Prostate Cancer Data in R

Jeremy Williams

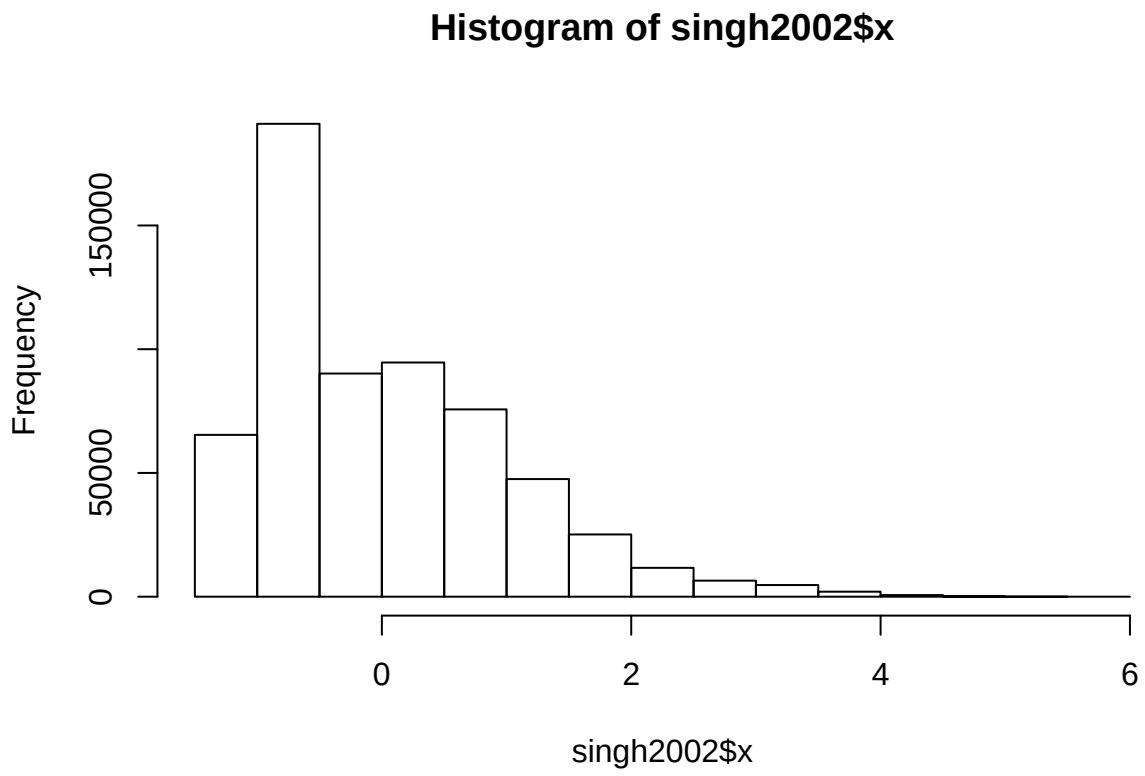
22 de febrero de 2018

```
# prostate data
#install.packages("sda")
# load sda library
suppressMessages(suppressWarnings(library("sda")))

# load Singh et al (2001) data set
data(singh2002)
dim(singh2002$x) # 102 6033
```

```
## [1]  102 6033
```

```
hist(singh2002$x)
```

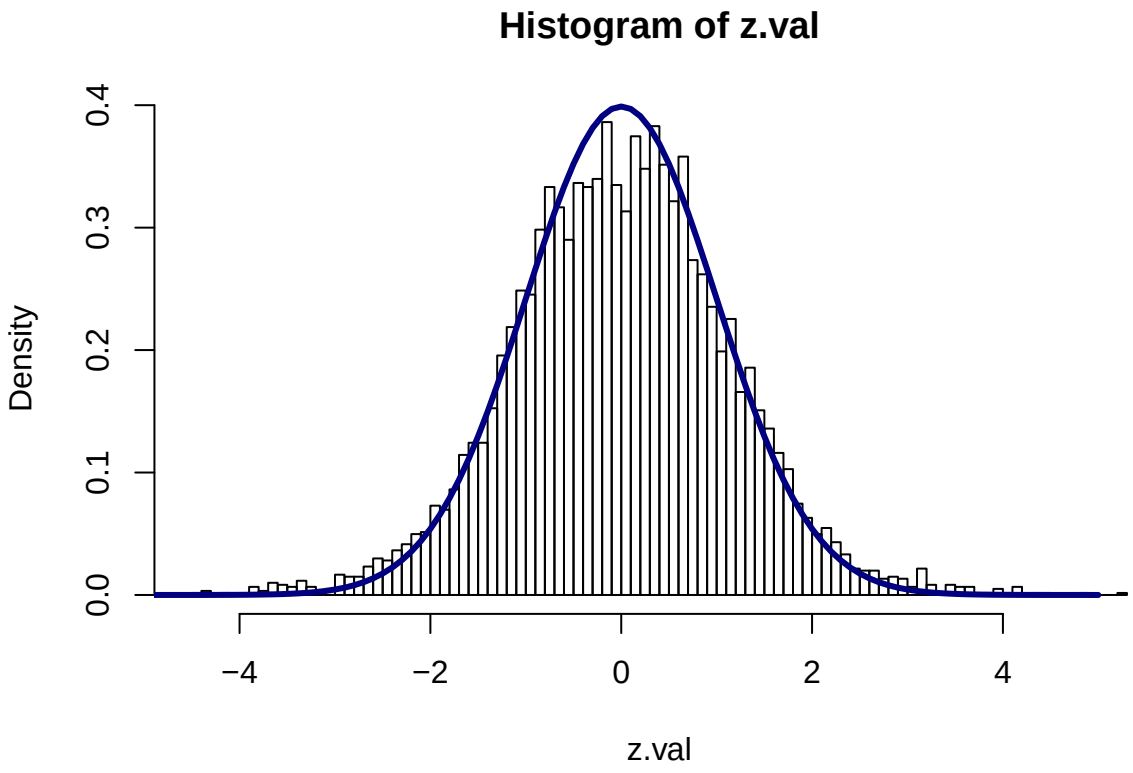



```
singh2002$y # 2 levels (healty/cancer)
```

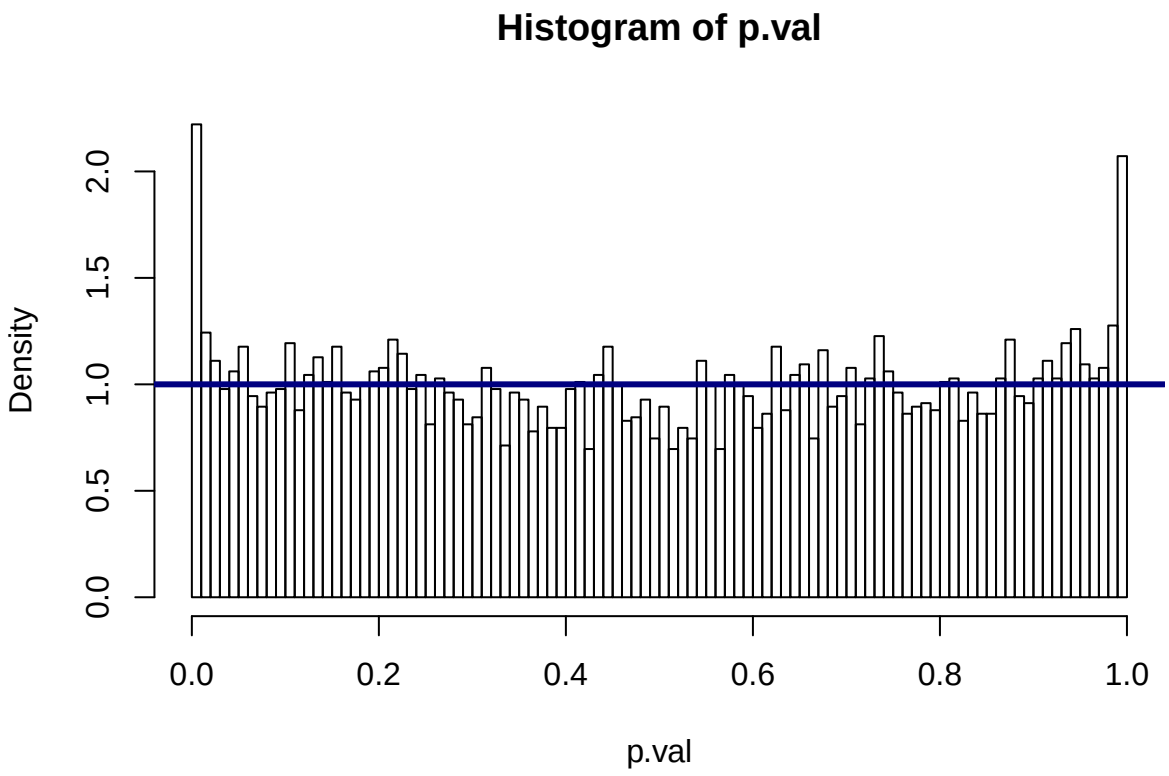
```
##   [1] healthy healthy healthy healthy healthy healthy healthy healthy
##   [9] healthy healthy healthy healthy healthy healthy healthy healthy
##  [17] healthy healthy healthy healthy healthy healthy healthy healthy
##  [25] healthy healthy healthy healthy healthy healthy healthy healthy
##  [33] healthy healthy healthy healthy healthy healthy healthy healthy
```

```
##  [41] healthy healthy healthy healthy healthy healthy healthy healthy
##  [49] healthy healthy cancer  cancer  cancer  cancer  cancer  cancer
##  [57] cancer  cancer  cancer  cancer  cancer  cancer  cancer  cancer
##  [65] cancer  cancer  cancer  cancer  cancer  cancer  cancer  cancer
##  [73] cancer  cancer  cancer  cancer  cancer  cancer  cancer  cancer
##  [81] cancer  cancer  cancer  cancer  cancer  cancer  cancer  cancer
##  [89] cancer  cancer  cancer  cancer  cancer  cancer  cancer  cancer
##  [97] cancer  cancer  cancer  cancer  cancer  cancer
## Levels: cancer healthy
```

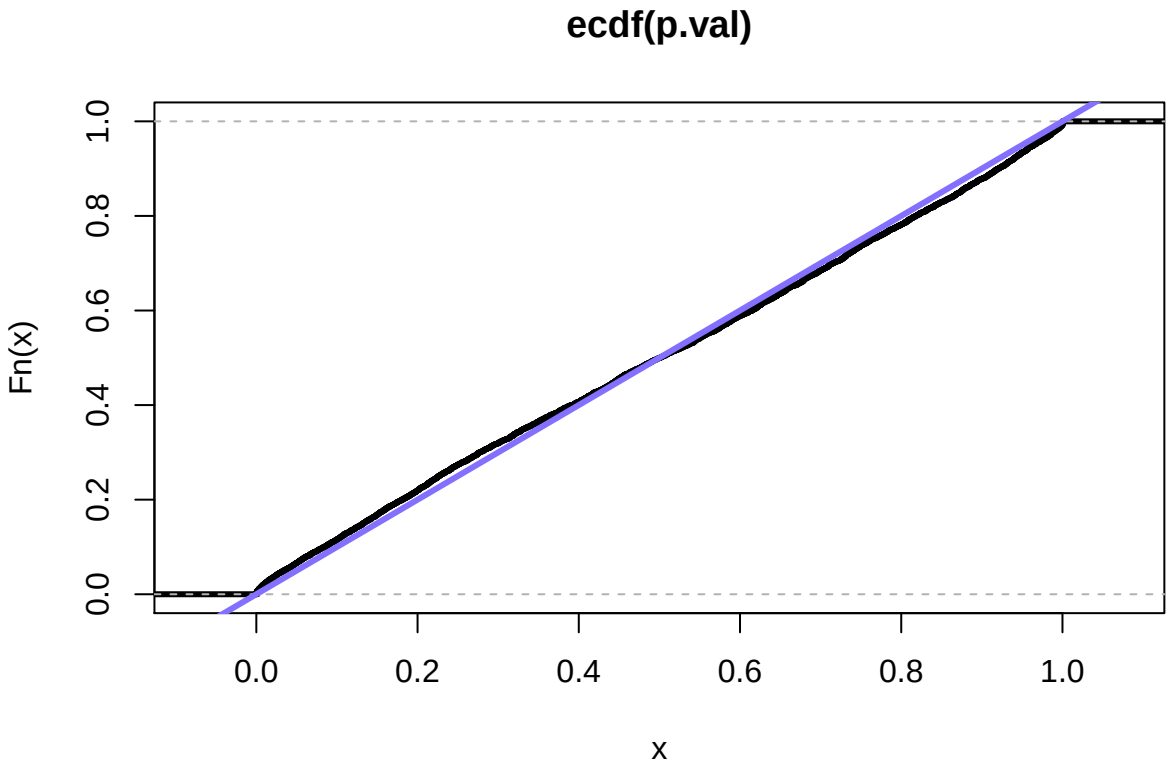
```
n<-dim(singh2002$x)[1]
N<-dim(singh2002$x)[2]
t.val <- apply(singh2002$x,2,function(x){t.test(x=x[51:102],y=x[1:50],alternative="less",var.equal=T)$s
z.val <- qnorm(pt(t.val,n-2))
hist(z.val,br=100,freq=FALSE)
u <- seq(-5,5,by=.1)
lines(u,dnorm(u),col="navy",lwd=3)
```




```
p.val <- apply(singh2002$x,2,function(x){t.test(x=x[51:102],y=x[1:50],alternative="less",var.equal=T)$p
hist(p.val,br=100,freq=FALSE)
abline(h=1,col="navy",lwd=3)
```




```
# ecdf: Empirical cumulative distribution function
plot(ecdf(p.val),lwd=3)
abline(a=0,b=1,col="lightslateblue",lwd=3)
```




```
ks.test(p.val,"punif")      #  rechazamos que sean uniformes...
```

```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  p.val
## D = 0.024958, p-value = 0.001088
## alternative hypothesis: two-sided
```

```
#####

###################
# Bonferroni

p.Bonf <- min(p.val)*N
p.Bonf
```

```
## [1] 0.02834847
```

```
alpha=0.05
tau=alpha/N
sum(p.val<=tau) ## how many p-vals are le tau?
```

```
## [1] 3
```

```
sum(p.adjust(sort(p.val), method="holm",N)<0.05)
```

```
## [1] 3
```

```r
sum(p.adjust(sort(p.val), method="bonferroni",N)<0.05)
```

```
## [1] 3
```