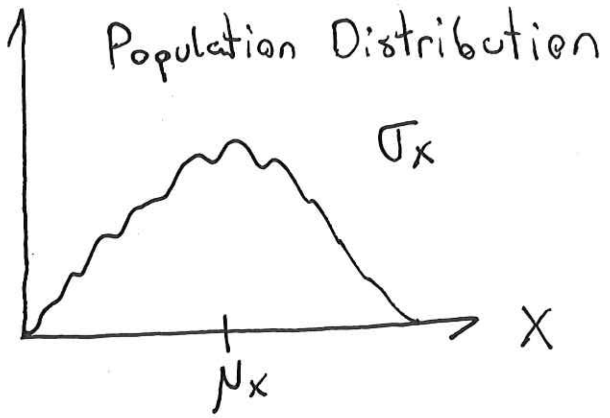
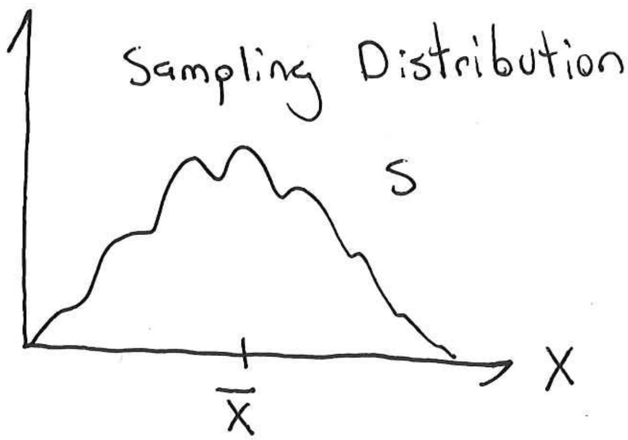
Many Hypotheses can be formulated as statements about population means. How can we estimate a population mean? How accurate is our estimate?



- Population mean (μ_X) and population standard deviation (σ_X) are <u>fixed parameters.</u>

random sampling n times

- sample mean ($\bar{x}$) and sample standard deviation (s) are <u>random variables.</u>

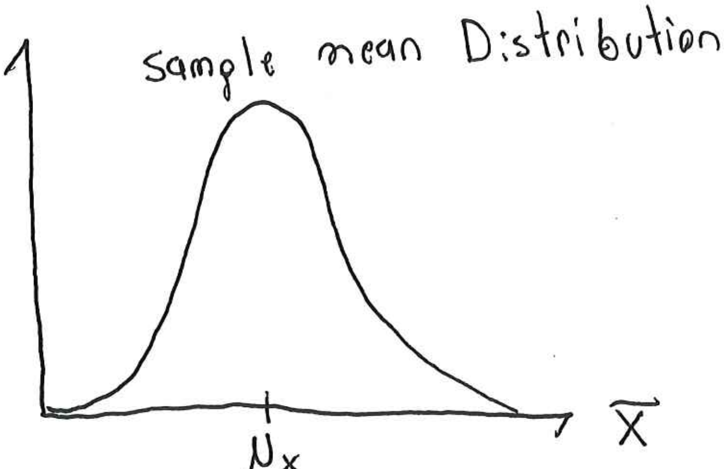


$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

CLT

$$\bar{x} \sim N\left(\mu_X, \frac{\sigma_X}{\sqrt{n}}\right)$$



$$\frac{\bar{x} - \mu_X}{SD(\bar{x})} \sim N(0,1)$$

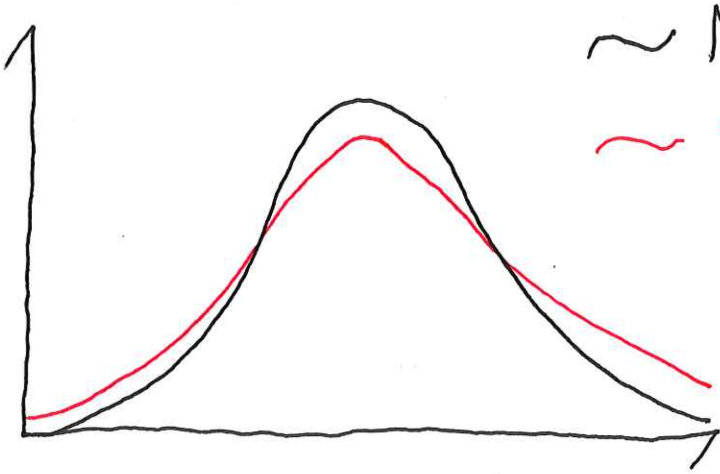
$\searrow$ Z-score

$$SD(\bar{x}) = \frac{\sigma_X}{\sqrt{n}}$$

What is the accuracy of our population mean estimator ($\bar{x}$) <u>if we don't know σ_x</u>?

$$
\begin{array}{cccc}
\sigma_x & SD(\bar{x}) & \dfrac{\bar{x}-\mu_x}{SD(\bar{x})} & \sim N(0,1) \\
\downarrow & \downarrow & \downarrow & \downarrow \\
s & SE(\bar{x}) & \dfrac{\bar{x}-\mu_x}{SE(\bar{x})} & \sim t_{n-1}
\end{array}
$$

- population standard deviation is estimated by sample standard deviation. $[\sigma_x \to s]$
- standard deviation of the estimator is replaced by Standard Error of the estimator $[SD(\bar{x}) \to SE(\bar{x})]$
- z-score is replaced by t-score $\left[\dfrac{\bar{x}-\mu_x}{SD(\bar{x})} \to \dfrac{\bar{x}-\mu_x}{SE(\bar{x})}\right]$
- Normal distribution is replaced by t distribution



$\sim N(0,1)$

$\sim t_{n-1}$

- t distro is fatter to reflect additional uncertainty
- $\lim_{n \to \infty} t_{n-1} = N(0,1)$