

NOME

DATA

PERÍODO

Materiais de apoio à família

Inferências estatísticas

Nesta unidade, o aluno vai usar uma pequena *amostra* de dados para estimar informações sobre um grupo maior, denominado *população*, e usará simulação para determinar um intervalo de valores para a estimativa. Uma população é todo o conjunto de assuntos de interesse para uma questão e uma amostra é um grupo menor, dentro dessa população.



Por exemplo, podemos querer descobrir o valor *médio* que as famílias nos Estados Unidos gastam com alimentação, todos os meses. A população inclui todas as famílias dos Estados Unidos, mas recolher informações de todos seria muito difícil e custaria muito dinheiro, por isso poderíamos começar a recolher dados com uma amostra de 50 famílias.

Uma questão importante a considerar quando se começa a recolher informações de uma amostra é como a amostra será selecionada. Os dados recolhidos podem ser muito diferentes se perguntar às famílias que estão a fazer compras num supermercado local, em comparação com as pessoas à porta de um restaurante chique. Da mesma forma, o valor gasto em alimentação em São Francisco é provavelmente muito diferente do valor gasto na

NOME

DATA

PERÍODO

zona rural do Iowa. Pode até haver alguns hábitos de consumo ocultos de formas sobre as quais ainda não pensamos. Então, como podes ter a certeza de que a tua amostra é representativa das famílias dos Estados Unidos, sem usar muitas famílias de grupos que não são típicos nos seus gastos?

A solução é usar *aleatoriedade*. Podemos selecionar 50 famílias usando um processo aleatório, como fazer com que um computador selecione aleatoriamente as famílias a partir de uma base de dados, sem considerar outros fatores. Isto deverá reduzir o *viés* que pode ser introduzido pelos humanos que tentam obter informações sobre as pessoas e irá, provavelmente, incluir proporções mais precisas dos diferentes tipos de famílias dos Estados Unidos. Embora a aleatoriedade possa não eliminar totalmente o viés da seleção da amostra, irá reduzir, significativamente, o viés presente quando comparada à seleção sem aleatoriedade.

Os investigadores fizeram estudos como esses e descobriram o valor médio gasto com alimentação, por mês. Um relatório diz que o valor médio gasto em alimentação, por mês, é de US\$ 600, com uma *margem de erro* (MoE) de US\$ 150. A margem de erro é usada para dizer que não esperamos que todas as famílias da amostra gastem exatamente US\$ 600.

É importante observar a margem de erro nos resultados estatísticos. É irresponsável discutir estatísticas sem fornecer uma margem de erro para descrever o quanto se espera que o valor varie. Muitos gráficos incluídos em reportagens irão reportá-lo em letras pequenas no gráfico. Procura algo como $\pm 3\%$ quando há um gráfico sobre o índice de aprovação de um funcionário ou pesquisas durante as próximas eleições. Isto significa que as percentagens mostradas no gráfico podem, na verdade, ser até 3% menores ou maiores que o número mostrado.

Aqui fica uma tarefa para experimentar com os alunos:

Uma cidade tem uma votação iminente sobre a possibilidade de aumentar o imposto sobre o rendimento corporativo em 2% para aumentar o financiamento de escolas públicas. O noticiário local mostra uma imagem que indica que 52% da população votante é a favor do aumento de impostos e no canto mostra “margem de erro $\pm 3.5\%$.” O repórter parece confiante de que os impostos sobre as empresas serão aumentados, porque qualquer valor superior a 50% dos votos a favor dos impostos aprovará a lei.

1. O repórter que descobriu o número de 52% chegou a ele ao conduzir até quatro dos 20 vários bairros da cidade e ao perguntar a opinião dos moradores. Há algo de errado com a forma como foi feito? Consegues pensar numa forma melhor de recolher dados?
2. O que significa a margem de erro nesta imagem?
3. Deverias ficar confiante de que os impostos vão aumentar? Explica o teu raciocínio.

Soluções:

NOME

DATA

PERÍODO

1. Ir a apenas 4 bairros da cidade poderia deixar de fora a opinião de muitos eleitores de outros bairros, onde o repórter não foi. Uma melhor forma de recolher informações poderia ser selecionar aleatoriamente vários agregados familiares, da cidade, para perguntar a opinião deles. É mais provável que a seleção aleatória evite qualquer preconceito do repórter sobre quais os bairros a visitar.
2. A margem de erro significa que a percentagem real a favor do aumento dos impostos pode ser 3,5% superior ou 3,5% inferior aos 52% reportados, com base na amostra. Isto significa que a percentagem real ficaria entre 48,5% e 55,5%.
3. Exemplo de resposta:
 - Acho que continua a haver uma boa hipótese de os impostos aumentarem. Embora a percentagem real possa ser tão baixa quanto 48,5%, com base na margem de erro, também pode chegar a 55,5%. A maioria dos percentuais possíveis está acima de 50%, por isso, acho que o aumento vai acontecer.
 - Acho que continua a não estar claro se o aumento vai acontecer. Com base na margem de erro, o percentual real poderia ser tão baixo quanto 48,5%, o que faria com que o aumento não acontecesse. Também não tenho a certeza sobre os métodos utilizados pelo repórter para recolher uma amostra neste relatório, por isso, o relatório pode não ser muito preciso.



© CC BY 2019 by Illustrative Mathematics®