

UMA REVISÃO SOBRE OS MÉTODOS DE AMOSTRAGEM EM INVESTIGAÇÃO

AUTORES

Jyotikumarie Juggernath: African Vision Research Institute (AVRI), Durban, South Africa

Kovin S Naidoo: Brien Holden Vision Institute (BHVI), African Vision Research Institute (AVRI), Durban, South Africa

Urmilla Bob: University of KwaZulu Natal (UKZN), Durban, South Africa

Vadivelu Moodley: African Vision Research Institute (AVRI), Durban, South Africa

REVISOR

James Loughman: Dublin Institute of Technology (DIT), Dublin, Ireland

AMOSTRAGEM

As amostras são usadas como um substituto para uma população poupar tempo, ou quando é impossível medir e contar toda a população. As amostras podem ser definidas como um subconjunto ou parte do total da população. As amostras podem ser classificadas em amostras não-probabilísticas e de probabilidade. Técnicas de amostragem probabilística representam subconjuntos de uma população com base matemática e são geralmente usadas em investigação quantitativa. O propósito de uma técnica de amostragem é o de obter uma amostra que represente a população a partir da qual foi elaborada, já que por vezes não é prático cobrir toda a população devido a diversas restrições. Desta forma, refletem-se resultados mais precisos. A população total a partir da qual uma amostra é elaborada é chamada uma base de amostragem. Uma amostra representativa deve ser cuidadosamente selecionada para representar as características da população total. A seleção da amostra e tamanho diferente de acordo com a investigação a ser conduzida. Amostragem não-probabilística significa apenas que as amostras do investigador não usam métodos de seleção aleatória. A amostragem não-probabilística envolve um grupo de técnicas de amostragem que ajudam o investigador a selecionar unidades de uma população que está interessada em estudar. Coletivamente, estas unidades formam a amostra que o investigador estuda. Uma característica nuclear de técnicas de amostragem não-probabilística é que as amostras são selecionadas com base no julgamento subjetivo do investigador, em vez da seleção aleatória.

CATEGORIAS DE AMOSTRAGEM

AMOSTRAGEM PROBABILÍSTICA

É uma seleção aleatória dos sujeitos, no qual cada indivíduo tem a mesma hipótese de ser selecionado.

AMOSTRAGEM ALEATÓRIA

Amostragem aleatória é a melhor forma de amostragem probabilística. Numa amostra aleatória cada membro da população tem a mesma probabilidade de ser incluído na amostra e cada amostra de um determinado tamanho tem a mesma hipótese de ser escolhida. Isto é verdade independentemente das semelhanças ou diferenças entre elas; no entanto, devem ser membros do mesmo universo. Uma amostra aleatória é realizada após ser construída uma base de amostragem adequada, sem ser tendenciosa para quaisquer características pessoais. A debilidade da amostragem aleatória é que a pessoa registrada mais de uma vez terá a mesma probabilidade aleatória de ser selecionada. Algumas pessoas não serão selecionadas se forem omitidas da lista. Em casos como estes, a amostragem por definição não será aleatória. Uma amostragem aleatória com uma base espacial reflete a amostragem da população usando atributos espaciais. A população é vista como um conjunto de pontos geográficos em diferentes áreas e cada ponto é selecionado aleatoriamente com diferentes áreas para formar a amostra

USAR UMA TABELA DE NÚMEROS ALEATÓRIOS

- Amostra de 10% de 500 clínicas na África do Sul
 - O número de cada clínica de 001 a 500
 - Ao trabalhar em qualquer direção selecionam-se três dígitos da tabela
 - Apenas os números que são menos de 500 são incluídos

Exemplo de uma tabela aleatória (Bless e Higson-Smith, 1995)

98	08	62	48	26	45	24	02	84	04
33	81	51	62	32	41	94	15	09	49
80	95	10	04	06	96	38	37	07	74
79	75	243	91	40	71	96	12	82	96
18	63	33	25	37	98	14	50	65	71

- Trabalho com os três primeiros dígitos do canto superior esquerdo da tabela e de seguida descendo:

980 338 809 797 186 862 151 ...

- Alternativamente, procede-se para o lado em vez de para baixo:

980 862 482 645 240 284 043 ...

Os números sublinhados são os seleccionados para a amostra! (Porquê? Porque os números sublinhados são menores que 500, o número de clínicas no nosso exemplo).

AMOSTRAGEM SISTEMÁTICA	<p>Na amostragem aleatória sistemática, as unidades são selecionadas em intervalos regulares e não são selecionadas aleatoriamente de uma tabela. O investigador primeiro escolhe aleatoriamente o primeiro número ou assunto da população e em seguida, seleciona cada enésimo termo da lista.</p> <p>Amostragem aleatória sistemática é muito fácil e pode ser feita manualmente. Os resultados são representativos da população a menos que determinadas características da população sejam repetidas para cada enésimo indivíduo. O processo de obter a amostra sistemática é semelhante como uma progressão aritmética.</p> <p>Por exemplo:</p> <ul style="list-style-type: none"> ▪ Cada enésimo elemento é amostrado ▪ Por exemplo, se uma população é de 1000 e precisamos de uma amostra de 100, cada décimo número deve ser selecionado. <ul style="list-style-type: none"> – O número que começa pode ser qualquer dígito entre 1 e 10 – Se o dígito inicial 5, os 100 elementos da amostra seriam 5, 15, 25,...485, 495, etc.
-------------------------------	---

AMOSTRAGEM PROBABILÍSTICA (cont.)

AMOSTRA ESTRATIFICADA	<p>Na amostragem estratificada, os subgrupos são representados proporcionalmente aos seus números dentro de uma população. Isto significa que a população é dividida em diferentes grupos (características diferentes). Um membro da população pertence a um (e somente um) estrato. Uma amostragem sistemática, é então, feita para cada estrato. Os estratos podem ser divididos com base em grupos espaciais ou socio-económico.</p> <p>Por exemplo: Ao realizar uma pesquisa na Flórida: a cidade pode ser dividida em três estratos, com base no estatuto social (rendimento alto = área 1; rendimento médio = área 2; rendimento baixo = área 3). A amostra será então selecionada aleatoriamente a partir de cada estrato ou ser mais estratificada de acordo com outras características e depois só depois ser caracterizada numa amostra.</p>
AMOSTRAGEM DE GRUPO	<p>Na amostragem de grupo, grupos de indivíduos e membros (clusters) constituem a unidade de amostragem, ao invés de uma única unidade ou elemento da população. A população é dividida em subgrupos heterogêneos ou agrupamentos e todas as unidades dentro dos agrupamentos são selecionadas. A população é dividida em grupos ou elementos geograficamente ou por terem alguns critérios uniformes. Os grupos podem ser selecionados aleatoriamente para um estudo mais aprofundado. A amostragem de grupo é usada mais frequentemente com redes sobrepostas numa grande área mapeada. A amostragem ocorre dentro de cada tabela. Este tipo de amostragem é normalmente usado por exemplo, em estudos de prevalência, tais como a prevalência da cegueira em diferentes regiões.</p>

AMOSTRAGEM NÃO-PROBABILÍSTICA

Amostragem não-probabilística não é realizada aleatoriamente. Amostragem é feita subjetivamente e com base em determinados critérios.

AMOSTRAGEM DE CONVENIÊNCIA	Os indivíduos são selecionados pelo investigador de acordo com a sua disponibilidade. O investigador geralmente seleciona os indivíduos que são mais fáceis de recrutar para o estudo. A amostragem por conveniência é frequentemente usada em estudos-piloto para obter tendências e dados básicos. No entanto, este tipo de amostragem não é representativo de toda a população. Pode ser visto como uma seleção tendenciosa porque qualquer pessoa disponível é selecionada.
AMOSTRAGEM INTENCIONAL	<p>A amostragem intencional é a seleção deliberada dos sujeitos a serem incluídos na amostra com base em alguns critérios predeterminados. A amostragem intencional é feita quando a população-alvo é desconhecida ou difícil de prever. No entanto, há uma necessidade de uma abordagem intencional mas sistemática.</p> <p>Por exemplo: quando são examinados pacientes pelos seus conhecimentos, atitudes e percepções sobre a cirurgia à catarata, pode-se questionar apenas aqueles pacientes que não tenham sido previamente submetidos a uma cirurgia à catarata.</p>
AMOSTRAGEM DE JULGAMENTO	<p>A amostragem de julgamento é semelhante à amostragem intencional. É a seleção deliberada da amostra com base em critérios predeterminados.</p> <p>Por exemplo: Enfermeiros especializados na identificação de condições visuais. O investigador pode escolher enfermeiros que estejam especializados na identificação de presbiopia.</p>
AMOSTRAGEM POR COTA	<p>A amostragem por cota também é uma forma de amostragem intencional. A população é dividida em categorias predeterminadas e os sujeitos dentro de cada categoria são deliberadamente selecionados até que se reúna um contingente específico para cada categoria.</p> <p>Por exemplo: Se precisar de uma amostra de 100 e uma população for 45% feminina e 55% masculina, continuará a amostragem até obter aquelas percentagens necessárias</p> <p>No entanto, se tiver uma amostra onde 45% são mulheres (cota atingida), mas não houver 55% homens, continuará até conseguir 55% homens.</p>
VARIAÇÃO MÁXIMA DE AMOSTRAGEM	<p>A amostragem de variação máxima procura representatividade através de probabilidade igual. É também uma forma de amostragem intencional onde categorias relacionadas ao tema são propositalmente identificados. O método procura indivíduos representando uma gama de diferenças no fenómeno em estudo.</p> <p>Por exemplo: Se deliberadamente tentar entrevistar uma seleção muito diferente de pessoas, as suas respostas agregadas podem ser próximas ou representar a visão de toda a população.</p>
AMOSTRAGEM DE BOLA DE NEVE	<p>Os indivíduos são identificados usando métodos probabilísticos ou não-probabilísticos. É um método semelhante à abordagem racional. É usado um sistema em rede para identificar ou aceder potenciais indivíduos. A amostra continua a crescer.</p> <p>Por exemplo: um grupo é identificado; Este grupo leva o investigador a um grupo semelhante e esse grupo aponta o investigador para outro grupo.</p>
AMOSTRAGEM AO LONGO DE UMA TRANSEÇÃO	Este método de amostragem é usado quando não há nenhum sistema de referência geográfica aparente. É usada uma transeção ao longo de uma linha para garantir uma boa amostragem.

TAMANHO DA AMOSTRA

Uma grande amostra é mais representativa da população total mas seria uma investigação de custo elevado, enquanto uma pequena amostra é menos dispendiosa e mais conveniente de investigar, mas podem ser imprecisa e não ser representativa de toda a população. Assim, é essencial que o tamanho da amostra seja representativo da população. Calcular o tamanho da amostra pode ser complicado ou envolver o uso de técnicas estatísticas complicadas de cálculo do tamanho da amostra ideal. Muitas vezes é necessário consultar alguém com formação estatística para trabalhar para deduzir o tamanho da amostra necessária para ser estatisticamente relevante. Os tamanhos da amostra são geralmente calculados com um nível de confiança de 95%. Dois indicadores, bastante precisos, em tamanhos de amostra indicam que cinco por cento da população são uma amostra adequada e uma amostra de mais de 30 (dependendo do tamanho da população e se a população é homogênea ou heterogênea) é estatisticamente representativa.

ANÁLISE DA POTÊNCIA ESTATÍSTICA

A análise da potência é geralmente feita antes de uma experiência (*apriori* da análise da potência) e é por vezes realizada após a recolha de dados (análise *post hoc* da potência) para determinar a potência no estudo. O objectivo de uma análise de potência é permitir que possa decidir aquando do processo da elaboração uma investigação, o tamanho necessário de uma amostra para permitir julgamentos estatísticos precisos e confiáveis e a probabilidade do teste estatístico em detectar os efeitos de um determinado tamanho numa situação particular. Quanto maior o tamanho da amostra, maior a potência estatística na análise. Os tamanhos de amostra maiores são também mais dispendiosos em termos de tempo, esforço e recursos. Uma análise estatística de potência permite que o investigador calcule um tamanho de amostra ideal, certificando-se que a análise é potente. Também ajuda, manter o tamanho da amostra o mais pequeno possível.

A potência estatística é a probabilidade de que uma análise estatística seja capaz de detectar falsas hipóteses nulas, o que significa simplesmente que a análise não fará um erro Tipo-II (o fracasso em rejeitar uma hipótese nula falsa. Noutras palavras, um falso positivo). A potência de uma análise estatística varia de acordo com a hipótese nula". Se a **hipótese nula** estiver errada por uma larga margem, vai ser fácil de detectar.

Em termos muito básicos, a potência estatística é a probabilidade de alcançar significância estatística. Por outras palavras, a potência estatística é a probabilidade de se obter um valor de p inferior a 0,05 ($p < 0,05$) com um tamanho de amostra dado.

Por exemplo: Imagine que sua hipótese de investigação é de que as pessoas ricas têm melhores cuidados visuais e, portanto, uma melhor qualidade de vida do que as pessoas pobres. Suponha que você tem uma medida válida e confiável do melhor cuidado visual e qualidade de vida e que a medida é um número entre 0 e 100, sendo 0 muito má e 100 muito bom. Agora imagine se testássemos todos os ricos e pobres e os ricos apontassem para uma média de 75 e os pobres apontassem para uma média de 35.

Imagine que fosse analisada um tamanho de amostra de 10 pessoas ricas e 10 pessoas pobres. Se assumirmos uma certa diferença entre os dois grupos como sendo verdadeiro (35 contra 75 neste exemplo), podemos então perguntar se é ou não uma amostra aleatória de um determinado tamanho (por exemplo, 10 em cada grupo) é suficiente para mostrar estatisticamente que essa diferença realmente existe.

Uma análise de potência *apriori*, é necessária para cada hipótese que vai ser testada pelo investigador, a fim de determinar o tamanho ideal da amostra. A análise de potência estatística é especialmente útil em questionários, experiências sociais e investigação médica, para determinar o número de indivíduos a testar necessários ao estudo ou teste.

TAMANHO DA AMOSTRA (cont.)

TABELA DE TAMANHO DE UMA AMOSTRA

Tabela 4.1: Determinar o tamanho da amostra de uma dada população

N	S	N	S	N	S	N	S	N	S
10	10	100	80	280	162	800	260	2800	338
15	14	110	86	290	165	850	265	3000	341
20	19	120	92	300	169	900	269	3500	246
25	24	130	97	320	175	950	274	4000	351
30	28	140	103	340	181	1000	278	4500	351
35	32	150	108	360	186	1100	285	5000	357
40	36	160	113	380	181	1200	291	6000	361
45	40	180	118	400	196	1300	297	7000	364
50	44	190	123	420	201	1400	302	8000	367
55	48	200	127	440	205	1500	306	9000	368
60	52	210	132	460	210	1600	310	10000	373
65	56	220	136	480	214	1700	313	15000	375
70	59	230	140	500	217	1800	317	20000	377
75	63	240	144	550	225	1900	320	30000	379
80	66	250	148	600	234	2000	322	40000	380
85	70	260	152	650	242	2200	327	50000	381
90	73	270	155	700	248	2400	331	75000	382
95	76	270	159	750	256	2600	335	100000	384

* Tamanho da população (N) **Amostra (S)

Adaptado de: Krejcie, R. V., Morgan, D. W., "Determining Sample Size for Research Activities", **Educational and Psychological Measurement**, 1970.