# Once more about the z-curve method

Erik van Zwet

20 januari, 2026

## Contents

This report accompanies a blog post Once more about the z-curve method.

# 1 Packages

```
suppressPackageStartupMessages({
  library(ggplot2)
  library(zcurve)
})
```

# 2 Introduction

Suppose we have an unbiased, normally distributed estimator $y$ with standard error $s$ of some (unknown) parameter of interest $\theta$. We assume that $s$ is known. The difference between knowing $s$ and having to estimate it, is the difference between a $t$-test and a $z$-test. That's a minor difference when there aren't any serious outliers and the sample size is not too small. So, we assume that $y$ has the normal distribution with mean $\theta$ and standard deviation $s$. The 95% confidence interval for $\theta$ is $y \pm 1.96 \cdot s$. All very standard.

The $z$-statistic is the ratio of the estimator to its standard error, so $z = y/s$. It follows that $z$ has the normal distribution with mean $\theta/s$ and standard deviation 1. We refer to $\theta/s$, i.e. the ratio of the true parameter to the standard error of its estimator, as the signal-to-noise ratio or SNR. If we define $\text{SNR} = \theta/s$. Then the $z$-statistic has the normal distribution with mean SNR and standard deviation 1.

If we have a collection of studies, then the distribution of the $z$-statistics is the convolution (sum) of the distribution of the SNRs of the studies and the standard normal distribution. If we've estimated the distribution of the $z$-statistics, we can get the distribution of the SNRs by *deconvolution*. Deconvolution is known to be very unstable. That means that we would need very many data points (studies) or very strong assumptions -– preferably both — to get an accurate result.

The $z$-curve method (Brunner and Schimmack (2020), Bartos and Schimmack (2022), Schimmack and Bartos (2023)) is based on the assumption that the absolute values of the SNRs have a discrete distribution supported on $0, 1, 2, \ldots, 6$. We denote the associated probabilities by $p_0, p_1, \ldots, p_6$. Note that SNR=0 corresponds to a "null effect" so $p_0$ is the proportion of null effects among the studies. The distribution of the absolute $z$-statistics is a mixture of the folded versions of normal distributions with means $0, 1, 2, \ldots, 6$ and unit variances.

To circumvent the effects of selection on statistical significance, $z$-curve uses only the absolute values of the z-statistics which exceed 1.96 in magnitude to estimate the $p_i$. The authors also truncate absolute $z$-statistics at 6. They then re-parameterize in terms of a mixture of folded normals truncated to $[1.96, 6]$. The new mixture weights $w_0, w_1, \ldots, w_6$ are

$$w_i = \frac{p_i \, P(1.96 < |z| < 6 \mid \text{SNR} = i)}{\sum_{j=0}^{6} p_j P(1.96 < |z| < 6 \mid \text{SNR} = j)}.$$

Clearly, there is a 1-1 relation between the $p_i$ and $w_i$ since

$$p_i = \frac{w_i / P(1.96 < |z| < 6 \mid \text{SNR} = i)}{\sum_{j=0}^{6} w_j / P(1.96 < |z| < 6 \mid \text{SNR} = j)}.$$

Therefore maximum likelihood estimation of the $w_i$ is equivalent to maximum likelihood estimation of the $p_i$.

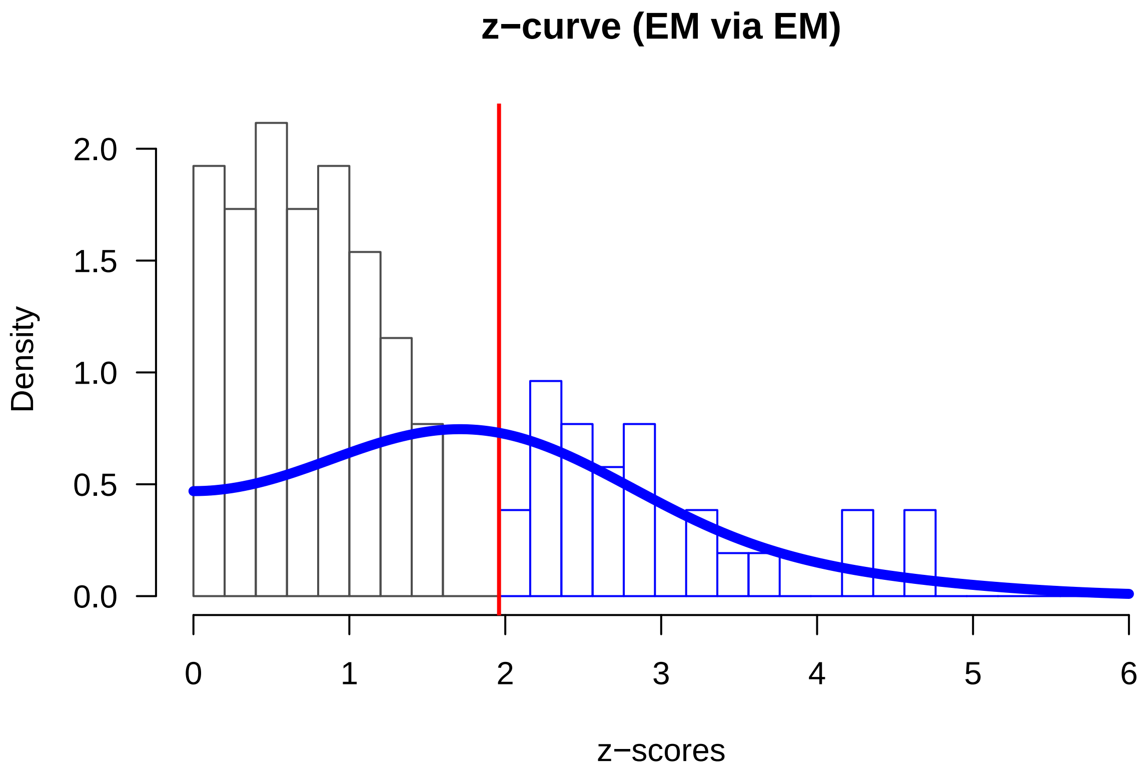
## 3 An example

The function `rmix()` samples from normal mixture distributions with weights `p`, means `m` and standard deviations `s`.

```
rmix = function(n,p,m,s){
  d=rmultinom(n,1,p)
  rnorm(n,m%*%d,s%*%d)
}
```

Suppose we sample from the normal mixture distribution specified below.

```
set.seed(111)
n=100
p=c(0.5,0.2,0.2,0.1,0,0,0)
m=0:6
s=rep(1,7)
z=rmix(n,p,m,s)
fit=zcurve(z)
plot(fit)
```



# 4 Coverage of the mixture weights

The model is parameterized in terms of the mixture weights $w_i$ of the truncated folded normals. These weights are

```
# prob = P(1.96 < |z| < 6 | SNR=m)
prob=pnorm(6, m, 1) - pnorm(1.96, m, 1) + pnorm(-1.96, m, 1) - pnorm(-6, m, 1)
w=p*prob; w=w/sum(w)
round(w,3)
```

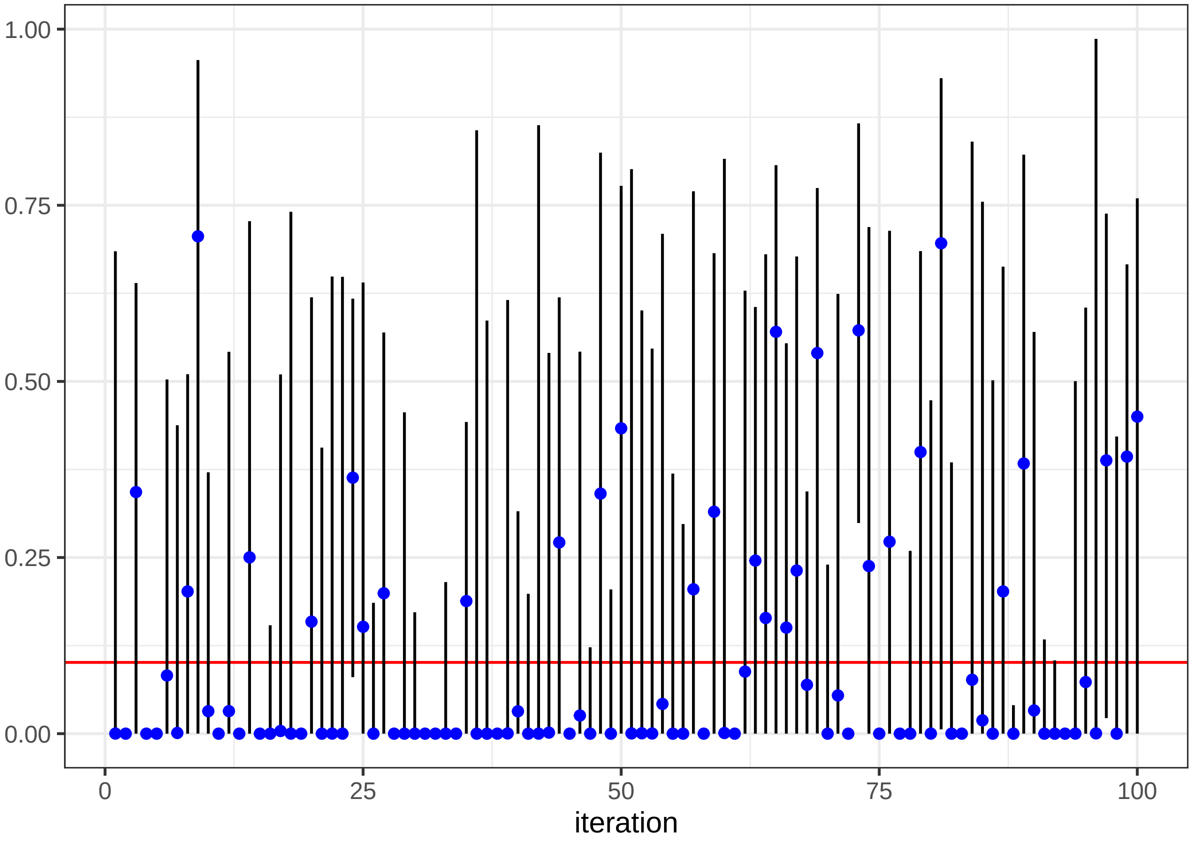
```
## [1] 0.101 0.138 0.418 0.344 0.000 0.000 0.000
```

The function `zcurve` provides a 95% confidence interval for the mixture weight of the zero-mean component which in this example is equal to 0.101. We check the coverage of these intervals across 100 simulations.

```
k=100
ci_w1=ci_EDR=ci_ERR=ci_FDR=NULL
for (i in 1:k){
  z=rmix(n,p,m,s)
  fit=zcurve(z)
  ci_w1=rbind(ci_w1,summary(fit,type="parameters")$coefficients[1,]) # CI for w[1]
  ci_ERR=rbind(ci_ERR,summary(fit, all=TRUE)$coefficients[1,])       # CI for ERR
  ci_EDR=rbind(ci_EDR,summary(fit, all=TRUE)$coefficients[2,])       # CI for EDR
}
ci=data.frame(iteration=1:k,estimate=ci_w1[,2],lower=ci_w1[,3],upper=ci_w1[,4])
collapse=(ci$upper - ci$lower < 0.0001)
mean((ci$lower < w[1]+0.0001) & (ci$upper > w[1]) )                  # coverage of w[1]
```

```
## [1] 0.78
```

We plot the confidence intervals. The horizontal red line is the true value, and the blue dots are the estimates.


1.00
0.75
0.50
0.25
0.00
0
25
50
75
100
iteration

# 5 Coverage of the EDR

The EDR is the “population mean true power”. It “equals the overall probability of a significant result” (Brunner and Schimmack, 2020).

$$\text{EDR} = P(|z| > 1.96) = \sum_{i=0}^{6} p_i \, P(|z| > 1.96 \mid \text{SNR} = i).$$

```
pmix = function(x,p,m,s){     # cdf of normal mixture (vector x)
  drop(p %*% sapply(x, function(x) pnorm(x,mean=m,sd=s)))
}
EDR=pmix(-1.96,p,m,s) + 1 - pmix(1.96,p,m,s)
EDR
```

```
## [1] 0.2472923
```

We find the EDR is 0.247. The coverage of the bootstrap intervals across 100 iterations is

```
ci=data.frame(iteration=1:k,estimate=ci_EDR[,1],lower=ci_EDR[,2],upper=ci_EDR[,3])
ci$collapse=collapse
mean((ci$lower < EDR-0.05) & (ci$upper > EDR+0.05) )     # coverage of EDR
```

```
## [1] 0.87
```

The coverage of the “robustified” bootstrap intervals across 100 iterations is

```
mean((ci$lower < EDR) & (ci$upper > EDR) )               # coverage of EDR
```

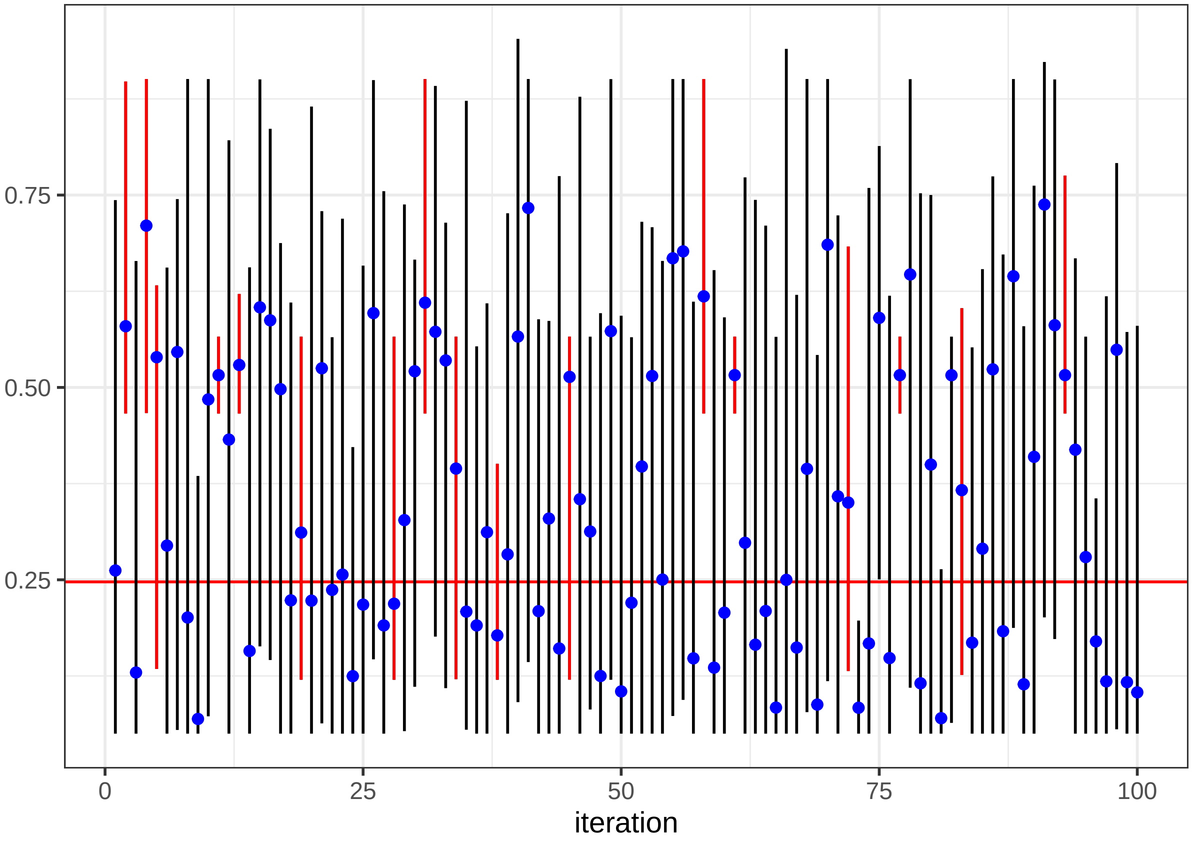
```
## [1] 0.89
```

The estimate of the EDR is biased

```
t.test(ci$estimate,mu=EDR)
```

```
##
##  One Sample t-test
##
## data:  ci$estimate
## t = 5.4999, df = 99, p-value = 2.975e-07
## alternative hypothesis: true mean is not equal to 0.2472923
## 95 percent confidence interval:
##  0.3149967 0.3914196
## sample estimates:
## mean of x
## 0.3532081
```

We plot the robust bootstrap intervals. The horizontal red line is the true EDR and the blue dots are the maximum likelihood estimates (MLEs). The red confidence intervals are the cases where the confidence interval of p0 collapsed to [0,0].


0.75
0.50
0.25
0
25
50
75
100
iteration