

CEFET/RJ
Bacharelado em Ciência da Computação
Inferência Estatística - Trabalho 04

Prof. Eduardo Bezerra (ebezerra@cefet-rj.br)

24/11/2018

Conteúdo

1	Distribuição χ^2	3
2	Teste χ^2 (I)	3
3	Teste χ^2 (II)	3
4	Teste χ^2 (III)	4

1 Distribuição χ^2

Suponha que uma variável aleatória Y siga a distribuição χ^2 com k graus de liberdade, isto é, $Y \sim \chi^2_{(k)}$. Por meio das funções apropriadas do R (**chisqp**, etc.), compute o que se pede a seguir.

- (i) $P(Y > 2,56)$, considerando que $k = 13$
- (ii) $P(2,56 < Y < 4,87)$, considerando que $k = 13$
- (iii) O valor de k tal que, $P(Y < k) = 0,95$

2 Teste χ^2 (I)

O R contém um conjunto de dados denominado **mtcars**. Para este conjunto de dados, descubra se as variáveis **cyl** e **carb** são ou não dependentes. Para isso, utilize o teste χ^2 . Apresente o desenvolvimento, isto é, os comandos em R que você utilizou para chegar à conclusão.

3 Teste χ^2 (II)

A fabricante das balinhas Zuzuba alega que cada pacote produzido dessas balinhas contém balinhas de diferentes cores (primeira coluna da tabela abaixo), e que as quantidades de balinhas de cada cor não difere significativamente das que são apresentadas na segunda coluna da tabela abaixo. Se a fabricante está fazendo seu trabalho corretamente, então não deve haver diferença significativa entre as proporções de diversas cores de Zuzubas que ela alega depositar em cada pacote e a que encontrei quando comprei um pacote de Zuzubas em uma loja perto de minha casa (terceira coluna na tabela abaixo).

cor	esperado	observado
vermelho	18	24
verde	19	16
roxo	16	13
azul	6	20
laranja	24	20
amarelo	17	14

- (i) Utilizando o R, apresente dois gráficos de setores (*pie charts*), um para as quantidades esperadas e outro para as quantidades observadas de cores.
- (ii) Também utilizando R, teste a hipótese nula de que a alegação da fabricante é verdadeira. Apresente os seguintes valores: graus de liberdade, valor da estatística e o valor- p . Apresente também a sua conclusão, contra ou a favor da hipótese nula, justificando sua resposta.

4 Teste χ^2 (III)

Considere um caso hipotético em que se deseja testar a eficácia de um medicamento para um determinado problema médico. Suponha que temos 105 pacientes em estudo e 50 deles foram tratados com a droga. Os restantes 55 pacientes foram mantidos como amostras de controle. O estado de saúde de todos os pacientes foi verificado após uma semana. Os dados e resultados para todos esse indivíduos podem ser encontrados no seguinte endereço: <https://raw.githubusercontent.com/selva86/datasets/master/treatment.csv>.

A tabela de contingência para o conjunto de dados fornecido pode ser produzida por meio dos comandos a seguir.

```
url <- "https://raw.githubusercontent.com/
selva86/datasets/master/treatment.csv"
df <- read.csv(url)
tbl = table(df$treatment, df$improvement)
```

- (i) Utilizando o R, apresente a tabela em um formato que exibe as quantidades observadas de forma gráfica (veja exemplo fornecido em aula).
- (ii) Também utilizando R, verifique a alegação de que as duas variáveis são dependentes.