L9

Basic Statistics for Economists

Spring 2020

Department of Statistics

Stockholm University

Summary of L8

- **Sample** of n **independent** X_i from the same distribution.

- **Statistics** which are r.v.
  - Sample space, **sampling distributions**, expected values, variances

If the observations do not follow a $N(\mu, \sigma^2)$ distribution, we can use **CLT**. CLT says that $\bar{X} \sim N(\mu, \sigma^2/n)$ approximately, as long as n is large enough.

- Rule of thumb $n = 30$ (in practice, larger sample may be required)

- Applications to $\hat{p}$ and the binomial distribution, rule of thumb $np(1-p) > 5$

Independent, identically distributed X_i

- We typically assume the simplest case, all $\boldsymbol{X_i}$ in the sample are
  - pairwise **independent**
  - Are drawn from the **same distribution**

- Abbreviated **iid** (sv. olf, oberoende likformigt fördelade)
  - **i** = *indepedent*
  - **i** = *identically*
  - **d** = *distributed*

- Example: "$\boldsymbol{X_i}$ **iid** $\boldsymbol{N(\mu, \sigma^2)}$, $\boldsymbol{i = 1, \ldots, n}$"

Summary of L8

1. If X_i **iid** $N(\mu, \sigma^2)$ then $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$ regardless of the size of n

2. If X_i **iid** ~~N~~(μ, σ^2) then, according to **CLT**, $\bar{X} \to N\left(\mu, \frac{\sigma^2}{n}\right)$ when $n \to \infty$
   - Rule of thumb: $n > 30$

   If X_i **iid** $Bernoulli(P)$ then, according to **CLT,** $\hat{p} \to N\left(P, \frac{P(1-P)}{n}\right)$
   - Rule of thumb: $nP(1-P) > 5$

3. In all cases: $\mu_{\bar{X}} = E(\bar{X}) = \mu \quad (\mu_{\hat{p}} = E(\hat{p}) = P)$

4. In all cases: $\sigma^2_{\bar{X}} = Var(\bar{X}) = \frac{\sigma^2}{n} \to 0$ as $n \to \infty$

Today – in more detail

Inference

- To draw conclusions about the population (model) based on a sample

Estimators

- An estimator is a function of the values in the sample
- **Point estimate** of the population's **mean** μ
- **Interval estimate** of the mean μ
  - interval of uncertainty around the point estimate
  - Point estimate ± uncertainty

Population parameters and model parameters

Definition: a **parameter** is a **numerical measure** that describes a **property** of a **population** or **model**

- E.g. the mean number of employees employed December 31, 2017 among all the telecom businesses in Sweden (finite population)
- E.g. the wait time before reaching telephone support among telecom businesses in Sweden (infinite population)
  - Maybe we can model the wait time using a model with a normally distributed r.v.?
- The parameter is **not random**, it is **a constant**
- The (true) parameter value is typically **unknown**

Estimate/estimation

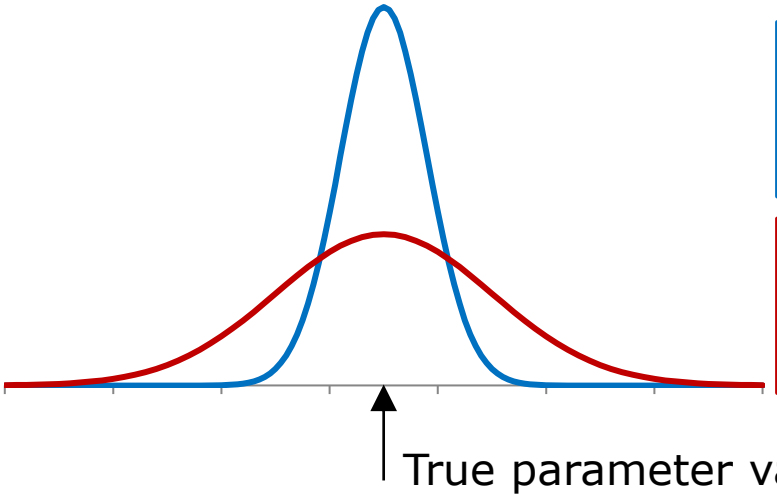
Definition: an **estimator** is a **statistic** that **estimates** an unknown value of some **parameter**.

- Since the **estimator** = statistic is a **random variable** it is easy to see that the **process of estimation is random**.
- We want the probability that our estimated value is close to the true parameter to be as high as possible.
- Depending on the properties of the estimator, the estimator may be good or bad. What are those properties?

Uncertainty in the estimation

Definition: **uncertainty** is the **risk** (or **probability**) that the estimated value will be **far from** the true parameter

- **Uncertainty = the variance of the estimator**
- Higher variance means higher uncertainty



Small variance = smaller probability of values far from the mean = less uncertanty

Big variance = greater probability of values far from the mean = greater uncertanty

Stockholm University

Estimators of μ and P

- An **estimator** (sv. *estimator*) a **statistic** which **estimates** a unknown value of a **parameter**.

Example	**Parameter**	**Estimator**
Mean	μ	$\overline{X}$
Proportion	P	$\hat{p}$
Variance	σ^2	S^2

Unknown constants **Random variables**

Called ***point estimates***, (a number). This is what you use to guess the value of the unknown parameter (also called **parameter estimates**).

Unbiased estimators

- An estimator is **unbiased** (sv. ***väntevärdesriktiga***) if it on average ”guesses” correctly, i.e. if the **expected value = parameter value**.

	Parameter	Estimator	Expected value
Mean	μ	$\bar{X}$	$E(\bar{X}) = \mu$
Proportion	p	$\hat{p}$	$E(\hat{p}) = p$
Variance	σ^2	S^2	$E(S^2) = \sigma^2$

All of these are unbiased!

Consistent estimators

$$P(|\bar{X}_n - \mu| \geq \varepsilon) \to 0 \text{ as } n \to \infty$$

- Somewhat simplified, the estimator is **consistent** if the **uncertanty** (variance) **decreases** as the sample size n **increases**.

	Parameter	Estimator	Variance
Mean	μ	$\bar{X}$	$Var(\bar{X}) = \frac{\sigma^2}{n}$
Proportion	p	$\hat{p}$	$Var(\hat{p}) = \frac{p(1-p)}{n}$
Variance	σ^2	S^2	Note $Var(S^2) = \frac{2\sigma^4}{n-1}$

$\to 0$ **as** $n \to \infty$

Note: if X_i are iid normally distributed. If not, things get a little more complicated.

Efficient estimators

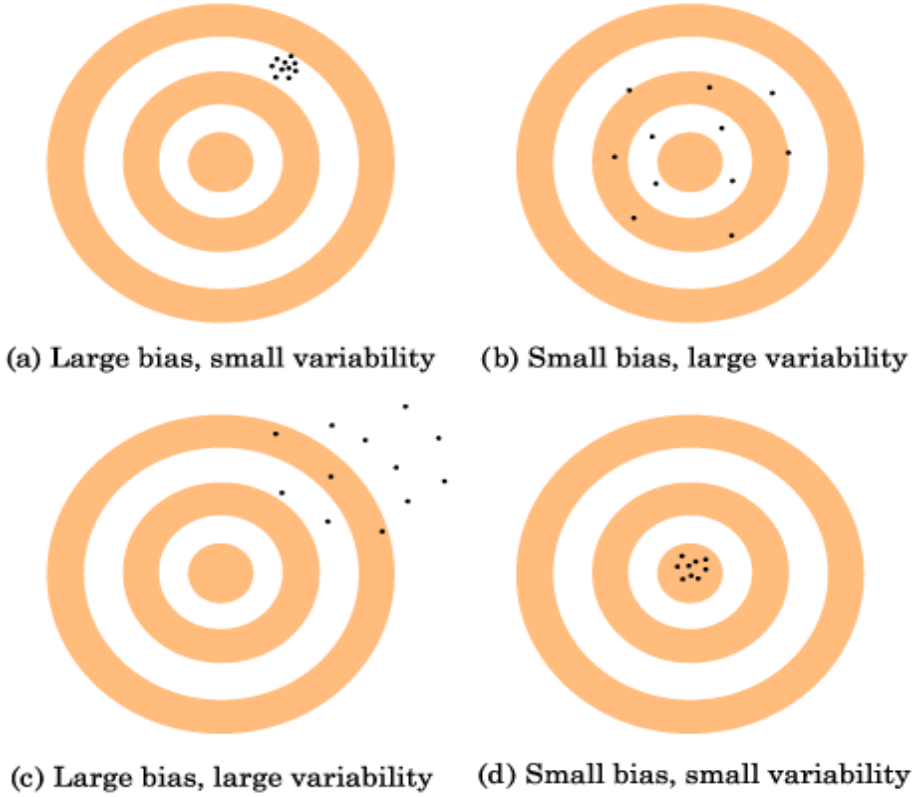
- Often, you have more than one possible estimator (statistic) that can be used to estimate the true parameter.
- In order to compare different estimators, you can check:
  - **unbiased** and **consistent**, or not
  - Their respective **variances**
- If $\boldsymbol{Var(estimator\ 1) < Var(estimator\ 2)}$

  $\Rightarrow$ then estimator 1 is more **efficient**.

e.g. $\bar{X}$ with $\boldsymbol{n = 20}$ and $\boldsymbol{n = 10}$ gives us $\dfrac{\sigma^2}{20} < \dfrac{\sigma^2}{10}$

Conclusion: greater $\boldsymbol{n}$ gives us a more efficient estimator.

The properties of estimators



Source: Moore et al. (2012) *Introduction to the Practice of Statistics*. New York: Freeman

Stockholm University

$\bar{X}$ as point estimate of μ

Expected value: $E(\bar{X}) = \mu$ *unbiased*

Variance: $Var(\bar{X}) = \frac{\sigma^2}{n}$ consistent, efficient, variance $\to 0$ as $n \to \infty$

Distribution:

1. If X_i **iid** $N(\mu, \sigma^2)$ then $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$ regardless of n
2. If X_i **iid** ~~N~~ (μ, σ^2) then, according to **CLT**, $\bar{X} \to N\left(\mu, \frac{\sigma^2}{n}\right)$ as $n \to \infty$
   - Rule of thumb: $n > 30$

Z-transformation, standardizing

- If $\bar{X}$ normally distributed of approximatively normally distributed according to CLT, we can calculate probabilities.
- We have to standardize correctly:

$$Z = \frac{\bar{X} - \mu_{\bar{X}}}{\sigma_{\bar{X}}} = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \qquad Z = \frac{X_i - \mu}{\sigma}$$

- I.e. suppose that $\boldsymbol{\mu = 10}$, $\boldsymbol{\sigma^2 = 25 \Rightarrow \sigma = 5}$, and $\boldsymbol{n = 100 \Rightarrow \sqrt{n} = 10}$

$$\boldsymbol{P(\bar{X} \leq 11)} = P\left(Z \leq \frac{11 - 10}{5/10}\right) = P\left(Z \leq \frac{1}{0{,}5}\right) = P(Z \leq 2) = \boldsymbol{0{,}97725}$$

$$\boldsymbol{P(\bar{X} \leq 11)} \neq \text{[wrong]} \neq P\left(Z \leq \frac{11-10}{\mathbf{5}}\right) = P(Z \leq 0{,}2) = \boldsymbol{0{,}57926}$$

Exercise 1

Calculate the probability that the sample mean $\bar{X}$ is **greater than** $\mu - \sigma/\sqrt{n}$ **and less than** $\mu + \sigma/\sqrt{n}$.

Note that I have not specified the values μ and σ!

Solution:

$$P\left(\bar{X} > \mu - \frac{\sigma}{\sqrt{n}} \cap \bar{X} < \mu + \frac{\sigma}{\sqrt{n}}\right) = P\left(\mu - \frac{\sigma}{\sqrt{n}} < \bar{X} < \mu + \frac{\sigma}{\sqrt{n}}\right) = \text{[standardize]}$$

$$= P\left(\frac{\mu - \sigma/\sqrt{n} - \mu}{\sigma/\sqrt{n}} < Z < \frac{\mu + \sigma/\sqrt{n} - \mu}{\sigma/\sqrt{n}}\right) = P(-1 < Z < 1) = \text{[draw!]}$$

$$= P(Z < 1) - P(Z < -1) = P(z < 1) - \big(1 - P(Z < 1)\big) = 2 \cdot P(Z < 1) - 1$$

$$= \text{[table 1]} = 2 \cdot 0,84134 - 1 = 0,68268$$

$\bar{X}$ will end upp in the interval $\mu \pm \sigma/\sqrt{n}$ with $68,2$ % probability.

TABELL 1. Normalfördelningen, standardiserad

$\Phi(z) = P(Z \leq z)$ där $Z \in N(0, 1)$.

För negativa värden, utnyttja att $\Phi(-z) = 1 - \Phi(z)$.

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,50000	0,50399	0,50798	0,51197	0,51595	0,51994	0,52392	0,52790	0,53188	0,53586
0,1	0,53983	0,54380	0,54776	0,55172	0,55567	0,55962	0,56356	0,56749	0,57142	0,57535
0,2	0,57926	0,58317	0,58706	0,59095	0,59483	0,59871	0,60257	0,60642	0,61026	0,61409
0,3	0,61791	0,62172	0,62552	0,62930	0,63307	0,63683	0,64058	0,64431	0,64803	0,65173
0,4	0,65542	0,65910	0,66276	0,66640	0,67003	0,67364	0,67724	0,68082	0,68439	0,68793
0,5	0,69146	0,69497	0,69847	0,70194	0,70540	0,70884	0,71226	0,71566	0,71904	0,72240
0,6	0,72575	0,72907	0,73237	0,73565	0,73891	0,74215	0,74537	0,74857	0,75175	0,75490
0,7	0,75804	0,76115	0,76424	0,76730	0,77035	0,77337	0,77637	0,77935	0,78230	0,78524
0,8	0,78814	0,79103	0,79389	0,79673	0,79955	0,80234	0,80511	0,80785	0,81057	0,81327
0,9	0,81594	0,81859	0,82121	0,82381	0,82639	0,82894	0,83147	0,83398	0,83646	0,83891
1,0	0,84134	0,84375	0,84614	0,84849	0,85083	0,85314	0,85543	0,85769	0,85993	0,86214
1,1	0,86433	0,86650	0,86864	0,87076	0,87286	0,87493	0,87698	0,87900	0,88100	0,88298
1,2	0,88493	0,88686	0,88877	0,89065	0,89251	0,89435	0,89617	0,89796	0,89973	0,90147
1,3	0,90320	0,90490	0,90658	0,90824	0,90988	0,91149	0,91309	0,91466	0,91621	0,91774
1,4	0,91924	0,92073	0,92220	0,92364	0,92507	0,92647	0,92785	0,92922	0,93056	0,93189
1,5	0,93319	0,93448	0,93574	0,93699	0,93822	0,93943	0,94062	0,94179	0,94295	0,94408
1,6	0,94520	0,94630	0,94738	0,94845	0,94950	0,95053	0,95154	0,95254	0,95352	0,95449
1,7	0,95543	0,95637	0,95728	0,95818	0,95907	0,95994	0,96080	0,96164	0,96246	0,96327
1,8	0,96407	0,96485	0,96562	0,96638	0,96712	0,96784	0,96856	0,96926	0,96995	0,97062
1,9	0,97128	0,97193	0,97257	0,97320	0,97381	0,97441	0,97500	0,97558	0,97615	0,97670
2,0	0,97725	0,97778	0,97831	0,97882	0,97932	0,97982	0,98030	0,98077	0,98124	0,98169
2,1	0,98214	0,98257	0,98300	0,98341	0,98382	0,98422	0,98461	0,98500	0,98537	0,98574

Exercise 2 (similar to 1)

Calculate the probability that the sample mean $\bar{X}$ is

greater than $\mu - 1,96 \cdot \sigma/\sqrt{n}$ and less than $\mu + 1,96 \cdot \sigma/\sqrt{n}$.

Note that I again have not specified the values μ and σ!

Solution:

$$P\left(\mu - 1,96\frac{\sigma}{\sqrt{n}} < \bar{X} < \mu + 1,96\frac{\sigma}{\sqrt{n}}\right) = \text{[standardize]}$$

$$= P\left(\frac{\mu - 1,96 \cdot \sigma/\sqrt{n} - \mu}{\sigma/\sqrt{n}} < Z < \frac{\mu + 1,96 \cdot \sigma/\sqrt{n} - \mu}{\sigma/\sqrt{n}}\right) = P(-1,96 < Z < 1,96)$$

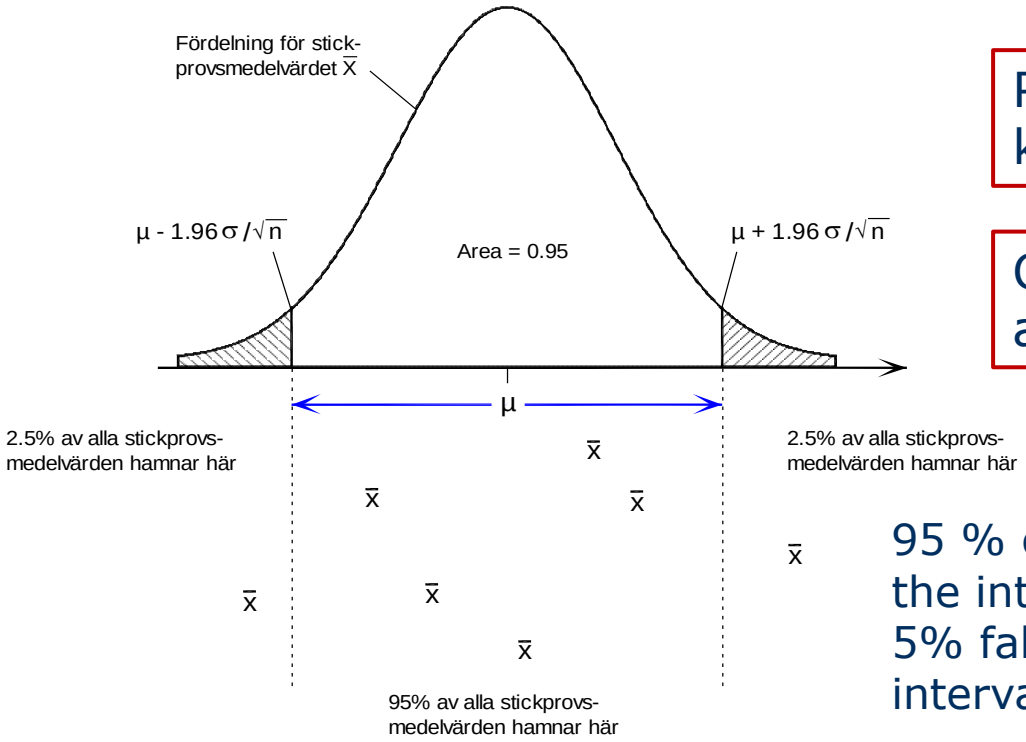
$$= \text{[the method is analogous to that of exercise 1]} = 2 \cdot P(Z < 1,96) - 1$$

$$= \text{[table 1]} = 2 \cdot 0,97500 - 1 = 0,95000$$

With 95 % probability, $\bar{X}$ will be in the interval $\mu \pm 1,96 \cdot \sigma/\sqrt{n}$

The probability that $\bar{X}$ falls within an interval

- **95** % probability that $\bar{X}$ falls within the interval $\mu \pm 1{,}96 \cdot \sigma/\sqrt{n}$



Problem: we do not know the true value of μ

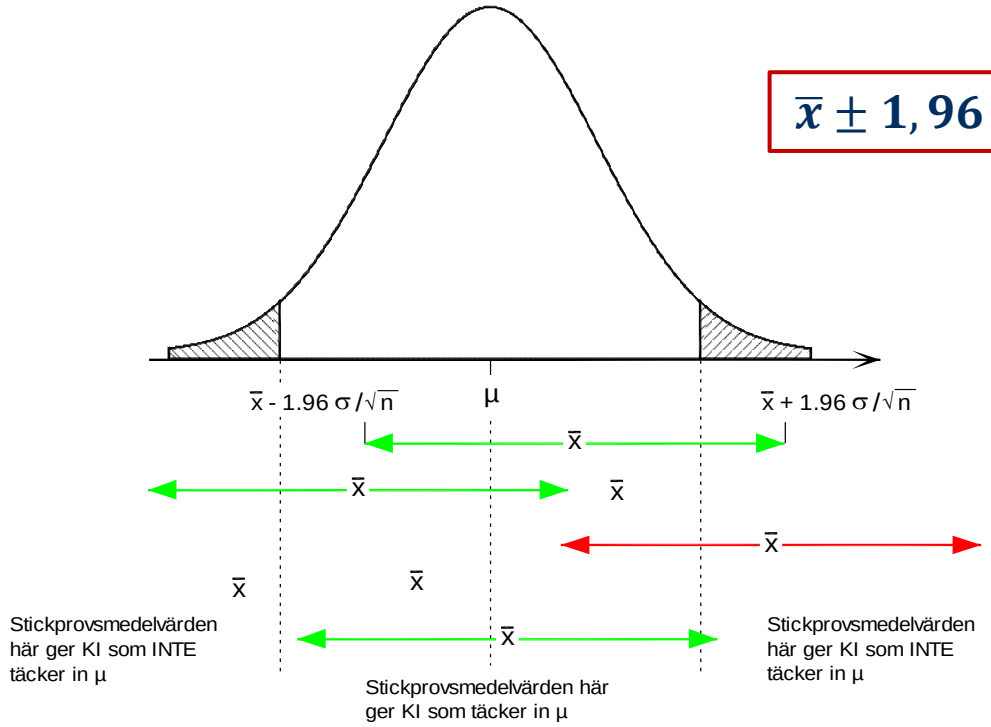
Construct an interval around $\bar{x}$ instead!

95 % of all $\bar{x}$ fall within the interval around μ, 5% fall outside this interval.

Stockholm University

Interval estimation

- Estimate the entire interval $\boldsymbol{\mu \pm 1,96 \cdot \sigma/\sqrt{n}}$, not just the point $\boldsymbol{\mu}$



95% of all intervals cover μ, 5% do not.

Stockholm University

Prediction and inference

- **Prediction**:

  Population, model ⟶ Sample, single observations

$$P\left(\boldsymbol{\mu} - 1{,}96\frac{\sigma}{\sqrt{n}} < \bar{\boldsymbol{X}} < \boldsymbol{\mu} + 1{,}96\frac{\sigma}{\sqrt{n}}\right) = 0{,}95$$

  A future sample mean has a 95% probability to fall within this interval.

- **Inference**:

  Sample ⟶ Population, model

$$\bar{\boldsymbol{x}} - 1{,}96\frac{\sigma}{\sqrt{n}} < \boldsymbol{\mu} < \bar{\boldsymbol{x}} + 1{,}96\frac{\sigma}{\sqrt{n}}$$

with **95 % confidence**

Confidence interval

- The interval $\bar{x} \pm 1,96\frac{\sigma}{\sqrt{n}}$

  is a **95% confidence interval** of μ (where $\bar{x}$ is an observed value of $\bar{X}$) if $\bar{X}$ is normally distributed and each observation is independent of the other observations.

In this case:
$1,96\frac{\sigma}{\sqrt{n}}$ = margin of error

- Interpretation: Before we drew the sample, there was a 95% probability that the confidence interval based on this sample would include the true value of the population parameter.
- **Half the interval width** is often called the **statistical margin of error**.

What influences the width of the interval?

$$\bar{x} \pm 1,96\frac{\sigma}{\sqrt{n}}$$

1. **Variance σ^2** – spread within the population
   - This is out of our control

2. **The sample size n**
   - greater the values of n mean narrower intervals (efficiency)

3. **Coverage, confidence level**
   - We have used 95 % and the value 1,96
   - Can we change it? What happens then?

Confidence level

- Confidence level or coverage is calculated as

  $\mathbf{100(1-\alpha)\%}$ (the Greek letter a, alpha)

- We have used the value $1{,}96$ which gives us $95\,\%$ confidence:

  $95 = 100(1-\alpha) \Rightarrow \boldsymbol{\alpha = 0{,}05}$

- $\boldsymbol{\alpha = 0{,}05}$ is the probability that $\bar{X}$ falls outside of $\boldsymbol{\mu \pm 1{,}96 \cdot \sigma/\sqrt{n}}$

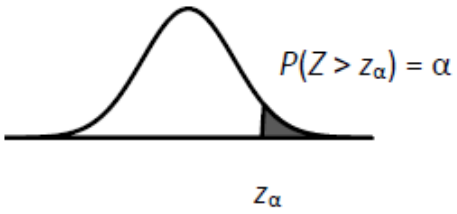
- Wider interval gives us an increased confidence level:
  - $\boldsymbol{\alpha/2 = 0{,}005}$: $\boldsymbol{\mu \pm 2{,}5758 \cdot \sigma/\sqrt{n}}$ gives **99%** confidence
- Narrower interval gives decreased confidence level:
  - $\boldsymbol{\alpha/2 = 0{,}05}$: $\boldsymbol{\mu \pm 1{,}6449 \cdot \sigma/\sqrt{n}}$ gives **90%** confidence

$$\bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

Table 2 – the inverse table for $N(0,1)$

- Suppose that you want the confidence level $\mathbf{100(1-\alpha)\% = 99\,\%}$
- Then $\alpha = 0{,}01$ and $\boldsymbol{\alpha/2 = 0{,}005}$
- Read from the table $\boldsymbol{z_{0{,}005} = 2{,}5758}$

Note! It says $\boldsymbol{\alpha}$ in the table, but you have to divide by two to get $\boldsymbol{\alpha/2}$, before you look up the value! Why?



α	z_α
0,25	0,6745
0,10	1,2816
0,05	1,6449
0,025	1,9600
0,010	2,3263
0,005	2,5758
0,0025	2,8070
0,0010	3,0902
0,0005	3,2905
0,00025	3,4808
0,00010	3,7190
0,00005	3,8906
0,000025	4,0556
0,000010	4,2649
0,000005	4,4172

- Notice that 95% confidence gives $\boldsymbol{\alpha/2 = 0{,}025}$ and $\boldsymbol{z_{0{,}025} = 1{,}96}$ as expected.

Stockholm University

Exercise 3

- A sample of size $\boldsymbol{n = 100}$ is drawn from population in which the variance is $\boldsymbol{\sigma^2 = 25}$. The sample mean is calculated: $\boldsymbol{\bar{x} = 40}$.
- Calculate the confidence interval of the population men $\boldsymbol{\mu}$ using a) 90, b) 95 and c) 99,9 % confidence level.
- Solution: nothing is said about the distribution of the population. We can assume that it is normally distributed or better, we can apply CLT and conclude that $\boldsymbol{\bar{X}}$ has a distribution that is approximately normally distributed.

Exercise 4, cont.

The interval is calculated as (assuming normal distribution or using CLT)

$$\mathbf{40} \pm \boldsymbol{z}_{\alpha/2} \cdot \frac{\sqrt{\mathbf{25}}}{\sqrt{\mathbf{100}}} = \mathbf{40} \pm \frac{\boldsymbol{z}_{\alpha/2}}{\mathbf{2}}$$

a) $\alpha = 0{,}10 \qquad \alpha/2 = 0{,}05 \qquad \boldsymbol{z}_{\mathbf{0{,}05}} = \mathbf{1{,}6449}$

90% CI renders $\mathbf{40} \pm \mathbf{0{,}82245}$, or in interval from: $(\mathbf{39{,}18}\,;\,\mathbf{40{,}82})$

b) $\alpha = 0{,}05 \qquad \alpha/2 = 0{,}025 \qquad \boldsymbol{z}_{\mathbf{0{,}025}} = \mathbf{1{,}9600}$

95% CI renders $\mathbf{40} \pm \mathbf{0{,}98}$, or in interval from: $(\mathbf{39{,}02}\,;\,\mathbf{40{,}98})$

c) Do it yourselves! Correct answer: $(\mathbf{38{,}35}\,;\,\mathbf{41{,}65})$

Unknown variance σ^2

- If the **variance** σ^2 for individual observations is **unknown** it follows that the variance of $\bar{X}$, i.e. σ^2/n, is unknown, too.
- Estimate σ^2 using the sample variance s^2.
- In this case, we get a different interval with t instead of z:

$$\bar{x} \pm t_{\nu,\alpha/2}\frac{s}{\sqrt{n}}$$

What are t and ν here?

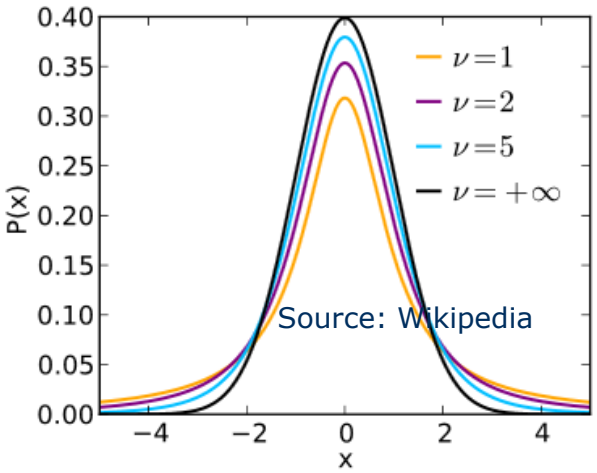
- Increased uncertainty when σ^2 is replaced by s^2.
- If X_i **iid** $\sim N(\mu, \sigma^2)$ then $t = \dfrac{\bar{X} - \mu}{s/\sqrt{n}} \sim t(n-1)$

Student's t-distribution

We write $t \sim t(\nu)$

t is **t-distributed** with **ν degrees of freedom**

- The sample space of t is (-∞,∞)
- Let $T \sim t(\nu)$. If ν > 1, then $E(T) = 0$
- If ν > 2, $Var(T) = \nu/(\nu-2)$
- The distribution is bell shaped and symmetrical
  (compare to the standard normal distribution)



- t-distribution → normal distribution as $\nu \to \infty$

(Greek n, "nu")

Stockholm University

t –table – inverse table for t

- **Table 3**
- Suppose that you want the confidence level $\mathbf{95}\,\%$
- Then $\boldsymbol{\alpha/2 = 0,025}$
- if $\boldsymbol{n = 10}$ then the degrees of freedom are $\boldsymbol{\nu = n - 1 = 9}$
- Read from the table:

$$\boldsymbol{t_{\nu,\alpha/2} = t_{9;0,025} = 2,262}$$

v	α = 0,1	0,05	0,025	0,010	0,005	0,0
1	3,078	6,314	12,706	31,821	63,657	127,
2	1,886	2,920	4,303	6,965	9,925	14,
3	1,638	2,353	3,182	4,541	5,841	7,
4	1,533	2,132	2,776	3,747	4,604	5,
5	1,476	2,015	2,571	3,365	4,032	4,
6	1,440	1,943	2,447	3,143	3,707	4,
7	1,415	1,895	2,365	2,998	3,499	4,
8	1,397	1,860	2,306	2,896	3,355	3,
9	1,383	1,833	2,262	2,821	3,250	3,
10	1,372	1,812	2,228	2,764	3,169	3,
11	1,363	1,796	2,201	2,718	3,106	3,
12	1,356	1,782	2,179	2,681	3,055	3,
13	1,350	1,771	2,160	2,650	3,012	3,
14	1,345	1,761	2,145	2,624	2,977	3,
15	1,341	1,753	2,131	2,602	2,947	3,
16	1,337	1,746	2,120	2,583	2,921	3,
17	1,333	1,740	2,110	2,567	2,898	3,
18	1,330	1,734	2,101	2,552	2,878	3,
19	1,328	1,729	2,093	2,539	2,861	3,
20	1,325	1,725	2,086	2,528	2,845	3,
21	1,323	1,721	2,080	2,518	2,831	3,

Exercise 4

- A sample of size $\boldsymbol{n = 10}$ is drawn from a population but the variance $\boldsymbol{\sigma^2}$ is unknown. The sample mean is calculated as $\boldsymbol{\bar{x} = 40}$ and the **sample variance** $\boldsymbol{s^2 = 25}$.

- Calculate the confidence interval for the population mean $\boldsymbol{\mu}$ with a) 90, b) 95 and c) 99 % confidence level.

- Solution: the distribution of the population is not mentioned. We assume that it is normally distributed. In order to use a $\boldsymbol{t}$-distribution the observations $\boldsymbol{X_i}$ have to be $\mathbf{iid} \sim \boldsymbol{N(\mu, \sigma^2)}$. The sample is too small for CLT.

Solution 4, cont.

The interval is calculated as follows (under assumption of normal dist.)

$$\mathbf{40 \pm t_{n-1;\alpha/2} \cdot \frac{s}{\sqrt{n}}} \;=\; \mathbf{40 \pm t_{9;\alpha/2} \cdot \frac{5}{\sqrt{10}}} \;=\; \mathbf{40 \pm t_{9;\alpha/2} \cdot 1,5811}$$

a) $\alpha = 0{,}10 \qquad \alpha/2 = 0{,}05 \qquad \mathbf{t_{9;0,05} = 1,833}$

90% CI renders $\mathbf{40 \pm 2,898}$ or in interval form: $\mathbf{(37,10\,;\,42,90)}$

b) $\alpha = 0{,}05 \qquad \alpha/2 = 0{,}025 \qquad \mathbf{t_{9;0,025} = 2,262}$

95% CI renders $\mathbf{40 \pm 3,576}$ or in interval form: $\mathbf{(36,42\,;\,43,58)}$

c) Do it yourselves! Correct answer: $\mathbf{(34,86\,;\,45,14)}$

Summary

- … saved for tomorrow …

But, we make time for a quick question:

- What happens to a confidence interval when we are force to estimate the variance?
- Does it get wider, narrower, or does it stay the same?
- For instance, compare $\boldsymbol{z_{0,025} = 1,9600}$ to $\boldsymbol{t_{n-1;0,025}}$ of different values of $\boldsymbol{n}$. What happens as $\boldsymbol{n \to \infty}$?

Next time

NCT sections 7.7 + 8.1-8.3

- More on confidence intervals
  - Of a proportion $\boldsymbol{P}$ using the estimate $\boldsymbol{\hat{p}}$
  - Differences between means $\boldsymbol{\overline{X} - \overline{Y}}$
  - pairwise differences $\boldsymbol{D_i = X_i - Y_i}$ and $\boldsymbol{\overline{D}}$
  - differences between proportions $\boldsymbol{P_x - P_y}$