



**PONTIFÍCIA UNIVERSIDADE
CATÓLICA DE MINAS GERAIS**

Técnicas de Amostragem

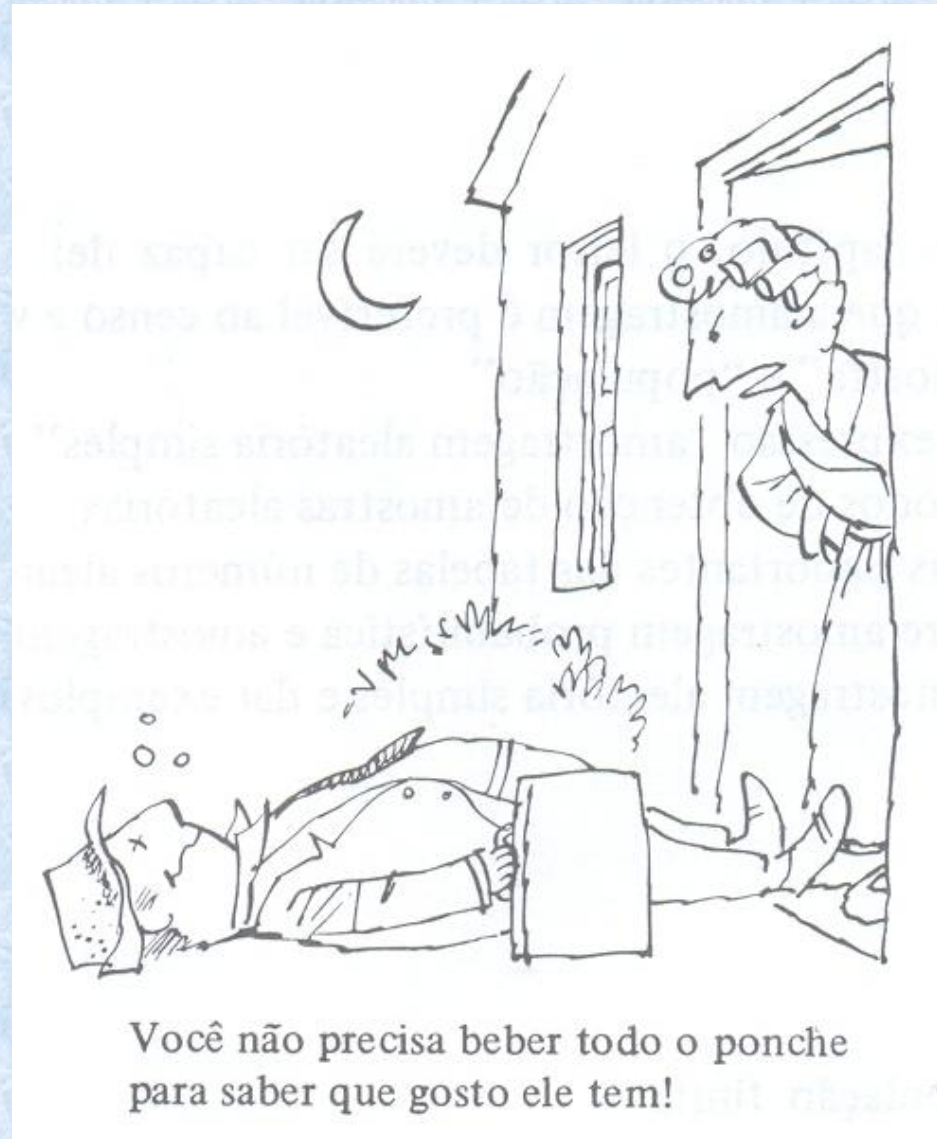
Unidade I – Aplicações em Amostragem

1.1 Introdução

- Um dos principais objetivos da maioria dos estudos, análises ou pesquisas estatísticas é fazer generalizações seguras – com base em amostras – sobre as populações das quais se extraíram as amostras.
- Por exemplo: Se quiséssemos estimar a despesa média de uma pessoa com suas férias, tomaríamos como amostra a despesa de passageiros da classe luxo em um cruzeiro de quatro semanas? Evidentemente que não. Mas quais e quantas pessoas em férias devemos incluir em nosso estudo?

1.2 Amostras e populações

- Um *censo* envolve um exame de todos os elementos de um dado grupo (*população*).
- Uma *amostragem* envolve o estudo de apenas uma parte (*amostra*) dos elementos da *população*.
- A finalidade da amostragem é fazer generalizações sobre todo um grupo sem precisar examinar cada um de seus elementos.



1.3 Amostragem *versus* censo

- À primeira vista pode parecer que a inspeção completa ou total de todos os itens de uma população seja mais conveniente do que a inspeção de apenas uma amostra deles.
- Na prática, o contrário é que é quase sempre válido: a amostragem é preferível ao censo.
- Vamos explorar as situações onde a amostragem é mais vantajosa:
 - ✓ A população pode ser *infinita*, e então o censo se tornaria impossível.

- ✓ Uma amostra pode ser mais *atualizada* do que um censo. Se você necessita de uma informação rapidamente, um estudo de toda a população, normalmente no caso de itens muito numerosos ou muito dispersos, pode consumir demasiado tempo e perder utilidade.
- ✓ O *custo* de um censo pode ser proibitivo, normalmente se o custo individual é elevado e se existem muitos itens na população. Por exemplo, o IBGE pediu R\$ 600 milhões para fazer o censo brasileiro do ano 2000, que foi realizado por 120 mil recenseadores.
- ✓ A *precisão* pode sofrer no caso de um censo de uma grande população. A amostragem envolve menor número de observações e, conseqüentemente, menor número de coletores de dados.

✓ *Testes destrutivos.* Os testes podem apresentar caráter destrutivo, ou seja, os itens examinados são destruídos no próprio ato do experimento. Nesse caso, o censo nos daria o panorama preciso de uma população que não mais existe.

✓ *O tipo de informação* pode depender da utilização de uma amostra ou de um censo.

Por exemplo, se tivermos uma população de fósforos e quisermos avaliar a porcentagem de falhas poderemos riscar um por um e ao final concluirmos que a falha é de, digamos, 2%.



- Há certas situações em que é mais vantajoso examinar todos os itens de uma população, ou seja, fazer um censo:
 - ✓ A *população* pode ser *tão pequena* que o custo e o tempo de um censo sejam pouco maiores que para uma amostra.
 - ✓ Se o *tamanho da amostra é grande em relação ao da população*, o esforço adicional requerido por um censo pode ser pequeno.
 - ✓ Se é exigida a *precisão completa*, então o censo é o único método aceitável. Em face da variabilidade amostral, nunca podemos ter certeza de quais sejam os verdadeiros parâmetros da população. Além disso, pode haver tendenciosidade ou erros na coleta dos dados.
 - ✓ Ocasionalmente, já se dispõe de informação completa, de modo que não há necessidade de amostra.

1.4 Técnicas de amostragem

- Para coletar dados não tendenciosos, é importante que a amostra seja representativa da população.
- Técnicas de amostragem apropriadas devem se usadas para garantir que as inferências sobre a população sejam válidas.
- Lembre-se de que, quando um estudo é feito com dados imprecisos, os resultados são questionáveis.

Importante

- o Uma amostra tendenciosa é aquela que não é representativa da população da qual foi extraída.
- o Dados coletados de forma descuidada podem ser tão inúteis que nenhum processamento estatístico consegue salvá-los.

1.4.1 Amostragem aleatória simples

- O processo de se selecionar uma amostra aleatória simples compreende os seguintes passos:
 - ✓ A amostra é escolhida elemento a elemento.
 - ✓ A população é numerada de 1 a N .
 - ✓ Escolhem-se na tábua de números aleatórios, n números compreendidos entre 1 e N .
 - ✓ Esse processo equivale a um sorteio no qual se colocam todos os números misturados dentro de uma urna.
 - ✓ Os elementos correspondentes aos números escolhidos formarão a amostra.

- A amostragem aleatória simples tem a seguinte propriedade:
 - ✓ Qualquer subconjunto da população, com o mesmo número de elementos, tem a mesma probabilidade de fazer parte da amostra.
- A seleção de uma amostra aleatória simples pode ser facilitada com o uso de números aleatórios, ou seja, números resultantes de sucessivos sorteios do conjunto $\{0, 1, 2, \dots, 9\}$, fazendo com que todo número com a mesma quantidade de algarismos tenha a mesma probabilidade de ocorrência. O quadro a seguir apresenta alguns números aleatórios gerados pelo Microsoft Excel:

Tabela de números aleatórios

1 8 6 3 4 1	7 3 8 1 4 1	1 0 6 6 6 8	9 9 7 0 3 2	4 0 2 2 1 2
2 3 5 9 5 6	5 5 9 3 7 6	4 7 7 9 7 8	3 6 9 3 9 0	5 5 3 4 4 1
6 2 2 4 0 6	6 5 4 0 2 1	6 0 8 6 4 5	2 9 2 4 9 9	9 4 1 5 5 3
4 5 3 0 7 7	1 1 2 0 1 7	5 2 5 3 1 9	6 9 0 7 5 9	0 7 0 9 0 0
6 6 9 6 4 5	3 7 7 6 3 1	8 9 0 9 4 1	9 4 2 7 5 2	3 5 3 6 5 4
1 4 7 7 2 8	4 7 3 1 6 7	0 5 2 1 8 6	9 7 6 7 1 3	6 0 1 2 9 6
2 9 4 4 9 6	1 2 8 8 4 2	6 8 5 3 1 8	3 3 5 8 7 6	7 8 7 0 5 7
0 2 1 5 7 1	1 1 8 9 7 3	6 5 0 0 9 9	8 9 4 3 4 2	9 6 7 7 9 4
9 3 8 2 5 6	2 1 9 6 2 0	3 5 8 2 0 8	6 1 6 2 6 4	6 0 6 2 8 6
6 0 3 2 5 1	1 8 6 3 7 1	6 3 2 1 2 5	7 7 6 3 3 8	0 8 0 1 6 6
3 1 9 9 5 8	4 4 6 6 3 1	7 3 1 6 1 0	7 4 6 7 4 0	9 9 1 9 0 0
5 2 5 2 3 7	3 6 8 5 3 8	1 2 8 6 8 8	1 0 5 0 5 3	2 5 7 9 1 8
1 9 1 2 4 0	0 4 0 6 1 6	6 5 3 9 8 0	0 3 7 5 6 1	5 9 4 9 3 4
5 8 3 7 9 6	8 6 6 1 8 2	4 9 1 7 9 6	2 2 9 7 4 5	3 7 8 4 3 5
4 1 6 7 6 8	8 3 1 4 2 8	8 0 4 5 8 7	2 4 3 5 7 5	3 6 7 2 3 5
3 3 3 8 9 6	8 4 6 4 8 9	5 5 6 7 3 4	1 9 3 6 2 8	9 0 5 4 0 5
7 4 9 2 3 4	7 3 2 3 1 6	8 8 8 9 0 0	6 4 5 1 6 1	8 9 6 4 5 6
4 9 3 1 5 0	2 3 9 6 1 0	1 5 2 0 9 9	1 1 8 5 8 7	1 1 8 9 3 0
2 1 4 0 4 1	7 6 9 7 5 0	6 6 5 9 5 3	0 9 1 0 5 2	1 9 3 0 5 0
3 7 5 0 0 0	9 7 9 6 7 1	6 3 5 9 2 2	4 0 9 6 9 3	1 3 9 6 4 1
2 6 2 6 8 5	7 2 8 8 7 6	1 1 3 6 4 3	2 8 8 4 0 3	1 0 3 1 8 1
0 6 8 0 3 5	8 0 8 7 5 9	8 9 9 5 7 7	7 9 5 1 0 8	2 6 3 2 6 3
8 6 2 5 3 2	1 2 3 0 9 7	4 3 7 2 3 3	0 5 6 5 0 3	5 0 8 4 6 4
7 1 1 8 6 8	0 7 2 3 6 3	1 9 6 8 1 5	5 6 3 5 9 4	7 5 3 7 0 2
2 7 2 7 2 4	5 4 3 7 1 6	0 5 2 4 1 8	6 1 4 6 0 6	5 8 9 9 3 8 ₁₁

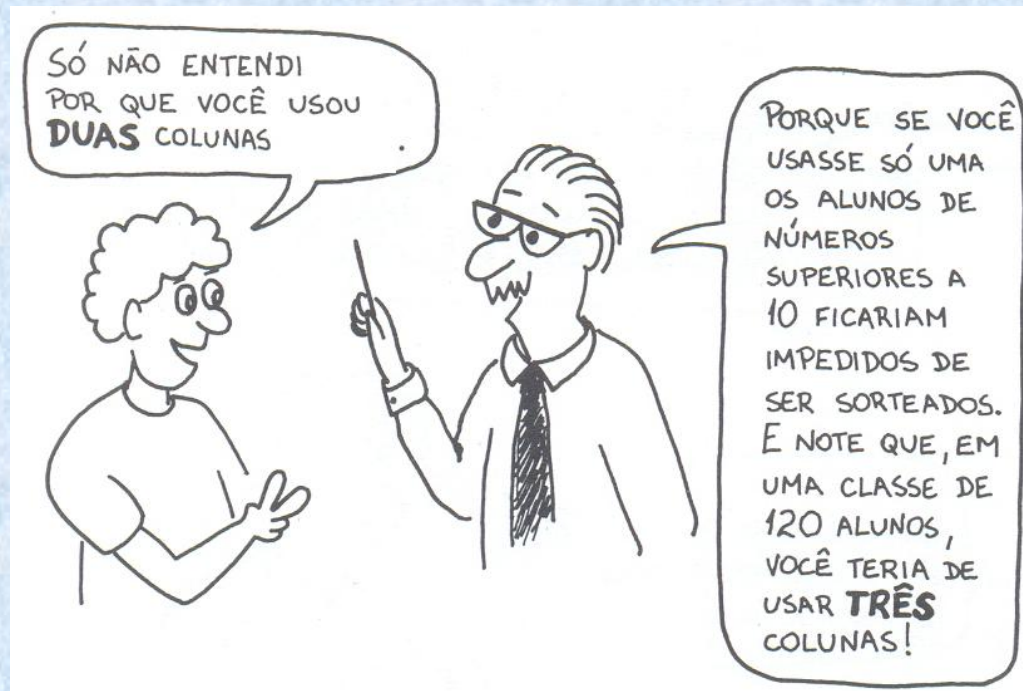
Exemplo 1: Amostragem aleatória simples

Considere o problema de extrair uma amostra aleatória simples de tamanho 6 da seguinte população de funcionários de uma empresa, na qual identificamos cada indivíduo com um número:

1 Aristóteles	8 Anastácia	15 Arnaldo	21 Bartolomeu	27 Bernardino
2 Cardoso	9 Carlito	16 Cláudio	22 Ermílio	28 Ercílio
3 Ernestino	10 Endevaldo	17 Francisco	23 Felício	29 Fabrício
4 Geraldo	11 Gabriel	18 Getúlio	24 Hiraldo	30 João da Silva
5 Joana	12 Joaquim	19 Joaquina	25 José da Silva	31 José de Souza
6 Josefa	13 Josefina	20 Maria José	26 Maria Cristina	32 Mauro
7 Paula	14 Paulo César			

Observe o processo:

1. A população (N) já está numerada de 1 a 32.
2. Temos que selecionar uma amostra (n) de 5 números na tabela de números aleatórios.
3. Temos que fazer a leitura na tabela de números aleatórios utilizando dois algarismos.



4. Você deve escolher o ponto de partida para a leitura da tabela de números aleatórios. Por exemplo, vamos começar a leitura da 1ª linha, 1ª coluna, da esquerda para a direita:

1	8	6	3	4	1	7	3	8	1	4	1	1	0	6	6	6	8	9	9	7	0	3	2	4	0	2	2	1	2
2	3	5	9	5	6	5	5	9	3	7	6	4	7	7	9	7	8	3	6	9	3	9	0	5	5	3	4	4	1
6	2	2	4	0	6	6	5	4	0	2	1	6	0	8	6	4	5	2	9	2	4	9	9	9	4	1	5	5	3
4	5	3	0	7	7	1	1	2	0	1	7	5	2	5	3	1	9	6	9	0	7	5	9	0	7	0	9	0	0
6	6	9	6	4	5	3	7	7	6	3	1	8	9	0	9	4	1	9	4	2	7	5	2	3	5	3	6	5	4
1	4	7	7	2	8	4	7	3	1	6	7	0	5	2	1	8	6	9	7	6	7	1	3	6	0	1	2	9	6



Sentido: Da esquerda para a direita.

Números aleatórios: 18 10 32 22 12

Amostra: {Getúlio, Endevaldo, Mauro, Ermílio, Joaquim}

- A dificuldade para fazer uma amostragem desse tipo é ter um arquivo com todos os elementos da população.

Exercício 1: Amostragem aleatória simples

Em uma empresa deseja-se fazer uma pesquisa com os funcionários sobre o grau de satisfação em relação a vários tópicos. O setor de recursos humanos forneceu a seguinte listagem com o nome de todos os funcionários:

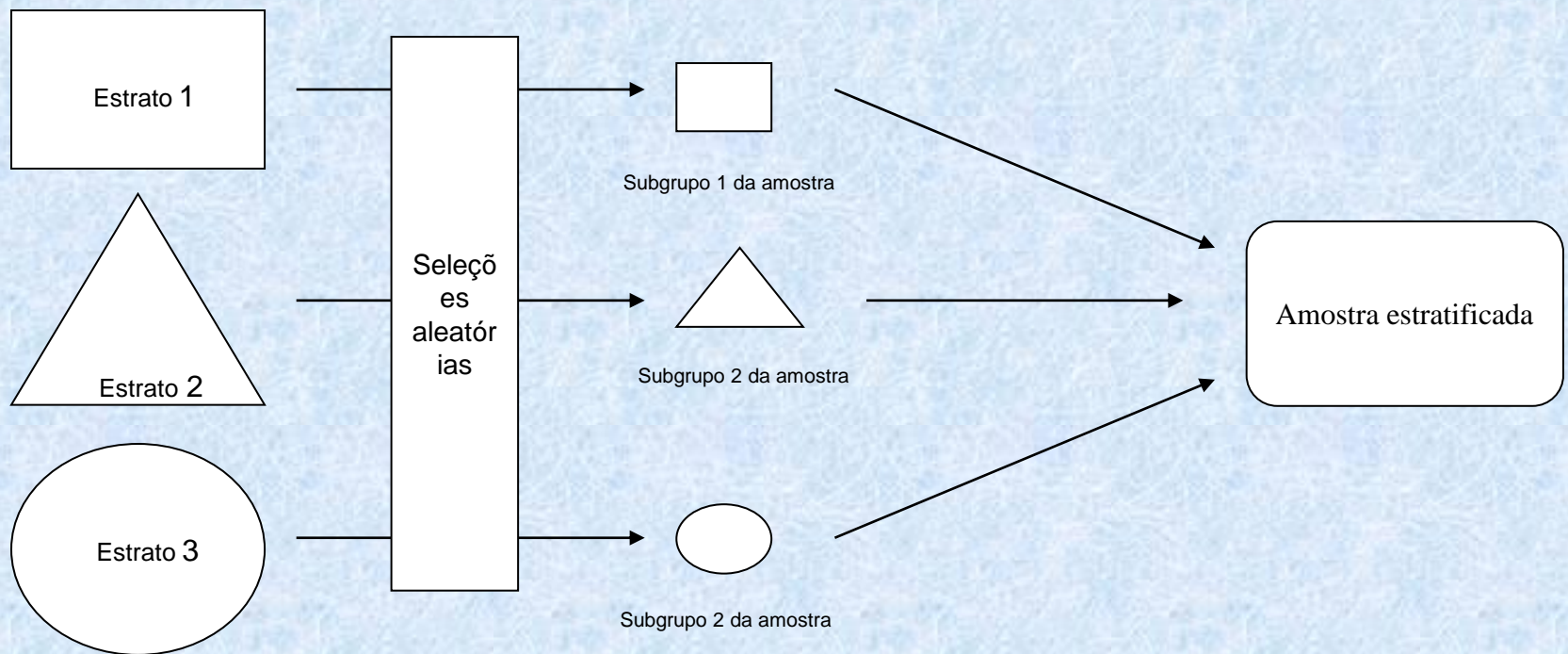
1 ALDOBERTON	16 ERNANE	31 LIVIA
2 ALEX	17 FABIANE	32 LUCIANA
3 ALINE	18 FELIPE	33 LUCYANA
4 ALINY	19 FELIPE PAULO	34 MARCELA
5 ANDERSON	20 FLAVIA	35 MARIA
6 BERENICE	21 FRANLIANE	36 MARIANA
7 BETHANIA	22 GENIANE	37 MARIANA LÚCIA
8 BRUNA	23 GREICIANE	38 MERCIA
9 BRUNO	24 HUARTALOS	39 NATALIA
10 CAMILA	25 KAMILA	40 OLIVIA
11 CAROLINA	26 KARINA	41 POLIANA
12 CHALINE	27 KEMMUEL	42 PRISCILLA
13 CHRISTIANE	28 KEULA	43 RAISSA
14 CRISTIANE	29 LEANDRO	44 SIMONE
15 DILIANE	30 LEIDILAINE	45 TALITA
		46 THAMILLES

Selecione uma amostra aleatória simples de 13 funcionários da empresa para participarem da pesquisa. Comece a leitura da última linha, 1ª coluna, da esquerda para a direita, de baixo para cima.

1.4.2 Amostragem estratificada

- Se dispomos de dados sobre a composição de uma população, e isto tem importância para nossa pesquisa, podemos eventualmente melhorar uma amostragem aleatória mediante estratificação.
- O processo consiste do seguinte:
 - ✓ Divide-se a população em subconjuntos diferentes, chamados estratos, que compartilham uma característica similar, como idade, gênero, etnia ou até mesmo preferência política.
 - ✓ A partir daí, uma amostra é selecionada aleatoriamente e separadamente de cada estrato.

- ✓ O emprego de uma amostra estratificada nos assegura que cada estrato ou segmento da população está representado.
- ✓ Observe o esquema a seguir:



Temos duas maneiras de se extrair amostras estratificadas:

➤ *Amostragem estratificada aleatória proporcional*: A proporcionalidade do tamanho de cada estrato da população é mantida na amostra. Por exemplo, se um estrato abrange 20% da população, ele também deverá abranger 20% da amostra.

➤ *Amostragem estratificada aleatória uniforme*: Selecionamos o mesmo número de elementos em cada estrato. É o processo usual quando se deseja comparar os diversos estratos.

Exemplo 2: Amostragem estratificada

Utilizando os mesmos funcionários do exemplo 1, suponha agora que os funcionários de 1 a 16 trabalham durante o dia e os funcionários de 17 a 32 trabalham no turno da noite.

Selecione amostras estratificadas aleatórias proporcionais de cada turno de tal maneira que a amostra final tenha tamanho 12:

Turno: Diurno

1 Aristóteles	9 Carlito
2 Cardoso	10 Endevaldo
3 Ernestino	11 Gabriel
4 Geraldo	12 Joaquim
5 Joana	13 Josefina
6 Josefa	14 Paulo César
7 Paula	15 Arnaldo
8 Anastácia	16 Cláudio

Turno: Noturno

1 Francisco	9 José da Silva
2 Getúlio	10 Maria Cristina
3 Joaquina	11 Bernardino
4 Maria José	12 Ercílio
5 Bartolomeu	13 Fabrício
6 Ermílio	14 João da Silva
7 Felício	15 José de Souza
8 Hiraldo	16 Mauro

Tamanho da população: $N=32$

Tamanho da amostra final: $n=12$

Como os dois estratos têm o mesmo tamanho ($N_1 = N_2 = 16$) então teremos que selecionar uma amostra aleatória de tamanho 6 para cada estrato.

Vamos começar retirando uma amostra aleatória dos trabalhadores do turno DIURNO. Se começássemos a leitura da tabela de números aleatórios da 1ª linha, 1ª coluna, da esquerda para a direita, teríamos:

Números aleatórios: 10 12 06 15 11 07

Amostra do estrato 1 (turno diurno): {Endevaldo, Joaquim, Josefa, Arnaldo, Gabriel e Paula}

Vamos agora selecionar uma amostra aleatória dos trabalhadores do turno NOTURNO. Se começássemos a leitura da tabela de números aleatórios da última linha, 1ª coluna, da esquerda para a direita, de baixo para cima, teríamos:

Números aleatórios: 16 05 06 07 15 02

Amostra do estrato 2 (turno noturno): {Mauro, Bartolomeu, Ermílio, Felício, José de Souza e Getúlio}

Exercício 2: Amostragem estratificada

Utilizando-se os mesmos dados do exercício 1. Sabe-se que os funcionários de 1 a 31 trabalham no setor de produção e o restante dos funcionários trabalham no setor administrativo. Como o objetivo da pesquisa é avaliar o grau de satisfação dos funcionários, a direção da empresa julgou necessário estratificá-los de acordo com o setor de trabalho. Selecione amostras estratificadas aleatórias proporcionais de cada setor de tal maneira que a amostra final tenha tamanho 20. Para a seleção do 1º estrato, comece a leitura da tabela de números aleatórios da quinta linha, 1ª coluna, da esquerda para a direita, de cima para baixo. Para a seleção do 2º estrato, comece a leitura da TNA da décima terceira linha, 1ª coluna, da esquerda para a direita, de cima para baixo. A lista dos funcionários encontra-se a seguir:

1 ALDOBERTON	16 ERNANE	31 LIVIA
2 ALEX	17 FABIANE	32 LUCIANA
3 ALINE	18 FELIPE	33 LUCYANA
4 ALINY	19 FELIPE PAULO	34 MARCELA
5 ANDERSON	20 FLAVIA	35 MARIA
6 BERENICE	21 FRANLIANE	36 MARIANA
7 BETHANIA	22 GENIANE	37 MARIANA LÚCIA
8 BRUNA	23 GREICIANE	38 MERCIA
9 BRUNO	24 HUARTALOS	39 NATALIA
10 CAMILA	25 KAMILA	40 OLIVIA
11 CAROLINA	26 KARINA	41 POLIANA
12 CHALINE	27 KEMMUEL	42 PRISCILLA
13 CHRISTIANE	28 KEULA	43 RAISSA
14 CRISTIANE	29 LEANDRO	44 SIMONE
15 DILIANE	30 LEIDILAINE	45 TALITA
		46 THAMILLES

1.4.3 Amostragem sistemática

- A amostra é escolhida de acordo com uma periodicidade de valores.
- O processo consiste dos seguintes passos:
 - ✓ Calcula-se o intervalo de seleção (k) obtido dividindo-se o tamanho da população (N) pelo tamanho da amostra (n), isto é, $k = N / n$.
 - ✓ Sorteia-se o primeiro elemento do conjunto $\{1, 2, \dots, k\}$, podendo-se utilizar a tabela de números aleatórios.
 - ✓ Partindo-se do primeiro elemento, seleciona-se todo k -ésimo elemento até que se consiga obter o tamanho da amostra que foi escolhido.

- Uma das vantagens da amostragem sistemática é a sua facilidade de uso.
- Esse tipo de amostragem deve ser evitado no caso de ocorrer regularmente um padrão nos dados.
 - ✓ Por exemplo, se inspecionássemos cada 40^a peça produzida por determinada máquina, os resultados seriam enganosos se, em virtude de uma falha regularmente recorrente, cada 10^a peça produzida pela máquina apresentasse defeito.

Importante

Se o tamanho da população é desconhecido, não podemos determinar exatamente o valor de k . Escolhemos intuitivamente um valor razoável para k .

Exemplo 3: Amostragem sistemática

Utilizando os mesmos funcionários do exemplo 1. Selecione uma amostra sistemática de 11 funcionários. Utilize a tabela de números aleatórios para escolher o 1º valor da amostra e, para isso, inicie a leitura pela 1ª linha e última coluna. A lista com os funcionários encontra-se a seguir:

1 Aristóteles	8 Anastácia	15 Arnaldo	21 Bartolomeu	27 Bernardino
2 Cardoso	9 Carlito	16 Cláudio	22 Ermílio	28 Ercílio
3 Ernestino	10 Endevaldo	17 Francisco	23 Felício	29 Fabrício
4 Geraldo	11 Gabriel	18 Getúlio	24 Hiraldo	30 João da Silva
5 Joana	12 Joaquim	19 Joaquina	25 José da Silva	31 José de Souza
6 Josefa	13 Josefina	20 Maria José	26 Maria Cristina	32 Mauro
7 Paula	14 Paulo César			

- A população tem tamanho 32 ($N=32$) e a amostra deve ser composta de 11 funcionários então, $k = 32 / 11 = 2,909$.
- Nesse caso, vamos utilizar $k = 3$ (utilizando a regra de arredondamento).
- Procedendo com a escolha do 1º valor na tabela de números aleatórios, teríamos o número 2.
- Dessa maneira:

Números escolhidos: 2, 5, 8, 11, 14, 17, 20, 23, 26, 29, 32.

Amostra escolhida: {Cardoso, Joana, Anastácia, Gabriel, Paulo César, Francisco, Maria José, Felício, Maria Cristina, Fabrício, Mauro}

Exercício 3: Amostragem sistemática

Utilizando-se os mesmos dados do exercício 1. Selecione uma amostra sistemática de 15 funcionários para participarem da pesquisa de avaliação da satisfação. Escolha o primeiro número utilizando a tabela de números aleatórios iniciando a leitura da última linha e última coluna.

1 ALDOBERTON	16 ERNANE	31 LIVIA
2 ALEX	17 FABIANE	32 LUCIANA
3 ALINE	18 FELIPE	33 LUCYANA
4 ALINY	19 FELIPE PAULO	34 MARCELA
5 ANDERSON	20 FLAVIA	35 MARIA
6 BERENICE	21 FRANLIANE	36 MARIANA
7 BETHANIA	22 GENIANE	37 MARIANA LÚCIA
8 BRUNA	23 GREICIANE	38 MERCIA
9 BRUNO	24 HUARTALOS	39 NATALIA
10 CAMILA	25 KAMILA	40 OLIVIA
11 CAROLINA	26 KARINA	41 POLIANA
12 CHALINE	27 KEMMUEL	42 PRISCILLA
13 CHRISTIANE	28 KEULA	43 RAISSA
14 CRISTIANE	29 LEANDRO	44 SIMONE
15 DILIANE	30 LEIDILAINE	45 TALITA
		46 THAMILLES