

CADERNO DE QUESTÕES



HOSPITAL DE
CLÍNICAS
PORTO ALEGRE RS

EDITAL N° 03/2024

DE PROCESSOS SELETIVOS (PS)

Cargo de Nível Superior

PS 20 - MÉDICO I
(Epidemiologia)

MATÉRIA	QUESTÕES	PONTUAÇÃO
Conhecimentos Específicos	01 a 25	0,40 cada

ATENÇÃO

Transcreva no espaço apropriado da sua FOLHA DE RESPOSTAS (Folha Óptica), com sua caligrafia usual, considerando as letras maiúsculas e minúsculas, a seguinte frase:

O que você faz hoje pode melhorar todos os amanhã.

Nome do Candidato: _____

Inscrição n°: _____

DIREITOS AUTORAIS RESERVADOS. PROIBIDA A REPRODUÇÃO, AINDA QUE PARCIAL, SEM A PRÉVIA AUTORIZAÇÃO DA FAURGS E DO HCPA.

INSTRUÇÕES

- 1 Verifique se este CADERNO DE QUESTÕES corresponde ao Processo Seletivo para o qual você está inscrito. Caso não corresponda, solicite ao Fiscal da sala que o substitua.
- 2 Esta PROVA consta de **25** (vinte e cinco) questões objetivas.
- 3 Caso o CADERNO DE QUESTÕES esteja incompleto ou apresente qualquer defeito, solicite ao Fiscal da sala que o substitua.
- 4 Para cada questão objetiva, existe apenas **uma** (1) alternativa correta, a qual deverá ser assinalada na FOLHA DE RESPOSTAS.
- 5 O candidato que comparecer para realizar a prova **não deverá, sob pena de ser excluído do certame, portar relógios, armas, malas, livros, máquinas calculadoras, fones de ouvido, gravadores, pagers, notebooks, telefones celulares, pen drives** ou quaisquer outros tipos de aparelhos eletrônicos, nem utilizar véus, bonés, chapéus, gorros, mantas, lenços, aparelhos/próteses auditivas, óculos escuros, ou qualquer outro adereço que lhes cubra a cabeça, o pescoço, os olhos, os ouvidos ou parte do rosto, **exceto em situações autorizadas pela Comissão do Concurso e/ou em situações determinadas em lei.** (conforme subitem 7.10 do Edital de Abertura)
- 6 **É de inteira responsabilidade do candidato comparecer ao local de prova munido de caneta esferográfica, preferencialmente de tinta azul, de escrita grossa, para a adequada realização de sua Prova Escrita. Não será permitido o uso de lápis, marca-textos, réguas, lapiseiras/grafites e/ou borrachas durante a realização da prova.** (conforme subitem 7.15.2 do Edital de Abertura)
- 7 Não será permitida nenhuma espécie de consulta em livros, códigos, revistas, folhetos ou anotações, nem o uso de instrumentos de cálculo ou outros instrumentos eletrônicos, exceto nos casos em que forem pré-estabelecidos no item 13 do Edital. (conforme subitem 7.15.3 do Edital de Abertura)
- 8 Preencha com cuidado a FOLHA DE RESPOSTAS, evitando rasuras. Eventuais marcas feitas nessa FOLHA a partir do número **26** serão desconsideradas.
- 9 Ao terminar a prova, entregue a FOLHA DE RESPOSTAS ao Fiscal da sala.
- 10 A duração da prova é de **duas horas e trinta minutos (2h30min)**, já incluído o tempo destinado ao preenchimento da FOLHA DE RESPOSTAS. Ao final desse prazo, a FOLHA DE RESPOSTAS será **imediatamente** recolhida.
- 11 **O candidato somente poderá se retirar da sala de prova uma hora (1h) após o seu início. Se quiser levar o Caderno de Questões da Prova Escrita, o candidato somente poderá se retirar da sala de prova uma hora e meia (1h30min) após o início. O candidato não poderá anotar/copiar o gabarito de suas respostas de prova.**
- 12 **Após concluir a prova e se retirar da sala, o candidato somente poderá utilizar os sanitários nas dependências do local de prova se for autorizado pela Coordenação do Prédio e se estiver acompanhado de um fiscal.** (conforme subitem 7.15.6 do Edital de Abertura)
- 13 Ao concluir a Prova Escrita, o candidato deverá devolver ao fiscal da sala a Folha de Respostas (Folha Óptica). Se assim não proceder, será excluído do Processo Seletivo. (conforme subitem 7.15.8 do Edital de Abertura)
- 14 A desobediência a qualquer uma das recomendações constantes nas presentes instruções poderá implicar a anulação da prova do candidato.



01. Considere as seguintes afirmações sobre a “curva ROC” (*Receiver Operator Characteristics*) para desenvolver modelos preditivos de desfechos em bases de dados com a utilização de aprendizado de máquinas.

- I - A curva ROC é usada para discriminação e calibração dos modelos.
- II - A curva ROC pode ser usada para comparar diferentes modelos para identificar qual dentre eles tem melhor desempenho.
- III - A curva ROC só é utilizada para testes de diagnóstico, não podendo ser utilizada para modelos preditivos de desfecho.

Quais estão corretas?

- (A) Apenas I.
- (B) Apenas II.
- (C) Apenas III.
- (D) Apenas I e II.
- (E) I, II e III.

02. Considere as seguintes afirmações sobre vieses em revisões sistemáticas e meta-análises.

- I - O viés é um erro sistemático, e a imprecisão refere-se ao erro aleatório.
- II - Estudos metaepidemiológicos demonstram que as estimativas de efeito foram mais baixas em ensaios registrados prospectivamente em comparação com ensaios não registrados ou registrados retrospectivamente.
- III - Dois tipos de vieses devem ser considerados ao se realizarem revisões sistemáticas e meta-análises: vieses nos estudos individuais incluídos na revisão sistemática e vieses na síntese dos resultados dos estudos incluídos.

Quais estão corretas?

- (A) Apenas I.
- (B) Apenas II.
- (C) Apenas III.
- (D) Apenas I e II.
- (E) I, II e III.

03. Qual é a afirmação **INCORRETA** a respeito de revisões sistemáticas e meta-análises?

- (A) Publicações duplicadas têm potencial de vieses se os estudos forem incluídos inadvertidamente mais de uma vez em uma meta-análise.
- (B) Ainda há falta de clareza na indexação de estudos primários de prognóstico de doenças, e existe variabilidade nos delineamentos utilizados; portanto, os resultados de revisões sistemáticas destes estudos não são confiáveis.
- (C) Revisões sistemáticas e meta-análises podem incluir estudos não aleatórios de intervenções.
- (D) A meta-análise é a combinação estatística de resultados de dois ou mais estudos separados.
- (E) É adequado incluir diferentes medidas de desfechos relatados pelo paciente em uma mesma meta-análise.

04. Numere a segunda coluna de acordo com a primeira, relacionando as situações descritas com os conceitos.

- (1) Viés de seleção.
 - (2) Fator de confusão.
 - (3) Viés de mensuração.
 - (4) Validade externa.
 - (5) Erro aleatório.
- () Um estudo sobre tabagismo e demência encontra menor taxa relativa de demência em fumantes com o aumento da idade.
 - () Estudo sobre neuroestimulação cerebelar em pacientes com encefalopatia crônica não evolutiva encontra melhora da função motora que não persistiu após finalização.
 - () Aplicação do resultado de um estudo a outros contextos.
 - () Baixa repetibilidade da mensuração.
 - () Não adesão do participante de pesquisa à intervenção do protocolo.

A sequência numérica correta de preenchimento dos parênteses da segunda coluna, de cima para baixo, é

- (A) 1 – 3 – 4 – 5 – 2.
- (B) 3 – 5 – 2 – 4 – 1.
- (C) 1 – 5 – 4 – 3 – 2.
- (D) 4 – 3 – 1 – 5 – 2.
- (E) 2 – 3 – 4 – 5 – 1.

05. No Reino Unido, entre 1999 e 2012, a prevalência e a incidência de lúpus eritematoso sistêmico foi de 97,04/100.000 pessoas e de 4,91/100.000 pessoas, respectivamente. Os dados foram obtidos retrospectivamente de uma base de dados com registros de dados assistenciais e mantiveram-se estáveis no período. Qual é a duração estimada da doença entre os acometidos?

- (A) 10,5 anos.
- (B) 4,9 anos.
- (C) 19,8 anos.
- (D) 50 anos.
- (E) Não é possível estimar.

06. Considere as seguintes afirmações sobre estudos epidemiológicos.

- I - Falta de acurácia que resulta de variação aleatória não pode ser reduzida com uma amostra maior para mensuração.
- II - A capacidade de um estudo em detectar pequenas diferenças entre dois grupos de pacientes (por exemplo, um grupo controle e um grupo de intervenção) é influenciada pelo número de participantes.
- III- Aumentar o número de comparações entre variáveis de um estudo diminui a probabilidade de falsas associações.

Quais estão corretas?

- (A) Apenas I.
- (B) Apenas II.
- (C) Apenas III.
- (D) Apenas I e II.
- (E) I, II e III.

07. Qual das seguintes bibliotecas é amplamente utilizada para manipulação de dados em R?

- (A) ggplot2
- (B) dplyr
- (C) shiny
- (D) purrr
- (E) modelr

08. No contexto de R, para que a função ggplot() é utilizada?

- (A) Manipular dados.
- (B) Exportar dados.
- (C) Importar dados.
- (D) Definir matrizes.
- (E) Criar visualizações gráficas.

09. Considere as seguintes afirmações sobre o tidyverse.

- I - É um conjunto de bibliotecas do R.
- II - Facilita a instalação e o carregamento de pacotes principais do tidyverse em um único comando.
- III- É exclusivo para visualização de dados.

Quais estão corretas?

- (A) Apenas I.
- (B) Apenas II.
- (C) Apenas III.
- (D) Apenas I e II.
- (E) Apenas II e III.

10. Relacione as funções do pacote dplyr com o que elas realizam.

- (1) mutate
- (2) select
- (3) filter
- (4) group_by
- (5) arrange

- () Agrupa dados de acordo com uma variável ou mais.
- () Seleciona colunas de uma base de dados.
- () Ordena uma base de dados de acordo com uma variável ou mais.
- () Seleciona linhas de uma base de dados de acordo com um ou mais critérios.
- () Adiciona novas variáveis na base de dados.

A sequência numérica correta de preenchimento dos parênteses da segunda coluna, de cima para baixo, é

- (A) 4 – 2 – 5 – 3 – 1
- (B) 5 – 3 – 4 – 1 – 2
- (C) 3 – 4 – 1 – 5 – 2
- (D) 1 – 5 – 2 – 3 – 4
- (E) 2 – 1 – 3 – 4 – 5

11. Qual é o propósito da função pivot_longer() no tidyr?

- (A) Transformar dados de formato longo para largo.
- (B) Combinar múltiplas colunas em uma única coluna.
- (C) Dividir uma coluna em múltiplas colunas.
- (D) Organizar dados em uma tabela de resumo.
- (E) Resumir múltiplas linhas em uma única linha.

12. Qual é o retorno da função abaixo escrita em python, caso passemos o valor 4?

```
def func (x):
    y = x % 2 + 3**2
    return y;
```

- (A) Retorna um erro.
- (B) Retorna o valor 9.
- (C) Retorna o valor 10.
- (D) Retorna o valor 11.
- (E) Retorna x.

13. Dados o vetor "v" de números inteiros e o vetor "z", informe a saída do código em python escrito abaixo.

```
v = [1, 2, 3, 4]
z = v[::-1]
print (z)
```

- (A) [1, 2, 3]
- (B) [2, 3, 4]
- (C) [-1, -2, -3, -4]
- (D) [0, 1, 2, 3]
- (E) [4, 3, 2, 1]

14. Considere o código em python descrito abaixo e a matriz "m" representada pelos números inteiros. Analise e responda o que a função "f" exerce sobre a matriz e qual a saída do código:

```
import numpy as np
m = np.array([1,2,3],[4,5,6],[7,8,9])
f = lambda x: x.max( ) - x.min( )
resultado = np.apply_along_axis(f,1,m)
print (resultado)
```

- (A) [1,4,7]
- (B) [3,6,9]
- (C) [2,2,2]
- (D) [3,3,3]
- (E) [6,6,6]

15. Considere o seguinte banco de dados relacional abaixo, denominado Tabela_diabetes. Realize uma consulta SQL (*Structured Query Language*) que retorne os pacientes diabéticos, do sexo feminino, ordenados por idade:

Tabela_diabetes

Prontuario	Nome	Data_Nascimento	Idade	Sexo	Peso	Glicose	Diabetes
123456	Carlos S S	06/05/1946	78	M	100	120	Não
564743	Pedro J C	01/03/2000	24	M	87	160	Sim
457800	Maria O S	05/08/1959	65	F	56	178	Sim
890765	Luis A G	03/12/1969	55	M	90	96	Não
786864	Lia N	13/04/1972	52	F	74	223	Sim
908675	Carla M X	01/06/1966	58	F	98	126	Sim

Considere o comando SQL abaixo e marque qual a alternativa correta.

- (A) FROM Tabela_diabetes SELECT Prontuario, Nome, Data_Nascimento, Idade, Sexo, Peso, Glicose, Diabetes WHERE Diabetes = 'Sim' and SEXO = 'F' ORDER BY Idade ASC;
- (B) SELECT Prontuario, Nome, Data_Nascimento, Idade, Sexo, Peso, Glicose, Diabetes FROM Tabela_diabetes WHERE Diabetes = 'Sim' and SEXO = 'F' ORDER BY Idade DESC;
- (C) SELECT Prontuario, Nome, Data_Nascimento, Idade, Sexo, Peso, Glicose, Diabetes WHERE Diabetes = 'Sim' and SEXO = 'F' ORDER BY Idade ASC;
- (D) FROM Tabela_diabetes SELECT Prontuario, Nome, Data_Nascimento, Idade, Sexo, Peso, Glicose, Diabetes WHERE Diabetes = 'Sim' and SEXO = 'F' AND Idade = "DESC";
- (E) SELECT Prontuario, Nome, Data_Nascimento, Idade, Sexo, Peso, Glicose, Diabetes FROM Tabela_diabetes WHERE Diabetes = 'Sim' and SEXO = 'F' AND Idade DESC;

16. Considere as afirmações em relação aos métodos não paramétricos no contexto de altas dimensões.

- I - Não é necessário fazer suposições adicionais para que seja possível estimar bem uma função de regressão.
- II - Support Vector Regression com o *kernel* Gaussiano tem bom desempenho sob a hipótese de redundância.
- III- Duas suposições usuais que frequentemente valem na prática são esparsidade e redundância nas covariáveis.

Quais estão corretas?

- (A) Apenas I.
- (B) Apenas II.
- (C) Apenas III.
- (D) Apenas I e II.
- (E) Apenas II e III.

17. Qual das seguintes abordagens é mais comumente utilizada para evitar o *overfitting* em modelos de aprendizado de máquina?

- (A) Aumentar o tamanho do conjunto de dados de treinamento.
- (B) Reduzir a complexidade do modelo.
- (C) Utilizar um modelo com mais parâmetros.
- (D) Aumentar o número de épocas durante o treinamento.
- (E) Ignorar a validação cruzada.

18. No contexto de aprendizado de máquina supervisionado, o que é um conjunto de validação?

- (A) O conjunto de dados utilizado para estimar o risco preditivo.
- (B) O conjunto de dados utilizado para treinar o modelo.
- (C) Um subconjunto do conjunto de treinamento usado para ajustar hiperparâmetros.
- (D) Dados que não são utilizados no processo de modelagem.
- (E) Dados que são rotulados, mas que não são utilizados na construção do modelo.

19. Qual é o principal objetivo de um modelo de Support Vector Machine (SVM) na classificação?

- (A) Minimizar a média dos erros de previsão.
- (B) Encontrar a linha ou o hiperplano que maximiza a margem entre as classes.
- (C) Agrupar dados semelhantes em classes.
- (D) Calcular a função de custo para otimização de parâmetros.
- (E) Transformar dados em alta dimensão em um espaço de menor dimensão.

20. Considerando os modelos de aprendizagem de máquina, associe os modelos (coluna 2) ao tipo de aprendizagem supervisionada e não supervisionada (coluna 1).

- (1) Aprendizado supervisionado.
- (2) Aprendizado não supervisionado.
- () Algoritmo vizinhos mais próximos (k-nearest neighbors - KNN).
- () Algoritmo de árvores aleatórias (random forest - RF).
- () Algoritmo K-means (K-médias).
- () Algoritmo máquina de vetor de suporte (support vector machine - SVM).

A sequência numérica correta de preenchimento dos parênteses da segunda coluna, de cima para baixo, é

- (A) 1 – 2 – 2 – 1.
- (B) 2 – 2 – 2 – 1.
- (C) 1 – 1 – 2 – 1.
- (D) 2 – 1 – 1 – 2.
- (E) 1 – 1 – 1 – 2.

21. Qual é um dos objetivos do aprendizado não supervisionado?

- (A) Prever rótulos para dados novos com base em dados rotulados.
- (B) Criar um número "pequeno" de variáveis Z_1, Z_2, \dots a partir de X_1, X_2, \dots que resumem bem as informações presentes nelas.
- (C) Minimizar a função de perda com dados previamente rotulados.
- (D) Ajustar um modelo linear a um conjunto de dados.
- (E) Estimar a densidade de probabilidade de um conjunto de dados.

22. Com relação às redes neurais, assinale **V** (verdadeiro) ou **F** (falso) nas afirmações abaixo.

- () As redes neurais são compostas por uma camada de entrada, uma camada oculta e uma camada de saída.
- () Uma estrutura de camadas que tem a camada de entrada diretamente conectada à camada de saída não é considerada uma rede neural.
- () O número de neurônios das camadas é definido pelo número de entradas.
- () O número de saídas é obrigatoriamente igual ao número de entradas da rede.

A sequência correta de preenchimento dos parênteses, de cima para baixo, é

- (A) V – V – V – F.
- (B) V – F – F – F.
- (C) F – F – V – V.
- (D) F – V – V – F.
- (E) V – F – V – V.

23. Considere as afirmações abaixo sobre métodos de classificação.

- I - Métodos baseados em "Ensemble", os comitês de votantes são características dos algoritmos árvores aleatórias (*random forest* - RF), *gradient boosting* e vizinhos mais próximos (k-nearest neighbors - KNN).
- II - Nas florestas aleatórias (*random forest*), cada árvore é treinada para corrigir os erros realizados pelas árvores anteriores.
- III- Métodos de *bagging* utilizam um subconjunto de dados aleatórios do banco de treino e têm como um dos objetivos reduzir o *overfitting*.

Quais são corretas?

- (A) Apenas I.
- (B) Apenas II.
- (C) Apenas III.
- (D) Apenas I e III.
- (E) I, II e III.

24. Considere as afirmações abaixo com relação ao algoritmo Multilayer Perceptron (MLP).

- I - MLP é um algoritmo de aprendizado não supervisionado.
- II - O algoritmo de *backpropagation* é utilizado para validar o modelo no *dataset* de teste.
- III- Para classificação, o MLP é treinado utilizando o gradiente descendente.
- IV- O MLP tem a vantagem de aprender modelos não lineares.

Quais estão corretas?

- (A) Apenas I e II.
- (B) Apenas I e III.
- (C) Apenas II e III.
- (D) Apenas III e IV.
- (E) I, II, III e IV.

25. Com relação às árvores de decisão, assinale **V** (verdadeiro) ou **F** (falso) nas afirmações abaixo.

- () Podem ser utilizadas para classificação e regressão.
- () Podem ser sensíveis ao *overfitting*.
- () Têm a vantagem de serem estáveis a modificações nos dados.
- () Podem ser consideradas um modelo de "caixa branca".

A sequência correta de preenchimento dos parênteses, de cima para baixo, é

- (A) V – V – V – F.
- (B) V – V – F – V.
- (C) F – F – V – V.
- (D) F – V – V – F.
- (E) V – F – V – V.



HOSPITAL DE
CLÍNICAS
PORTO ALEGRE RS

EDITAL Nº 03/2024 DE PROCESSOS SELETIVOS

GABARITO APÓS RECURSOS

PROCESSO SELETIVO 20

MÉDICO I (Epidemiologia)

01.	B	11.	B	21.	B
02.	E	12.	B	22.	B
03.	ANULADA	13.	E	23.	C
04.	ANULADA	14.	C	24.	D
05.	C	15.	B	25.	B
06.	B	16.	E		
07.	B	17.	B		
08.	E	18.	A		
09.	D	19.	B		
10.	A	20.	C		