

### Lista nº 4 – Comparações de Médias

#### Instruções:

Há características comuns a todos os exercícios que se seguem que, para abreviar são listadas aqui em vez de em cada questão individual:

- Quando forem fornecidas medidas reais, em vez de somente estatísticas de resumo, espera-se que você *crie os gráficos apropriados aos dados* e comente sobre qualquer coisa de interesse que veja nos gráficos, em particular se observar algum comportamento não normal que possa fazê-lo sentir-se pouco à vontade para usar os testes e os intervalos de confiança padrão. Para lembra-lo desta exigência padrão de *sempre* olhar os dados, será usada com frequência uma frase tal como: “Que aspectos interessantes você vê nestes dados? Isto não significa que haverá sempre algo em particular que dava ser notado, mas apenas que você é encorajado a sempre olhar os dados.
  - Quando o exercício perguntar se ocorre algum efeito ou se há qualquer evidência da existência de um efeito ou diferença, espera-se que seja utilizado um teste de significância, enquanto que quando o exercício pergunta qual o tamanho do efeito ou da diferença, espera-se que seja utilizado um intervalo de confiança. Mais frequentemente a questão perguntará ao mesmo tempo se ocorre um efeito e qual o tamanho de tal efeito, e espera-se que você use ambas as ferramentas na sua resposta.
- 1) Como parte de um estudo para comparar programas de educação física em duas escolas no Canadá, os tempos em corridas (em seg.) de percurso de uma distância fixada foram registrados para amostras independentes de estudantes da sexta série. Os dados são fornecidos no arquivo: *runtime.txt*.
    - a) Há alguma diferença real entre as duas escolas nos tempos médios em corridas? Quão grande é a diferença? Há algo nos dados que o faça duvidar da aplicabilidade das técnicas da teoria normal?
    - b) Use uma técnica não-paramétrica para comparar os tempos medianos em corridas. Isso o conduz às mesmas conclusões?
    - c) As crianças da Escola Primária Gooscap foram treinadas por um técnico especialista em corrida de curta distância. Temos evidência de que as marcas das crianças treinadas por um especialista são melhores do que aquelas treinadas por um professor generalista em educação física? Em que medida um estudo como este responde a tal questão?
  - 2) Um padrão de medida comum usado para testar a efetividade de propagandas de televisão é o escore de lembrança após 24 horas. Os anunciantes são solicitados a lançar ao ar o comercial e entrevistar por telefone, no dia seguinte, uma amostra de telespectadores, digamos 200, que assistiram ao comercial. Pergunta-se aos entrevistados se eles se lembram de ter visto um comercial de uma determinada marca na noite anterior e, caso se lembrem, faz-se uma série de perguntas de acompanhamento. A partir das respostas é calculado um escore de lembranças. Esse é um procedimento um tanto vagaroso, complicado e caro. Se certos tipos de atividades cerebrais fossem fortemente preditivas de lembrança posterior, os profissionais de propaganda poderiam obter muito mais *feedback* imediato sobre o que funciona e o que não funciona, pela medição da atividade cerebral de sujeitos experimentais, enquanto estão assistindo aos comerciais.

A atividade de onda alfa do cérebro é característica de se estar desperto, mas desatento. Com aumento do estímulo, a atividade alfa tende a diminuir à medida que

as ondas alfa são substituídas por ondas de frequência mais alta. Os dados do arquivo *alfa.txt* provêm de um estudo descrito por Appel et al. (1979). As medidas índice de atividade total são médias, para grupos de sujeitos experimentais, de medidas de *redução* da atividade da onda cerebral alfa ao assistirem ao comercial em comparação com a mesma atividade diante de uma tela vazia. Para cada tipo de marca/produto foram usados dois comerciais. Em cada caso, um comercial teve um escore consideravelmente maior do que o outro.

Uma das hipóteses dos pesquisadores antes de começar o estudo era que os índices seriam mais altos para comerciais de alta lembrança (um índice mais alto significa uma maior redução na atividade de onda alfa) do que para comerciais de baixa lembrança.

- a) Há alguma evidência de que os comerciais de alta lembrança tendem a ter índices mais altos de atividade do que os comerciais de baixa lembrança?
  - b) Um resultado significativo em (a) significaria necessariamente que escores de lembrança após 24 horas podem ser descartados e em vez deles usados escores de índice de atividade total?
  - c) Crie um gráfico de dispersão de cada índice de atividade de baixa lembrança contra o valor correspondente de alta lembrança. Que padrão, se qualquer, você vê? Há uma relação forte entre os escores?
  - d) Os experimentadores fizeram quatro cliques. O clique A tinha as marcas 1-5 na ordem 1a, 1b, 2b, 2a, ...; o clique B tinha as marcas 5-1 na ordem 5b, 5a, 4a, 4b, ...; o clique C tinha as marcas 6-10 na ordem 6b, 6a, 7a, 7b, ...; enquanto o clique D tinha as marcas 10-6 na ordem 10a, 10b, 9b, 9a, ... (1a significa um comercial de alta lembrança para a marca 1, 1b significa um comercial de baixa lembrança). Eles dividiram seus 30 sujeitos experimentais em grupos de 5. Cada grupo via uma sequência de cliques rotulada AC, DB, CA, BD, CB ou AD. Cada sujeito foi primeiro exposto a 30 segundos de tela vazia (X), então exposto a cada um dos pares de cliques três vezes, e finalmente exposto a outros 30 segundos de tela vazia. Por exemplo, o grupo AC viu X-A-C-A-C-A-C-X. Por que os pesquisadores tiveram todo este trabalho? Quais são alguns dos fatores contra os quais eles tentaram se resguardar?
- 3) No arquivo *house.txt* estão os preços de moradias (em milhares de dólares) extraídos entre as anunciadas por corretora na cidade de Hamilton, em três momentos diferentes.
- a) Represente graficamente os dados e obtenha médias e desvios padrão para cada um dos três períodos de tempo. O que o gráfico e as estatísticas de resumo sugerem sobre o movimento de preços e moradias ao longo do tempo?
  - b) Faça o teste F. O que os resultados lhe dizem?
  - c) Obtenha intervalos de confiança apropriados para as verdadeiras diferenças entre os preços médios de venda nos três períodos de tempo
  - d) Você vê alguma coisa nos resultados de (a) que o faça duvidar da aplicabilidade do teste F ou de intervalos de confiança para as diferenças, obtidas a partir de um programa de análise de variância?
  - e) Em vez de representar os próprios preços de moradias, represente os logaritmos dos preços de moradias. Os logaritmos dos preços de moradia seguem mais de perto as suposições da teoria normal do que as observações não transformadas?
  - f) Usando os logaritmos dos dados, há qualquer evidência real de que os preços de moradias anunciadas pela corretora mudaram em média? De que maneira mudaram?

Para os propósitos da próxima parte, suponha que os dados precedentes são os preços de venda no lugar dos preços anunciados. Em vez de estar interessado no movimento do preço médio das casas atualmente à venda, as pessoas tendem a estar mais interessadas no que aconteceu com o valor de sua própria casa.

- g) Liste qualquer razão que apoie a ideia de que um aumento no preço médio das casas vendidas pela corretora pode não implicar que qualquer casa em particular na cidade teve seu valor aumentado. Qual destas razões ainda seria aplicável se os números se referissem a todas as moradias que foram vendidas na cidade (e não apenas àquelas anunciadas pela corretora em questão)? Que tipo de informação você desejaria a fim de ter uma ideia melhor sobre as mudanças no valor de moradias específicas?
- 4) Para cada um dos seguintes experimentos, responda as estas questões:
- (i) Especifique o desenho mais apropriado para o experimento. (Escolha um dos seguintes: uma amostra, duas amostras independentes, mais de duas amostras independentes, dados emparelhados.) Especifique que método estatístico – a saber, um teste de hipótese ou intervalo de confiança – é mais apropriado.
  - (ii) Especifique a(s) hipótese(s) que devem ser satisfeitas de modo que você possa aplicar seu método escolhido.
  - (iii) Descreva brevemente um problema potencial com o experimento.
- a) Um instituto de pesquisa está investigando um novo medicamento para aliviar enxaquecas. Eles têm disponíveis 40 sujeitos que sofrem de enxaqueca aguda. O medicamento sendo investigado tem por objetivo fazer efeito imediato e é anunciado com tendo um efeito ponderável em 60 minutos. O grupo de pesquisa decide solicitar a cada paciente que registre a porcentagem de enxaquecas que eles acharam que foram aliviadas em uma hora. Eles querem descobrir se o medicamento é efetivo ou não.
- b) Um horticultor está comparando dois métodos (A e B) de cultivar batatas. Enxertos padrão de batata serão plantados em pequenos canteiros do terreno. As variáveis de interesse (variáveis-resposta) são o número de tubérculos de batata e o peso fresco (peso logo após a colheita) de vegetais cultivados por planta. O horticultor tem 20 canteiros disponíveis para o experimento e está interessado em determinar o tamanho da diferença, se houver, entre os métodos de plantação no peso resultante do vegetal cultivado.
- c) Biólogos estão interessados em executar um experimento para determinar a melhor cor para atrair besouros de folhas de cereais para placas nas quais serão capturados. Quatro cores diferentes devem ser comparadas: amarelo, branco, verde e azul. A variável-resposta é o número de besouros capturados. Uma placa será fixada em cada um de 16 postes igualmente espaçados na parte central de um campo de aveia (16 placas no total, 4 por cor).

Fonte: MOORE, D. S.; MCCABE, G. P; DUCKWORTH, W. M.; SCLOVE, S. L. *A Prática da estatística empresarial: como usar dados para tomar decisões*. Rio de Janeiro: LTC, 2006

### Referência:

APPEL, V.; WEINSTEIN, S.; WEINSTEIN, C. Brain activity and recall of TV advertising. *Journal of Advertising Research*, v. 19, n. 4, p. 7-15, 1979.