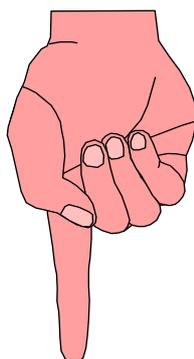




Sondagens

CADERNO DE EXERCÍCIOS



É uma amostra representativa!

1. Uma equipa da Segurança Social realizou uma sondagem às famílias que beneficiaram, em anos anteriores, de ajuda no âmbito de um determinado projecto de luta contra a pobreza, com o objectivo de inquirir vários aspectos relacionados com as suas condições de vida actuais, nomeadamente apurar qual a taxa de desemprego actual nessas famílias. Construiu, a partir dos registos existentes, uma base de sondagem com 7500 famílias que beneficiaram dessa ajuda e seleccionou uma amostra aleatória simples de 750 famílias, das quais não foi possível localizar 250 (já não residiam no mesmo local e não existia contacto de local de trabalho). Verificados os registos constatou-se que essas 250 famílias pertenciam maioritariamente ao escalão de rendimento mais baixo.

Tendo exclusivamente por base a informação apresentada acima indique, das seguintes frases, quais as verdadeiras (**V**) ou falsas (**F**), justificando sucintamente:

- a) A população inquirida deste estudo é composta pelas 750 famílias seleccionadas.
 - b) Um dos parâmetros em estudo é uma proporção.
 - c) Esta amostra não tem validade estatística porque a taxa de amostragem é baixa.
 - d) As 250 famílias não localizadas constituem uma fonte de erro não quantificável que poderá enviesar as conclusões do estudo.
2. Pretendendo avaliar a percentagem dos “sem abrigo” da cidade de Lisboa que sofrem de perturbações do foro psiquiátrico realizou-se um estudo junto a uma amostra dos “sem abrigo” que procuraram cuidados médicos no decorrer do 1º semestre deste ano no Hospital de São José. Com base na informação recolhida nessa amostra obteve-se uma estimativa de 33% para essa percentagem.
 - a) Identifique a população objectivo e a população inquirida e descreva a metodologia amostral subjacente ao estudo.
 - b) Faça uma análise crítica sobre a metodologia utilizada para a selecção desta amostra e comente a validade da estimativa obtida para o parâmetro em estudo.

3. Da ficha técnica de uma sondagem (telefónica) publicada num diário da capital foi retirado o seguinte excerto:

“O universo alvo é composto pelos indivíduos com mais de 18 anos residentes em Portugal continental em habitações com telefone. Os números de telefone foram escolhidos aleatoriamente das listas telefónicas nacionais, e em cada habitação foi inquirida a próxima pessoa a fazer anos. Obtiveram-se 916 inquéritos válidos, o que corresponde a uma taxa de resposta de 61,3 por cento. 52,5 por cento dos inquiridos eram do sexo feminino e a margem de erro máximo da amostra é de 3,24 por cento, com um nível de confiança de 95 por cento”.

Caracterize de forma sumária a metodologia de amostragem utilizada nesta sondagem e explicito o que leva a inquirir, em cada habitação, a próxima pessoa a fazer anos. Explicito como se obtém o valor de 3,24% para o erro máximo da amostra, com um nível de confiança de 95%.

4. Apresenta-se em seguida a transcrição das fichas técnicas de duas sondagens realizadas no decorrer da campanha eleitoral .

Sondagem 1

.... O universo alvo é composto pelos indivíduos com 18 ou mais anos recenseados eleitoralmente e residentes em Portugal Continental. Foram seleccionadas aleatoriamente 48 freguesias do país, tendo em conta a distribuição da população recenseada eleitoralmente por regiões NUT II (2001) e por freguesias com mais e menos de 3200 recenseados. A selecção aleatória das freguesias foi sistematicamente repetida até que os resultados eleitorais das eleições legislativas de 2002 e 2005 e europeias de 2009 nesse conjunto de freguesias, ponderado o número de inquéritos a realizar em cada

uma, estivessem a menos de 1% dos resultados nacionais dos cinco maiores partidos. Os domicílios em cada freguesia foram seleccionados por caminho aleatório e foi inquirido em cada domicílio o mais recente aniversariante recenseado eleitoralmente na freguesia. A intenção de voto foi recolhida através de boletim simulado e voto em urna. Foram obtidos 4367 inquéritos válidos, sendo que 55% dos inquiridos eram do sexo feminino, 38% da região Norte, 19% do Centro, 30% de Lisboa e Vale do Tejo, 7% do Alentejo e 5% do Algarve. Todos os resultados obtidos foram depois ponderados de acordo com a distribuição da população com 18 ou mais anos residente no Continente por sexo (2007) e escalões etários (2007), na base dos dados do INE, e por região e habitat na base dos dados do recenseamento eleitoral. A taxa de resposta foi de 68%.* A margem de erro máximo associado a uma amostra aleatória de 4367 inquiridos é de 1,5%, com um nível de confiança de 95%...

Sondagem 2

... um dos objectivos da sondagem é a análise da intenção de Voto para a Assembleia da República. O Universo deste estudo é constituído por indivíduos de ambos os sexos com idade igual ou superior a 18 anos, residentes em Portugal Continental, 8 357 491 Indivíduos Fonte: DGAI dados de 29 de Fevereiro de 2008. A amostra deste estudo é constituída por um total de 809 inquiridos, distribuída geograficamente da seguinte forma: Grande Lisboa 157

entrevistas, Grande Porto 91, Litoral Centro 132, Litoral Norte 155, Interior Norte 182 e Sul 92. Quanto às variáveis sexo e idade, a amostra apresenta-se distribuída da seguinte forma: Masculino 385 entrevistas e Feminino 424; *Idade* - 18/34 anos 257 entrevistas, 35/54 anos 276 e mais de 54 anos 276.

A selecção dos lares a serem contactados foi efectuada aleatoriamente, a partir da base de telefones residenciais disponíveis em Portugal Continental. Esta base foi construída, a partir da base de telefones residenciais disponíveis em Portugal Continental dos operadores Portugal Telecom e Cabovisão. Os telefones seleccionados foram extraídos aleatoriamente desta base, através de um software proprietário desenvolvido para o efeito, de forma proporcional ao número de entrevistas a realizar em cada Região/Distrito/Concelho. Os entrevistadores não tiveram qualquer participação na selecção dos números de telefone. Existe um condicionalismo que impõe que o número seleccionado não tenha sido utilizado neste mesmo estudo num período recente (12 meses). A selecção dos entrevistados, em em cada lar, foi efectuada através do método de quotas tendo em consideração as seguintes variáveis: sexo, idade e região do entrevistado.

A amostra prevista era de 800 entrevistas e a amostra final do estudo é de 809 entrevistas válidas, tendo sido utilizados 18478 números de telefones, e realizado um total de 28036 contactos telefónicos. A recolha da informação foi efectuada através do método da entrevista telefónica com recurso ao Sistema CATI (Computer Assisted Telephone Interview). Foi utilizado um questionário estruturado contendo perguntas fechadas e abertas ...

Faça uma análise crítica das duas metodologias (apresentando nomeadamente, as principais semelhanças e/ou diferenças, vantagens e inconvenientes de cada uma delas, se os houver).

5. Uma amostra aleatória simples sem reposição de 30 domicílios foi seleccionada numa zona urbana com 18.850 domicílios. No quadro seguinte apresenta-se a distribuição do número de pessoas por domicílio para a amostra observada:

Nº pessoas/domicílio	1	2	3	4	5	6	7
Nº domicílios	1	4	12	9	2	1	1

Estime o número total de pessoas que vivem nessa zona e calcule a precisão relativa associada a esta estimativa.

6. A empresa de contabilidade M&Ec Lda. tem a seu cargo a auditoria de um grande armazém com filiais em vários pontos do país. A auditoria envolve, entre outros aspectos, um exame às *contas a receber*. Quase todas estas contas representam pagamentos feitos com cartões de crédito. Para proceder à auditoria de cada conta a M&Ec deverá examinar todos os débitos e créditos feitos no ano, e certificar-se que estes estão correctamente contabilizados. No dia 31 de Dezembro desse ano existiam 25520 contas no total. Tornando-se extremamente dispendioso examiná-las todas a auditoria será baseada numa amostra.
- a) Existe uma diferença de opinião entre dois dos membros da equipa que vai realizar o trabalho, sobre o modo de seleccionar as contas a examinar. O senhor A entende que: “a selecção deverá ser ao acaso pois, só assim, a amostra não terá enviesamentos”. O senhor B contrapõe por seu lado: “examinam-se somente as maiores contas, pois, havendo algum erro, especialmente erros que possam ser gravosos para o Armazém, eles ocorrerão, mais provavelmente, nessas contas. Examinar as contas pequenas é pura perda de tempo”. Tendo em conta as opiniões apresentadas, explique qual a metodologia que deve ser seguida para a selecção da amostra.
- b) Suponha que se vai seleccionar uma amostra PISR com o objectivo de estimar a proporção de contas com erros. Qual deverá ser a dimensão da amostra se se pretender garantir, com uma confiança de 95%, um erro máximo de $\pm 0,03$ na proporção estimada, sabendo-se que a verdadeira percentagem de contas erradas não ultrapassa os 10%.
- c) Seleccionou-se uma amostra aleatória PISR de 300 contas tendo-se observado:
- | | | | |
|--------------------------------------|-------------|------------------|--|
| Percentagem de contas com erro: | 8% | | |
| Valor contabilístico, por conta (VC) | média: 10,9 | variância: 490,5 | |
| Valor real, por conta (VR) | média: 10,7 | variância: 473,8 | |
- c1) Construa um intervalo de confiança a 90% para o valor real da totalidade das *contas a receber* no final desse ano.
- c2) Obtenha uma estimativa para o desvio (VR-VC) da totalidade das *contas a receber* e apresente como calcularia a precisão amostral associada a esta estimativa.
7. De uma lista de 468 cursos profissionais foi seleccionada uma amostra aleatória simples sem reposição de 100 cursos. Os dados sobre o número de alunos (y_j) e de professores (x_j) foram os seguintes:

$$\sum y_i = 44\,988 ; \sum x_i = 3\,099 ; \sum y_i^2 = 36\,247\,904 ; \sum x_i^2 = 144\,209 ; \sum x_i y_i = 2\,160\,380$$

Construa um intervalo de confiança a 90% para o rácio alunos/professor dos cursos profissionais da lista em análise.

8. Uma associação de defesa do consumidor sediada na Região Centro pretende averiguar o impacto da proximidade do Natal nos preços de um determinado cabaz de produtos, através da observação dos preços praticados numa amostra aleatória simples, sem reposição, das 2050 lojas que na região vendem esses produtos. Foram contratados estudantes que visitaram e registaram, em cada uma das 50 lojas da amostra, os preços desse cabaz de produtos em dois períodos de tempo distintos, na primeira semana do mês de Outubro (X) e na terceira semana do mês de Dezembro (Y).

a) Proponha, tendo em conta o modo como foi recolhida a informação, duas alternativas para estimar a variação relativa do preço do cabaz provocada pela proximidade do Natal. Justifique sucintamente, por qual optaria.

b) Com base na informação recolhida obtiveram-se os seguintes resultados:

$$\sum_{i=1}^{50} x_i = 5614,42 \quad \sum_{i=1}^{50} y_i = 6075,85 \quad s_x^2 = 719,73 \quad s_y^2 = 2262,16 \quad r_{xy} = 0,974 .$$

Calcule um intervalo de confiança a 95% para a variação relativa no preço desse cabaz provocada pela proximidade do Natal.

9. Pretende-se estimar a taxa de incidência actual de dois tipos de doenças (A e B), sobre uma população numerosa. Estudos anteriores sobre essas doenças permitem afirmar que a doença A é muito comum e chegou a afectar cerca de 60% da população, enquanto que a B nunca atingiu mais de 10% dos indivíduos. Qual deverá ser a dimensão da amostra a seleccionar de forma a garantir, com um grau de confiança de 95%, um erro máximo de 2,5% para a taxa de incidência de cada uma das doenças?
10. A REVDIST, empresa distribuidora de revistas, pretende avaliar o consumo mensal de revistas de informática numa região onde residem cerca de 800 mil agregados familiares. Para tal realizou um primeiro inquérito telefónico junto a uma amostra de 500 agregados familiares seleccionados, aleatoriamente e sem reposição, da lista telefónica dessa região. Inquirindo-se qual o número de revistas de informática adquiridas no último mês (Y_i), obtiveram-se os seguintes resultados:

$$\sum_{i=1}^{500} y_i = 260 \quad e \quad \sum_{i=1}^{500} y_i^2 = 4175$$

- a) Identifique a população objectivo e a população inquirida.
- b) Qual deveria ser a dimensão de uma amostra aleatória simples, sem reposição, se pretender estimar o total de revistas de informática consumidas com uma precisão relativa de 10% e uma confiança de 95%.
11. Uma empresa imobiliária deseja estimar, com base numa amostra aleatória simples, sem reposição, o número de lugares de estacionamento necessários para um prédio de escritórios a construir. Sabe-se que no prédio irão trabalhar 5000 pessoas, e que em empresas do mesmo tipo das que lá se vão instalar a proporção de trabalhadores que utilizam carro próprio na sua deslocação diária para o emprego não é superior a 30%.
Qual deverá ser a dimensão da amostra (a seleccionar dos potenciais futuros ocupantes do edifício) para que seja possível estimar a 90% o número de lugares de estacionamento necessários com uma margem de erro de ± 150 .
12. Uma empresa pretende estimar o número de horas de trabalho perdidas mensalmente devido a acidentes ocorridos no local de trabalho e, para tal, vai observar uma amostra de trabalhadores durante o próximo mês de Março. Havendo a percepção de que as taxas incidência são diferentes consoante a categoria decidiu-se agrupar os 260 trabalhadores existentes em três classes: trabalhadores fabris, técnico/administrativos e quadros dirigentes.
Com base nos registos feitos no último ano calculou-se a variância do número de horas de trabalho perdidas devido a acidente, por trabalhador e por mês para cada uma das categorias consideradas.
No quadro abaixo apresentam-se estes dados bem como o número de tabalhadores existentes em cada classe:

	Trab. Fabris	Técnicos e Adm.	Quadros dirigentes
Variância	36	25	4
Nº trabalhadores	133	95	32

- a) Qual a população objectivo e a população inquirida.
- b) Diga, justificando, que afixação recomendaria para este problema.
- c) A empresa tem um orçamento de 1000€ para efectuar o estudo. Sabe-se que o custo fixo para o estudo é de 750€ e que o custo da recolha de informação para cada empregado seleccionado é de 5€, qualquer que seja a categoria a que pertence.
- c1) Calcule, justificando, qual o número óptimo de empregados a inquirir total e por estrato.
- c2) Utilizando as dimensões das amostras por estrato calculadas em b) calcule as probabilidades de selecção do i -ésimo trabalhador de cada um dos estratos, isto é, qual a probabilidade de selecção de um trabalhador fabril, de um técnico/administrativo e de um quadro dirigente. (Nota: Se não resolveu c1) suponha que a amostra era composta por 30 trabalhadores fabris, 15 técnicos/administrativos e 5 quadros dirigentes).
13. Pretende-se estimar, através de um inquérito por amostragem, a produção de cereais das 5000 explorações agrícolas agrupadas em três zonas geográficas. No quadro seguinte apresentam-se informações sobre o total de explorações agrícolas existentes em cada zona e os resultados referentes a uma amostra estratificada de 100 explorações.

Zona Geográfica	Nº de explorações Agrícolas existentes	Resultados referentes à amostra		
		Explorações Inquiridas	Produção por exploração	
			Média	D. pad. corrigido
1	2 500	50	1,0	0,5
2	1 800	36	5,0	1,1
3	700	14	15,0	3,5
Total	5 000	100	4,4	4,9

- a) Calcule um intervalo de confiança a 95% para a produção de cereais dessa população.
- b) Determine a dimensão mínima da amostra a recolher, total e por zona, de forma a garantir com uma confiança de 95% um erro máximo de 950, para a produção total de cereais.
14. O volume de vendas diárias de um certo estabelecimento comercial varia com o dia da semana e do mês. Pretende-se seleccionar uma amostra de 52 dias dos 365 que compõem o ano, para observar nesses dias qual o comportamento de uma variável que se sabe estar fortemente correlacionada com o volume de vendas.
- Da análise dos dados anteriores sobre os volumes de vendas diários sabe-se que:
- ao Sábado e Domingo o volume de vendas é quádruplo do dos restantes dias;
 - no 1º dia do mês é três vezes superior ao dos outros dias;
 - se o primeiro dia do mês for Sábado ou Domingo o montante de vendas é dez vezes superior.
- Se necessário, suponha que 4 dos 104 Sábados/Domingos desse ano são 1º dia do mês.
- a) Diga que estratificação recomendaria para este problema, e tendo em conta a informação disponível, calcule a dimensão da amostra dos estratos.
- b) Diga, justificando, se a sua proposta de estratificação proporcionaria melhores resultados do que seleccionar um dia em cada uma das 52 semanas.
15. Num inquérito por amostragem vão-se utilizar dois estratos com pesos W_1 e W_2 . Baseando-se em estudos anteriores pode também considerar-se que:
- os estratos têm idêntica variabilidade ($\sigma_1 = \sigma_2$),
 - o custo por inquérito no estrato 2, c_2 , situa-se entre $2c_1$ e $4c_1$, (sendo c_1 o custo por inquérito no estrato 1).

Embora preferisse utilizar a afixação proporcional, não deseja com isso incorrer num grande incremento da

$$\frac{V(\hat{\mu}_{EP})}{V(\hat{\mu}_{EO})} = \frac{W_1 c_1 + W_2 c_2}{(W_1 \sqrt{c_1} + W_2 \sqrt{c_2})^2}$$

variância, comparativamente com a utilização da afixação óptima. Para um dado valor orçamentado C para o custo, $C = c_1 n_1 + c_2 n_2$, e desprezando as c.p.f., mostre que:

Admitindo que $W_1 = W_2$, calcule os incrementos relativos na variância decorrentes da utilização da afixação proporcional quando $c_2/c_1 = 2$ e quando $c_2/c_1 = 4$.

16. Numa estratificação com dois estratos, de grande dimensão, os valores dos pesos e das variâncias dos estratos são os seguintes:

Estrato	Peso	Desvio padrão
1	0,8	2
2	0,2	4

Calcule as grandezas das amostras de cada estrato, n_1 e n_2 , necessárias para que se satisfaçam as diferentes condições apresentadas nas alíneas seguintes.

- a) O erro padrão do valor médio estimado da população deve ser 0,1 e a grandeza total da amostra, $n = n_1 + n_2$, deve ser mínima.
- b) O erro padrão do valor médio estimado em cada estrato deve ser 0,1.
17. Mostre que, quando o coeficiente de variação da característica em estudo é aproximadamente igual para os diferentes estratos, a afixação óptima corresponde à situação em que a dimensão das amostras de cada estrato (n_h) é proporcional ao total da característica em cada estrato.
18. Para estudar o estado de saúde dos clientes de um ginásio foi seleccionada uma amostra aleatória PISR de 100 dos cerca de 5000 clientes constantes dos ficheiros. Entre vários aspectos inquiridos/observados registou-se o peso de cada um deles, tendo-se apurado um peso médio de 61,7 kgs e uma variância corrigida de 189,1.
- a) Calcule um IC a 95% para o peso médio de um cliente desse ginásio.
- b) O gerente do ginásio ficou muito surpreendido com o valor obtido para o peso médio/por cliente, pois os seus clientes são maioritariamente homens (55%) e resolveu averiguar a composição da amostra ventilando os resultados por sexo.

Sexo	Nº de clientes Na amostra	Peso registado	
		Média (\bar{x})	Variância (s^2)
Masculino	30	82	16,5
Feminino	70	53	8,4
Total	100		

Obtenha, tendo em conta a informação agora disponibilizada, um novo IC a 95% para o peso médio por cliente. Analise os resultados comparando-os com os resultados obtidos na alínea anterior.

- c) O ginásio tem um ficheiro informatizado dos seus clientes onde constam, entre outros, dados sobre: data de nascimento, sexo, profissão e registo de principais problemas de saúde se os houver. Sugira, concisa e justificadamente, um desenho amostral adequado para estudos futuros desta natureza.
19. Uma dada distribuidora de revistas pretende avaliar com uma boa precisão o mercado total de revistas de actualidades de uma região onde residem cerca de 800 mil agregados familiares. Para tal realizou um inquérito telefónico junto a uma amostra uma amostra aleatória PISR de 4100 agregados tendo-se apurado para a quantidade de revistas de actualidades adquiridas nos três últimos meses por agregado, uma média de 0,62 e uma variância corrigida de 5,0. Os resultados ventilados segundo o tipo de habitat, urbano e rural, foram os seguintes:

Tipo de habitat	Nº de respostas	Nº de revistas de actualidades adquiridas por agregado	
		Média	Variância corrigida
Urbano	2450	0,78	6,9
Rural	1650	0,37	2,1

- a) Considerando a metodologia utilizada para a selecção da amostra calcule um intervalo de confiança a 95% para a procura trimestral de revistas de actualidades nessa região.
- b) Diga, justificadamente, se haverá um efeito significativo do tipo de habitat na procura deste tipo de revistas..
- c) Sabendo que cerca de 25% dos agregados desta região pertencem a zonas rurais, construa um intervalo de confiança a 95% para a procura trimestral de revistas nessa região. Explique as vantagens e/ou desvantagens da metodologia utilizada, bem como as hipóteses subjacentes à sua resolução. Proceda a uma análise comparativa com o resultado obtido na alínea a).
20. Seleccionou-se uma amostra aleatória simples, sem reposição, de 30 dos 86 concelhos pertencentes à Região Norte e procedeu-se ao levantamento dos cursos de formação profissional que decorreram no ano passado em cada um desses concelhos registando-se o número de alunos neles inscritos (Y_i).

Obtiveram-se os seguintes dados: $\sum_{i=1}^{30} y_i = 8745$, $\sum_{i=1}^{30} y_i^2 = 12150435$.

- a) Determine uma estimativa para o total de alunos inscritos em cursos de formação profissional na região e calcule a precisão associada.
- b) Dos dados do censo de 2001 retirou-se informação sobre a população residente no escalão etário 15 a 59 anos tanto para a totalidade da Região Norte (cerca de 2 milhões) como para os 30 concelhos seleccionados (cerca de 800 mil residentes). Calcule uma estimativa para o total de alunos inscritos em cursos de formação profissional na Região Norte aproveitando esta informação disponível e indique como calcularia a precisão associada.
21. Um biólogo quer estimar a idade média das árvores existentes numa área florestal. Para calcular a idade exacta de uma árvore é necessário contar o número de anéis existentes no cerne da árvore, tornando-se um processo complicado e muitas vezes destrutivo. No entanto, em princípio, quanto mais velha é a árvore maior é o perímetro do seu tronco. Nesse sentido foi medido o perímetro de cada uma das 1050 árvores existentes nessa zona, tendo-se obtido um perímetro médio de 2,3. Seguidamente escolheu-se uma amostra aleatória simples sem reposição de 20 árvores e procedeu-se à medição da idade (Y_i) e do perímetro (X_i) de cada árvore. Obtiveram-se os seguintes resultados:

$$\sum_{i=1}^{20} x_i = 37,9 \quad \sum_{i=1}^{20} y_i = 1433 \quad \sum_{i=1}^{20} x_i^2 = 74,45 \quad \sum_{i=1}^{20} y_i^2 = 109593 \quad \sum_{i=1}^{20} x_i y_i = 2826$$

- a) Determine, utilizando plenamente toda a informação disponível, a estimativa para a idade média das árvores dessa área e calcule a precisão associada. Justifique.
- b) Qual deveria ser a dimensão da amostra se se desejasse garantir, com um grau de confiança de 95%, uma precisão relativa de 5% para a estimativa da idade média.
22. Um agricultor fez um cálculo do peso da fruta de cada uma das 200 laranjeiras do seu pomar, originando o valor de 6000 Kgs para a produção total. Colheram-se e pesaram-se os frutos de uma amostra PISR de 25 laranjeiras desse pomar comparando-se para cada árvore o peso observado (Y_i) e o peso “palpite” do agricultor (X_i).

Os resultados obtidos foram os seguintes:

$$\sum_{i=1}^{25} y_i = 660 \quad \sum_{i=1}^{25} y_i^2 = 18200 \quad \sum_{i=1}^{25} y_i x_i = 18987$$

$$\sum_{i=1}^{25} x_i = 700 \quad \sum_{i=1}^{25} x_i^2 = 19994 \quad s_y^2 = 32,33 \quad s_x^2 = 16,42 \quad s'_{xy} = 21,125$$

- a) Calcule a estimativa por índice para o peso total da fruta e calcule o erro padrão associado.
- b) O que poderá afirmar acerca da “qualidade” dessa estimativa relativamente à estimativa de 5280 Kgs obtida com base na amostra observada.
- c) Construa um intervalo de confiança a 95% para a diferença entre o peso médio real por árvore e o resultante do palpite do agricultor. O que pode afirmar sobre a “qualidade” do palpite do agricultor.
- d) Utilizando como estimador para a produção média por laranjeira:

$$\hat{\mu}_y = \bar{Y} + (30 - \bar{X}),$$

calcule, com base na amostra recolhida, a estimativa para a produção total do pomar bem como o erro padrão associado.

- e) Obtenha a estimativa pela regressão para a produção total do pomar e calcule o erro padrão associado.
- f) Teça algumas considerações sobre as propriedades dos estimadores associados às estimativas das alíneas d) e e) comentando os resultados obtidos.

23. De uma lista de 468 cursos profissionais foi seleccionada uma amostra aleatória simples sem reposição de 100 cursos: 54 eram ministrados em estabelecimentos oficiais e 46 em particulares. Os dados sobre o número de alunos (y_i) e de professores (x_i) são apresentados abaixo:

Estabelecimentos	n	$\sum y_i$	$\sum x_i$	$\sum y_i^2$	$\sum x_i^2$	$\sum x_i y_i$
Oficiais	54	31 281	2 024	29 881 219	111 090	1 729 349
Particulares	46	13 707	1 075	6 366 785	33 119	431 041
TOTAL	100	44 988	3 099	36 247 904	144 209	2 160 380

- a) Sabendo que nessa lista existem 251 estabelecimentos oficiais obtenha uma estimativa para o número de alunos em cursos profissionais desses estabelecimentos. Qual o erro padrão associado a essa estimativa?
- c) Compare a estimativa obtida na alínea anterior com a que obteria se não conhecesse o total de estabelecimentos oficiais na população. Neste caso qual seria o erro padrão.
24. Seleccionou-se uma amostra aleatória simples sem reposição de 400 agregados familiares dos 3600 existentes numa área urbana. Cada família foi sujeita a um inquérito mais ou menos extenso, de onde se salientam duas questões sobre a habitação onde residem: Regime de propriedade (casa alugada, casa própria ou outra) e despesa mensal (renda ou mensalidade do empréstimo/condomínio ...). Os resultados foram os seguintes:

Regime de propriedade	Número de Agregados	Despesa mensal por agregado	
		Média	Desvio padrão
Casa própria	250	216,0	58,4
Casa alugada	145	195,5	49,3
Outro	5	50,5	30,9
Total	400	206,5	59,5

- a) Construa um intervalo de confiança a 95% para o número de agregados familiares dessa área que residem em casa alugada.
- b) Poder-se-á afirmar que não existe diferença, estatisticamente significativa, para a despesa mensal média com habitação, entre os agregados que residem em casa própria e os que residem em casa alugada.
- c) Calcule uma estimativa para o montante global da despesa mensal com habitação do conjunto dos agregados dessa região que residem em casa própria e determine qual a precisão associada.
25. Pretende-se realizar um estudo sobre uma amostra de 300 explorações agrícolas com o objectivo de avaliar a produção média de leite por exploração (μ). O número total de explorações recenseadas é de 5000, distribuídas por duas regiões: 4000 na Região A e as restantes na Região B. O último recenseamento disponível contém informação sobre as explorações agrícolas, fornecendo a base de sondagem e permitindo a obtenção de valores para os desvios padrão da produção de leite por exploração em 20 e 40 para as regiões A e B, respectivamente.

O estudo tem por objectivo a actualização dos dados do recenseamento, admitindo-se, contudo, que os desvios padrão são da mesma ordem de grandeza. Utilize a região como variável de estratificação, sendo as amostras em cada região seleccionadas sem reposição.

- a) Justifique o método de amostragem escolhido.
- b) Utilizando a afixação proporcional calcule uma aproximação da precisão associada às estimativas da produção média de leite por exploração, por região e para o conjunto das regiões.
- c) Qual a repartição que proporciona a melhor precisão para a estimação de μ . Calcule uma estimativa para essa precisão.
- d) O custo unitário de realização de uma entrevista é de 3 u.m. na Região A e 5 u.m. na Região B. Calcule a dimensão óptima da amostra, global (n) e por região (n_A e n_B), para um orçamento máximo disponível de 1300 u.m.. (Suponha que a amostra é seleccionada com reposição).

e) Sabe-se que, para além da produção de leite por exploração, também interessa estimar a quantidade média de leite produzida por animal pelo que foi recolhida informação sobre o número de animais produtores de leite existentes na exploração. Diga qual a metodologia a seguir para a obtenção da estimativa pretendida. Qual a expressão do estimador e o seu erro padrão.

26. Um jornal tem cerca de 40000 assinantes que, para efeitos da distribuição diária da publicação, se encontram agrupados em 2000 zonas. Com o objectivo de obter o perfil dos seus leitores para assim melhor promover a venda dos seus espaços publicitários, o jornal levou a efeito um inquérito por amostragem junto dos seus assinantes.

Para a selecção da amostra procedeu-se ao sorteio com PISR de 40 das 2000 zonas existentes e inquiriram-se todos os assinantes pertencentes às zonas sorteadas totalizando 800 entrevistas. Um das questões referia-se à propriedade ou não do seu alojamento, tendo-se apurado o seguinte:

$$\sum_{k=1}^{40} y_k = 360 \quad \text{e} \quad \sum_{k=1}^{40} y_k^2 = 3750 \quad \text{onde}$$

y_k – número de assinantes pertencentes à zona k residentes em habitação própria.

- Identifique o método de amostragem utilizado e calcule as estimativas, pontual e por intervalo de confiança a 95%, para o número total de assinantes do jornal com casa própria.
- Se o valor observado (360 assinantes residindo em habitação própria) tivesse sido obtido com base numa amostra aleatória simples sem reposição de 800 dos assinantes, qual seria o intervalo de confiança a 95% para o número total de assinantes do jornal com habitação própria?
- Explique as semelhanças e/ou diferenças constatadas nos intervalos de confiança obtidos nas alíneas anteriores. [NOTA: Se não resolveu a alínea a) utilize o intervalo de confiança (15700; 20300)].

27. Pretende-se saber a quantidade de cimento vendida pelos comerciantes que é utilizado na construção para habitação. Para tal realizou-se um inquérito por amostragem aos empreiteiros que se abastecem junto a esses comerciantes, revendedores de cimento.

Encontra-se disponível a listagem dos comerciantes clientes das cimenteiras no ano anterior, em número de 450, e a quantidade de cimento comprada por cada um deles. Assim, seleccionou-se uma amostra aleatória de 25 revendedores, com probabilidade proporcional à sua dimensão e inquiriram-se todos os empreiteiros seus clientes sobre a quantidade total de cimento consumido em toneladas, no ano passado, em construção para habitação (X), tendo-se obtido os seguintes resultados:

$$\sum_{k=1}^{25} \bar{x}_k = 4750 \quad , \quad \sum_{k=1}^{25} (\bar{x}_k)^2 = 2116790$$

onde \bar{x}_k é a quantidade média de cimento consumido em construção para habitação, por empreiteiro, dos clientes do k -ésimo revendedor da amostra.

- Descreva o plano amostral subjacente a este inquérito identificando, nomeadamente, população objectivo e inquirida e método de amostragem utilizado.
 - Determine o intervalo de confiança a 95% para o consumo médio de cimento para habitação de um empreiteiro dessa população.
 - Mantendo a mesma metodologia de amostragem, qual deverá ser a dimensão da amostra se se pretender estimar o consumo médio de cimento para habitação, por empreiteiro, com um erro máximo de 50 toneladas e uma confiança de 90%.
28. Na Faculdade de Medicina Dentária existe, desde há três anos, uma consulta de estomatologia para pessoas com idade compreendida entre os 3 e os 18 anos. Existe um ficheiro, contendo 2500 fichas (uma ficha por paciente), onde se encontram registados todos os tratamentos dentários prestados. Pretende-se estimar o número médio de dentes careados dos utentes desta consulta e, para tal, recorreu-se a vários planos de amostragem.

- Seleccionou-se uma amostra aleatória PISR de 40 fichas e registou-se o número de dentes careados por pessoa (Y_i), tendo-se obtido: $\sum_{i=1}^{40} y_i = 105$ e $\sum_{i=1}^{40} y_i^2 = 470$. Determine a estimativa para o número médio de dentes careados e o erro padrão associado.

- Extraíram-se do ficheiro oito amostras sistemáticas, com 5 elementos, obtendo-se:

Amostra	1	2	3	4	5	6	7	8
---------	---	---	---	---	---	---	---	---

Média de dentes careados 1,4 5,2 1,8 2,6 1,4 2,8 2,2 3,4

Diga como procederá para obter estas amostras e calcule, tendo por base este plano de amostragem, a estimativa para o número médio de dentes careados bem como o erro padrão associado.

- c) Dividiu-se a população em quatro classes etárias, de acordo com a idade na consulta mais recente, e seleccionaram-se amostras casuais simples dentro de cada classe, tendo-se obtido:

Classe etária	3-6 anos	7-10 anos	11-14 anos	15-18 anos
Percentagem de utentes	10%	20%	40%	30%
Dimensão da amostra	4	8	16	12
Dentes careados: Média	0,5	1,5	2,5	3,5
DP corr.	0,8	1,6	1,1	2,5

- c1) Identifique o tipo de amostragem que está subjacente a este plano e calcule um intervalo de confiança a 95% para o número médio de dentes careados.
- c2) Determine as dimensões das amostras a seleccionar em cada classe etária de modo a constituir uma amostra com igual dimensão mas que forneça um estimador de precisão máxima.
- d) Faça uma análise comparativa dos estimadores utilizados nos três planos de amostragem e apresente as vantagens e os inconvenientes associados a cada uma das metodologias.

29. Numa das linhas de produção da fábrica "ELECTRIK" produzem-se diariamente 5000 embalagens, com 10 fusíveis cada uma.

- a) O controlo de qualidade é efectuado semanalmente e tem por base o seguinte procedimento: selecção de uma amostra aleatória simples sem reposição de 100 embalagens das 25000 produzidas semanalmente, e em cada uma delas, também aleatoriamente e sem reposição, selecção de metade dos fusíveis e registo do número de fusíveis defeituosos (y_i). Numa determinada semana obtiveram-se os seguintes dados:

y_i	0	1	2	3
Nº de embalagens	65	29	5	1

- a1) Descreva, pormenorizadamente, o plano amostral subjacente a este processo de controlo de qualidade, nomeadamente, composição/quantificação da população e da amostra.

- a2) Calcule o intervalo de confiança a 95% para a proporção de fusíveis defeituosos produzidos nessa semana.

- b) Se esse controlo fosse feito a partir de uma amostra aleatória PISR de 50 embalagens da produção semanal, sendo observados todos os seus fusíveis e registado o número de defeituosos por embalagem, quais as semelhanças/diferenças e as vantagens/ desvantagens face ao procedimento de controlo descrito nas alíneas anteriores.

- c) A partir do início do corrente ano, embora continuem a ser observados 500 fusíveis por semana, a metodologia de amostragem utilizada no controlo de qualidade foi modificada, passando a selecção dos fusíveis a ser feita antes do acondicionamento nas embalagens e numa base diária. Assim, diariamente, são seleccionados de forma automática 100 fusíveis: a partir de um primeiro (sorteado aleatoriamente entre os primeiros quinhentos produzidos) selecciona-se um em cada 500 produzidos.

- c1) Identifique e descreva a metodologia subjacente a este plano semanal de amostragem.

- c2) Os resultados obtidos com as amostras recolhidas nos cinco dias da última semana foram os seguintes:

Dia	1	2	3	4	5
% de fusíveis defeituosos	3%	3%	1%	2%	0%

Com base nesta informação calcule uma estimativa para o total de fusíveis defeituosos produzidos nessa semana bem como o erro padrão associado.

30. Realizou-se um inquérito por amostragem aos empreiteiros, clientes de empresas de pré-fabricados, para apurar o valor médio dos pré-fabricados utilizados em obras públicas. Não existe base de sondagem desta população mas dispõe-se da listagem completa das 650 empresas de pré-fabricados existentes, bem como dos volumes de produção de cada uma delas. Assim, seleccionou-se, com probabilidade proporcional aos volumes de produção, uma amostra aleatória de 30 empresas de pré-fabricados e foi-lhes pedido que indicassem a lista dos nomes dos empreiteiros seus clientes. Com base nessa lista, procedeu-se à selecção de uma amostra aleatória de vinte empreiteiros.

Perguntou-se a estes empreiteiros qual o valor total dos pré-fabricados utilizados em obras públicas (X), no ano passado, tendo-se obtido os seguintes resultados:

$$\sum_{K=1}^{30} \bar{x}_k = 1029 \quad \sum_{K=1}^{30} (\bar{x}_k)^2 = 77830$$

onde, \bar{X}_k é o valor médio dos pré-fabricados utilizados em obras públicas por empreiteiro, para os clientes da k-ésima empresa de pré-fabricados, seleccionados para a amostra.

- Identifique o tipo de amostragem que está subjacente a esta recolha de informação e calcule um intervalo de confiança a 95% para o valor médio dos pré-fabricados utilizados em obras públicas de um empreiteiro dessa população.
 - Mantendo a metodologia inicial de selecção, qual deverá ser a dimensão da amostra se pretender estimar o valor médio de pré-fabricados para obras públicas, por empreiteiro, com um erro máximo de ± 10 e uma confiança de 90%.
31. Uma equipa de sociólogos está a estudar vários aspectos sobre comportamentos sociais da população estudantil em colaboração com o Ministério da Educação, que disponibiliza toda a informação existente sobre a população escolar. Um dos aspectos inquiridos diz respeito ao consumo de bebidas alcoólicas. Foi efectuado um primeiro estudo na Região A, onde existem 50 escolas com cerca de 21800 alunos. A amostra foi seleccionada seguindo a seguinte metodologia: em primeiro lugar seleccionou-se uma amostra de cinco escolas, com probabilidade proporcional ao número de estudantes matriculados, tendo seguidamente inquirido uma amostra aleatória simples de 10% dos alunos de cada uma das escolas seleccionadas. Obtiveram-se os seguintes dados:

Escolas seleccionadas	Total de estudantes	Nº de estudantes inquiridos	Consumidores de bebidas alcoólicas
E5	600	60	6
E18	50	5	0
E23	250	25	1
E37	1100	110	22
E40	1500	150	36

- Diga qual o método de amostragem utilizado e identifique qual a população inquirida e quais as unidades estatísticas subjacentes a este inquérito.
 - Diga, justificando cuidadosamente, se se pode afirmar que, com a metodologia seguida, cada aluno tem idêntica probabilidade de vir a integrar a amostra.
 - Calcule uma estimativa para o total de estudantes dessa região que consomem bebidas alcoólicas e determine a precisão relativa associada.
32. Segundo os dados do último censo existem cerca de 9600 agregados familiares no Concelho A. Para levar a efeito um inquérito às despesas das famílias que lá residiam, dividiu-se o concelho em **100** quarteirões e seleccionaram-se aleatoriamente e sem reposição **4**. Em cada quarteirão escolhido, listaram-se todos os agregados existentes e seleccionaram-se **50** agregados familiares, também aleatoriamente e sem reposição. No quadro abaixo apresentam-se alguns dos dados recolhidos:

Quarteirão	Nº de agregados do quarteirão	Gasto médio mensal em saúde, por agregado	Nº de agregados com automóvel
Q012	144	166 €	10
Q048	68	153 €	16
Q089	84	158 €	11
Q099	70	219 €	13

- Poder-se-á afirmar que, com a metodologia seguida, cada um dos agregados familiares desse concelho tem idêntica probabilidade de vir a integrar a amostra? Justifique.
- Determine uma estimativa para o gasto médio mensal em saúde de um agregado familiar dessa cidade. Como calcularia o erro padrão associado a essa estimativa?
- Construa um intervalo de confiança a 90% para o número de agregados com automóvel nessa cidade.

33. A Câmara Municipal de uma pequena cidade quer saber qual a proporção de votantes que se opõem à construção de uma incineradora para queima de lixo, numa zona situada no limite da cidade. Para tal seleccionaram-se aleatoriamente 150 números de telefone residenciais dos 3000 existentes na lista telefónica referente à cidade. Cada residência seleccionada foi contactada telefonicamente tendo sido perguntado: (a) qual número de votantes existentes na residência e (b) qual a opinião de cada um deles acerca do assunto em questão. Assim apurou-se um total de 240 votantes dos quais 93 recusaram responder à questão. Dos restantes 147 votantes, 121 manifestaram-se contrários à construção.

Com base nos dados anteriores foi afirmado o seguinte: “Como conclusão do inquérito realizado podemos afirmar, com uma confiança de 95%, que a percentagem de eleitores desta cidade contrários à construção da incineradora para a queima de lixo, se situa entre os 76,1% e os 88,5%”.

- Com base na afirmação feita calcule a estimativa pontual para a proporção de votantes contrários à construção dessa incineradora bem como o erro padrão que lhe está associado, apresentando as hipóteses subjacentes à afirmação produzida.
- Tendo por base os procedimentos seguidos para recolha da informação, comente a validade das estimativas calculadas, justificando cuidadosamente a metodologia de extrapolação seguida (caso aceite a validade) ou apresentando formas alternativas (caso não aceite) .

34. Numa determinada zona existem 40000 estabelecimentos comerciais e pretende-se estimar qual o gasto mensal médio com a segurança (μ), por estabelecimento. Para tal, foi enviado, por via postal, um questionário a 1000 estabelecimentos seleccionados ao acaso, tendo-se recebido cerca de 600 respostas. De entre os não respondentes (restantes 400), seleccionou-se uma amostra de 60 e procedeu-se à recolha da informação pretendida através de entrevista directa.

- Proponha um estimador não enviesado para μ , que tenha em conta toda a informação recolhida, e indique qual o erro padrão associado.

b) Sabendo que:

- o custo por questionário enviado é de 1,00 € ;
- o custo do tratamento das respostas recebidas é de 2,50 € por questionário;
- o custo unitário associado à recolha e tratamento da informação obtida através de entrevista directa é de 25 €;
- a variabilidade da característica em estudo é idêntica nos dois grupos, respondentes (σ_R) e não respondentes (σ_{NR}), e corresponde a um coeficiente de variação de 50% ($\sigma_R = \sigma_{NR} = \sigma = 0,5\mu$).

b1) Qual deveria ser a taxa de insistência óptima junto aos não respondentes?

b2) Se se pretender estimar μ , com uma confiança de 90% e uma precisão relativa de 3%, que dimensões devem ter a amostra a inquirir por via postal e a subamostra a inquirir por entrevista directa. Qual o custo associado a este dimensionamento?

- Se tiver a garantia de que, para além do já referido acerca da variabilidade, também não existem diferenças significativas na média da característica em estudo, entre os que respondem e os que não respondem, como dimensionaria a amostra a inquirir de modo a, com uma confiança de 90%, obter uma precisão relativa de 3% para μ .

35. Numa escola C+S, com 1500 alunos, foi efectuado um inquérito a uma amostra aleatória simples sem reposição de 200 alunos com o objectivo de estudar os hábitos de lazer. Somente foi possível obter a resposta de 150 alunos e a partir da informação recolhida calculou-se, entre outras variáveis, para cada aluno inquirido o número de horas despendidas semanalmente a ver televisão tendo-se obtido uma média de 32,7 horas e um desvio padrão corrigido de 17,3 horas.

- Calcule, justificando cuidadosamente, um intervalo de confiança a 95% para o número médio de horas que um aluno daquela escola despende semanalmente a ver televisão.

- Os investigadores encarregados do estudo ao analisar a informação recolhida verificaram que tanto a taxa de resposta como o tempo despendido frente à televisão pareciam variar consoante a idade do aluno e, então, agruparam os alunos em três classes etárias apurando os seguintes resultados:

Classe etária	Número de alunos na:		Número de Respostas	Tempo despendido (horas)	
	População	Amostra		Média (\bar{x})	D. Pad. (s')
13 ou menos	350	58	58	40,3	15,1

14 a 16	600	70	62	31,5	18,2
17 ou mais	550	72	30	20,5	10,5
Total	1500	200	150		

Obtenha, tendo em conta esta informação, uma estimativa para o número médio de horas que um aluno daquela escola depende semanalmente a ver televisão e calcule qual a precisão associada.

c) Faça uma análise crítica dos resultados obtidos nas alíneas anteriores.

36. Uma cadeia de supermercados encomendou a uma empresa de sondagens um inquérito a realizar nas suas lojas durante uma semana, com o objectivo de estudar o impacto nos seus clientes da animação sonora das suas lojas: música ambiente, campanhas, promoções, anúncios e etc.. Os questionários são preenchidos no local, por entrevista directa aos clientes sendo os resultados expressos em percentagens simples do tipo: "x% dos clientes recordam-se da promoção feita ao produto A".

A priori, tudo leva a crer que a dimensão da loja tem influência nas respostas recolhidas, pois o ambiente e a clientela são diferentes nas "pequenas" e nas "grandes" lojas.

Designe por N_k o número de clientes que frequentam, na semana em questão, a loja k , $k=1,2,\dots, M$ (M é o número de lojas da empresa). Naturalmente que N_k é desconhecido, mas dispõem-se de informações recentes e precisas sobre o volume médio de vendas semanais X_k (e também das vendas médias por dia da semana).

Admite-se a hipótese de que existe uma forte correlação positiva entre as vendas X_k e o número de clientes N_k , de modo que se verifica:

$$\frac{N_k}{N} \approx \frac{X_k}{X} \quad \text{onde} \quad N = \sum_{k=1}^M N_k \quad \text{e} \quad X = \sum_{k=1}^M X_k$$

O contrato prevê a realização de 500 entrevistas distribuídas por 10 das M lojas da empresa, tendo sido solicitado que a amostra fosse "representativa", no sentido em que "todos os clientes devem ser considerados equivalentes e ter a mesma hipótese de ser inquiridos".

Admita que está a trabalhar nessa empresa de sondagens e é responsável pela elaboração de uma proposta de plano de amostragem para o estudo pretendido.

Elabore o plano que deverá apresentar, nomeadamente, o método de amostragem a seguir, as expressões dos estimadores para as proporções bem como para as suas variâncias. Deverá ainda, qualquer que seja a metodologia seguida, explicitar as regras de selecção dos clientes a inquirir de forma a satisfazer o solicitado.

37. Vamos definir a taxa de actividade como a proporção de indivíduos, com idade entre os 20 e os 64 anos, que trabalham pelo menos 20 horas por semana.

Para estimar o valor da taxa de actividade de uma dada região, onde habitam cerca de um milhão de pessoas desse escalão etário, a informação foi recolhida de acordo com o seguinte esquema.

As 75 freguesias que constituem essa região foram ordenadas de acordo com o número de habitantes que pertencem a esse escalão etário, e depois agrupadas em três estratos de igual dimensão. Assim,

Estrato 1 - as 25 freguesias com menor população no escalão etário em análise;

Estrato 2 - ...

Estrato 3 - as 25 freguesias com maior número de pessoas nesse escalão.

Seleccionaram-se nove freguesias, amostra aleatória simples sem reposição de três freguesias em cada estrato, e em cada uma delas inquiriu-se uma amostra aleatória PISR de 100 indivíduos pertencentes ao grupo etário referido, sobre a sua situação face ao emprego. Obtiveram-se os seguintes resultados:

Estratos	Freguesias da amostra	População com idade entre os 20 e os 64 anos	Nº de pessoas na amostra que trabalham pelo menos 20 horas por semana
1	A	1460	72
	B	2230	72
	C	2820	74
2	D	3350	75
	E	5820	68
	F	4540	73
3	G	9510	75
	H	11760	70
	I	30140	69

- a) Descreva a metodologia subjacente a esta recolha de informação indicando, nomeadamente, a população objectivo, o tipo de variável a estudar e a metodologia de amostragem utilizada.
- b) Deduza o estimador para a taxa de actividade dessa região, bem como para a variância desse estimador e calcule, com base na informação recolhida, a estimativa para essa taxa.
- c) Faça uma análise crítica do desenho utilizado nesta recolha e proponha, caso ache justificável, uma alternativa que, utilizando o mesmo número total de freguesias na amostra, lhe pareça mais eficiente.
38. A Direcção Comercial de uma empresa de Telecomunicações realiza regularmente sondagens de opinião junto dos seus clientes, com o objectivo de avaliar a imagem da empresa e os hábitos de consumo em produtos de telecomunicações.
- Os clientes estão agrupados, consoante a região de residência, em 150 Grupos de Redes e, nas sondagens já realizadas, a amostra de clientes tem sido seleccionada de acordo com a seguinte metodologia:
- seleccionam-se, em primeiro lugar, 15 dos 150 Grupos de Redes, sendo a selecção feita com probabilidade proporcional ao número de clientes existentes em cada um dos Grupos de Redes;
 - em cada um destes 15 Grupos de Redes, selecciona-se, com probabilidade igual e sem reposição, uma amostra de 200 clientes.
- a) O novo Director Comercial da empresa ao tomar conhecimento da metodologia utilizada achou estranho que se inquirisse o mesmo número de clientes (200) em todos os Grupos de Redes pois, segundo ele, “não faz sentido fazer o mesmo número de entrevistas em Grupos de Redes com dimensões tão distintas” e, assim, pôs em causa a “representatividade” da amostra e a validade dos resultados.
- Como “especialista em amostragem” coube-lhe a tarefa de escrever uma resposta, curta mas convincente, explicando e defendendo a metodologia até agora utilizada. Redija essa resposta.
- b) Os resultados dos anos anteriores mostram que as opiniões expressas são muito homogéneas nas zonas rurais e muito variáveis nas zonas fortemente urbanizadas. Elabore um novo desenho para a amostra de forma a integrar esta informação. Supondo que quer estimar uma proporção, por exemplo, a proporção de clientes que possuem um atendedor de chamadas, indique as expressões tanto para o seu estimador como para a precisão associada.
- c) Os resultados mostram também que os hábitos de consumo em produtos de Telecomunicações variam bastante em função da idade e da categoria social dos inquiridos. Com base na informação do último censo é possível saber a repartição da população de cada Grupo de Redes por classe etária e por categoria socioprofissional.
- Formule sugestões, tanto na elaboração dos questionários como nos tratamentos a efectuar posteriormente, que lhe permitam ter em conta estas informações complementares.
39. Vai ser feito um estudo que tem por objectivo caracterizar a população dos “sem abrigo” de Lisboa quanto a vários aspectos, nomeadamente, questões relacionadas com a saúde, opinião sobre os apoios oferecidos e principais carências/necessidades sentidas. Segundo os seus promotores esse estudo será feito com base na informação recolhida junto de uma amostra que se pretende “representativa dos sem abrigo da cidade” de forma a poder garantir “validade estatística aos apuramentos feitos”. Para tal foi disponibilizada toda a informação existente sobre o assunto tanto pela Câmara Municipal de Lisboa como por outras entidades de solidariedade social. Assim, com base em estudos anteriores estima-se em cerca de 1366 o número de “sem abrigo” da cidade utilizando, a grande maioria, os centros de acolhimento/apoio

existentes. Sabe-se também a caracterização desses centros quanto a uma série de variáveis tais como: localização na cidade, tipo de apoio prestado e número de frequentadores.

Como membro da equipa encarregada do estudo compete-lhe fazer um primeiro esboço do plano de amostragem apresentando, nomeadamente: população e unidades estatísticas, método de amostragem, metodologias para a selecção dos inquiridos e recolha de informação bem como informação necessária à boa execução do estudo.

40. O ISEG deseja realizar um inquérito por amostragem junto dos estudantes e licenciados recentes pela Escola visando tanto proceder à caracterização social, demográfica e cultural da população discente da Escola como aferir das suas expectativas, motivações e aspirações. Pretende-se nomeadamente conhecer:
- . *Caracterização social e geográfica da sua população escolar,*
 - . *Condições de vida,*
 - . *Contextos e práticas escolares, percursos e expectativas relativamente ao curso que frequentam,*
 - . *Actividades extra-escolares,*
 - . *Expectativas face ao mercado de emprego e percurso/realização profissional.*
 - . *Perfis dos alunos das várias licenciaturas bem como da pós-graduação.*

Este estudo é promovido pelo próprio Conselho Directivo pelo que os Serviços Académicos disponibilizarão toda a informação existente, não confidencial, sobre os alunos inscritos nos últimos quatro anos.

Como parte integrante da equipa que vai levar a efeito este estudo elabore uma primeira versão da proposta contendo nomeadamente:

A) Concepção e desenho da amostra, identificando justificadamente qual a informação que é necessário obter junto dos Serviços. B) Informação a recolher, exemplificando algumas perguntas a incluir no questionário e metodologias a seguir tanto para a selecção dos inquiridos e recolha de informação, como para obtenção resultados.

RESULTADOS

1. a) F, b) V, c) F, d) V.
5. 65347; 6,44%
6. b) 379 c1) (220529,96 , 325598,04) c2) - 5104
7. (13,581; 15,453)
8. b) acréscimo relativo entre 3,25% e 13,19%.
9. 1537
10. b) 11339
11. 561
12. b) 30 ; 18 ; 2 c2) 0,2256 ; 0,1895 ; 0,0625 (0,09375)
13. a) (20536,52 , 23463,48) b) 131 ; 29, 46, 56 (57)
14. a) 3 ; 29 ; 2 ; 18
15. 2,94% ; 11,1%
16. a) 384 ; 192 b) 400 ; 1600
18. a) (59,032 , 64,368) b) (68,25 , 69,65)
19. a) (441 383,56 , 550 616,44) b) $z_{obs} = 6,412$ logo existe diferença c) (483 531,46 , 600 468,54)
20. a) 25069 ; 7290,342 b) 21862,5
21. a) 86,963 ; 2,45 (ou 88,664 ; 2,425) b) 24
22. a) 5657,143 ; 99,64 b) 5280 ; 212,76 c) (- 2,535 , - 0,665)
d) 5680 ; 95,394 e) 5794,616 ; 84,927
23. a) 145399 (145398,72); 14254,57 b) 146395 (146,395,08) ; 18697,62
24. a) (1144,901 , 1465,099) b) $z_{obs} = 3,922$ existe diferença c) 486000; 19423,21
25. b) 1,252 ; 5,009 ; 1,4160 c) 200 ; 100 ; 1,339 d) 365; 263 ; 102
26. a) 18000 ; (15781,2 , 20218,8) b) (16634 , 19366)

27. b) (101,826 , 278,174) c) 55 (49).
28. a) 2,625 ; 0,3502 b) 2,6; 0,4420 c1) (1,873 , 2,927) c2) 2 ; 8; 11 ;19
29. a2) (0,059 , 0,109) c2) 4500; 1450,31
30. a) (20,595 , 48,005) b) 38
31. b) 163,828 c) (1269,926 , 3092,074)
32. b) não c) 2528,8 ; 39,5%
33. a) 0,823 ; 0,032
34. b1) 0,4082 b2) 1167 ; 191 ; 7692,50€ c) 1230 (1254)
35. a) (30,073 , 35,327) b) 29,52 ; 1,167