Trabalho 1. SME0807 Tecnicas de Amostragrem

Parte 1. Resolver os seguintes exercícios Capitulo 1 de Bolfarine e Bussab (2005)

1.1 Apresente uma questão ligada à sua área de interesse e que poderia ser respondida por um levantamento amostral. Aproveite para definir claramente quais seriam os seguintes conceitos na sua pesquisa:

a. unidade de pesquisa;

b. população;

c. instrumento de coleta de dados;

d. unidade respondente;

e. possível sistema de referência;

f. unidade amostral mais provável;

g. unidades amostrais alternativas.

Discuta também como você fixaria o tamanho da amostra a outros tópicos que achar relevantes.

1.10 Um pesquisador pretende estimar o consumo médio de água por domicílio em uma cidade. Discuta as vantagens e desvantagens em usar as seguintes UPA's:

a. Unidade domiciliar;

b. Blocos de domicílios: casa, prédio de apartamentos, vilas, etc;

c. Quarteirões.

Parte 2. Revisar os segintes exemplos do Capitulo 2 Bolfarine e Bussab (2005)

Exemplo 2.1 Considere a população formada por três domicílios $\mathcal{U} = \{1, 2, 3\}$ e que estão sendo observadas as seguintes variáveis: nome (do chefe), sexo, idade, fumante ou não, renda bruta familiar e número de trabalhadores. A população está descrita na Tabela 2.1.
Portanto, para os dados descritos na Tabela 2.1, os seguintes parâmetros populacionais podem ser definidos:

i. para a variável idade,

$$\mathbf{D} = (20, 30, 40) = \mathbf{Y};$$

Tabela 2.1: População de três domicílios

Variável	Valores			Notação
unidade	1	2	3	i
nome do chefe	Ada	Beto	Ema	A_i
sexo[1]	0	1	0	X_i
idade	20	30	40	Y_i
fumante[2]	0	1	1	G_i
renda bruta familiar	12	30	18	F_i
nº de trabalhadores	1	3	2	T_i

[1] 0: feminino; 1: masculino.

[2] 0: não fumante; 1: fumante.

ii. para o vetor $(F_i, T_i)'$,

$$\mathbf{D} = \begin{pmatrix} 12 & 30 & 18 \\ 1 & 3 & 2 \end{pmatrix}.$$

Com relação à funções paramétricas populacionais, tem-se:

i. idade média,

$$\theta(\mathbf{Y}) = \theta(\mathbf{D}) = \frac{20 + 30 + 40}{3} = 30;$$

ii. média das variáveis renda e número de trabalhadores,

$$\theta(\mathbf{D}) = \begin{pmatrix} \frac{12+30+18}{3} \\ \frac{1+3+2}{3} \end{pmatrix} = \begin{pmatrix} 20 \\ 2 \end{pmatrix};$$

iii. renda média por trabalhador,

$$\theta(\mathbf{D}) = \frac{12 + 30 + 18}{1 + 3 + 2} = 10.$$

Exemplo 2.2 Seja $\mathcal{U} = \{1, 2, 3\}$. Os vetores $\mathbf{s}_1 = (1, 2)$, $\mathbf{s}_2 = (2, 1)$, $\mathbf{s}_3 = (1, 1, 3)$, $\mathbf{s}_4 = (3)$ e $\mathbf{s}_5 = (2, 2, 1, 3, 2)$ são exemplos de amostras ordenadas de $\mathcal{U}$.

Exemplo 2.3 Usando as amostras do Exemplo 2.2, tem-se para a variável freqüência f que $f_1(\mathbf{s}_1) = 1$, $f_2(\mathbf{s}_1) = 1$, $f_3(\mathbf{s}_1) = 0$, $f_1(\mathbf{s}_5) = 1$, $f_2(\mathbf{s}_5) = 3$ e $f_3(\mathbf{s}_5) = 1$. Com relação a variável presença δ, temos, por exemplo, que $\delta_1(\mathbf{s}_1) = 1$, $\delta_2(\mathbf{s}_1) = 1$, $\delta_3(\mathbf{s}_1) = 0$, e $\delta_1(\mathbf{s}_5) = 1$, $\delta_2(\mathbf{s}_5) = 1$, $\delta_3(\mathbf{s}_5) = 1$.

Exemplo 2.4 Usando os dados do Exemplo 2.3, observa-se que:

$$n(\mathbf{s}_1) = 1 + 1 + 0 = 2, \quad \text{enquanto} \quad \text{que} \quad \nu(\mathbf{s}_1) = 1 + 1 + 0 = 2.$$

Também,

$$n(\mathbf{s}_5) = 1 + 3 + 1 = 5 \quad \text{enquanto} \quad \text{que} \quad \nu(\mathbf{s}_5) = 1 + 1 + 1 = 3.$$

Verifique que: $n(\mathbf{s}_2) = 2$ e $\nu(\mathbf{s}_2) = 2$, enquanto que $n(\mathbf{s}_4) = 1$ e $\nu(\mathbf{s}_4) = 1$.

Exemplo 2.5 (Continuação do Exemplo 2.4) Como $\mathcal{U} = \{1, 2, 3\}$, então:

$$\mathcal{S}(\mathcal{U}) = \{(1), (2), (3), (1,1), (1,2), (1,3), (2,1), \ldots, (2,2,1,3,2), \ldots\}$$

e

$$\mathcal{S}_2(\mathcal{U}) = \{(1,1), (1,2), (1,3), (2,1), (2,2), (2,3), (3,1), (3,2), (3,3)\}.$$

Quando não houver dúvidas em relação ao universo, usa-se a notação simplificada:

$$\mathcal{S} = \{1, 2, 3, 11, 12, 13, 21, \ldots, 22132, \ldots\}$$

e

$$\mathcal{S}_2 = \{11, 12, 13, 21, 22, 23, 31, 32, 33\}.$$

Exemplo 2.6 Considere $\mathcal{U} = \{1, 2, 3\}$ e o respectivo $\mathcal{S}(\mathcal{U})$ construído no Exemplo 2.5. Considere os seguintes exemplos de planejamentos amostrais:

- **Plano A**,

$$\begin{aligned}
&P(11) = P(12) = P(13) = 1/9 \\
&P(21) = P(22) = P(23) = 1/9 \\
&P(31) = P(32) = P(33) = 1/9 \\
&P(\mathbf{s}) = 0, \text{ para as demais } \mathbf{s} \in \mathcal{S};
\end{aligned}$$

- **Plano B**,

$$\begin{aligned}
&P(12) = P(13) = P(21) = P(23) = P(31) = P(32) = 1/6 \\
&P(\mathbf{s}) = 0, \text{ para as demais } \mathbf{s} \in \mathcal{S};
\end{aligned}$$

- **Plano C**,

$$\begin{aligned}
&P(2) = 1/3 \\
&P(12) = P(32) = 1/9 \\
&P(112) = P(132) = P(332) = P(312) = 1/27 \\
&P(111) = P(113) = P(131) = P(311) = 1/27 \\
&P(133) = P(313) = P(331) = P(333) = 1/27 \\
&P(\mathbf{s}) = 0, \text{ para as demais } \mathbf{s} \in \mathcal{S};
\end{aligned}$$

- **Plano D**,

$$\begin{aligned}
&P(12) = 1/10 \qquad && P(21) = 1/6 \\
&P(13) = 1/15 && P(31) = 1/12 \\
&P(23) = 1/3 && P(32) = 1/4 \\
&P(\mathbf{s}) = 0, \text{ para as demais } \mathbf{s} \in \mathcal{S};
\end{aligned}$$

- **Plano E**,

$$\begin{aligned}
&P(12) = P(32) = 1/2 \\
&P(\mathbf{s}) = 0, \text{ para as demais } \mathbf{s} \in \mathcal{S}.
\end{aligned}$$

Exemplo 2.7 Seja $\mathcal{U} = \{1, 2, 3\}$, como no Exemplo 2.6, e a seguinte regra de sorteio:

i. Sorteia-se com igual probabilidade um elemento de $\mathcal{U}$, e anota-se a unidade sorteada;

ii. Este elemento é devolvido à população e sorteia-se um segundo elemento do mesmo modo.

Com estas regras, a probabilidade de ocorrer a amostra 11, será

$$\begin{aligned} P(11) &= P(1 \text{ no } 1^{\text{o}} \text{ sorteio})P(1 \text{ no } 2^{\text{o}} \text{ sorteio}|1 \text{ no } 1^{\text{o}} \text{ sorteio}) \\ &= \frac{1}{3} \times \frac{1}{3} = \frac{1}{9}. \end{aligned}$$

De modo análogo, conclui-se que só terão probabilidades não nulas, as amostras de $\mathcal{S}_2$, isto é,

$$\mathcal{S}_2 = \{11, 12, 13, 21, 22, 23, 31, 32, 33\}.$$

Quanto ao planejamento amostral, este será dado por

$$P(\mathbf{s}) = \begin{cases} 1/9, & \text{se } i \in \mathbf{s} \\ 0, & \text{se } i \notin \mathbf{s} \end{cases}.$$

Observe que este é o mesmo plano amostral descrito no Exemplo 2.6, plano A. Incidentalmente, este plano amostral, um dos mais simples, é conhecido como Amostragem Aleatória Simples, com reposição, e será estudado detalhadamente no próximo capítulo.

Parte 3. Resolver os seguintes exercícios Capitulo 2 de Bolfarine e Bussab (2005)

2.1 Usando um pacote computacional conveniente, simule uma população de tamanho $N = 100$, onde a característica de interesse Y é gerada a partir da distribuição normal com média 50 e variância 16, que denotamos por $N(50, 16)$. Encontre o total τ, a média populacional μ e a variância populacional S^2, da população que foi simulada.

2.2 Considere a população dada na Tabela 2.8, onde X denota o número de apartamentos nos condomínios observados e Y denota o número de apartamentos alugados. Os espaços em branco devem ser interpretados como zero. Encontre:

a. μ_Y, τ_Y e S_Y^2;

b. μ_X, τ_X e S_X^2;

c. a proporção P de condomínios com mais de 20 apartamentos alugados e a variância populacional correspondente a variável W_i que assume o valor 1 se o i-ésimo condomínio possui mais que 20 apartamentos alugados e 0 caso contrário, $i = 1, \ldots, 180$.

2.3 Para cada um dos planos amostrais B, C, D e E do Exemplo 2.6:

a. Construa as distribuições das variáveis f_i e δ_i;

b. Calcule $E[\delta_i]$ e $Var[\delta_i]$;

c. Encontre π_i e π_{ij}, para todo i e j.

2.4 Para os planos amostrais B, C, D e E definidos no Exemplo 2.6, calcule $EQM[r]$, $Cov\left[r, \overline{f}\right]$ e $Corr\left[r, \overline{f}\right]$, usando os dados do Exemplo 2.10.

Tabela 2.8: População de 180 domicílios

i	Y_i	X_i	i	Y_i	X_i	i	Y_i	X_i	i	Y_i	X_i	i	Y_i	X_i	i	Y_i	X_i
1	19	23	31	47	53	61	67	110	91	34	48	121	1	3	151	6	37
2	17	18	32	27	28	62	44	57	92	13	24	122	22	37	152	4	11
3	25	33	33	80	90	63	43	81	93	16	27	123	25	30	153	9	24
4	84	89	34	52	68	64	15	23	94	21	32	124	2	3	154	54	102
5	91	114	35	90	99	65	17	25	95	12	14	125	4	4	155	50	82
6	48	66	36	78	89	66	29	59	96	10	18	126	7	13	156	9	24
7	48	61	37	46	48	67	18	27	97	50	61	127	15	24	157	6	18
8	20	25	38	35	48	68	14	22	98	58	65	128	10	19	158	5	18
9	34	46	39	59	62	69	24	29	99	17	25	129	5	17	159	1	3
10	42	58	40	27	33	70	35	44	100	41	68	130	8	13	160	1	6
11	35	44	41	33	43	71	48	53	101	3	8	131	8	18	161		1
12	55	66	42	27	37	72	20	27	102	4	12	132		1	162	2	7
13	42	61	43	9	14	73	24	28	103	18	27	133	4	10	163	2	8
14	36	45	44	9	15	74	55	62	104	1	3	134	1	4	164	3	12
15	13	20	45	12	21	75	43	56	105	1	3	135	3	9	165	1	4
16	7	16	46	49	68	76	13	22	106	3	6	136		5	166	6	8
17	8	15	47	60	81	77	19	22	107	6	14	137	14	20	167	3	9
18	18	26	48	35	59	78	48	57	108	5	15	138	3	5	168	3	7
19	20	22	49	11	23	79	44	57	109	5	14	139	5	13	169	5	12
20	18	22	50	21	32	80	36	46	110	4	9	140		1	170	3	10
21		2	51	22	36	81	3	8	111		1	14 1	11	23	171		1
22	23	29	52	10	16	82	2	4	112		4	142	19	39	172		1
23		3	53	9	15	83	13	18	113	7	12	143	5	9	173		1
24	19	29	54	7	16	84	34	42	114	7	22	144		2	174	2	4
25	11	21	55	3	8	85	28	32	115	3	11	145	3	5	175		1
26	11	15	56	5	25	86	23	28	116	12	27	146	12	26	176		1
27	42	54	57	2	11	87	8	14	117	11	20	147	4	10	177		2
28	28	42	58	8	9	88	69	76	118	27	38	148	14	35	178	1	1
29	8	13	59	14	19	89	2	19	119	14	31	149		4	179		1
30		2	60	5	5	90	5	9	120	2	4	150	20	38	180		1

Data de entrega do trabalho: 13 de Março