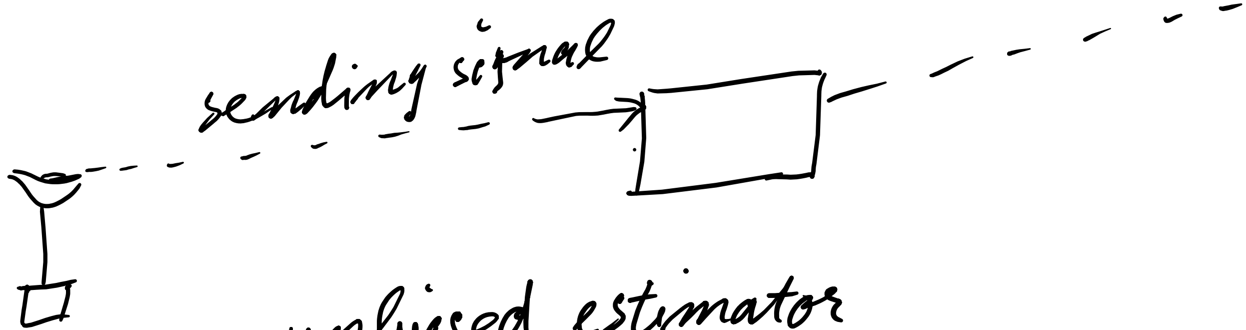
Lecture 11



$$\mu_{\bar{X}} = E(\bar{X}) = \mu$$

unbiased estimator

μ → population mean

biased estimator

$$s^2 = \sigma^2_{\bar{X}} = \frac{(X_1-\bar{X})^2 + (X_2-\bar{X})^2 + \dots + (X_n-\bar{X})^2}{n}$$

$$= \frac{n-1}{n}\sigma^2$$

σ^2 → population variance

$$\underbrace{E(\sigma^2_{\bar{X}})} = E\left(\frac{n-1}{n}\sigma^2\right) = \frac{n-1}{n}\cdot\sigma^2$$

$$E(\sigma^2_{\bar{X}}) \neq \sigma^2$$

$$E(\hat{s}^2) = \sigma^2$$

$$\hat{s}^2 = \frac{n}{n-1}s^2 = \frac{\cancel{n}}{\cancel{n-1}}\cdot\frac{\cancel{n-1}}{\cancel{n}}\cdot\sigma^2$$

$$E(\hat{s}^2) = E(\sigma^2) = \sigma^2$$

$\hat{s}^2$ → unbiased estimator

$$
\begin{tikzcd}[column sep=small]
 & & & P \arrow[r] \arrow[dr] \arrow[dll] \arrow[drr] & \mu & & \\
 & S_1 \arrow[dl] \arrow[dr] & & & \sigma^2 & S_2 \arrow[dl] \arrow[dr] & \\
\mu_{S_1} \arrow[d, phantom, "\parallel"] & & \sigma^2_{S_1} & & \mu_{S_2} \arrow[d, phantom, "\parallel"] & & \sigma^2_{S_2} \\
\mu & & & & \mu & &
\end{tikzcd}
$$

If $\sigma^2_{S_1} < \sigma^2_{S_2}$

$\rightarrow$ S_1 is a more efficient estimator

* Unbiased estimator of standard deviation:

- Normal distribution:

$$E(s) = \sqrt{\frac{2}{n-1}} \cdot \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)} \cdot \sigma$$

$$\hat{s} \,! \rightarrow E(\hat{s}) = \sigma$$

- other distribution:

$$\hat{s} = \sqrt{\frac{1}{n - 1{,}5 - \frac{1}{4}\gamma_2} \sum_{i=1}^{n} (x_i - \bar{x})^2}$$

γ_2: population excess kurtosis

In-class exercise:

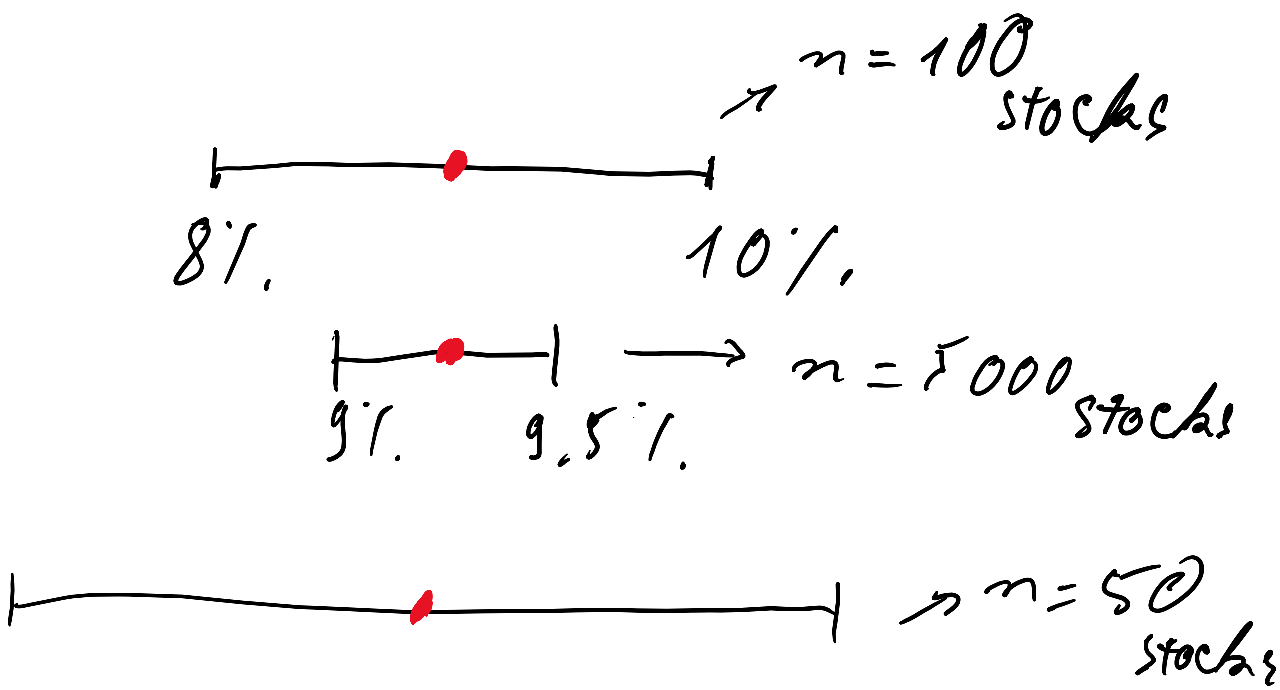
Unbiased estimator of mean: 67.45 inches

" variance: $\frac{100}{99}(8.5275)$

$$= 8.61$$
$$= \hat{s}^2$$

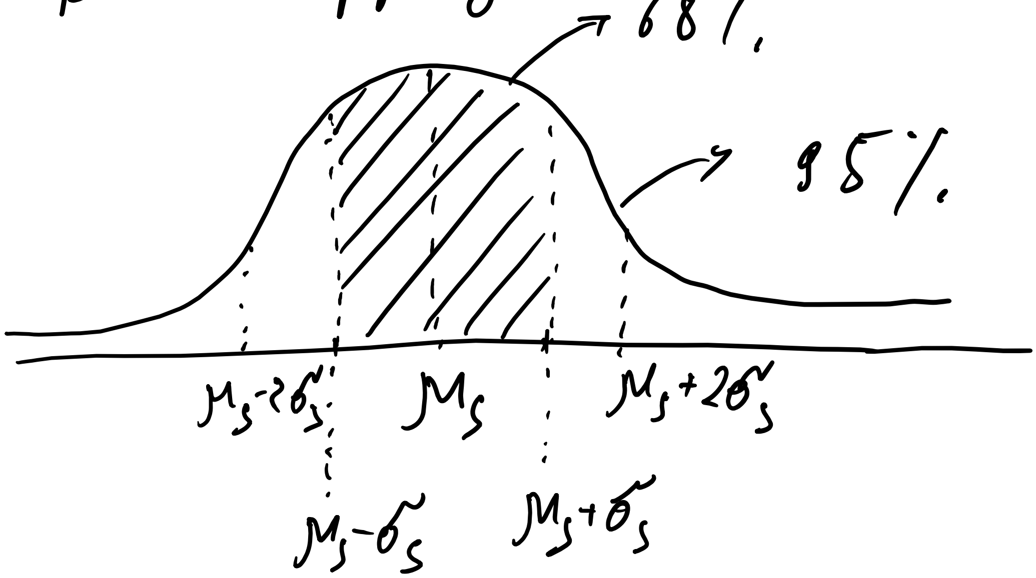
$$\hat{s} = \sqrt{8.61}$$
$$= 2.93$$

biased estimator of standard deviation



Methods for calculating confidence interval:

- Informal
- Normal-based
- Bootstrapping



$[s - \sigma_s, s + \sigma_s]$: confidence interval

You are 68% μ_s is lying in this range

$s - \sigma_s$ μ_s $s + \sigma_s$

confidence limit

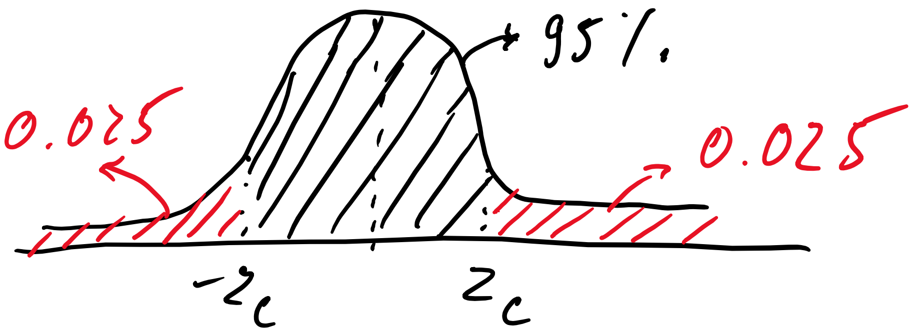
$[s - 2\sigma_s, s + 2\sigma_s]$

confidence level

95% confident that μ_s is lying in this range

$s - 2\sigma_s$ μ_s $s + 2\sigma_s$

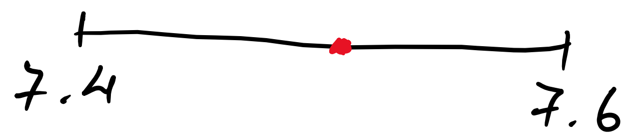
Determine the critical value z_c of 95% confidence level



$0.95/2 = 0.475 \longrightarrow z_c = 1.96$

$99\% \rightarrow z_c = 2.58$

$\mu_{Sample} = 7.5\%$

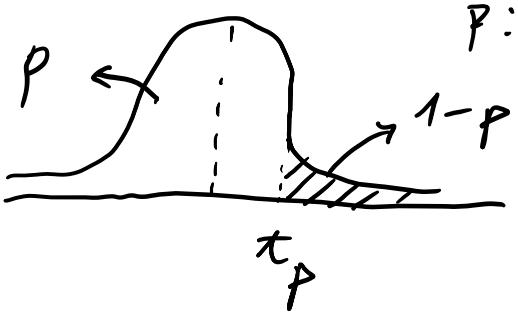


$\sigma = 1\% \longrightarrow \mu = 7.5 \pm \frac{1}{\sqrt{100}}$

$n = 100$

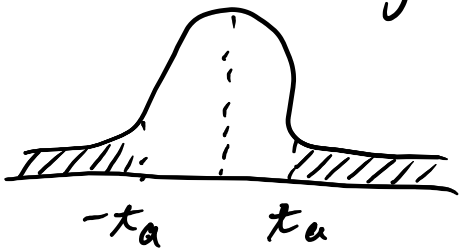
Student t-distribution

p: percentile



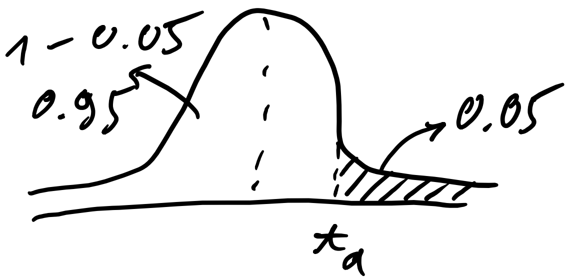
$V = n - 1$

Example: degree of freedom $v = 9$



Determine value of t_a if

a) Shaded area on the right = 0.05

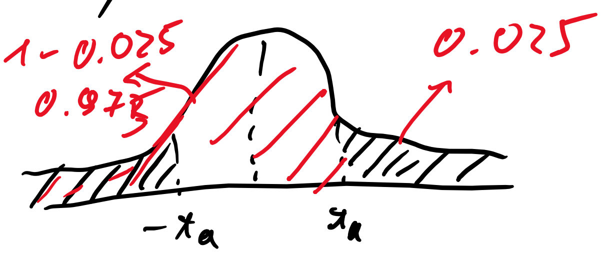


→ 95th percentile!

→ $t_{0.95}$

$t_a = 1.83$

b) The total shaded area = 0.05

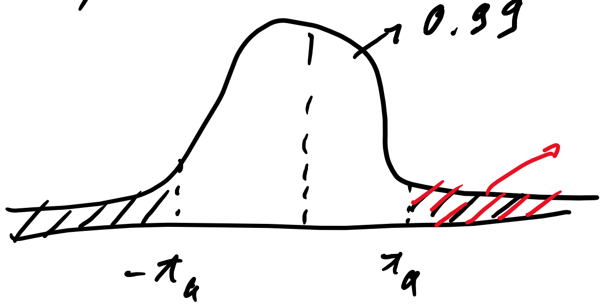


→ 97.5th percentile!

→ $t_{0.975}$

$t_a = 2.26$

c) The total unshaded area = 0.99



$(1-0.99)/2 = 0.005$

→ 99.5th percentile!

→ $t_{0.995}$!

$t_a = 3.25$

In-class exercise:

99% confidence level

→ $z_c = 2.58$

$\bar{X} = 67.45$ $\sigma = 2.93$ $n = 100$

$67.45 \pm 2.58 \dfrac{2.93}{\sqrt{100}} \rightarrow 66.69 < \mu < 68.21$

$\bar{X} = 67.45$ $\sigma = 2.93$ $n = 20$

$\downarrow$

$v = n - 1 = 19$

99% confidence level

$67.45 \pm t_{0.995} \cdot \dfrac{2.93}{\sqrt{20}}$

$$\ln(a^x) = x \cdot \ln(a)$$

$$\ln(x \cdot y) = \ln(x) + \ln(y)$$

$$\sum_{k=1}^{n} (x_k - \mu) = (x_1 - \mu) + (x_2 - \mu) + \dots (x_n - \mu)$$

$$= (x_1 + x_2 + \dots x_n) - (n\mu)$$

$$= \left(\sum_{k=1}^{n} x_k \right) - n\mu$$

$$f(x_k, \underset{\mu}{\underset{\downarrow}{\theta_1}}, \underset{\sigma^2}{\underset{\downarrow}{\theta_2}}) = \frac{1}{\sqrt{2\pi\theta_2}} e^{-\frac{(x_k - \theta_1)^2}{2\theta_2}}$$

$$L = \frac{1}{(2\pi)^{n/2} \theta_2^{n/2}} \cdot \exp\left(-\frac{1}{2\theta_2} \sum_{i=1}^{n} (x_1 - \theta_1)^2\right)$$

$$\ln L = \, !$$

$$\frac{\partial(\ln L)}{\partial \theta_1} = 0 \qquad \frac{\partial(\ln L)}{\partial \theta_2} = 0$$