Research Methods for Political Science

MT week 4, lecture 2

Univariate statistics: confidence intervals & significance testing

TRINITY COLLEGE DUBLIN
COLÁISTE NA TRÍONÓIDE, BAILE ÁTHA CLIATH
THE UNIVERSITY OF DUBLIN

Dr. Thomas Chadefaux

Assistant Professor in Political Science

Thomas.chadefaux@tcd.ie

Last time: standard error

The standard error of the mean is
the standard deviation of
the sampling distribution of
sample means

Confidence interval

For a given statistic calculated for a sample of observations (e.g. sample mean) the confidence interval is a range of values around that statistic that are believed to contain, with a certain probability (e.g. 95%), the true value of that statistic (i.e. the population value).

Confidence interval

- The confidence interval is chosen so that it will contain the population mean 95% of the times.
- I.e., there is a 95% chance that the confidence interval contains the true population mean.

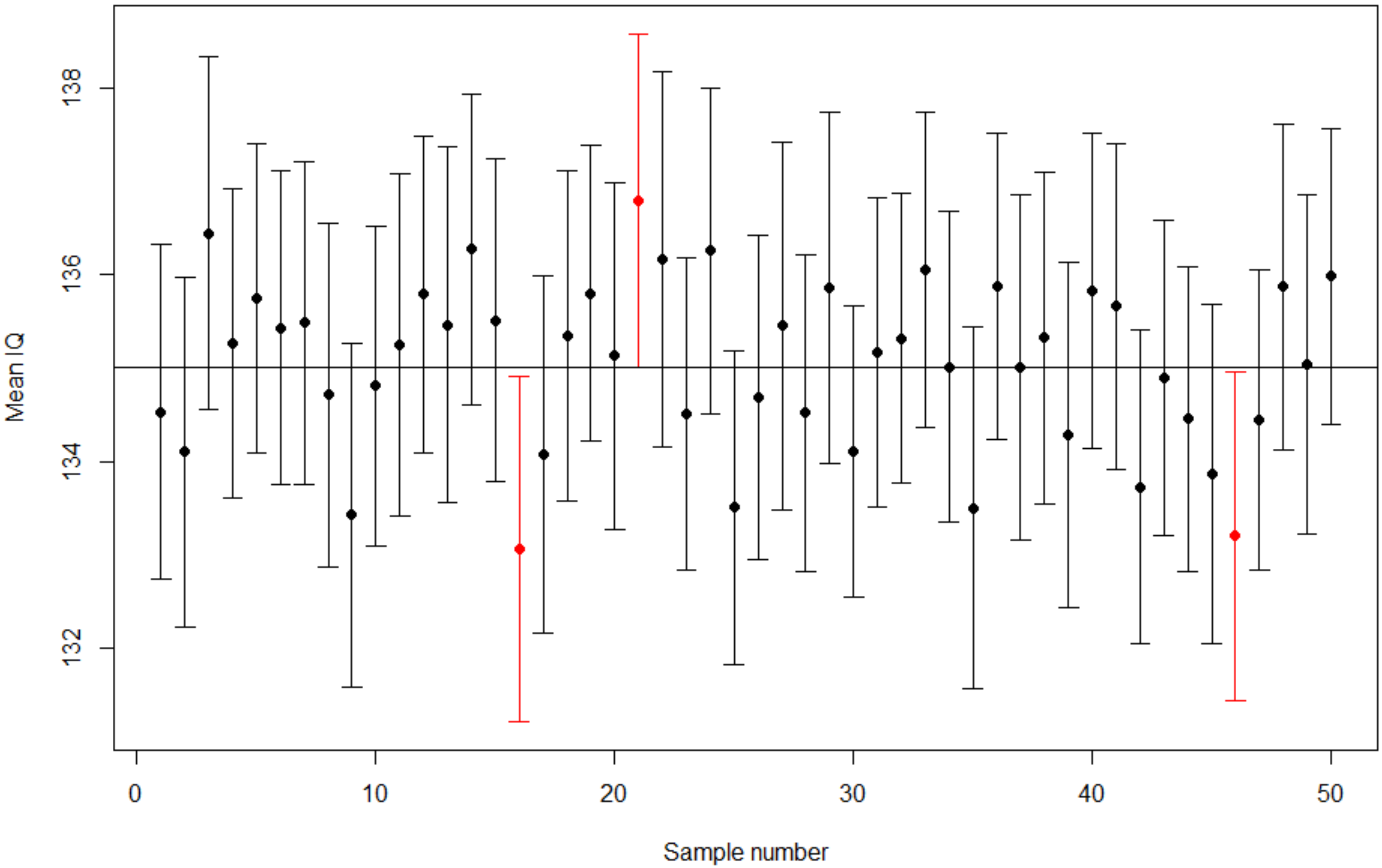
Interpreting the confidence interval

- **Wrong Interpretation**: "95% of the x are between 13 and 15."
  - Why it's wrong: the confidence interval is about the *population* mean. It is not about the sample
- **Wrong Interpretation** : "There is a 95% chance that the mean x is between 13 and 15".
  - Why it's wrong: the population mean value is fixed. So either it is or it is not in the interval, but there is no probability that it is.
- **Correct Interpretation**: we are 95% confident that the mean value of x is between 13 and 15

Example:

We know that the IQ of university students is 135 on average.

We draw 50 samples of 125 students and measure their IQ. Then we calculate the mean and 95% confidence interval for each sample.


Mean IQ
132
134
136
138
0
10
20
30
40
50
Sample number

A 95% confidence interval will contain the population mean 19 out of 20 times. But there is a 1 in 20 chance that the confidence interval for our sample would not include the population mean.

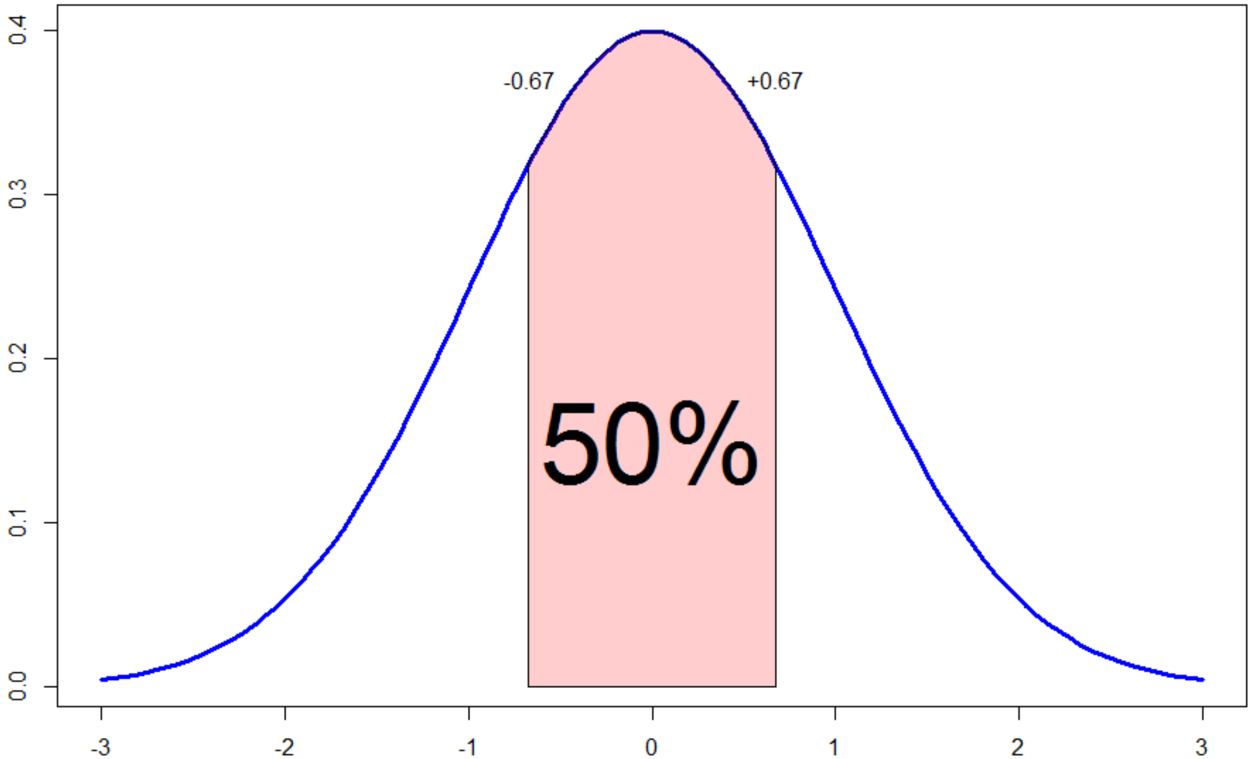
If we wish to be more confident about our results, we have to increase the width of our interval. E.g. a 99% confidence interval will contain the population mean in 99 of 100 cases.

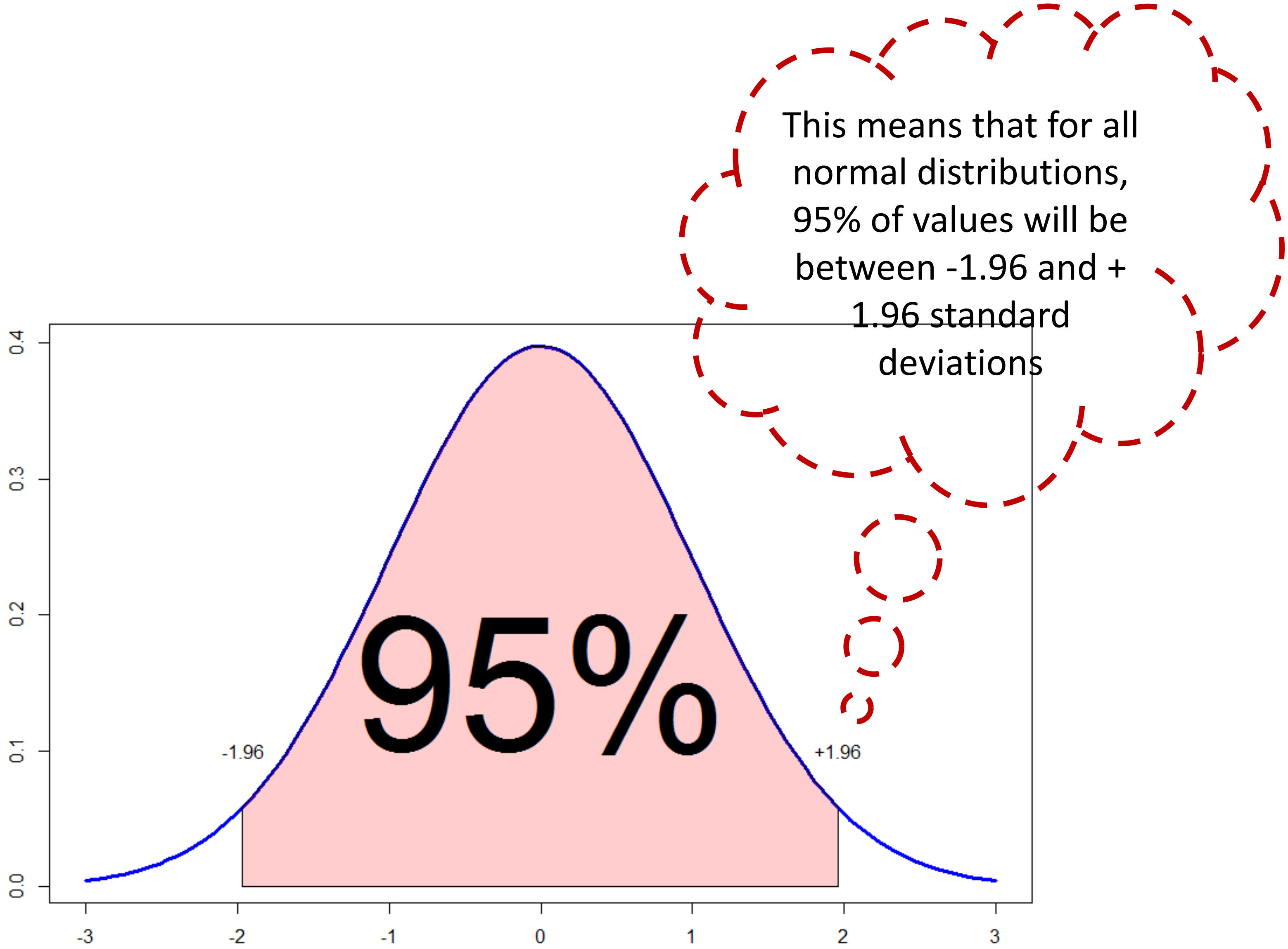
Why not always use the 99% CI, then?

HOW TO CALCULATE THE CONFIDENCE INTERVAL?

Standard Normal distribution

Mean = 0, Standard deviation = 1




This means that for all normal distributions, 95% of values will be between -1.96 and + 1.96 standard deviations
95%
-1.96
+1.96
0.0
0.1
0.2
0.3
0.4
-3
-2
-1
0
1
2
3

Standardizing Variables

- We know a lot about the standard normal distribution (mean = 0, SD = 1). But what if you data comes from a normal distribution with another mean or standard distribution?
- Luckily you can easily standardize your variable:

Z = (X-mean)/SD

- 90% probability: z-value = ± 1.68
- 95% probability: z-value = ± 1.96
- 99% probability: z-value = ± 2.58

Calculating the confidence interval

- A 95% confidence interval (large sample) is given by

$$CI = \bar{X} \boxed{\pm 1.96 * SE}$$

Why this? Because we want to find
the range within which 95% of
the data falls.
Remember that the Standard error is
the standard deviation of your statistic
(i.e. here the mean)

Example:

For one of our student samples we find a mean of 134 and a standard deviation of 10.

Calculate the 95% confidence interval.

Example (II)

- Sample mean = 134
- Sample standard deviation = 10
- Sample size (N) = 125

- Standard error = $\frac{s}{\sqrt{n}} = \frac{10}{\sqrt{125}} = \frac{10}{11.2}$= 0.89

Example (III)

$$CI = \bar{X} \pm 1.96 * SE$$

CI = 134 ± 1.96 * 0.89

CI_{low}= 134 – 1.96 * 0.89 = 132.25

CI_{high}= 134 + 1.96 * 0.89 = 135.75

Another example (I)

A random sample of 144 newspaper editorials is analysed and the political position of each editorial is estimated for each article. The position is measured on an interval scale, where -10 is very liberal and +10 is very conservative. The average score found was -1 with a standard deviation of 8.

Calculate a 95% confidence interval.

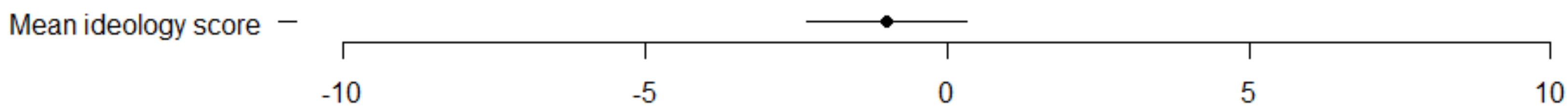
Another example (II)

- $CI = \bar{X} \pm 1.96 * SE$
- Standard error = $\frac{s}{\sqrt{n}} = \frac{8}{\sqrt{144}} = \frac{8}{12}$= 0.67

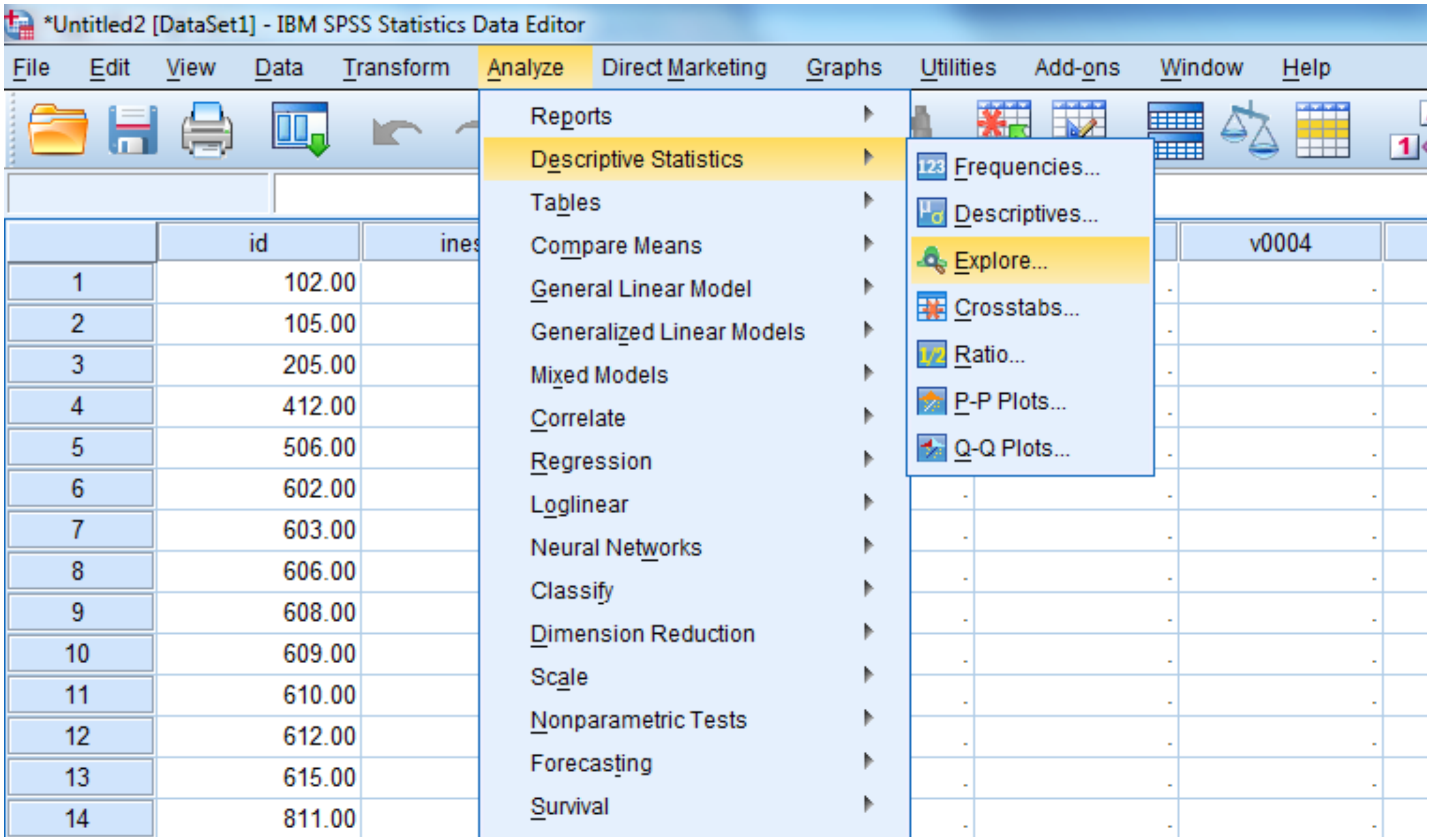
CI_{low}= -1 - 1.96 * 0.67 = -2.30

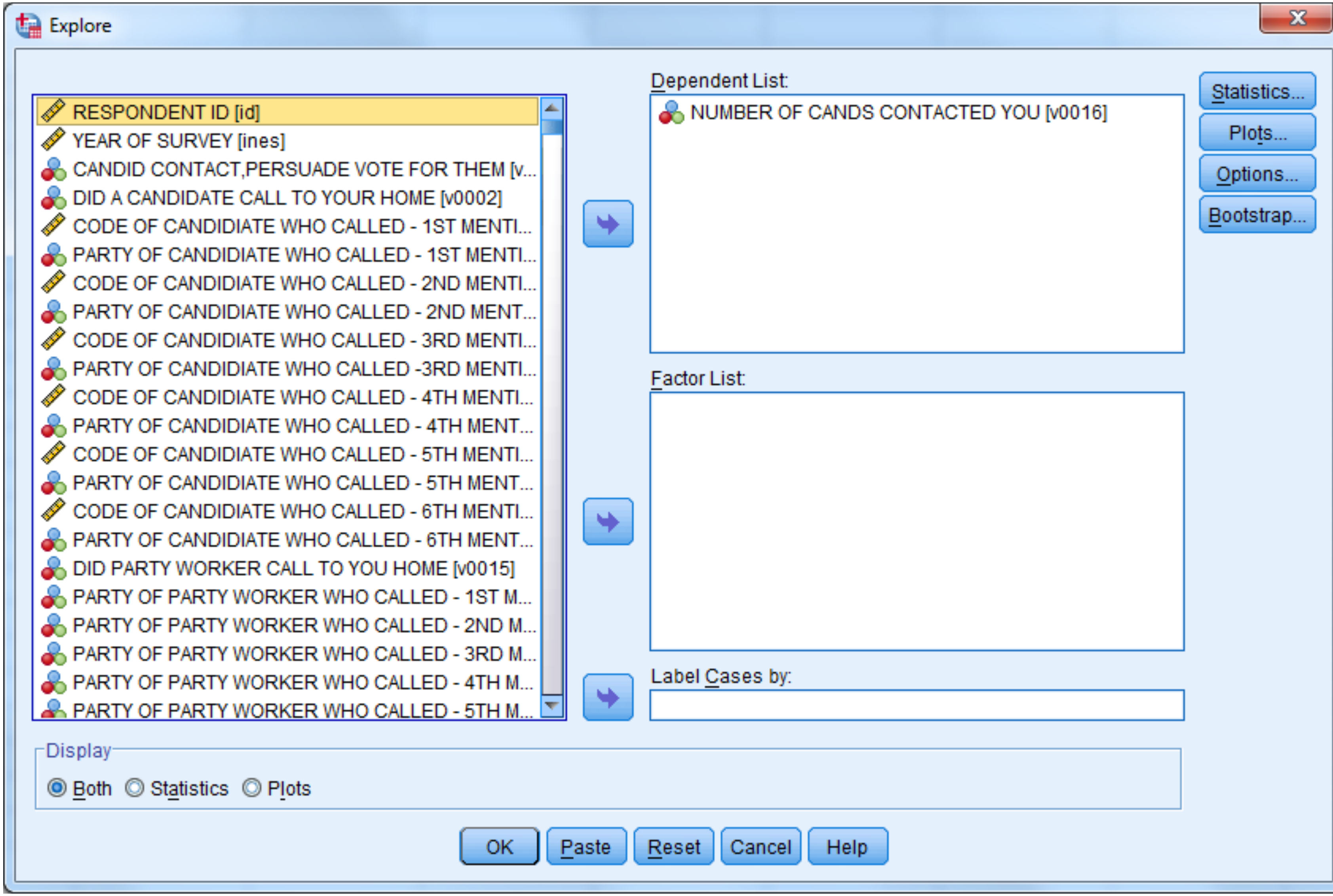
CI_{high}= -1 + 1.96 * 0.67 = 0.31

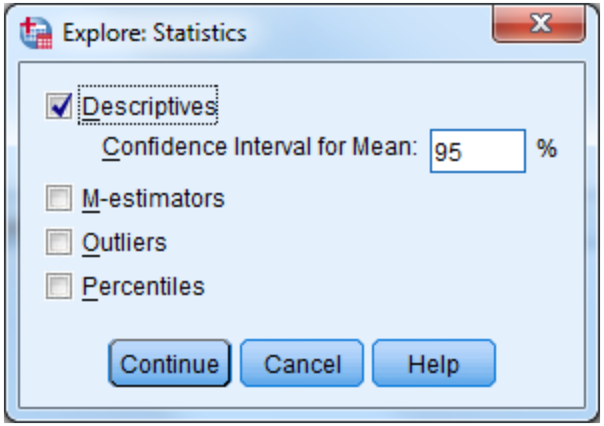
Biased newspaper?



Confidence intervals in SPSS




Explore
RESPONDENT ID [id]
YEAR OF SURVEY [ines]
CANDID CONTACT,PERSUADE VOTE FOR THEM [v...
DID A CANDIDATE CALL TO YOUR HOME [v0002]
CODE OF CANDIDATE WHO CALLED - 1ST MENTI...
PARTY OF CANDIDATE WHO CALLED - 1ST MENTI...
CODE OF CANDIDATE WHO CALLED - 2ND MENTI...
PARTY OF CANDIDATE WHO CALLED - 2ND MENT...
CODE OF CANDIDATE WHO CALLED - 3RD MENTI...
PARTY OF CANDIDATE WHO CALLED -3RD MENTI...
CODE OF CANDIDATE WHO CALLED - 4TH MENTI...
PARTY OF CANDIDATE WHO CALLED - 4TH MENT...
CODE OF CANDIDATE WHO CALLED - 5TH MENTI...
PARTY OF CANDIDATE WHO CALLED - 5TH MENT...
CODE OF CANDIDATE WHO CALLED - 6TH MENTI...
PARTY OF CANDIDATE WHO CALLED - 6TH MENT...
DID PARTY WORKER CALL TO YOU HOME [v0015]
PARTY OF PARTY WORKER WHO CALLED - 1ST M...
PARTY OF PARTY WORKER WHO CALLED - 2ND M...
PARTY OF PARTY WORKER WHO CALLED - 3RD M...
PARTY OF PARTY WORKER WHO CALLED - 4TH M...
PARTY OF PARTY WORKER WHO CALLED - 5TH M...
Dependent List:
NUMBER OF CANDS CONTACTED YOU [v0016]
Factor List:
Label Cases by:
Statistics...
Plots...
Options...
Bootstrap...
Display
Both
Statistics
Plots
OK
Paste
Reset
Cancel
Help


Explore: Statistics
Descriptives
Confidence Interval for Mean:
95
%
M-estimators
Outliers
Percentiles
Continue
Cancel
Help

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
NUMBER OF CANDS CONTACTED YOU	816	7.8%	9607	92.2%	10423	100.0%

Descriptives

			Statistic	Std. Error
NUMBER OF CANDS CONTACTED YOU	Mean		1.8186	.04110
	95% Confidence Interval for Mean	Lower Bound	1.7380	
		Upper Bound	1.8993	
	5% Trimmed Mean		1.7456	
	Median		2.0000	
	Variance		1.378	
	Std. Deviation		1.17393	
	Minimum		.00	
	Maximum		6.00	
	Range		6.00	
	Interquartile Range		1.00	
	Skewness		1.250	.086
	Kurtosis		1.790	.171

Check by hand

			Statistic	Std. Error
NUMBER OF CANDS CONTACTED YOU	Mean		1.8186	.04110
	95% Confidence Interval for Mean	Lower Bound	1.7380	
		Upper Bound	1.8993	

CI_{low}= Mean – 1.96 * SE

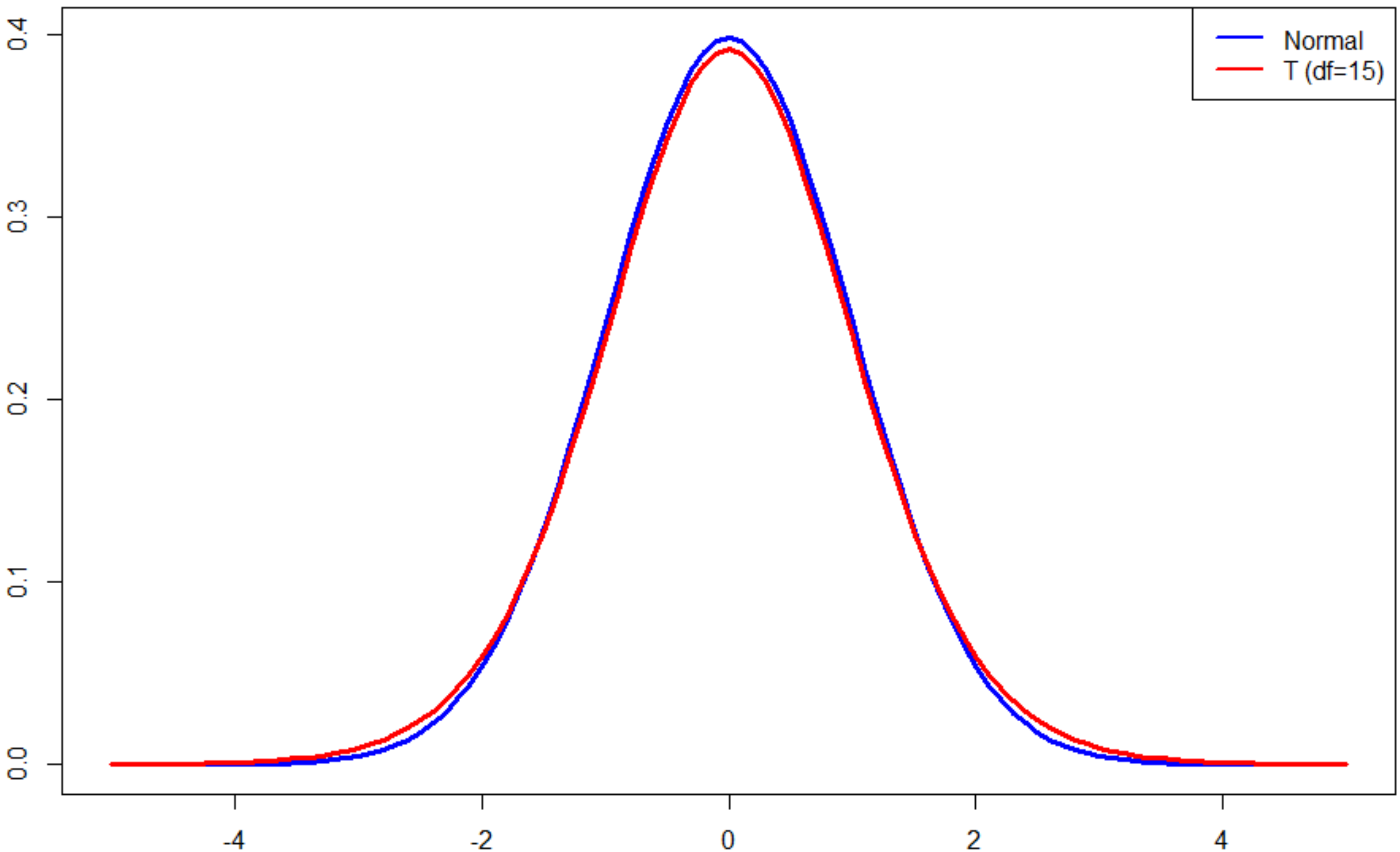
CI_{low}=1.82 – 1.96 * 0.04 = 1.74

CI_{high}= Mean + 1.96 * SE

CI_{high}=1.82 + 1.96 * 0.04 = 1.90

Normal or t distribution?

- SPSS uses the t-distribution to calculate confidence intervals
- Reason: we use *s* (not σ) to calculate the standard error; matters for small samples
- The t-distribution is (virtually) equal to the normal distribution when N > 100
- For smaller N the t-distribution is a better approximation of the sampling distribution
- To be safe: always use t-distribution (SPSS does)


Normal
T (df=15)
0.0
0.1
0.2
0.3
0.4
-4
-2
0
2
4

Example (small N)

Number of casualties in random selection of months in Iraq (2003-2012):

N = 20

Mean = 1280.60

Standard deviation = 865.88

Standard error

- We have to use the *t* distribution with N – 1 degrees of freedom: t(19) = 2.093
- Intuitively, 2.093 means that 95% of the data is within (-2.093, +2.093)
- SE = s / sqrt(N) = 865.88 / sqrt(20) = 193.62

CI = 1280.60 ± 2.093 * 193.62

CI = {875.35; 1685.85}

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
bodycount	20	100.0%	0	0.0%	20	100.0%

Descriptives

			Statistic	Std. Error
bodycount	Mean		1280.6000	193.61567
	95% Confidence Interval for Mean	Lower Bound	875.3577	
		Upper Bound	1685.8423	
	5% Trimmed Mean		1251.5556	
	Median		1116.5000	
	Variance		749740.568	
	Std. Deviation		865.87561	
	Minimum		275.00	
	Maximum		2809.00	
	Range		2534.00	
	Interquartile Range		1708.25	
	Skewness		.551	.512
	Kurtosis		-1.053	.992

The t-test

William Sealy Gosset (1876-1937)

Testing hypotheses

- Testing hypotheses
  - Null hypothesis (H_0)
  - Alternative hypothesis (H_1)

- E.g.:
  - H_0: TCD students have the same IQ as UCD students
  - H_1: TCD students' IQ differs from UCD students' IQ

Null hypothesis

- We ask: **if the null hypothesis were true**, how likely would we be to collect the data we have?

The one-sample case

Comparing a sample statistic against a known population parameter.

E.g. general population IQ is 100. Random sample of 125 high school pupils. We find M = 125, SD = 10.

H_0: Pupils' IQ equals general population IQ.

H_1: Pupils' IQ differs from general population IQ.

Example: IQ

<u>Step 1: stating the null hypothesis</u>

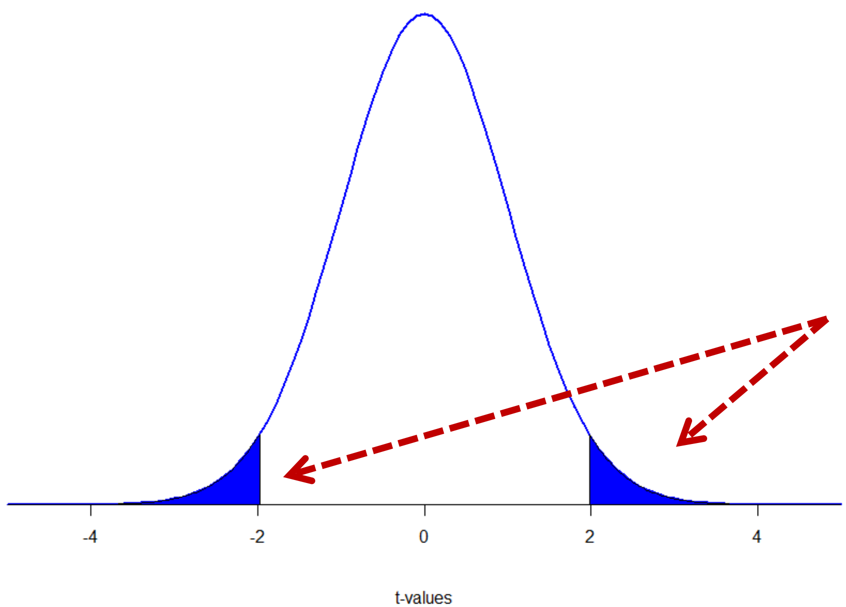
H0: $\mu = 100$

(H1: $\mu \neq 100$)

Example: student IQ

<u>Step 2: Selecting the Sampling Distribution and Establishing the Critical Region</u>

Population SD is unknown: t-distribution with n – 1 degrees of freedom (n=125, df=124). $\alpha = 0.05$



Critical regions:
T > 1.96 or
T < -1.96

APPENDIX

A.2. Critical values of the *t*-distribution

	Two-Tailed Test		One-Tailed Test	
df	0.05	0.01	0.05	0.01
1	12.71	63.66	6.31	31.82
2	4.30	9.92	2.92	6.96
3	3.18	5.84	2.35	4.54
4	2.78	4.60	2.13	3.75
5	2.57	4.03	2.02	3.36
6	2.45	3.71	1.94	3.14
7	2.36	3.50	1.89	3.00
8	2.31	3.36	1.86	2.90
9	2.26	3.25	1.83	2.82
10	2.23	3.17	1.81	2.76
11	2.20	3.11	1.80	2.72
12	2.18	3.05	1.78	2.68
13	2.16	3.01	1.77	2.65
14	2.14	2.98	1.76	2.62
15	2.13	2.95	1.75	2.60
16	2.12	2.92	1.75	2.58

Step 3: Computing the test statistic

The general form of the t-statistic is:

T= (observed value – Expected value under H0) / (estimate of the standard error)

Note: compare to formula for standardized values:

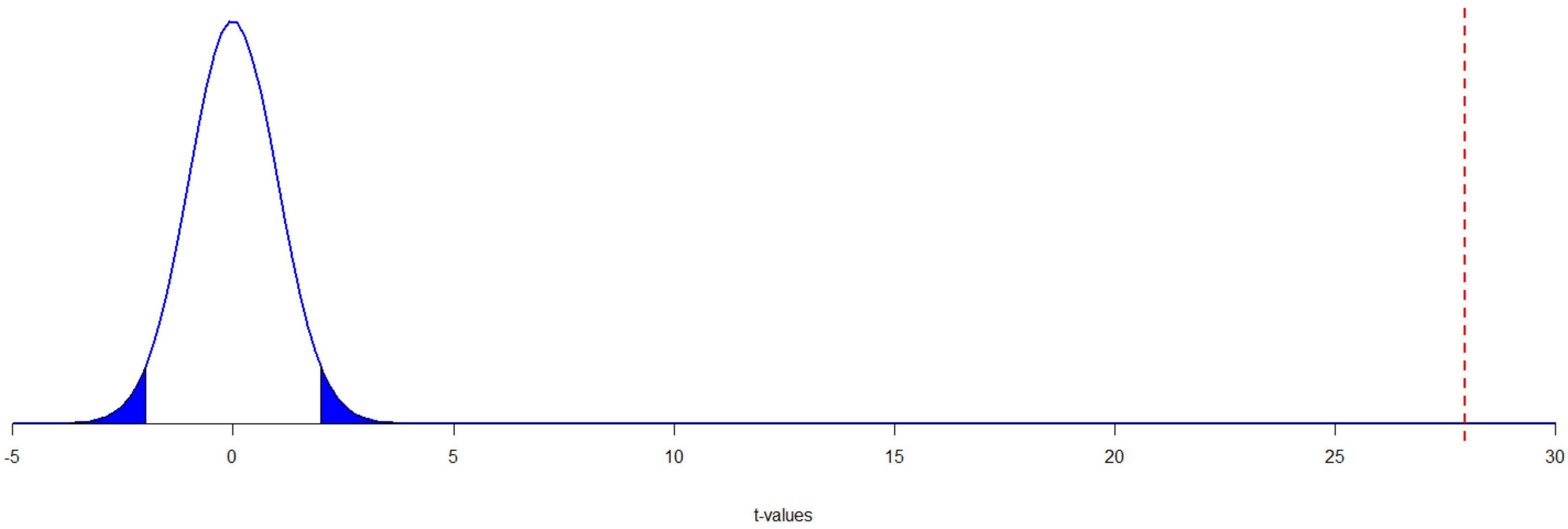
Z = (observed value – mean)/ standard deviation

Step 3: Computing the test statistic

For the one-sample case

$$t = \frac{X - \mu}{s / \sqrt{n}}$$

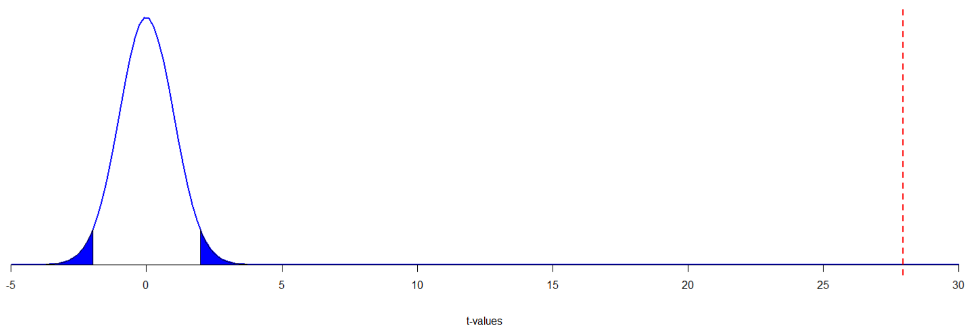
$$t = \frac{125 - 100}{10 / \sqrt{125}} = 29.95$$


-5
0
5
10
15
20
25
30
t-values

Step 5: Making a decision

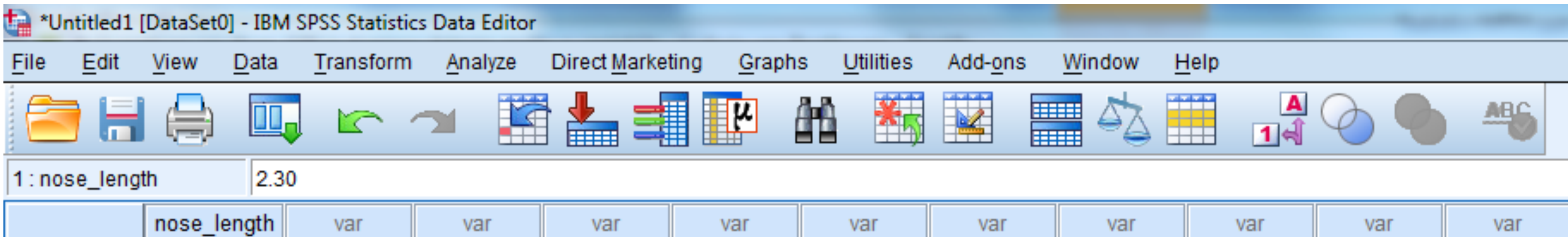
t(obtained) = 29.98

t(critical) = +/- 1.96



As t(obtained) fell in the *critical region,* we have to reject the null hypothesis.

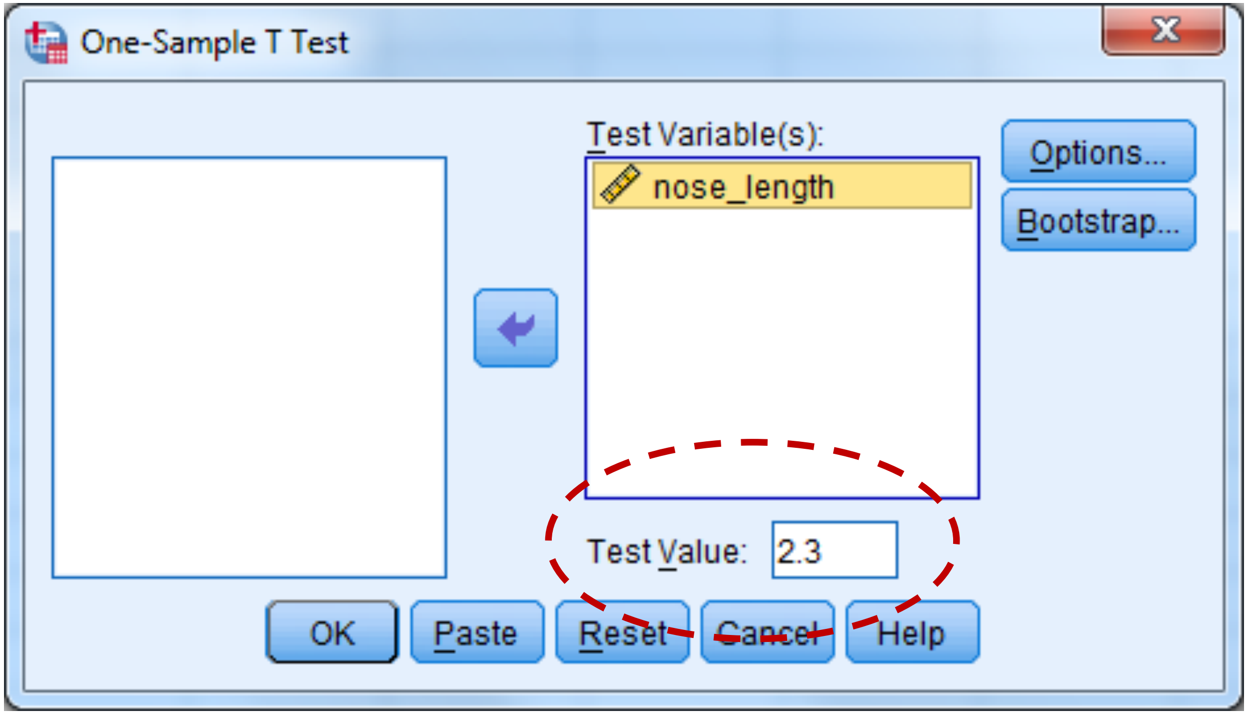
One-sample t-test in SPSS



	nose_length	var	var	var	var	var	var	var	var	var	var
1	2.30										
2	2.40										
3	2.20										
4	2.10										
5	2.40										
6	2.30										
7	2.20										
8	2.50										
9	1.70										
10	1.90										
11	2.00										
12	2.30										
13	2.40										
14	2.50										
15	2.30										
16	2.10										
17	2.00										
18	1.94										
19	1.99										

One-sample t-test in SPSS

- Analyze .. Compare Means ... One sample t-test



One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
nose_length	20	2.1890	.21665	.04844

One-Sample Test

	Test Value = 2.3					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
nose_length	-2.291	19	.034	-.11100	-.2124	-.0096

P-value

- SPSS reports a “Sig. (2 tailed)” value. This is often called the **p-value**, and denotes the probability that this value of the mean (or more extreme) would have been observed by chance, had the null hypothesis been true
- E.g., $p=0.05$ means that there is a 5% chance that this value would have been observed by sheer chance, had the null hypothesis been true

Wrap-up: an example

- Suppose a group of 100 people is exposed to a TV ad promoting diversity, after which this group's attitude towards diversity is evaluated. We know that the overall attitude in the population is +1. The mean of the treated group is +1.3, with a sample standard deviation of 1. Do you think the ad had an effect?
- H0: treatment has no effect: mean=1, even with treatment
- H1: treatment has an effect: mean not =1 with the treatment.

- SE = s/sqrt(100) = 1/10 = 0.1
- t test: (1.3 – 1)/0.1 = 0.2/0.1 = 3
- Check p-value in your book. Because 3> 1.96 (the value in the book), we can reject the null hypothesis at the 0.05 significance level

A.2. Critical values of the *t*-distribution

	Two-Tailed Test		One-Tailed Test	
df	0.05	0.01	0.05	0.01
1	12.71	63.66	6.31	31.82
2	4.30	9.92	2.92	6.96
3	3.18	5.84	2.35	4.54
4	2.78	4.60	2.13	3.75
5	2.57	4.03	2.02	3.36
6	2.45	3.71	1.94	3.14
7	2.36	3.50	1.89	3.00
8	2.31	3.36	1.86	2.90
9	2.26	3.25	1.83	2.82
10	2.23	3.17	1.81	2.76
11	2.20	3.11	1.80	2.72
12	2.18	3.05	1.78	2.68
13	2.16	3.01	1.77	2.65
14	2.14	2.98	1.76	2.62
15	2.13	2.95	1.75	2.60
16	2.12	2.92	1.75	2.58

Testing hypotheses

- Testing hypotheses
  - Alternative hypothesis (H_1)
  - Null hypothesis (H_0)

Null hypothesis

- We ask: if the null hypothesis were true, how likely would we be to collect the data we have?

The one-sample case

Comparing a sample statistic against a known population parameter.

E.g. population IQ is 100. Random sample of 125 high school pupils. We find M = 105, SD = 10.

H_1: Pupils' IQ differs from population IQ.

H_0: Pupils' IQ differs equals population IQ.

Example: IQ

Step 1: Assumptions

Random sampling

Level of measurement interval-ratio

Sampling distribution is normal

Example: IQ

<u>Step 2: stating the null hypothesis</u>

H0: $\mu = 100$

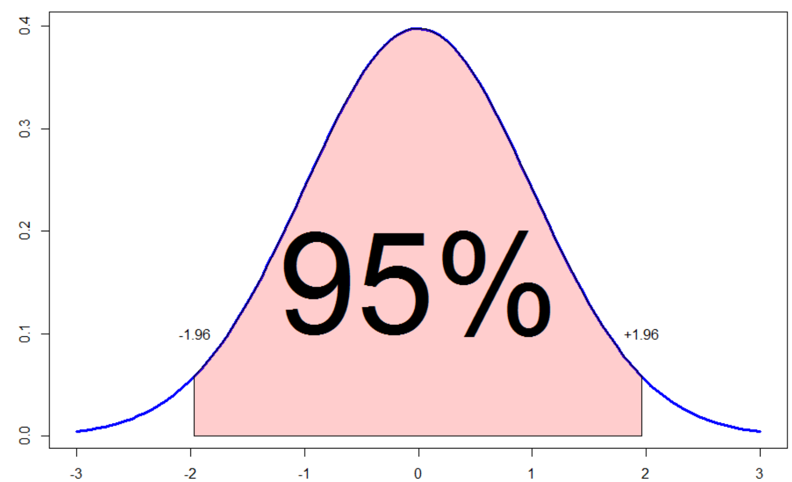
(H1: $\mu \neq 100$)

Example: student IQ

<u>Step 3: Selecting the Sampling Distribution and Establishing the Critical Region</u>

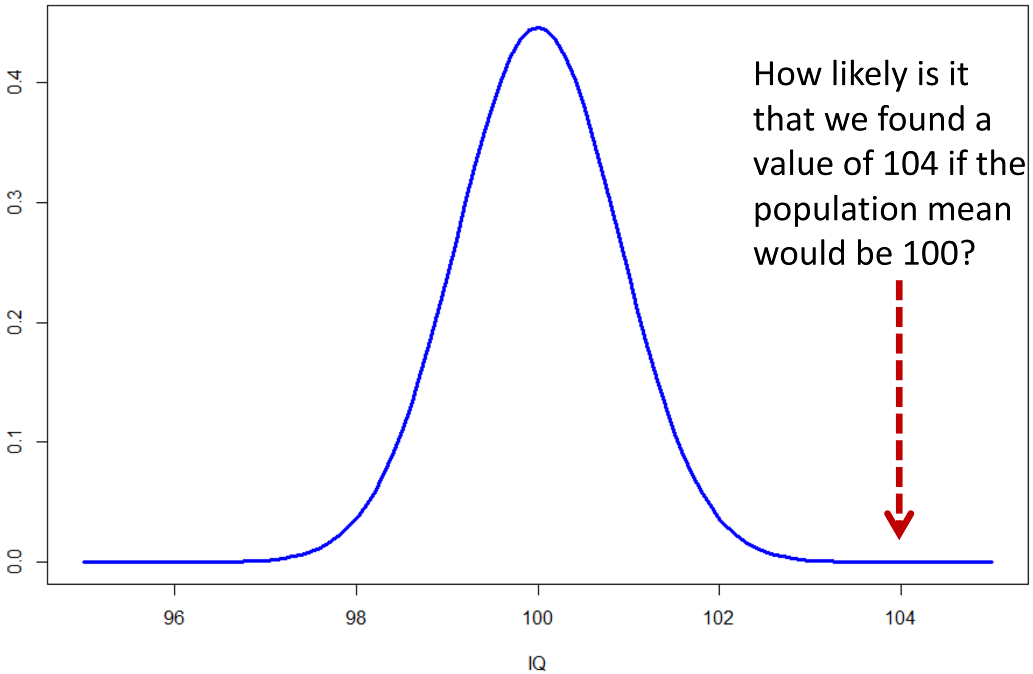
N is large, so we use the (standard) normal distribution

+/- 1.96 is the <u>critical value</u>



Step 4: Computing the test statistic

If H_0 is true, then the sampling distribution of the mean looks like this:



Mean = 100

Standard deviation = Standard error =

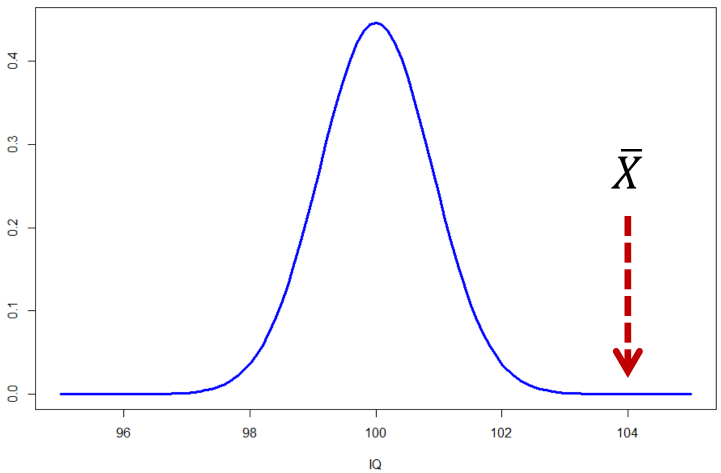
$= \frac{s}{\sqrt{n}} = \frac{10}{\sqrt{125}} = 0.89$

Standardizing variables

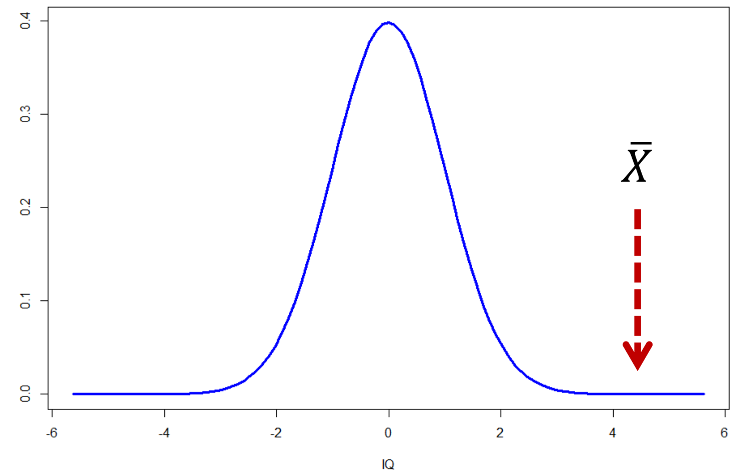
We know a lot about the <u>standard</u> normal distribution (mean = 0, standard deviation=1), so we must standardize our variables before the analysis.

$$Z = \frac{X - \mu}{s/\sqrt{N}}$$

Unstandardized values



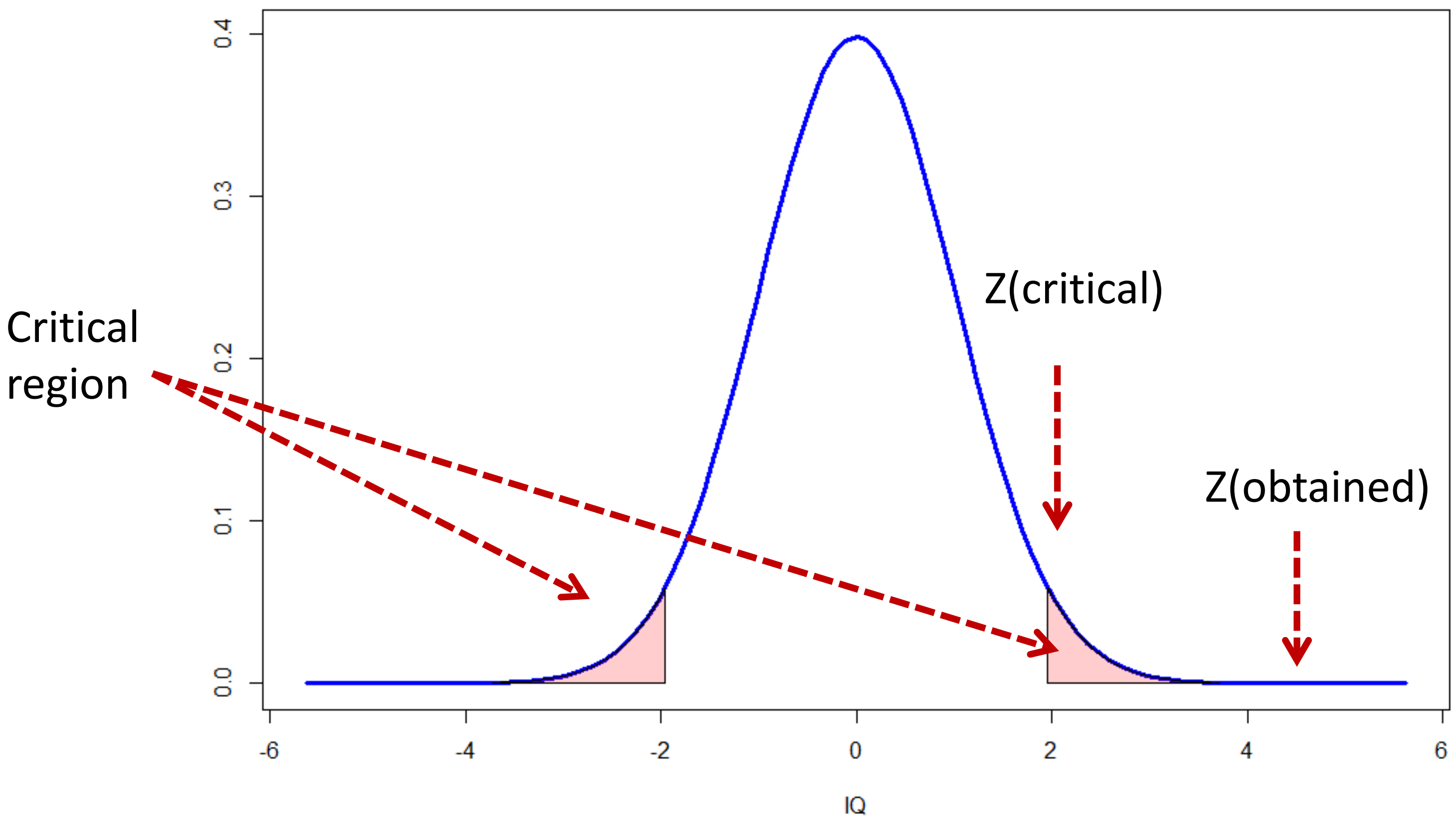
Standardized values



$\mu_{\bar{x}}$ = 100 $\dashrightarrow$ $Z = \frac{X - \mu}{s/\sqrt{N}}$ $\dashrightarrow$ $Z(\mu_{\bar{x}}) = \frac{100-100}{10/\sqrt{125}} = 0$

$\bar{X}$=104 $\dashrightarrow$ $Z = \frac{X - \mu}{s/\sqrt{N}}$ $\dashrightarrow$ $\mathrm{Z}(\bar{X}) = \frac{104-100}{10/\sqrt{125}} = 4.47$

This is called the test statistic, or Z(obtained)


Critical region
Z(critical)
Z(obtained)
0.4
0.3
0.2
0.1
0.0
-6
-4
-2
0
2
4
6
IQ

<u>Step 5: Making a decision</u>

Z(obtained) = 4.47

Z(critical) = +/- 1.96

As Z(obtained) fell in the *critical region,* we have to reject the null hypothesis.

Confidence interval and testing

These are two sides of the same coin: both are an estimate of the probability that our data would originate from a sampling distribution with a certain mean.