Teste Qui-Quadrado de Pearson

SME0810-Métodos não paramétricos

18-09-2015

- Giovani Carrara Rodrigues 7151669
- Vitor Bonini 8065859

Exercício

O recenseamento de 320 famílias com 5 filhos revelam os dados abaixo:

Número de Filhos	5M 0H	4M 1H	3M 2H	2M 3H	1M 4H	0M 5H	Total
Número de Famílias	18	56	110	88	40	8	320

Onde:

M: Mulher
H: Homem

Teste a hipótese de que o nascimento de homens e mulheres é igualmente provável. Caso H_0 seja rejeitada, estimar o parâmetro da distribuição e refazer o teste.

Resolução

Criando o dataset

```
# criando o dataset
clas1=c(5,0,18);clas2=c(4,1,56);clas3=c(3,2,110)
clas4=c(2,3,88);clas5=c(1,4,40);clas6=c(0,5,8)

data = data.frame(clas1,clas2,clas3,clas4,clas5,clas6,c(0,0,sum(clas1[3],clas2[3],clas3[3],clas4[3],clas
names(data)=c("clas1","clas2","clas3","clas4","clas5","clas6","total")
row.names(data)=c("m","h","nfami")
attach(data)
```

```
## The following objects are masked _by_ .GlobalEnv:
##
##     clas1, clas2, clas3, clas4, clas5, clas6
```

```
data
```

```
##       clas1 clas2 clas3 clas4 clas5 clas6 total
## m         5     4     3     2     1     0     0
## h         0     1     2     3     4     5     0
## nfami    18    56   110    88    40     8   320
```

Seja X: o número de homens em 5 filhos, então $X \sim Bin(5, 0.5)$, pois queremos testar se a probabilidade de nascer um homem ou uma mulher é igual. Sendo assim temos as seguintes hipóteses

$H_0 : p = 0.5$
$H_1 : p \neq 0.5$

Onde as frequências observadas e experadas são dadas respectivamente por

```
#frequencias observadas
fob=data[3,1:6]
fob
```

```
##       clas1 clas2 clas3 clas4 clas5 clas6
## nfami    18    56   110    88    40     8
```

```
#frequencias experadas
aux=0:5
prob=dbinom(aux,5,0.5)
fexp=prob*320
fexp
```

```
## [1]  10  50 100 100  50  10
```

Sendo

E_i: frequências experadas
O_i: frequências observadas

Onde uma hipótese equivalente é

$H_0 : E_i - O_i = 0$
$H_1 : E_i - O_i \neq 0$

A estatística de teste é dado por

$$Q = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

onde $Q \sim X^2_{k-1}$

Portanto, temos

```
n=320 # numero de familias

Ei=fexp # frequencias esperadas

Oi=fob # frequências Observadas

k=length(Oi) # quantidade de frequências esperadas

Q=sum(((Oi-Ei)^2)/Ei) # Estatística de teste
Q
```

```
## [1] 11.96
```

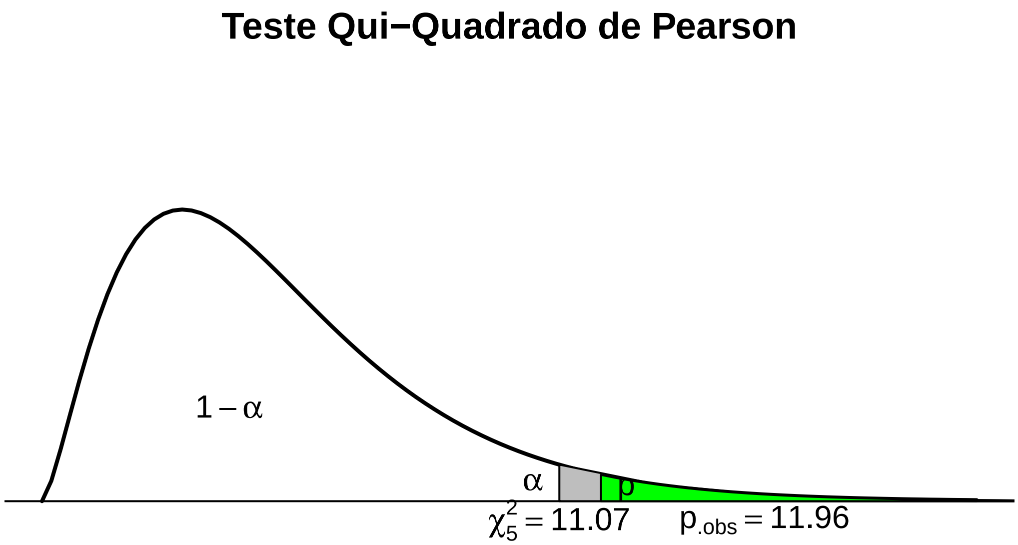
```r
p_valor=pchisq(Q,k-1,lower=F) # calculando o p-valor
p_valor
```

```
## [1] 0.03534
```

```r
a=0.05
qa=qchisq(1-a,k-1) # quantil teorico com alpha=0.05
qa
```

```
## [1] 11.0705
```

Vemos que $Q_{obs} > Q_{0.05}$. portanto rejeitamos H_0, ou seja,temos indicios a probabilidade de nascer um homem e a probabilidade de nascer uma mulher são diferentes para um $\alpha = 0.05$.
Abaixo podemos ver o gráfico do teste como p-valor.



Estimando o parâmetro e refazendo o teste

Temos que o estimador de máxima verossimilhança para p na distribuição $binomial(n, p)$ é dado por

$\hat{p} = \frac{Totaldehomens}{Totaldehomens+Totaldemulheres}$

```r
totalh=sum(data[3,1:6]*data[2,1:6]) # total de homens

totalm=sum(data[3,1:6]*data[1,1:6]) # total de mulheres

fh=totalh/(totalh+totalm) # porcentagem de homens
fh
```

```
## [1] 0.4625
```

Portano recalculando as novos frequências esperadas temos

```r
# recalculando as frequencias experadas
aux2=0:5
prob2=dbinom(aux,5,fh)
prob2
```

```
## [1] 0.04486342 0.19301702 0.33216883 0.28581969 0.12296894 0.02116210
```

```r
fexp2=prob2*320
fexp2 # nova frequencia experada
```

```
## [1]  14.356293  61.765448 106.294026  91.462302  39.350060   6.771871
```

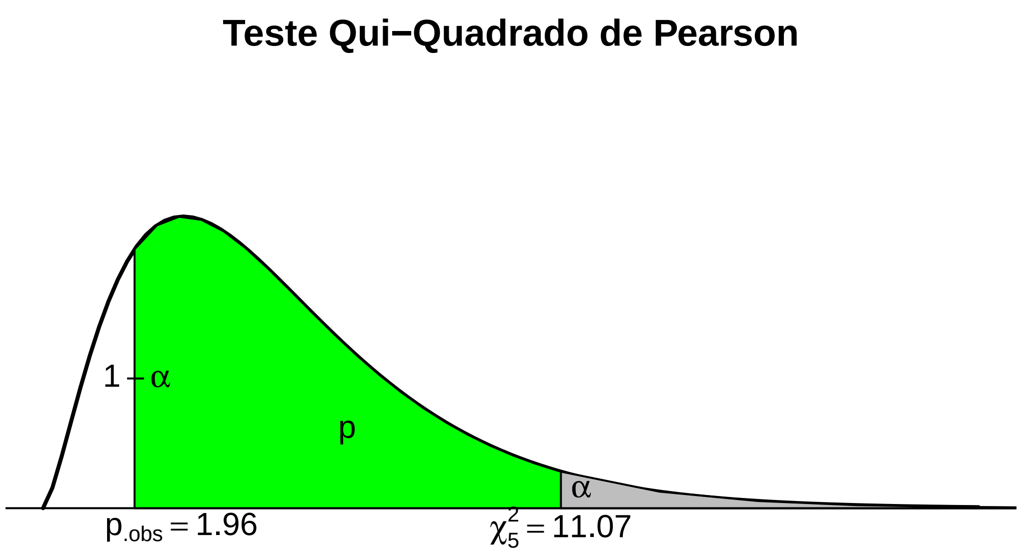
Ficamos com o seguinte teste de hipóteses:

$H_0 : p = 0.4625$
$H_1 : p \neq 0.4625$

Refazendo as contas com a nova frequência experada chegamos no seguinte gráfico

```
## [1] 0.8551028
```

```
## [1] 11.0705
```



Vemos que o p-valor é maior que um $\alpha = 0.05$, ou seja não reitamos H_0, isso implicaque que para um $\alpha = 0.05$ temos indicios que a probabilidade de nascer um homem não é diferente de que 0,4625. Como a probabilidade de nascer uma mulher é complementar à de um homem podemos dizer que temos suspeita de que a probabilide de nascer um homem é diferente da probabilidade de nascer uma mulher.

Você pode acessar esse trabalho através do link "rpubs.com/Giovani/exer1non", lá os gráficos e as linhas de código estarão coloridos.