Revision section: the basics

R1.3 Confidence intervals: a way of precisely representing uncertainty

A whole range of descriptive statistics can be produced by SPSS for a variable, including the confidence interval. Below we give an example for the exam results described in Box R1.1.

SPSS COMMANDS FOR BOX R1.1 **Descriptive statistics and confidence intervals for single samples**	
Syntax	`examine EXAMRES.`
Menu route	Analyze > Descriptive Statistics > Explore `EXAMRES` → Dependent List

SPSS OUTPUT FOR BOX R1.1 **Descriptive statistics and confidence intervals for single samples**

Explore

Total sample

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
EXAMRES	30	60.0%	20	40.0%	50	100.0

Descriptives

			Statistic	Std. Error
EXAMRES	Mean		50.33	2.895
	95% Confidence Interval for Mean	Lower Bound	44.41	
		Upper Bound	56.25	
	5% Trimmed Mean		50.41	
	Median		50.50	
	Variance		251.402	
	Std. Deviation		15.856	
	Minimum		18	
	Maximum		83	
	Range		65	
	Interquartile Range		26.75	
	Skewness		–.182	.427
	Kurtosis		–.577	.833

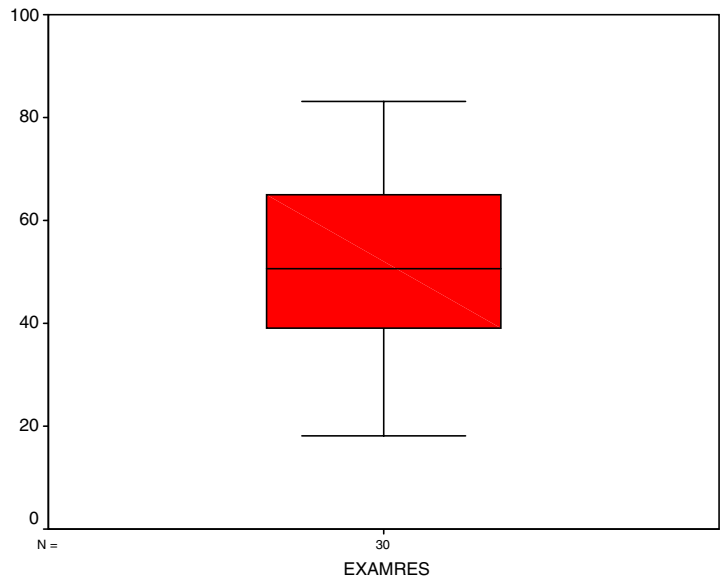
Examres

```
EXAMRES Stem-and-Leaf Plot

 Frequency    Stem &  Leaf

     1.00        1 .  8
     4.00        2 .  5889
     3.00        3 .  469
     5.00        4 .  15679
     7.00        5 .  0011369
     8.00        6 .  12556778
     1.00        7 .  1
     1.00        8 .  3

 Stem width:        10
 Each leaf:       1 case(s)
```



SPSS produces a wealth of information in addition to the confidence interval. However, you can see that under '95% Confidence Interval for Mean' is the same interval as that produced in Box R1.1, and calculated in the main text. SPSS has also drawn a Stem-and-leaf plot, which indicates the four quartiles of the data.

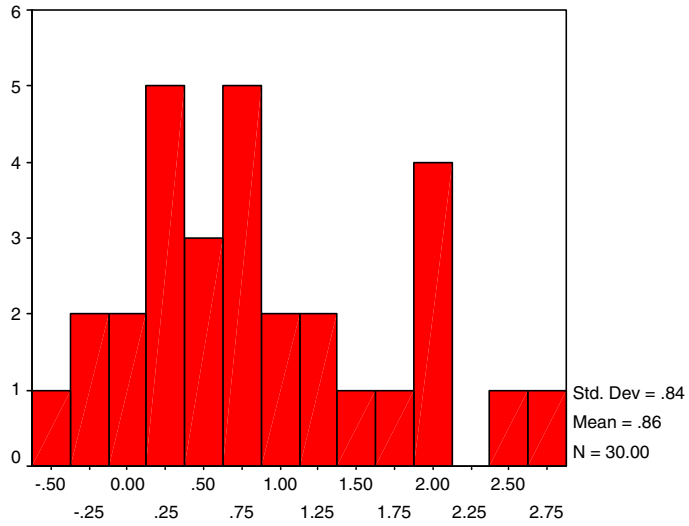
R1.4 The null hypothesis — taking the conservative approach

In this section, a one sample t-test is conducted on the data variable `DIFF`. First, a histogram of the data is plotted.

SPSS COMMANDS FOR FIGURE R1.6	
Syntax	`graph` `  /histogram DIFF.`
Menu route	Graphs > Histogram `DIFF` → Variable

SPSS OUTPUT FOR FIGURE R1.6

Graph



Then the one sample t-test is conducted:

SPSS COMMANDS FOR BOX R1.2 **One sample t-test**

Syntax	`t-test`
	`  /testval 0`
	`  /variables DIFF.`

Menu route	Analyze > Compare Means > One-Sample T Test `DIFF` → Test Variable(s)

SPSS OUTPUT FOR BOX R1.2 **One sample t-test**

T-Test

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
DIFF	30	.8621973	.83785035	.15296985

One-Sample Test

	Test Value = 0					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
DIFF	5.636	29	.000	.8621973	.5493389	1.1750558

R1.5 Comparing two means

Two sample t-test

The two sample t-test is conducted on the *squirrels* dataset. First the data are transformed to natural logarithms.

SPSS COMMANDS TO TRANSFORM DATA

Syntax

```
compute LMALE= ln(MALE).
compute LFEM = ln(FEMALE).
execute .
```

Menu route

Transform > Compute

`LMALE` → Target Variable

`ln(MALE)` → Numeric Expression

Transform > Compute

`LFEM` → Target Variable

`ln(FEMALE)` → Numeric Expression

Both the raw data and the logged data are plotted — but here we just show how to do this for the logged data. The SCALERANGE subcommand allows you to specify the bar divisions, and so ensure the two histograms are plotted on the same X scale. In the menu route, in the initial 'Assign Variables' window, the variables are moved to the X axis pane by dragging and dropping (rather than clicking on an arrow as in most other windows).

SPSS COMMANDS FOR FIGURE R1.7 **Histograms of Log_e body mass of male and female squirrels**

Syntax

```
igraph
  /x1 = var(LFEM) type = scale
  /y = $count
  /scalerange = var(LFEM) min=-1.25 max=0.15
  /histogram x1interval width = 0.1 .

igraph
  /x1 = var(LMALE) type = scale
  /y = $count
  /scalerange = var(LMALE) min=-1.25 max=0.15
  /histogram x1interval width = 0.1 .
```

Menu route

Graphs >Interactive > Histogram

Assign Variables

LFEM → (X Axis)

Histogram

– ☑ Set interval size automatically

0.1 → Width of intervals

Options

Scale Range

Variable: LFEM ▼

– ☑ Auto

–1.25 → Minimum

Graphs >Interactive > Histogram

Assign Variables

LMALE → (X Axis)

Histogram

– ☑ Set interval size automatically

0.1 → Width of intervals

Options

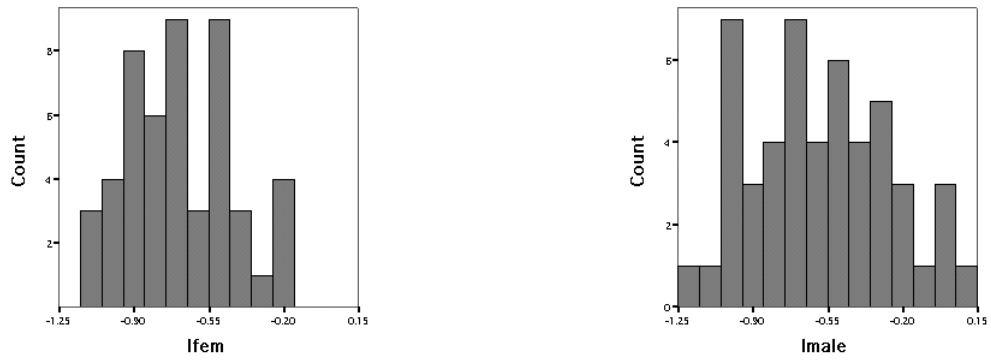
Scale Range

Variable: LMALE ▼

– ☑ Auto

–1.25 → Minimum

SPSS OUTPUT FOR FIGURE R1.7 **Histograms of Log_e body mass of male and female squirrels**



Finally, the two sample t-test. In this case, as in the main text, we have used the 'pooled' method to calculated the standard errors. In other words, we have assumed that one error variance can be estimated for the two samples. SPSS needs to have the male and female data in the same variable, and to have a second variable containing a code for sex.

To do this, create a new variable called `LOGWT`, by going to the 'Variable View' window using the tab at the bottom left of the Data Window. Type the new variable name in the first cell of the first empty row in the worksheet. Then go back to the 'Data View' window, select the male values from 1 to 50, and copy-and-paste into cell 1 of `LOGWT`. Then select the female values from 1 to 50, and copy-and-paste into cell 51 of `LOGWT`. The extra variable is created by computing SEX and giving it the value `1+MISSING(MALE)`. This gives the value of 1 to the first fifty datapoints, and 2 to the rest, because `MISSING()` equals 1 for rows where `MALE` is missing and zero otherwise. Then we can do the test.

SPSS COMMANDS FOR BOX R1.6 **Two sample t-test**

Syntax

```
compute SEX = 1 + MISSING(MALE).
execute.

t-test groups=SEX(1.0,2.0)
  /variable LOGWT.
```

Menu route

Transform > Compute

SEX → Target Variable

1 + MISSING(MALE) → Numeric Expression

Analyze > Compare Means > Independent-Samples T Test

LOGWT → Test Variable

SEX → Grouping Variable

Define Groups

1.0 → Group 1

2.0 → Group 2

SPSS OUTPUT FOR BOX R1.6 **Two sample t-test**

T-Test

Group Statistics

	SEX	N	Mean	Std. Deviation	Std. Error Mean
LOGWT	1.00	50	–.5786	.32481	.04594
	2.00	50	–.6878	.24526	.03468

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
LOGWT	Equal variances assumed	3.974	.049	1.898	98	.061	.1092	.05756	–.00500	.22345
	Equal variances not assumed			1.898	91.166	.061	.1092	.05756	–.00511	.22356

One and two-tailed tests

SPSS will not perform one-tailed t-tests, and so we need to do it by hand. As explained in the main text the only difference comes into play when we convert the t-ratio into a p-value. The SPSS Syntax Guide states in the documentation of T_TEST that "Probability levels are two-tailed. To obtain the one-tailed probability, divide the two-tailed probability by 2.". However, this advice ignores direction and, taken literally, would double your chance of making a Type I error (see main text). Therefore, if you wish to conduct a one-tailed test in SPSS you should follow these steps: 1) Set up your alternative hypothesis (before doing the experiment!); 2) Calculate your t-ratio as if for a two-tailed test; 3) Compare the sign of your t-ratio with your alternative hypothesis. If the difference is in the opposite direction to H_A, then you cannot reject your null hypothesis regardless of the size of your t-ratio; 4) If the difference is in the same direction to H_A, then divide the two-tailed probability by 2.