Z-curve analysis is a statistical method used to estimate the true discovery rate (TDR) of a set of studies, which reflects the proportion of significant results that are true positives. The method is useful for meta-analysis, especially when dealing with studies that may have low statistical power or may suffer from publication bias.

The assumptions of z-curve analysis include that the studies being analyzed are independent, and that the true effect sizes are normally distributed. The method assumes that the distribution of observed p-values is a mixture of two normal distributions, with one reflecting true positive results and the other reflecting null results.

The output of z-curve analysis includes several estimates, including the TDR and the expected replication rate (ERR). The TDR is the proportion of significant results that are true positives, while the ERR is the proportion of studies with true effects that are significant. The method also provides a confidence interval for the TDR and the ERR.

Additionally, z-curve analysis produces a plot that shows the distribution of true effect sizes, as estimated from the observed p-values. The plot also shows the estimated density of the null distribution and the cut-off for statistical significance. The plot can be used to visualize the distribution of significant results and to evaluate the potential impact of publication bias on the results.

In summary, z-curve analysis is a useful tool for meta-analysis that can help to estimate the TDR and the ERR of a set of studies. The method assumes that the studies being analyzed are independent, and that the distribution of true effect sizes is normal. The output includes estimates of the TDR and the ERR, as well as a plot that visualizes the distribution of true effect sizes and the potential impact of publication bias.

Glossary

**Power** is defined as the long-run relative frequency of statistically significant results in a series of exact replication studies with the same sample size when the null-hypothesis is false.

**Unconditional power** is the long-run frequency of statistically significant results in studies where the null-hypothesis is true, regardless of effect size.

**Mean (unconditional) power** is the arithmetic mean of the power of individual studies.

**Discovery rate** is the relative frequency of statistically significant results, without distinguishing between true or false discoveries.

**Selection bias** is the process that favors the publication of statistically significant results, resulting in a higher percentage of statistically significant results in the published literature than in the population of all studies that were conducted.

**Observed discovery rate (ODR)** is the percentage of statistically significant results in an observed set of studies.

**Expected discovery rate (EDR)** is the mean power before selection for significance.

**Expected replication rate (ERR)** is the mean power after selection for significance, based on the power to obtain a statistically significant result in the same direction.

Example

Zcurve analysis of 20 p-values.

Mean effect size estimate: 0.50

Estimated discovery rate (EDR):

95% lower bound: 0.10

Maximum: 0.52

95% upper bound: 0.79

Number of true null hypotheses: 5 (25%)

Number of false null hypotheses: 15 (75%)

Area under the curve (AUC): 0.86

The output of the Z-curve analysis contains several pieces of information:

The mean effect size estimate: This is an estimate of the average effect size across all of the studies included in the analysis. In this example, the mean effect size estimate is 0.50.

The expected discovery rate (EDR): This is the estimated proportion of significant results that are true positives (i.e., the proportion of findings that represent true effects). The EDR is reported as a range of values: the 95% lower bound, the maximum, and the 95% upper bound. In this example, the EDR ranges from 0.10 to 0.79.

The number of true null hypotheses and false null hypotheses: These are the estimated numbers of true and false null hypotheses, respectively. A null hypothesis is considered true if there is no real effect, and false if there is a real effect. In this example, the analysis suggests that there are 5 true null hypotheses and 15 false null hypotheses.

The area under the curve (AUC): This is a measure of the overall accuracy of the Z-curve analysis. The AUC ranges from 0 to 1, with higher values indicating greater accuracy. In this example, the AUC is 0.86, indicating a reasonably accurate analysis.

Interpreting the results of the Z-curve analysis involves considering the estimated discovery rate (EDR) and the number of true and false null hypotheses. The EDR provides an estimate of the proportion of significant results that are true positives, while the numbers of true and false null hypotheses provide estimates of the overall proportion of effects that are real.

In this example, the EDR ranges from 0.10 to 0.79, indicating some uncertainty about the proportion of true effects among the significant results. The estimated number of true null hypotheses is 5, which suggests that about 25% of the findings in the analysis represent true effects. The estimated number of false null hypotheses is 15, indicating that the majority of the findings are likely false positives.

Overall, these results suggest that while there may be some real effects in the data, there is also a high likelihood of false positives. This highlights the importance of interpreting the results of any single study with caution, and of conducting replications and meta-analyses to gain a more accurate understanding of the true effects.

The main take-home implications of a z-curve analysis output are:

1. Estimating the true discovery rate: One of the primary purposes of z-curve analysis is to estimate the true discovery rate, which is the percentage of statistically significant results that are true positives. This is important because it provides an estimate of the reliability of the results in a field. If the true discovery rate is low, it suggests that a significant proportion of the statistically significant results are likely false positives.
2. Identifying the distribution of true effect sizes: The shape of the z-curve can provide information about the distribution of true effect sizes. If the z-curve is steep, it suggests that there are many studies with large effect sizes, and fewer studies with small effect sizes. If the z-curve is relatively flat, it suggests that the studies in the field are relatively evenly distributed across effect sizes.
3. Evaluating the presence of publication bias: The shape of the z-curve can also provide information about the presence of publication bias. If the z-curve has a sharp drop-off, it suggests that there may be a significant amount of publication bias in the field, with many small, non-significant studies going unpublished.
4. Assessing the replicability of results: The mean expected replication rate (ERR) can be used to estimate the likelihood that a study's results can be replicated. If the ERR is high, it suggests that the results are likely to be replicable in future studies. If the ERR is low, it suggests that the results are less likely to be replicable, and may be due to random chance or other factors.

Plans moving forward

1. Github (platform)
2. För att göra en Z-curve analysis: vad behöver vi?
   1. Preregga: Inklusions & exklusionskriterier (specificera)
   2. Hur söka the Schimmackian way
   3. X antal studier (X = hur många? titta på tidigare z-curve analysis)
   4. Litteratursök med syfte att få fram:
      1. Sample size
      2. Test statistics (e.g., t-values, F-values)
      3. Degrees of freedom (df)
      4. Significance level
   5. Sammanställ i datafile:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Jn.Number | Article.Number | Year | Journal | DOI | Studies | z | type.stat | df1 | df2 | Test.Stat |

1. Förstå och tolka output. + förstå och tolka figur

Raver to do: doc med tolkning av output, excelfil

Elli to do: github

Tolka output

Call:

zcurve(z = data$z, bootstrap = 0)

model: EM via EM

Estimate

ERR 0.730

EDR 0.194

Model converged in 79 + 1000 iterations

Fitted using 158 z-values. 301 supplied, 204 significant (ODR = 0.68, 95% CI [0.62, 0.73]).

Q = -192.00

Checklist review article according to Siddaway et al.,