Estatística computacional - Estatística - ce223 (www.leg.ufpr.br/ce223)
Prof. Walmes Marques Zeviani - Departamento de Estatística - UFPR
Segunda avaliação - Probabilidade e inferência estatística (31/05/2011)
Acadêmico: Gabarito da Prova Tirei Dez Matrícula: GRR2011xxxx

Instruções:
- Você pode consultar a internet e o cartão de referência do R. Você não pode consultar o caderno nem trocar informações com colegas;
- Todas as questões devem ser respondidas na folha de prova, com os códigos empregados e os resultados obtidos;
- É altamente recomendável que você faça a prova no computador do laboratório (linux). Isso porque a importação de dados pode gerar problemas devido a codificação de caracteres. Caso faça a prova no seu computador pessoal, a responsabilidade de fazer a importação total e correta dos conjuntos de dados é sua.

1. Considere a função abaixo e obtenha, numericamente

$$f(x) = \begin{cases} c \cdot x^2 & \text{se } 0 < x \leq 2 \\ 0 & \text{outros casos.} \end{cases} \tag{1}$$

a) o valor da constante c para que $f(x)$ seja uma função densidade de probabilidade;
b) qual o valor esperado de X (use o valor de c encontrado a partir de agora);
c) qual a variância de X;
d) faça o gráfico da função $f(x)$;
e) represente com uma reta vertical a $\mathrm{E}(X)$.

```
> #-----------------------------------------------------------------------------------------
> # função escrita em código R, inicialmente com c=1
>
> c <- 1
> fx <- function(x){ sapply(x, function(xx) ifelse(xx>=0 && xx<=2, c*xx^2, 0)) }
> #-----------------------------------------------------------------------------------------
> # valor de c para que integral da função seja 1
>
> c <- 1/integrate(fx, 0, 2)$value
> c
[1] 0.375
> #-----------------------------------------------------------------------------------------
> # nova função com o valor de c atualizado, veja que a integral dá 1
>
> fx <- function(x){ sapply(x, function(xx) ifelse(xx>=0 && xx<=2, c*xx^2, 0)) }
> integrate(fx, 0, 2)$value
[1] 1
> #-----------------------------------------------------------------------------------------
> # esperança
>
> efx <- function(x){ sapply(x, function(xx) ifelse(xx>=0 && xx<=2, xx*c*xx^2, 0)) }
> ex <- integrate(efx, 0, 2)$value
> ex
[1] 1.5
> #-----------------------------------------------------------------------------------------
> # variancia
>
> vfx <- function(x){ sapply(x, function(xx) ifelse(xx>=0 && xx<=2, (xx-ex)^2*c*xx^2, 0)) }
> vx <- integrate(vfx, 0, 2)$value
> vx
[1] 0.15
> #-----------------------------------------------------------------------------------------
> # gráfico com a linha média
>
> pdf("fig1.pdf", w=5, h=5)
> par(mar=c(5.1,4.1,2.1,1))
> curve(fx, -1, 3, ylab="f(x)", main="Função densidade de probabilidade de X")
> abline(v=ex)
> mtext("E(X)", side=1, line=0, at=ex)
> dev.off()
null device
          1
>
+ #-----------------------------------------------------------------------------------------
```

2. Encontre as probabilidades para os eventos abaixo:

a) se $X \sim \text{Normal}(\mu = 3, \sigma = 2)$, qual $\text{P}(X < 2{,}22)$;

b) se $X \sim \text{Poisson}(\lambda = 5{,}5)$, qual $\text{P}(X = 2)$;

c) se $X \sim \text{Poisson}(\lambda = 12{,}1)$, qual $\text{P}(X > 5)$;

d) se $X \sim \text{binomial}(n = 21, p = 0{,}28)$, qual $\text{P}(X \leq 6)$;

e) se $X \sim \text{hipergeometrica}(m = 22, n = 16, k = 5)$, qual $\text{P}(X = 3)$;

f) se X tem dsitribuição representada pela $f(x)$ acima, qual $\text{P}(0{,}5 < X < 1{,}25)$.

```
> #-----------------------------------------------------------------------------------------
> # abaixo os códigos para calculas as probilidades de cada evento
>
> pnorm(2.22, 3, 2)              # a)

[1] 0.3482683

> dpois(2, 5.5)                  # b)

[1] 0.06181242

> 1-ppois(5, 12.1)               # c)

[1] 0.9808965

> pbinom(6, 21, 0.28)            # d)

[1] 0.6304882

> dhyper(3, 22, 16, 5)           # e)

[1] 0.36817

> integrate(fx, 0.5, 1.25)$value # f)

[1] 0.2285156

>
+ #-----------------------------------------------------------------------------------------
```

3. O intervalo de confiança para uma variância populacional de uma v.a. de distribuição normal é obtido pela expressão

$$\text{IC}_{1-\alpha}(\sigma^2) = \left(\frac{(n-1) \cdot s^2}{\chi^2_{1-\alpha/2}} ; \frac{(n-1) \cdot s^2}{\chi^2_{\alpha/2}} \right), \qquad (2)$$

em que s^2 é a variância amostral, n é o número de elementos da amostra, α é o nível de significância do teste e χ^2_q representa o quantil q (que acumula q de área a sua direita) da distribuição χ^2 com $n-1$ graus de liberdade. Construa uma função para obter o esse intervalo de confiança. Se $s^2 = 3{,}25$, $n = 16$ e $\alpha = 0{,}9$, qual é o intervalo para σ^2?

```
> #-----------------------------------------------------------------------------------------
> # função para obter o IC para o parâmetro sigma^2
>
> ic.sigma2 <- function(s2, n, alpha){
+   ## s2 é a variância amostral
+   ## n é o número de elementos
+   ## alpha é o nível de significancia, 1-alpha é o nível de confiança
+   ## a função retorna um vetor com os limites inferior e superior
+   chi <- qchisq(c(inf=1-alpha/2, sup=alpha/2), df=n-1)
+   sq <- (n-1)*s2
+   lim <- sq/chi
+   return(lim)
+ }
> ic.sigma2(3.25, 16, 0.05)

     inf      sup
1.773476 7.784881

>
+ #-----------------------------------------------------------------------------------------
```

4. Com a função $f(x)$ obtenha

a) a função de distribuição acumulada $F(x)$;

b) crie a função para gerar números aleatórios da distribuição de X (use o princípio da inversa da função de distribuição);

c) obtenha uma amostra da v.a. X usando o procedimento que você criou;

d) faça o histograma para a sua amostra e adicione o gráfico da função $f(x)$;

e) obtenha a média, a variância, e os quantis 0,05, 0,25, 0,50, 0,75 e 0,95.

```
> #-----------------------------------------------------------------------------------------
> # a função de distribuição acumulada de X, F(x), é a integral de f(x), em código R é
>
> Fx <- function(x){ sapply(x, function(xx) ifelse(xx<=0, 0, ifelse(xx<=2, c*xx^3/3, 1))) }
> Fx(seq(-0.5, 2.5, by=0.5))
[1] 0.000000 0.000000 0.015625 0.125000 0.421875 1.000000 1.000000
> pdf("fig2.pdf", w=5, h=5)
> par(mar=c(5.1,4.1,2.1,1))
> curve(Fx(x), -1, 3, ylab="F(x)", main="Função de distribuição acumulada de X")
> dev.off()
null device
          1
> #-----------------------------------------------------------------------------------------
> # a inversa da função Fx é iFx, em código R é
>
> rfx <- function(u){
+   ## u é um número uniforme entre 0 e 1
+   ## o resultado da função é uma realização da v.a. X~f(x)
+   ## a função abaixo é a inversa da Fx, iFx
+   (3*u/c)^(1/3)
+ }
> #-----------------------------------------------------------------------------------------
> # uma amostra aleatória da distribuição X
>
> u <- runif(100)
> x <- rfx(u)
> str(x)
 num [1:100] 1.89 1.84 1.84 1.32 1.57 ...
> #-----------------------------------------------------------------------------------------
> # histograma com a função f(x)
>
> pdf("fig3.pdf", w=5, h=5)
> par(mar=c(5.1,4.1,2.1,1))
> hist(x, freq=FALSE, ylab="Densidade", main="histograma de uma amostra de X")
> curve(fx, add=TRUE)
> legend("topleft", legend="f(x)", lty=1, bty="n")
> box()
> dev.off()
null device
          1
> #-----------------------------------------------------------------------------------------
> # média, variância e quantis da amostra
>
> mean(x)
[1] 1.539956
> var(x)
[1] 0.1348235
> quantile(x, probs=c(5,25,50,75,95)/100)
      5%      25%      50%      75%      95%
0.790995 1.374379 1.574146 1.835528 1.957436
>
+ #-----------------------------------------------------------------------------------------
```