Teste de Conhecimento de Conceitos

2020

Instruções

- As respostas às perguntas deste teste fazem parte da lista de tarefas para avaliação. Há outras tarefas para avaliação além desta. Verifique as instruções completas de avaliação no wiki da disciplina [1].
- Todas as questões são de identificação da única alternativa **incorreta**. Portanto marque apenas a alternativa que julgar incorreta em cada caso.
- Escolha **cinco** das nove questões para responder. Inclua suas respostas no documento com as respostas às outras tarefas da lista, como está explicado nas instruções da avaliação.

Questões: identifique a alternativa incorreta

1. Uma ecóloga usou armadilhas de intercepção e queda (*pitfall*) para amostrar uma espécie de sapo em uma área. Com esta metodologia ela obteve (I) o número de indivíduos capturados em cada armadilha; (II) o número de machos em relação ao total capturado em cada armadilha; (III) a massa corporal de cada indivíduo adulto capturado; (IV) o número de dias até a captura do primeiro indivíduo em cada armadilha. Supondo que as capturas são eventos independentes, são descrições adequadas para estas medidas as variáveis aleatórias:

   (a) Uniforme para (II).
   (b) Poisson ou binomial negativa para (I).
   (c) Geométrica para (IV).
   (d) Normal para (III).

2. No contexto desta disciplina os modelos estocásticos ou probabilísticos devem ser entendidos como:

   (a) Uma formulação matemática que incorpora elementos relativos à incerteza sobre o comportamento de variáveis observadas.
   (b) Uma descrição do comportamento de variáveis em termos de distribuições de probabilidade.

[1] http://cmq.esalq.usp.br/BIE5781/doku.php?id=historico:2020:avaliacao

(c) Uma distribuição de probabilidade de uma variável resposta cujos parâmetros são funções de variáveis preditoras.

(d) Qualquer formulação matemática a partir de uma teoria ou hipótese.

3. A média amostral:

   (a) É uma variável aleatória.

   (b) É o mesmo que a esperança estatística.

   (c) É o estimador de máxima verossimilhança do parâmetro μ da normal.

   (d) É o estimador de máxima verossimilhança do parâmetro λ da Poisson.

4. Considere a função de densidade de probabilidade da variável aleatória exponencial:

$$f(x) = \lambda e^{-\lambda x}$$

   Podemos afirmar que:

   (a) $\ln(0,5\, e^{-0,5\, x}) = \ln 0,5 - 0,5x$ é a função de log-verossimilhança para o parâmetro $\lambda = 0,5$.

   (b) $f(x) = 0,5\, e^{-0,5\, x}$ nos dá a densidade probabilística de uma variável exponencial com parâmetro $\lambda = 0,5$.

   (c) $g(\lambda) = \lambda\, e^{-\lambda\, 1,2}$ é a função de verossimilhança para a observação $x = 1,2$.

   (d) A probabilidade de x estar entre 1 e 2 é a integral da função de densidade neste intervalo:

$$P(1 \leq x \leq 2) = \int_1^2 f(x)dx$$

5. O modelos lineares generalizados Poisson:

   (a) Impõem que a variância da resposta seja maior que seu valor esperado.

   (b) Podem ser expressos como $Y \sim \text{Poisson}(\lambda = f(X))$.

   (c) Podem ser considerados um caso particular de um modelo de variável binomial negativa.

   (d) São adequados para contagens de eventos independentes.

6. A função de verossimilhança:

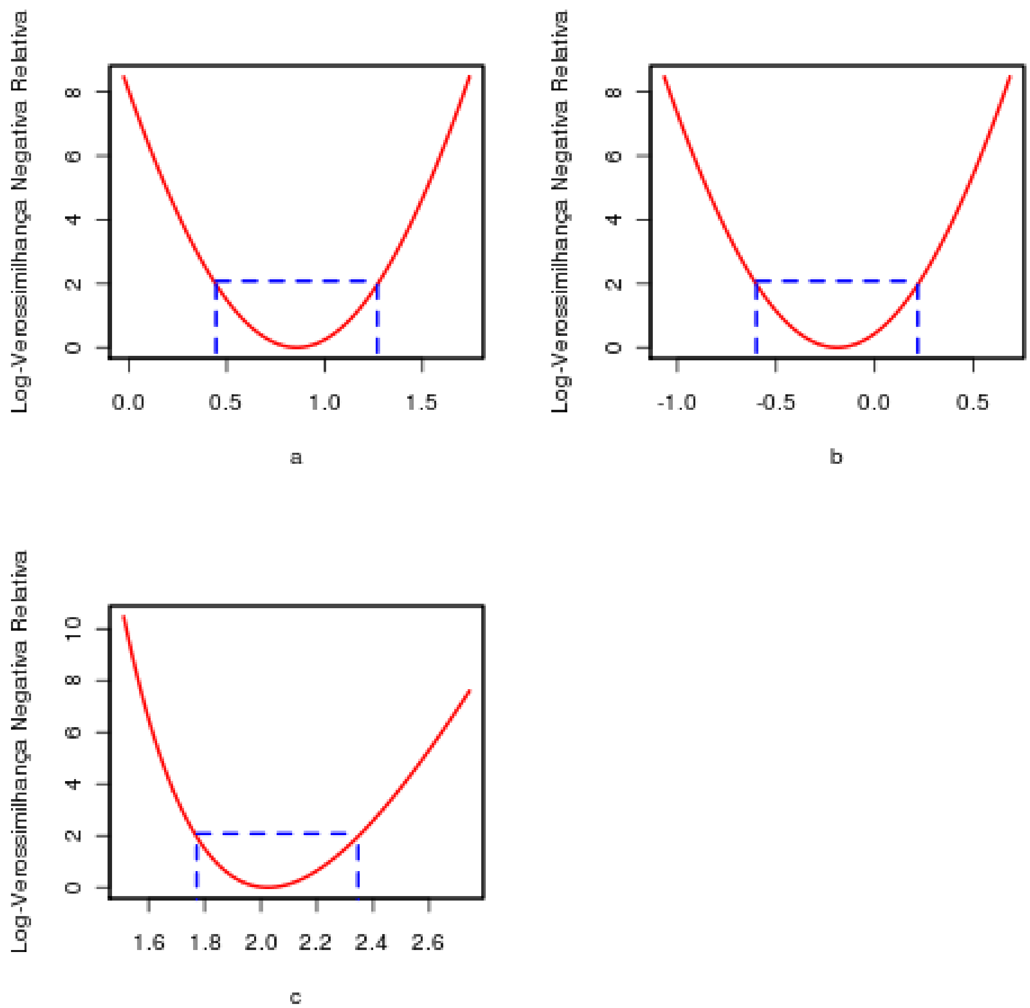
   (a) É proporcional à função de densidade probabilística em um modelo de variável aleatória contínua.

   (b) Expressa a probabilidade do modelo ser verdadeiro.

   (c) É proporcional ao produto das probabilidades atribuídas pelo modelo às observações independentes.

   (d) Ao ser maximizada gera as as estimativas de máxima verossimlhança (MLEs).

7. Considere o modelo $Y \sim$ Normal$(\mu = a + bX, \sigma = c)$, sendo X uma variável preditora contínua e os perfis de log-verossimilhança negativa dos parâmetros desse modelo para um certo conjunto de dados a figura a seguir:



Podemos afirmar que:

(a) As projeções das linhas pontilhadas no eixo X delimitam os valores dos parâmetros que resultam em modelos até e^2 tão plausíveis quanto o obtido com as MLEs.

(b) A estimativa de máxima verossimilhança (MLE) de cada parâmetro do modelo está no centro do seu intervalo de plausibilidade.

(c) É plausível que X não tenha efeito sobre o valor esperado de Y.

(d) Neste modelo o valor previsto de Y é independente de sua variância.

8. O Critério de Informação de Akaike

   (a) É uma medida da distância relativa de um modelo proposto ao modelo verdadeiro.

   (b) É uma medida da discrepância do ajuste do modelo proposto aos dados.

   (c) Expressa a probabilidade de um modelo ser mais plausível que outro.

   (d) É função da máxima log-verossimilhança do modelo proposto.

9. Considerando a tabela abaixo, podemos afirmar que:

Modelo	**AIC**
(1) $Y \sim$ Gauss($\mu = a, \sigma = b$)	3790
(2) $Y \sim$ Gauss($\mu = \beta_0 + \beta_1 X, \sigma = b$)	485
(3) $Y \sim$ Gauss($\mu = \beta_0 + \beta_1 X, \sigma = \alpha_0 X^{\alpha_1}$)	252
(4) $Y \sim$ Log-Normal($\mu = \beta_0, \sigma = c$)	3808
(5) $Y \sim$ Log-Normal($\mu = \beta_0 + \beta_1 \ln(X), \sigma = c$)	398

   (a) É plausível um efeito linear de X sobre o valor esperado de Y.

   (b) A distribuição Gaussiana é um modelo melhor que a Log-Normal para descrever o comportamento de Y.

   (c) O melhor modelo corresponde a uma regressão linear Gaussiana simples.

   (d) É plausível que a variância de Y na sua escala original não seja constante em relação a X.