MFin Econometrics I
Session 2: Sampling distributions of estimators, Tests of hypotheses

Thilo Klein

University of Cambridge
Judge Business School

Point and Interval estimation

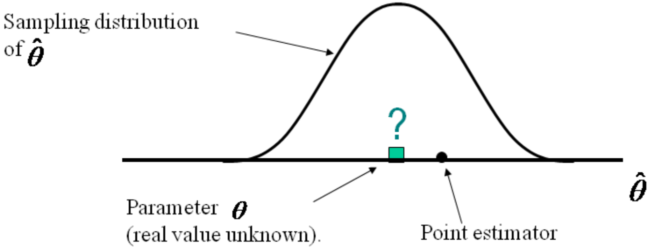
Point estimation

- A point estimator estimates the value of an unknown parameter in a population using a *single value*
- But we deal with random variables and therefore cannot have certainty
- Way forward?

CAMBRIDGE Judge Business School

Point and Interval estimation

Point estimation - illustration



Point and Interval estimation

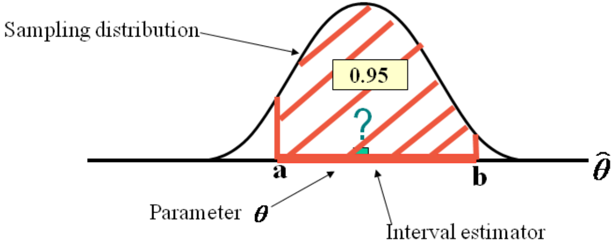
Interval estimation

- An interval estimator estimates the unknown parameter using a (small) *interval*, ...
- ...and the associated (high) probability that the population parameter is contained in that interval
- This takes account of sampling. The sample (to which the estimator is applied to obtain an estimate) is random
- Q: What is the the smallest interval with a sufficiently high probability - the most informative *interval*?

CAMBRIDGE Judge Business School

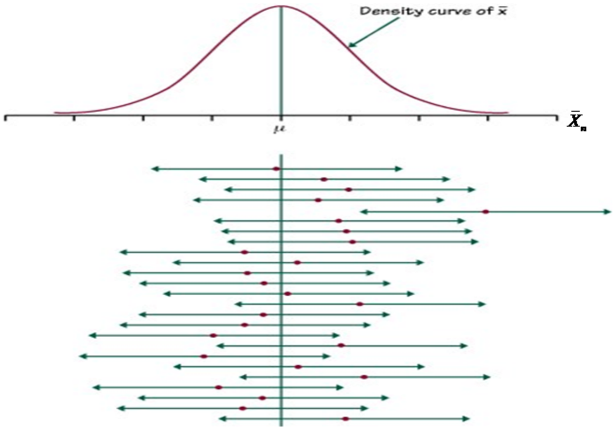
Point and Interval estimation

Interval estimation - illustration



Point and Interval estimation

Interval estimation - illustration 2



CAMBRIDGE Judge Business School

Point and Interval estimation

Q: Interval estimation

- Question: Why not define an interval we can be *certain* of containing the true value?
- The only certain interval is $[-\infty, +\infty]$!

CAMBRIDGE Judge Business School

Point and Interval estimation

Confidence interval for a parameter θ ...

- ... is an interval on the line (the space in which θ can lie) that, given the sampling distribution of the estimator $\hat{\theta}$, contains θ with a specified (sufficiently high) probability
  - e.g., What is the interval $[a, b]$ that will contain θ with probability of, say, 0.95 (i.e., $a \leq \theta \leq b$ with probability 0.95)?
  - Find a and b, and you have an interval estimate: $[a, b]$ is the 95% confidence interval for θ
  - The price of defining a (small) interval $[a, b]$, and not $(-\infty, \infty)$ is the (5%) probability that you necessarily allow that your interval estimate may be wrong (does not contain θ)
  - This probability split is yours to make: 99% – 1% or 90% – 10%, any other, depending on the *probability of being wrong* that you can live with

CAMBRIDGE Judge Business School

Point and Interval estimation

Confidence interval and Critical region for a test of hypothesis

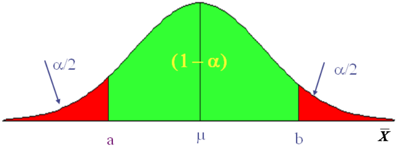
- So, to test your (well reasoned) hypothesis about the unknown θ, you need to fix a and b;
- the region in the parameter space (the real line) outside $[a, b]$ is the critical region for your test
  - What you are really asking is: Is the difference between your hypothesized θ and the estimated $\hat{\theta}$ attributable to the randomness of sampling?
  - Or is the difference between θ and $\hat{\theta}$ too large for it to be merely due to sampling variation?
  - If so, what should you do with your pet theory?

CAMBRIDGE Judge Business School

Point and Interval estimation

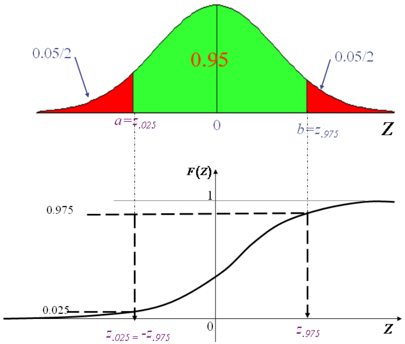
Confidence interval for μ

- Assume $X \sim N(\mu, \sigma)$ and that σ is known (or that n is large)
- $\bar{X} \sim N(\mu, \frac{\sigma}{\sqrt{n}})$ So the test statistic $\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} = Z \sim N(0,1)$
- What probability (α) that your best interval estimate is wrong can you live with? (in testing hypotheses, α will be referred to as the size of the test). Let us fix $\alpha = 5\%$
- What are that values a and b, (with reference to $\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} = Z \sim N(0,1)$) such that $Pr(a \leq \mu \leq b) = 1 - \alpha$?



Point and Interval estimation

95% Confidence interval for μ (2)



$$\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} = Z \sim N(0,1); \qquad \alpha = 5\%;$$

find: $Pr(-z_{\alpha/2} \leq \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \leq z_{\alpha/2}) = 0.95$

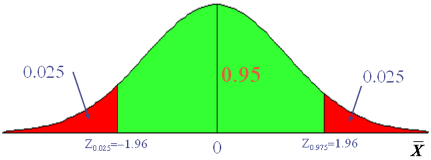
CAMBRIDGE Judge Business School

Point and Interval estimation

95% Confidence interval for μ (3)

$$Pr(-z_{\alpha/2}\frac{\sigma}{\sqrt{n}} + \bar{X} \leq \mu \leq z_{\alpha/2}\frac{\sigma}{\sqrt{n}} + \bar{X}) = 0.95$$

- From the standard Normal Table:
  $z_{.025} = -1.96 \quad z_{.975} = 1.96$
- $Pr(-1.96\frac{\sigma}{\sqrt{n}} + \bar{X} \leq \mu \leq 1.96\frac{\sigma}{\sqrt{n}} + \bar{X}) = 0.95$



Tests of hypothesis

Old example to illustrate types of errors in testing hypotheses

- A rare disease infects 1 person in a 1000
- There is good but imperfect test
- 99% of the time, the test identifies the disease
- 2% of uninfected patients also return a positive test result
- Null Hypothesis H_0: Patient has the disease
- Alternate Hypothesis H_a: Patient does not
    - Q: Why not choose as H_0: Patient does not has the disease?

CAMBRIDGE Judge Business School

Tests of hypothesis

Test of hypotheses example: Joint distribution

	A: patient has disease	$\overline{A}$: patient does not have disease	
B: patient tests positive	0.00099	0.01998	$P(B) = 0.02097$
$\overline{B}$: patient does not test positive	0.00001	0.97902	$P(\overline{B}) = 0.97903$
	$P(A) = 0.001$	$P(\overline{A}) = 0.999$	1

H_0: Patient has the disease H_a: Patient does not

CAMBRIDGE Judge Business School

Tests of hypothesis

Test of hypotheses example: correct decisions

	A: Patient has disease	$\overline{A}$: Patient does not have disease
B: Tests positive	If you do not reject H_0: **Correct decision**	
$\overline{B}$: patient does not test positive		If you reject H_0: **Correct decision**

H_0: Patient has the disease H_a: Patient does not

CAMBRIDGE Judge Business School

Tests of hypothesis

Test of hypotheses example: Type I error

	A: Patient has disease	$\overline{A}$: Patient does not have disease
B: Tests positive	Correct decision	
$\overline{B}$: Does not test positive	If you rejected (the true) H_0, Type I error	Correct decision

H_0: Patient has the disease H_a: Patient does not

CAMBRIDGE Judge Business School

Tests of hypothesis

Test of hypotheses example: Type II error

	A: Patient has disease	$\overline{A}$: Patient does not have disease
B: Tests positive	Correct decision	If you did not reject (the false) H_0, Type II error
$\overline{B}$: Does not test positive	Type I error	Correct decision

H_0: Patient has the disease H_a: Patient does not

CAMBRIDGE Judge Business School

Tests of hypothesis

H_0: Patient has the disease H_a: Patient does not

	Patient has disease	Patient does not
B: Tests positive	Correct decision	Prob(Type II error) $=0.01998/0.999 = .02 = 2\%$
$\overline{B}$: Does not test positive	Prob(Type I error) $=0.00001/0.001 = 0.01 = 1\%$	Correct decision

- $P(\text{Error type I}) = P(\text{Reject } H_0 \mid H_0 \text{ true}) = \alpha =$ *Size of test*
- $P(\text{Error type II}) = P(\text{Not Reject } H_0 \mid H_0 \text{ false}) = \beta$
- $1 - P(\text{Error type } II) = 1 - \beta =$ *Power of test*

CAMBRIDGE Judge Business School

Tests of hypothesis

Hypothesis tests - two points

- First:
  - Probability of Type I error, is the size of the test, α and is 1% in this example
  - Can it be changed? How?
- Second:
  - We can never have enough evidence to *accept* a null hypothesis
  - We suspend judgement if the evidence is against the alternative
  - We can only *reject* or *not reject* the null

CAMBRIDGE Judge Business School

Test of hypothesis: an example

A hypothesis about the impact of discounts on sales of automobiles

- Increased sales of Citreons after discount
- $\mu_0 = 1200$ *hypothesized* increase in UK sales of Citreons with discount
- $\sigma = 300$ *assumed known* population st. dev. of increase in sales with discount
- X random variable - increase in sales of Citreons after discount
- A sample of 100 discount episodes observed: $\bar{X} = 1265$
- Frame a test:

CAMBRIDGE Judge Business School

Test of hypothesis: an example

Possible hypothesis tests for the mean: One or two tailed

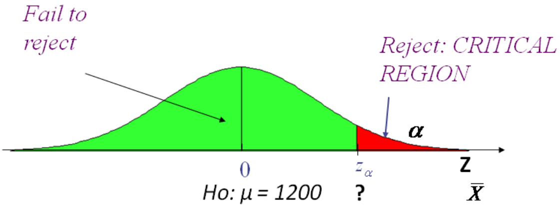
- $H_0: \mu = \mu_0$ Vs. $H_a: \mu > \mu_0$ (1-sided, $>$)
- $H_0: \mu = \mu_0$ Vs. $H_a: \mu < \mu_0$ (1-sided, $<$)
- $H_0: \mu = \mu_0$ Vs. $H_a: \mu \neq \mu_0$ (2-sided)

CAMBRIDGE Judge Business School

Test of hypothesis: an example

Q: Sales of automobiles

- $H_0 : \mu = 1200$, $H_a : \mu > 1200$
- Find a and b, using sample estimate $\bar{X}$, such that $Pr(a \leq \mu \leq b) = 1 - \alpha$
- i.e., a and b, such that Under H_0: Pr(test statistic *does not lie* in the critical region)=$1 - \alpha$
- $Pr\left(\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \leq z_\alpha\right)$=1-$\alpha$ *i.e.,* $Pr\left(\frac{\bar{X}-1200}{\frac{300}{\sqrt{100}}} \leq 1.645\right)$=0.95



CAMBRIDGE Judge Business School

Test of hypothesis: an example

Q: Sales of automobiles

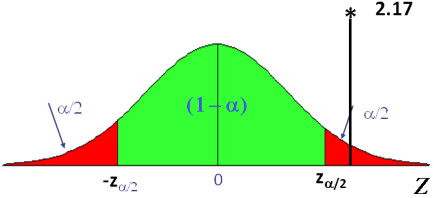
- The critical region is: $Pr\left(\frac{\bar{X}-1200}{\frac{300}{\sqrt{100}}} > 1.645\right)$
- Note that this is a one-tail test, so all 5% is on the right tail
- In this case: $(1265 - 1200)/30 = 2.17 > 1.645$
- So we reject the null hypothesis

CAMBRIDGE Judge Business School

Test of hypothesis: an example

Two tailed test: Sales of automobiles

- $H_0 : \mu = 1200$, $H_a : \mu \neq 1200$
- From the SND table, $z_{\alpha/2} = z_{.025} = 1.96$

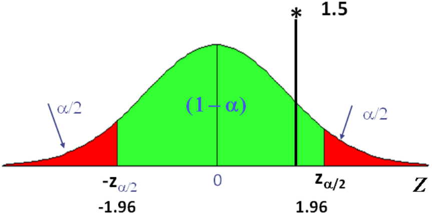


CAMBRIDGE Judge Business School

Test of hypothesis: an example

Two tailed test: Sales of automobiles (2)

- Suppose our sample estimate is different: $\bar{X} = 1245$



CAMBRIDGE Judge Business School

Test of hypothesis: an example

p-value

- The significance level of a test is the pre-specified probability of incorrectly rejecting the null, when the null is true
- e.g., if the pre-specified significance level is 5% (size of test):
    - you reject the null hypothesis in a two-tailed test if $|\text{standardised test statistic}| \geq 1.96$
- p-value = *probability* of drawing a statistic (e.g. $\bar{Y}$) *at least as adverse to the null* as the value actually computed with your data, assuming that the null hypothesis is true
    - If significance level is 5%, you reject the null hypothesis if $p \leq 0.05$
    - The p-value is sometimes called the *marginal significance level*
    - It is better to report the p-value, than simply whether a test rejects or not
    - p-value contains more information than “reject/not reject”

CAMBRIDGE Judge Business School

Test of hypothesis: an example

Different Confidence Intervals

1-α	Confidence interval
0.5	$(\bar{X} - 0.67\frac{\sigma}{\sqrt{n}}), \bar{X} + 0.67\frac{\sigma}{\sqrt{n}})$
0.9	$(\bar{X} - 1.64\frac{\sigma}{\sqrt{n}}), \bar{X} + 1.64\frac{\sigma}{\sqrt{n}})$
0.95	$(\bar{X} - 1.96\frac{\sigma}{\sqrt{n}}), \bar{X} + 1.96\frac{\sigma}{\sqrt{n}})$
0.99	$(\bar{X} - 2.57\frac{\sigma}{\sqrt{n}}), \bar{X} + 2.57\frac{\sigma}{\sqrt{n}})$
0.999	$(\bar{X} - 3.27\frac{\sigma}{\sqrt{n}}), \bar{X} + 3.27\frac{\sigma}{\sqrt{n}})$

The more the degree of certainty (lower Pr(Type I error)) needed, the larger the interval

CAMBRIDGE Judge Business School

Tests of hypothesis: Power of the test

Type II errors: simulation

- You commit a type II error if you *do not reject a false Null*
- Error type II occurs if you *do not reject a false Null*
- $P[\text{Reject} \quad H_0|H_0 \ \text{ false}] = 1 - P[\text{Error type II}] =$ Power of the test
- Experiment to illustrate:
  - Generate data through i.i.d. draws from $N(\mu, 1)$ (simple random sampling)
  - Keep $\sigma^2 = 1$; but different values of μ in the interval $[-2, 2]$
  - Always test the null: $H_0 : \mu = 0$ against alternative: $H_a : \mu \neq 0$
  - Aim: determine the power of the test, i.e., the prob of not making type II errors, prob. of not rejecting the Null when it is false

CAMBRIDGE Judge Business School

Tests of hypothesis: Power of the test

Probability of not making type II errors: simulation (2)

- Sample mean ($\bar{Y}$) is the estimator of μ
- 3 sample sizes: 10, 100 and 1000, used for estimating ($\bar{Y}$)
- Recall: Samples are from $N(\mu, 1)$ where μ is in $[-2, 2]$
- Critical region:
  - Size of the test fixed at 5%
  - We reject the null ($\mu = 0$) if $|\bar{Y}| > c$, where c is determined by $P[-c \leq \bar{Y} \leq c] = 0.95$, for $\mu = 0$
  - As $\sigma = 1$, and the test is for $\mu = 0$, we have $c = 1.96/\sqrt{n}$
- Note: in most cases in this experiment, the null is false
- 10000 runs of each test. The proportion of times when H_0 is rejected is reported

CAMBRIDGE Judge Business School

Tests of hypothesis: Power of the test

Pr(Reject H_0) reported in percentages

DGP: $Y_i \sim N(\mu, 1)$ for $\mu \in [-2, 2]$, including $\mu = 0$

$H_0 : \mu = 0; H_a : \mu \neq 0$

	population mean (actual)	n=10	n=100	n=1000
1-P(EII)	μ = -2	100	100	100
	μ = -1	89	100	100
	μ = -0.2	9.7	51.2	100
	μ = -0.1	6.4	17.3	88.5
	μ = -0.05	5.4	7.3	35.5
EI →	**μ = 0**	**5**	**4.7**	**4.8**
1-P(EII)	μ = 0.05	5.2	7.7	34.4
	μ = 0.1	6.4	16.7	88.3
	μ = 0.2	9.5	51.6	100
	μ = 1	88.4	100	100
	μ = 2	100	100	100

Tests of hypothesis: Power of the test

Probability of not making type II errors: simulation (3)

- The power of the test increases with the sample size
- The power of the test increases the further away is the true μ from the Null hypothesis μ
- For $n = 1000$ the null is rejected nearly always if DGP has $\mu < -0.1$ or $\mu > 0.1$
- Also: The smaller the probability of a Type 1 error, the greater the probability of Type II error (Show)
- Lesson: Choose the level of significance with care, and use as large a sample as possible

CAMBRIDGE Judge Business School