Propriedades

Os experimentos estatístico que segue uma lei binomial negativa possuem as seguintes propriedades:

- cada tentativa pode resultar em dois resultados possíveis, o sucesso ou fracasso;
- a probabilidade de sucesso, denotada por p, é igual para cada tentativa;
- as tentativas são independentes;
- o experimento continua até que k sucessos sejam observados, onde k é especificado antecipadamente.

Definição

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

Definição

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

Definição

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

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

Usualmente utilizá-se a notação $X \sim BN(p,k)$.

Esperânça e Variância

A esperânça matemática da distribuição binomial negativa:

$$E(X) = \frac{k}{p} \quad (1)$$

Esperânça e Variância

A esperânça matemática da distribuição binomial negativa:

$$E(X) = \frac{k}{p} \tag{1}$$

A variância da distribuição binomail negativa:

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

Esperânça e Variância

A esperânça matemática da distribuição binomial negativa:

$$E(X) = \frac{k}{p} \tag{1}$$

A variância da distribuição binomail negativa:

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

Mais informações sobre a distribuição binomial negativa podem ser encontradas no livro de Magalhães [**?**].

Usando a Ferramenta *R*

- No *R*, ver [**?**], para uma distribuição binomial negativa, pode-se determinar a função de probabilidade acumulada, função de distribuição, função quartil e função de geração aleatória com parâmetros $size = k$ e $prob = p$.

Usando a Ferramenta R

- No R, ver [**?**], para uma distribuição binomial negativa, pode-se determinar a função de probabilidade acumulada, função de distribuição, função quartil e função de geração aleatória com parâmetros $size = k$ e $prob = p$.

A chamada da função de geração aleatória

rnbinom(n, size, prob)

- n - tamanho da amostra;
- $size$ - número de sucessos;
- $prob$ - probabilidade de sucesso.

Usando a Ferramenta *R*

- No *R*, ver [**?**], para uma distribuição binomial negativa, pode-se determinar a função de probabilidade acumulada, função de distribuição, função quartil e função de geração aleatória com parâmetros $size = k$ e $prob = p$.

A chamada da função de geração aleatória

rnbinom(n, size, prob)

- n - tamanho da amostra;
- *size* - número de sucessos;
- *prob* - probabilidade de sucesso.

Este comando retorna uma amostra de tamanho n, com as variáveis aleatórias $x - k$ de uma distribuição binomial negativa com probabilidade p.

Máxima Verossimilhança

$$\hat{p} = \frac{nk}{\sum_{i=1}^{n} X_i} \tag{3}$$

Estimador por Momentos

Outro estimador possível, é calculado pelo momento central de ordem 2. Calcula-se igualando a variância ao momento central de ordem dois, segue que

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

Estimador por Momentos

Um terceiro estimador pode ser encontrado utilizando a seguinte expressão: $Var(X_i) = E(X_i^2) - [E(X_i)]^2$ e substituindo a $Var(X_i)$ e a $E(X_i)$ por $\frac{k(1-p)}{p^2}$ e $\frac{k}{p}$, respectivamente, resulta:

$$\widehat{p}_3 = \frac{-k + \sqrt{k^2 + 4E(X_i^2)(k^2 + k)}}{2E(X_i^2)} \tag{5}$$

Tabela: 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

Tabela: 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

Tabela: Comparação dos estimadores $\hat{p}_2$ e $\tilde{p}_2$.

Comparação dos Estimadores $\hat{p}_2$ e $\tilde{p}_2$			
n	p	$\lvert B(\hat{p}_2)\rvert > \lvert B(\tilde{p}_2)\rvert$	$EQM(\hat{p}_2) > EQM(\tilde{p}_2)$
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

Tabela: 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

Estimador por máxima verossimilhança

1. 10 dentre os 12 viés do estimador por máxima verossimilhança foram reduzidos, em valor absoluto, com o uso da metodologia *bootstrap*,

Conclusões

Estimador por máxima verossimilhança

1. 10 dentre os 12 viés do estimador por máxima verossimilhança foram reduzidos, em valor absoluto, com o uso da metodologia *bootstrap*, o mesmo ocorreu com os EQM;

Conclusões

Estimador por máxima verossimilhança

1. 10 dentre os 12 viés do estimador por máxima verossimilhança foram reduzidos, em valor absoluto, com o uso da metodologia *bootstrap*, o mesmo ocorreu com os EQM;
2. os 2 viés e EQM que não melhoraram com *bootstrap* foram para uma amostra de cem mil (100000) variáveis aleatórias e $p \leq 0.2$;

Conclusões

Estimador por máxima verossimilhança

1. 10 dentre os 12 viés do estimador por máxima verossimilhança foram reduzidos, em valor absoluto, com o uso da metodologia *bootstrap*, o mesmo ocorreu com os EQM;
2. os 2 viés e EQM que não melhoraram com *bootstrap* foram para uma amostra de cem mil (100000) variáveis aleatórias e $p \leq 0.2$;
3. para uma amostra muito grande de variáveis aleatórias, obdecendo uma lei binomial negativa com o valor de p pequeno, a metodologia *bootstrap* não "melhora" o estimador, quando calculado por máxima verossimilhança.

Conclusões

- a aplicação de *bootstrap* para melhorar o estimador calculado utilizando o momento central amostral de ordem 2, $\hat{p}_2$, comportou-se de forma inconclusiva;

Conclusões

- a aplicação de *bootstrap* para melhorar o estimador calculado utilizando o momento central amostral de ordem 2, $\hat{p}_2$, comportou-se de forma inconclusiva;
  - melhorou 9 entre os 12 viés;
  - melhorou somente 1 EQM.

Conclusões

- a aplicação de *bootstrap* para melhorar o estimador calculado utilizando o momento central amostral de ordem 2, $\hat{p}_2$, comportou-se de forma inconclusiva;
    - melhorou 9 entre os 12 viés;
    - melhorou somente 1 EQM.
- o estimador por momento amostral de ordem 2, $\hat{p}_3$, conseguiu ganhos em seus viés e EQM em todos os casos estudados com o uso da metodologia *bootstrap*;

Conclusões

- a aplicação de *bootstrap* para melhorar o estimador calculado utilizando o momento central amostral de ordem 2, $\hat{p}_2$, comportou-se de forma inconclusiva;
    - melhorou 9 entre os 12 viés;
    - melhorou somente 1 EQM.
- o estimador por momento amostral de ordem 2, $\hat{p}_3$, conseguiu ganhos em seus viés e EQM em todos os casos estudados com o uso da metodologia *bootstrap*;
- a metodologia *bootstrap* aplicada para obter ganho de qualidade em estimadores de uma amostra de variáveis aleatória independente e identicamente distribuídas segundo uma lei binomial negativa, é eficaz nos seguintes casos:
    - para estimadores calculados por máxima verossimilhança, para amostra pequenas;
    - para estimadores calculados pelo momento amostral de ordem 2.

Referências Bibliográficas