

Chapter 29

Robust Testing (33 Patients)

1 General Purpose

Robust tests are tests that can handle the inclusion into a data file of some outliers without largely changing the overall test results. The following robust tests are available.

1. Z-test for medians and median absolute deviations (MADs).
2. Z-test for Winsorized variances.
3. Mood's test.
4. Z-test for M-estimators with bootstrap standard errors.

The first three can be performed on a pocket calculator and are reviewed in *Statistics on a Pocket Calculator Part 2, Chapter 8, Springer New York, 2011*, from the same authors. The fourth robust test is reviewed in this chapter.

2 Schematic Overview of Type of Data File

Outcome
.
.
.
.
.
.
.

3 Primary Scientific Question

Is robust testing more sensitive than standard testing of imperfect data.

4 Data Example

The underneath study assesses whether physiotherapy reduces frailty. Frailty score improvements after physiotherapy are measured. The data file is underneath.

Frailty score improvements after physiotherapy

-8,00
-8,00
-8,00
-4,00
-4,00
-4,00
-4,00
-1,00
0,00
0,00

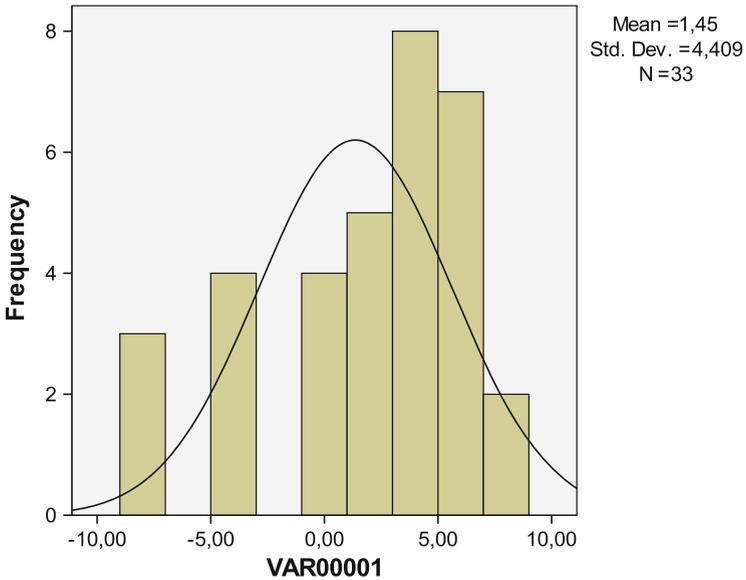
The above data give the first 10 patients, the entire data file is in “chapter29robusttesting”, and is in extras.springer.com. First, we will try and make a histogram of the data.

5 Data Histogram Graph

Command:

Graph. ...Legacy Dialogs. ...Histogram. ...Variable: frailty score improvement. ...

Mark: Display normal Curve. ...click OK.



The above graph suggests the presence of some central tendency: the values between 3,00 and 5,00 are observed more frequently than the rest. However, the Gaussian curve calculated from the mean and standard deviation does not fit the data very well with outliers on either side. Next, we will perform a one sample t-test to see if the calculated mean is significantly different 0. For analysis the statistical model One Sample T-Test in the module Compare Means is required.

Command:

Analyze. ...Compare Meams. ...One Sample T-Test. ...Test Variable: frailty score improvement. ...click OK.

One-sample test

	Test value = 0					
					95 % confidence interval of the difference	
	t	df	Sig. (2-tailed)	Mean difference	Lower	Upper
VAR00001	1,895	32	,067	1,45455	-,1090	3,0181

The above table shows that the t-value based on Gaussian-like t-curves is not significantly different from 0, $p = 0,067$.

6 Robust Testing

M-estimators is a general term for maximum likelihood estimators (MLEs), which can be considered as central values for different types of sampling distributions.

Huber (Proc 5th Berkeley Symp Stat 1967) described an approach to estimate MLEs with excellent performance, and this method is, currently, often applied. The Huber maximum likelihood estimator is calculated from the underneath equation (MAD = median absolute deviation, * = sign of multiplication)

$$\frac{\sum 0.6745 * (x - \text{median})}{\text{MAD}}$$

Command:

Analyze... Descriptives... Explore: enter variable into box dependent list...
Statistics: mark M-estimators... click OK.

In the output sheets the underneath result is given.

Huber's M-estimator = 2,4011
Huber's standard error = not given.

Usually, the 2nd derivative of the M-estimator function is used to find the standard error. However, the problem with the second derivative procedure in practice is that it requires very large data files in order to be accurate. Instead of an inaccurate estimate of the standard error, a bootstrap standard error can be calculated. This is not provided in SPSS. Bootstrapping is a data based simulation process for statistical inference. The basic idea is sampling with replacement in order to produce random samples from the original data. Standard errors are calculated from the 95 % confidence intervals of the random samples [95 % confidence interval = (central value \pm 2 standard errors)]. We will use "R bootstrap Plot – Central Tendency", available on the Internet as a free calculator tool.

Enter your data.

Then command: compute.

The bootstrap standard error of the median is used.

Bootstrap standard error = 0,8619.

The z-test is used.

z-value = Huber's M-estimator/bootstrap standard error

z-value = 2,4011/ 0,8619 = 2,7858

p-value = 0,005

Unlike the one sample t-test, the M-estimator with bootstraps produces a highly significant effect. Frailty scores can, obviously, be improved by physiotherapy.

7 Conclusion

Robust tests are wonderful for imperfect data, because they often produce statistically significant results, when the standard tests do not.

8 Note

The robust tests that can be performed on a pocket calculator, are reviewed in *Statistics on a Pocket Calculator Part 2, Chapter 8*, Springer New York, 2011, from the same authors.