



FIOCRUZ

Concurso Público Fiocruz 2023

Pesquisador em Saúde Pública

Prova Discursiva

PE64 - Métodos Quantitativos



Questão 01

Uma equipe de pesquisa dedicada à investigação das condições relacionadas à saúde da mulher tem como objetivo a construção de um modelo preditivo para óbito em mulheres diagnosticadas com câncer de mama. Inicialmente, a equipe optou por empregar um algoritmo de árvore de decisão devido à sua simplicidade e representação intuitiva. Diversas variáveis explicativas, tais como idade, nível socioeconômico, comorbidades, estágio do câncer, tipo de tratamento, histórico familiar, índice de massa corporal, tipo histológico do tumor, status hormonal, presença de metástases, resposta ao tratamento inicial, marcadores genéticos, entre outras, foram disponibilizadas como preditores para o modelo.

Com base no contexto fornecido, desenvolva um texto com o mínimo de 50 linhas e o máximo de 150 linhas, considerando os itens abaixo:

- a) Explique de maneira detalhada o funcionamento de um algoritmo de árvore de decisão, descrevendo seus principais componentes e etapas de construção do modelo (divisão da amostra em treinamento e teste, critério de divisão, critério de parada, ajuste de hiperparâmetros, avaliação do modelo etc.).
- b) Uma abordagem para lidar com o problema de alta variância frequentemente observada em modelos de árvores de decisão é o uso de técnicas de *ensemble* (combinação de modelos). Detalhe um método que poderia ser utilizado, explicando seus princípios básicos, vantagens e desvantagens.
- c) Qual outro método de classificação poderia ser utilizado para o problema em questão? Justifique.

Questão 02

Um grupo de pesquisadores elaborou um modelo teórico para avaliação do TSH (hormônio estimulador da tireoide) em recém-nascidos (RN) pré-termos (idade gestacional menor de 37 semanas) e/ou com muito baixo peso ao nascer (menos de 1500 g). Este hormônio é um importante indicador do funcionamento da tireoide e encontra-se envolvido em inúmeros processos relacionados ao crescimento, metabolismo entre outros. Para análise dos dados, os RN foram classificados segundo as curvas de crescimento da Organização Mundial de Saúde (OMS), as quais permitem avaliar, em termos de desvios padrão da curva normal, a diferença entre o peso atual do RN e o que seria esperado considerando a idade gestacional e o sexo. Esta comparação fornece uma medida padronizada conhecida como z-score. Assim, os RN com z-score menor de -1,28 foram classificados como “Pequenos para a Idade Gestacional” (PIG) e os demais como sendo “Adequados para a Idade Gestacional” (AIG). Desta forma, o modelo teórico para avaliação dos níveis de TSH (aferido em mU/l) considerou variáveis como o sexo (masculino/feminino), gestação múltipla (sim/não), peso de nascimento (g), idade gestacional (semanas), adequação do nascimento (PIG/AIG), tipo de parto (vaginal/cesariana), sepse (sim/não), enterocolite necrosante (sim/não), broncodisplasia pulmonar (sim/não), hemorragia intracraniana (sim/não), dieta zero (dias) e taxa metabólica basal (kcal/dia).

Onde mU/l – miliunidades por litro; g – gramas; kcal/dia – quilo calorias por dia

Com base no texto, atenda ao solicitado nos itens abaixo utilizando o mínimo de 50 e o máximo de 150 linhas.

- a) Considerando que o modelo teórico é composto por variáveis categóricas e numéricas, descreva as hipóteses que podem ser formuladas a partir do conjunto de variáveis consideradas para a avaliação dos níveis de TSH no grupo de RN estudados, detalhando, para cada hipótese, o teste estatístico e seus respectivos pressupostos.
- b) Partindo da premissa que os pesquisadores irão abordar o problema a partir de um modelo linear generalizado (MLG), descreva e justifique uma possível família de distribuição de probabilidade que possa ser utilizada para descrever o TSH em função das variáveis de exposição, quais funções de ligação podem ser utilizadas, avaliação dos pressupostos, análise e interpretação dos coeficientes, qualidade do ajuste do modelo etc.?
- c) A correlação existente entre o peso de nascimento e idade gestacional ou condição de nascimento com peso, idade gestacional e sexo pode contribuir para a ocorrência de um fenômeno conhecido como multicolinearidade. Desta forma, explique como este fenômeno pode impactar no ajuste do modelo final, quais as formas de avaliação deste fenômeno e como lidar com ele.

Rascunho da Questão 01

RASCUNHO

Rascunho da Questão 02

RASCUNHO

Instruções - Questões Discursivas

1. Cada questão discursiva deverá ter um Limite mínimo de 50 linhas e máximo de 150 linhas.
2. Transcreva sua resposta para a parte pautada no Caderno de Respostas. Não assine, rubrique ou coloque qualquer marca que o identifique, sob pena de ter sua prova anulada. Assim, a detecção de qualquer marca identificadora no espaço destinado à transcrição do texto definitivo acarretará nota ZERO na respectiva prova discursiva.
3. O tempo total de duração da prova será de 4 (quatro) horas, incluindo o tempo para o preenchimento da Resposta Definitiva da Questão Discursiva. Nenhum rascunho **SERÁ LEVADO EM CONTA**.
4. Verifique se a prova é para o **PERFIL** para o qual concorre.
5. Somente após autorizado o início da prova, verifique se este Caderno de Questões está completo e em ordem. **Folhear o Caderno de Questões antes do início da prova implica na eliminação do candidato.**
6. Verifique, no **Caderno de Respostas**, se seu nome, número de inscrição, identidade e data de nascimento estão corretos. Caso contrário, comunique ao fiscal de sala.
7. O rascunho do **Caderno de Questões** poderá ser utilizado para anotações, mas somente as respostas assinaladas no **Caderno de Respostas** serão objeto de correção.
8. Observe as seguintes recomendações relativas ao **Caderno de Respostas**:
 - . não haverá substituição por erro do candidato;
 - . não pode ser dobrado, amassado, rasurado, manchado ou conter qualquer registro fora dos locais destinados às respostas;
9. O fiscal não está autorizado a alterar quaisquer dessas instruções.
10. Você só poderá retirar-se da sala após 60 minutos do início da prova.
11. Quaisquer anotações só serão permitidas se feitas no **Caderno de Respostas**.
12. Os três últimos candidatos deverão permanecer na sala até que o último candidato entregue o **Caderno de Respostas**.
13. Ao terminar a prova, entregue ao fiscal de sala, **obrigatoriamente**, o **Caderno de Questões** e o **Caderno de Respostas**.