# Problem of calibrating benchmark distributions for z > 5

Authors: Sebastian Kranz and Peter Pütz   BCH write on p. 3650:

> We assume that the observed test statistic distribution above z = 5 should be free of p-hacking or publication bias—the incentives to p-hack in a range so far above the traditional significance thresholds are plausibly zero. We then produce a non-central t-distribution for each method that closely fits the observed distribution in the range z > 5 by calibrating the degrees of freedom and non-centrality parameter.

We just now realized (after submitting our revision) that the argumentation that the observed distribution for z above 5 is free of p-hacking and publication bias is problematic. Even if publication bias only occurs for small absolute z, it will also affect the observed density for large z. That is because the total area under the density will be normalized to 1. So if publication bias takes away density for small z, this will automatically add density for the large z not afflicted by publication bias.

Let us illustrate this with a simple simulation in R. We assume that the latent distribution of z-statistics is a t-distribution. At the same time we assume that tests with $abs(z) \leq 1.9$ only have a 50% publication probability.

Densities
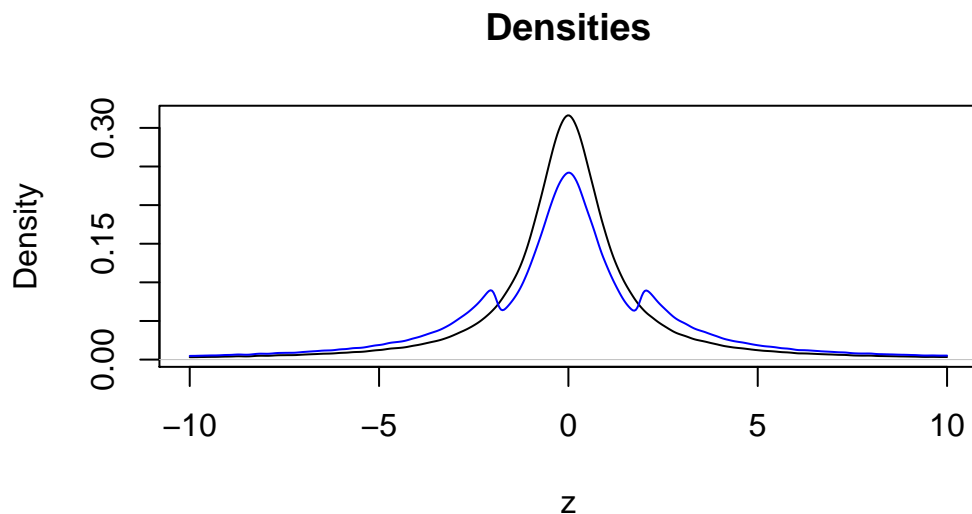
The black line is the original density of z without publication bias and the blue line the observed density after publication bias.

Let us now plot the ratio's of the observed to the latent density:

Density Ratios: Observed / Latent
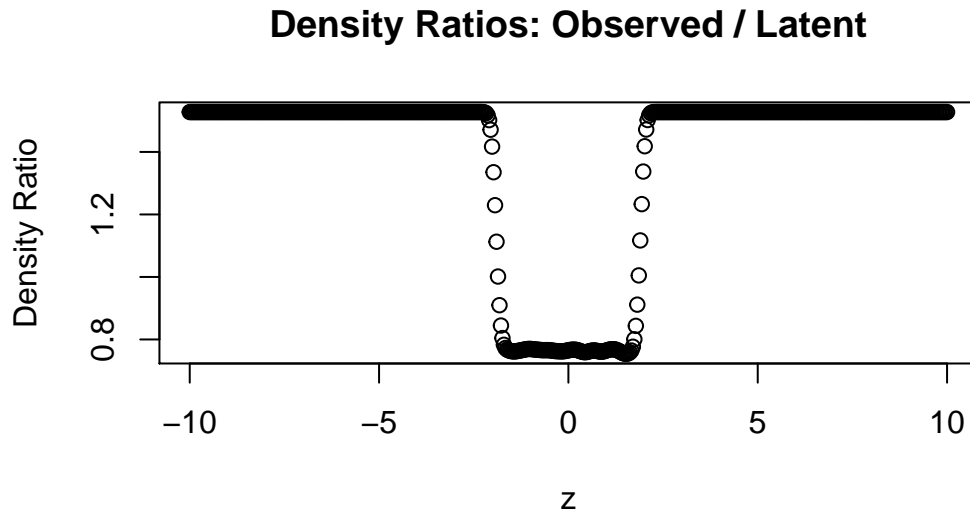
We see that the density ratio is not only different from 1 for low absolute z-statistics where publication bias takes place, but also in the range where no publication bias takes place. In particular also for $abs(z) > 5$ both densities don't match, the density of observed z-statistics is roughly 50% higher than the latent original density.

Thus, calibrating a t-distribution on the observed data that matches the tails with $abs(z) > 5$ would not recover in this example the latent distribution of z-statistics.
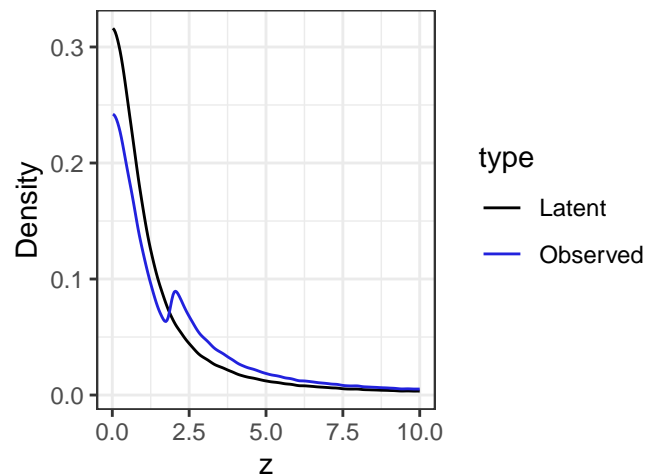
Nicer plot with ggplot

These plots are shown in the working paper. They only show z >= 0.

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union

## Warning: Removed 512 row(s) containing missing values (geom_path).
```
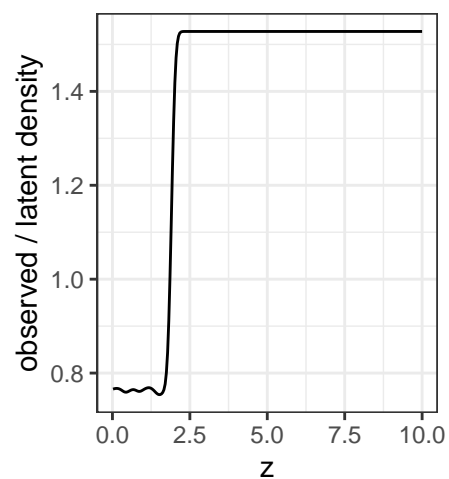
Ratio

```
## Warning: Removed 256 row(s) containing missing values (geom_path).
```



```
## <ggproto object: Class ScaleDiscrete, Scale, gg>
##     aesthetics: colour
##     axis_order: function
##     break_info: function
##     break_positions: function
##     breaks: waiver
##     call: call
##     clone: function
##     dimension: function
##     drop: TRUE
##     expand: waiver
##     get_breaks: function
##     get_breaks_minor: function
##     get_labels: function
##     get_limits: function
##     guide: legend
##     is_discrete: function
```

```
##     is_empty: function
##     labels: waiver
##     limits: NULL
##     make_sec_title: function
##     make_title: function
##     map: function
##     map_df: function
##     n.breaks.cache: NULL
##     na.translate: TRUE
##     na.value: grey50
##     name: waiver
##     palette: function
##     palette.cache: NULL
##     position: left
##     range: <ggproto object: Class RangeDiscrete, Range, gg>
##         range: NULL
##         reset: function
##         train: function
##         super:  <ggproto object: Class RangeDiscrete, Range, gg>
##     rescale: function
##     reset: function
##     scale_name: manual
##     train: function
##     train_df: function
##     transform: function
##     transform_df: function
##     super:  <ggproto object: Class ScaleDiscrete, Scale, gg>
```