Comparação de estimadores pelo método Monte Carlo das distribuições: $\mathcal{G}_{\mathrm{I}}^{0}$, Uniforme, Binomial negativa, Gamma e Exponencial

Antônio Marcos Larangeiras Alisson Nascimento Paulo Bruno Normande Juliana Leal

Universidade Federal de Alagoas
Instituto de Computação

Maceió-AL, Dezembro de 2013

Introdução

- Calcular o estimador por máxima verossimilhança;
- Calcular estimador pelo primeiro momento;
- Calcular estimador pelo segundo momento;
- Utilizar o Método Monte Carlo força bruta.

Distribuição Uniforme

Admitimos que em um dado problema, números reais cobrem de forma uniforme um segmento de reta $[a - -b]$ de tal maneira que quando se observa qualquer subintervalo contenha o mesmo número de pontos, e portanto , equiprovável.
Sua função de distribuição é dada por:

$$f(x) = \begin{cases} \frac{1}{a+b} & a \leqslant x \leqslant b \\ 0 & \text{caso contrário.} \end{cases} \tag{1}$$

Distribuição Uniforme

Aplicando a função 1 a uma variável aleatória $X \sim U(x; 0, \theta)$, com $\theta > 0$.
Sua função de distribuição é esta:

$$f(x) = \begin{cases} \frac{1}{\theta} & 0 \leqslant x \leqslant \theta \\ 0 & \text{caso contrário.} \end{cases}$$

Distribuição $U(x; 0, \theta)$

A esperânça matemática da distribuição $U(x; 0, \theta)$:

$$E(X) = \frac{\theta}{2} \tag{2}$$

Distribuição $U(x; 0, \theta)$

A esperânça matemática da distribuição $U(x; 0, \theta)$:

$$E(X) = \frac{\theta}{2} \tag{2}$$

A variância da distribuição $U(x; 0, \theta)$:

$$Var(X) = \frac{\theta^2}{12} \tag{3}$$

Estimadores da Distribuição $U(x; 0, \theta)$

Pela Máxima Veressimilhança:

$$\hat{\theta} = \frac{1}{\theta^n}\mathbb{1}_{\max\{X_1,\ldots,X_n\}\leqslant\theta}, \text{portanto } \hat{\theta} = \max\{X_1,\ldots,X_n\} \leqslant \theta \quad (4)$$

Estimadores da Distribuição $U(x; 0, \theta)$

Pela Máxima Veressimilhança:

$$\hat{\theta} = \frac{1}{\theta^n}\mathbb{1}_{\max\{X_1,\ldots,X_n\}\leqslant\theta}, \text{portanto } \hat{\theta} = \max\{X_1,\ldots,X_n\} \leqslant \theta \quad (4)$$

Pelo momento primeiro momento:

$$\hat{\theta}_2 = 2\overline{X} \quad (5)$$

Estimadores da Distribuição $U(x; 0, \theta)$

Pelo segundo momento:

$$\hat{\theta}_3 = \sqrt{\frac{3}{n}\sum_{i=1}^{n} X_i^2} \tag{6}$$

Gráficos da distribuição $U(x; 0, \theta)$

Distribuição $\mathcal{G}_{\mathrm{I}}^{0}$

Denotaremos esta distribuição por $Z \sim \mathcal{G}_{\mathrm{I}}^{0}(z; \alpha, 1, 1)$, com $-\alpha, z > 0$. Sua densidade é

$$f_Z = (z; \alpha, 1, 1) = \frac{\Gamma(L - \alpha)}{\Gamma(-\alpha)(1 + z)^{1-\alpha}}, \text{com} - \alpha \text{ e } z > 0.$$

Distribuição $\mathcal{G}_{\mathrm{I}}^0$

A esperança matemática da distribuição $\mathcal{G}_{\mathrm{I}}^0$:

$$E(Z) = \frac{\gamma}{\alpha - 1} \tag{7}$$

Distribuição $\mathcal{G}_{\mathrm{I}}^{0}$

A esperança matemática da distribuição $\mathcal{G}_{\mathrm{I}}^{0}$:

$$E(Z) = \frac{\gamma}{\alpha - 1} \tag{7}$$

A variância da distribuição $\mathcal{G}_{\mathrm{I}}^{0}$:

$$Var(Z) = \frac{1}{(\alpha-1)^2} + \frac{1}{(\alpha-2)(\alpha-1)^2} + \frac{1}{\alpha-1} \tag{8}$$

Estimadores da Distribuição $\mathcal{G}_{\mathrm{I}}^{0}$

Pela Máxima Veressimilhança:

$$\hat{\alpha} = \frac{n}{\displaystyle\sum_{i=1}^{n} \ln(1 + z_i)} \tag{9}$$

Estimadores da Distribuição $\mathcal{G}_I^0$

Pela Máxima Veressimilhança:

$$\hat{\alpha} = \frac{n}{\sum_{i=1}^{n} \ln(1 + z_i)} \tag{9}$$

Pelo momento central de ordem 2:

$$\hat{\alpha}_2 = \overline{Z} \tag{10}$$

Estimadores da Distribuição $\mathcal{G}_{\mathrm{I}}^{0}$

Pelo segundo momento:

$$\hat{\alpha}_3 = \frac{\frac{3}{k}\sum_{i=1}^{k} Z_i^2 + 2 + \sqrt{\frac{1}{k^2}\sum_{i=1}^{k} Z_i^4 + 4}}{\frac{2}{k}\sum_{i=1}^{k} Z_i^2} \tag{11}$$

Gráficos da distribuição $\mathcal{G}_{\mathrm{I}}^{0}$

Distribuição Binomial Negativa

Considere a situação de observar um evento dicotômico Y_i independentes e identicamente distribuídos segundo uma lei de Bernoulli de probabilidade p. Suponha que se registre Y, o número de ensaios até obter exatamente k sucessos.

Distribuição Binomial Negativa

Seja uma variável aleatória que fornece o numero de ensaios até o k-ésimo sucesso. Assim, Y tem uma distribuição binomial negativa com parâmetro $p \in (0,1)$, se sua função de probabilidade é dada por:

$$P_r(Y = y) = \begin{cases} \binom{y-1}{k-1} \cdot p^k \cdot (1-p)^{y-k} & \text{se y=k,k+1, } \ldots. \\ 0 & \text{, caso contrário} \end{cases}$$

Distribuição $BN(p, k)$

Usualmente sua função de probabilidade denota-se:
$Y \sim BN(p, k)$.

Distribuição $BN(p, k)$

Usualmente sua função de probabilidade denota-se:
$Y \sim BN(p, k)$.
A esperança matemática da distribuição binomial negativa:

$$E(X) = \frac{k}{p} \tag{12}$$

Distribuição $BN(p, k)$

Usualmente sua função de probabilidade denota-se:
$Y \sim BN(p, k)$.
A esperança matemática da distribuição binomial negativa:

$$E(X) = \frac{k}{p} \tag{12}$$

A variância da distribuição binomail negativa:

$$Var(X) = \frac{k(1-p)}{p^2} \tag{13}$$

Estimadores da Distribuição $BN(p, k)$

Pela Máxima Veressimilhança:

$$\hat{p} = \frac{nk}{\sum_{i=1}^{n} Y_i} \tag{14}$$

Estimadores da Distribuição $BN(p, k)$

Pela Máxima Veressimilhança:

$$\hat{p} = \frac{nk}{\sum_{i=1}^{n} Y_i} \tag{14}$$

Pelo momento central de ordem 2:

$$\hat{p}_2 = \frac{-k + \sqrt{k^2 + 4k\,Var(X_i)}}{2\,Var(X_i)} \tag{15}$$

Estimadores da Distribuição $BN(p, k)$

Pelo segundo momento:

$$\widehat{p}_3 = \frac{-k + \sqrt{k^2 + 4E(X_i^2)(k^2 + k)}}{2E(X_i^2)} \tag{16}$$

Gráficos da distribuição $BN(p, k)$

Distribuição $\Gamma(w; k, \theta)$

A Distribuição Gamma é caracterizada por dois valores, denominados *shape* (k) e *scale* (θ). Ela possui a seguinte função densidade probabilidade:

$$f(w; k, \theta) = \frac{w^{k-1} e^{-\frac{w}{\theta}}}{\theta^k \Gamma(w)}$$

para, $w > 0$ e $k, \theta > 0$

Distribuição $\Gamma(w; k, \theta)$

Esperança da distribuição $\Gamma(w; k, \theta)$:

$$E[W] = k\theta \tag{17}$$

Distribuição $\Gamma(w; k, \theta)$

Esperança da distribuição $\Gamma(w; k, \theta)$:

$$E[W] = k\theta \tag{17}$$

A variância da distribuição $\Gamma(w; k, \theta)$:

$$Var[W] = k\theta^2 \tag{18}$$

Em nosso estudo usaremos sempre $\theta = 1$ para simplicidade dos cálculos.

Estimadores da Distribuição $\Gamma(w; k, 1)$

Para estimar pela Máxima Verossimilhança é preciso encontrar o máximo da função log-verossimilhança de $\Gamma(w; k, 1)$

$$log(p(W|k,1)) = n(k-1)\overline{log(x)} - nlog(\Gamma(k)) - nklog(\overline{x}) + nklog(a) - nk$$

Que podemos resolver numericamente iterando sobre k em:

$$\frac{1}{k} = \frac{1}{k_0} + \frac{\overline{log(x)} - log(\overline{x}) + log(k_0) - \psi(k_0)}{k_0^2(\frac{1}{k_0} - \psi'(k_0))}$$

Estimadores da Distribuição $\Gamma(w; k, 1)$

Quando $k \approx k_0$ então encontramos o estimador $\hat{k}$

Como k inicial podemos usar a seguinte aproximação

$$\hat{k} = \frac{0.5}{log(\overline{x}) - \overline{log(x)}}$$

Estimadores da Distribuição $\Gamma(w; k, 1)$

Estimador pelo primeiro momento:

$$E[W] = k\theta$$

$$\hat{k}_1 = \frac{1}{n}\sum_{i=1}^{n} w_i$$

Estimadores da Distribuição $\Gamma(w; k, 1)$

Estimador pelo segundo momento central:

$$
\begin{aligned}
Var[W] &= k\theta^2 \\
\hat{k}_2^0 &= Var(w)
\end{aligned}
$$

Gráficos da distribuição $U(x; 0, \theta)$

Distribuição $BN(p,k)$

Table : Viés dos estimadores $\hat{p_1}$ e $\tilde{p_1}$.

Comparação dos viés dos Estimadores $\hat{p_1}$ e $\tilde{p_1}$				
n	p	$B(\hat{p_1})$	$B(\tilde{p_1})$	$\lvert B(\hat{p_1})\rvert > \lvert B(\tilde{p_1})\rvert$
50	0.1	0.01158347	0.01108683	TRUE
100	0.1	0.01147125	0.01124116	TRUE
150	0.1	0.01129354	0.01113929	TRUE
100000	0.1	0.01110255	0.01110307	FALSE
50	0.2	0.05069611	0.04952346	TRUE
100	0.2	0.05113692	0.05050091	TRUE
150	0.2	0.05021439	0.04981072	TRUE
100000	0.2	0.05000023	0.05000045	FALSE
50	0.3	0.13303390	0.13062580	TRUE
100	0.3	0.12839159	0.12709425	TRUE
150	0.3	0.12910398	0.12821978	TRUE
100000	0.3	0.12858837	0.12858547	TRUE

Distribuição $BN(p, k)$

Table : EQM dos estimadores $\hat{p_1}$ e $\tilde{p_1}$.

		Comparação dos EQM do Estimador $\hat{p_1}$ e $\tilde{p_1}$		
n	p	$EQM(\hat{p_1})$	$EQM(\tilde{p_1})$	$EQM(\hat{p_1}) > EQM(\tilde{p_1})$
50	0.1	0.0001897657	0.0001792158	TRUE
100	0.1	0.0001601663	0.0001552416	TRUE
150	0.1	0.0001461537	0.0001429903	TRUE
100000	0.1	0.0001232947	0.0001233068	FALSE
50	0.2	0.0028992563	0.0027791147	TRUE
100	0.2	0.0027753925	0.0027122546	TRUE
150	0.2	0.0026332882	0.0025933471	TRUE
100000	0.2	0.0025001601	0.0025001831	FALSE
50	0.3	0.0188113709	0.0181701887	TRUE
100	0.3	0.0170132255	0.0166791609	TRUE
150	0.3	0.0170230361	0.0167979328	TRUE
100000	0.3	0.0165355188	0.0165347806	TRUE

Distribuição $BN(p, k)$

Table : Comparação dos estimadores $\hat{p_2}$ e $\tilde{p_2}$.

Comparação dos Estimadores $\hat{p_2}$ e $\tilde{p_2}$			
n	p	$\|B(\hat{p_2})\| > \|B(\tilde{p_2})\|$	$EQM(\hat{p_2}) > EQM(\tilde{p_2})$
50	0.1	TRUE	FALSE
100	0.1	TRUE	FALSE
150	0.1	TRUE	FALSE
100000	0.1	TRUE	FALSE
50	0.2	TRUE	TRUE
100	0.2	TRUE	FALSE
150	0.2	FALSE	FALSE
100000	0.2	TRUE	FALSE
50	0.3	TRUE	FALSE
100	0.3	TRUE	FALSE
150	0.3	FALSE	FALSE
100000	0.3	FALSE	FALSE

Distribuição $BN(p, k)$

Table : Comparação dos estimadores $\hat{p_3}$ e $\tilde{p_3}$.

Comparação dos Estimadores $\hat{p_3}$ e $\tilde{p_3}$			
n	p	$\lvert B(\hat{p_3})\rvert > \lvert B(\tilde{p_3})\rvert$	$EQM(\hat{p_3}) > EQM(\tilde{p_3})$
50	0.1	TRUE	TRUE
100	0.1	TRUE	TRUE
150	0.1	TRUE	TRUE
100000	0.1	TRUE	TRUE
50	0.2	TRUE	TRUE
100	0.2	TRUE	TRUE
150	0.2	TRUE	TRUE
100000	0.2	TRUE	TRUE
50	0.3	TRUE	TRUE
100	0.3	TRUE	TRUE
150	0.3	TRUE	TRUE
100000	0.3	TRUE	TRUE

Conclusões

Referências Bibliográficas