Lista de Exercicios IPE#4C

Taiguara Melo Tupinambas

Entrega: 19 de junho de 2017

Exercício 1

Letra a

O estimador de máxima verossimilhança é o que maximiza a função $L(\theta|X) = \prod_{i=1}^{n} P(X_i|\theta)$. No caso da PMF data, temos que:

$$L(\theta|X) = \left(\frac{2\theta}{3}\right)^{N_0} \left(\frac{\theta}{3}\right)^{N_1} \left(\frac{2(1-\theta)}{3}\right)^{N_2} \left(\frac{(1-\theta)}{3}\right)^{N_3} \tag{1}$$

Em que N_i, com $i = 1, 2, 3, 4$ é a quantidade de observações repetidas.

Como a função logarítmica é monotonicamente crescente, o resultado da maximização é o mesmo, se derivarmos o logaritmo da expressão matemática.

Desta forma:

$$logL(\theta|X) = N_0(log(2) + log(\theta) - log(3)) + N_1(log(\theta) - log(3)) + \tag{2}$$
$$N_2(log(2) + log(1-\theta) - log(3)) + N_3(log(1-\theta) - log(3)) \tag{3}$$

Derivando e igualando a 0, temos:

$$\frac{dlogL(\theta|X)}{d\theta} = \frac{1}{\theta}(N_0 + N_1) - \frac{1}{1-\theta}(N_2 + N_3) = 0 \tag{4}$$

Isolando θ, encontramos a seguinte expressão do estimador de máxima verossimilhança, baseado nas observações:

$$\hat{\theta} = \frac{N_0 + N_1}{N_0 + N_1 + N_2 + N_3} \tag{5}$$

Para as dez observações obtidas, o resultado pelo MV seria:

```
obs<-c(3,0,2,1,3,2,1,0,2,1)
n<-integer(4)
for (i in 0:3) {
        n[i+1]<-sum(obs==i)
}

Theta_MV<-(n[1]+n[2])/(sum(n))

cat("Theta estimado pelo MV:", Theta_MV)
```

```
## Theta estimado pelo MV: 0.5
```

O estimador do métodos dos momentos para o único parâmetro θ compara o valor esperado da VA $E[X]$ em função de θ e compara com a a média amostral. Ou seja:

$$E[X] = 0.\frac{2\theta}{3} + 1\frac{\theta}{3} + 2\frac{2(1-\theta)}{3} + 3\frac{(1-\theta)}{3} \tag{6}$$

$$E[X] = \frac{7}{3} - 2\theta \tag{7}$$

$$\theta = \frac{7}{6} - \frac{E[X]}{2} \tag{8}$$

$$\hat{\theta} = \frac{7}{6} - \frac{\hat{\mu}}{2} \tag{9}$$

Para as dez observações obtidas, o resultado pelo MM seria:

```r
Theta_MM<-7/6-mean(obs)/2

cat("Theta estimado pelo MM:", Theta_MM)
```

```
## Theta estimado pelo MM: 0.4166667
```

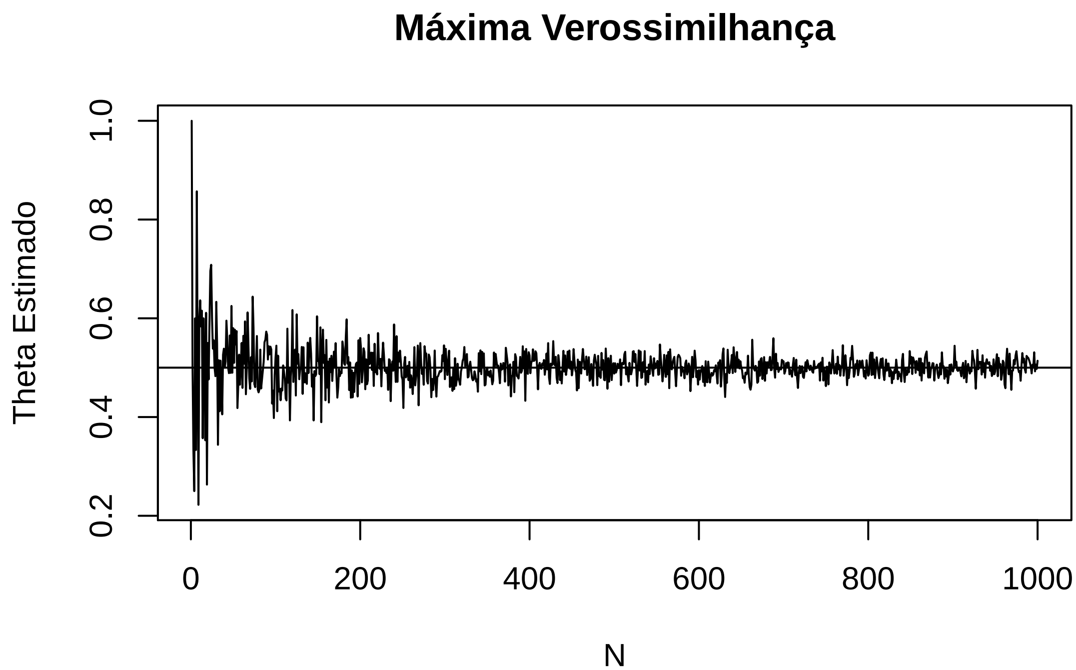
Letra b

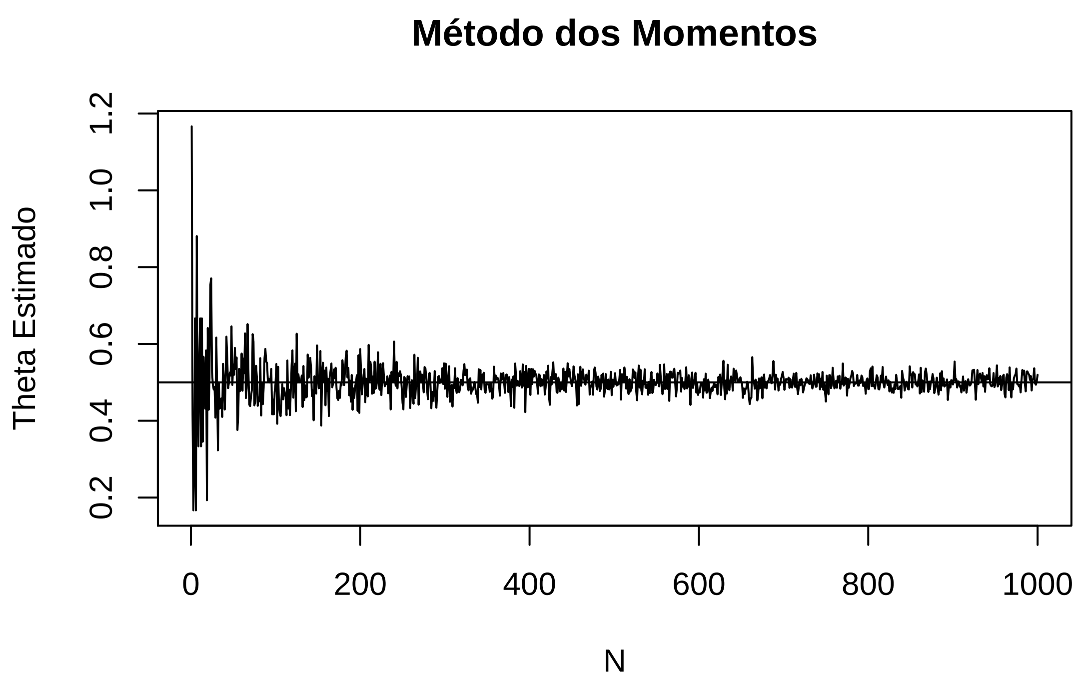
Os métodos foram aplicados para um N (quantidade de observações) crescente e o resultado da estimação de ambos é apresentada em gráficos, onde nota-se que para o aumento do tamanho da amostra os estimadores tendem a covergir.

O valor de θ escolhido para a simulação foi de 0.5.

```r
set.seed(1)

theta<-1/2
N<-1000

Theta_MV<-numeric(N)
Theta_MM<-numeric(N)

for (j in 1:N) {
        x<-integer(j)
        for (i in 1:j) {
                r=runif(1)
                if (r<2*theta/3) {
                        x[i]<-0
                } else if (r<theta) {
                        x[i]<-1
                } else if (r<theta+2*(1-theta)/3) {
                        x[i]<-2
                } else {
                        x[i]<-3
                }
        }

        # Maximum Likelihood
```

```
        n<-integer(4)
        for (k in 0:3) {
                n[k+1]<-sum(x==k)
        }
        Theta_MV[j]<-(n[1]+n[2])/(sum(n))

        # Método dos Momentos
        Theta_MM[j]<-(7/3-mean(x))/2
}

par(mfrow=c(2,1))

plot(Theta_MV,type="l",ylab="Theta Estimado",xlab="N",main="Máxima Verossimilhança")
abline(h=theta)
plot(Theta_MM,type="l",ylab="Theta Estimado",,xlab="N",main="Método dos Momentos")
abline(h=theta)
```





Exercício 2

O estimador de MV dado um conjunto de n observações é:

$$log(L(\sigma|X)) = \sum_{i=1}^{n} \left[-log(2) - log(\sigma) - \frac{|x_1|}{\sigma} \right] \tag{10}$$

$$\frac{dlog(L(\sigma|X))}{d\sigma} = \sum_{i=1}^{n} \left[-\frac{1}{\sigma} + \frac{|x_i|}{\sigma^2} \right] = 0 \tag{11}$$

$$-\frac{n}{\sigma} + \sum_{i=1}^{n} \left[\frac{|x_i|}{\sigma^2} \right] = 0 \tag{12}$$

$$\hat{\sigma}_{MV} = \sum_{i=1}^{n} \left[\frac{|x_i|}{n} \right] \tag{13}$$

$$\tag{14}$$

Para o estimador MM do laplaciano, é necessário calcular o segundo momento, já que $E[x] = 0 = \mu$. Dessa forma, temos que:

$$Var[X] = 2\sigma^2 = E[X^2] - E[X]^2 E[X^2] = 2\sigma^2 \tag{15}$$

Logo, para estimar σ com o MM a partir de um conjunto de n observações:

$$\hat{E}[X^2] = \sum_{n} \frac{|x_i|^2}{n} = 2\hat{\sigma}^2 \tag{16}$$

$$\tag{17}$$

$$\hat{\sigma} = \sqrt{\frac{\sum |x_i|^2}{2N}} \tag{18}$$

Para avaliar a convergência via simulação, foi definido um valor de $\sigma = 2$ e foram geradas observações utilizando a CDF inversa do laplaciano, dada por:

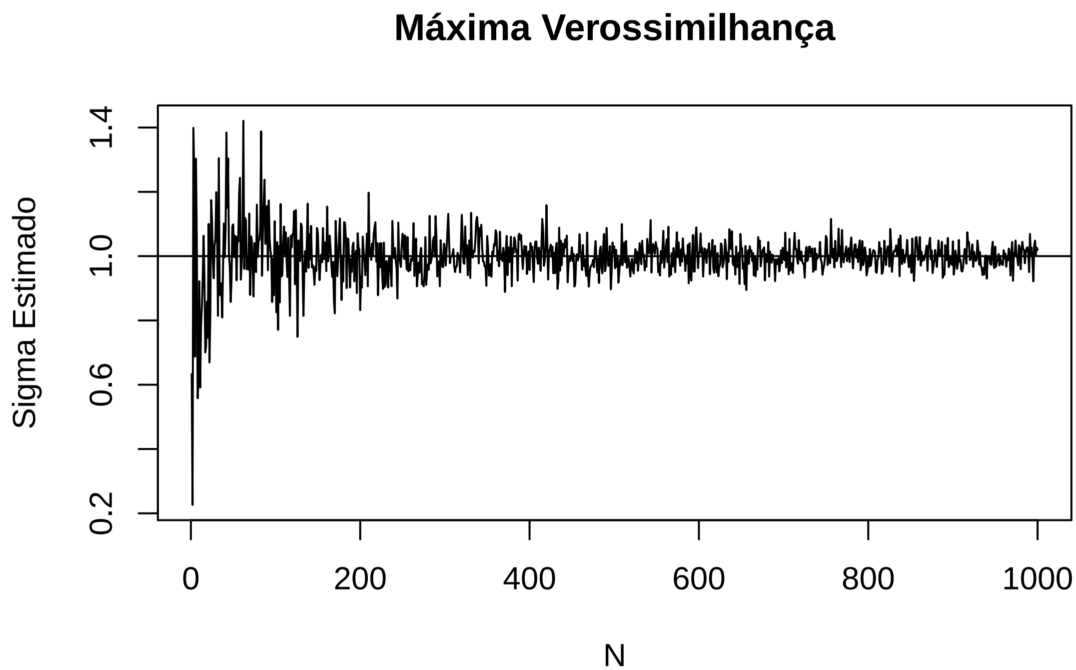
$$F^{-1}(x) = \mu - \sigma sgn(x - 0.5) ln(1 - 2|x - 0.5|) \tag{19}$$

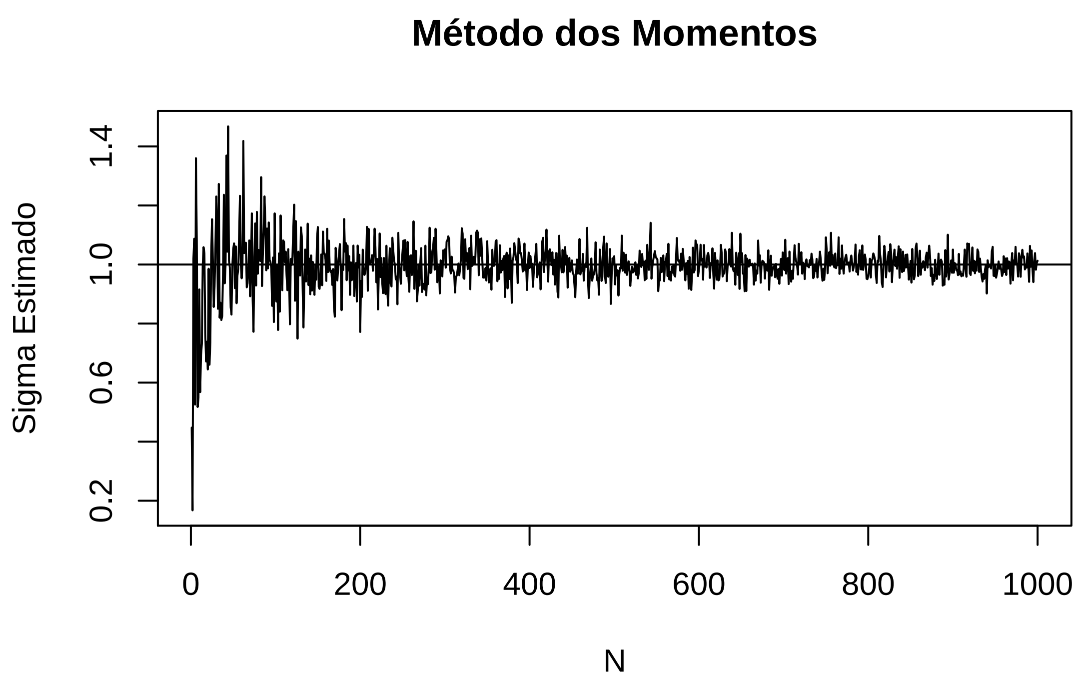
Em que $\mu = 0$.

Mais uma vez, nota-se que para o aumento do tamanho da amostra os estimadores tendem a covergir.

```
set.seed(1)

sigma<-1
N<-1000

sigma_MV<-numeric(N)
sigma_MM<-numeric(N)

for (j in 1:N) {
        u <- runif(j)
        x <- -sigma*sign(u-0.5)*log(1-2*abs(u-0.5))
```

```
        # Máxima Verossimilhança
        sigma_MV[j]<-sum(abs(x))/j

        # Método dos Momentos
        sigma_MM[j]<-sqrt(sum(abs(x)^2)/(2*j))
}

par(mfrow=c(2,1))

plot(sigma_MV,type="l",ylab="Sigma Estimado",xlab="N",main="Máxima Verossimilhança")
abline(h=sigma)
plot(sigma_MM,type="l",ylab="Sigma Estimado",,xlab="N",main="Método dos Momentos")
abline(h=sigma)
```





Exercício 3

O estimador de MV de $\theta = (\mu, \sigma)$ dado um conjunto de n observações é dado por:

$$log(L(\theta|X)) = \sum_{i=1}^{n}\Bigg[-log\sigma - \frac{1}{2}log2\pi - \frac{1}{2\sigma^2}\big(x_i-\mu\big)^2\Bigg] \tag{20}$$

$$log(L(\theta|X)) = -nlog\sigma - \frac{n}{2}log2\pi - \frac{1}{2\sigma^2}\sum_n\big(x_i-\mu\big)^2 \tag{21}$$

$$\tag{22}$$

$$\frac{\delta log(L(\theta|X))}{\delta\mu} = -\frac{1}{\sigma^2}\sum_n\big(x_i-\mu\big) = 0 \tag{23}$$

$$\frac{\delta log(L(\theta|X))}{\delta\sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_n\big(x_i-\mu\big)^2 = 0 \tag{24}$$

$$\tag{25}$$

Desta forma, resolvendo o sistema, temos que:

$$\hat{\mu}_{MV} = \bar{x} \tag{26}$$

$$\hat{\sigma}_{MV} = \sqrt{\frac{1}{n}\sum_n\big(x_i-\bar{x}\big)^2} \tag{27}$$

Desenvolvendo o somatório temos que:

$$\frac{1}{n}\sum_n\big(x_i-\bar{x}\big)^2 = \frac{1}{n}\sum_n\big(x_i^2-2x_i\bar{x}+\bar{x}^2\big) = \frac{1}{n}\sum_n x_i^2 - 2\bar{x}\frac{1}{n}\sum_n x_i + \bar{x}^2 \tag{28}$$

$$= \frac{1}{n}\sum_n x_i^2 - 2\bar{x}^2 + \bar{x}^2 = \frac{1}{n}\sum_n x_i^2 - \bar{x}^2 \tag{29}$$

Logo:

$$\hat{\mu}_{MV} = \bar{x} \tag{30}$$

$$\hat{\sigma}_{MV} = \sqrt{\frac{1}{n}\sum_n x_i^2 - \bar{x}^2} \tag{31}$$

O estimador pelo MM de $\theta = (\mu, \sigma)$ dado um conjunto de n observações é dado por:

$$E[X] = \mu \tag{32}$$

$$\hat{\mu}_{MM} = \bar{x} \tag{33}$$

$$\tag{34}$$

$$E[X^2] = Var[X] - \mu^2 = \sigma^2 - \mu^2 \tag{35}$$

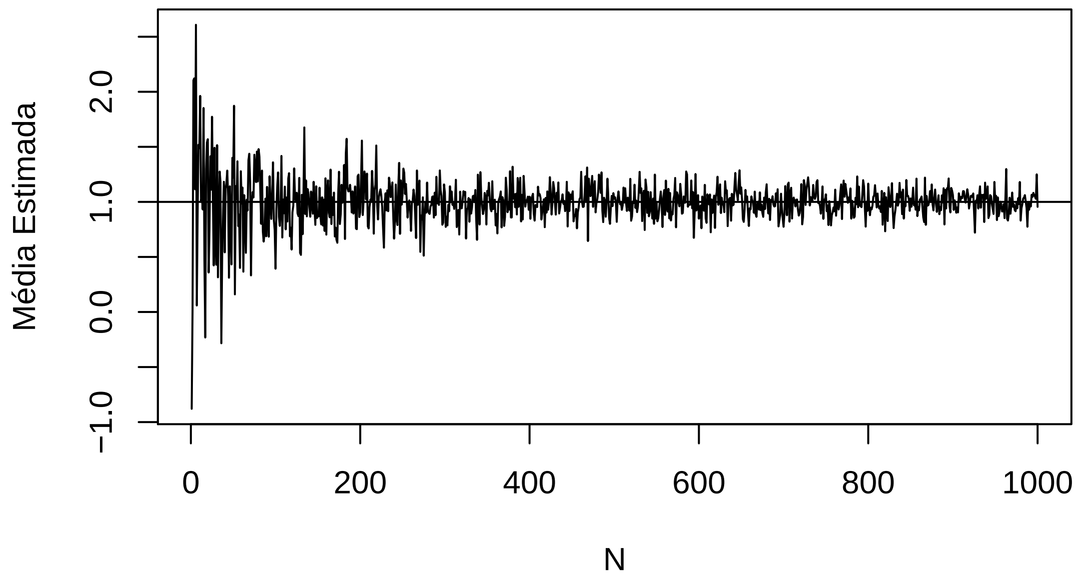
$$\sigma^2 = E[X^2] - \mu^2 \tag{36}$$

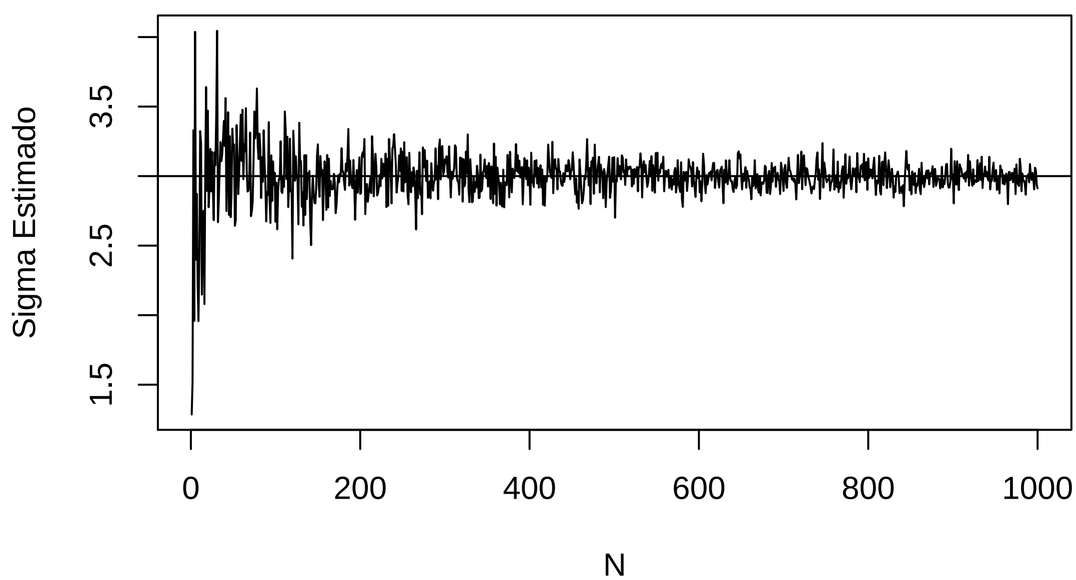
$$\hat{\sigma}_{MM} = \sqrt{\frac{1}{n}\sum_n x_i^2 - \bar{x}^2} \tag{37}$$

Desta forma, temos que $\theta_{MV} = \theta_{MM}$, c.q.d.

Para estimar via simulação foi considerado um $\theta = (\mu = 1, \sigma = 3)$ e o resultado obtido mostra que os estimadores parece convergir a medida que o tamanho amostral aumenta.

```
set.seed(1)

mu<-1
sigma<-3
N<-1000

mu_est<-numeric(N)
sigma_est<-numeric(N)

for (j in 1:N) {
        x <- rnorm(j,mean=mu,sd=sigma)

        # Máxima Verossimilhança
        mu_est[j]<-sum(x)/j

        # Método dos Momentos
        sigma_est[j]<-sqrt(1/j*sum(x^2)-mu_est[j])
}

par(mfrow=c(2,1))

plot(mu_est,type="l",ylab="Média Estimada",xlab="N")
abline(h=mu)
plot(sigma_est,type="l",ylab="Sigma Estimado",,xlab="N")
abline(h=sigma)
```


Média Estimada
−1.0
0.0
1.0
2.0
0
200
400
600
800
1000
N


Sigma Estimado
1.5
2.5
3.5
0
200
400
600
800
1000
N