



PROVA I — MATD44

23 DE OUTUBRO DE 2023

Professor: Raydonal Ospina**E-mail:** raydonal@castlab.org**Site do curso:** <https://castlab.org/courses/matd44/>**Regras:** Leia com atenção as perguntas. **Todas as questões devem ser detalhadas.** A prova deve ser claramente resolvida. Seja claro e organizado.**Atenção:** Você deverá encaminhar a sua prova resolvida digitalizada em formato **PDF** no e-mail acima no dia 7/11/2023 até as 15:00h (GMT-3 - Horário de Brasília)

I ▶ **(Amostragem Bernoulli)** Seja s uma amostra obtida de um desenho amostral de tipo Bernoulli com probabilidades de inclusão $\pi_k = \pi$ para todo $k \in U$ (população). Seja n_s o tamanho de amostra da amostra s . Mostre que a probabilidade condicional de se obter s dado n_s é a mesma que a probabilidade obtida por uma amostragem aleatória simples sem substituição de tamanho fixado n_s de N (Tamanho da população).

2 ▶ **(Amostragem Bernoulli)** Na elaboração de um estudo de rede de escritório, o seguinte esquema de amostragem sequencial foi proposto para selecionar uma amostra aleatória de dois intervalos de horários de escritório não adjacentes dentre os cinco intervalos 9-10, 10-11, 11-12, 12-13, 13-14 (rotulados de 1 a 5). Seleciona-se o primeiro intervalo de uma hora para a amostra com igual probabilidade a partir dos cinco intervalos. Em seguida, seleciona-se o segundo intervalo de uma hora, sem reposição, com igual probabilidade dos intervalos que não são adjacentes no tempo ao primeiro intervalo selecionado no primeiro sorteio.

- Determine as probabilidades de inclusão de primeira ordem.
- Determine as probabilidades de inclusão de segunda ordem e determine se esse plano amostral induzido pelo esquema de amostragem proposto é mensurável?

- Determine as covariâncias dos indicadores de pertencimento à amostra.

3 ▶ **(Amostragem estratificada)** A Tabela 1 contem as informações do gasto mensal em serviços públicos de uma amostra aleatória estratificada de 120 famílias na cidade de Salvador a qual foi geograficamente dividida em três estratos: Norte, Centro e Sul.

Estatísticas	Estratos		
	Norte (1)	Centro (2)	Sul (3)
N_h	4.000	6.000	10.000
W_h	0,3	0,2	0,5
n_h	40	36	44
\bar{y}_h	1,2	2,4	0,6
\bar{Y}_h	9600	7200	6000
s_h^2	0,36	1,21	0,04
$\text{Var}(\bar{y}_h)$	0,000993	0,004404	0,000226

Tabela 1: Informações do gasto familiar mensal em serviços públicos (em salários mínimos) a partir de uma amostra aleatória simples estratificada na cidade de Salvador.

- Estime o gasto médio e o gasto total de toda a população. Estabeleça um intervalo de confiança de 95% para essas medidas.
- Suponhamos que os custos de coletar a informação por família para cada um dos estratos são: $C_1 = \text{R\$}5.000$, $C_2 = \text{R\$}3.000$ e $C_3 = \text{R\$}1.000$, respectivamente. Assuma que existem informações prévias em relação a variabilidade de cada estrato e que estas informações correspondem as estimativas s_h^2 da tabela 1. Se o orçamento da coleta de informação não pode ser superior a R\$2,5 milhões qual deve ser o tamanho da amostra global e como ele deve ser dividido entre os diferentes estratos? Suponha que a função de custo é linear na forma

$$C = C_0 + \sum_{h=1}^H C_h n_h,$$



em que C é o orçamento total para a coleta de informação, C_0 é o custo fixo que não depende do número de unidades amostrais a serem selecionadas e C_h é o custo de amostrar uma unidade amostral no estrato h .

4 ▶ **(Estimação em subpopulações finitas (domínios))** Uma amostra aleatória simples e sem substituição de 56 pessoas foi selecionada de uma população de 1000 trabalhadores da empresa LINCATECH. Foram coletadas informações sobre a renda mensal em miles de reais (Renda) e o sexo (Sexo) do trabalhador. Com as informações da Tabela 2 estime:

- A renda média dos trabalhadores. Estabeleça um intervalo de confiança de 95% para a renda média.
- A renda total dos trabalhadores. Estabeleça um intervalo de confiança de 95% para a renda total.
- Estime a proporção e o número total de mulheres na empresa. Estabeleça intervalos de confiança de 95% para a proporção e total de mulheres na empresa.
- Podem ser consideradas válidas aproximações pela distribuição normal no item anterior? Explique.
- Considera que as amostras, tanto de homens como de mulheres poderiam ser assumidas como amostras aleatórias simples e sem substituição das respectivas subpopulações de homens e mulheres da empresa? Explique.
- Como poderia ser estimada a renda média e o total das mulheres para toda a empresa se não se conhecesse o número total delas?
- Qual das duas subpopulações (homens, mulheres) é mais homogênea em relação a renda?

ID	Sexo	Renda	ID	Sexo	Renda
1	Fem	2094.90	15	Mas	2939.21
2	Mas	2386.14	16	Mas	1722.62
3	Mas	1562.82	17	Fem	2739.79
4	Mas	1781.41	18	Mas	1821.61
5	Mas	1603.14	19	Mas	1742.40
6	Mas	479.73	20	Mas	1845.22
7	Mas	2196.85	21	Mas	1916.60
8	Mas	2365.21	22	Mas	1329.28
9	Mas	2016.32	23	Mas	2143.04
10	Mas	1322.23	24	Mas	2618.97
11	Fem	2589.08	25	Fem	1399.85
12	Fem	2896.28	26	Mas	1610.41
13	Mas	1370.55	27	Mas	2300.84
14	Mas	975.94	28	Mas	1192.90
29	Mas	2715.62	43	Mas	1802.16
30	Mas	2042.58	44	Mas	2444.68
31	Mas	2235.73	45	Mas	2644.75
32	Mas	2223.33	46	Fem	1431.53
33	Mas	2618.16	47	Mas	1094.02
34	Mas	2206.57	48	Mas	1548.96
35	Mas	2432.08	49	Mas	2410.51
36	Mas	1340.94	50	Mas	2286.37
37	Mas	2321.37	51	Mas	1589.07
38	Mas	1922.91	52	Fem	1646.21
39	Mas	2520.83	53	Mas	3358.63
40	Mas	2063.78	54	Mas	1369.42
41	Mas	2335.66	55	Mas	2047.63
42	Mas	2357.94	56	Mas	1719.34

Tabela 2: Tabela de Informações dos empregados na amostra

BOA PROVA