

# Validation of a transformed pdf in R

Marek Rychlik

## Purpose

In this file there is R Code in the form of an “R Notebook” (a .Rmd file), which illustrates how data transformation affects the density. We use an example of squaring uniformly distributed data. We plot both the theoretical density and the histogram of the original and transformed samples.

One purpose this file could serve is to verify that the theoretically derived density of  $Y = g(X)$  based on the formula

$$f_Y(y) = \sum_{x \in g^{-1}(y)} \frac{f_X(x)}{|g'(x)|}$$

is correctly derived. If it is correct, there should be no significant discrepancy between the histogram and the theoretical density.

With some R skills (sampling and writing functions), you could reuse this Notebook to validate your own derivation.

## The steps

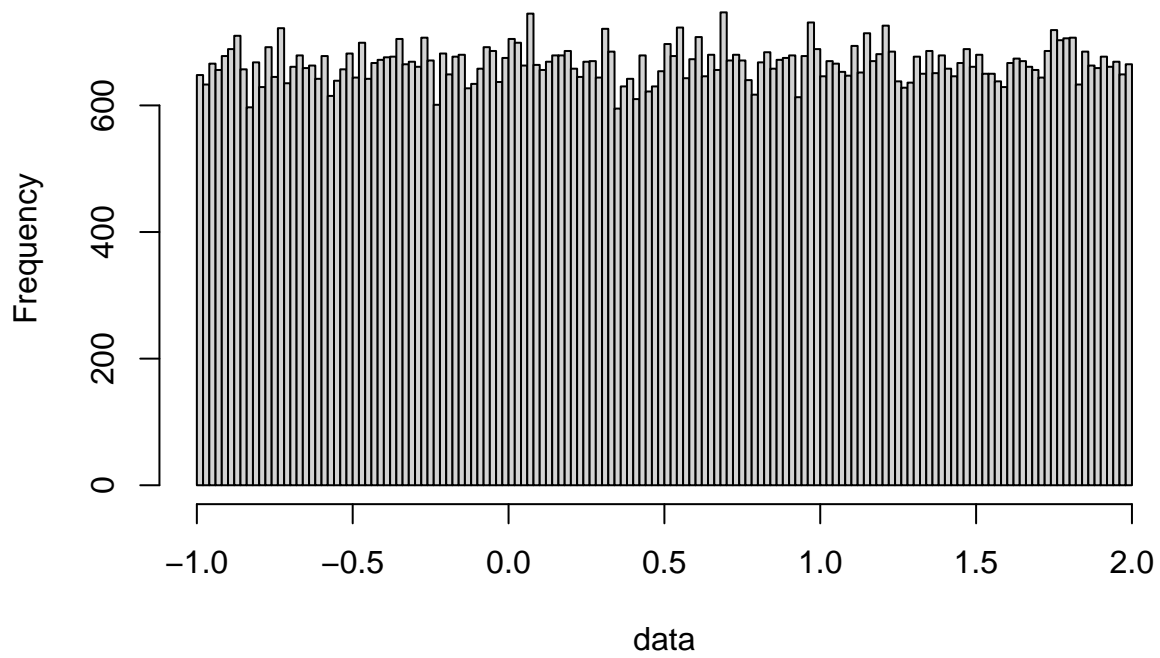
Below we conduct a numerical experiment with the following steps:

- generate a uniformly distributed sample on the interval  $[-1, 2]$ ; call the generated sample ‘data’;
- Transform the data using the mapping  $y = g(x^2)$ ; called the transformed sample ‘xformed’;
- Plot the histograms of the original data ‘data’ and transformed data ‘xformed’

## Implementation

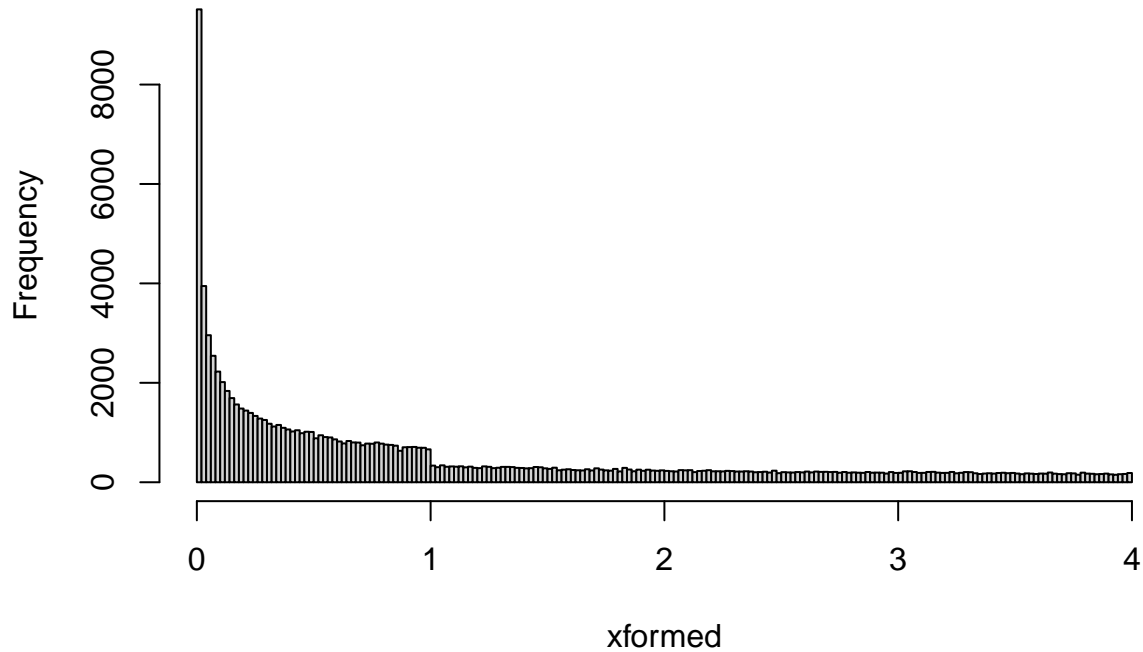
```
n=100000
nbreaks=200
data=runif(n,-1,2)
hist(data,breaks=nbreaks,main="Original Data")
```

## Original Data



```
g<-function(x)x^2  
xformed=g(data)  
hist(xformed,breaks=nbreaks,main="Transformed Data")
```

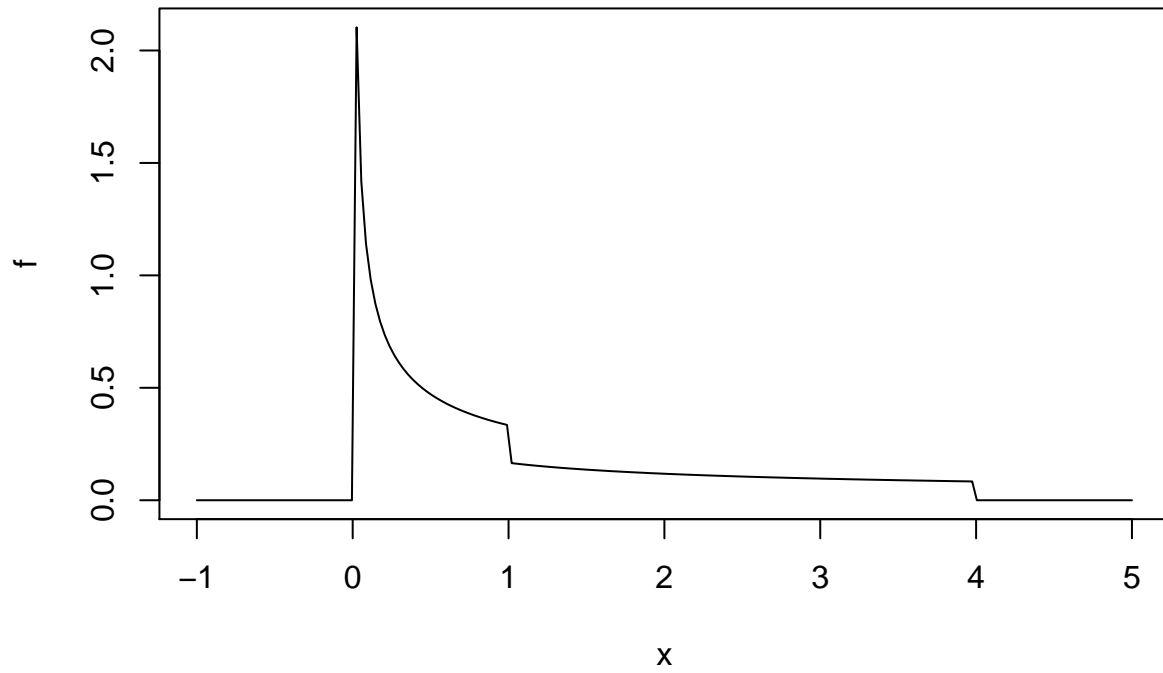
## Transformed Data



Plot transformed data over the distribution

```
f <- Vectorize(function(y) {  
  if(y<=0) 0  
  else if(y<1) 1/3/sqrt(y)  
  else if(y<=4) 1/6/sqrt(y)  
  else 0;  
})  
  
# Standalone plot of the density  
plot(f,-1, 5,n=200,main="Transformed pdf")
```

## Transformed pdf



```
p1=plot(f,-1, 5,n=200,col=rgb(1,0,0,1),main="Transformed pdf + histogram")  
hist(xformed,freq=FALSE,add=TRUE,col=rgb(1,0,0,1/4))
```

### Transformed pdf + histogram

