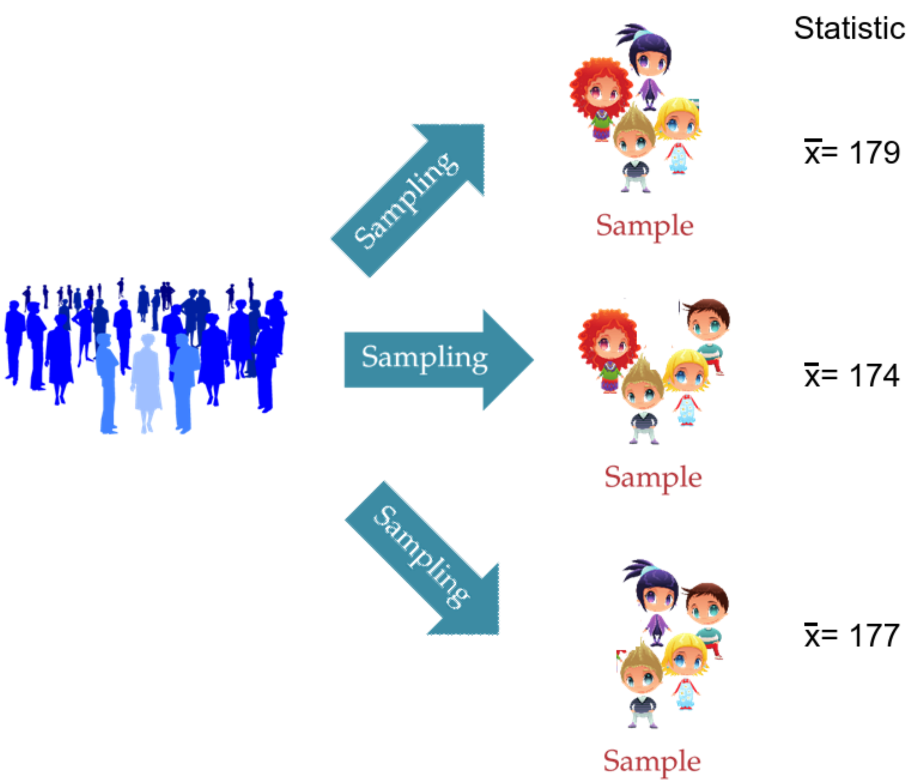
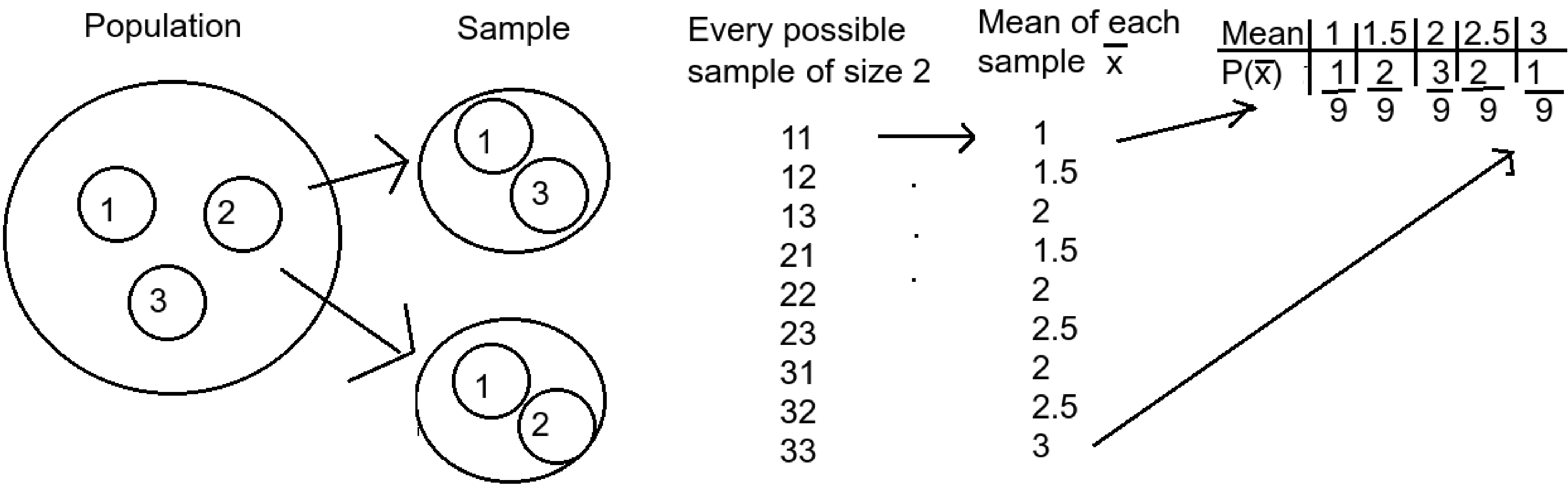
Sampling Distributions



- The mean of a sample might be different for each sample, therefore we can think of the sample mean as random variable.
- How are the sample mean and variance related to the population mean and variance.

Calculating the sample distribution



- We calculate the population mean and variance and the of the sampling distribution:

Population : $\mu = (1+2+3)/3=2$ $\sigma^2 = ((1-2)^2+(2-2)^2+(3-2)^2)/3 = 2/3$

Sample Dist.: $\mu_{\bar{x}} = 1\frac{1}{9}+1.5\frac{2}{9}+2\frac{3}{9}+2.5\frac{2}{9}+3\frac{1}{9}=2$ $\sigma^2_{\bar{x}} = (1-2)^2\frac{1}{9}+(1.5-2)^2\frac{2}{9}+(2-2)^2\frac{3}{9}+(2.5-2)^2\frac{2}{9}+(3-2)^2\frac{1}{9}=\frac{1}{3}$

For a better explanation: https://www.youtube.com/watch?v=z0Ry_3_qhDw

Relation between population and sample

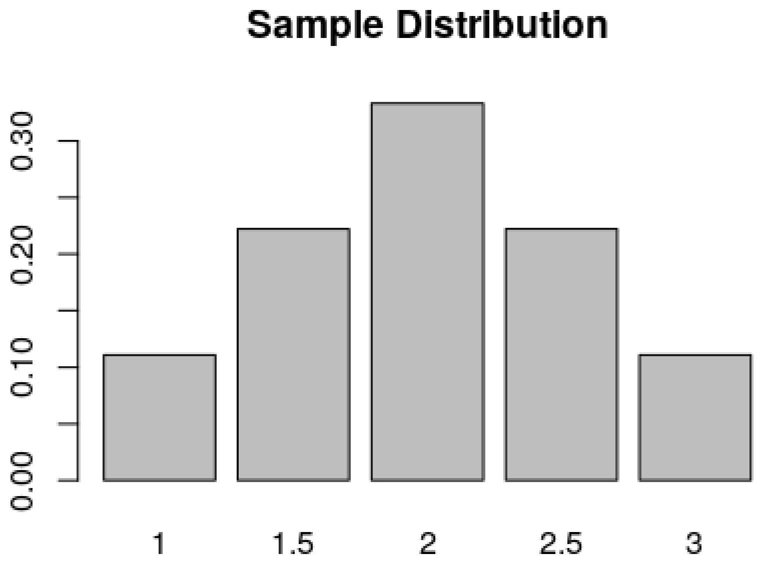
Suppose random samples of size n are drawn from a population with mean μ and standard deviation σ. The mean $\mu_{\overline{X}}$ and standard deviation $\sigma_{\overline{X}}$ of the sample mean $\overline{x}$ satisfy

$$\mu_{\overline{X}} = \mu \quad \text{and} \quad \sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$$

- If we could take every possible sample from the population, this distribution would center around the population mean. Averages computed from samples vary less than individual measurements on the population

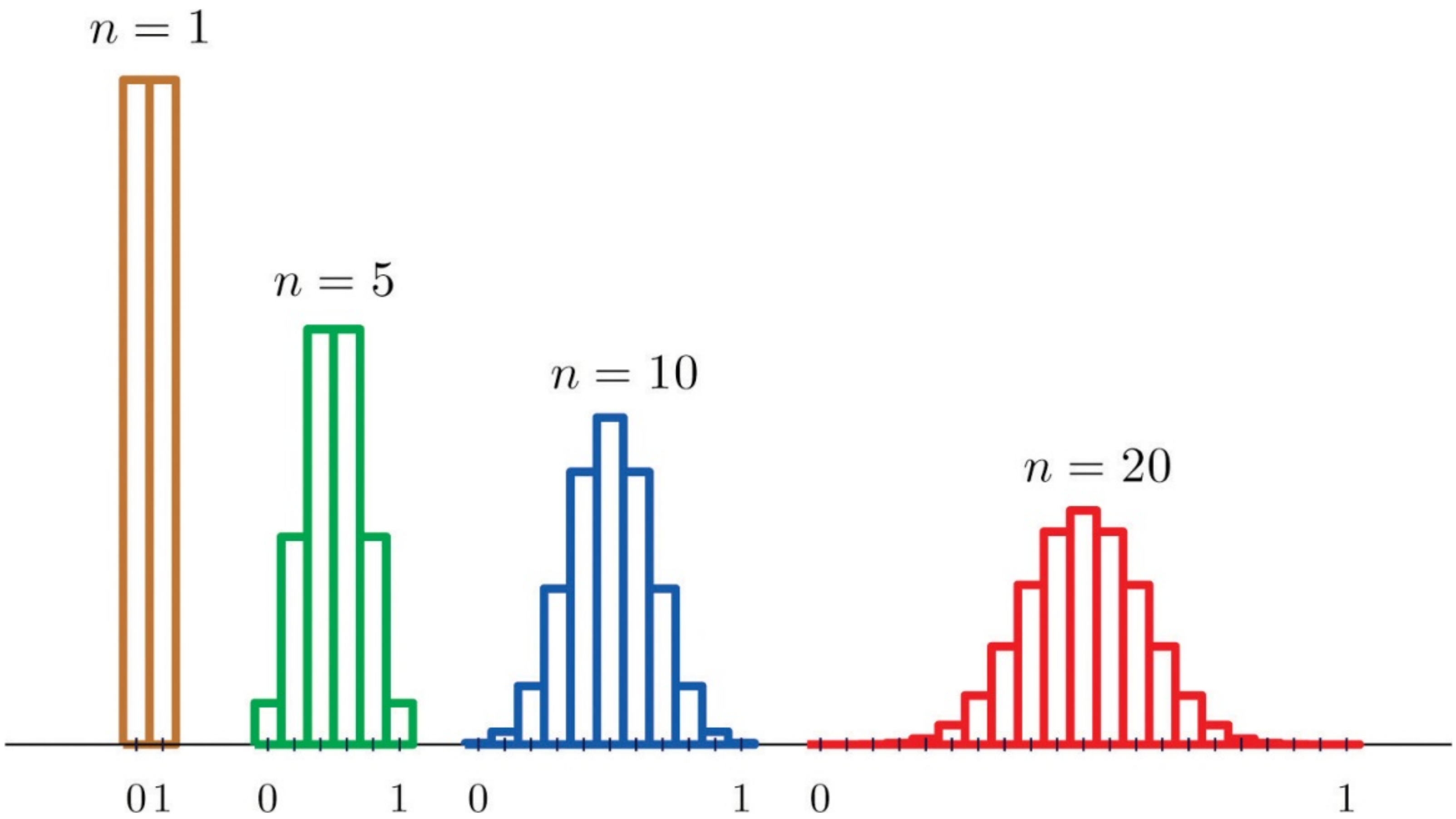
The central limit theorem

- Although, the population distribution was uniform, the sampling distribution becomes bell shaped.



The central limit theorem

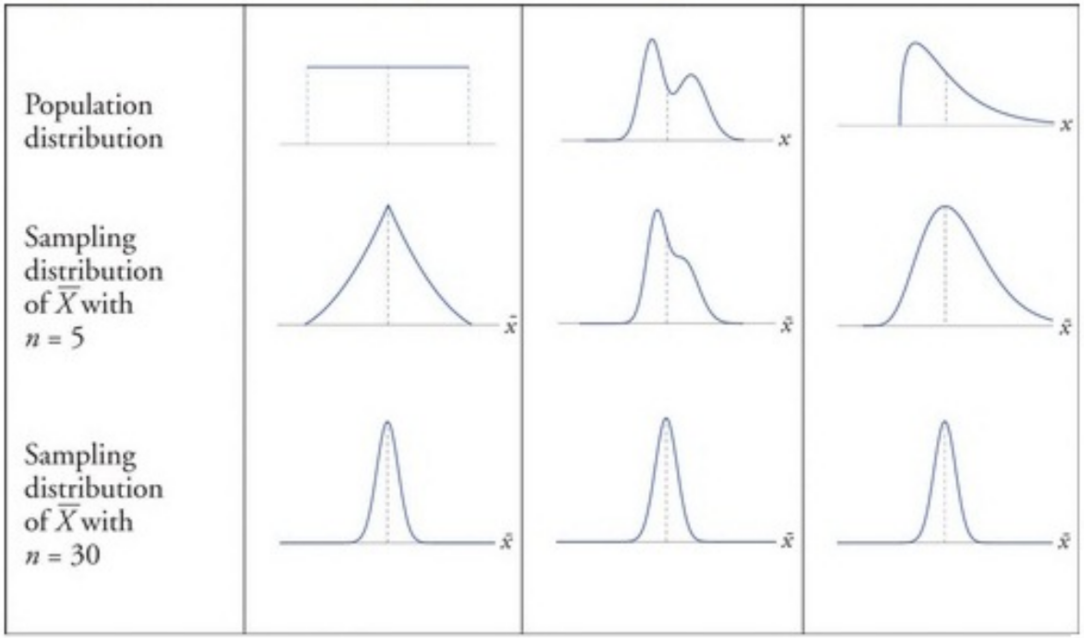
- With an increase in sample size, the sampling distribution becomes more bell shaped.



The Central Limit Theorem

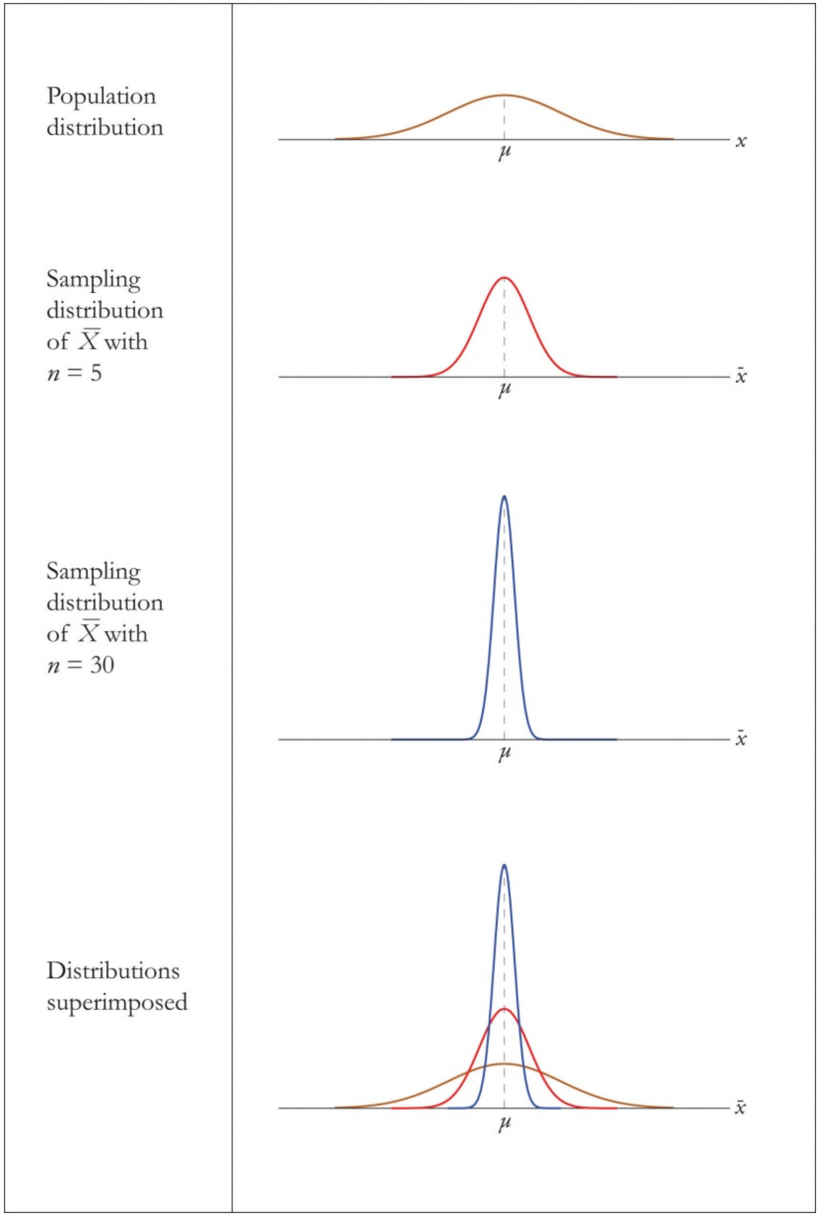
The Central Limit Theorem

For samples of size 30 or more, the sample mean is approximately normally distributed, with mean $\mu_{\overline{X}} = \mu$ and standard deviation $\sigma_{\overline{X}} = \sigma/\sqrt{n}$, where n is the sample size. The larger the sample size, the better the approximation.



Normally distributed populations

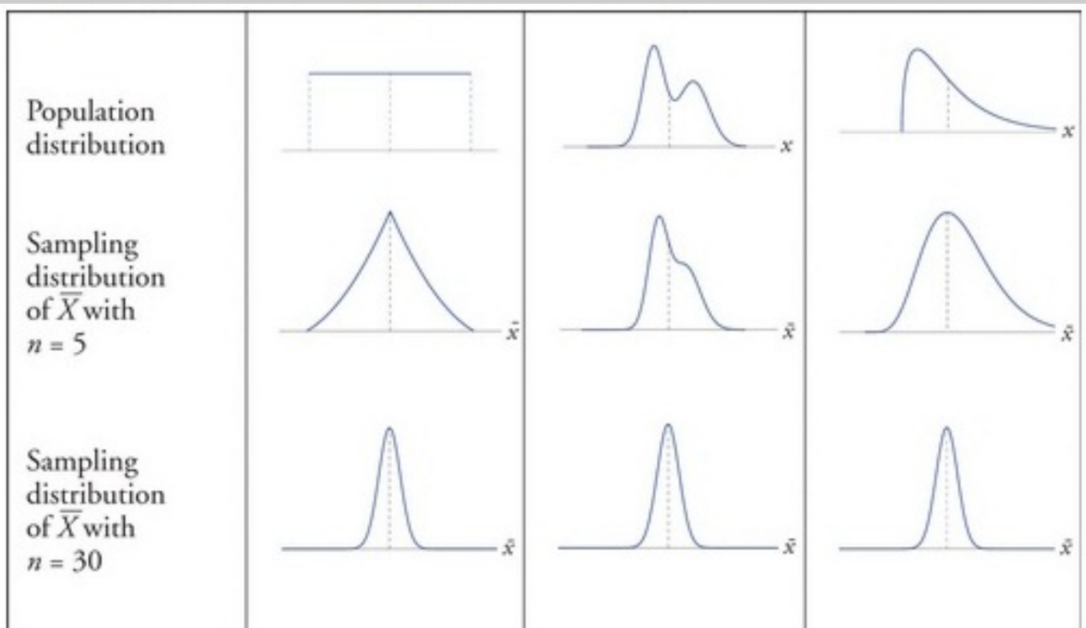
- If the population is distribution is normal, then the sampling distribution is also normal, regardless of the sample size.
- Well, its apparently normal.



The Central Limit Theorem

The Central Limit Theorem

For samples of size 30 or more, the sample mean is approximately normally distributed, with mean $\mu_{\overline{X}} = \mu$ and standard deviation $\sigma_{\overline{X}} = \sigma/\sqrt{n}$, where n is the sample size. The larger the sample size, the better the approximation.



Example

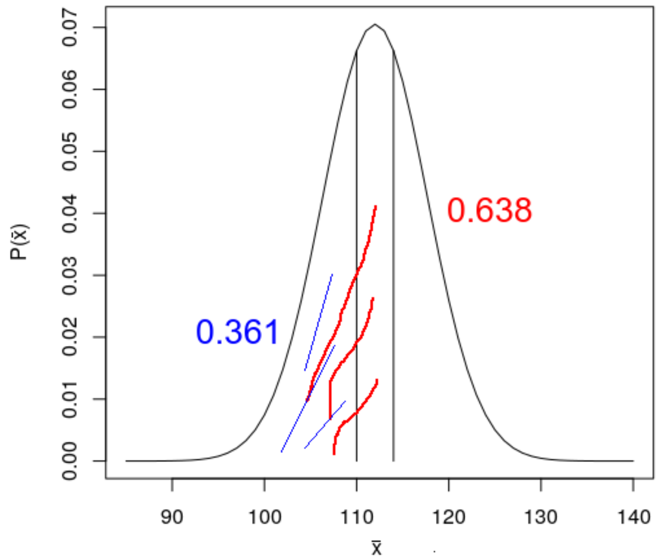
- Let $\bar{x}$ be the mean of a random sample of size 50 drawn from a population with mean 112 and standard deviation 40.
- Find the mean and standard deviation of $\bar{x}$.

$\mu_{\bar{x}} = \mu = 112$ $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{40}{\sqrt{50}} = 5.66$

- Find the probability that $\bar{x}$ assumes values 110 and 114.
- $P(110<\bar{X}<114) = 0.638-0.361=0.277$

- Note, that we can make statements

  about $\bar{X}$ but not about X.



R Exercise

- Will this distribution have a normal sample distribution ?