Wyara Vanesa Moura e Silva

Professora Dione Maria Valença

Inferência

Novembro, 2015

Questão 1 - Lista 3

Será realizado uma simulação para verificar a veracidade do teorema dado abaixo, o algoritmo será desenvolvido na linguagem do R.

Teorema: Sejam, $X_1, \cdots, X_n$ amostra aleatória de $X \sim f(x;\theta)$ com $\theta \in \Theta$. Seja θ_0 o verdadeiro valor de θ. Então (sob condições de regularidades) temos que:

$$\lim_{n\to\infty} P_{\theta_0}[L(\theta_0; X) > L(\theta; X)] = 1 \text{ para todo } \theta \in \Theta - \{\theta_0\}$$

Simulação: Os passos realizados estão com os comentários no script.

```
> # 1) Fixe  \theta_0 para algum f(x|\theta)
> theta0=2
>
> # 2) criar sequencia de valores para \theta
> theta=seq(1,5, 0.01)
>
> # 3) fixar n e gerar X de X~f(x|\theta)
> n=10
> x=rexp(n, theta0)
>
> # obter L(\theta, x) para da \theta seq
> # função de verossimilhança e log-verossimilhança
> lexp<-function(x, n, theta){(theta^n)*exp(-theta*sum(x))}
> loglexp<-function(x, n, theta){(n*log(theta))-(theta*sum(x))}
> l=loglexp(x, n, theta)
>
> # representar graficamente \theta x L(\theta, x) para \theta seq
> plot(theta, l, type="l", col=2)
>
```

```
> # função geral para gráfico da log verossimilhança
> par(mfrow=c(1,4))
> for (i in 1:4)
+ {
+   theta0=2
+   theta=seq(1,5, 0.01)
+   n=10^i
+   x=rexp(n, theta0)
+   loglexp<-function(x, n, theta){(n*log(theta))-(theta*sum(x))}
+   l=loglexp(x, n, theta)
+   plot(theta, l, type="l", ylab="log-verossimilhança", col=2,
+   main=paste("n= ", n))
+   abline(v =2, col = "gray60")
+ }
>
```

Os gráficos que encontram-se na Figura 1 são os retornos do código acima.

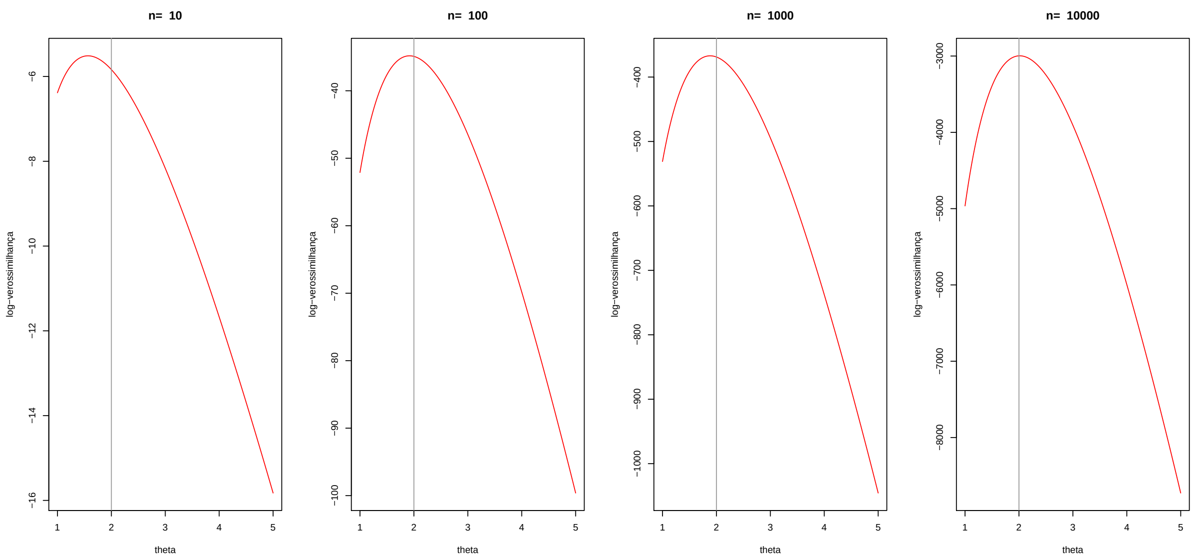

Figura 1: Gráficos da função de verossilhança em relação a uma sequência de $\theta's$ gerados

Como pode ser observado a partir dos gráficos acima, quando maior o tamanho da amostra (n), o máximo da função se aproxima cada vez mais do verdadeiro valor do parâmetro θ.

Agora, será desenvolvida um função que simula réplicas, para verificar a veracidade do teorema enunciado no início da questão.

```
> replicas<-function(n){
+ nrep=1000
+ theta0=2
+ x=NULL
+ a=0
+ b=0
+ for (i in 1:nrep)
+ {
+  theta=seq(1,10, 0.1)
+  x=rexp(n,theta0)
+  loglexp<-function(x, n, theta){(n*log(theta))-(theta*sum(x))}
+  l=loglexp(x, n, theta0)
+  l1=loglexp(x, n, sample(theta,1, replace = FALSE))
+  if (l>l1)
+   {a=a+1}
+     else {b=b+1}
+ }
+ prob=(1/nrep)*sum(a)
+ prob
+ }
>
> replicas(10)
[1] 0.877
> replicas(100)
[1] 0.967
> replicas(1000)
[1] 0.986
> replicas(10000)
[1] 0.992
>
```

Observando então, com os resultados obtidos no algoritmo da função `replicas` o que o teorema enuncia, ou seja, quando maior o tamanho da amostra (n), a densidade da amostra sob o verdadeiro valor do parâmetro θ_0, excede qualquer outra sob um valor de $\theta \neq \theta_0$, com alta probabilidade.