

-- CONHECIMENTOS ESPECÍFICOS --

Considerando um espaço de probabilidade no qual estejam definidos os eventos aleatórios A_x e B_x e, para $x \in \{1, 2, 3, \dots\}$, e as probabilidades $P(A_x \cup B_x) = 1$, $P(A_x \cap B_x) = 0,6^x$ e $P(B_x) = 0,8^x$, julgue os itens a seguir.

- 51 O conjunto $\{1, 2, 3, \dots\}$ representa o espaço amostral da situação em questão.
- 52 A sequência de eventos $B_1, B_2, B_3, \dots$ constitui uma distribuição de probabilidades.
- 53 $P(A_x) = 0,75^x$
- 54 $P(A_x | B_x) = 0,75^x$
- 55 A_x e B_x são eventos independentes.

Julgue os próximos itens, considerando uma variável aleatória contínua X e uma constante real apropriada θ , tal que

$$P(X > x) = \begin{cases} 1, & \text{se } x < 0, \\ \theta x^2 - x + 1, & \text{se } 0 \leq x \leq 2, \\ 0, & \text{se } x > 2. \end{cases}$$

- 56 O valor esperado de X é igual a $\frac{2}{3}$.
- 57 A variável aleatória $Y = X - \theta X^2$ segue a distribuição uniforme contínua no intervalo $[0, 1]$.
- 58 A moda da distribuição X é igual a zero.
- 59 A função de densidade de probabilidade de X é

$$f(x) = \begin{cases} 1 - 2\theta x, & \text{se } 0 \leq x \leq 2, \\ 0, & \text{se } x < 0 \text{ ou } x > 2. \end{cases}$$

- 60 θ representa o parâmetro da distribuição X .

Uma auditoria realizada por um tribunal de contas identificou divergências entre as despesas previstas e as despesas efetivamente realizadas em contratos administrativos de determinado órgão público.

Para a avaliação estatística dessas divergências, considerou-se a variável aleatória Y , definida como a diferença, em milhares de reais, entre a despesa realizada e a despesa prevista em cada contrato. Valores positivos indicam estouro orçamentário, enquanto valores negativos indicam economia.

Deseja-se verificar se há, em média, evidência de estouro orçamentário populacional (μ) por meio do teste de hipóteses $H_0: \mu = 0$ versus $H_1: \mu > 0$.

Uma amostra aleatória simples de 40 contratos resultou em uma média amostral $\bar{y} = 12$ e desvio padrão amostral $s = 30$.

Sabendo que Y segue uma distribuição aproximadamente normal, julgue os próximos itens.

- 61 Como Y segue distribuição normal, a estatística do teste dada pela expressão $\frac{\bar{y}}{30/\sqrt{40}}$ segue uma distribuição normal padrão.
- 62 Supondo que o nível descritivo do teste seja, aproximadamente, igual a 0,008, a hipótese nula (H_0) é rejeitada, o que leva à conclusão de que, estatisticamente, $\mu = 12$.
- 63 Se o intervalo de 95% de confiança para a média populacional for, aproximadamente, $12 \pm 9,6$, e se o nível de significância do teste de hipóteses em questão for de 2,5%, então a hipótese nula do teste em questão deve ser rejeitada, uma vez que $\bar{y} = 12 > 9,6$.

- 64 À medida que o valor verdadeiro da média populacional μ se afasta de zero no sentido positivo, o poder do teste de hipóteses em questão aumenta.

- 65 A variância amostral é um estimador não viciado da variância populacional, no entanto, ela não apresenta erro quadrático médio mínimo entre outros possíveis estimadores.

Considere um modelo de regressão linear simples na forma

$$y = \beta_0 + \beta_1 x + \varepsilon,$$

no qual x representa o valor total, em milhões de reais, contratado para execução de obras públicas e y denota o percentual de aditivos contratuais (em percentagem), entendido como o aumento percentual do valor inicialmente contratado. O termo ε refere-se ao erro aleatório com média zero e variância σ^2 . Por meio de uma amostra aleatória de 27 contratos, obtiveram-se os seguintes resultados, fornecidos pelo método dos mínimos quadrados ordinários.

estimativa de β_1	1,5
coeficiente de determinação do modelo	81%
soma de quadrados total	500

Com base nessas informações, julgue os itens subsequentes.

- 66 A razão entre o desvio padrão amostral de y e o desvio padrão amostral de x é igual a $5/3$.
- 67 A estimativa da variância do erro aleatório (σ^2) é igual a 3,8.
- 68 A estatística do teste t de Student para as hipóteses $H_0: \beta_1 = 0$ versus $H_1: \beta_1 \neq 0$ é superior a 10.
- 69 A correlação linear entre y e x é superior a 0,95.
- 70 A variância amostral de y foi inferior a 20.

Um tribunal de contas deseja realizar um levantamento do andamento das obras públicas em determinada região a fim de estimar os percentuais de obras adiantadas e atrasadas em relação ao cronograma. Cada obra encontra-se classificada em uma única categoria. Devido à grande quantidade de categorias existentes, a equipe técnica decidiu concentrar o estudo em quatro delas: (i) portos e aeroportos, (ii) energia elétrica, (iii) estradas e rodovias e (iv) saneamento básico. Essa escolha foi motivada pelo elevado volume de reclamações e denúncias recebidas ao longo de um ano. De cada uma dessas categorias, será selecionada uma amostra aleatória simples de 100 obras para avaliar o andamento em relação ao cronograma.

Com respeito a essa situação hipotética, julgue os itens a seguir.

- 71 A probabilidade de seleção de cada obra na amostra é igual entre as categorias, uma vez que foram selecionadas 100 obras de cada uma delas.
- 72 O procedimento descrito caracteriza uma amostragem por conglomerados, pois foram selecionados grupos específicos de obras para compor a amostra.
- 73 Cada categoria de obra constitui uma unidade amostral.
- 74 A fração amostral referente a cada categoria de obras é igual a $1/100$, em que o denominador corresponde ao tamanho da amostra.
- 75 Na estimação do percentual de obras adiantadas no saneamento básico por amostragem aleatória simples, o erro padrão do estimador deve ser, no máximo, igual a 10%.

Acerca de dados, informação, conhecimento e gestão de dados, julgue os itens a seguir.

- 76 Na recuperação de dados, o ETL e o ELT são processos idênticos.
- 77 O armazenamento de dados é uma etapa isolada no ciclo de gestão, isto é, sem relação com as etapas de coleta, tratamento e recuperação das informações.
- 78 A recuperação de dados estruturados pode ser otimizada com a criação de índices.
- 79 Dados abertos são disponibilizados publicamente para que qualquer pessoa possa acessá-los, utilizá-los e redistribuí-los livremente, sem restrições de licença.
- 80 Dados não estruturados, como *emails* e vídeos, têm formato fixo e podem ser facilmente armazenados em tabelas de banco de dados relacional.

No que se refere a bancos de dados relacionais, julgue os itens seguintes.

- 81 Um índice, em um banco de dados relacional, funciona de forma semelhante ao índice de um livro: ele permite localizar registros de maneira muito rápida, sem a necessidade de percorrer linha por linha de toda a tabela.
- 82 Em um banco de dados relacional, os dados são organizados em tabelas compostas por linhas e colunas, em que cada linha representa um registro único, e cada coluna, um atributo desse registro.
- 83 Uma visão (*view*) é uma tabela virtual baseada no resultado de uma consulta SQL, não armazena dados fisicamente e apresenta informações de uma ou mais tabelas de forma organizada e reutilizável.
- 84 Metadados são os próprios dados armazenados nas tabelas do banco, ou seja, são os registros de clientes, produtos e transações que o sistema utiliza no dia a dia.
- 85 A chave primária de uma tabela pode conter valores nulos e repetidos, pois sua função é apenas identificar o nome da tabela dentro do banco de dados.

A respeito da modelagem dimensional, julgue os itens que se seguem.

- 86 Fatos aditivos, semiaditivos e não aditivos referem-se ao mesmo tipo de medida em uma tabela fato, diferindo entre si apenas na granularidade com que foram registrados, sendo todos igualmente aplicáveis a qualquer tipo de agregação entre dimensões.
- 87 Uma dimensão SCD (*slowly changing dimension*) do tipo 2 preserva o histórico completo de mudanças de um atributo, criando um novo registro para cada alteração, mantendo inalterado o registro anterior e permitindo análises históricas precisas.
- 88 Em um *data warehouse* modelado dimensionalmente, a tabela fato armazena métricas numéricas do negócio e se relaciona com as tabelas de dimensão por meio de chaves estrangeiras, formando a estrutura central do modelo.
- 89 No modelo estrela (*star schema*), as tabelas de dimensão são normalizadas em múltiplos níveis, o que reduz a redundância de dados e melhora o desempenho das consultas analíticas, em comparação ao modelo floco de neve (*snowflake schema*).

Julgue os itens subsequentes, relativos a mineração de dados.

- 90 A modelagem preditiva por regressão linear assume que existe uma relação linear entre as variáveis independentes e a variável-alvo, o que a torna adequada para qualquer tipo de dado, até mesmo séries temporais com sazonalidade acentuada e dados com relações altamente não lineares.
- 91 Na detecção de anomalias, um ponto é considerado *outlier* não apenas por ter um valor absoluto extremo, mas também por apresentar um comportamento estatisticamente improvável em relação à distribuição dos demais dados, o que significa que um valor numericamente moderado pode ser uma anomalia, a depender do contexto e da distribuição da variável analisada.
- 92 O algoritmo *apriori*, utilizado na geração de regras de associação, opera em duas etapas: primeiro, identifica todos os *itemsets* frequentes que atingem o suporte mínimo; em seguida, gera regras a partir desses *itemsets*, avaliando sua confiança. Seu principal gargalo de desempenho, entretanto, caracteriza-se pela necessidade de realizar múltiplas varreduras na base de dados e pelo crescimento exponencial dos *itemsets* candidatos em bases com muitos itens distintos.
- 93 Na etapa de pré-processamento, a normalização e a padronização são técnicas distintas: a normalização muda a escala dos valores para um intervalo fixo (geralmente entre 0 e 1), enquanto a padronização transforma os dados para que tenham média zero e desvio padrão igual a 1, sendo esta segunda mais adequada quando os dados possuem *outliers* severos.
- 94 A clusterização é uma técnica supervisionada de mineração de dados, pois depende de rótulos previamente definidos para agrupar os registros em *clusters* homogêneos, funcionando de forma similar à classificação.
- 95 Em tarefas de classificação, ocorre *overfitting* quando o modelo aprende com excessiva precisão os padrões do conjunto de treinamento, incluindo o ruído, e, por isso, apresenta alta acurácia nos dados de treino, mas desempenho significativamente inferior em dados novos, evidenciando baixa capacidade de generalização.

Em relação a modelagem preditiva, aprendizado de máquina e mineração de texto, julgue os próximos itens.

- 96 As redes neurais artificiais são inspiradas no funcionamento do cérebro humano e organizam-se nas camadas de entrada, que recebe os dados brutos, ocultas (*hidden layers*), que extraem representações progressivamente mais abstratas, e de saída, que produz a predição final, de forma que o aumento indiscriminado de camadas sempre melhora o desempenho do modelo.
- 97 Em modelagem preditiva, os dados são divididos em conjuntos de treino, validação e teste: o conjunto de treino é utilizado para ajustar os parâmetros do modelo; o de validação, para comparar algoritmos diferentes, ajustar hiperparâmetros e decidir quando parar o treinamento; e o de teste, para estimar o desempenho real do modelo escolhido em dados completamente novos.
- 98 Na mineração de texto, a técnica TF-IDF (*term frequency-inverse document frequency*) atribui maior peso às palavras que aparecem com mais frequência em todos os documentos da coleção, pois, quanto mais uma palavra se repete no *corpus* inteiro, mais relevante ela é para distinguir o conteúdo de um documento específico.
- 99 O aprendizado supervisionado requer que os dados de treinamento sejam previamente rotulados com as respostas corretas, de maneira a permitir que o modelo aprenda a relação entre as variáveis de entrada e a variável-alvo; o aprendizado não supervisionado, por sua vez, busca padrões nos dados sem nenhuma indicação prévia de resultado esperado.

100 Em mineração de texto, a tokenização divide o texto em unidades menores (palavras ou frases), *stopwords* são palavras de alta frequência e baixo valor semântico, que são removidas, e *stemming* reduz as palavras à sua raiz morfológica, podendo gerar formas que não sejam palavras válidas no idioma.

Considerando os fundamentos da auditoria do setor público e os instrumentos de fiscalização, julgue os itens a seguir.

- 101** Devido ao princípio da independência plena do controle externo, é vedado a auditor de tribunal de contas considerar ou utilizar em seu planejamento os resultados de trabalhos produzidos pela unidade de auditoria interna da entidade fiscalizada, sob risco de anulação do parecer técnico por vício de parcialidade.
- 102** A auditoria do setor público cumpre papel essencial na governança democrática, pois fornece informações independentes e baseadas em evidências sobre a gestão de recursos públicos, visando aumentar o grau de confiança dos cidadãos e demais usuários previstos nos relatórios da administração pública.
- 103** O objetivo da auditoria do setor público limita-se à verificação da legalidade e da conformidade contábil, não abrangendo a avaliação da eficácia e da eficiência dos programas governamentais, visto que o mérito administrativo é insuscetível de controle externo.

Acerca dos tipos de auditoria e das normas de auditoria, julgue os itens que se seguem.

- 104** O conceito de risco de auditoria na auditoria financeira, conforme a ISSAI 200, refere-se à probabilidade de o auditor ser processado judicialmente por erros em seu relatório, independentemente de as demonstrações financeiras conterem ou não distorções relevantes.
- 105** A ISSAI 100 estabelece a estrutura de princípios fundamentais que regem a auditoria do setor público, fornecendo a base para as normas específicas de auditoria financeira, operacional e de conformidade, e é aplicável a todos os trabalhos de auditoria realizados pelas instituições superiores de controle (ISC).
- 106** Na auditoria de conformidade, o trabalho do auditor pode ser realizado sob duas perspectivas: a de atestação, em que o responsável pela gestão declara a conformidade e o auditor verifica essa declaração; ou a de reporte direto, na qual o próprio auditor avalia o objeto e relata a conformidade diretamente aos usuários.
- 107** A auditoria operacional deve abster-se de emitir recomendações propositivas ao final do relatório, limitando-se a apontar as falhas de eficiência encontradas, a fim de garantir a neutralidade do tribunal de contas e evitar interferência no mérito administrativo do órgão auditado.

Em relação ao planejamento da auditoria governamental, julgue os itens seguintes.

- 108** Na auditoria do setor público, a determinação da materialidade deve fundamentar-se em critérios quantitativos, sendo a avaliação de aspectos qualitativos, como a natureza social ou política do objeto, restrita à fase de comunicação dos resultados.
- 109** No planejamento de auditoria, se o auditor elevar o nível de materialidade para uma conta contábil, ele assumirá maior tolerância a erros, o que, tecnicamente, permitirá uma redução na extensão dos testes substantivos de detalhe sem comprometer o risco de auditoria.
- 110** A matriz de planejamento é o instrumento que organiza as questões de auditoria, os respectivos procedimentos de coleta e análise de dados, as fontes de informação e as possíveis limitações, servindo como roteiro estruturado para a fase de execução.
- 111** A utilização da amostragem não estatística é considerada pela ISSAI 100 falha técnica grave que prejudica a validade das conclusões do relatório, uma vez que a referida norma exige a aplicação de métodos probabilísticos em todos os testes de conformidade do setor público.
- 112** Na definição do escopo da auditoria, são estabelecidos os limites do trabalho, incluindo-se o período de tempo abrangido, as unidades administrativas e os processos específicos que serão submetidos ao exame do auditor.

Julgue os itens subsequentes, relativos à execução e ao monitoramento da auditoria do setor público.

- 113** O monitoramento de uma auditoria operacional deve ser considerado concluído e o processo deve ser arquivado após o gestor auditado protocolar o plano de ação, uma vez que a responsabilidade pela implementação das melhorias sugeridas é exclusiva da administração, cessando a jurisdição do órgão de controle sobre o caso.
- 114** A inspeção física de ativos imobilizados no curso da execução da auditoria permite ao auditor obter evidência direta e fidedigna sobre a existência e a condição física do bem, contudo tal procedimento, isoladamente, é insuficiente para atestar a propriedade jurídica do ativo pela entidade pública, sendo necessária complementação por meio do exame documental.
- 115** A técnica de revisão analítica consiste na conferência aritmética detalhada de cálculos de impostos e somas de documentos fiscais, com o propósito de atestar a exatidão matemática de processos de liquidação da despesa pública.
- 116** A circularização, também denominada confirmação externa, consiste na obtenção de resposta escrita direta de um terceiro independente sobre informações relativas a saldos ou transações, sendo considerada técnica de alta confiabilidade por originar-se de fonte externa à entidade auditada.
- 117** Caso os testes de observância revelem que os controles internos da entidade são ineficazes para prevenir ou detectar distorções relevantes, o auditor deve, por imposição normativa, abster-se de realizar testes substantivos e deve declarar a impossibilidade de emitir opinião sobre o objeto, sob pena de vício na caracterização dos achados.
- 118** Por serem documentos que fundamentam a opinião do auditor, os papéis de trabalho são de propriedade exclusiva da entidade auditada, devendo o tribunal de contas proceder à sua devolução integral logo após a homologação do relatório final.

Acerca das evidências de auditoria e da comunicação dos resultados de uma auditoria governamental, julgue os próximos itens.

- 119** No âmbito da auditoria do setor público, a caracterização de um achado depende fundamentalmente da identificação de uma discrepância entre a situação encontrada pelo auditor, denominada condição, e o padrão de desempenho ou legalidade esperado, denominado critério.
- 120** A evidência de auditoria é considerada adequada quando fornece suporte razoável para a caracterização dos quatro atributos de um achado, sendo sua confiabilidade influenciada pela sua fonte e natureza, conforme as circunstâncias específicas em que foi obtida pelo auditor no curso dos testes.

Espaço livre
