UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE MATEMÁTICA
CADERNOS DE MATEMÁTICA E ESTATÍSTICA
SÉRIE H: CADERNOS DE EXERCÍCIOS

MAT02251 - AMOSTRAGEM I
LISTA DE EXERCÍCIOS E TRABALHOS

Elsa Cristina de Munsdstock

Digitação parcial: Bolsista Anderson do Couto Soares

SÉRIE H: Nº 1
Porto Alegre, março de 2000

I. EXERCÍCIOS DE AULA

DISTRIBUIÇÃO DA MÉDIA AMOSTRAL

1 Amostragem com reposição
Considere uma população hipotética formada pelos elementos 2,4,6.
a) Calcule a média μ e a variância σ^2 dessa população.
b) Selecione todas as amostras aleatórias simples com reposição de tamanho n=2 e calcule a média de cada amostra, $\bar{X}$.
c) Calcule a média $\mu_{\bar{x}}$ e a variância $\sigma^2_{\bar{x}}$ das médias amostrais.
d) Qual é a relação entre a média da população μ e a média das médias amostrais $\mu_{\bar{x}}$?
e) Qual é a relação entre a variância da população σ^2 e a variância das médias amostrais $\sigma^2_{\bar{x}}$?
f) Observe que as conclusões de a) e b) são válidas só para amostragem **com reposição**.

2 Amostragem sem reposição
Repita o exercício 1.1 para amostragem **sem reposição**.

II.TRABALHOS USANDO O COMPUTADOR

1. DISTRIBUIÇÃO DA MÉDIA AMOSTRAL

1.1 Utilizando a população em arquivo, de tamanho N=500,
a) Calcule μ e σ^2. Faça um histograma da população
b) Selecione 50 amostras aleatórias simples de tamanho n=5, com reposição. Para cada amostra calcule a média.
c) Calcule a média e a variância das 50 médias amostrais selecionadas em b. Compare com os valores indicados pela teoria.
d) Faça um histograma das 50 médias amostrais
e) Compare o histograma das médias amostrais com o histograma da população: o que você observa?
f) Selecione 50 amostras aleatórias simples de tamanho n=20, com reposição. Para cada amostra calcule a média.
g) Calcule a média e a variância das 50 médias amostrais selecionadas em f. Compare com os valores indicados pela teoria.
h) Faça um histograma das 50 médias amostrais
i) Compare o histograma das médias amostrais com o histograma da população: o que você observa?

j) Compare os histogramas das médias amostrais de tamanho n=5 com o histograma das médias amostrais de tamanho n= 20: o que você observa?
k) Extrair uma conclusão geral das observações feitas, em relação ao teorema Central do Limite.

1.2 MATERIAL A ENTREGAR:
a) Relatório;
b) Disquete comos arquivos usados, planilhas, etc;
c) Uma folha com seu nome contendo:
- histograma da população na parte superior, identificando a distribuição, a média e a variância
- histograma das médias para amostras de tamanho 5 na parte central, com $\mu_{\bar{x}}$ e $\sigma^2_{\bar{x}}$
- histograma das médias para amostras de tamanho 20 na parte inferior, com $\mu_{\bar{x}}$ e $\sigma^2_{\bar{x}}$

1.3 COMANDOS DO MINITAB
a) PARA SELECIONAR 50 AMOSTRAS PSEUDO-ALEATÓRIAS DE TAMANHO 5, CALCULAR AS MÉDIAS E COLOCÁ-LAS EM UMA COLUNA

a.1) Cria-se um arquivo chamado, por exemplo, medias, no disquete:

```
Store 'a:medias'
Sample  5 c1 c9;
Replace.
Mean c9 k1
Stack k1 c10 c10
End
```

Explicação: São selecionadas amostras de tamanho 5 da coluna c1 e colocadas em c9 (cada amostra substitui a anterior). As médias são empilhadas em c10 sucessivamente.

a.2) Executam-se os comandos do arquivo 'médias' 50 vezes

```
Execute 'a:medias' 50
```

b) PARA CALCULAR A MÉDIA E A VARIÂNCIA DAS MÉDIAS E FAZER O HISTOGRAMA DAS MÉDIAS (Supondo que as médias estão na coluna 10)

```
mean c10
stdev c10 k2
let  k3=k2**2
print k3
```

Fazer o histograma:

i) em modo gráfico de alta resolução controlando os intervalos de classe (e por tanto o número de classes):

```
Histogram c10;
CutPoint 20:80/5;
Bar.
```

Obs: CutPoint 20:80/5 significa definir os limites de classe de 20 até 80 em intervalos de 5.

ii) em modo gráfico padrão de maneira que apareçam as frequências das classes:
GSTD
Histogram c10;
Increment 5;
Start 22.5.

2. -AMOSTRAGEM ALEATÓRIA SIMPLES

Exercício com o cadastro da Produção Agrícola Municipal do RGS em 1993:
a) Calcular média, variância e coeficiente de correlação para as variáveis quantitativas contínuas do cadastro, e a proporção de área plantada com arroz.
b) Calcular o tamanho da amostra necessário para estimar a média com um erro relativo não maior que 5% e coeficiente de confiança 95% para cada variável do item a). Decidir um tamanho de amostra para a "pesquisa". Justifique.
c) Selecionar uma amostra aleatória simples do tamanho escolhido em b).
d) Estimar a média de cada variável e a proporção de área plantada com arroz, por ponto e por intervalo, com um coeficiente de 99% de confiança.
e) Calcule o erro máximo esperado com n = 200 para cada variável e compare com o erro fixado em b). Explique as diferenças.
f) O que você pode afirmar em relação à confiança nas estimativas para cada variável?

INSTRUÇÕES PARA RESOLVER ESTE EXERCÍCIO USANDO SPSS
Primeiro, ler o arquivo de dados do cadastro no SPSS (Os arquivos .sav são arquivos de dados do SPSS)

a) Calcular média e variância da população no menu
STATISTICS / SUMMARIZE / DESCRITIVES/Variable(s):<selecione aqui as variávies quantitativas>/OPTIONS/ mean; std. deviation/Continue/OK

Calcular a proporção de área plantada com arroz pelo menu:
STATISTICS / SUMMARIZE / FREQUENCIES/produto/OK

b) Calcular o tamanho da amostra manualmente ou usando
TRANSFORM / COMPUTE/Target Variable=<nome da variável nova>/Numeric Expression:<fórmula para o cálculo da variável nova>/OK

c) Selecionar uma amostra aleatória simples de tamanho n=1000 no menu
DATA / SELECT CASES / RANDOM SAMPLE OF CASES
Dentro desse menu entrar em SAMPLE e marcar

- EXACTLY 1000 CASES FROM THE FIRST 11327(tamanho da população) CASES

Observe que os casos não selecionados ficam riscados na planilha de dados.

d) Para calcular a média e variância amostrais, entre de novo em
STATISTICS / SUMMARIZE / DESCRITIVES
O SPSS utilizará as observações não riscadas
O intervalo de confiança pode ser calculado manualmente ou usando
TRANSFORM / COMPUTE

e) Calcular erro máximo esperado manualmente ou pelo menu
TRANSFORM / COMPUTE

- OBS: NÃO LISTAR A POPULAÇÃO NO RELATÓRIO, NEM A AMOSTRA!!!

3. AMOSTRAGEM ESTRATIFICADA

1) Dividir o cadastro em 3 estratos pela variável Área Colhida, usando o critério de $\sqrt{f(x)}$
Calcular σ^2, σ_E^2 e σ_D^2
2) Repetir 1) para 4, 5, 6 e 7 estratos
3) Analisar o comportamento de σ_E^2 e σ_D^2 para diferentes números de estratos (Fazer gráfico com Nº de estratos no eixo X e σ_D^2 no eixo Y)
4) Utilizando 6 estratos e n=1000,

4a) Calcular o tamanho da amostra nos estratos pela distribuição proporcional e pela de Neymann. Comparar os dois métodos (DICA: Use $\sigma_{\bar{x}}^2$)
4b) Selecionar a amostra estratificada proporcional de n=250, estimar a média de área plantada, área colhida, quantidade, rendimento e preço por ponto e com IC95%. Estimar a proporção de área plantada com arroz por ponto e com IC95%

PARA TRABALHAR NO SPSS:

1) Criar uma variável ESTRATO para identificar os estratos através de
TRANSFORM / COMPUTE / IF (uma vez para cada estrato)
ou RECODE INTO DIFFERENT VARIABLE

2) Para calcular as variâncias ENTRE e DENTRO (Suponha que ARC= Área Colhida), usar
DATA / AGGREGATE / Break variable ESTRATO,
Aggregate variables :

Name N, Function N(ARC)
Name MEDARC , Function MEAN (ARC)
Name DPARC, Function SD(ARC)
etc (REPETE PARA CADA VARIÁVEL)

Explicação: Com este comando se cria um arquivo novo (com nome que será dado em File...) com a estrutura de uma linha por estrato. Nessa linha teremos o tamanho do estrato, a média, desvio padrão, etc., ou seja, as variáveis que foram criadas. Alerta: Aqui você já pode incluir a variável Área Plantada com arroz também

3) Use PASTE para salvar os comandos em arquivo e poder repeti-los para 4, 5 6 e 7estratos.

4)Para selecionar a amostra pela distribuição proporcional (Por exemplo, 20% em cada estrato) ou pela distribuição de Neyman, faça uma seleção independente em cada estrato. Volte a usar AGGREGATE para calcular as estimativas em base à amostra.

4. NORMAS DE APRESENTAÇÃO DOS RELATÓRIOS

ROTEIRO SUGERIDO:
- identificação (disciplina, nome e número do trabalho, nome do aluno, semestre, data, etc);
- objetivos do trabalho a ser realizado;
- metodologia;
- resultados;
- conclusões, comentários sobre os resultados obtidos;
- software utilizado;
- bibliografia utilizada;
- anexos (pode ser os comandos utilizados, etc)

APRESENTAÇÃO
O relatório deve preencher os seguintes requisitos:
- conter todas as informações de maneira suficientemente clara e simples para que possa ser entendido por alguém que não é da área de estatística (cliente);
- conter todas as informações técnicas necessárias para que uma pessoa da área de estatística possa acompanhar o que foi feito;
- ser o mais claro e simples possível.

AVALIAÇÃO
Na avaliação do relatório serão levados em conta os seguintes aspectos:
- apresentação;
- métodos utilizados;
- resultados: corretos ou não;
- clareza;
- se está completo ou não;
- data de entrega.

III. LISTA DE EXERCÍCIOS

1. CONCEITOS BÁSICOS

1.1) Mostre a diferença entre população objetivo e população amostrada. Explique em que caso os estimadores obtidos para uma população amostrada podem ser considerados válidos para a população objetivo.

1.2) Uma pessoa selecionou uma amostra de 30 floriculturas na cidade para estimar: valor das vendas, quantidade de flores vendidas, preferência dos clientes por determinadas cores e consumo médio por cliente. As floriculturas foram selecionadas sem o uso de qualquer procedimento probabilístico (tabela de números aleatórios, sorteio, etc.), mas foram incluídas floriculturas de diferentes tamanhos (pequenas, médias e grandes), e de todos os bairros da cidade. A pesquisa foi levantada, os dados foram tabulados e suponha que agora a pessoa responsável pela pesquisa procura você para orientá-la no cálculo de estimativas por ponto e por intervalo para médias, totais e proporções das variáveis pesquisadas. Explique detalhadamente a orientação que você daria a essa pessoa (não inclua fórmulas). Justifique amplamente a sua resposta.

1.3) Defina amostragem probabilística e amostragem não probabilística. Descreva uma situação prática na qual a amostragem probabilística é mais apropriada para selecionar uma amostra. Idem para amostragem não probabilística.

1.4) Um agrônomo plantou milho em parcelas (retângulos) com 5 linhas de plantas cada parcela. Para estudar características morfológicas da planta ele descartou as 2 linhas externas de cada parcela e fez uma seleção nas 3 linhas internas da seguinte maneira: tirou um número entre 1 e 6 da tabela de números aleatórios. Suponha que foi 4. Selecionou a 4ª planta da linha 2. A partir daí, selecionou de 6 em 6 plantas, ou seja, a nº 10, 16, 22 etc. Quando esgotou a 2ª linha passou para a 3ª e depois para a 4ª linha.
Esse método de seleção é probabilístico? Justifique.

1.5) Para estimar o número de leucócitos, taxa de ferro e nível de hemoglobina no sangue, um laboratório tirou uma amostra de aproximadamente 5 ml de sangue de uma paciente.
a) Identifique a população.
b) Essa amostra é probabilística? Justifique.
c) É representativa? Justifique.
d) Como poderia ser avaliada a confiança nas estimativas obtidas?
Explique.

2. AMOSTRAGEM ALEATÓRIA SIMPLES

2.1) Com a população 1, 3, 5, 7:
a) Selecione todas as amostras possíveis com reposição de tamanho 2;
a.1) Calcule $\overline{X}$ e s^2 para cada amostra;
a.2) Faça um histograma de $\overline{X}$;
a.3) Mostre que $\mu_{\overline{X}} = \mu$ e $\sigma_{\overline{X}} = \dfrac{\sigma}{\sqrt{n}}$.
b) Selecione todas as amostras possíveis sem reposição de tamanho 2;
b.1) Calcule $\overline{X}$ e s para cada amostra;
b.2) Faça o histograma de $\overline{X}$;
b.3) Mostre que $\mu_{\overline{X}} = \mu$ e $\sigma_{\overline{X}} = \sqrt{\dfrac{N-n}{N-1}}\dfrac{\sigma}{\sqrt{n}}$.

2.2) Considere a população: {1, 3, 5, 5, 7}. Na tabela a seguir encontra-se a distribuição das amostras possíveis, com reposição, de tamanho n = 3.

Tipo de Amostra:	111	113	115	117	133	135	137	155	157	177
Freq.(prob. x 125):	1	3	6	3	3	12	6	12	12	3
Tipo de Amostra:	333	335	337	355	357	377	555	557	577	777
Freq.(prob. x 125)	1	6	3	12	12	3	8	12	6	1

Total de amostras = 125

a) Para cada uma das amostras possíveis calcule $\overline{X}$ e s^2.

b) Calcule $\mu_{\overline{X}}$ e $\sigma_{\overline{X}}$.

c) Ache as percentagens de médias que estão nos intervalos;
$(\mu_{\overline{X}} \pm \sigma_{\overline{X}})$, $(\mu_{\overline{X}} \pm 2\sigma_{\overline{X}})$, $(\mu_{\overline{X}} \pm 3\sigma_{\overline{X}})$

d) Explique o que representam os resultados em c.

e) Calcule todos os intervalos de confiança 95% para a média da população.

f) Ache a proporção de intervalos que contém o parâmetro estimado.

g) Explique os resultados obtidos em f.

2.3) Para uma população de tamanho N, mostre que $\sum_{i=1}^{N}(X_i - \mu) = 0$

2.4) Defina amostragem aleatória simples. Explique como se seleciona uma amostra aleatória simples na prática.

2.5) Mostre que, na amostragem sem reposição,

$$E(s^2) = S^2 \qquad s^2 = \frac{\sum_{i}^{n}(X_i - \overline{X})^2}{n-1}$$

2.6) Ache um estimador não viciado para σ^2 na amostragem sem reposição.

2.7) Mostre algebricamente que, na amostragem sem reposição (populações finitas), a variância da média amostral tem a seguinte expressão:

$$\sigma^2_{\overline{X}} = \frac{N-n}{N}\frac{S^2}{n}$$

Mostre um estimador não viciado para essa variância.

2.8) Numa população com N = 6, temos X = (8, 2, 2, 11, 4, 7). Para amostras aleatórias simples de tamanho n = 2 sem reposição:

a) Ache todas as amostras possíveis e calcule $\overline{X}$ para cada amostra.

b) Mostre que $E(\overline{X}) = \mu$.

c) Mostre que $\sigma^2_{\overline{X}}$ é dada por $(1-f)\frac{S^2}{n}$, $f = \frac{n}{N}$.

2.9) Dois dentistas, D1 e D2 fazem uma pesquisa por amostragem sobre o estado dos dentes das 200 crianças de certa escola estadual de determinada localidade.

O Dr. D1 selecionou uma amostra aleatória simples de 20 crianças e conta o número de dentes cariados para cada criança com os seguintes resultados;

- Número de dentes cariados:	0	1	2	3	4	5	6	7	8	9	10
- Número de crianças:	8	4	2	2	1	1	0	0	0	1	1

O Dr. D2, usando o mesma técnica dental, examina as 200 crianças da escola, mas anota somente o número de crianças com dentes cariados. Ele encontra um total de 60 crianças sem nenhuma cárie. Estime o número de dentes cariados nas crianças da escola, quando se utiliza:

a) Somente os resultados do Dr. D1;

b) Os resultados do Dr. D1 e do Dr. D2;
Encontre a variância dos estimativas obtidas em a e b. Compare a precisão dos dois métodos.

2.10) A tabela abaixo apresenta as populações dos 30 maiores municípios do Brasil, segundo o censo de 1980 (em 10.000 habitantes):

MUNICÍPIO	POPULAÇÃO
1- SÃO PAULO (SP)	849,3
2- RIO DE JANEIRO (RJ)	509,3
3- BELO HORIZONTE (MG)	178,1
4- SALVADOR (BA)	150,6
5- FORTALEZA (CE)	130,8
6- RECIFE (PE)	120,4
7- BRASÍLIA (DF)	117,7
8- PORTO ALEGRE (RS)	112,5
9- NOVA IGUAÇU (RJ)	109,4
10- CURITIBA (PR)	102,5
11- BELEM (PA)	93,4
12- GOIÂNA (GO)	71,7
13- CAMPINAS (SP)	66,4
14- MANAUS (AM)	63,4
15- SÃO GONÇALO (RJ)	61,4
16- DUQUE DE CAXIAS (RJ)	57,5
17- SANTO ANDRÉ (SP)	55,2
18- GUARULHOS (SP)	53,2
19- OSASCO (SP)	47,3
20- SÃO LUIS (MA)	44,9
21- S. BERNARDO DO CAMPO (SP)	42,5
22- NATAL (RN)	41,7
23- SANTOS (SP)	41,6
24- NITERÓI (RJ)	40,1
25- MACEIÓ (AL)	40,0
26- SÃO JOÃO DO MERITI (RJ)	39,8
27- TERESINA (PI)	37,8
28- CAMPOS (RJ)	34,9
29- JABOATÃO (PE)	33,1
30- JOÃO PESSOA (PB)	33,0

a) Encontre T, μ e S^2 para a população acima.

b) Selecione uma amostra aleatória simples de n = 10 municípios e calcule estimadores para μ e T (com as respectivas estimativas das variâncias dos estimadores).

c) Compare as estimativas das variâncias dos estimadores com as verdadeiras variâncias (calculadas para a população toda).

2.11) Considere uma população de tamanho N=10 onde a variável de interesse toma os valores: 8, 15, 12, 6, 9, 10, 15, 6, 8, 10.

a) Calcule os parâmetros dessa população: μ, σ^2 e S^2

b) Calcule $\sigma^2_{\bar{X}}$ para amostras sem reposição de tamanho n= 4.

c) Selecione uma amostra aleatória simples sem reposição de tamanho n=4 e calcule $\overline{X}$, s^2 e $\hat{\sigma}^2_{\bar{X}}$.

d) Compare os valores dos parâmetros e suas correspondentes estimativas. Comente.
e) Ache um IC 95% para μ usando a amostra selecionada. Interprete-o.
f) O intervalo obtido contem o parâmetro estimado?
Em face dessa constatação como você explica a interpretação dada em e?
g) Repita os itens B até F para amostras com reposição.

2.12) Como você explicaria a um produtor agrícola, o que significa dizer que uma amostra aleatória simples de 20 propriedades agrícolas forneceu o seguinte intervalo de confiança de 99% para a produção total de trigo do estado na próxima safra: 2,5 a 2,8 milhões de toneladas.

2.13) Considere uma amostragem aleatória simples com reposição de tamanho n=2 da população do exercício 2.9.
a) Encontre a distribuição de $\overline{X}$. Mostre que $E(\overline{X}) = \mu$.
b) Encontre $\sigma^2_{\overline{X}}$.
c) Suponha que uma amostra aleatória simples com reposição de tamanho n = 10 retirada da população apresenta $\overline{X}$=5,435 e s^2=3,6. Encontre um intervalo de confiança para μ com $\alpha = 0{,}02$.

2.14) Uma pesquisa foi conduzida com o objetivo de se estudar o índice de ausência ao trabalho em uma determinada indústria. Da população de 36.000 operários foi selecionada uma amostra aleatória simples sem reposição de 1000. Na amostra foi registrado o número de faltas não justificadas num período de 6 meses. Os resultados obtidos foram:

Faltas:	0	1	2	3	4	5	6	7	8
Número de trabalhadores:	451	162	187	112	49	21	5	11	2

a) Calcule a estimativa por ponto de μ fornecida por essa amostra.
b) Calcule s^2 .
c) Ache um intervalo 95% para μ .
d) Usando o intervalo para μ, calcule um intervalo 95% para T.
e) Interprete os intervalos obtidos.

2.15) Uma amostra aleatória simples sem reposição com n = 30 foi observada em uma área da cidade contendo 14848 residências. O número de pessoas por residência na amostra observada foi: 5, 6, 3, 3, 2, 3, 3, 3, 4, 4, 3, 2, 7, 4, 3, 5, 4, 4, 3, 3, 4, 3, 3, 1, 2, 4, 3, 4, 2, 4.
a) Encontre uma estimativa por ponto do número médio de pessoas por residência na população e uma estimativa para a variância da média.
b) Encontre um intervalo de 90% de confiança para μ . Interprete-o.
c) Suponha que seja de interesse uma estimativa duas vezes mais precisa que a obtida com a amostra acima. Qual o tamanho da amostra necessário para a tal precisão?

2.16) Na tabela a seguir, considere a população dada, 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.

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	14	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	5	14	140	5	13	170	5	12
21		2	51	22	36	81	3	8	111		1	141	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	12	27	147	12	26	177		1
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

Encontre:

a) μ_y, T_y e S_y^2; b) μ_x, T_x e S_x^2;

c) A proporção P de condomínios com mais de 20 apartamentos alugados e a variância populacional correspondente.

d) Selecione uma amostra aleatória simples de tamanho 20 e outra de tamanho 30 sem reposição e construa intervalos de confiança para μ e T com coeficiente de confiança $1-\alpha = 0{,}95$.

d) Considere a amostra de tamanho 30. Qual o tamanho necessário da amostra para que tenhamos uma estimativa duas vezes mais precisa?

2.17) Em uma amostra de 200 colégios particulares de uma população de 2.000, 120 colégios eram favoráveis a certa proposição, 57 contra e 23 eram indiferentes. Encontre o tamanho da amostra necessário para estimar o número de colégios favoráveis à proposição com um erro não maior do que 20 colégios, com probabilidade igual a 0,95. Justifique o procedimento utilizado.

2.18) Uma amostra piloto de tamanho 30 selecionada de uma população de 1.000 empresas estimou a porcentagem de empresas exportadoras em 34%.
a) calcule o tamanho de amostra necessário para estimar a proporção de empresas exportadoras com um erro máximo absoluto de 8% e 95% de confiança. Suponha que a nova amostra selecionada estimou a proporção em 32,5%. Calcule o intervalo de confiança para P.
b) Calcule o tamanho de amostra necessário para estimar a proporção de empresas exportadoras com um erro máximo relativo de 8% e 95% de confiança. Suponha que a amostra selecionada com esse tamanho estimou o proporção em 35,5%. Calcule o intervalo de confiança para P.
c) Comente os resultados obtidos em a) e b).

2.19) Considere uma população com N=6, onde D'= (0, 0, 1, 1, 1, 1). Queremos estimar P (proporção de uns na população) utilizando uma amostragem aleatória simples sem reposição de n = 4 unidades.
a) Encontre a distribuição da média amostral $\overline{X}$ e mostre que $\overline{X}$ é um estimador não viciado de P.
b) Sugira um estimador para $\sigma^2_{\overline{X}}$. Verifique se seu estimador é não viciado.

2.20) A Receita Federal deseja estimar a sonegação total do ICM entre 10.000 contribuintes de determinada região. Admite-se um erro relativo de 2,5% e uma confiança de 95%. Experiências anteriores indicam que o coeficiente de variação das sonegações é aproximadamente 50%. Que tamanho de amostra deve ser utilizado para uma amostra aleatória simples?

2.21) Considerando as fórmulas da amostragem aleatória simples com reposição, comente qual é o efeito:
a) Do valor da variância sobre o tamanho da amostra;
b) Do valor do erro máximo admissível sobre o tamanho da amostra, para um coeficiente de confiança determinado.
c) Do tamanho da amostra sobre o tamanho do intervalo de confiança, para um determinado valor do coeficiente de confiança e variância da população conhecida.

2.22) Suponha que você quer estimar a média de uma população com um erro não maior que 5% da média verdadeira, com uma confiança de 95%, sabendo que o coeficiente de variação é 0,4. Qual é o tamanho de amostra necessário se a população é de tamanho:
a) N = 5.000;
b) N = 2.000.000.

Qual a conclusão que pode ser feita a partir dos resultados anteriores?

3. AMOSTRAGEM ESTRATIFICADA

3.1) Mostre a decomposição algébrica da variância total na amostragem estratificada e:
a) Explique o que mede cada um dos termos dessa decomposição;
b) Explique como a distribuição das unidades da população nos estratos influencia nos valores de cada um dos termos dessa decomposição;
c) Explique a importância de cada um dos termos na precisão do estimador da média (efeito deles na variância da média).

3.2) Seja a população 2, 3, 3, 4, 12, 14, 15, 15. Considere a estratificação:
Estrato 1: 2, 3, 3, 4; Estrato 2: 12, 14, 15, 15.
Calcule σ^2, μ, σ^2_E, σ^2_D:

3.3) Ache uma expressão para a probabilidade de uma unidade da população estar na amostra estratificada.

3.4) Uma pesquisa foi realizada para estimar a média de uma população. Foi selecionada uma amostra estratificada com distribuição de Neyman, e os resultados amostrais foram os seguintes:

Estrato 1: $\overline{X} = 7{,}5$ $s^2 = 7$ $n_1 = 25$ $N_1 = 300$

Estrato2: $\overline{X} = 30{,}5$ $s^2 = 18$ $n_2 = 30$ $N_2 = 600$

Ache um IC 95% para estimar a média de cada estrato e a média geral.

3.5) Foram selecionadas duas amostras probabilísticas de n = 1000 da mesma população. A primeira foi uma amostra aleatória simples que forneceu uma média de 12,3 e uma variância de 17,1. A segunda amostra foi estratificada e resultou nos seguintes valores:

ESTRATO	PONDERAÇÃO	MÉDIA	VARIÂNCIA
1	0,5	10,2	10,82
2	0,3	12,6	10,82
3	0,2	17,1	10,82

Com base nesses resultados você recomendaria a estratificação? Justifique.

3.6) Em uma pesquisa de mercado para avaliar a aceitação de um novo produto foram obtidos os seguintes resultados:

CIDADE	Nº HABITANTES	AMOSTRA	ACEITAM O PRODUTO
A	2.000.000	400	150
B	500.000	100	35
C	400.000	80	20

Calcule a estimativa por intervalo para a proporção de pessoas que aceitam o produto, com uma confiança de 90%:
a) Para cada cidade;
b) Para o conjunto das 3 cidades.
Interprete os intervalos obtidos.

3.7) Compare, algebricamente, a precisão da amostragem aleatória simples com a da amostragem estratificada com distribuição de Neyman.

3.8) Considere a seguinte amostra aleatória selecionada de uma população com $N_1 = N_2 = 200$.

Estrato 1 : $a_1 = 8$ $n_1 = 30$

Estrato 2: $a_2 = 12$ $n_2 = 50$

Estime P_1, P_2 e P com um intervalo de confiança de 95%.

3.9) Um pesquisador quer estudar o nível de ferro no sangue das vacas de uma rebanho no qual há 2 raças. Para isto, ele separa os animais em 2 estratos: cada raça um estrato. Os resultados da pesquisa foram:

Raça A: 1,5 2,1 1,8 2,6

Raça B: 2,5 1,6 2,3 1,6

a) Estime o nível médio de ferro no rebanho, com uma confiança de 95%, sabendo que havia 50 animais em cada estrato.

b) Qual foi o ganho que o pesquisador teve em termos de precisão com a amostragem estratificada se comparada com amostragem aleatória simples?

3.10) Uma população de 2.000 famílias foi dividida em 3 estratos de acordo com a renda familiar (em s.m.)

ESTRATO	W_i	S_i
1	0,35	2,0
2	0,55	3,3
3	0,10	11,3

a) Ache o tamanho de amostra necessário para estimar a renda com um erro máximo de 0,27 s.m. e uma confiança de 95%. A amostra será repartida nos estratos pela distribuição proporcional. Quantas famílias devem ser selecionadas em cada estrato? Calcule $\sigma^2_{\bar{x}}$.

b) Ache o tamanho de amostra necessário para estimar a renda com um erro máximo de 0,27 s.m. e uma confiança de 95%. A amostra será repartida nos estratos pela distribuição de Neyman. Quantas famílias devem ser selecionadas em cada estrato? Calcule $\sigma^2_{\bar{x}}$.

c) Compare os resultados em a) e b). Comente.

3.11) Suponha a seguinte população:

11 5 25 30 6 8 32 35

Para estimar a média da população, pretende-se selecionar uma amostra de tamanho n = 4 e existem duas sugestões. A primeira é selecionar uma amostra aleatória simples; a segunda sugestão é dividir a população em 2 estratos de tamanho 4 cada e selecionar uma amostra de tamanho 2 em cada estrato.

a) Faça a distribuição dessa população em 2 estratos. Justifique-a.

b) Mostre qual das duas sugestões oferece maior precisão.

3.12) Uma pesquisa para estudar o nível de renda familiar dos alunos da UFRGS considerou a população dividida em 3 sub-populações:
Campus do Vale, Campus Médico e Campus Central. Foram selecionados 100 alunos em cada Campus.
a) Qual foi o delineamento amostral utilizado?
b) Você considera apropriado o critério utilizado? Justifique.
c) Proponha um outro esquema de amostragem que você considere conveniente para o objetivo da pesquisa.

3.13) Suponha uma população de tamanho 10, dividida em 2 estratos de tamanho 6 e 4. Considere o caso de selecionar uma amostra de tamanho 5 tomando 3 unidades do primeiro estrato e 2 do segundo estrato. Quantas amostras diferentes podem ser selecionadas?

3.14) Uma caixa contém 200 ovos de páscoa de chocolate embalados em papel vermelho, 500 embalados em papel azul e 300 com papel amarelo. Proponha um delineamento amostral para estimar a proporção de ovos com peso inferior ao indicado na embalagem, que é de 200g.

3.15) Em uma população de propriedades agrícolas, suponha que foi selecionada a amostra a seguir:

Estrato	N_h	n_h	Amostra
1	26	6	40 50 20 70 90 70
2	25	5	150 140 130 150 170
3	25	5	260 250 280 280 290
4	24	4	350 380 370 390

a) Calcule um IC95% para a média de cada estrato e para a média geral;
b) Calcule os erros absolutos e os erros relativos para cada estrato e para toda a população.