Cut points – a statistical perspective

Simon Cowen (LGC, Teddington, UK)

Science

for a safer world

http://www.lgcgroup.com/

Outline

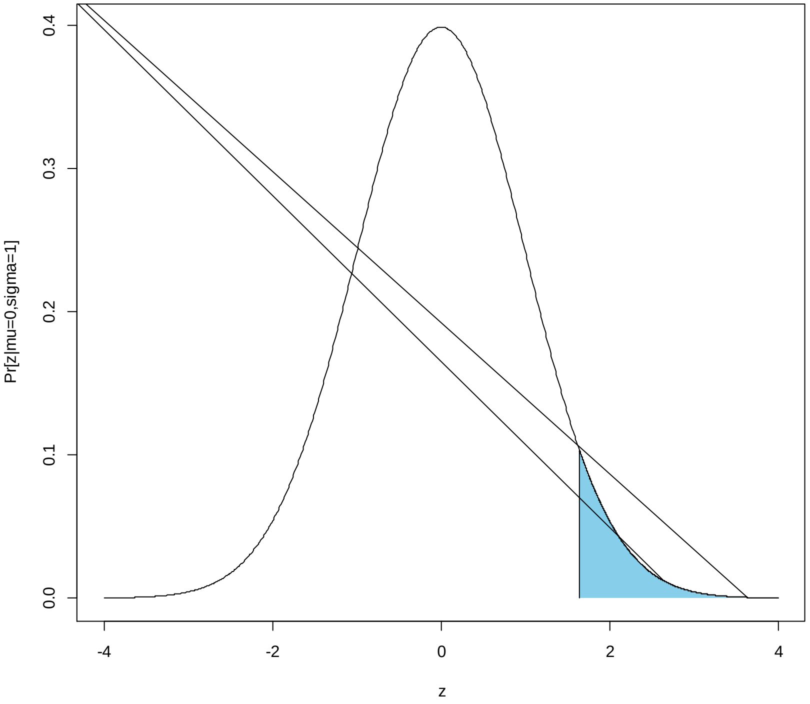
- Different approaches
  - Using screening assay as the example
- Dealing with outliers

Setting cut points

The cut point

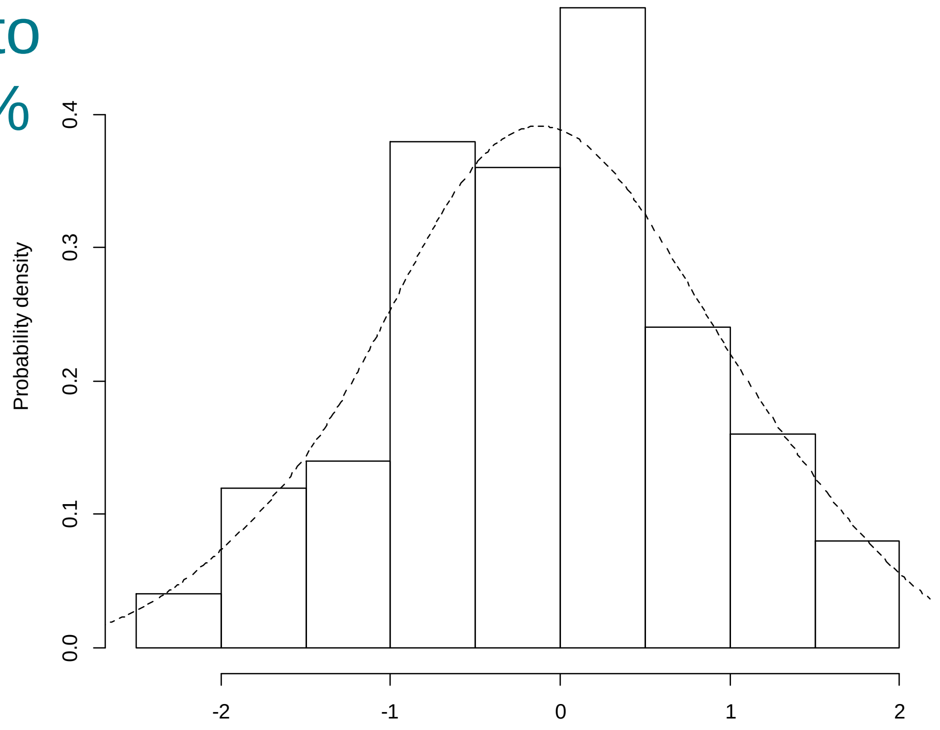
- Intended to provide a cutoff value above or below which a sample is considered positive
- Corresponds to a particular percentile of the *underlying background population*

- As such, we need an appropriate estimate available from our method validation data



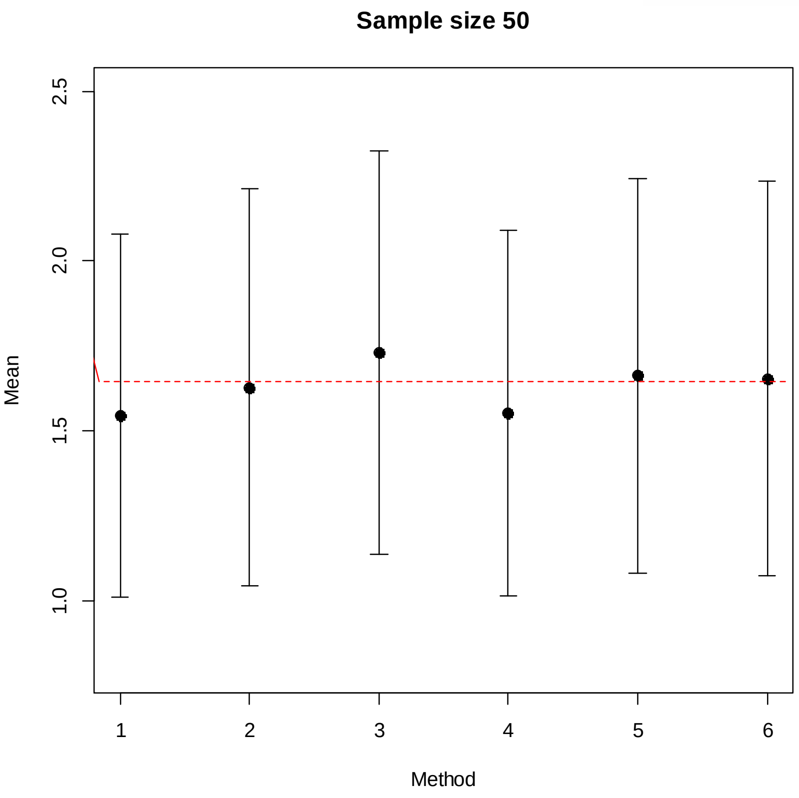
Knowing the underlying distribution is key

- We are essentially trying to estimate a percentile (95% or 0.1%) from a data set
  - A lot of data needed, or
  - Knowledge/prior information about the distribution



Direct estimation of quantiles

- Needs a lot of data for good results
- Sensitive to the calculation method used
- Difficult for a 5% tail, even more so for a 1% or 0.1% tail
- I would say use as a last resort



The distribution is often not normal

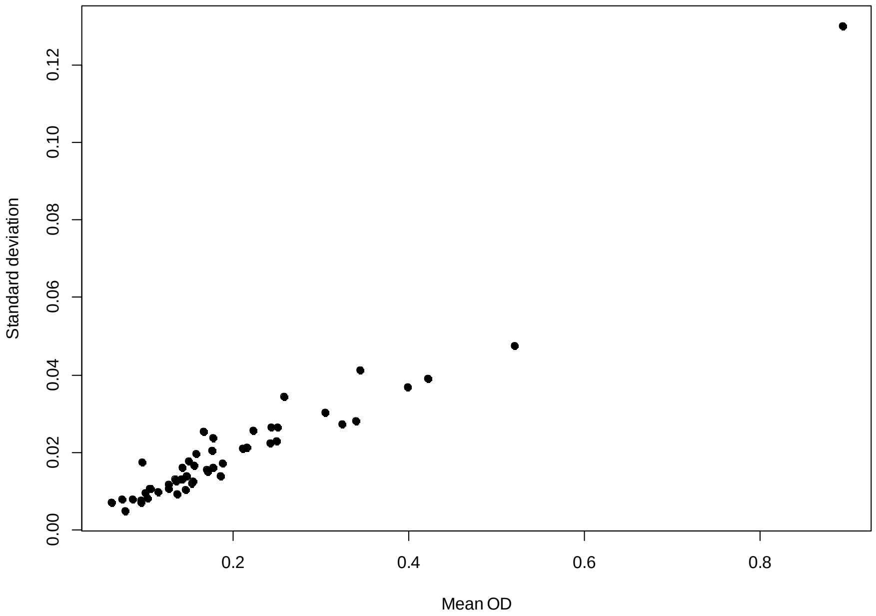
- Continuous, lower limit of zero
- Skewed
- Outlier-contaminated

- It is usually possible [in our experience] to transform the data (e.g. log) to produce an approximately normal distribution

- Important, because a lot of the calculations assume normality

Distribution of response data

- Response is usually lognormal (which is sometimes very close to normal)
  - relative SD observed to be constant
  - log transformation produces a normal distribution



- For studies which are ‘similar’, it is probably better to proceed on the basis of prior knowledge of the way analytical data are distributed, rather than on a case-by-case basis

- Departures from lognormality or similar are likely caused more by outliers and other anomalies
  - Purely based on experience, but not a rule as such

Method #1 – Assume normal distribution

$$Y = \bar{x} + Z_{0.95} s$$

$$Y = \bar{x} + Z_{0.001} s$$

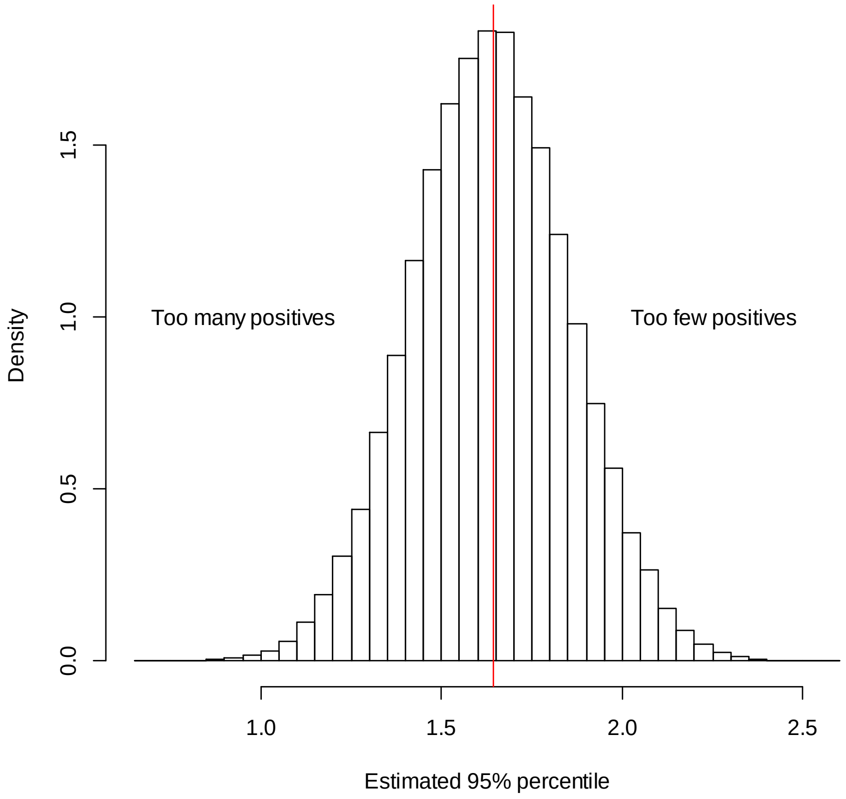
- Slightly biased since *s* is not an unbiased estimate of σ
- Can correct for bias with a multiplying factor
  - Near 1 for large degrees of freedom, e.g. 1.0051 for *n* = 50

- This is an unbiased estimate of the 95th (or 0.1th) percentile, given the data
  - It is our best estimate of the true value

Method #1 – implications

- Approximately half of the time, the estimate will correspond to below 5% of the population

- There is a significant risk that our validation study will produce a cut point which covers too small a fraction of the population



Method #2 – prediction limit

- Provides a range for the next expected value: given the data, at which value is there a 5% chance that the next observed result will exceed it?
- Differs from method #1 in that it describes the distribution of a future individual data point rather than the underlying population parameters

- Does not seem appropriate, since cut points will tend to be set too high

Method #3 – Shen *et al*

- M Shen *et al*, *J Biopharm Stat*, 25 (2015) 269

Journal of Biopharmaceutical Statistics, 25: 269–279, 2015
ISSN: 1054-3406 print/1520-5711 online
DOI: 10.1080/10543406.2014.979196

Taylor & Francis
Taylor & Francis Group

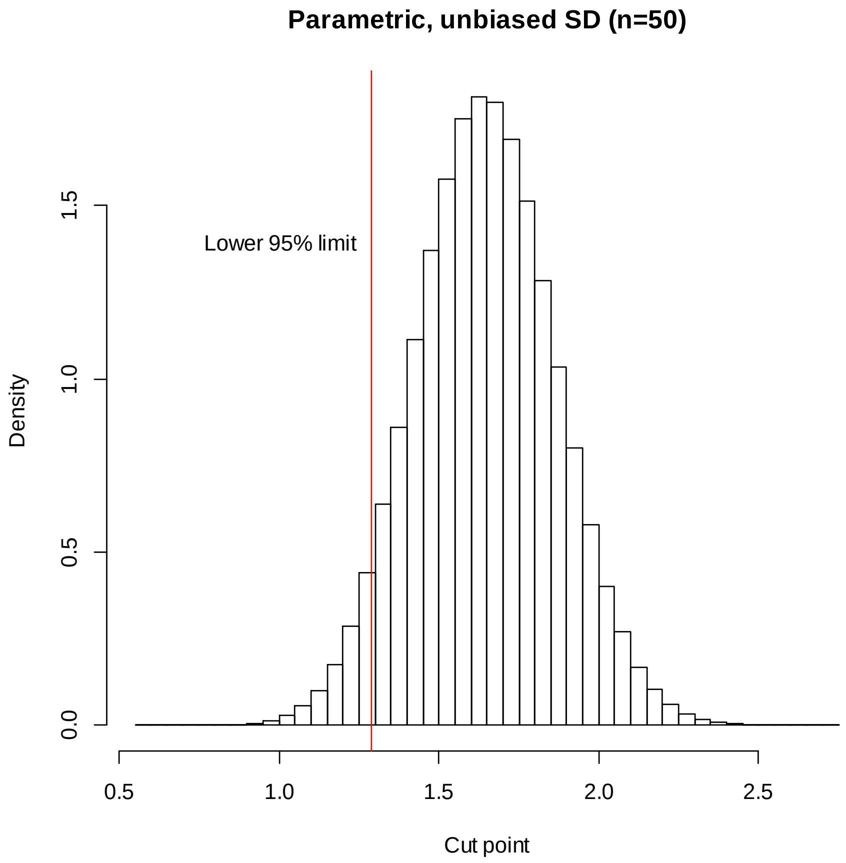
STATISTICAL EVALUATION OF SEVERAL METHODS FOR CUT-POINT DETERMINATION OF IMMUNOGENICITY SCREENING ASSAY

Meiyu Shen, Xiaoyu Dong, and Yi Tsong
Office of Biostatistics/Office of Translational Sciences, Center for Drug Evaluation and Research, U.S. Food and Drug Administration, Silver Spring, Maryland, USA

The cut point of the immunogenicity screening assay is the level of response of the immunogenicity screening assay at or above which a sample is defined to be positive and below which it is defined to be negative. The Food and Drug Administration Guidance for

- Contained in latest draft of FDA guidance

Method #3 - implications



- FP rate depends strongly on sample size
- Complex/mathematical if "exact *t*" is used

Method #4 – tolerance limits

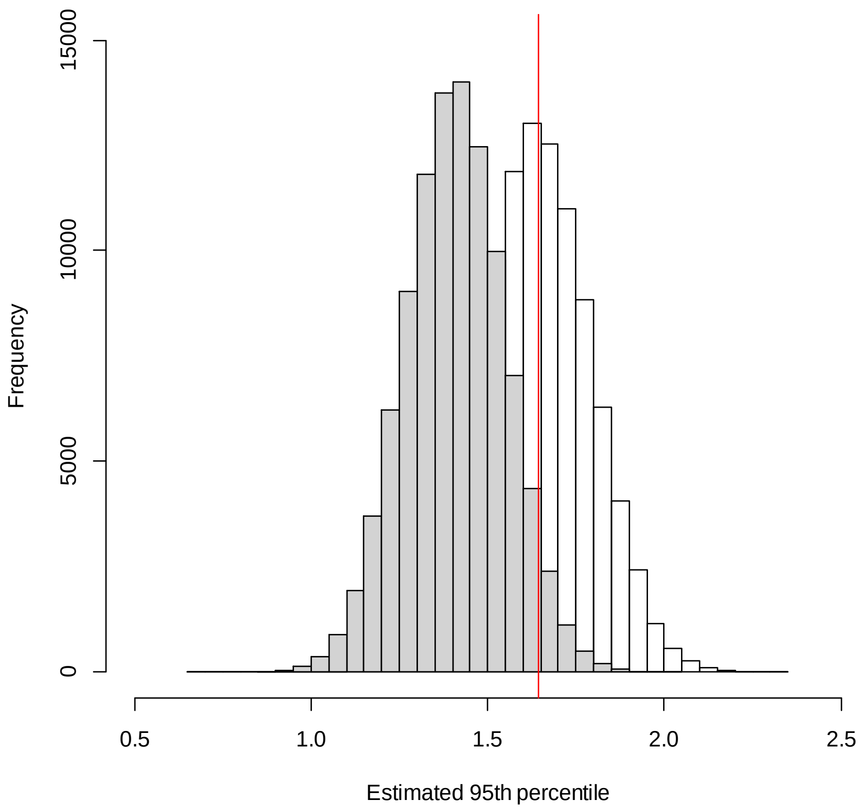
- Think of a tolerance interval as a confidence interval on a proportion of the population

- The tolerance interval covers (say) 95% of the population with 95% confidence
  - Over many repeated experiments, at least 95% of the population is covered 95% of the time

- Can also express in terms of the tail, e.g. "at least 5% of the population lies above the limit, with 95% confidence"

Method #4 – expected positives

- We may want our validation study to deliver a cut point which has a low probability of covering less than 5% of the population
- A tolerance interval will do this at the cost of a biased estimate
- Expect to get about 9% with *n* = 50



- Tolerance intervals can be a bit tricky to calculate, but approximations available

- See http://www.itl.nist.gov/div898/handbook/prc/section2/prc26.htm

Method #5 – Robust estimation

$$Y = \mathrm{med}(x) + Z_{0.95}\mathrm{MAD_E}$$

- Median is an unbiased estimator of μ
- $\mathrm{MAD_E}$ (=1.483 × MAD) is a consistent estimator of s *if the underlying distribution is normal*
  - As $n \to \infty$, $\mathrm{MAD_E} \to s$

- $\mathrm{MAD_E}$ is biased unless n is large
- Expect more positives than using s and Z

- But it is not sensitive to outliers

Comparison of methods by simulation

	***n* = 10**		***n* = 50**	
Method	*Expected FPR*	*Pr [FPR < 5%]*	*Expected FPR*	*Pr [FPR < 5%]*
Parametric	5.5%	0.46	5.1%	0.48
Parametric with unbiased SD	5.0%	0.49	5.0%	0.50
Shen *et al* (approx. *Z*)	20%	0.017	9.9%	0.033
Tolerance limit	15%	0.052	9.3%	0.050
Robust estimation	6.7%	0.40	5.3%	0.46

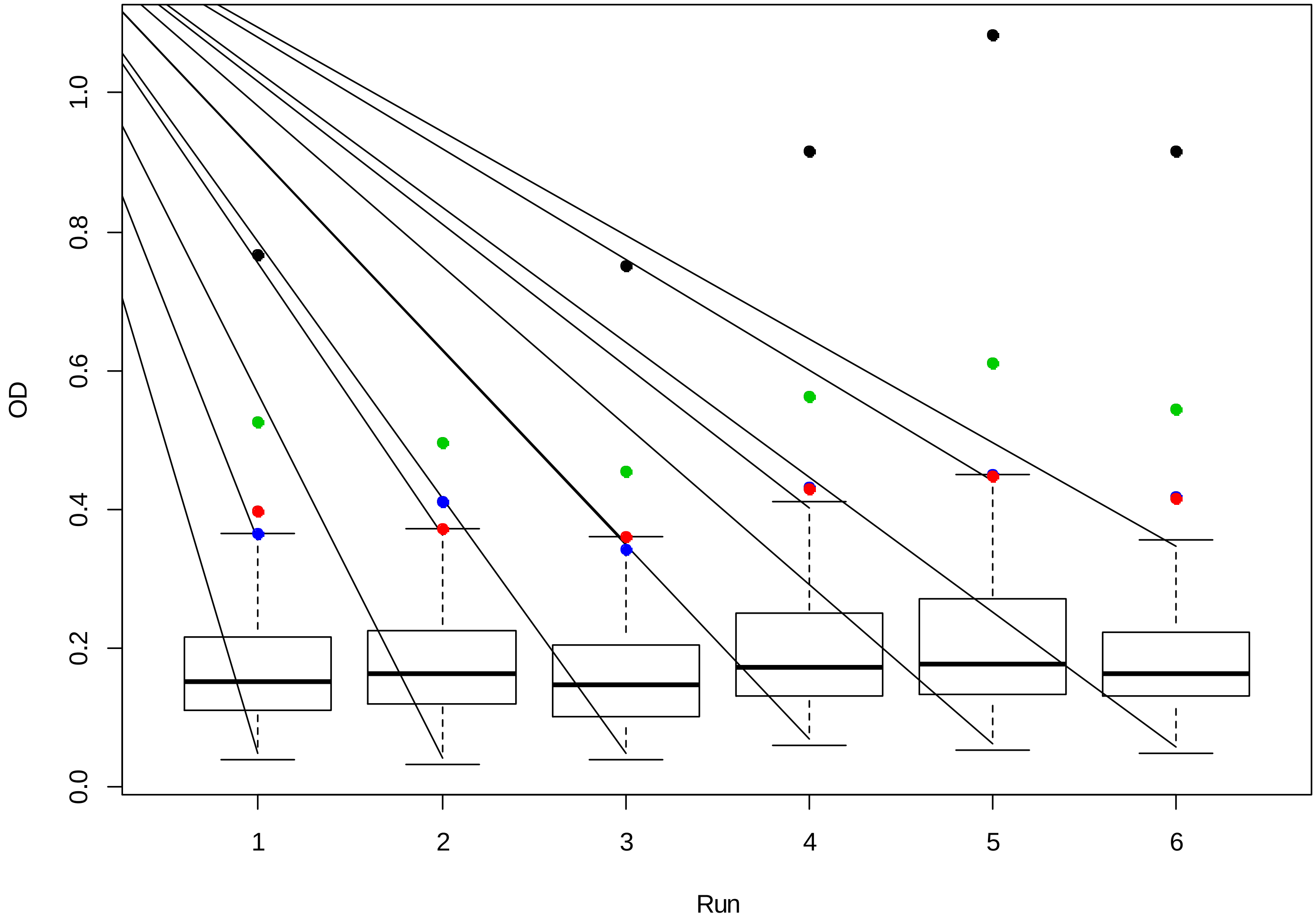
Dealing with outliers

Testing for outliers

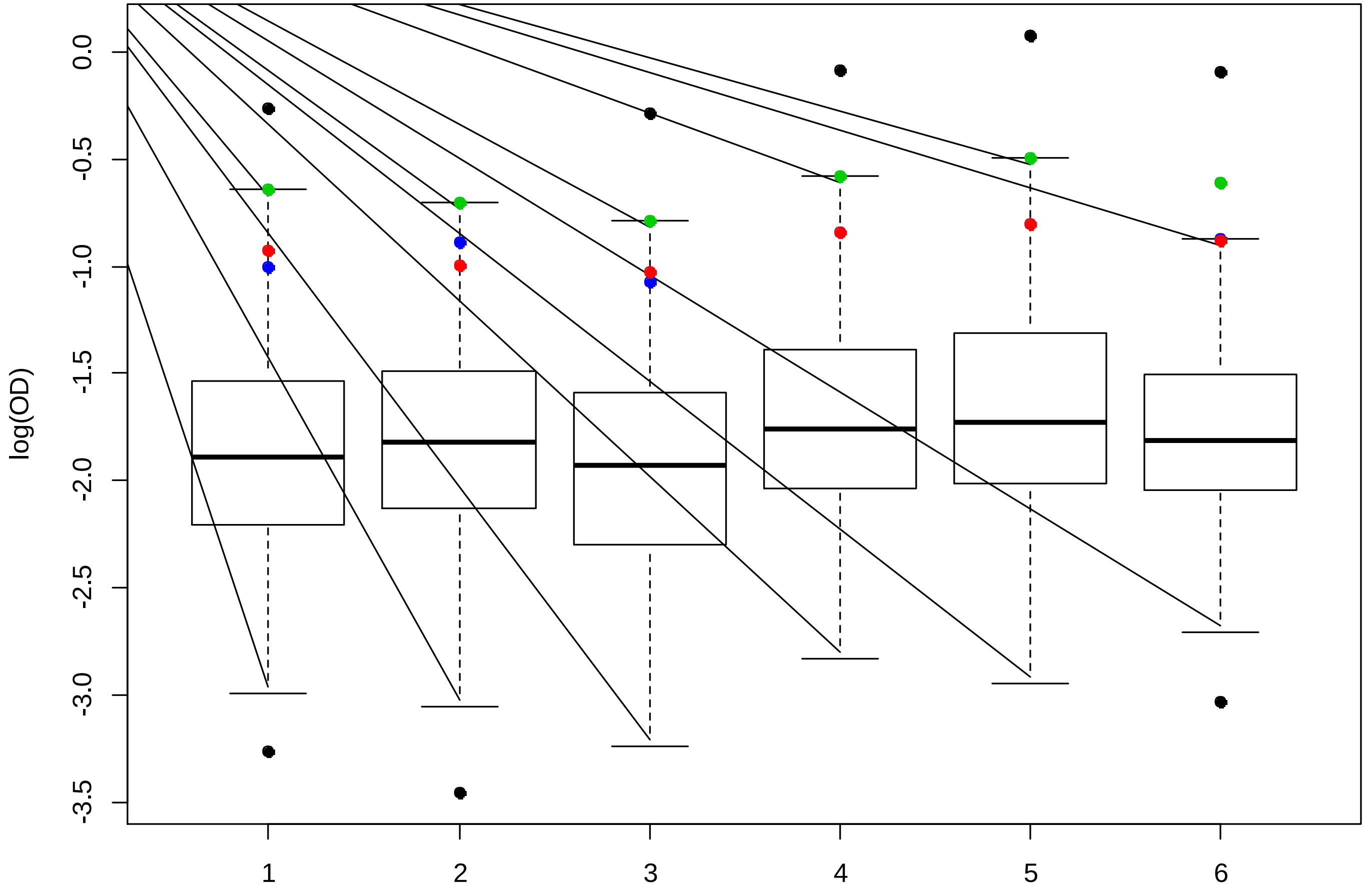
- Generally, we should aim to avoid biasing the data set unduly by excessive removal
- Problem is to distinguish data points which should not be present from those which are just part of the underlying population

- Analytical outliers
- Biological outliers – part of the underlying population, or a separate population?

Example #1

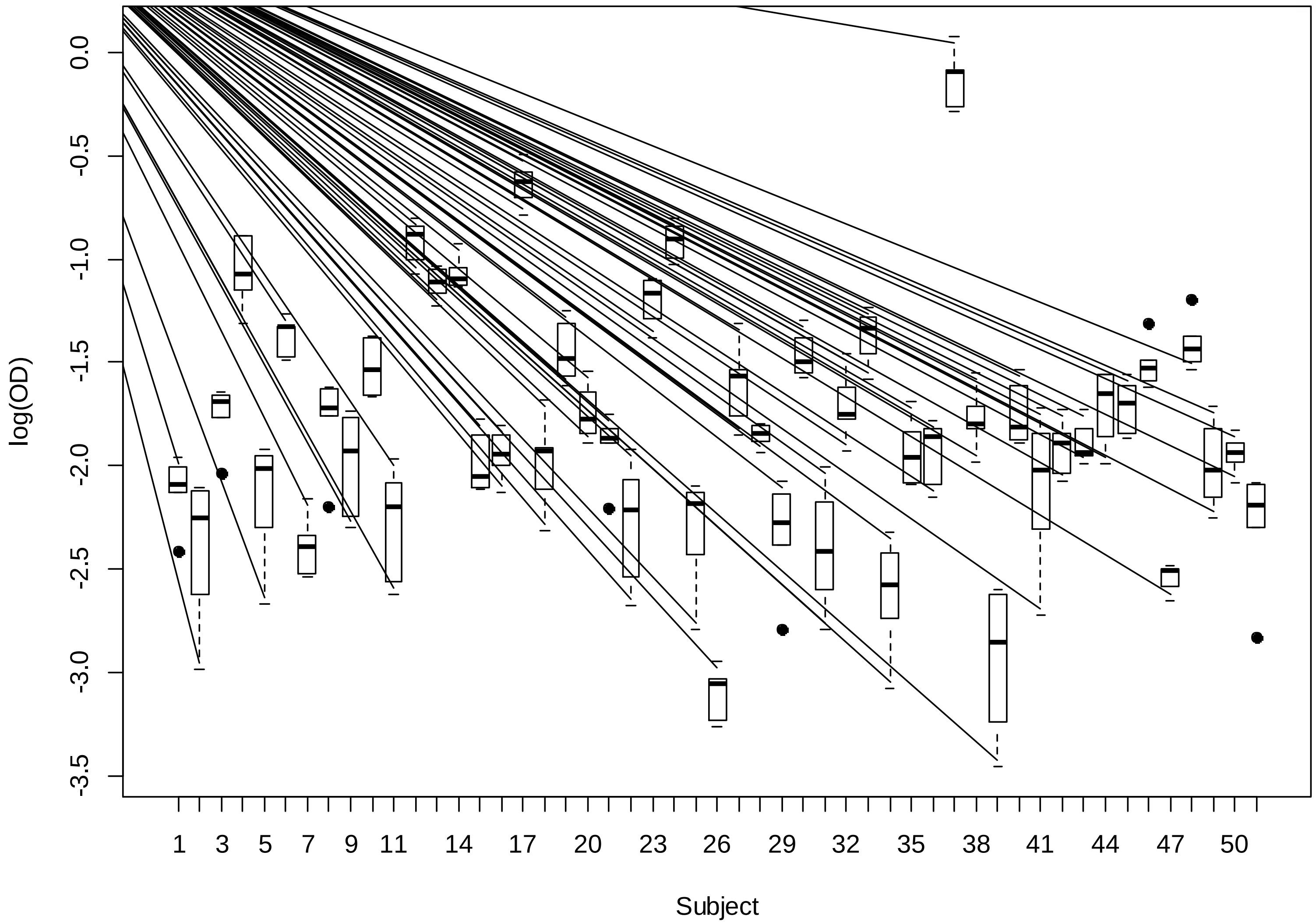


Transform to log scale



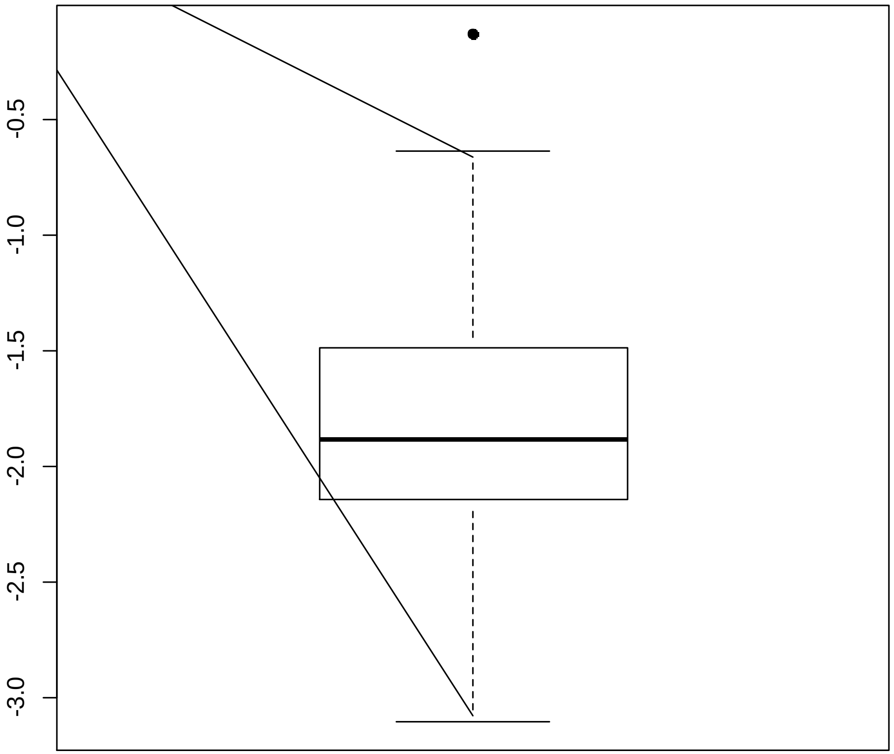
- Most seem to be just part of the distribution

Shown by subject



Testing visually – box plots

- Equivalent to removal of data at $p = 0.007$ (or $p = 0.0035$ if one-sided)
- Not a test of extreme values
- Potentially truncates the data set regardless of distribution
- This data set tests as normal ($p = 0.49$, SW)

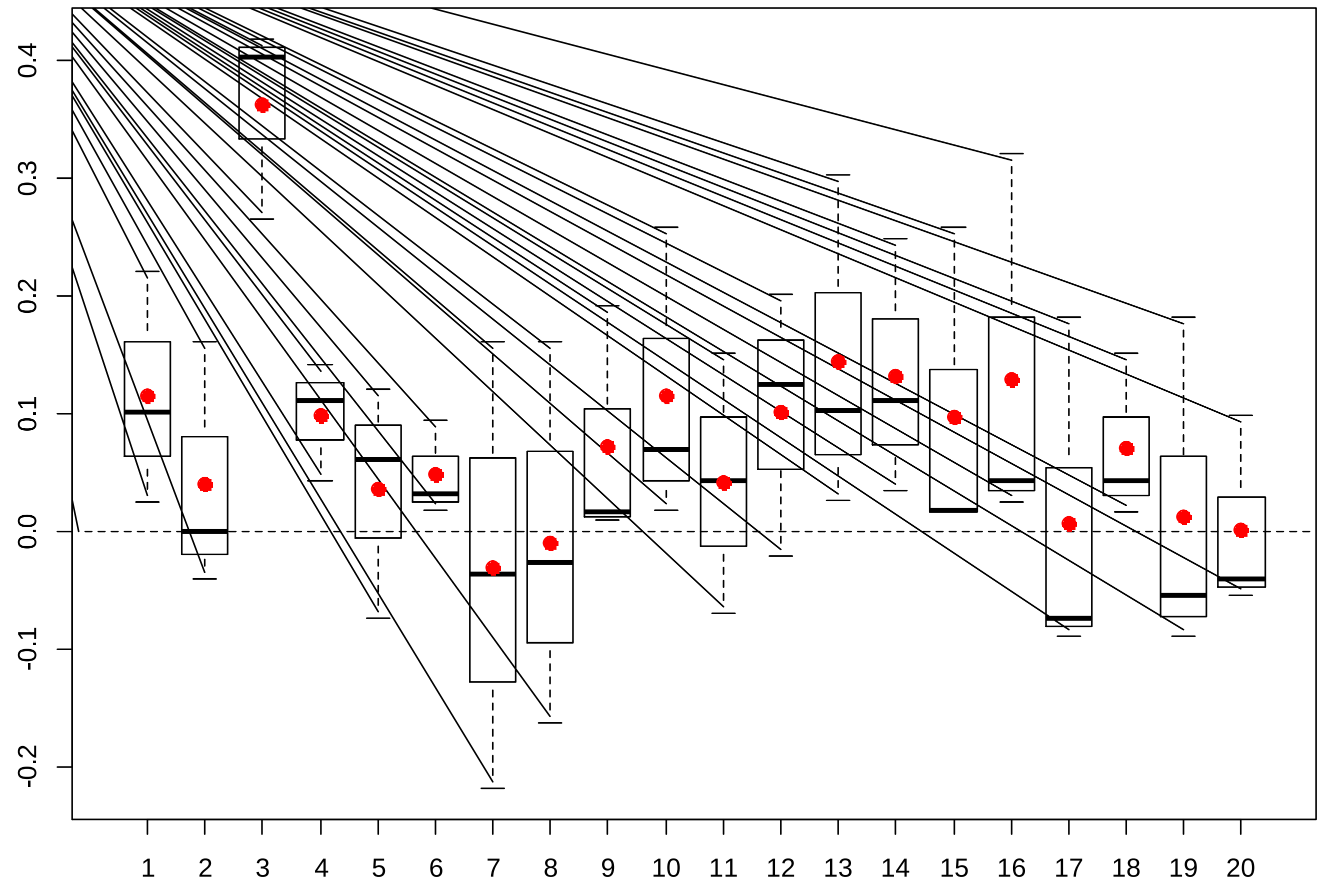


Statistical testing – Grubbs' test

- Tests extreme observations against the sample
  - F Grubbs, *Ann Math Stat*, 21 (1950) 27
- Assumes an underlying normal distribution
- Sequential removal of one or two data points

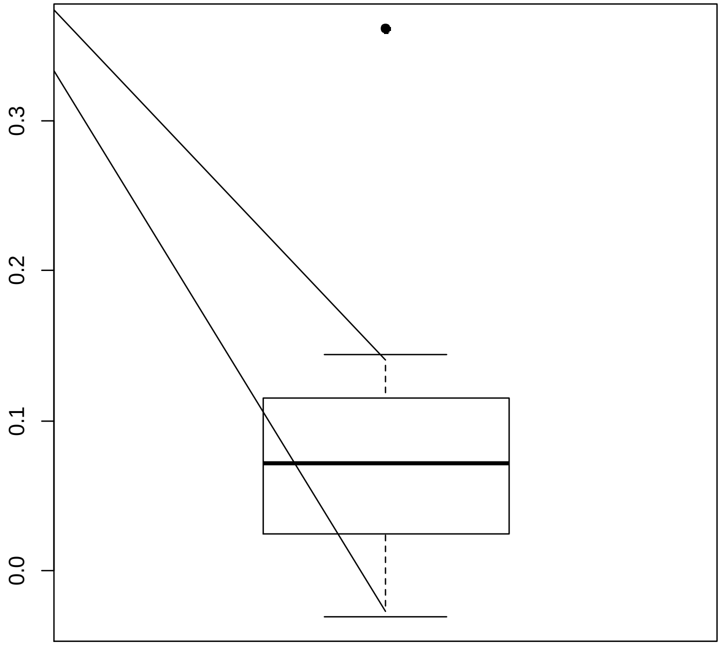
- This example doesn't test as a strong outlier (p = 0.045)
  - At 95% retain unless reason to exclude
  - At 99% exclude unless reason to retain

Example #2



Data from Hoffman & Berger, *J Immunol Methods*, 373 (2011) 200

Outlier tests



- Grubbs' test produces p = 0.00034
- Clearly not part of an underlying distribution

To conclude

- You have to have good information about the underlying distribution, for cut point setting and outlier assessment

- The method you choose depends very much on what you want in terms of risk

Acknowledgements

- Colleagues at LGC’s Fordham laboratory
- David Egging for help and advice