Lecture Five: Parameter Estimation

chap. 8 of your textbook is
NOT a fun chapter

but it's an IMPORTANT chapter because
it introduces the t-distribution

→ In population
mean μ
variance σ^2

BUT it's RARE to
get data on the
entire population

→ In practice, we ~~usually~~ almost always get a SAMPLE
from which we have to estimate
the population mean + variance

→ Point Estimates

$\bar{X} = 28$ years
$s = 8$ years
$N = 25$

$\frac{\#\text{ male}}{N} = 30\%$

$\frac{\#\text{ female}}{N} = 70\%$

note: the standard deviation of the
sample is 8 years
the std. error of the estimated mean
is: $s_{\bar{X}} = \frac{s}{\sqrt{N}} = \frac{8}{\sqrt{25}} = \frac{8}{5} = 1.6$

→ Suppose we knew with absolute certainty that the true value of the std. deviation is 8

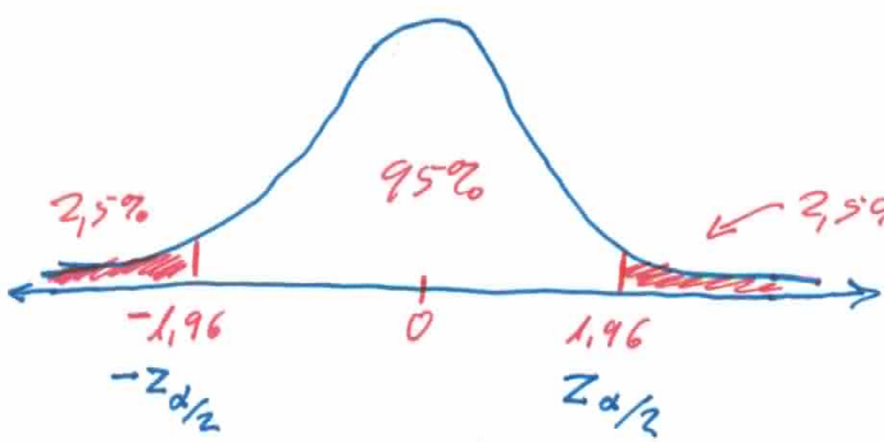
but we do not know the true value of the average age

→ In that case, how confident can we be that the average age is 28 years?

→ Answer: ZERO

→ There is zero probability that the average age is 28 years

→ BUT we can state with 95% confidence that the average age is between 24,864 years and 31,136 years



$\alpha = 0{,}95$

$$24{,}864 \leq 28 \leq 31{,}136$$

$$28 - z_{\alpha/2} \cdot \frac{8}{\sqrt{25}} \leq 28 \leq 28 + z_{\alpha/2} \cdot \frac{8}{\sqrt{25}}$$

$$28 - 3{,}136 \leq 28 \leq 28 + 3{,}136$$

→ we could also have chosen a 99% confidence interval

$\alpha = 0{,}99$ $\qquad z_{\alpha/2} = 2{,}576$

$$28 - z_{\alpha/2} \cdot \frac{8}{\sqrt{25}} \le 28 \le 28 + z_{\alpha/2} \cdot \frac{8}{\sqrt{25}}$$

$$28 - 2{,}576 \cdot 1{,}6 \le 28 \le 28 + 2{,}576 \cdot 1{,}6$$

$$28 - 4{,}121 \le 28 \le 28 + 4{,}121$$

$$23{,}879 \le 28 \le 32{,}121$$

→ But it's absurd to assume that we know the true population std. deviation

→ so we have to work with the t-distribution (which is approximately normal for large samples),

→ the normal distribution depends on the z-score where: $z = \frac{x - \mu}{\sigma}$

→ the t-distribution is very similar, but it also asks "How much information do we have?"

→ the answer to the "How much information?" question is the number of degrees of freedom

$$df = n - \text{number of estimated parameters}$$

→ In this case, we're only estimating one parameter (i.e. the mean), so $df = n-1$
$= 25 - 1 = 24$

→ So the 95% confidence interval is

$\alpha = 0{,}95$ $\quad t_{\frac{\alpha}{2}, df} = 2{,}064$

$df = 24$

$$28 - t_{\frac{\alpha}{2}, df} \cdot \frac{8}{\sqrt{25}} \leq 28 \leq 28 + t_{\frac{\alpha}{2}, df} \cdot \frac{8}{\sqrt{25}}$$

$$28 - 2{,}064 \cdot 1{,}6 \leq 28 \leq 28 + 2{,}064 \cdot 1{,}6$$

$$28 - 3{,}302 \leq 28 \leq 28 + 3{,}302$$

$$24{,}698 \leq 28 \leq 31{,}302$$