COMPORTAMENTO ASSITÓTICO DE ESTIMADORES DE MÁXIMA VEROSSIMILHANÇA

Felipe Matheus Gonçalves Costa (1); Divanilda Maia Esteves (2)

1 Universidade Estadual da Paraíba; felipematheusem@hotmail.com.br

2 Universidade Estadual da Paraíba; diana.maia@gmail.com

1. INTRODUÇÃO

Frequentemente, na estatística deseja-se estudar uma determinada característica de uma população. Tendo que, geralmente, é inviável o estudo sobre o grupo como um todo, considera-se apenas uma amostra do mesmo. Observa-se o que acontece nessa amostra e infere-se à população através de técnicas adequadas derivadas de dois ramos principais: a estimação de parâmetros e os testes de hipóteses.

Cada técnica tem sua importância na conquista da confiabilidade dos resultados obtidos. Muitos procedimentos na estimação de parâmetros dependem de um tamanho suficientemente grande da amostra, para que os estimadores tenham propriedades ótimas, dado que ele é um fator de grande influência no ajuste de um modelo probabilístico aos dados. Grande parte da modelagem estatística usa resultados assintóticos em amostras finitas, sendo assim deve-se saber qual o tamanho mínimo que uma amostra deve ter para que o uso de tais resultados seja possível.

Sendo assim, este artigo desenvolve-se no estudo do comportamento assintótico dos estimadores de máxima verossimilhança, simulando dados para devidas ilustrações.

2. METODOLOGIA

O método de máxima verossimilhança é uma das técnicas utilizadas na obtenção de estimadores paramétricos. Ele é bastante difundido por produzir equações convenientes do ponto de vista computacional.

Sejam X_1, ..., X_r variáveis aleatórias independentes e identicamente distribuídas, com função de densidade f (x; θ), $\theta \in \Theta \subset \mathbb{R}$. A função de verossimilhança definida como:

$$L(\theta) = \prod f(X_i;\theta)$$

considerada como função de θ. O estimador de máxima verossimilhança de θ é o valor que maximiza L(θ; x_1, ..., X – r).

O estimador de máxima verossimilhança pode ser encontrado, na maioria das vezes, através dos seguintes passos:

- Encontrar a função de verossimilhança;
- Aplicar a função ln;

- Derivar em relação ao parâmetro θ;
- Igualar o resultado a zero;
- Verificar que este estimador é o ponto de máximo.

O seguinte teorema indica condições sob as quais os estimadores de máxima verossimilhança possuem uma distribuição assintótica Normal.

Teorema: Sejam $X_1, \ldots, X_r$ variáveis aleatórias independentes e identicamente distribuídas com função de densidade f (x; θ), $\theta \in \Theta \subset \mathbb{R}$ satisfazendo as seguintes condições:

i) $\frac{\partial}{\partial\theta} f(x;\theta)$ e $\frac{\partial^2}{\partial\theta^2} f(x;\theta)$ existem em quase toda parte e são tais que $|\frac{\partial}{\partial\theta} f(x;\theta)| \leq H_1(x)$ e $|\frac{\partial^2}{\partial\theta^2} f(x;\theta)| \leq H_2(x)$ onde $\int_R H_j(x)dx < \infty, j = 1,2$.

ii) $\frac{\partial}{\partial\theta} logf(x;\theta)$ e $\frac{\partial^2}{\partial\theta^2} logf(x;\theta)$ existem em quase toda parte e são tal que:

a) $0 < I_\theta = E_\theta \left\{\frac{\partial}{\partial\theta} logf(X_1;\theta)\right\}^2 = \int_R \{\frac{\partial}{\partial\theta} f(x;\theta)\}^2 \{f(x;\theta)\}^{-1} dx < \infty$, ou seja, X_1 tem informação de Fisher finita;

b) $E_\theta \left\{\max_{h:|h|\leq\delta} \left|\frac{\partial^2}{\partial\theta^2} logf(X_1;\theta+h) - \frac{\partial^2}{\partial\theta^2} logf(X_1;\theta)\right|\right\} = \psi_\delta \to 0$ quando $\delta \to 0$.

Então o estimador de máxima verossimilhança de θ é tal que $\sqrt{n}(\hat{\theta}_r - \theta) \xrightarrow{D} N(0, I_\theta^{-1})$.

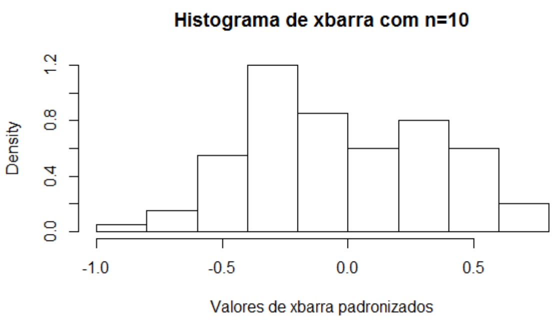
Para a visualização do comportamento dos estimadores de máxima verossimilhança, utilizou-se do método da simulação de dados no software R. Onde foram geradas amostras de diferentes tamanhos (10, 30, 100 e 500), das distribuições Poisson e Exponencial, onde para cada tamanho repetiu-se a amostragem 100 vezes.

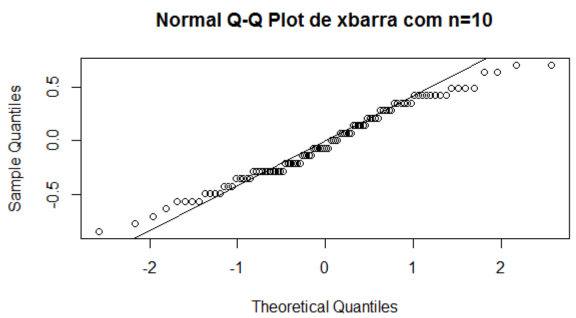
Sabendo que tanto no caso da Distribuição de Poisson, quanto na Distribuição Exponencial, o estimador de máxima verossimilhança de θ é x-barra, o calculamos para cada amostra obtida, e verificamos sua proximidade com a Distribuição Normal através de histogramas e de Q-Q Plots.

3. RESULTADOS E DISCUSSÃO

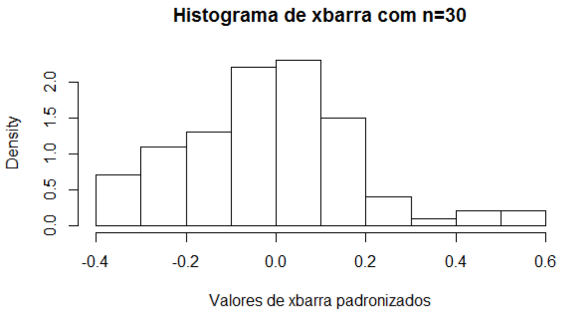
Primeiramente verificamos o comportamento de x-barra na Distribuição de Poisson:

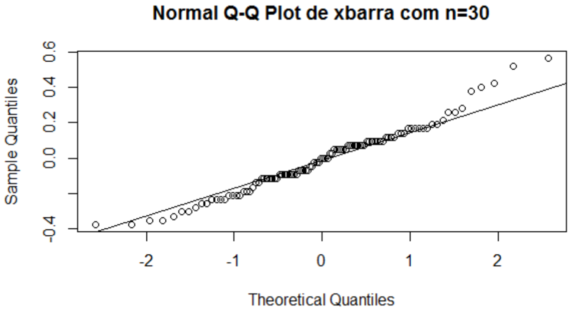
- Para n = 10:



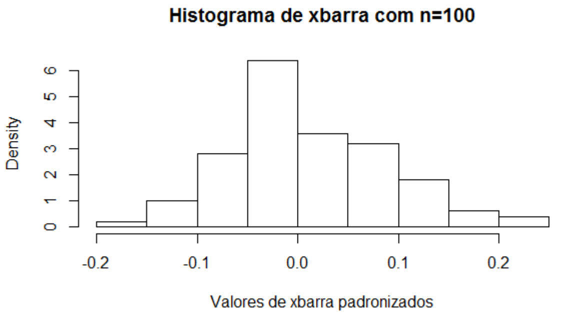


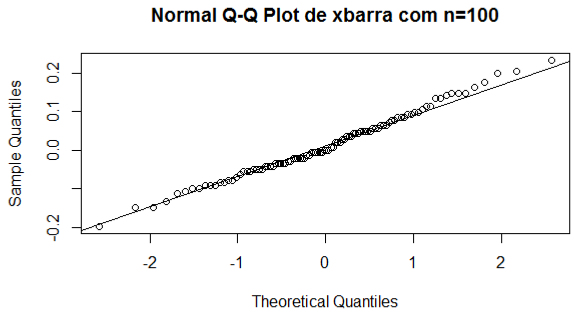
- Para n = 30:



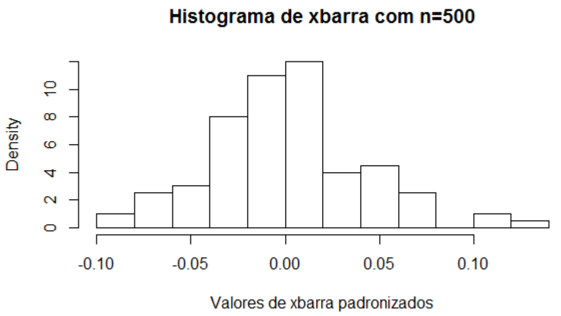


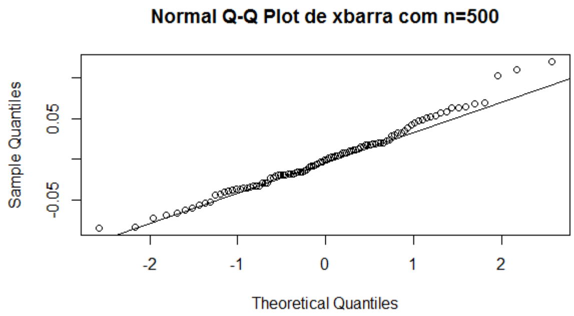
- Para n = 100:





- Para n = 500:

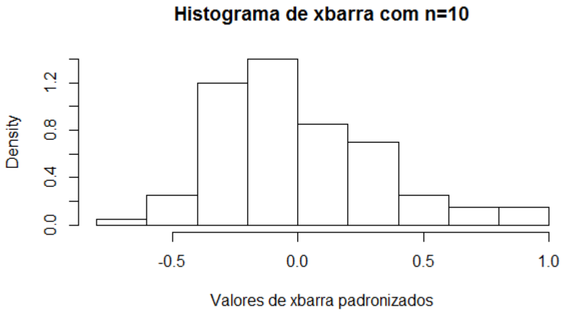


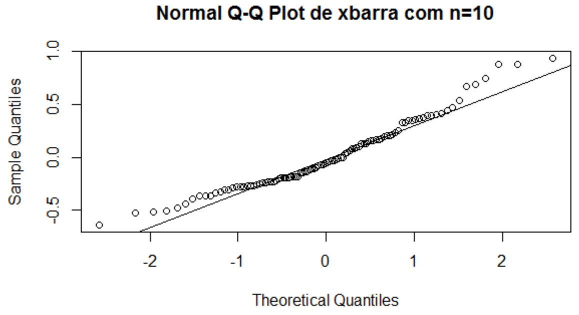


Notamos através do histograma que, à medida que o tamanho da amostra aumenta, o desvio padrão de x-barra padronizado diminui e que, pelo Q-Q Plot, cada vez mais, a distribuição de x-barra se aproxima de uma distribuição Normal.

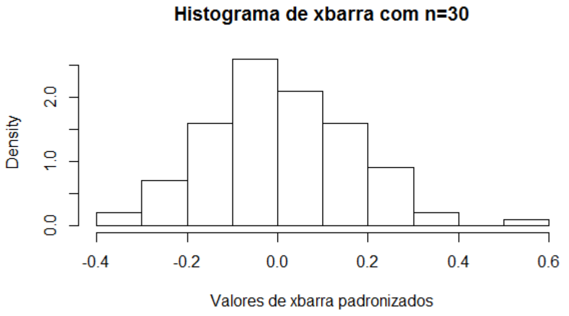
Agora, fazendo o mesmo para a distribuição Exponencial, temos:

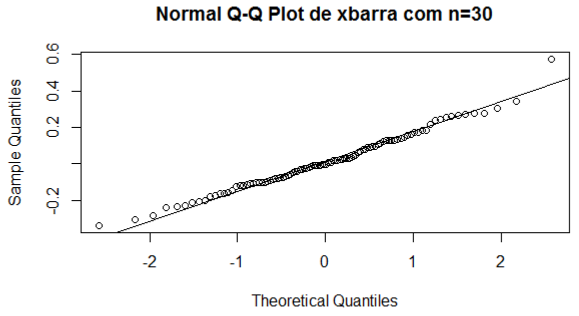
- Para n = 10:



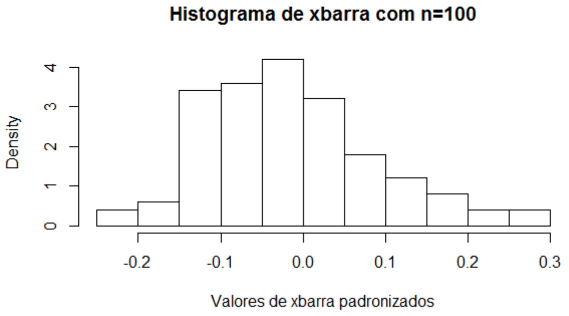


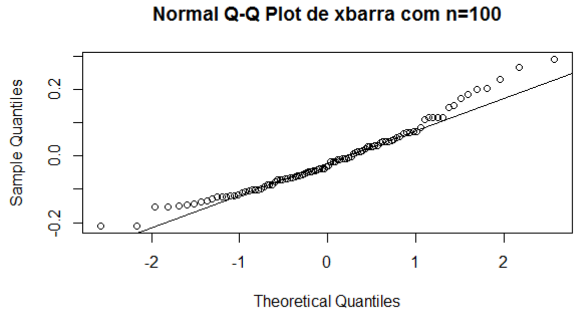
- Para n = 30:



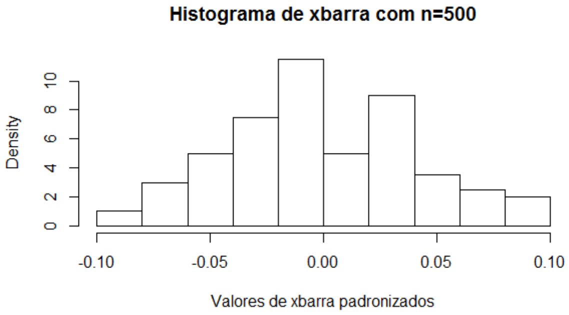


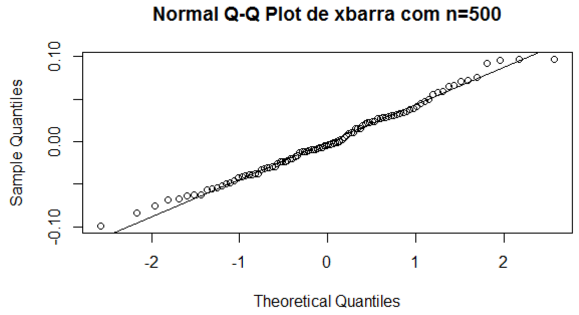
- Para n = 100:





- Para n = 500:





Observando os histogramas notou-se que, assim como na Distribuição de Poisson, na Exponencial o desvio padrão diminuiu quando se aumentou o tamanho da amostra. E da mesma forma, através dos Q-Q Plots, nota-se que a distribuição de x-barra se aproxima da Normal a medida que aumentamos o valor de n.

4. CONCLUSÕES

Através das simulações, percebemos que a distribuição dos estimadores de máxima verossimilhança nos dois casos considerados parece se aproximar de uma normal, o que concorda com os resultados assintóticos teóricos estudados.

REFERÊNCIAS BIBLIOGRÁFICAS

CORDEIRO, G. M. Introdução à Teoria Assintótica. In: 22º colóquio brasileiro de matemática. [S.l.: s.n.], 1999.

LEHMANN, E. L.; CASELLA, G. Theory of Point Estimation. 2ª ed. [S.l.]: Springer, 1998.

LEITE, J. G. ao; MOTTA SINGER, J. da. Métodos Assintóticos em Estatística. In: 9º simpósio brasileiro de probabilidade e estatística. [S.l.: s.n.], 1990.