CE-315: Inferência Bayesiana — 1ª Prova (10/09/2026)

GRR: _____ Nome: _____

Nas questões abertas, quando pertinente, utilize o R para realizar os cálculos (por exemplo, obtenção de probabilidades, quantis, intervalos ou resumos da distribuição a posteriori), indicando os comandos utilizados.

1 Questão

Suponha que o tempo em minutos necessário para atender um cliente em um banco possui uma distribuição exponencial com parâmetro θ. Para fins de planejamento de serviços deseja-se conhecer melhor o processo, o que é feito por inferências sobre θ. Baseada em conhecimento prático e subjetivo, a incerteza sobre θ é descrita por uma distribuição (*a priori conjugada*) de probabilidades para os tempos de atendimento. Para conhecer melhor o processo vai se coletar uma amostra registrando o tempo gasto em n atendimentos.

a) Declare o modelo bayesiano para o problema, escrevendo também as expressões das distribuições envolvidas.
b) Escreva a função de verossimilhança para uma amostra $(y_1, ..., y_n)$ de n observações.
c) Reconheça uma distribuição de probabilidades proporcional a esta expressão da verossimilhança. Qual é a distribuição e com quais valores de parâmetros?
d) Baseado em conhecimento prático e subjetivo acredita-se que a incerteza sobre θ é *a priori* descrita por uma distribuição Gama com média 0,5 e desvio padrão 0,25. Encontre os parâmetros da priori conjugada.
e) Obtenha a expressão da posteriori e indique qual a distribuição de probabilidades.
f) Obtenha a expressão da distribuição marginal (preditiva a priori). Identifique a distribuição como uma das listadas no formulário.
g) Obtenha o valor esperado da distribuição marginal.
h) Descreva ao menos duas formas de simular valores da distribuição marginal.
i) Qual(is) quantidade(s) calculada(s) da amostra são necessárias para a obtenção da verossimilhança e posteriori?
j) O estudo foi feito em uma primeira agência e o tempo médio observado para atender 20 clientes foi de 10,5 minutos. Forneça a expressão da distribuição a posteriori para θ com estes valores.
k) Faça um gráfico mostrando as distribuições priori, posteriori e também a verossimilhança padronizada.
l) Qual o valor esperado da distribuição a posteriori para θ?
m) Qual a probabilidade (a posteriori) de $\theta \leq 1/6$?
n) Obtenha um intervalo de credibilidade (95%) para θ.
o) Mostre que o valor esperado da posteriori é uma média ponderada entre o valor esperado da priori e o estimador de máxima verossimilhança.
p) Qual seria a distribuição preditiva para uma nova observação após obter esta primeira amostra?
q) Qual o valor esperado da distribuição marginal $[Y]$? Descreva três formas de obter este valor.
r) Em uma segunda agência, o estudo foi feito com 12 clientes e o tempo médio de atendimento foi de 8,5 minutos. Qual a posteriori após considerar também os dados desta segunda agência? Que suposição é feita?
s) Faça um gráfico mostrando a distribuição marginal e as preditivas após a primeira amostra e após a segunda amostra.

2 Questão

Dentre usuários que se cadastram em uma plataforma, 20% declaram ter menos de 25 anos e acredita-se que destes, 10% utiliza perfil falso. Entre 25 e 50 anos são 50% dos usuários, com 20% deles sendo perfis falsos. Acima de 50 anos são 30% dos usuários e 5% deles com perfis falsos. Toma-se um usuário ao acaso e verifica-se que possui perfil falso.

a) Sem fazer contas, qual faixa etária você espera ser a deste usuário? Justifique.
b) Qual a probabilidade de ser de cada uma das faixas etárias? Resolva o problema com fundamentos de probabilidade/Teorema de Bayes.
c) Considere agora o problema do ponto de vista de inferência bayesiana.
   - Qual é a variável resposta (Y) e sua distribuição?
   - Qual é o parâmetro (θ) e sua distribuição a *priori*?
   - Qual a verossimilhança? E a *posteriori*? E a marginal (preditiva)?
d) Faça um gráfico da priori e da posteriori.
e) Interprete os resultados.
f) Agora que resolveu o problema, a qual faixa etária você espera que o usuário pertença? Justifique.

g) As proporções de perfis masculinos são de 15%, 30% e 55% para as faixas etárias de menos de 25 anos, entre 25 e 50 anos e acima de 50 anos, respectivamente. Se o usuário de perfil falso selecionado for do sexo masculino, qual a probabilidade de pertencer a cada uma das faixas etárias?

3 Questão

O número diário de acidentes em um trecho de rodovia vai ser modelado por uma distribuição de Poisson com priori "half-normal" (meia-normal), com $\sigma = 2$. O conjunto de dados é $0, 3, 2, 1, 0, 2$.

a) Mostre a expressão da posteriori (a menos de uma constante) e indique se ela é (ou não) compatível com alguma distribuição conhecida.
b) Opta-se por fazer uma análise com aproximação discreta com cinco valores para o parâmetro: 0,5, 1,5, 2,5, 3,5 e 4,5. Obtenha a posteriori (discreta), na forma de uma tabela e também um gráfico com priori e posteriori. Calcule na posteriori a média, moda e a probabilidade de o parâmetro ultrapassar 2.
c) Implemente computacionalmente a solução: aproximação de 5 pontos e depois discretização com mais pontos (ao menos 50). Forneça os códigos para obter o gráfico da posteriori. Volte a calcular as quantidades pedidas no item (b).