

Amostragem

Amostragem – Conceitos básicos

- Universo ou população
 - Indivíduos com os quais obteremos a informação necessária ao alcance dos objetivos de pesquisa.
- Amostra
 - Parcela da população de interesse selecionada.
 - Estrutura da amostra: listagem de todos os elementos da população.
- Unidade de análise

Estratégias de amostragem

- Amostragem probabilística

Adota parâmetros estatísticos para garantir que todos os elementos da população terão igual probabilidade de serem selecionados.

- Amostragem não-probabilística

A amostra é selecionada por critérios estabelecidos pelo pesquisador, de acordo com os objetivos do estudo.

Técnicas de amostragem probabilística

- Amostragem probabilística simples
 - Escolha aleatória dos elementos que compõem a amostra.
 - Todos os elementos têm igual probabilidade.
 - Vantagem: resultados podem ser projetados.
 - Desvantagem: baixa precisão e baixa garantia de representatividade.
- Amostragem probabilística sistemática
 - Elementos são selecionados a partir de uma listagem existente, aleatoriamente, estabelecendo-se um intervalo entre eles.
 - Ponto de partida aleatório.
 - Intervalo = tamanho da população / tamanho da amostra.
 - Vantagem: maior representatividade.
 - Desvantagem: necessidade de um mailing que sirva de base para seleção dos elementos.

Técnicas de amostragem probabilística

- Amostragem probabilística estratificada
 - Subdivide a população em subconjuntos exclusivos com base em fatores que se pretende mensurar: classe social, sexo, idade, escolaridade etc.
 - Elementos que compõem estratos são selecionados pela técnica da amostragem probabilística simples.
 - Número de estratos: de 2 a 6.
 - Vantagem: inclui todas as subpopulações importantes; oferece maior precisão.
 - Desvantagem: dificuldade na escolha de variáveis relevantes, é cara e inviável com muitas variáveis.

Técnicas de amostragem probabilística

- Amostragem probabilística por conglomerado ou agrupamento
 - População é dividida em subpopulações mutuamente excludentes
 - Conglomerado de um estágio: todos os elementos dos subgrupos escolhidos são incluídos na amostra.
 - Conglomerado de dois estágios: tira-se uma amostra probabilística simples de cada um dos subgrupos escolhidos.
 - Vantagem: maior viabilidade pelo baixo custo.
 - Desvantagem: produz amostras imprecisas pela dificuldade de se produzir grupos heterogêneos; dificuldade de interpretação de dados.

Técnicas de amostragem não-probabilística

- Amostragem não-probabilística por conveniência
 - O entrevistador estabelece com base em sua conveniência as unidades de amostragem.
 - Vantagem: barata e rápida.
 - Desvantagem: amostra não é representativa pois elementos possuem características diferentes da população-alvo.
- Amostragem não-probabilística por julgamento
 - Elementos da amostra são selecionados pelo pesquisador com base em critérios que julga representar a população.
 - Vantagem: baixo custo, rápida
 - Desvantagem: sofre viés do pesquisador; não permite generalização.

Técnicas de amostragem não-probabilística

- Amostragem não-probabilística por cota
 - Variação da amostragem por julgamento que busca refletir composição da população.
 - Identifica-se categorias entre os elementos da população.
 - Estima-se a proporção desta categoria dentro a população, com base em fontes secundárias.
 - Elementos são selecionados por conveniência ou julgamento desde que se encaixe nas características de controle.
 - Desvantagem: pode não espelhar características relevantes da população; pode produzir viés na seleção dos elementos.
 - Vantagem: baixo custo, fácil de obter.

Cálculo Amostral

Cálculo amostral

- Cálculos estatísticos dependem do tamanho do universo.
- Elementos considerados no cálculo da amostra:
 - Confiança: grau de precisão do cálculo estatístico
 - Margem de erro: diferença esperada entre o parâmetro da população e o resultado a ser obtido pela pesquisa.
Margem de erro comumente utilizada varia de 3 a 5%
 - Distribuição: indica a proporção de elementos distribuídos com relação ao parâmetro da população.

Equações para o cálculo da amostra

POPULAÇÕES FINITAS ou $N < 1000$

$$n = \frac{NZ^2pq}{E^2(N-1) + Z^2pq}$$

POPULAÇÕES INFINITAS OU $N > 1000$

$$n = \frac{Z^2pq}{E^2}$$

Legenda

N= Universo

n = Amostra

Z = Confiança

E = Erro

pq = Distribuição dos elementos da amostra

p = Nível de aprovação

q = Nível de reprovação

Nível de confiança

Níveis de confiança utilizados:

| Nível de Confiança % | Índice Z |
|---------------------------------|---------------------|
| 68 | 1 |
| 90 | 1,645 |
| 95 | 1,96 |
| 95,5 | 2 |
| 99 | 2,575 |

Distribuição dos elementos da amostra

Geralmente não se conhece a % dos que respondem favoravelmente, considerando-se $p=50%$ e $q=50%$

| p | q | p.q |
|----|----|------|
| 10 | 90 | 900 |
| 20 | 80 | 1600 |
| 30 | 70 | 2100 |
| 40 | 60 | 2400 |
| 50 | 50 | 2500 |
| 60 | 40 | 2400 |
| 70 | 30 | 2100 |
| 80 | 20 | 1600 |
| 90 | 10 | 900 |

P= Nível de aprovação Q= Nível de reprovação

Tamanho da amostra

- O tamanho da amostra é determinado pelo nível de confiança desejado, o erro de estimação aceitável e a variância da população.
- O cálculo estatístico da amostra pode apresentar uma variância com relação ao parâmetro da população.
- Essa diferença é a medida de confiabilidade dos resultados.
- A estimação, baseada na estatística da amostra, possui um intervalo que reflete sua variância e o nível de confiança.

Tamanho da amostra – Universo, erro, nível de confiança

A tabela é baseada a fórmula para cálculo de amostra de pequenas populações considerando 50-50 de distribuição.

| Tamanho da População (N) | Tamanhos de Amostras | | | | | |
|--------------------------|----------------------------------|------------|-------------|----------------------------------|------------------|-------------|
| | Nível de Confiança de 95% | | | Nível de Confiança de 99% | | |
| | +3% | +5% | +10% | +3% | +5% | +10% |
| 500 | 250 ^a | 218 | 81 | 250 ^a | 250 ^a | 125 |
| 1.000 | 500 ^a | 278 | 88 | 500 ^a | 399 | 143 |
| 1.500 | 624 | 306 | 91 | 750 ^a | 460 | 150 |
| 2.000 | 696 | 323 | 92 | 959 | 498 | 154 |
| 3.000 | 788 | 341 | 94 | 1.142 | 544 | 158 |
| 5.000 | 880 | 357 | 95 | 1.347 | 586 | 161 |
| 10.000 | 965 | 370 | 96 | 1.556 | 622 | 164 |
| 20.000 | 1.014 | 377 | 96 | 1.687 | 642 | 165 |
| 50.000 | 1.045 | 382 | 96 | 1.777 | 655 | 166 |
| 100.000 | 1.058 | 383 | 96 | 1.809 | 659 | 166 |

Tamanho da amostra – Homogênea / Heterogênea

A tabela foi calculada com confiança 95%.

| POPULAÇÃO | ERRO AMOSTRAL 3% | | ERRO AMOSTRAL 5% | | ERRO AMOSTRAL 10% | |
|-------------|------------------|-------|------------------|-------|-------------------|-------|
| | 50/50 | 80/20 | 50/50 | 80/20 | 50/50 | 80/20 |
| 100 | 92 | 87 | 80 | 71 | 49 | 38 |
| 250 | 203 | 183 | 152 | 124 | 70 | 49 |
| 500 | 341 | 289 | 217 | 165 | 81 | 55 |
| 750 | 441 | 358 | 254 | 185 | 85 | 57 |
| 1.000 | 516 | 406 | 278 | 198 | 88 | 58 |
| 2.500 | 748 | 537 | 333 | 224 | 93 | 60 |
| 5.000 | 880 | 601 | 357 | 234 | 94 | 61 |
| 10.000 | 964 | 639 | 370 | 240 | 95 | 61 |
| 250.000 | 1.023 | 665 | 378 | 243 | 96 | 61 |
| 50.000 | 1.045 | 674 | 381 | 245 | 96 | 61 |
| 100.000 | 1.056 | 678 | 383 | 245 | 96 | 61 |
| 1.000.000 | 1.066 | 678 | 383 | 245 | 96 | 61 |
| 100.000.000 | 1.067 | 683 | 384 | 246 | 96 | 61 |