

Compatibilidade das amostragens convencional e sequencial na estimação da proporção da distribuição binomial

Diullay C. V. Amaral¹, Carla R. G. Brighenti^{2†}

¹Graduanda em Zootecnia, UFSJ. E-mail: diullay.amaral@hotmail.com.

²Professora UFSJ/UFLA.

Resumo: *Objetivou-se verificar a compatibilidade entre as abordagens convencional e sequencial para a estimação do parâmetro proporção da distribuição Binomial. Foi desenvolvido no software R, scripts para realizar a simulação de dados e posterior estimação do parâmetro proporção sob as duas abordagens. Diferentes cenários de simulação foram avaliados a partir dos parâmetros (0,3; 0,4; 0,5; 0,6; 0,7; 0,8 e 0,9) e taxas de erro $\alpha = 5\%$ com $n=20, 50, 100$ e 200 para a abordagem convencional e $\alpha = 5\%$ e $\beta=5\%$ para a abordagem sequencial, totalizando 28 cenários. A amostragem convencional é caracterizada por necessitar de um número de amostras fixo determinado antes do processo de amostragem, sendo assim, seus resultados foram avaliados após o término de todas as amostras simuladas. Já no caso da amostragem sequencial, foi necessário utilizar um número variável de unidades amostrais, então, resultados parciais foram obtidos no decorrer da simulação da amostragem a fim de auxiliar na decisão de concluir o experimento. Para verificação de desvios na validação das abordagens, realizou-se avaliação de tendências, cálculo do erro percentual, amplitude das variâncias e cálculo da esperança matemática. Percebeu-se que há maior compatibilidade entre as estimativas convencional e sequencial quando os tamanhos de amostra são menores. As maiores porcentagens de erro entre abordagens foram obtidas nos cenários com valores de proporção próximos ao limite inferior do domínio de existência desta. Houve em média 51% de correlação entre as estimativas obtidas nos cenários avaliados.*

Palavras-chave: Bernoulli; Simulação; SPRT; Correlação.

Abstract: *The objective was to verify the compatibility between the conventional and sequential approaches for the estimation of the Binomial distribution ratio parameter. It was developed in the R software, scripts to perform the data simulation and subsequent estimation of the proportion parameter under the two approaches. Different simulation scenarios were evaluated from the parameters (0.3, 0.4, 0.5, 0.6, 0.7, 0.8 and 0.9) and $\alpha = 5\%$ error rates with $n = 20, 50, 100$ and 200 for the conventional approach and $\alpha = 5\%$ and $\beta = 5\%$ for the sequential approach, totaling 28 scenarios. Conventional sampling is characterized by requiring a fixed number of samples determined before the sampling process, thus, its results were evaluated after the completion of all simulated samples. Already in the case of sequential sampling, it was necessary to use a variable number of sample units, then partial results were obtained during the simulation of the sample in order to assist in the decision to conclude the experiment. To verify deviations in the validation of the approaches, the evaluation of trends, calculation of the percentage error, amplitude of the variances and calculation of the mathematical expectation were carried out. The highest error percentages between approaches were obtained in scenarios with proportion values close to the lower limit of the domain of existence of this one. There was an average of 51% correlation between the estimates.*

Keywords: Bernoulli; Simulation; SPRT; Correlation.

Introdução

A amostragem consiste em observar ou mensurar uma parte de uma população para obter estimativas representativas de todo o restante desta mesma população. É uma forma de estimar, com confiança, parâmetros populacionais sem precisar mensurar toda a população. A quantidade de objetos analisados é, em geral, muito pequena em relação ao tamanho da população que representa (Garthwaite et al, 1995).

†Autora correspondente: carlabrighenti@ufsj.edu.br.

Marotti et al. (2008) cita que uma das dificuldades nas etapas de uma pesquisa é calcular o tamanho de uma amostra e como fazer a amostragem de maneira correta a fim de que os resultados informem de maneira clara e precisa os objetivos da pesquisa. O número de unidades amostrais a serem mensuradas deve ser baseado em um limite de erro estabelecido pelo pesquisador, conforme um determinado nível de probabilidade, para a variável de interesse (no caso da distribuição binomial, a proporção). No caso da amostragem convencional, através de uma pré-amostragem, calcula-se o número de unidades amostrais necessárias para compor a amostra representativa da população dentro do limite de erro estabelecido. Assim, a amostragem convencional caracteriza-se, principalmente, por apresentar um número fixo de unidades amostrais, determinado antes de se iniciar o procedimento de amostragem, e os resultados são avaliados após o seu término.

Já a amostragem sequencial é um método estatístico caracterizado por utilizar um número variável de observações na amostra, ou seja, o tamanho amostral é uma variável aleatória. A amostragem sequencial é um procedimento no qual observações são tiradas em sequência e as decisões são feitas após a leitura de cada unidade amostral (LIU et al., 2009). Assim, de acordo com Wald (1947), o momento de terminar a amostragem e tomar uma decisão depende, em cada etapa, dos resultados obtidos a cada momento. Barbosa (1992) cita que a amostragem sequencial caracteriza-se por utilizar amostras de tamanho variável, ao contrário da amostragem convencional, que utiliza amostras de tamanho fixo. A razão principal do crescente interesse pela amostragem sequencial é que a avaliação pode classificar rapidamente a população quando há consistência de observações (COCCO et al., 2015).

As principais vantagens são: confiabilidade, facilidade, e principalmente rapidez na amostragem. Principalmente, no caso de lotes muito bons ou muito ruins, a decisão pode ser tomada com base em tamanho de amostra menor que o clássico. Liu e colaboradores (2009) cita que a amostragem sequencial se caracteriza por utilizar amostras de tamanho variável, ao contrário da amostragem convencional, que utiliza amostras de tamanho fixo.

Quando se deseja comparar a eficiência ou compatibilidade de diferentes métodos, pode-se utilizar de modelos de simulação. O objetivo é utilizá-los como suporte ao processo de tomada de decisão. No entanto, para que os resultados dessas experimentações possam realmente ser utilizados, o modelo de simulação empregado deve passar pelos procedimentos de verificação e validação (STRAPASSON et al., 2009).

A distribuição binomial é a distribuição de probabilidade discreta do número de sucessos numa sequência de tentativas nas quais se tem exclusivamente como resultado duas possibilidades, sucesso ou fracasso, onde cada tentativa é independente das demais e a probabilidade de sucesso a cada tentativa permanece constante independente das outras. Então, a variável de interesse ou pretendida, é a proporção p , correspondente ao número de sucessos p nas n tentativas. Assim, o objetivo deste trabalho foi verificar a compatibilidade entre as amostragens convencional e sequencial em diferentes cenários de simulação de obtenção da população binomial.

Metodologia

Inicialmente foram desenvolvidos scripts para gerar amostras da distribuição Bernoulli com diferentes parâmetros (0,3; 0,4; 0,5; 0,6; 0,7; 0,8 e 0,9). A partir de cada um destes cenários, retiradas amostras considerando as abordagens convencional e sequencial. Para a abordagem convencional foi considerada a simulação de amostras a partir de uma população binomial com parâmetros n e p , em que n refere-se ao tamanho amostral e p a probabilidade de sucesso do evento Bernoulli de interesse. Utilizou-se $n = 20, 50, 100$ ou 200 .

$$X \sim \text{Bin}(n, p) : f(X|p) = f_x(X) = \binom{n}{x} p^x (1-p)^{n-x} \quad 0 \leq p \leq 1$$

Sigmae, Alfenas, v.8, n.2, p. 315-322, 2019.

64ª Reunião da Região Brasileira da Sociedade Internacional de Biometria (RBRAS).

18º Simpósio de Estatística Aplicada à Experimentação Agrônômica (SEAGRO).

Considerou-se o estimador pontual de máxima verossimilhança do parâmetro p dado por $\hat{p} = \frac{x}{n}$. Em que x é o número de sucessos observados. Para o erro α adotou-se o nível de 5%, não havendo valor estabelecido para o erro tipo II (β). Registrou-se para cada cenário, com base nas 2000 simulações obtidas em cada caso, o coeficiente de variação (CV) e os respectivos intervalos de confiança para a proporção a 95%.

No caso da amostragem sequencial, a sua essência é o teste de razão de probabilidades de Wald (1947), em que as seguintes probabilidades foram calculadas como:

$p_0(n, x)$ = a probabilidade de que ocorra a quantidade x de sucesso na amostra, se $p = p_0$ (aceitável).

$p_1(n, x)$ = a probabilidade de que ocorra a quantidade x de sucesso, se $p = p_1$ (não aceitável).

O teste de razão de probabilidades examina a razão: $\frac{p_1(n, x)}{p_0(n, x)}$.

Registrando-se o tamanho da amostra quando era satisfeita a condição $\frac{p_1(n, x)}{p_0(n, x)} \leq \frac{\beta}{1-\alpha}$ ou

$\frac{p_1(n, x)}{p_0(n, x)} \geq \frac{1-\beta}{\alpha}$. Caso contrário, uma nova amostra da distribuição de Bernoulli (p_0) era obtida.

O sistema de decisão do plano de amostragem é então definido em três regiões: aceitação de H_0 , rejeição e continuação do teste. Para obtenção desse plano de avaliação é necessária a construção das retas de aceitação $A_N = c + s * x$ e rejeição: $R_N = -c + s * x$ sendo c os coeficientes lineares calculados conforme Shirose & Mori (1996) e s o coeficiente angular, dados por:

$$C = \frac{\log\left(\frac{\beta}{1-\alpha}\right)}{\log\left(\frac{p_{1*}(1-p_0)}{p_{0*}(1-p_1)}\right)} \text{ e } S = \frac{\log\left(\frac{1-p_0}{1-p_1}\right)}{\log\left(\frac{p_{1*}(1-p_0)}{p_{0*}(1-p_1)}\right)}$$

em que p é o parâmetro de proporção adotado da distribuição Binomial, α e β são os erros. E x a soma de sucessos de uma distribuição Binomial. Para o erro α adotou-se o nível de significância de 5% e para β adotou-se 5%. Para p_1 adotou-se, em relação ao p_0 , os valores 0,1; 0,2; 0,2; 0,2; 0,4; 0,6 e 0,7 respectivamente.

Na abordagem sequencial foram registrados, além da estimativa pontual, o intervalo de confiança a 95%, o tamanho médio amostral obtido com respectivo desvio padrão e o coeficiente de variação das 2000 simulações de cada cenário.

A compatibilidade entre as diferentes abordagens foi realizada pelo emprego de procedimentos estatísticos tais como cálculo das medidas de posição e dispersão, coeficiente de variação, amplitude e total, coeficiente de correlação. Para verificação de desvios na validação das abordagens, avaliou ainda, conforme sugerido por Silva (2002), as seguintes características:

- *Cálculo da esperança matemática*: No caso da amostragem convencional (fixa), a expressão que determina o número de unidades de amostras necessário para atingir determinado nível de precisão, a dado nível de probabilidade, é dada por $s^2 = \frac{pq}{e^2}$ e para o caso sequencial é dada por:

Sigmae, Alfenas, v.8, n.2, p. 315-322, 2019.

64ª Reunião da Região Brasileira da Sociedade Internacional de Biometria (RBRAS).

18º Simpósio de Estatística Aplicada à Experimentação Agrônômica (SEAGRO).

$$E(N_0) = \frac{\left(\frac{(1-\alpha) * (\log \beta)}{(1-\alpha)} \right)}{\left(p_1 * \log \left(\frac{p_1}{p_0} \right) \right)} + \frac{\left(\frac{(1-\beta) * \log(1-\beta)}{\alpha} \right)}{\left((1-p_1) * \log \left(\frac{1-p_1}{1-p_0} \right) \right)}$$

- *Cálculo do erro percentual:* $Erro = \frac{|\bar{X}_1 - \bar{X}_2|}{X_2} * 100$

em que:

\bar{X}_1 = é a média dos dados simulados da abordagem 1; e

\bar{X}_2 = é a média dos dados da abordagem 2.

Da mesma forma foi avaliada a amplitude das variâncias.

- *Avaliação de tendências:* Consiste em proceder à obtenção do coeficiente de correlação tendo por dados os gerados pelo modelo e os coletados do sistema real;

- *Teste de proporção:* Construiu-se o intervalo de confiança para diferença de proporção, verificando-se a aceitação ou rejeição de compatibilidade entre as estimativas obtidas. Na inferência com proporções é comum construir-se intervalos de confiança, seja no caso de uma proporção seja para a diferença de duas proporções. Em geral, as proporções estão relacionadas com a distribuição binomial de parâmetros n e p . O método de estimação mais conhecido é o de Wald, que se baseia na aproximação assintótica normal do estimador da proporção (Carari et al., 2010).

Depois de calculados os limites de confiança das medidas de posição, foram testados para cada nível de probabilidade. Se os valores máximos e mínimos obtidos estavam contidos no intervalo de confiança, os procedimentos testados eram aceitos, caso contrário é rejeitado. Todos os procedimentos de simulação e validação proposto neste trabalho foram registrados simultaneamente em script elaborado no software R (R Core Team, 2018).

Resultado e Discussão

A partir das simulações realizadas obtiveram-se os resultados em relação às estimativas do parâmetro proporção p nos diferentes cenários. A tabela 1 apresenta o resultado considerando diferentes tamanhos amostrais na abordagem convencional e sequencial.

A amostragem sequencial é um modelo que não precisa de um número fixo de amostras pré-estabelecidas antes do experimento. Dessa forma, a decisão de terminar a simulação e tomar uma decisão no decorrer da amostragem, vai depender dos resultados de cada etapa, levando em consideração os conjuntos de dados parciais obtidos na simulação. O que não acontece com os métodos de amostragem não sequenciais, por exemplo, na amostragem convencional.

Tabela 1 – Parâmetros da distribuição binomial (p) considerados para as hipóteses nula (p_0) e alternativa (p_1), fixado o erro $\alpha = 0,05$, com respectivo tamanho amostral do teste convencional e médias com desvio padrão obtido no teste sequencial, em 2000 simulações de cada cenário.

P		Convencional		Sequencial			
p_0	n	$\hat{P}_{convencional}$	CV (%)	p_1	$\hat{P}_{sequencial}$	CV (%)	n_{seq}
0,3	20	0,2973 (0,097; 0,498)	34,07	0,2	0,3351 (0,242; 0,428)	22,95	98,25 (72,74)
	50	0,3002 (0,173; 0,427)	21,49		0,3348 (0,242; 0,428)	23,60	99,14 (72,14)
	100	0,3015 (0,212; 0,391)	15,02		0,3345 (0,241; 0,428)	23,71	98,18 (71,12)
	200	0,2992 (0,236; 0,363)	10,66		0,3342 (0,242; 0,426)	23,79	101,26 (73,25)
0,4	20	0,3988 (0,184; 0,613)	27,53	0,2	0,4602 (0,279; 0,641)	31,03	29,16 (21,78)
	50	0,4006 (0,265; 0,536)	17,00		0,4630 (0,280; 0,646)	29,31	28,41 (20,53)
	100	0,3991 (0,303; 0,495)	12,18		0,4651 (0,284; 0,647)	30,64	29,04 (22,06)
	200	0,3995 (0,332; 0,467)	8,69		0,4608 (0,277; 0,644)	31,21	28,35 (20,62)
0,5	20	0,4997 (0,281; 0,719)	22,41	0,1	0,6382 (0,269; 1,008)	39,47	6,49 (4,78)
	50	0,4978 (0,359; 0,636)	14,22		0,6308 (0,261; 1,000)	41,04	6,55 (4,75)
	100	0,5010 (0,403; 0,599)	10,14		0,6368 (0,268; 1,005)	39,40	6,54 (4,83)
	200	0,5012 (0,432; 0,570)	7,10		0,6275 (0,261; 0,994)	39,36	6,68 (4,90)
0,6	20	0,6001 (0,385; 0,815)	18,86	0,4	0,6485 (0,494; 0,803)	19,29	36,55 (25,49)
	50	0,5969 (0,461; 0,733)	11,81		0,6462 (0,494; 0,799)	18,31	37,75 (26,42)
	100	0,6012 (0,505; 0,697)	8,25		0,6492 (0,494; 0,804)	18,53	36,33 (25,35)
	200	0,5990 (0,531; 0,667)	5,79		0,6472 (0,495; 0,800)	18,22	37,73 (26,09)
0,7	20	0,7004 (0,500; 0,901)	14,48	0,6	0,7243 (0,647; 0,801)	8,37	129,75 (86,99)
	50	0,6998 (0,573; 0,827)	9,43		0,7243 (0,647; 0,802)	8,79	127,35 (88,70)
	100	0,7000 (0,610; 0,790)	6,50		0,7254 (0,647; 0,803)	9,07	125,64 (89,18)
	200	0,7006 (0,637; 0,764)	4,67		0,7255 (0,647; 0,804)	9,08	125,20 (89,48)
0,8	20	0,7971 (0,621; 0,973)	11,13	0,2	0,8600 (0,552; 1,168)	19,69	4,88 (2,68)
	50	0,8017 (0,691; 0,912)	7,02		0,8659 (0,562; 1,17)	18,93	4,83 (2,72)
	100	0,7987 (0,720; 0,877)	4,96		0,8692 (0,570; 1,169)	17,77	4,87 (2,70)
	200	0,8002 (0,745; 0,856)	3,47		0,8666 (0,565; 1,168)	18,30	4,87 (2,71)
0,9	20	0,9007 (0,770; 1,032)	7,43	0,7	0,9225 (0,817; 1,028)	8,68	24,85 (15,37)
	50	0,8987 (0,815; 0,982)	4,70		0,9179 (0,810; 1,026)	9,52	24,95 (15,32)
	100	0,9013 (0,843; 0,960)	3,33		0,9239 (0,818; 1,030)	8,68	24,05 (14,61)
	200	0,8998 (0,858; 0,941)	2,39		0,9182 (0,810; 1,026)	9,97	24,77 (15,44)

De acordo com Fernandes et al, (2002), a amostragem sequencial tem se mostrado mais rápida e confiável que a amostragem convencional e, de acordo com Barbosa (1992), caracteriza-se por utilizar número variável de unidades amostrais, ao contrário da amostragem convencional que utiliza número fixo de unidades amostrais. Na amostragem sequencial, a avaliação pode classificar rapidamente a população do organismo tratado em grandes categorias como baixa, média ou alta, ou ainda indicar se é necessário adotar alguma medida de controle ou não. O tamanho da amostra não é constante, dependendo do tamanho da população (Kogan & Herzog, 1980).

Assim, a decisão de terminar a amostragem e tomar uma decisão depende, em cada etapa, dos resultados obtidos até então e, por este motivo, a amostragem sequencial requer, em média, amostras com um terço, ou menos, do tamanho que seria utilizado com a amostragem de tamanho fixo (Wald,

1947). Tal fato pode ser observado na Tabela 2, em que o tamanho esperado em alguns casos, chega a ser apenas 10% do tamanho na amostragem convencional.

Tabela 2 – Tamanho de amostra em função da magnitude do erro amostral igual a 5%.

Erros	Parâmetro (p)	0,3	0,4	0,5	0,6	0,7	0,8	0,9
$\alpha = 5\%$ $\beta = 5\%$	Convencional (n)	322,69	368,80	384,17	368,80	322,70	245,87	138,29
	Sequencial (En)	101,12	27,21	5,57	35,12	131,87	3,42	24,48

Em relação à comparação dos resultados obtidos nas duas abordagens, foram obtidos os erros, coeficiente de correlação e intervalo de confiança para diferença de proporção (Tabela 3).

Tabela 3 – Comparações entre estimativas clássica e sequencial resultantes de estimação para erros $\alpha=0,05$, em 2000 simulações.

p_0	p_1	$n_{\text{convencional}}$	Erro (%)	Correlação	IC (dif)
0,3	0,2	20	11,28	0,672	-0,043; -0,032
		50	10,33	0,7748	-0,039; -0,030
		100	9,86	0,6536	-0,037; -0,029
		200	10,47	0,5071	-0,038; -0,031
0,4	0,2	20	13,34	0,7781	-0,069; -0,053
		50	13,47	0,5681	-0,069; -0,0557
		100	14,19	0,3887	-0,072; -0,059
		200	13,30	0,3434	-0,067; -0,0547
0,5	0,1	20	21,62	0,4408	-0,150; -0,126
		50	21,31	0,2806	-0,146; -0,124
		100	21,58	0,2078	-0,149; -0,126
		200	20,83	0,1519	-0,142; -0,120
0,6	0,4	20	7,46	0,8247	-0,055; -0,041
		50	7,62	0,6498	-0,055; -0,043
		100	7,39	0,4781	-0,053; -0,042
		200	7,67	0,3455	-0,055; 0,0441
0,7	0,6	20	3,29	0,5547	-0,029; -0,018
		50	3,38	0,7794	-0,028; -0,020
		100	3,50	0,7429	-0,028; -0,021
		200	3,43	0,5877	-0,028; -0,021
0,8	0,2	20	7,31	0,4174	-0,071; -0,054
		50	7,41	0,2642	-0,071; -0,056
		100	8,11	0,201	-0,077; -0,06
		200	7,66	0,1834	-0,073; -0,059
0,9	0,7	20	2,36	0,7479	-0,026; -0,017
		50	2,09	0,528	-0,023; -0,014
		100	2,44	0,3968	-0,026; -0,018
		200	2,56	0,2476	-0,027; -0,019

Os maiores tamanhos de amostra forneceram as menores correlações entre as estimativas, principalmente quando houve maior diferença entre os valores de p_0 e p_1 adotados (Tabela 2). Dos oito cenários de proporções, em cinco deles houve um comportamento decrescente da correlação com o aumento do tamanho da amostra, indicando que há maior compatibilidade entre as estimativas clássica e sequencial quando os tamanhos de amostra são menores. No entanto, para os três casos em que tal fato não foi observado, percebem-se que a diferença entre os parâmetros p_0 e p_1 são iguais ou inferiores a 0,1, indicando que este parâmetro utilizado na simulação, teve maior influência nos resultados do que o tamanho amostral.

Deve-se ressaltar ainda, que, apesar de não ocorrer tal redução na correlação, houve certa estabilização na correlação em torno de 60%, sendo este valor maior do que a média geral dos 28 cenários que foi de 51%. As maiores porcentagens de erro foram obtidas nos cenários de simulação com valores de proporção próximos ao limite inferior do domínio de existência da proporção. Todos os valores-p foram significativos ($p < 0,001$), indicando diferenças entre as estimativas.

Conclusão

A aplicação da amostragem sequencial permitiu diminuição, em alguns casos, no tamanho da amostra quando comparada a amostragem convencional.

Houve compatibilidade entre as amostragens sequencial e convencional em 51% dos casos simulados, considerando a distribuição binomial.

Agradecimentos

Agradecimentos especiais à FAPEMIG, à CNPq e a Universidade Federal de São João Del rei.

Referências Bibliográficas

BARBOSA, J.C. A amostragem seqüencial, In: FERNANDES, O.A.; CORREIA, A.C.B.; de BORTOLI, S.A.(ed.), *Manejo integrado de pragas e nematóides*. Jaboticabal: FUNEP,1992. p. 205-211.

CARARI, M. L.; LIMA, P. C.; FERREIRA, D. F.; CIRILLO, M. A. Estimate of differences among two binomial proportions through bootstrap. *Revista Brasileira de Biometria*, v. 28, n.3, p.112-134, 2010.

COCCO, A.; SERRA, G.; LENTINI, A.; DELIPERI, S.; DELRIO, G. Spatial distribution and sequential sampling plans for *Tuta absoluta* (Lepidoptera: Gelechiidae) in greenhouse tomato crops. *Pest Management Science*, v. 71, p.1311-1323, 2015.

FERNANDES, M.; BUSOLI, A. C.; BARBOSA, J. C. Amostragem sequencial de *Spodoptera frugiperda* (J. E. Smith, 1797) (Lepidoptera, Noctuidae) em algodoeiro. *Revista Brasileira Agrociência*, v. 8, n. 3, p. 213-218, 2002.

GARTHWAITE, P. H.; JOLLIFFE, I.; BYRON, J. *Statistical inference*. 2.ed. London: Prentice Hall, 1995. p.114-131.

KOGAN, M.; HERZOG, D.C. Sequential sampling. In:KOGAN, M. & HERZOG, D.C. *Sampling methods in soybean entomology*. New York: Springer-Verlag, 1980, p.65-97.

Sigmae, Alfenas, v.8, n.2, p. 315-322, 2019.

64ª Reunião da Região Brasileira da Sociedade Internacional de Biometria (RBRAS).

18º Simpósio de Estatística Aplicada à Experimentação Agrônômica (SEAGRO).

LIU, Y.; CHEN, Y.; CHENG, J. A comparative study of optimization methods and conventional methods for sampling design in fishery-independent surveys. *ICES Journal of Marine Science*, storebo, v. 66, p. 1873-1882, 2009.

MAROTTI, J.; GALHARDO, A. P..M.; FURUYAMA, R. J.; PIGOZZ, M. N.; CAMPOS, T. N.; LAGANÁ, D. C. Amostragem em pesquisa clínica: tamanho da amostra. *Revista de Odontologia da Universidade Cidade de São Paulo*, v. 20, n.2, p. 186-194, 2008.

R CORE TEAM. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2018. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.

SILVA, L C. Stochastic simulation of the dynamic behavior of grain storage facility. *Tese* (Doutorado em Engenharia Agrícola.) Universidade Federal de Viçosa. Viçosa: MG.2002.

STRAPASSON, E.; COLOSIMO, E.A; GIOLO, S.R. A simulation study to compare imputation methods to handle grouped data survival, *Revista Brasileira de Biometria*, 2009.

WALD, A. *Sequential analysis*. New York: J. Wiley & Sons, Inc., 1947. 212p.

Sigmae, Alfenas, v.8, n.2, p. 315-322, 2019.

64^a Reunião da Região Brasileira da Sociedade Internacional de Biometria (RBRAS).
18^o Simpósio de Estatística Aplicada à Experimentação Agrônômica (SEAGRO).