Exercício Aula 2

Diego G. de Paulo (10857040) Vitor Gratiere Torres (10284952)

27/04/2020

EX 1

Motivação: O intuito deste exercício é gerar uma amostra pseudo-aleatória de $f(x)$ dada por $f(x) \propto q(x) = e^{\frac{-|x|^3}{3}}$.

Para gerar tal amostra, foi selecionado o método da rejeição. Este método é descrito por:

A seleção de uma variável aleatória Y, com função de densidade dada por $g(y)$ amostrável. Além disso, há a suposição:

- $\frac{f(x)}{g(x)} \leq m$, $1 \leq m < \infty$

E, por recomendação, toma-se $m = max_x \left(\frac{f(x)}{g(x)}\right)$

Para o exercício em questão, seleciona-se Y, tal que $Y \sim Laplace(0,1)$ que tem a função de probabilidade dada por: $g(y) = \frac{1}{2}e^{-|y|}, y \in \mathbb{R}$. Para obter m, tem-se:

$$m = max_x \left(\frac{f(x)}{g(x)}\right) \iff \frac{d\left(\frac{e^{\frac{-|x|^3}{3}}}{\frac{1}{2}e^{-|x|}}\right)}{dx} = 0$$

Calculando a derivada de $\frac{f(x)}{g(x)}$:

$$\frac{d\left(\frac{f(x)}{g(x)}\right)}{dx} = \frac{d\left(\frac{e^{\frac{-|x|^3}{3}}}{\frac{1}{2}e^{-|x|}}\right)}{dx} = \frac{2e^{\frac{-|x|^3+3|x|}{3}}x(-x^2+1)}{|x|}$$

Igualando a zero:

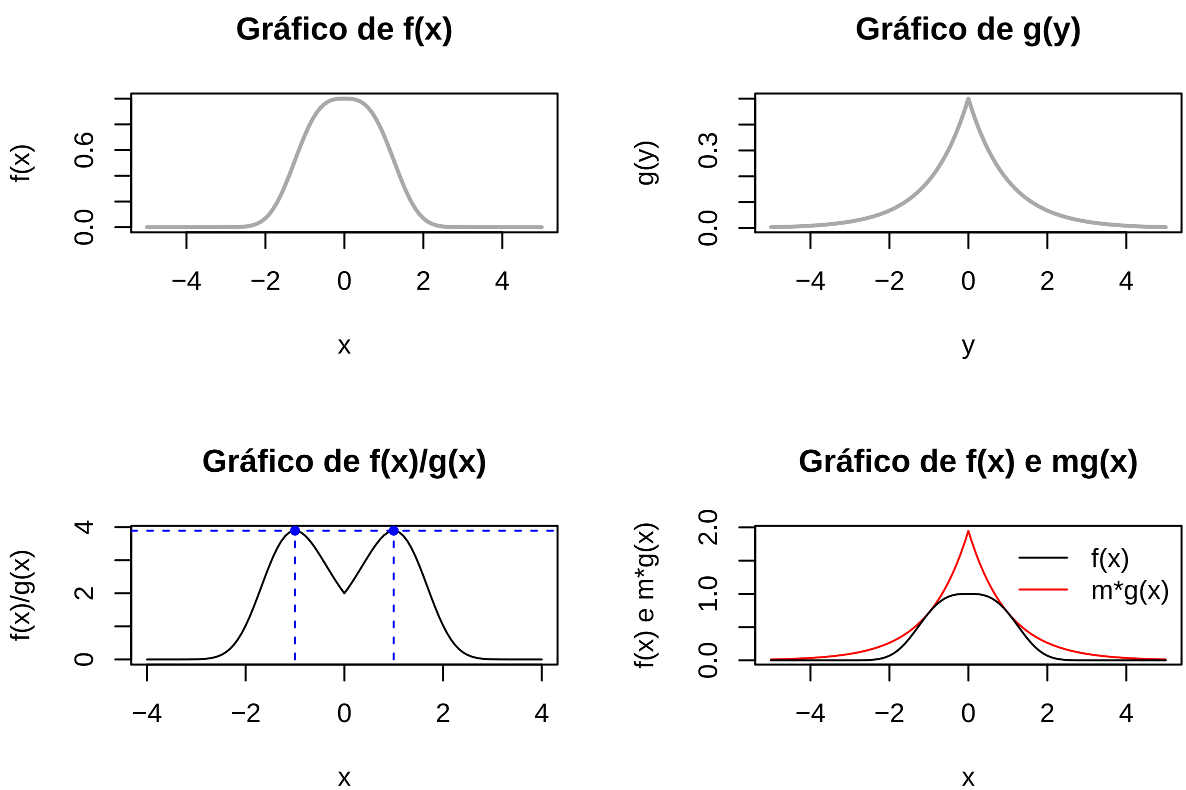
$$\frac{2e^{\frac{-|x|^3+3|x|}{3}}x(-x^2+1)}{|x|} = 0$$

Como solução para esta equação tem-se: $x = -1, 0, 1$, afim de não postergar o cálculo e partir para o gerador de amostra pseudo-aleatória, seleciona-se, dos pontos críticos, apenas os pontos de máximo em $x = -1, 1$. Sendo assim obtém-se: $m = max_x \left(\frac{f(x)}{g(x)}\right) = 2e^{\frac{2}{3}}$. Finalizada a etapa de seleção das variáveis, segue a apliacação das etapas:

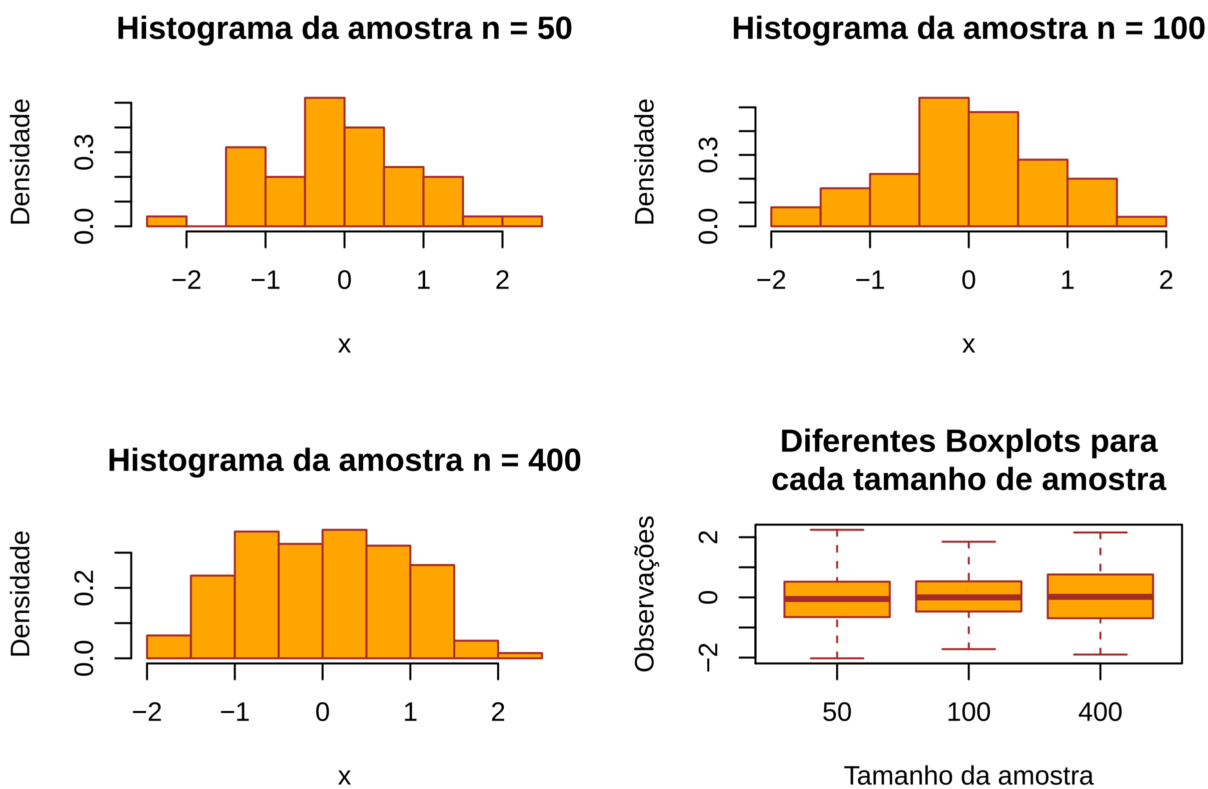
- 1º: Gerar uma amostra de Y
- 2º: Gerar $u \sim U(0,1)$

- 3º: Se $u \leq \frac{f(y)}{mg(y)}$ faça $x = y$, Caso contrário retornar ao primeiro passo.
- 4º: Repita os passos anteriores até obter n observações necessárias.

Abaixo, para melhor visualização dos resultados obtidos matematicamente, estão exibidos os gráficos das funções $f(x)$, $g(y)$, assim como os gráficos de $\frac{f(x)}{g(x)}$, para que os pontos critícos possam ser observados, e o gráfico sobreposto de $f(x)$ e $mg(x)$, para notar o envolapamento de $f(x)$ por $mg(x)$.



A seguir, é possível observar histogramas e boxplots para cada tamanho de amostra $n = (50, 100, 400)$ das amostras pseudo-aleatórias geradas da $f(x)$. É possível notar, observando o gráfico, que a medida em que se aumenta o tamanho da amostra mais próximo da simetria observada na função, dispõe-se a amostra.



EX 2

EX 3

Motivação: Neste presente exercício busca-se trabalhar com a distribuição poisson e estudar os Erros do Tipo I e II para esta distribuição utilizando o Método de Monte Carlo.

Para facilitar a compreensão, será indicado alguns aspectos importantes de tal distribuição. Dado $X \sim Poisson(\lambda)$, $X_1, ..., X_n$ é uma amostra aleatória de X de tamanho n.

- $\mathbb{P}(X = k) = \frac{e^{-\lambda}\lambda^x}{x!}$, para $\lambda > 0$ e $x = 0, 1, 2, ...$;
- $\mathbb{E}(X) = \lambda$;
- $\text{Var}(X) = \lambda$;
- $\hat{\lambda}_{EMV} = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}$, em que k_i é a i-ésima observação amostral de X_i;

Pela propriedade da eficiêcnia de um estimador de máxima verossimilhança, tem-se o resultado:

$$\sqrt{n}(\hat{\lambda}_{EMV} - \lambda) \xrightarrow{D} \mathcal{N}(0, \mathcal{IF}(\lambda)^{-1})$$

em que $\mathcal{IF}(\lambda) = \frac{1}{\lambda}$

Para o teste de hipótese:

- $H_0 : \lambda = 2$
- $H_A : \lambda > 2$

A Estatística do Teste T pode ser definida a partir do resultado da efiência do estimador de máxima verossimilhança.

$$T = \sqrt{n}(\hat{\lambda}_{EMV} - 2) \xrightarrow[\text{Sob } H_0]{D} \mathcal{N}\left(0, \frac{1}{2}\right)$$

```r
r <- 1000
lambda <- 2
po_sample <- function(r, lambda, n) {

  pvalue <- c()
  aceptance <- c()
  medias_v <- c()
  poder_dif <- seq(2.2, 4, length.out = 5) - 2
  power_1 <- c()


  for (i in 1:r) {

    sample_p <- rpois(n, lambda)
    media <- mean(sample_p)

    p_value <- pnorm(media, mean = lambda, sd = sqrt(1 / n * lambda))
    pvalue[i] <- p_value

    medias_v[i] <- media

    if (pvalue[i] >= 0.05) {
      aceptance[i] <- 1
    }
    else {
      aceptance[i] <- 0
    }
  }

  poder_1 <- power.t.test(n = n, delta = poder_dif[1],
                          sd = sqrt(1 / n * lambda),
                          sig.level = .05,
                          alternative = "two.sided",
                          type = "one.sample")
  power_1[1] <- (as.numeric(unlist(poder_1[5])))

  poder_2 <- power.t.test(n = n, delta = poder_dif[2],
                          sd = sqrt(1 / n * lambda),
                          sig.level = .05,
```

```r
                          alternative = "one.sided",
                          type = "one.sample")
  power_1[2] <- as.numeric(unlist(poder_2[5]))

  poder_3 <- power.t.test(n = n, delta = poder_dif[3],
                          sd = sqrt(1 / n * lambda),
                          sig.level = .05,
                          alternative = "one.sided",
                          type = "one.sample")
  power_1[3] <- as.numeric(unlist(poder_3[5]))

  poder_4 <- power.t.test(n = n, delta = poder_dif[4],
                          sd = sqrt(1 / n * lambda),
                          sig.level = .05,
                          alternative = "one.sided",
                          type = "one.sample")
  power_1[4] <- as.numeric(unlist(poder_4[5]))

  poder_5 <- power.t.test(n = n, delta = poder_dif[5],
                          sd = sqrt(1 / n * lambda),
                          sig.level = .05,
                          alternative = "one.sided",
                          type = "one.sample")
  power_1[5] <- as.numeric(unlist(poder_5[5]))

  results <- list(pvalor = pvalue, aceita = aceptance,
                  media = medias_v, poder = power_1)

  return(results)
}

a <- po_sample(r, lambda, 10)
b <- po_sample(r, lambda, 30)
c <- po_sample(r, lambda, 75)
d <- po_sample(r, lambda, 100)
```

Na tabela subsequente, é possível notar a tendência esperada, conforme o o tamanho n da amostra aumenta, mais próximo de 95% fica a taxa de aceitação de H_0. Este resultado é esperado pois, pela definição do teste de hipótese estatístico, ao fixarmos $\alpha = 0.05$ espera-se que em 95% dos casos testados sejam em direção a aceitação de H_0 se a suposição de normalidade dos dados realmente é válida.

```r
freqa <- table(a$aceita) * 100 / sum(table(a$aceita))
freqb <- table(b$aceita) * 100 / sum(table(b$aceita))
freqc <- table(c$aceita) * 100 / sum(table(c$aceita))
freqd <- table(d$aceita) * 100 / sum(table(d$aceita))

tabela <- rbind("Amostra n = 10" = freqa, "Amostra n = 30" = freqb,
```

```
                 "Amostra n = 75" = freqc, "Amostra n = 100" = freqd)
tabela_1 <- data.frame(tabela)

knitr::kable(tabela_1, booktabs = T, align = "c",
             caption = "Tabela de Erro Tipo I (em %)",
             col.names = c("Rejeita Hipótese Nula", "Não Rejeita Hipótese Nula"),
             format = "latex", escape = T) %>%
kable_styling(position = "center", latex_options = c("hold_position"))
```

\begin{table}[!h]

\caption{Tabela de Erro Tipo I (em %)}

	Rejeita Hipótese Nula	Não Rejeita Hipótese Nula
Amostra n = 10	3.1	96.9
Amostra n = 30	5.1	94.9
Amostra n = 75	4.1	95.9
Amostra n = 100	4.9	95.1

\end{table}

Agora, partindo para análise do poder, notamos outro resultado esperado na tabela seguinte, a tendência apresentada é apresentada pelo gráfico da seguinte forma: conforme cresce a diferença entre o valor testado do paramêtro ($\lambda = 2$) e/ou cresce o tamanho da amostra maior é o poder do teste, ou seja, torna-se cada vez mais sensível aos desvios do valor de $\hat{\lambda}_{EMV}$.

```
tabela_2 <- rbind("Amostra n = 10" = a$poder, "Amostra n = 30" = b$poder,
                  "Amostra n = 75" = c$poder, "Amostra n = 100" = d$poder)
p <- seq(2.2, 4, length.out = 5)
df2 <- data.frame(round(tabela_2, 2))

knitr::kable(df2, booktabs = T, align = "c",
             caption = "Tabela do Poder do Teste $\\lambda \\in[2,2;4]$",
             col.names = c(paste("$\\lambda =$", p[1]),
                           paste("$\\lambda =$", p[2]),
                           paste("$\\lambda =$", p[3]),
                           paste("$\\lambda =$", p[4]),
                           paste("$\\lambda =$", p[5])),
             format = "latex", escape = F) %>%
kable_styling(position = "center", latex_options = c("hold_position"))
```

Table 1: Tabela do Poder do Teste $\lambda \in [2, 2; 4]$

	$\lambda = 2.2$	$\lambda = 2.65$	$\lambda = 3.1$	$\lambda = 3.55$	$\lambda = 4$
Amostra n = 10	0.24	0.99	1	1	1
Amostra n = 30	0.98	1.00	1	1	1
Amostra n = 75	1.00	1.00	1	1	1
Amostra n = 100	1.00	1.00	1	1	1