Chi-squared test

Peter Nandori

3/14/2022

Chi-squared test in R

Problem

The response time of a computer system to a request for a certain type of information is hypothesized to have an exponential distribution with parameter 1 sec (so if X = response time, the pdf of X under H_0 is e^{-x} for $x > 0$)

(a)

If you had observed and wanted to use the chi-squared test with five class intervals having equal probability under H_0, what would be the resulting class intervals? Hint: use *qexp(1/5)*, etc.

(b)

Carry out the chi-squared test using the following data resulting from a random sample of 40 response times:

```
vec<-c(.10,.99,1.14,1.26, 3.24,.12,.26,.80,.79,1.16,
1.76,.41,.59,.27,2.22,.66,.71, 2.21,.68, .43,
.11,.46,.69,.38,.91, .55,.81,2.51,2.77,.16,
1.11,.02,.13,.19,1.21,1.13,2.93,2.14,.34,.44)
```

Hint: Construct a vector **count** so that **count**[**i**] contains the number of sample elements that belong to the ith class interval found in part (a) for $i = 1, ..., 5$. Then perform the chi-squared test using **chisq**.**test**(**count**). As always, you can type **help**(**chisq**.**test**) in RStudio to learn more about the chi-squared test in R.

Solution

The following list contains the endpoints of the intervals

```
quantiles <- c(qexp(0),qexp(1/5),qexp(2/5),qexp(3/5),qexp(4/5), qexp(1))
```

One possible solution is

```
bins = 5
count = rep(0,5)
for(i in 1: length(vec)){
  count[ceiling(bins*pexp(vec[i]))] = count[ceiling(bins*pexp(vec[i]))] +1
}
chisq.test(count)
```

```
##
##  Chi-squared test for given probabilities
##
## data:  count
## X-squared = 0.75, df = 4, p-value = 0.945
```

Let X be an $m \times n$ data matrix such that $X^T X$ is invertible, and let $M = I_m - X\left(X^T X\right)^{-1} X^T$. Add a column $\mathbf{x}_0$ to the data and form $W = \left[\begin{array}{ll} X & \mathbf{x}_0 \end{array}\right]$ Compute $W^T W$. The $(1,1)$ -entry is $X^T X$. Show that the Schur complement (Exercise 15) of $X^T X$ can be written in the form $\mathbf{x}_0^T M \mathbf{x}_0$. It can be shown that the quantity $\left(\mathbf{x}_0^T M \mathbf{x}_0\right)^{-1}$ is the $(2,2)$ -entry in $\left(W^T W\right)^{-1}$. This entry has a useful statistical interpretation, under appropriate hypotheses.