

Resposta aleatória: uma técnica de pesquisa para eliminar o viés de resposta evasiva Autor(es): Stanley L. Warner Trabalho(s) revisado(is): Fonte:

Jornal do Associação Estatística Americana, Vol. 60, nº 309 (mar., 1965), pp. 63-

69

Publicado por: Associação Estatística Americana URL

estável: http://www.jstor.org/stable/2283137 .

Acessado: 11/12/2012 12:32

Seu uso do arquivo JSTOR indica sua aceitação dos Termos e Condições de Uso, disponíveis em http://www.jstor.org/page/info/about/policies/terms.jsp

JSTOR é um serviço sem fins lucrativos que ajuda acadêmicos, pesquisadores e estudantes a descobrir, usar e desenvolver uma ampla variedade de conteúdo em um arquivo digital confiável. Usamos tecnologia da informação e ferramentas para aumentar a produtividade e facilitar novas formas de bolsa de estudos. Para obter mais informações sobre o JSTOR, entre em contato com support@jstor.org.



A American Statistical Association está colaborando com a JSTOR para digitalizar, preservar e ampliar o acesso ao Journal of the American Statistical Association.

RANDOMIZED RESPONSE: A SURVEY TECHNIQUE FOR ELIMINATING EVASIVE ANSWER BIAS

RESPOSTA RANDOMIZADA: UMA TÉCNICA DE PESQUISA PARA ELIMINAR O VIES DE RESPOSTA EVASIVA Claremont Graduate School

For various reasons individuals in a sample survey may prefer not to confide to the afterviewer fire to fire the survey may prefer not to confide to the afterviewer fire to fire the survey may prefer not to confide to the afterviewer fire to fire the survey at all or to reply with the first of the first

sobre a população subjaçonto DUCTION

For reasons of modesty, fear of being thought bigoted, or merely a reluctance to confide secrets to strangers, many individuals attempt to evade certain questions put to them by interviewers. In survey vernacular, these people increase and considerate and the second people in the putgishtastquequeuryayedtasrpenosentievastadores: weved hauti purposeduirs oxidina prosta a severato de prosta de la como de la 33-61, 65, 465, 267-50, that our revealing. Innocuous questions ordinarily receive good response units questions requiring nervotal sa contraversial expertions excite resistance when resistance is engountered the usual modification of the survey methodas gim plygen added effert es the part of the sate provent resistence confdence of the interviewee of here is mooneyer a neature histograf of the general índividual to confide certain things to any one met alone partranger and there is also a natural reluctance to have confidential statements on a paper containar ing his pame and address. For some guestions at least aprehably only limited a gains are possible through triving to be smade the interview set hat he surronders little by confiding to the interview of ovavelmente apenas ganhos limitados são possive paper suggests an alternate method for increasing concention. The method is pull on the premise that cooperation should be naturally better if the questions allow answers which reveal less even to the interviewer foscertions ruthe controls in a dream the device at the operation of the second of the control of the macuous the interviewes posted de vite answers the taturnish information enly on a probability basis of carporally one application distributed and carporally one interviewee's contractions have a developed a series of the contraction of the contractio 1. In this case, even the interviewer would know only the probability that the base has probabilitied. Por exemple, uma aplicação pode envolver o entrevistado apenas given answerowas true venasoelish as this twice of any probablists deviseling then The same was a required to be truthful with perabability of pithis suggested that resposta dada fosse verdadeira. Na medida em quenesse tipo de resposta é menos revelador do que uma resposta exigida para ser verdadeira com probabilidade 1, sugere-se que esse

type of approach may encourage greater cooperation for certain survey probfems. As another more related about a few and of the randomarced response method. the following section outlines a marticular model for estimating a population presporta o combo un sulting catimate a cama taga compose de vita can areationa la estica o mater inder agricus assumptions about the cooperation of athose interviewed. estimativas resultantes são então comparadas com estimativas convencionais sob várias suposições sobre a cooperação dos entrevistados.

Suppose that every person in a population belongs to either Group A or Group B and it is recorded to restimate by servey the proportion belonging to Group An Assimple random sample of a people is drawn with replacement from the procession earth provisions made forceach persone to be enterviewed and serious the ainterviewes each jessevieven inducein hed ovide an identical enjoyes eviteurs fana markeelsoatbata tee emieveranoi atrato deselattevistas viiteanimbeviilitus menede tort bei letteru British annot havilite karceda Teronia each isternie exothe batery iewea isoaskeobatonianie tar enimperialno beenvedo baptika de tervienra open ennethorusta. w bethevisea doct shie manipaeri periodo to tao the letters crapses entire a the pero unitevisibiely e the interviewee belongs of the total property of the contraction of th ngar cerdius tauw better corotatae spinner points to the correct seques bender ac confrommer share reported by the spines points. Hinder the assumption that a these averand a gorner to are made that the ully essex in any likelihand estimates of the terma perpulativa presentativa sue straightforte maxima da verdadeira proporção da pologiação são diretas.

m= the true probability of A in the population, p= the probability that the spinner points to A, and v7__Ja viertheeith promobled lement savo veração. p = a propatilitatiet de same ple palementos as para A, e I se o i-

ésimo elemento da amostra disser sim 0l se o i-ésimo Then

Então

elemento da amostra disser não.
$$P(X_i=1) = \pi p + (1-\pi)(1-p),$$

 $P(X_i = 0) = (1 - \pi)p + \pi(1 - p),$

and arranging the indexing of the sample so that the first n_1 report "yes" while the second (n P(N)) report "The," the likelihood of the sample is

organizand Daindexação (da amp) (na de p) ded (que o p) imetro m(irepost) sim! enquant (p) segundo (n-n1) reporte "não", o probabilidade da amostra é The log of the likelihood is L = [irp + (1 p)]n-f(1 r)p + 7r(1-1)

$$L = [\text{irp} + (1 \quad p)] \cdot [(1 - \pi)] \cdot [(1$$

O logaritmo da probabilidade é $(n-n_1)\log \left[(1-\pi)p + \pi(1-p)\right],$ (2)

necessárias em $\frac{\log\left[(1\text{ - r})p + r(l\text{ -}p)\right], \text{ e as condições}}{(n-n_1)(2p-1)}$ $\frac{(n-n_1)(2p-1)}{(1-\pi)p+\pi(1-p)} = \frac{n_1(2p-1)}{\pi p+(1-\pi)(1-p)}$

(n - n1)(2p - 1)

or $\frac{(1-X)p + r(1-p) - rp + (-r)(}{\pi p + (1-\pi)} (1-p) = \frac{n_1}{n} \cdot$ (3)ou

$$7rp + (1 - 7r)(1 - p) = n$$
 (3)

RANDOMIZED RESPONSE

65

Then, supposing $p \neq 1/2$, the maximum likelihood estimate of π is RESPOSTA ALEATÓRIA

65

p-1 n_1 Então, supondo p 1/2, a esti \widehat{m} ativ $\widehat{n_2}$ de \widehat{m} axir $\widehat{n_2}$ veros $\widehat{s_n}$ ilhança de 7r é (4)

The expected value of the estimate is (4)

O valor esperadE $\hat{\mathbf{da}}$ $\overline{\mathbf{estimativa}}$ e^{p-1} e^{p-

$$= \frac{1}{2p-1} \left[p^{-1} + 1 \left(\frac{1}{p} \right) \pi p^{-1} \right] (1-\pi)(1-p)$$
 (5)

$$= \pi,$$
=- p_1 [p-1 + 7rp + (1-lr)(1-p)] (5)

$$= \pi,$$

$$= p_{1} [p-1] + 7rp + (1-lr)(1-p)]$$
and the variance of $\hat{\pi}$ is
$$= a \text{ variaWeira } \hat{\pi} l e^{\pm} e^{\frac{\pi}{2} \frac{\sqrt{4r} X_{i}}{(2p-1)^{2}n^{2}}}$$

$$= \frac{\left[\frac{n}{\pi} p^{\sqrt{4r}} (1-\pi)(1-p)\right] [(1-\pi)p + \pi(1-p)]}{(2p-1)Em}$$

$$= \frac{\left[\frac{n}{2} p^{4} + \frac{1}{2} \frac{i}{2} \frac{j}{2} \frac{j}{2}$$

Expression (5) shows his an unbiase (x estanate of the true population prop 6) tion π . Moreover, 1640ce 14218 Amaximum likelihood estimate and any useful n's are apt to be large, \$\phi\$ may be assumed normally distributed about \$\pi\$ with the expression (5) hiestra que e una estimativa imparcial da verdadeira proporção da variance indicated in expression (6). Thus all the usual confidence intervals pobliação it. Alem disso como 1 e una estimativa de maxima verossimilimança are easily established. Expression (6) also sets out the separate dependence of quaisquer no uteis são capazes de ser grande, / pode, ser assumido normalmente the variance of \$\phi\$ upon the choice of \$p\$. In fact, identifying distributed em forno de 71 com a variancia indicada na expressão (6). Assim, todos os intervalos de confiança usuals são facilmente estabelecidos. A expressão (6) também estabelece a dependência separada da Variância de resobre a escolha de p. Na verdade, identificar

as the variance due to sampling and writing expression (6) as

como a variação devido à
$$\frac{1}{\exp res são}$$
 de amostrage $\frac{1}{6(pscrita/2)somo_4}$ $\frac{1}{n}$, (7)

it is clear that the pariance of a can be expressed as the sum of the variance does to sampling plus the variance due to the random device.

é dayo que divadantes tiens pour em the estimation method implied by A. First, who likely are people to cooperate expressa como a soma da variancia devida a mostragem mais a variancia devida ao dispositivo aleatorio.

Du Electros idition of incident in the control of t مة المستورة المستورة

possibilidade de 79 obter valores fora da faixa de 0-1 não pode ser descartada, mas essa possibilidade é remota em grandes amostras.

the manner described? Second, how large a sample is required to obtain various dégrees of precision by this estimate as compared to the conventional estimate?

degrees of precision by this estimate as compared to the conventional estimate? The first question is primarily an empirical question but the rationale for damaneira descrita? Em sequindo lugar, qual e o tamanno da amostra necessaria para obter expecting better cooperation is clear. The individual being interviewed is asked yanos graffs de precisao por essa estimativa em comparação, com a estimativa convencional? for less. The matter of how much less is summarized by the parameter m. Note Apimeira questao e principalmente uma questao empirica, mas a justificativa para esperar first from expression (1) that if p=1/2, the likelihood function does not even uma melhor cooperação e clara. O indivíduo que esta sendo entrevistado e solicitado por dependo on matalo de quanto menos e resumida pelo parametro p. Observe primeiro a partir mation at all. Then note that if p=1, the entire procedure would reduce to the da expressão (1) que se p = 1/2, a função alterator de aconventional procedure or requiring the indivídual to state unreservedit whether or not he belonged to Group A for m s between 1/2 and 1 or between que o indivíduo declare sem reservas se en perece ou nação o grupo A. Para p se primeiro que o indivíduo declare sem reservas se en perece ou nação o grupo A. Para p se primeiro 2 a valor do un entre 1/2 e 0) a pessoa entrevistada fornece informações uteis, mas não absolutes informações contexto, o popo esta para pos entre vistado esta exatamente em qual grupo e a esta. Nesse contexto, o popo esta posa do como describing the nature of the cooperation between the interviewer and the interviewer. As poses from 1 to 1/2 the burden of cooperating passes from the 1/2, o onus da cooperação entre o entrevistado para o entrevistado. A medida que p val de 1 para viewee. As p goes from 1 to 1/2 the burden of cooperating passes from the 1/2, o onus da cooperação entre o entrevistado para o entrevistado. Portanto, parece interviewe to the interviewer. It therefore seems reasonable to expect that for razovel esperar que

The question of the sample size required for a given level of precision also depends on the parameter p. If a p close to 1 (or close to 0) is adequate to insure cooperation, then a smaller sample size is required than if a η close to 1/2 is A questão do tamanho amostral necessario para um determinado nível de precisão required to insure cooperation. Values of η close to 1/2 convey less information the modern depende do parameiro p. Se um p proximo de 1 (ou proximo de 0) foi adequação parameiro de amostra menor do que se um parameiro de amostra menor do que se um parameiro or, a larger sample size. Substituting values of p in expression (6) sets out the proximo de 1/2 for necessario para garantica copperação. Valores de p proximos a 1/2 proximo de 1/2 for mecessatio para garantifa cooperação. Valores de proximos a 1/2 un me precise relation. As an example, supposing a $\pi=5$ and a p halfway between the zero and full information points i.e. a p of 75, the variance shown by (6) is variancia maior da estimativa ou em um tanamino amostral major. A substituição dos valores de 1/2. This would imply that the sample size should be about 400 in order to pria expressão (6) estaberece a relação precisa. Como exemplo, supondo um 71 = 5 e um pa secure a standard deviation of 05. By way of comparison, the conventional mejo caminho entre os pontos de informação zero eccelo, ou seja, um p de 75, a variancia estimation method (equivalent to a p=1) way of comparison that a sample of only mostrada por (6) e 1/2. Isso implicana que o tananho da amostra devena ser de cerça de 400 about 100 would be sufficient for a standard deviation of 05—provided that para garantir um desvio padrão de 0.05. A titulo de comparação, o metodo de estimação convencional (equivalent e p=1) implicana que uma amostra de apenas cerca de 100 seria ... The more pertinent comparisons are between the randomized estimates and sufficience para difficulty comparisons are between the randomized estimates and suitcente para and desvio padrae de 0.05 desde que todos os entrevistados dissessem a regular estimates under the assumption that the regular estimates are handicapped by less than 100 per cent truthfulness. Suppose that in a regular survey as comparato estimation per the safe entre as estimations e as estimations entre estimations entre estimation estimation entre estimation entre estimation estimatio is or is not in Group A, the conventional estimate of the true population os membros do Grupo B dizem a verdade apenas com probabilidade 16. Enião, se 7 = 1 ou o proportion π is conforme o i-esimo membro da amostra relatar que ele está ou não no Grupo A, a estimativa convencional da verdadeira proporção populacional r é $\hat{\pi} = \frac{\sum_{n} Y_i}{n} \cdot$

$$\hat{\pi} = \frac{\sum Y_i}{n^n} \,. \tag{8}$$

The expected value, response bias [3, p. 89], and variance of this regular estimate are given by

O valor esperado, xir que resposta 13.10 ... 89 p. 8 variância desta estimativa regular são (9)dadas por

$$EV = 7rTa + [(1 - 7r)(1 - Tb)],$$
 (9)

RANDOMIZED RESPONSE

Bias $\hat{\pi} = E(\hat{\pi} - \pi)$ RESPOSTA ALEATÓRIA = $\pi[T_a + T_b - 2] + [1 - T_b]$, and (10)

 $\begin{array}{l} \text{Viés'} - \text{E}(\text{Tr}_{a} + (1 - \pi)(1 - T_{b})] [1 - \pi T_{a} - (1 - \pi)(1 - T_{b})] \\ \text{Var } \hat{\pi} = \frac{[\pi T_{a} + \text{Tb-2}] + [\text{1-Tb}], \text{ e [rTa + (1 - ir)(1 - T_{b})]}}{\pi (1 - \pi)(1 - T_{b})} . \end{array}$ (14)

Tb) NO1 - rT4 -(1- 7r)(1 - Tb)]
Tables/d and 2 then compare the mean square errors (the variance plus the square of the bias) of the randomized and regular methods of estimation under thousanntion that the interviewed individuals tell the trath in the randomized methed double the truth in the simplicated or the use habilities which are the use habilities which is also also the probabilities are the use the use

-método aleatório com probabilidades dadas por Ta e Tb. As duas colunas à esquerda de cada tabelifican Comparison of RANDOMIZED AND REGULAR ESTIMATES
FOR TRUE PROBABILITY OF A = .6 AND n=1000

TABELA 1. COMPARAÇÃO DE ESTIMATIVAS RANDOMIZADAS E REGULARES PARA VERDADEIRA PROBABILIDADE									
DE A=.6 ANDr n=1000 Regular Estimates			Mean Square Error Randomized						
Probability of Truth		D		Mean Square Error Regular Erro Quadrado Médio Randomizado					
T _a Est	imativas R egular	s Bias	p = .6	p = .7 ro Quadrático Médio	p=.8 Regular	p = .9			
Prob abil idade		Bias 03	$^{5.45}_{\stackrel{p=0,6}{1.62}}$	1.36	,60 p=0,8, c	p = 0.33			
.90	1.00	06	1.62	.40	.18	.10			
.70 .50	1.00	$\frac{18}{-03}$.19	1.36	.02	.01 .00 .60			
1.00 1.00	1.00	$\frac{03}{06}$ 30	1.62	.02	.01	.00			
1.700	1,000	1804	9.82	2.44	1.08	.60 .01			
1.50	1,090	3012	3.41	.85 ,92	.37	.21 .03			
12:000	:55	95	9,826	2,444	1.082	.03 .69			
1,005	.95	- :81	18,425	4.84	2.00	1 71			
1,990	.3%	02	9.70	2.41	1.05	. 33			
1,000	. 5 %	66	1.62	.46	.98	ďĥ.			
	.50 .90	01 02	18,2 5 1 9,70	4,5 4 5 2,41	2.00 7 1.06	1.11 .04 .59			
,70	,70	06	1,62	,40	.18	.10			

TABLE 2. COMPARISON OF RANDOMIZED AND REGULAR ESTIMATES FOR TRUE PROBABILITY OF A = .5 AND n = 1000

TABELA 2. CON	IPARAÇÃO DE E gu Par^aEstin a	STIMATIVAS RAN	DOMIZADAS E REGULARES PARA VERDADEIRA PROBABILIDADE Mean Square Error Randomized				
Probability	y of ${ m Truth}_{ m mativas}$ Regula	es Bias	p = .6	Mean Square Quadrado Médic p = .7 Irro Quadrático Médic	Randomizado $p = .8$	p = .9	
Probabilidade	da verdante 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	baix 03 05 15 05 00 15 00 15 00 25 00 00	7.15 P2.28 .28 7.150 25.200 25.000 25.000 25.000	1.79 p=0.57 .07 1.73 6.25 6.25 6.25 6.25 6.25	.79 p=0.25 .03 .01 2.78 2.78 2.78 2.78 2.78	.45 p9 .14 .022 .01 1.56 1.56 1.56 1.56	
.70 .50	,70 ,50	.00 .00	25,00 25,00	6,25 6,25	2,78 2,78	1.56 1.56	

paired values for T_a and T_b . The third column shows the bias of the non-random filethod, and the remaining collings extinct the third soft the mean square errors values and an incomparison of the resultance of the resulta

Both tables are constructed under the assumption that the p in each case is low enough to induce full cooperations of the randomized approach. Thus the advantages efither and emised method, shown by the propertion in the tables that are described a are in the capture of potential advantement at that depend upon those geomeratio pleatural machines obtast por antemized anethodo Navantholes othere is the clarification that the read on the discount of the contraction the residence ethorniene ovarient profinite mainte extrable dravit budestad and ut o motofor example artificity and a sequence of the mineral state of the sequence of the reserted irectly appredig to their extrapole west being minaritems to the being whompetan ingrograface the assay afference of But Eberbier supertendicat of monoda the omtion free consequences are represented the properties of the system. of performancion and the control of Appeyer greater improvement is possible if it is the larger population that hesitates to identify itself openly. This latter case is exemplified by the row in whish Thoria ainda Thaidré possível se for a população maior que hesita em se ide More gape relive it is to be a loss exact the tempercent of such a rases, where the bias **db**the regular estimate is 0 or negligible—there appear to be sizable potential gains through the randomized response Lt. should also checkent in minde that the netiential advantegos of its admerities succeens larger for larger campiles id frevers annheia samaleoire afe 200 haw pallo emide v that the montain in Table antechno poterioris La vandoni barres strom de 62 atores 4 a Thurst the sandomized presterodois uto de pares ferrandianthie ingtonen rean ifue a entlavare fieberari rodute es une congretie ve 1.62) Safet Ostion Assintil in portoral televibre terestros refice a deservado response queill peny notice sanst unnervisor with reservisor estimating properties, the method set a quite in a Bention e Bris em la contre sobre any a profesibilité de l'étais in tare atting sto avoite indisinference tino to the mathement is ally many alert operates the energy is setting to Raisseranished by a simply describing sassibilitaries we interest the construction of iespurovetle prohobilitymatem to cunich coedurate reseaues beas in filmes in this models then text a vive concerning and the control of the latter a vive and the control of the latter and the control of the control of the latter and the l which mointe de atoro orithering opposed will be a land ten calle emitely supposed lity hbvamente on a lea instervie some makes a give to mentather tals to unerafolm a constructions toethadvav" toensving of Josiphied de sve balasiaally chip providing per to bensuitand. a differentation for transfirm test Section 2 deutet a outpust is a become right of the teorno models graceri valenti. Esconstianon lekelikopal esti var tenforotten latter ordrenere basthe sursefare, and the properacion established through a forestion of Societies of Societies

A estimativa de máxima verossimilhança para o último esquema tem a mesma forma e a mesma forma e a mesma a forma e a mesma a forma e a mesma da de si furnishes no triormatian, a p of p furnish information depending on how far they are from p to the dichotomous nature of the population that telling the truth p of the time is equivalent to telling the truth p of the time.

² Como antes, ap de i não fornece nenhuma informação, ap de 1 fornece informações completas e outros p's fornecem informações dependendo de quão longe eles estão de J. É uma característica da natureza dicotômica da população dizer a verdade .2 do tempo é equivalente a dizer a verdade 0,8 do tempo.

69 RANDOMIZED RESPONSE

thus a question as to which of these or other equivalent randomized models is ft be breferred from the standpoint of increasing cooperation.

no Finally with should be anoted that title coast to extend the randomized response technique to activate distributions other than that appropriate to a simple dichatamente vacioble Astanguezample ethentechaigunica uld respondied ta castiparte es firme clustrious pes dintributio u abureus la bue de viene fiere proprieta de la companya del companya del companya de la companya del la companya de la companya the armine sach descentated year the continue of Sarties 3-la this case each intenviouse might be simply eaked to make if ve sagaratave adomired armar Just as with the proper timbers here it is short that other randomized response methodas representation de la contra metro constituir de la contra methodas representation problems con de la contra methodas representations de la contra methoda de la Assimilation proportion are beneather of estimate of estimate the proportion are beneather than the proportion are beneather the contract of the proportion are beneather than the proportion are beneather that the proportion are beneather the proportion are beneather than the proportion are Bream Rupoviernisse matematematers paraies su viotienta de estimativa mais geral. E. assim como no problema da proporção, a questão de qual técnica específica se mostrará superior é uma questão para investigação empírica.

I am indebted to the referee for helpful suggestions.

4. AGRADECIMENTOS

REFERENCES
Estou em dívida com o árbitro por sugestões úteis.
[1] Cochran, W. G., Sampling Techniques, Second Edition. New York: John Wiley and Sons, Inc., 1963. REFERÊNCIAS

Sons, Inc., 1963.

[2] Deming, W. E., Some Theory of Sampling. New York: John Wiley and Sons, Inc., 1950.

[3] Gocksen, W.G., Sampling Hochwidges, Segundad Sons, Inc., 1953.

[4] Farrier, W. S., Sampling Hochwidges, Segundad Sons, Inc., 1953.

[4] Farrier, W. S., Houng, Tenniada Amostragan, Journal Sons, Inc., 1953.

[4] Farrier, W. Houng, Tenniada Amostragan, Journal Sons, Inc., 1953.

[5] Helling, W. S., Houng, Tenniada Amostragan, Journal Sons, Methods and Theory, Volume

[5] Helling, W. S., Houng, J. Sons, Inc., 1953.

[6] Helling, W. S., W. S.,

[5] Stephan, FF, e McCarthy, PJ, Sampling Opinions. Nova York: John Wiley e Filhos, Inc., 1963.