Sampling Error

Data Science for Quality Management: Sampling Distributions, Error and Estimation

with **Wendy Martin**

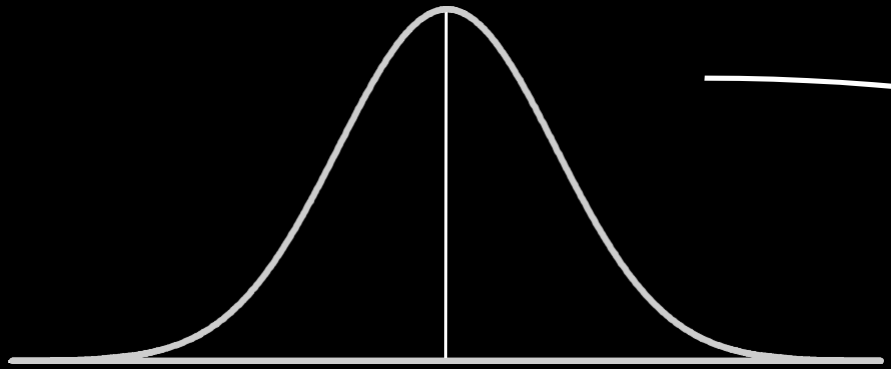
Learning objectives:

Describe the concept of sampling error

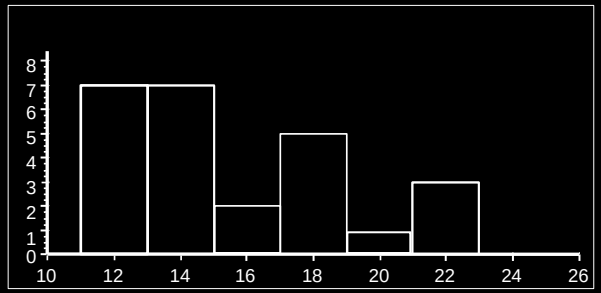
Create a vector of normally distributed random numbers

Sampling Distributions and Estimation

Are you ***really*** interested in the sample?



Population



Sample

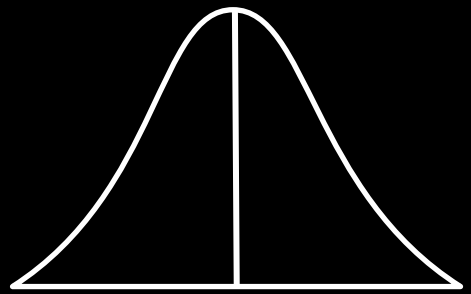
Parameters and Statistics

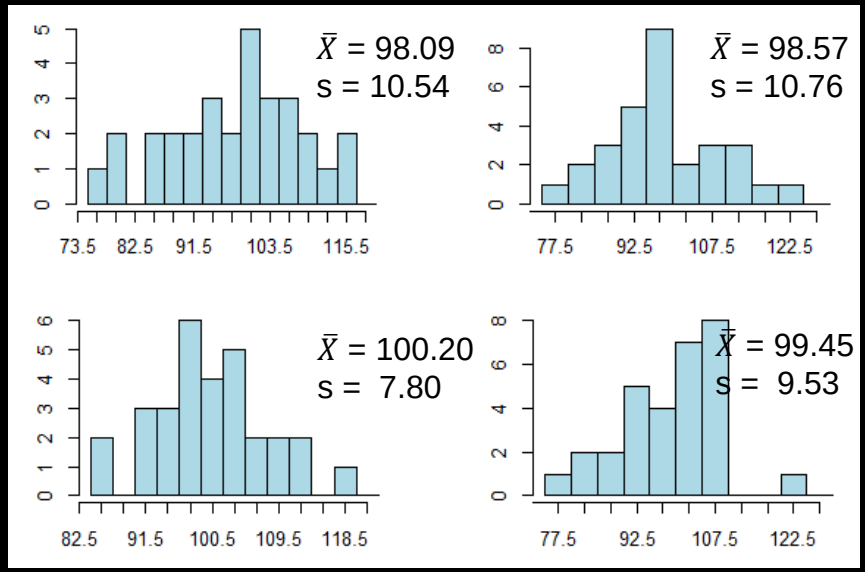
	Sample	Population
Definitions	Subgroup or portion of the population chosen for evaluation or study	Collection of all items produced or considered
Characteristics	Statistics	Parameters
Size	n	N
Mean	$\bar{X}$	μ
Median	$\tilde{X}$	M
Standard Deviation	s	σ
Variance	s^2	σ^2
Skewness	g_3	γ_3
Kurtosis	g_4	γ_4
Proportion	p	π
Rate	$\bar{c}$	λ

Sampling Distributions and Statistical Inference

μ = 100

σ = 10





Creating Random Numbers in R

In R / Rstudio:

> rnorm( )

> rexp( )

> rpois( )

> rbinom( )

Sampling Error

- Repeated samples may not be identical
- Descriptive statistics calculated from repeated sampling (with replacement) will not be exactly the same, even though the population is unchanged.

Sampling Error

- This is an expected phenomenon since we are not measuring all of the subjects or units for the entire population.
- Statistical methods allow us to account for sampling error, and make appropriate decisions.

Sampling Error

- In spite of the presence of sampling error, random sampling allow us to use sample statistics as point estimators of population parameters; however

Sampling Error

- Even when unbiased, sample statistics will probably not exactly equal their associated true population parameters.

Sampling Error

- An observed difference between a true parameter value and its associated sample descriptive statistic is caused by sampling error.

Sampling Error Defined

- The expected and quantifiable discrepancy between a population parameter and its associated descriptive statistic due to the sample size employed, and in the case of some descriptive statistics, the variability of the population.

Sampling Error & Probability

- Sampling Error is quantifiable using Random Sampling Distributions (RSDs).

- These distributions, like all probability distributions, are based on the principles of classical probability.

Sources

- Luftig, J. An Introduction to Statistical Process Control & Capability. Luftig & Associates, Inc. Farmington Hills, MI, 1982