

MBA EM **DATA SCIENCE E ANALYTICS**

MATERIAL COMPLEMENTAR

Fundamentos de Estatística

Professor Wilson Tarantin Junior

LISTA DE EXERCÍCIOS COMPLEMENTARES

1. Na análise de concessão de empréstimos, uma variável potencialmente importante é a renda da pessoa. O gerente de um banco coleta uma base de dados de seus correntistas e extrai a variável “renda mensal (R\$)” para 50 pessoas. Embora se trate de uma variável quantitativa, deseja realizar uma análise por meio de tabela de frequências. Neste sentido, pede-se:

a) Classifique os correntistas em faixas de renda, sendo: 0-2.000; 2.001-4.000; 4.001-6.000; 6.001-8.000; 8.001-10.000 e 10.001-12.000.

b) Em seguida, elabore a tabela de frequências para as faixas de renda acima.

O banco de dados está na planilha Lista de Exercício Complementares: aba Exercício 1.

2. Um analista do mercado acionário coletou os retornos mensais de duas ações que pretende indicar aos seus clientes. Calcule as estatísticas descritivas para as duas variáveis. O banco de dados com os retornos percentuais mensais está na planilha Lista de Exercício Complementares: aba Exercício 2. Posteriormente, analise o coeficiente de correlação de Pearson entre os retornos.

3. Em certo jogo, probabilidade de vitória (sucesso) a cada nova jogada é $1/6$. Se forem feitas 10 jogadas, quais são as seguintes probabilidades:

a) Ter vitória em 4 jogadas.

b) Ter vitória em pelo menos 7 jogadas.

4. (Fonte: Fávero e Belfiore, 2024, Cap. 5) Suponha que um aluno acerte três questões a cada cinco testes. Seja X o número de tentativas até o décimo segundo acerto. Determine a probabilidade de que o aluno precise fazer 20 questões para acertar 12.

5. Suponha que o número de acidentes de trânsito em determinada cidade seja, em média, de 5 acidentes por dia e que siga uma distribuição Poisson. Calcule

a probabilidade de que, em tal cidade, ocorram 9 acidentes de trânsito em um dia.

6. Nos últimos meses, foram feitas medições do tempo decorrido entre o início e finalização de uma das etapas do processo de produção de certo produto. O tempo médio foi calculado em 26,5 minutos e o desvio padrão foi de 4,0 minutos. Sabendo que tal variável segue uma distribuição normal, identifique as seguintes informações:

a) $P(X > 37)$

b) $P(X < 20)$

c) $P(22 < X < 28)$

7. (Fonte: Fávero e Belfiore, 2024, Cap. 8) Um grupo de 60 leitores fez uma avaliação de três livros de romance e, ao final, escolheram uma das três opções. Teste a hipótese nula de que não há diferença na preferência dos leitores, ao nível de significância de 5%.

Os dados estão disponíveis na planilha Lista de Exercício Complementares: aba Exercício 7.

8. Foram coletados dados sobre a quantidade de chuva por dia (em milímetros), durante 14 dias, para dois locais distintos. O pesquisador notou que o local A aparenta ter maior variabilidade na quantidade chuva, comparativamente ao local B. Neste sentido, deseja testar a hipótese de que a variabilidade na quantidade de chuva é significativamente maior em A do que em B. Realize o teste F para testar tal hipótese.

Os dados estão disponíveis na planilha Lista de Exercício Complementares: aba Exercício 8.

9. (Fonte: Fávero e Belfiore, 2024, Cap. 7) Deseja-se comparar o tempo médio de espera para atendimento (min) em 2 hospitais. Para isso, coletou-se uma amostra com 20 pacientes em cada hospital. Verifique se há diferenças entre os tempos médios de espera nos dois hospitais. Considere $\alpha = 1\%$.

Adicionalmente, calcule os intervalos de confiança para o tempo médio de espera nos dois hospitais.

Os dados estão disponíveis na planilha Lista de Exercício Complementares: aba Exercício 9.

10. (Fonte: Fávero e Belfiore, 2024, Cap. 7) Um fabricante de iogurtes desnatados afirma que a quantidade de calorias em cada pote é 60 cal. Para verificar se essa informação procede, uma amostra aleatória com 36 potes é coletada, observando-se que a quantidade média de calorias é de 65 cal com desvio-padrão 3,5. Aplique o teste adequado e verifique se a afirmação do fabricante é verdadeira, considerando o nível de significância de 5%.

11. Em uma indústria alimentícia, para a preparação de certo produto, a quantidade padrão de sal são 8g com desvio padrão igual a 1g. Na análise de uma amostra de 85 produtos, foi observada uma quantidade média de sal de 8,25g. Avalie se a quantidade observada na amostra pode ser considerada estatisticamente maior, ao nível de significância de 5%.

12. Em uma amostra de 425 medições, observou-se que o peso médio dos produtos que foram medidos é 226g, com desvio padrão de 10g. Estabeleça o intervalo de confiança para esta média, sabendo que o desvio padrão observado é exatamente seu valor histórico. Considere um nível de confiança de 95%.

Referência

Fávero, Luiz Paulo; Belfiore, Patrícia. (2024). Manual de análise de dados: estatística e machine learning com Excel®, SPSS®, Stata®, R® e Python®. 2 ed. Rio de Janeiro: LTC.