

O QUE É AMOSTRAGEM? PARTE II



- Principais métodos aleatórios
- A extensão da amostra
- Margem de erro

Francisco Cavalcante(f_c_a@uol.com.br)

- Administrador de Empresas graduado pela EAESP/FGV.
- É Sócio-Diretor da Cavalcante & Associados, empresa especializada na elaboração de sistemas financeiros nas áreas de projeções financeiras, preços, fluxo de caixa e avaliação de projetos. A Cavalcante & Associados também elabora projetos de capitalização de empresas, assessora na obtenção de recursos estáveis e compra e venda de participações acionárias.
- O consultor Francisco Cavalcante já desenvolveu mais de 100 projetos de consultoria, principalmente nas áreas de planejamento financeiro, formação do preço de venda, avaliação de empresas e consultoria financeira em geral.

Paulo Dragaud Zeppelini(f_c_a@uol.com.br)

- Administrador de Empresas com MBA em finanças pelo Instituto Brasileiro de Mercado de Capitais - IBMEC.
- Executivo financeiro com carreira desenvolvida em instituições financeiras do segmento de mercado de capitais. Atualmente é consultor da Cavalcante & Associados, empresa especializada na elaboração de sistemas financeiros nas áreas de projeções financeiras, preços, fluxo de caixa e avaliação de projetos.

ÍNDICE

	PÁG
APRESENTAÇÃO	3
PRINCIPAIS MÉTODOS ALEATÓRIOS	4
A EXTENSÃO DA AMOSTRA	7
CÁLCULO DA EXTENSÃO DA AMOSTRA	9
DETERMINAÇÃO DA MARGEM DE ERRO	10

APRESENTAÇÃO

No **Up-To-Date® 168** mostramos que a teoria da amostragem estuda as relações existentes entre uma população e as amostras extraídas dessa população. Enfatizamos que é útil para avaliação de grandezas desconhecidas da população, ou para determinar se as diferenças observadas entre duas amostras são devidas ao acaso ou se são verdadeiramente significativas.

Amostragem é o processo de determinação de uma amostra a ser pesquisada. A amostra é uma parte de elementos selecionada de uma população estatística.

Enquanto um **censo** envolve o exame de **todos** os elementos de um dado grupo, a amostragem envolve o estudo apenas de **uma parte** dos elementos. A finalidade da amostragem é fazer generalizações sem precisar examinar todos os elementos de um dado grupo.

Neste **Up-To-Date®** vamos finalizar este assunto mostrando os principais métodos aleatórios de amostragem e os problemas envolvidos na determinação do tamanho da amostra.

PRINCIPAIS MÉTODOS ALEATÓRIOS

➤ **Método Aleatório Simples**

É aquele em que cada elemento da população tem a mesma chance de ser escolhido para a amostra. A população é numerada seqüencialmente e depois a amostra é sorteada aleatoriamente ou relacionada a uma tabela de números aleatórios.

Este tipo de amostragem é chamado também de simples, ao acaso, aleatória, casual, elementar, randômica, etc., é equivalente a um sorteio lotérico.

➤ **Amostragem Estratificada**

Se a população pode ser dividida em subgrupos que consistem, todos eles, em indivíduos bastante semelhantes entre si, podemos obter uma amostra representativa entrevistando uma amostra aleatória de pessoas de cada grupo. Este processo pode gerar amostras bastante precisas, mas só é viável quando a população pode ser dividida em grupos homogêneos, ou estratos, e de cada um deles toma-se a amostra aleatória. Alguns exemplos: sexo (dois), grupos etários, as ocupações ou outras divisões em grupos que possam ser relevantes a um estudo em questão.

Uma das vantagens da estratificação é a diminuição das variâncias tendo em vista a maior homogeneidade dentro de cada estrato do que em comparação com a média da população total.

➤ **Amostragem por Conglomerado (ou em cacho)**

Neste método a população é dividida em diferentes conglomerados, extraindo-se uma amostra apenas dos conglomerados selecionados, e não toda a população. O ideal seria que cada conglomerado representasse tanto quanto possível o total da população. Na prática, selecionam-se os conglomerados geograficamente, escolhe-se aleatoriamente algumas regiões, em seguida algumas sub-regiões e finalmente alguns lares. Este processo assegura que as pessoas da amostra vivam em conglomerados, possibilitando ao pesquisador entrevistar poucas pessoas para obter o resultado desejado.

Aplica-se este método quando não se dispõe de uma lista dos membros da população ou de unidades estatísticas sobre as quais se aplicaria a seleção. Em tais casos é mais fácil obter-se o registro de grupos, entidades em que a população estatística esteja agrupada. Ao invés de se selecionar itens individuais da população, seleciona-se um grupo ou entidade e examina-se a totalidade dos componentes deste grupo. Exemplo: todos os alunos de uma escola, todos os componentes de uma família. Como estes aglomerados podem ser heterogêneos, pode-se estratificá-los de acordo com categorias que interessem à pesquisa.

➤ **Método de Amostragem por Estágios Múltiplos**

Sempre que não é possível atribuir a cada elemento da amostra um número para que todos tenham a mesma chance de seleção, como por exemplo, numa população estatística muito grande, pode-se optar por seleção preliminar de áreas geográficas, localidades, grupos ou conglomerados e, depois, numerar e sortear membros destas partes escolhidas que integram o universo.

Assim, num inquérito epidemiológico a ser aplicado sobre uma área muito extensa, pode-se sortear previamente alguns municípios, dentro destes, alguns distritos, e nestes últimos, um certo número de casas ou escolas onde serão

examinados todos os indivíduos ou se seleciona uma amostra aleatória para estudo.

Pontos importantes:

1. Na prática, é muito dispendioso selecionar amostras puramente aleatórias; assim, empregam-se dois outros métodos de amostragem, a amostragem por conglomerados e a amostragem estratificada;
2. Na amostragem por conglomerado, a população é dividida em grupos, ou conglomerados; selecionam-se aleatoriamente alguns conglomerados, e deles se extraem aleatoriamente os elementos que irão compor a amostra;
3. Na amostragem estratificada, a população é dividida em extratos, ou camadas tão semelhantes entre si quanto possível;
4. Não é apropriado aplicar a análise estatística à amostras que consistam em pessoas ou outras unidades às quais houve maior facilidade de acesso do entrevistador.

A EXTENSÃO DA AMOSTRA

Uma das perguntas mais freqüentes em estatística é: "Qual o tamanho da amostra que devemos tomar? Para que a amostra represente com fidelidade o universo é necessário que ela agrupe um número suficiente de casos. Este número, por sua vez, depende dos seguintes fatores:

a) amplitude do universo:

A extensão da amostra tem a ver com a extensão do universo, para tanto, os universos são classificados em finitos e infinitos. De acordo com Gil (1991:84) “ universos finitos são aqueles cujo número de elementos não excede 100.000. Universos infinitos, por sua vez, são aqueles que apresentam elementos em número superior a esse. São assim denominados porque, acima de 100.000, qualquer que seja o número de elementos do universo, o número de elementos da amostra a ser selecionada será praticamente o mesmo”.

b) O nível de confiança estabelecido:

De acordo com as normas gerais de probabilidade, a distribuição das informações coletadas tende, geralmente, a ajustar-se à curva “ normal “ (curva de Gauss) , que apresenta valores centrais elevados e valores extremos reduzidos. O nível de confiança refere-se à distância entre um ponto em relação à média. Numa curva normal, a área compreendida por um desvio padrão à direita e um desvio padrão à esquerda da média cobre 68% de seu total. Da mesma forma dois desvios-padrão cobrem

95,5% e três desvios-padrão cobrem 99,7%. Isso significa que, quando na seleção de uma amostra são considerados dois desvios-padrão, trabalha-se com um nível de confiança de 95,5%.

c) Erro máximo permitido:

Os resultados obtidos a partir de uma amostra não são exatamente exatos em relação ao universo de onde foram extraídos. O resultado apresenta um erro de medição que diminui à proporção que a amostra aumenta. O erro de medição é uma tolerância dentro da qual se considera a amostra normal.

d) Percentagem com que o fenômeno se verifica:

A estimativa prévia da porcentagem que um fenômeno se verifica é importante para definir o tamanho da amostra conforme veremos a seguir.

CÁLCULO DA EXTENSÃO DA AMOSTRA

O cálculo de extensão da amostra pode ser obtido através da fórmula:

$$n = \frac{s^2 \cdot p \cdot q \cdot N}{e^2 (N - 1) + s^2 \cdot p \cdot q}$$

onde: n = tamanho da amostra

s^2 = nível de confiança escolhido, expresso em números de desvio padrão

p = percentagem com a qual o fenômeno se verifica

q = percentagem complementar ($100 - p$)

N = tamanho da população

e^2 = erro máximo permitido.

Exemplo: na pesquisa de 10.000 sindicalizados , quantos são analfabetos? Presume-se que o número seja superior a 30% e deseja-se um nível de confiança de 95% (dois desvios) com a tolerância de erro de até 3%.

$$n = \frac{2^2 \cdot 30 \cdot 70 \cdot 10000}{9 \cdot (9999) + 4 \cdot 30 \cdot 70} = \frac{84.000.000}{98.391} = 853,74 \quad \text{ou} \quad \approx 854 \rightarrow \text{ sindicalizados serão pesquisados.}$$

DETERMINAÇÃO DA MARGEM DE ERRO

A margem de erro de uma amostra varia em função do valor encontrado, portanto, sua determinação só é possível após a obtenção dos dados. É calculada a partir da seguinte fórmula:

$$S_p = \sqrt{\frac{p \cdot q}{n}}$$

onde: S_p = erro-padrão ou desvio da percentagem com que se verifica determinado fenômeno;

p = percentagem com que se verifica o fenômeno;

q = percentagem complementar ($100 - p$);

n = número de elementos incluídos na amostra.

Exemplo: em uma pesquisa realizada com 1.000 pessoas adultas, se verificou que 30% bebem café pelo menos uma vez por dia, qual a possibilidade do resultado ser verdadeiro para todo o universo?

$$s = \sqrt{\frac{30 \cdot 70}{1000}} = 1,45$$

Como o valor encontrado (1,45) corresponde a um desvio, para dois desvios teremos o dobro (2,95). Isto significa que para um nível de confiança de 95% (dois desvios), o resultado da pesquisa apresentará uma margem de erro de 2,95%. É provável que o número de consumidores de café esteja entre 27,05% ($30 - 2,95$) e 32,95% ($30 + 2,95$) do total da população.