

Proposta de resolução dos Exercícios 5, 6 e 7 - Projeto Computacional PE 21/22

Conceição Amado

2022

Pergunta 5 (exemplo)

O tempo de atendimento (em minutos), de doentes graves num determinado hospital, é modelado por uma variável aleatória X com distribuição Exponencial de valor esperado $1/\lambda = 1/0.32$, i.e. $X \sim \text{Exp}(\lambda = 0.32)$.

Usando o R e fixando a semente em 767, gere uma amostra de dimensão $n = 630$ dessa variável. Para essa amostra, calcule a função de distribuição empírica e com base nessa função estime a probabilidade do tempo de atendimento de um doente grave, selecionado ao acaso, ser maior que 5. Calcule ainda o valor teórico dessa probabilidade.

Indique o valor absoluto da diferença entre o valor teórico e o estimado dessa probabilidade utilizando 6 casas decimais.

Código para confirmação (Resposta: valor numérico)

```
seed.dada<-767
n.dado<-630
lambda.dado<-0.32
set.seed(seed.dada)
vx<-5

amostra<-rexp(n.dado,rate=lambda.dado)
### dis empirica
Fn <- ecdf(amostra)
### prob estimada e teorica no ponto vp0
p.teo<-1-pexp(vx,rate=lambda.dado)
p.est<-1-Fn(vx)
dif.p<-abs(p.teo-p.est)
dif.p<-round(dif.p,6)
dif.p
```

```
[1] 0.009215
```

Pergunta 6 (exemplo)

Para cada um dos três valores de $n = 4, 29$ e 65 , fixe a semente de geração em 1712 e gere 1250 amostras de uma população, X , com distribuição Uniforme contínua no intervalo $[14,18]$.

Para cada valor de n : Calcule a média de cada uma das amostras obtendo assim valores da distribuição da média \bar{X}_n .

Faça o histograma de frequência relativa associado aos valores obtidos da distribuição da média \bar{X}_n . Sobreponha no gráfico uma curva com a distribuição normal com valor esperado $E(X)$ e variância $Var(X)/n$.

Submeta um ficheiro em formato PDF, com uma única página, que inclua: Valores dos parâmetros: semente, dimensões das amostras e os parâmetros da distribuição uniforme.

- O código em R.
- Para cada n o gráfico construído.
- Comentários sobre os resultados obtidos.

Exemplo de código (há muitas alternativas para fazer este código... , com ciclos, com replicate, etc)

Cotação:

- Código da geração - 0.9 valor
- Código dos gráficos e qualidade dos gráficos (legendas, título, apresentação geral), verificar se usam os valores teóricos de $E(X)$ e $V(X)/n$ para a curva da normal - 0.9 valor
- Comentário - ex. “o ajuste melhora com o aumento da dimensão da amostra” - 0.2 valor

NOTAS sobre a avaliação:

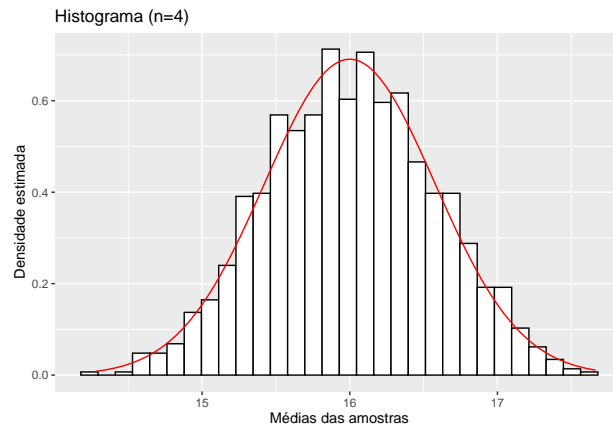
1. quem fez os histogramas usando a frequência relativa mas não com área total unitária, deverá ter a cotação completa, pois era o que estava no enunciado.
2. quem fez os histograms usando a função “hist” não deve ser penalizado, pois não é pedido para fazer com a package ggplot
3. Quem apresentar o código para um “n” (correcto!!) e disser que é igual para os restantes valores de “n” não deve ser penalizado pois pode não ter espaço numa página A4 para colocar toda a informação.

```
set.seed(1712)
n1<-4
n2<-29
n3<-65
a<-14
b<-18
n.amostras<-1250
media.unif <- (a+b)/2
var.unif<-((b-a)^2)/12

### para uniforme
amostras1<-matrix(runif(n1*n.amostras,a,b),n.amostras)
amostras2<-matrix(runif(n2*n.amostras,a,b),n.amostras)
amostras3<-matrix(runif(n3*n.amostras,a,b),n.amostras)
medias.amostras1<- apply(amostras1,1,mean)
medias.amostras2<- apply(amostras2,1,mean)
medias.amostras3<- apply(amostras3,1,mean)
```

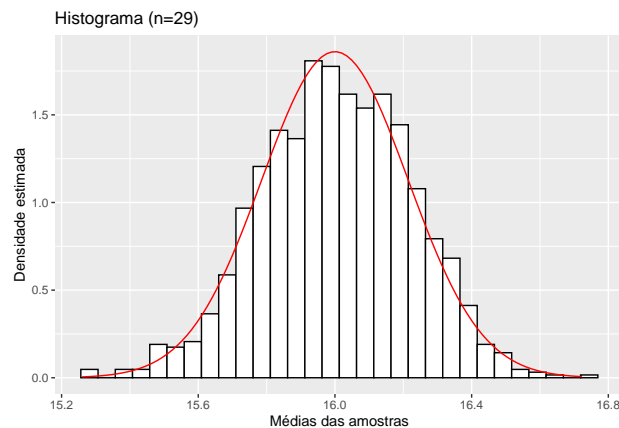
```
library(ggplot2)
df <- data.frame(PF = medias.amostras1)
p1<-ggplot(df, aes(x = PF)) +
  geom_histogram(aes(y = ..density..),
    colour = "black",
    fill = "white") +
  stat_function(fun = dnorm, args = list(mean = media.unif, sd = sqrt(var.unif/n1)), col="red")

p1+labs(title="Histograma (n=4)",x="Médias das amostras", y = "Densidade estimada")
```



```
## n2
df <- data.frame(PF = medias.amostras2)
p2<-ggplot(df, aes(x = PF)) +
  geom_histogram(aes(y = ..density..),
    colour = "black",
    fill = "white") +
  stat_function(fun = dnorm, args = list(mean = media.unif, sd = sqrt(var.unif/n2)), col="red")

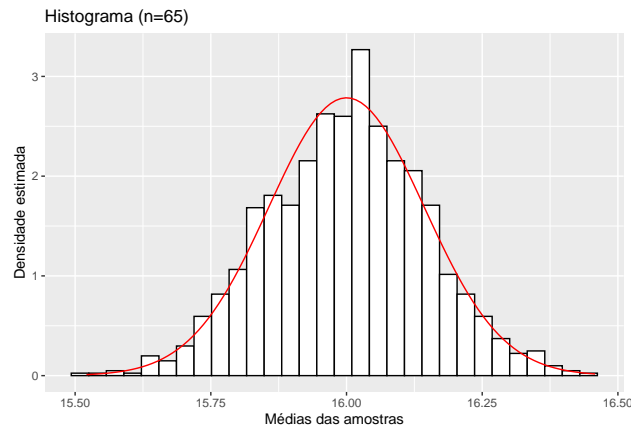
p2+labs(title="Histograma (n=29)",x="Médias das amostras", y = "Densidade estimada")
```



```
### n3
df <- data.frame(PF = medias.amostras3)
p3<-ggplot(df, aes(x = PF)) +
```

```
geom_histogram(aes(y = ..density..),
               colour = "black",
               fill = "white") +
  stat_function(fun = dnorm, args = list(mean = media.unif, sd = sqrt(var.unif/n3)), col="red")

p3+labs(title="Histograma (n=65)", x="Médias das amostras", y = "Densidade estimada")
```



Pergunta 7 (exemplo)

Usando o R e fixando a semente em 1328, gere 1280 amostras de dimensão 104 de uma população, X , com distribuição Binomial(49,0.47). Calcule a média de cada uma dessas amostras, obtendo uma amostra de médias.

Indique o valor absoluto da diferença entre a média da amostra de médias e o valor esperado da distribuição teórica de X utilizando 6 casas decimais.

Código para confirmação (Resposta: valor numérico)

```
seed.dada<-1328
n.replicas<-1280
n.amostra<-104
n.bin<-49
p.bin<-0.47
media.binom<-n.bin*p.bin

set.seed(seed.dada)
amostras<-matrix(rbinom(n.amostra*n.replicas,size=n.bin,prob=p.bin),n.replicas)
medias.amostras<- apply(amostras,1,mean)
media.medias<-mean(medias.amostras)
dif.medias<-abs(media.medias-media.binom)
dif.medias<-round(dif.medias,6)
dif.medias
```

```
[1] 0.00125
```