

INE 5111- ESTATÍSTICA APLICADA I - TURMA 05324 - GABARITO
LISTA DE EXERCÍCIOS SOBRE AMOSTRAGEM E PLANEJAMENTO DA PESQUISA

1. Analise as situações descritas abaixo e decida se a pesquisa deve ser feita por amostragem ou por censo, justificando sua resposta.

a) Numa linha de produção de empacotamento de café, observar o peso dos pacotes produzidos.
Se houver possibilidade de agregar uma balança automática ao processo produtivo pode-se utilizar censo. Pois como não se trata de teste destrutivo, e peso dos pacotes é importante para a imagem da empresa (e para não haver desperdício) todos os pacotes podem ser medidos.

b) Em uma sala de aula composta por 40 alunos, analisar suas idades.

Censo, porque a população é pequena, apenas 40 elementos.

c) Observar se a água de uma lagoa está contaminada.

Amostragem. É no mínimo contraproducente retirar toda a água da lagoa para exame de sua contaminação.

d) Verificar a carga horária diária de trabalho dos funcionários da cozinha de um restaurante.

Censo, por razões políticas, para que ninguém se sinta prejudicado, além da população ser pequena e acessível.

e) Num lote de cabos de aço, verificar a resistência dos mesmos à tração.

Amostragem. O teste é destrutivo.

f) Pesquisa de opinião eleitoral para governador do estado de São Paulo.

Amostragem. A necessidade de rápido processamento exige amostragem, devido ao grande tamanho da população que tornaria muito lenta a obtenção dos resultados, além de muito cara.

2. Analise os planos de amostragens apresentados abaixo. Você concorda com a maneira como foram elaborados? Justifique. Apresente as soluções que julgar necessárias.

a) Para analisar os laboratórios de pesquisa quanto ao seu investimento na aquisição de tecnologia de ponta, foram enviados questionários e analisadas as respostas daqueles que responderam o questionário.

Não. Apenas os que responderam serão considerados. Deveriam ser entrevistados os responsáveis pelos laboratórios sorteados em uma amostragem probabilística.

b) Para analisar o perfil dos clientes de um banco, foram analisados ao longo de um mês 4 clientes por dia, retirados da fila do caixa, variando sistematicamente o horário da coleta.

Sim. Mostra o perfil dos clientes ao longo do mês e mesmo ao longo do dia, mas apenas dos clientes que usam os serviços de caixa "convencional"

c) Para ser conhecida a opinião dos estudantes da UFSC sobre o Jornal Universitário, foram colhidas as opiniões de 40 estudantes da última fase do curso de Jornalismo daquela instituição.

Não. Se o objetivo é conhecer a opinião dos alunos da UFSC sobre o jornal, alunos de todos os cursos deveriam ser ouvidos, o plano está incoerente com o objetivo da pesquisa. Deveria ser selecionada uma amostra contendo integrantes de todos os cursos da UFSC, podendo ser estratificada por fase predominante (quanto mais tempo na UFSC maior a probabilidade de ter contato com o jornal e acompanhar sua evolução, o que pode ser relevante).

d) Há interesse em medir o índice de luminosidade das salas de aula da UFSC. A coleta de dados será feita em todos os centros da UFSC, durante os períodos diurno e noturno, nas salas que estiveram desocupadas no momento da pesquisa. Cada centro será visitado apenas uma vez.

Não. O problema reside em medir a luminosidade apenas nas salas desocupadas no momento da pesquisa, e em não retornar aos centros mais de uma vez. É preciso medir os índices de luminosidade nos períodos matutino, vespertino e noturno, levando em conta também a posição solar das salas (selecionar as salas a pesquisar com as plantas dos centros de ensino, que devem estar disponíveis no ETUSC ou na Prefeitura Universitária).

e) As constantes reclamações dos usuários motivaram a direção da Biblioteca Central da UFSC a realizar uma pesquisa sobre o nível de ruído em suas dependências. O ruído será medido em todas as seções da Biblioteca, na primeira e na penúltima semanas do semestre, de segunda à sábado, durante todo o horário de funcionamento.

Não. Por que apenas na primeira e na penúltima semana do semestre? O ideal seria realizar medidas em todas as semanas.

f) No controle de qualidade de uma fábrica de peças, que trabalha 24 horas por dia, sete dias por semana, um item produzido é retirado de cada máquina, a cada meia hora, para avaliação. O procedimento é feito durante todo o dia, ao longo da semana.

Sim. Com este procedimento é possível retratar a qualidade do processo ao longo do tempo (levando em conta os diferentes turnos, operários, condições ambientais), e identificar eventuais problemas, possibilitando a sua prevenção e correção.

g) O Comando de um Batalhão da PMSC quer conhecer a opinião das pessoas que residem em sua área de atuação, no intuito de formular novas escalas de policiamento ostensivo. Para tanto serão feitas entrevistas com as pessoas que se passarem a pé pela frente do Batalhão, de segunda à sexta das 8:30 às 12:00 horas e das 14:00 às 17:30 horas durante duas semanas.

Não. Os que não passarem a pé pela frente do batalhão, e nos horários estabelecidos, não terão a menor possibilidade de serem ouvidos, o que pode comprometer a confiabilidade dos resultados da pesquisa. Seria necessário obter uma listagem das residências, utilizar algum mecanismo de sorteio aleatório para selecionar algumas e entrevistar os moradores sobre a necessidade de policiamento ostensivo.

h) Com a finalidade de estudar o perfil dos consumidores de um supermercado, observaram-se os consumidores que compareceram ao supermercado no primeiro sábado do mês.

Não. E os que forem ao supermercado em outros dias da semana, ou mesmo em outros sábados do mês? Não terão chance de participar da pesquisa, o que poderá comprometer os resultados. A coleta de dados precisa ser feita todos os dias, ao longo do horário de atendimento, e por pelo menos um mês (para avaliar o efeito das diferentes datas de pagamento no consumo dos clientes).

i) Com a finalidade de estudar o perfil dos consumidores de um supermercado, fez-se a coleta de dados durante um mês, tomando a cada dia um consumidor de cada fila de cada caixa, variando-se sistematicamente o horário de coleta dos dados.

Possibilita captar todo o movimento mensal, e o movimento durante o dia (ver item h).

j) Para avaliar a qualidade dos itens que saem de um linha de produção, observaram-se todos os itens das 14 às 14 horas e trinta minutos.

Não. Suponha que algo aconteça às 14h31min. Um eventual problema somente seria notado quando a avaliação fosse repetida, e então os prejuízos poderiam ser muito grandes.

k) Para avaliar a qualidade dos itens que saem de uma linha de produção, observou-se um item a cada meia hora, durante todo o dia.

Sim. Com este plano todo o processo produtivo é monitorado (de forma semelhante à vista no item f), possibilitando a detecção mais rápida de eventuais problemas.

l) Para estimar a porcentagem de empresas que investiram em novas tecnologias no último ano, enviou-se um questionário a todas as empresas de um estado. A amostra foi formada pelas empresas que responderam o questionário.

Não. Provavelmente os respondentes são justamente as empresas que investiram em tecnologia, o que poderia dar a falsa impressão de que a situação está bem melhor do que realmente está. Uma forma de contornar este problema seria conduzir a pesquisa por telefone: obtém-se uma listagem das empresas por porte e ramo de atuação (via FIESC, Fecomércio, CDL's), seleciona-se uma amostra aleatória (estratificada por ramo e porte) e aplica-se um questionário com aqueles que poderiam fornecer informações confiáveis sobre o investimento em tecnologia (ver item a).

3. Analise as situações abaixo e determine qual é o tipo de amostragem a ser usado em cada caso, e explique por quê.

a) Parte da população é inacessível e trata-se de um estudo preliminar.

Amostragem não probabilística, a esmo. Como não há acesso a toda a população seria impossível aplicar uma amostragem probabilística, e como se trata de estudo preliminar a não probabilística é aceitável.

b) Todos os elementos da população podem ser pesquisados, mas não há recursos para a sua listagem total. Sabe-se também que a população subdivide-se em subgrupos semelhantes (para os quais há uma listagem).

Amostragem probabilística por conglomerados. Há acesso a toda a população (mas não há recursos para listar todos os elementos) e a população divide-se em grupos homogêneos (que podem ser listados).

c) Sabe-se que toda a população é acessível, e que é homogênea. A amostra deve ser obtida rapidamente.

Amostragem probabilística sistemática. Há acesso a toda a população (pressupõe-se que haja listagem), que é homogênea. A amostragem sistemática, com o sorteio do ponto de partida e a retirada de elementos a intervalos regulares possibilita um processamento mais rápido do que a aleatória simples.

d) Uma empresa atua em três mercados distintos. Dispõe de uma listagem com os nomes e endereços de todos os clientes. Pretende pesquisar qual seria a eventual demanda de um novo produto. Precisa fazer isso rapidamente (não há tempo para censo).

Amostragem probabilística estratificada proporcional. Não há tempo para um censo. Há acesso a toda a população (há listagem), que pode ser considerada dividida em três estratos (mercados). Como se deseja conhecer a demanda por um novo produto é preciso obter informações precisas, o que pode ser obtido com uma amostra proporcional ao tamanho de cada estrato.

e) A reitoria da UFSC quer conhecer as diferenças básicas entre as idéias de professores, servidores e alunos sobre a instituição. Dispõe de listas com todos os professores, alunos e servidores.

Amostragem probabilística estratificada uniforme. Há acesso a toda a população (listagem), e supõe-se que há uma divisão em 3 estratos (embora dentro dos estratos suponha-se uma certa homogeneidade). Como há interesse em comparar os estratos (as opiniões dos seus integrantes) não há necessidade de obter uma amostra proporcional, bastando retirar a mesma quantidade de cada estrato.

4. Uma população é composta por 2960 elementos que estão ordenados. Se devesse ser retirada uma amostra sistemática de 20 elementos desta população, como você procederia?

Na amostragem sistemática o primeiro passo é obter o intervalo de retirada k , sabendo o tamanho da população ($N = 2960$) e o tamanho da amostra ($n = 20$).

$$k = \frac{N}{n} = \frac{2960}{20} = 148$$
 Como k é inteiro não há necessidade de ajustes. A cada 148 elementos da

população retira-se um para fazer parte da amostra, até completar os 20 que devem compô-la.

Deve-se sortear o ponto de partida das retiradas, para garantir que todos terão chance de pertencer à amostra: os primeiros 148 elementos devem receber números de 1 a 148. Vamos supor que a função ALEATÓRIOENTRE do Calc (com argumentos de 1 a 148) desse o resultado 44): então o ponto de partida será o quadragésimo quarto elemento da lista. A amostra então será composta pelos números: 44, 192, 340, 488, 636, 784, 932, 1080, 1228, 1376, 1524, 1672, 1820, 1968, 2116, 2264, 2412, 2560, 2856 (totalizando 20 elementos).

5. Uma amostragem entre os estudantes de ciências da computação, engenharia de produção e engenharia de automação foi realizada da seguinte maneira: considerou-se cada curso como um estrato, e fez-se retirada proporcional. O curso de computação possui 350 alunos, o de engenharia

de produção 450 alunos e o curso de automação possui 200 alunos. Do curso de computação foram analisados 42 alunos. Qual o número total de alunos analisados na amostra ?

Trata-se de amostragem probabilística estratificada proporcional, indicando que o número de elementos de cada estrato que fazem parte da amostra é proporcional ao tamanho do estrato na população. A população é composta por 1025 alunos (350 + 475 + 200), com os seguintes percentuais para cada curso:

computação = $(350/1025) \times 100 = 34,15\%$ engenharia de produção = $(475/1025) \times 100 = 46,34\%$

automação = $(200/1025) \times 100 = 19,51\%$

Então, os 42 alunos de computação correspondem a 34,15% da amostra, fazendo uma regra de três simples podemos obter o total de elementos da amostra:

42 -----> 34,15%

? -----> 100% E encontramos 122,98, praticamente igual a 123.

6. Numa sala de aula temos 36 homens e 28 mulheres. Faça uma amostragem estratificada proporcional de tamanho 16 considerando o sexo como variável estratificadora. Quantos de cada sexo serão analisados ?

Temos que avaliar as proporções de cada estrato no total de 64 (36 + 28).

Homens = $(36/64) \times 100 = 56,25\%$ Mulheres = $(28/64) \times 100 = 43,75\%$

Basta aplicar os percentuais acima ao tamanho da amostra (16):

Homens = $0,5625 \times 16 = 9$ Mulheres = $0,4375 \times 16 = 7$

Então serão pesquisados 9 homens e 7 mulheres. A retirada dentro de cada estrato pode ser feita com amostra aleatória simples.

7. A biblioteca do CEPM apresenta-se dividida em sete grandes seções (que podem ser considerados como estratos ou conglomerados, dependendo do caso): direito (com 2000 títulos), técnicas policiais (com 1200 títulos), psicologia/sociologia (com 900 títulos), documentos sobre a PMSC (com 900 títulos), língua portuguesa (com 600 títulos), obras de referência- enciclopédias e dicionários (com 400 títulos), e outros- matemática, estatística, ciências (com 1000 títulos). A direção quer conhecer a idade média dos títulos, através do ano de edição, para determinar se há necessidade de renovação do acervo.

Vamos executar vários planos de amostragem, todos podem ser feitos através da função ALEATÓRIOENTRE do Calc.

Há um total de 7000 títulos: 1 a 2000 (direito - D), 2001 a 3200 (técnicas policiais - TP), 3201 a 4100 (psicologia/sociologia - P/S), 4101 a 5000 (documentos sobre a PMSC - DP), 5001 a 5600 (língua portuguesa - LP), 5601 a 6000 (obras de referência - R), 6001 a 7000 (outros - O).

a) Selecionar uma amostra aleatória simples de 70 títulos. Suponha os títulos numerados de acordo com a ordem do enunciado: de 1 a 2000 títulos de metodologia de ensino, de 2001 a 3200 títulos de metodologia de ensino especial, etc.

Amostragem aleatória simples. Precisamos encontrar 70 números entre 1 e 7000 na tabela de números aleatórios (agrupando 4 dígitos, pois o valor máximo é 7000). Vamos registrar, pelo número, a que categoria o livro pertence. Usando a função ALEATÓRIOENTRE do Calc, com argumentos de 1 a 7000, constituindo 70 células, poderíamos obter:

6248 (O), 2645 (TP), 2402 (TP), 4499 (DP), 0947 (D), 3407 (P/S), 3544 (P/S), 1318 (D), 1851 (D), 6232 (O), 4194 (DP), 1509 (D), 4989 (DP), 4354 (DP), 5419 (LP), 5487 (LP), 3043 (TP), 1004 (D), 0696 (D), 3827 (P/S), 0774 (D), 2015 (TP), 1233 (D), 0162 (D), 5298 (LP), 4611 (DP), 1282 (D), 2591 (TP), 2205 (TP), 3900 (P/S), 3875 (P/S), 1863 (D), 3325 (P/S), 3798 (P/S), 1450(D),6571 (O), 3101 (TP), 0246 (D), 4556 (DP), 1427 (D), 3674 (P/S), 0294 (D), 3902 (P/S), 5252(LP), 2180 (TP), 1748 (D), 5417 (LP), 1180 (D), 0533 (D), 5129 (LP), 6956 (O), 1271 (D), 3604 (P/S), 0903 (D), 2411 (TP), 6644 (O), 5207 (LP), 2759 (TP), 3817 (P/S), 1539 (D), 0997 (D), 3334(P/S), 4088 (P/S), 4612 (DP), 3356 (P/S), 4832 (DP), 4779 (DP), 2831 (TP), 2496 (TP), 4710 (DP).

Repare que nenhum livro de obras de referência (entre 5601 e 6000) foi selecionado, não obstante o sorteio não ser viciado: 22 livros de direito, 12 de técnicas policiais, 14 de psicologia/sociologia, 10 de documento da PMSC, 7 de língua portuguesa e 5 de outros.

b) Repetir o item anterior, usando amostragem sistemática.

Sabemos o tamanho da população ($N = 7000$) e o tamanho da amostra ($n = 70$), calculamos então o intervalo de retirada: $k = N/n = 7000/70 = 100$. A cada 100 elementos, retiramos um para fazer parte da amostra. Precisamos sortear o ponto de partida: escolher, por sorteio, um número entre 1 e 70. Podemos usar a função ALEATÓRIOENTRE do Calc, com argumentos de 1 a 100, poderíamos obter 39, que passa a ser o ponto de partida. Então a amostra será composta por:

39 (D), 139 (D), 239 (D), 339 (D), 439 (D), 539 (D), 639 (D), 739 (D), 839 (D), 939 (D), 1039 (D), 1139 (D), 1239 (D), 1339 (D), 1439(D), 1539(D), 1639(D), 1739(D), 1839(D), 1939(D), 2039 (TP), 2139 (TP), 2239(TP), 2339(TP), 2439(TP), 2539(TP), 2639(TP), 2739 (TP), 2839 (TP), 2939 (TP), 3039(TP), 3139(TP), 3239(P/S), 3339(P/S), 3439(P/S), 3539(P/S), 3639(P/S), 3739(P/S), 3839 (P/S), 3939(P/S), 4039(P/S), 4139(DP), 4239(DP), 4339(DP), 4439(DP), 4539(DP), 4639(DP), 4739(DP), 4839(DP), 4939(DP), 5039(LP), 5139(LP), 5239(LP), 5339(LP), 5439(LP), 5539(LP), 5639(R), 5739(R), 5839(R), 5939(R), 6039(O), 6139(O), 6239(O), 6339(O), 6439(O), 6539(O), 6639(O), 6739(O), 6839(O), 6939(O).

c) Selecionar uma amostra estratificada proporcional de 70 títulos.

O primeiro passo é definir o quanto cada estrato representa da população, para saber quantos elementos da amostra ($n = 70$) virão de cada um deles.

Direito = $(2000/7000) \times 100 = 28,57\%$ (então 28,57% da amostra virá deste estrato, 20 livros).

Técnicas policiais = $(1200/7000) \times 100 = 17,14\%$ (12 livros da amostra virão desta seção).

Psicologia/sociologia = $(900/7000) \times 100 = 12,86\%$ (9 livros da amostra virão desta seção).

Documentos sobre a PMSC = $(900/7000) \times 100 = 12,86\%$ (9 livros da amostra virão desta seção).

Língua portuguesa = $(600/7000) \times 100 = 8,57\%$ (6 livros da amostra virão desta seção).

Obras de referência = $(400/7000) \times 100 = 5,71\%$ (4 livros da amostra virão desta seção).

Outros = $(1000/7000) \times 100 = 14,28\%$ (10 livros da amostra virão desta seção).

Dentro de cada estrato podemos selecionar a amostra utilizando a amostragem sistemática, descrita no item b: basta definir os intervalos de retirada para cada estrato (com o tamanho de cada estrato sendo N e os tamanhos de amostra definidos acima como n).

d) Extrair uma amostra aleatória de 70 títulos através de amostragem de conglomerados em dois estágios: no primeiro selecionar 2 seções, e no segundo estágio 35 títulos de cada conglomerado selecionado.

Seleção de conglomerados. Consideremos as seções como conglomerados, há 7 deles, vamos numerá-los: Direito - 1 Técnicas policiais - 2 Psicologia/sociologia - 3 Documentos sobre a PMSC - 4 Língua portuguesa - 5 Referência - 6 Outros - 7

Temos que selecionar dois: precisamos obter dois números de 1 dígito (entre 1 e 7) através da função ALEATÓRIOENTRE do Calc. Poderíamos ter encontrado 6 e 7: os conglomerados sorteados serão Referência e Outros. Dentro de cada um deles podemos selecionar 35 elementos, vejamos o procedimento no conglomerado Referência, que tem 400 títulos (numerados de 5601 a 6000).

$N = 400$ $n = 35$ $k = N/n = 400/35 = 11,43$

O valor de k é fracionário, e como não podemos modificar n devemos descartar por sorteio 15 elementos do conglomerado de Referência, restando 385 elementos, resultando em $k = 11$. O sorteio pode ser realizado procurando por 15 números de 4 dígitos entre 5601 e 6000 através da função ALEATÓRIOENTRE do Calc. Após remover estes elementos reordena-se a lista e retiram-se os 35 elementos da amostra propriamente dita, usando o procedimento visto na letra b. Observe que os conglomerados sorteados não contêm exatamente os livros mais importantes da biblioteca.

e) Compare os resultados e comente sobre as semelhanças e ou diferenças. Qual dos itens terá um resultado mais preciso, por quê?

Os resultados das amostragens sistemática e estratificada proporcional são bastante razoáveis, pois as amostras obtidas incluíram vários elementos de todas as seções. Não obstante o resultado mais preciso é o da estratificada proporcional, por preservar na amostra a proporcionalidade das seções na população, e por garantir que todos os estratos estejam presentes na amostra: esta retirou 20 livros de direito, enquanto a sistemática retirou apenas 11, por exemplo. A amostragem aleatória simples apresenta títulos de quase todas as áreas, mas é ainda menos precisa do que a sistemática (uma área não teve elementos selecionados). Mas a menos precisa foi a de conglomerados, que sorteou livros apenas das seções de referência e outros: isso ocorreu meramente por acaso, o mesmo poderia ter acontecido na aleatória simples, pois ambas supõem que a população é homogênea.

8. O mapa seguinte simboliza os domicílios de um bairro. Os quadros grandes correspondem aos quarteirões, divididos em duas localidades (estratos) do bairro. Os números dentro dos quadros menores (domicílios), correspondem ao número de cômodos do domicílio, que é a variável a ser observada, numa amostra de domicílios.

4	3	4	5
5		5	
3	2	2	4

2	4	5	4
6		3	
3	4	3	2

4	2	5	5
5		6	
4	5	3	4

5	3	2	5
5		7	
3	2	4	4

2	4	6	4
3		5	
3	4	3	2

4	4	5	5
5		6	
4	6	5	4

Estrato A

Estrato B

8	8	7	5
13		8	
8	9	7	6

5	6	5	8
13		9	
8	4	9	4

9	7	7	5
12		10	
7	6	7	8

Para resolver este exercício numere os domicílios de cada quarteirão, da esquerda para a direita, e de cima para baixo, resultando em domicílios de 1 a 90 (onde os domicílios de 1 a 60 estão no estrato A, e os de 61 a 90 no estrato B).

a) Ignorando os estratos, selecionar uma amostra aleatória simples de 9 domicílios. Anotar o número de cômodos dos domicílios selecionados na amostra.

Obtendo 9 números aleatórios de 2 dígitos (entre 1 e 90) pela função ALEATÓRIOENTRE do Calc, poderíamos encontrar: 08, 62, 48, 26, 45, 24, 02, 84, 04. Os números de cômodos serão então: 2, 8, 4, 6, 3, 5, 3, 5, 5.

b) Selecionar uma amostra estratificada proporcional de 9 domicílios. Anotar o número de cômodos dos domicílios selecionados na amostra.

O estrato A tem 60 domicílios, representando $\frac{2}{3}$ da população, e o estrato B tem 30, representando $\frac{1}{3}$. Como a amostra deve ter 9 domicílios, 6 virão do estrato A e 3 do estrato B.

Através da função ALEATÓRIOENTRE do Calc podemos obter 6 números aleatórios de 2 dígitos (entre 1 e 60), por exemplo: 54, 11, 48, 09, 08, 28. Então, do estrato A teremos: 5, 2, 4, 2, 2 e 5 cômodos.

Através da função ALEATÓRIOENTRE do Calc podemos obter 3 números aleatórios de 2 dígitos (entre 61 e 90) por exemplo: 82, 63, 83. Então, do estrato B teremos: 7, 7, e 7 cômodos. Amostra completa: 5, 2, 4, 2, 2, 5, 7, 7, 7.

c) Extrair uma amostra aleatória de tamanho 9, através de uma amostragem de conglomerados em dois estágios. No primeiro estágio selecione 3 quarteirões e, no segundo estágio, 3 domicílios em cada conglomerado selecionado. Anotar o número de cômodos dos domicílios selecionados na amostra.

Atribuem-se números de 1 a 9 aos quarteirões, da esquerda para a direita e de cima para baixo. Através da função ALEATÓRIOENTRE do Calc podemos obter 3 números de 1 dígito (de 1 a 9), por exemplo: 2, 5. Foram selecionados os quarteirões 2 e 5 (ambos do estrato A...). Vamos então selecionar 3 domicílios em cada um deles.

O conglomerado 2 contém os domicílios numerados de 11 a 20. Através da função ALEATÓRIOENTRE do Calc podemos obter 3 números de 2 dígitos (de 11 a 20), por exemplo: 18, 11, 19. Então, do conglomerado 2 teremos: 4, 2, e 3 cômodos.

O conglomerado 5 contém os domicílios numerados de 41 a 50. Através da função ALEATÓRIOENTRE do Calc podemos obter 3 números de 2 dígitos (de 41 a 50), por exemplo: 48, 45, 44. Então, do conglomerado 5 teremos: 4, 3 e 4 cômodos. Amostra completa: 4, 2, 3, 4, 3, 4.

d) Comparar os resultados e comentar sobre as eventuais semelhanças ou discrepâncias. Qual dos itens terá um resultado mais preciso, por quê?

Os resultados mais precisos serão obtidos com a amostragem estratificada proporcional, que garante a proporcionalidade dos estratos na amostra, tornando-a mais representativa. Os resultados da amostragem aleatória simples são bastante próximos: 7 domicílios do estrato A e 2 do B (contra 6 do A e 3 do B na estratificada). A amostragem por conglomerados obteve o pior resultado, pois foram sorteados apenas conglomerados do estrato A: mas, tal resultado ocorreu apenas por acaso, na aleatória simples também poderíamos ter obtido apenas domicílios do estrato A, ou apenas do estrato B. A vantagem da estratificada proporcional é a garantia de que todos os estratos estarão presentes na amostra, quando a população NÃO é homogênea (o que ocorre neste problema).

9. Será feito um levantamento por amostragem de uma população de 2000 famílias, para a realização de uma pesquisa.

a) Calcule o tamanho mínimo de uma amostra para que se tenha um erro amostral máximo de 5%.

$$n_0 = \frac{1}{e_0^2} = \frac{1}{(0,05)^2} = 400 \quad n = \frac{N \times n_0}{N + n_0} = \frac{2000 \times 400}{2000 + 400} = 333,33 \cong 334$$

b) Supondo a população dividida em 2 estratos iguais, qual o tamanho mínimo de amostra para se ter um erro amostral máximo de 5% em cada estrato?

O valor de N será igual a 1000 (2000/2):

$$n_0 = \frac{1}{e_0^2} = \frac{1}{(0,05)^2} = 400 \quad n = \frac{N \times n_0}{N + n_0} = \frac{1000 \times 400}{1000 + 400} = 285,71 \cong 286 \text{ (592 elementos)}$$

c) Qual seria o erro amostral em cada estrato do item b, se o tamanho da amostra em cada estrato fosse simplesmente o valor definido no item a dividido por 2?

Vamos fazer o caminho inverso: a partir de n (que vale 334/2 = 167), supondo N = 1000, encontrar o e₀.

$$167 = \frac{1000 \times n_0}{1000 + n_0} \quad 167000 + 167n_0 = 1000n_0 \quad n_0 = 167000/833 = 200,48 \cong 201$$

$$n_0 = \frac{1}{e_0^2}, \text{ isolando } e_0: e_0 = \sqrt{\frac{1}{n_0}} = 0,0705 \text{ (7,05\%)}$$

10. Repetir o exercício 9 para os seguintes tamanhos de população:

- 4000 famílias.
- 20000 famílias
- 400000 famílias

e quatro regiões no interior. Devido ao maior percentual de habitantes decidiu-se realizar a pesquisa apenas na Capital, e adotou-se a estratégia de pontos de fluxo: foram colocados entrevistadores em 3 pontos estratégicos da cidade durante 3 dias com instruções sobre que tipos de pessoas (sexo, faixa etária, etc) deveriam abordar: em cada ponto deveriam ser entrevistadas 400 pessoas, das quais 25% deveriam ser homens com até 30 anos, 25% mulheres com até 30 anos, 25% homens com mais de 30 anos e 25% mulheres com mais de 30 anos. Estas pessoas deveriam responder ao questionário abaixo:

Você não concorda que a atual administração estadual está dando mais atenção ao problema da saúde? SIM NÃO NÃO SABE

Você acredita que a atual administração estadual está aplicando melhor o dinheiro público?

SIM NÃO NÃO SABE

Você não concorda que um bom governante tenha o direito de se reeleger?

SIM NÃO NÃO SABE

Após realizar a pesquisa o governo divulgou que 72,6% dos habitantes estavam satisfeitos com a atual administração estadual.

a) Você concorda com a forma como a pesquisa foi conduzida? Por quê? Você mudaria algo? JUSTIFIQUE.

Não concordo com o modo como a pesquisa foi conduzida. Deveriam ter sido realizadas entrevistas nas 5 regiões do estado, e a metodologia de passantes eventuais simplesmente impede que aqueles que não passaram pelos pontos de coleta sejam selecionados, IMPOSSIBILITANDO a generalização estatística dos resultados obtidos para a população de eleitores.

b) O questionário aplicado é adequado? Por quê? Se achar que não elabore um questionário (máximo de 10 questões).

O questionário não é adequado. As questões 1 e 3 estão confusas, e todas estão induzindo o entrevistado a manifestar-se favoravelmente ao governo: bastaria retirar o "não" nas perguntas 1 e 3, o "mais" na questão 1, o "melhor" na questão 2 e o "bom" na questão 3.

c) Qual é a sua opinião sobre o resultado da pesquisa? É confiável? JUSTIFIQUE.

Os resultados de tal pesquisa não são confiáveis. A amostragem usada não é probabilística, é viesada (por selecionar pessoas apenas na capital do estado), e o questionário está confuso e induzindo os entrevistados.