

Exercícios sobre Verossimilhança - I

- Seja X uma v.a. com distribuição binomial com $n = 12$. Obtenha a função de verossimilhança para cada uma das observações a seguir e desenhe todas em um mesmo gráfico, escalonado se necessário.
 - $x = 5$
 - $x \leq 10$
 - $3 \leq x \leq 7$
- A fim de se obter uma estimativa do público de um jogo sem utilizar dados de venda de ingressos ou registros das roletas do estádio, foram distribuídas camisas especiais para 300 torcedores sob a condição que estes a utilizassem durante um jogo. Durante o jogo foram selecionados ao acaso 250 torcedores verificando-se 12 destes possuíam a camisa.
 - Obtenha a função de verossimilhança para o número total de torcedores
 - Obtenha a estimativa pontual e a intervalar, esta última por pelo menos dois métodos diferentes.
 - Repita e compare os resultados caso fossem 500 camisas e 20 com camisas dentre os 250.
- Para $X \sim N(\theta, 1)$ construa e compare as verossimilhanças para cada uma das seguintes observações. *Obs: use um escalonamento adequado das funções de verossimilhança para permitir uma melhor visualização.*
 - $x = 2.45$
 - $0.9 < x < 4$
 - somente o máximo de uma amostra de tamanho cinco é fornecido $x_{(5)} = 3.5$
- Para $X \sim N(\theta, 1)$. Construa a função de verossimilhança para dados provenientes de três amostras que forneceram as informações descritas a seguir. Mostre no mesmo gráfico as verossimilhanças individuais e conjunta. Se preciso use um escalonamento adequado das funções de verossimilhança para permitir uma melhor visualização.
 - $\bar{x} = 10$ em uma amostra de tamanho $n = 3$;
 - o máximo $x_{(5)} = 8$ em uma amostra de tamanho $n = 5$;
 - o mínimo $x_{(1)} = 7.5$ em uma amostra de tamanho $n = 4$
- Obtenha estimativas de máxima verossimilhança para os dois problemas anteriores

6. Seja X uma v.a. com distribuição uniforme $X \sim U(0, \theta)$.

Uma amostra aleatória forneceu os seguintes valores: 5,5; 3,2; 4,8, 5,3; 3,8 e 5,0. Obtenha a função de verossimilhança e um (ou mais) intervalo(s) adequado(s) especificando a forma de obtenção.

7. Seja X uma v.a. de uma distribuição de Poisson ($X \sim P(\lambda)$) para a qual foi obtida a seguinte amostra aleatória: (3, 1, 0, 2, 1, 1, 0, 0)

(a) Obtenha a função de verossimilhança, sua aproximação quadrática e intervalos para λ .

(b) Repita a questão anterior para a reparametrização $\theta = \log(\lambda)$.

(c) Obtenha ainda (por pelo menos dois métodos diferentes) intervalos de confiança para o parâmetro λ a partir da função de verossimilhança (aproximada ou não) de θ .

3. Sejam $\phi(\cdot)$ e $\Phi(\cdot)$ a densidade e densidade acumulada da normal padrão, respectivamente. As verossimilhanças para cada uma das observações são calculadas da seguinte forma:

$$L(\theta; x) = \phi(x - \theta) \equiv \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}(x - \theta)^2\right\};$$

$$L_1 = L(\theta; x = 2.45) = \phi(x - \theta) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}(2.45 - \theta)^2\right\};$$

$$L_2 = L(\theta; 0,9 < x < 4) = \Phi(4 - \theta) - \Phi(0,9 - \theta);$$

$$L_3 = L(\theta; x_{(5)} = 3.5) = n\{\Phi(x_{(n)} - \theta)\}^{n-1}\phi(x_{(n)} - \theta).$$

Note que a última verossimilhança decorre de um argumento multinomial e com

$$F(y) = P(X_{\{n\}} \leq y) = P[X_{\{i\}} < y \forall i \neq n \text{ e } X_{\{n\}} = y]$$

Os códigos para obtenção das verossimilhanças são mostrados a seguir.

```
> theta.vals <- seq(-1, 6, l=201)
> L1 <- function(theta) (1/sqrt(2*pi)) * exp(-0.5 * (2.45 - theta)^2)
> ## ou, usando a função de densidade da normal do R ...
> L1 <- function(theta) dnorm(2.45, m=theta, sd=1)
> L1.vals <- L1(theta.vals)
> plot(theta.vals, L1.vals/max(L1.vals), ty="l", xlab=expression(theta),
+       ylab=expression(L(theta)))
> ##
> L2 <- function(theta) pnorm(4, mean=theta, sd=1) - pnorm(0.9, mean=theta, sd=1)
> L2.vals <- L2(theta.vals)
> lines(theta.vals, L2.vals/max(L2.vals), ty="l", lty=2)
> ##
> L3 <- function(theta) 5*pnorm(3.5, m=theta, s=1)^4 * dnorm(3.5, m=theta, s=1)
> L3.vals <- L3(theta.vals)
> lines(theta.vals, L3.vals/max(L3.vals), ty="l", lty=3)
> #
> legend("topright", c("x=2.45", "0,9<X<4", "x[5]=3.5"), lty=1:3)
```

Pode-se ainda considerar a função de verossimilhança conjunta das três observações que, assumindo independência, é dada pelo produto da verossimilhanças individuais $L(\theta) = L_1 \cdot L_2 \cdot L_3$.

```

> L4 <- function(theta){
+   L3 <- function(theta) 3*pnorm(3.5, m=theta, s=1)^2 * dnorm(3.5, m=theta, s=1)
+   L1(theta) * L2(theta) * L3(theta)
+ }
> L4.vals <- L4(theta.vals)
> lines(theta.vals, L4.vals/max(L4.vals), ty="l", lty=4)

```

Curvas da função de log-verossimilhança $l(\theta) = \log[L(\theta)]$ podem ser obtidas notando que, e geral, este é um cálculo computacionalmente mais adequado e estável.

```

> l1 <- function(theta) -0.5 * (log(2*pi) + (2.45 - theta)^2)
> ## ou ...
> l1 <- function(theta) dnorm(2.45, m=theta, sd=1, log=TRUE)
> l1.vals <- l1(theta.vals)
> plot(theta.vals, l1.vals-max(l1.vals), ty="l", xlab=expression(theta),
+       ylab=expression(l(theta)))
> ##
> l2 <- function(theta) log(pnorm(4, mean=theta, sd=1) - pnorm(0.9, mean=theta, sd=1))
> l2.vals <- l2(theta.vals)
> lines(theta.vals, l2.vals-max(l2.vals), ty="l", lty=2)
> ##
> l3 <- function(theta) log(5) + 4*pnorm(3.5, mean=theta, sd=1, log=T) +
+       dnorm(3.5, mean=theta, sd=1, log=T)
> l3.vals <- l3(theta.vals)
> lines(theta.vals, l3.vals-max(l3.vals), ty="l", lty=3)
> #
> l4 <- function(theta) l1(theta) + l2(theta) + l3(theta)
> l4.vals <- l4(theta.vals)
> lines(theta.vals, l4.vals- max(l4.vals), ty="l", lty=4)
> #
> legend("topright", c("x=2.45", "0,9<X<4", "x[5]=3.5", "conjunta"), lty=1:4)

```

Outra alternativa é traçar curvas da função deviance $D(\theta) = -2[l(\theta) - l(\hat{\theta})]$. Nos gráficos a seguir utilizamos valor máximo computado para a sequência de valores para o parâmetro como uma aproximação de $l(\hat{\theta})$.

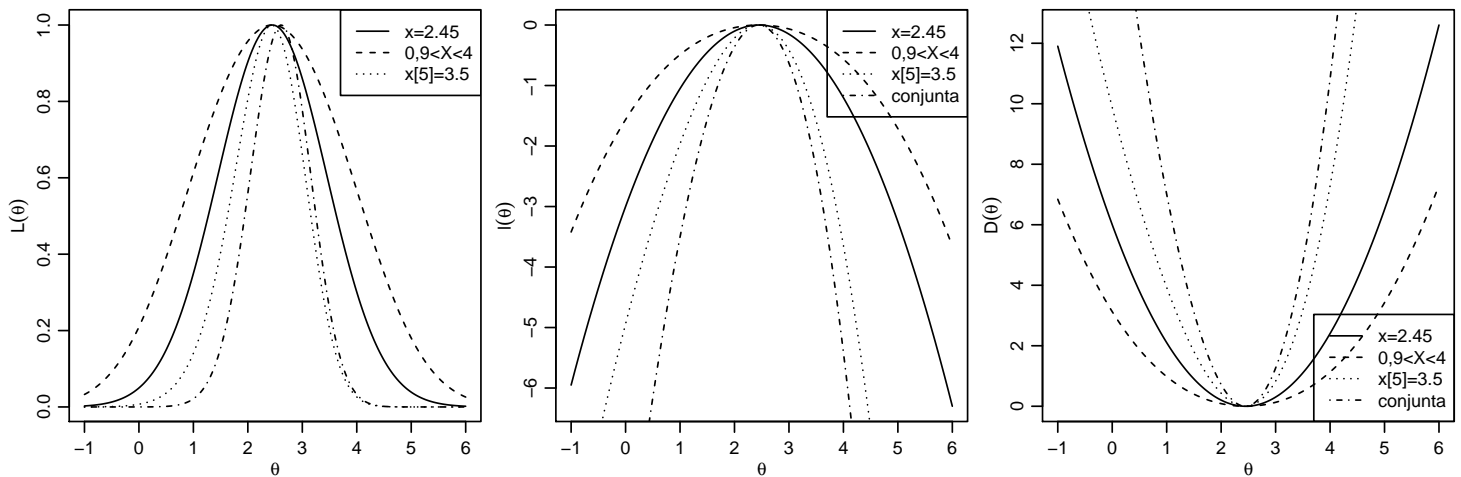


Figura 1: Verossimilhanças, log-verossimilhanças e deviances para as observações individuais e a conjunta.

```
> plot(theta.vals, -2*(l1.vals - max(l1.vals)), ty="l", xlab=expression(theta), ylab=L(theta))
> lines(theta.vals, -2*(l2.vals - max(l2.vals)), ty="l", lty=2)
> lines(theta.vals, -2*(l3.vals - max(l3.vals)), ty="l", lty=3)
> lines(theta.vals, -2*(l4.vals - max(l4.vals)), ty="l", lty=4)
> legend("bottomright", c("x=2.45", "0,9<X<4", "x[5]=3.5", "conjunta"), lty=1:4)
```

Estimativa para Exercício 3:

```
5. > ll.ex3 <- function(theta){
+   l1 <- -0.5 * (log(2*pi) + (2.45 - theta)^2)
+   l2 <- log(pnorm(4, mean=theta, sd=1) - pnorm(0.9, mean=theta, sd=1))
+   l3 <- log(5) + 4*pnorm(3.5, mean=theta, sd=1, log=T) +
+     dnorm(3.5, mean=theta, sd=1, log=T)
+   return(l1+l2+l3)
+ }
> est <- optimize(ll.ex3, interval=c(0,5), maximum=TRUE)
> unlist(est)

maximum objective
2.442739 -1.544042
```

O gráfico da função de log-verossimilhança indicando o valor da estimativa é mostrado na gráfico à esquerda da figura 2 e a função deviance à direita. Note que a deviance requer o valor da verossimilhança maximizada.

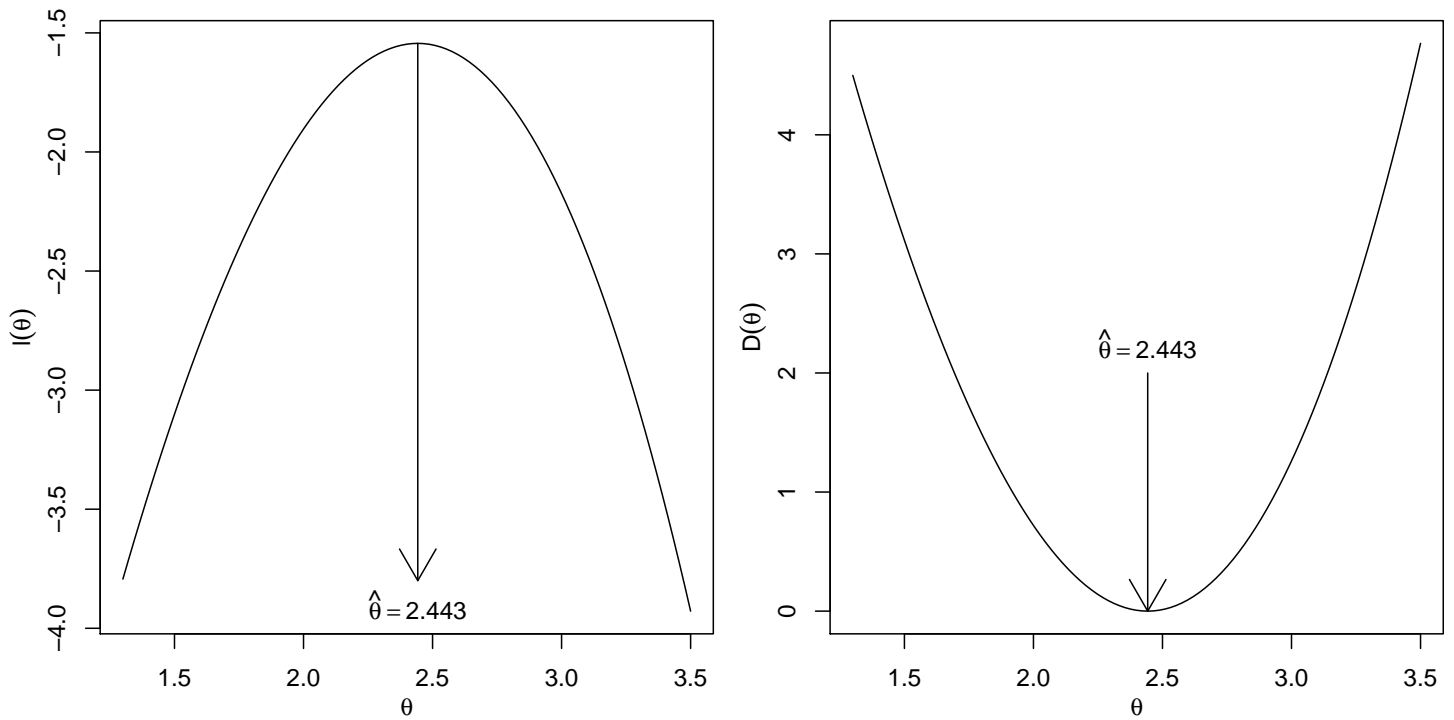


Figura 2: Função de verossimilhanças e função deviance.

```

> curve(l1.ex3, from=1.3, to=3.5, xlab=expression(theta),
+       ylab=expression(l(theta)))
> arrows(est[[1]], est[[2]], est[[1]], -3.8)
> text(est[[1]], -3.9, substitute(hat(theta)==a, list(a=round(est[[1]], dig=3))))
> #
> f <- function(x) -2*(l1.ex3(x) - est[[2]])
> curve(f, from=1.3, to=3.5,
+       xlab=expression(theta), ylab=expression(D(theta)))
> arrows(est[[1]], 2, est[[1]], 0)
> text(est[[1]], 2, substitute(hat(theta)==a, list(a=round(est[[1]], dig=3))), pos=3)

```