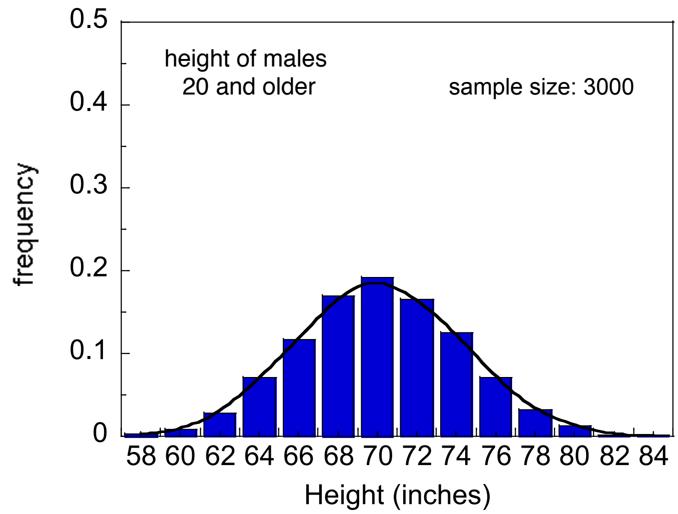


Statistics



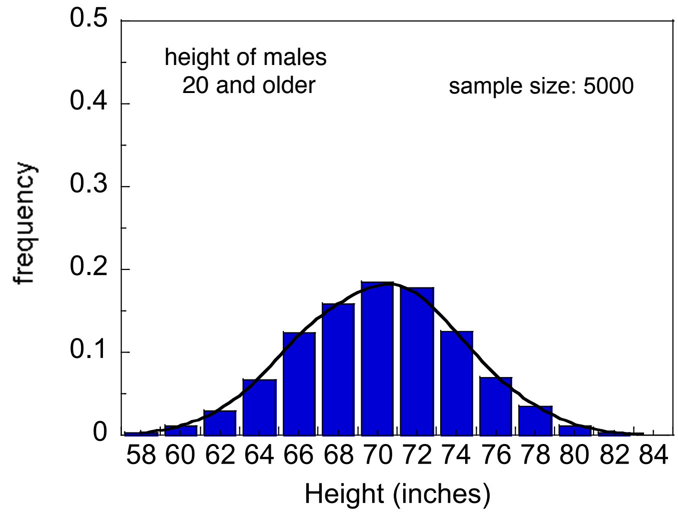
Lara-Curzio. Metodología de la Investigación. Escuela Politécnica Nacional. Quito, Ecuador 2015

Statistics

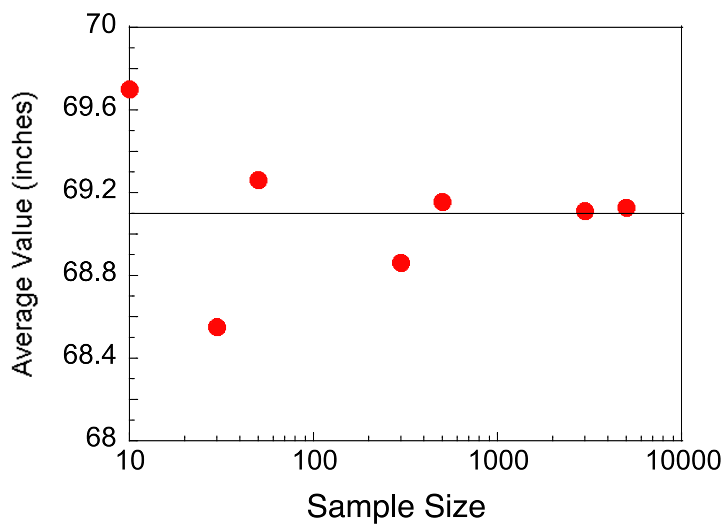


Lara-Curzio. Metodología de la Investigación. Escuela Politécnica Nacional. Quito, Ecuador 2015

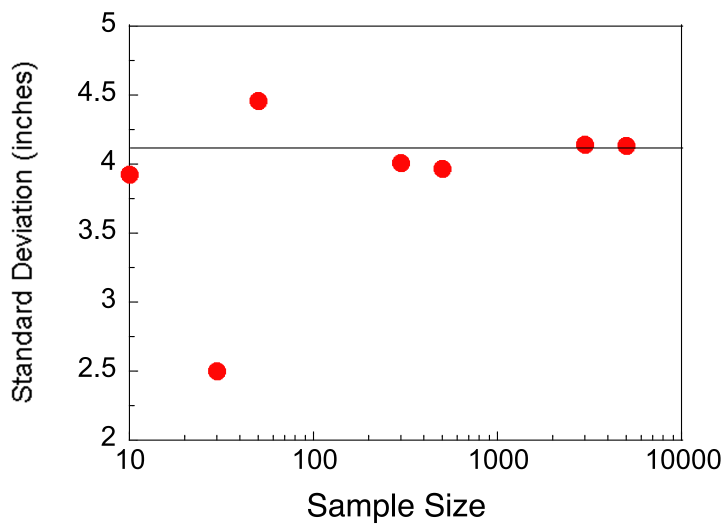
Statistics



Statistics



Statistics



Lara-Curzio. Metodología de la Investigación. Escuela Politécnica Nacional. Quito, Ecuador 2015

Statistics

The precision of our ability to estimate the parameters of the distribution increases with the sample size.

The actual value of the parameters can only be known by sampling the entire population!

Lara-Curzio. Metodología de la Investigación. Escuela Politécnica Nacional. Quito, Ecuador 2015

Statistics

The CDC measured 7,943 people.

Statistics

The CDC measured 7,943 people.

- Why using a normal distribution?

Statistics

To analyze and manipulate data, it is desirable to use mathematical representations of those data.

Those mathematical representations are known as distributions.

Statistics

The normal distribution does a good job at describing the variability of results in many practical applications.

Normal Distribution

In 1809 Gauss introduced the normal distribution, a.k.a Gaussian distribution. This distribution is characterized by two parameters, commonly known as the mean and standard deviation.

Carl Friedrich Gauss

Normal distribution

We are interested in two forms of the distribution:

- The probability density function
- The cumulative distribution function

Lara-Curzio. Metodología de la Investigación. Escuela Politécnica Nacional. Quito, Ecuador 2015

Normal distribution

pdf

$$f_{Gauss}(x;\mu,\sigma) = \frac{1}{\sigma\,\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$$

cdf

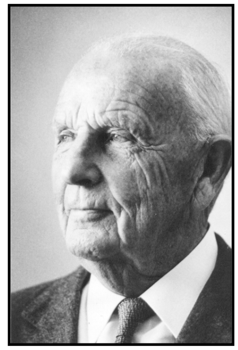
$$F_{Gauss}(x;\mu,\sigma) = \frac{1}{\sigma\,\sqrt{2\pi}}\int_{-\infty}^{x}\exp\left(-\frac{1}{2}\left(\frac{u-\mu}{\sigma}\right)^2\right)du$$

Distributions

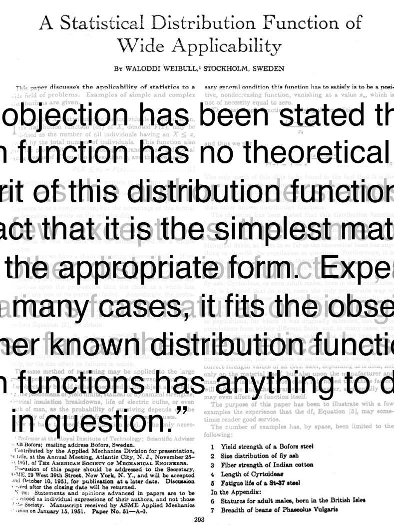
However, the normal distribution is not the only distribution available. There are dozens of distributions that are used in science and engineering every day.

Weibull distribution

One of those distributions is the Weibull distribution, which was introduced by the Swedish engineer Waloddi Weibull in 1951.



Weibull's famous paper



A Statistical Distribution Function of Wide Applicability

By WALODDI WEIBULL, STOCKHOLM, SWEDEN

1 Yield strength of a Bofors steel
2 Size distribution of fly ash
3 Fiber strength of Indian cotton
4 Length of Cyrtoideae
5 Fatigue life of a St-37 steel
In the Appendix:
6 Statures for adult males, born in the British Isles
7 Breadth of beans of Phaseolus Vulgaris

"The only objection has been stated that this distribution function has no theoretical basis.

"The only merit of this distribution function is to be found in the fact that it is the simplest mathematical expression of the appropriate form. Experience has shown that, in many cases, it fits the observations better than other known distribution functions."

distribution functions has anything to do with the population in question."

Weibull distribution

pdf

$$f_{Weibull}(x;m,\sigma_o) = \frac{m}{\sigma_o^m} x^{m-1} \exp\left(-\left(\frac{x}{\sigma_o}\right)^m\right)$$

cdf

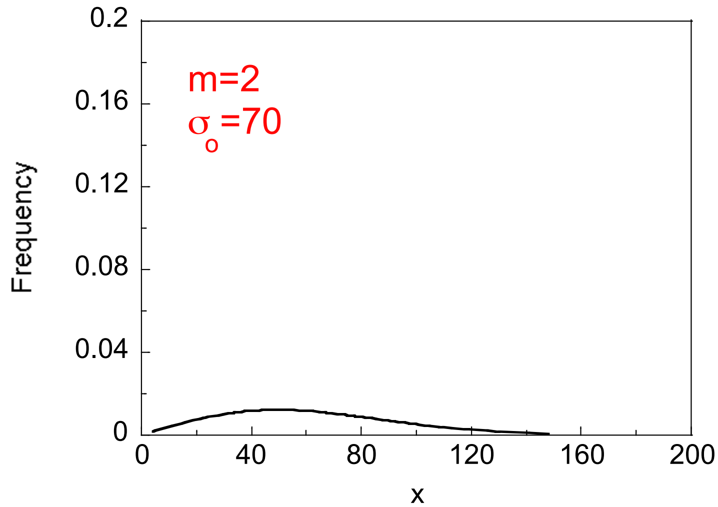
$$F_{Weibull}(x;m,\sigma_o) = 1 - \exp\left(-\left(\frac{x}{\sigma_o}\right)^m\right)$$

Weibull distribution

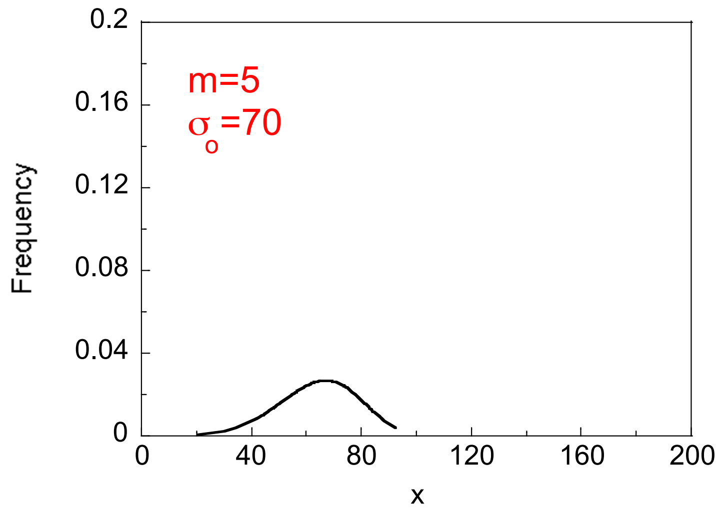
The parameters of the Weibull distribution are *m* and σ_o.

- m: Weibull modulus
- σ_o: characteristic value

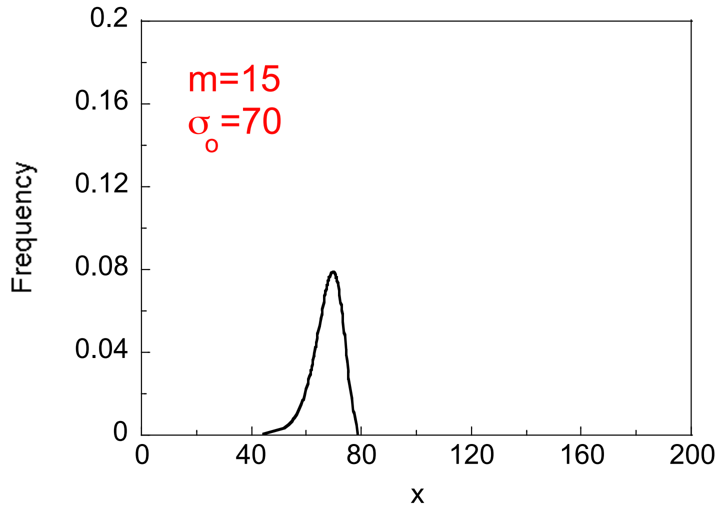
Statistics



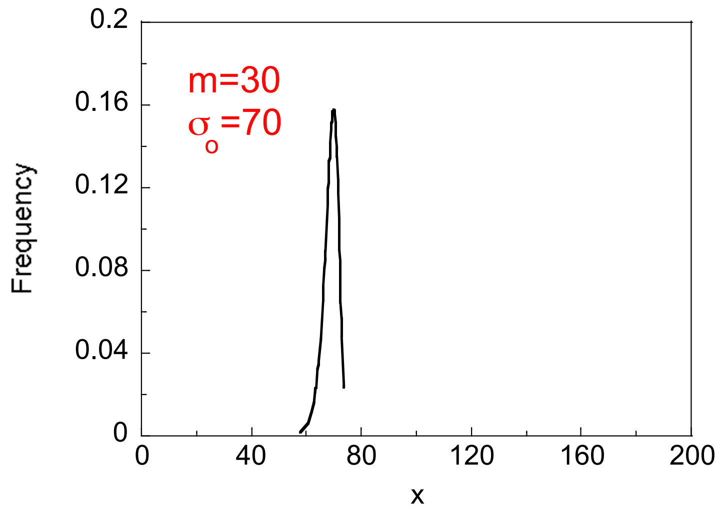
Weibull distribution



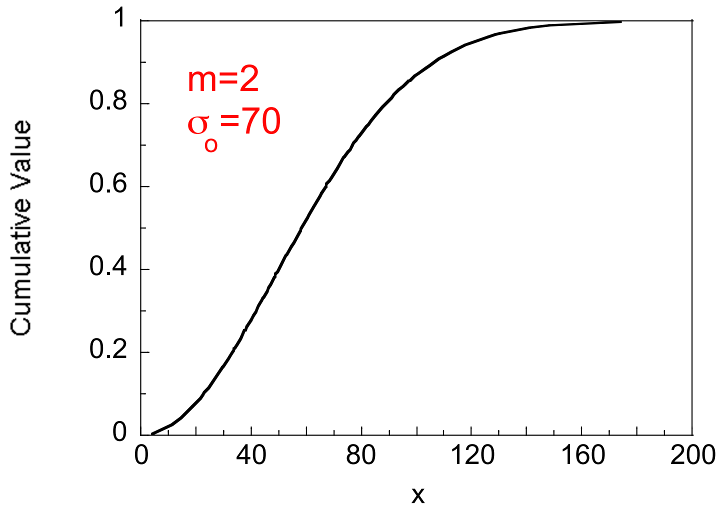
Weibull distribution



Weibull distribution

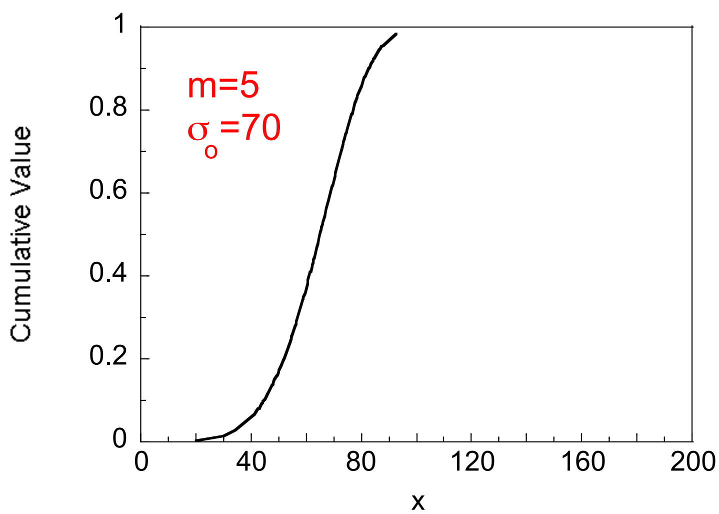


Weibull distribution



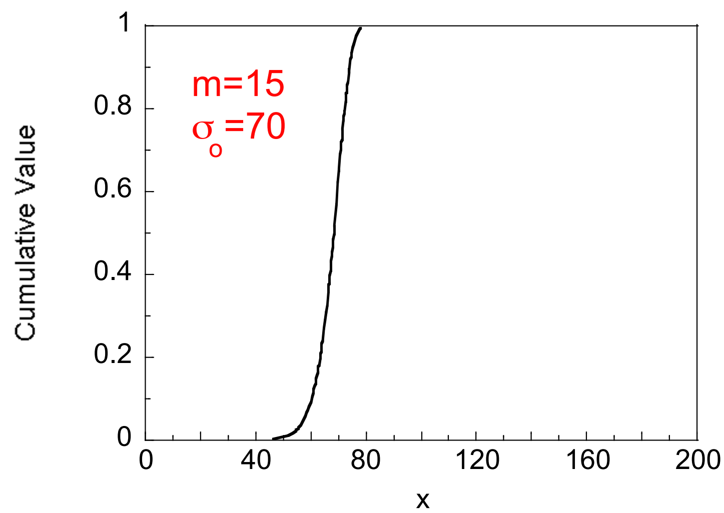
Lara-Curzio. Metodología de la Investigación. Escuela Politécnica Nacional. Quito, Ecuador 2015

Weibull distribution

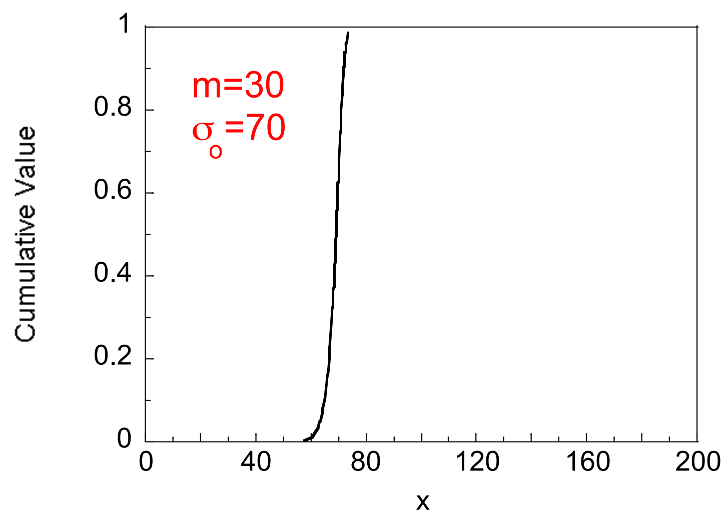


Lara-Curzio. Metodología de la Investigación. Escuela Politécnica Nacional. Quito, Ecuador 2015

Weibull distribution



Weibull distribution



Weibull distribution

The Weibull modulus, m, provides a measure of the spread of the distribution.

As m increases, the width of the distribution decreases. In the limit, it describes a process that is deterministic.

Weibull distribution

Weibull demonstrated the applicability of his distribution for:

- Yield strength of a Bofors® steel
- Size distribution of fly ash.
- Fiber strength of Indian cotton.
- Length of cytoidea (Worm length for ancient sedimentary deposits).
- Fatigue life of a St-37 steel
- Statures for adult males, born in the British Isles

Weibull distribution

The Weibull distribution is consistent with the weakest-link principle when the distribution is used to characterize the strength of materials and the dependence of strength on volume

$$\frac{\sigma_2}{\sigma_1} = \left(\frac{V_1}{V_2}\right)^{\frac{1}{m}}$$

Weibull distribution

The Weibull distribution describes the strength of most ceramic and metallic materials.

Ceramics: $3 < m < 15$

Metals: $m > 30$

Lara-Curzio. Metodología de la Investigación. Escuela Politécnica Nacional. Quito, Ecuador 2015