1^a Lista de Exercícios - EST088 Estatística Bayesiana.
Prof. Vinícius D. Mayrink - Estatística/ICEx/UFMG

1. Uma empresa de consultoria econômica criou um método de previsão de recessão. O método tem a capacidade de prever com probabilidade 0.8 uma recessão que de fato irá ocorrer. O método também falha, prevendo com probabilidade 0.1 uma recessão que de fato não vai acontecer. Saiba que a probabilidade de ocorrer uma recessão é igual a 0.2. Em certo dia, o método foi usado e indicou que uma recessão irá acontecer. Qual é a probabilidade de que a recessão aconteça realmente?

2. Alice tem duas moedas no bolso. Uma delas é honesta (face 1 = cara e face 2 = coroa) e a outra moeda é viciada (face 1 = cara e face 2 = cara). Alice coloca a mão no bolso, pega uma das moedas (sem ver) e lança para cima. Saiba que o resultado do lançamento foi cara. Qual é a probabilidade de que Alice tenha lançado a moeda honesta?

3. Alice agora tem cinco moedas no bolso. Duas delas são do tipo A (honestas, cada uma tem probabilidade de cara = 0.5), duas são do tipo B (viciadas, cada uma tem probabilidade de cara = 0.6) e uma do tipo C (viciada, ela tem probabilidade de cara = 0.9). Alice coloca a mão no bolso, pega uma das moedas (sem ver) e lança para cima. O resultado do lançamento foi cara. Diante deste resultado, qual é a probabilidade de que Alice tenha lançado a moeda A? e a moeda B? e a moeda C?

4. Em um exame oral você tem que resolver exatamente um problema, o qual pode ser de três tipos: A, B ou C. A probabilidade da comissão avaliadora selecionar um problema tipo A, B e C são 0.30, 0.20 e 0.50, respectivamente. Durante a sua preparação para o exame, você conseguiu resolver 9 de 10 problemas tipo A, 2 de 10 problemas tipo B e 6 de 10 problemas tipo C.

   a - Qual é a probabilidade de que você resolva o problema aplicado no exame?

   b - Dado que você resolveu o problema proposto no exame, qual é a probabilidade dele ter sido do tipo A?

5. Considere duas urnas denominadas A e B. A urna A contém 75 bolas vermelhas e 25 bolas brancas. A urna B contém 55 bolas vermelhas e 45 bolas brancas. Um dado de 6 faces será lançado e se o resultado for face 1 ou 2, uma bola será retirada da urna A. Se o resultado do lançamento for 3, 4, 5 ou 6, então uma bola será retirada da urna B.

   a - Uma bola foi retirada de uma das duas urnas e a cor observada foi vermelho. Qual é a probabilidade de que esta bola tenha vindo da urna A?

   b - Uma bola foi retirada de uma das duas urnas e a cor observada foi vermelho. Qual é a probabilidade de que esta bola tenha vindo da urna B?

6. Tom foi ao médico e descobriu que em sua urina havia pequenos vestígios de sangue (visível apenas no microscópio). O médico explicou ao Tom que este sintoma ocorre em 10% de todas as pessoas e em 95% das pessoas com câncer nos rins. Mesmo não tendo nenhum outro sintoma, Tom ficou preocupado com a possibilidade de ter câncer nos rins. Saiba que o câncer de rins ocorre em 14 de cada 100 mil pessoas. Dado o diagnóstico de sangue na urina, qual é a probabilidade do Tom ter câncer nos rins?

7. Das pacientes de uma clínica de ginecologia com idade acima de 40 anos, 60% são ou foram casadas e 40% são solteiras. Sendo solteira, a probabilidade de ter tido um distúrbio hormonal no último ano é de 10%, enquanto que para as demais essa probabilidade aumenta para 30%. Uma paciente foi escolhida ao acaso e observou-se que ela tem distúrbio hormonal. Qual é a probabilidade dela ser solteira?

8. Em uma região, a probabilidade de chuva em um dia qualquer de primavera é de 0.1. Um meteorologista da TV acerta suas previsões em 80% dos dias em que chove e em 90% dos dias em que não chove.

   a - Qual é a probabilidade do meteorologista acerta uma previsão?

   b - Se houve acerto na previsão, qual é a probabilidade de ter sido um dia de chuva?

9. Três candidatos disputam as eleições para uma prefeitura. O candidato do partido A tem 30% da preferência eleitoral, o do partido B tem 30% e o do partido C tem 40%. Em sendo eleito, a probabilidade de dar efetivamente prioridade para a Educação e Saúde é de 0.4, 0.6 e 0.9 para os candidatos A, B e C, respectivamente.

   a - Qual é a probabilidade de não ser dada prioridade a essas áreas no próximo governo?

   b - Se a área teve prioridade, qual é a probabilidade do candidato A ter ganho a eleição?

10. Um médico desconfia que um paciente tem tumor no abdômen, pois isto ocorreu em 70% dos casos similares que tratou. Se o paciente de fato tiver o tumor, o exame de ultra-som consegue detectar com probabilidade 0.9. Entretanto, se ele não tiver o tumor, o exame ultra-som pode indicar sua existência com probabilidade 0.1. Se o exame detectou um tumor, qual é a probabilidade do paciente realmente ter a anomalia?

11. Um teste de HIV fornece resultado positivo com probabilidade 0.98, quando o paciente realmente está infectado pelo vírus. Por outro lado, o teste fornece resultado negativo com probabilidade 0.99, quando o paciente de fato não está infectado. Em uma cidade, temos 10% dos indivíduos infectados pelo HIV. Se uma pessoa é selecionada aleatoriamente nesta cidade e o teste indica positivo para o HIV, qual é a probabilidade desta pessoal realmente estar infectada?

12. Três fábricas A, B e C fornecem equipamentos de precisão para o laboratório de química de uma universidade. Apesar de serem aparelhos de precisão, existe uma pequena chance de subestimação ou superestimação das medidas efetuadas. Os percentuais de subestimação, estimação exata e superestimação dos equipamentos são como segue: 1%, 98% e 1% para a fábrica A, 0.5%, 98% e 1.5% para a fábrica B e 0%, 99% e 1% para a fábrica C. As fábricas fornecem, respectivamente, 20%, 30% e 50% dos aparelhos utilizados no laboratório. Um aparelho foi escolhido ao acaso e verificou-se que ele subestima as medidas. Qual é a probabilidade deste equipamento ter sido produzido pela fábrica A?

13. Uma família viaja ao litoral para passar as férias. A probabilidade de congestionamento na estrada é 0.6. Havendo congestionamento, a probabilidade dos dois filhos brigarem no carro é de 0.8. Sem

congestionamento, a briga ocorre com probabilidade 0.4. Quando há briga e não há congestionamento, a probabilidade do pai perder a paciência com os filhos é de 0.7. Quando há briga e existe congestionamento, a probabilidade do pai perder a paciência com os filhos é também 0.7. Quando os filhos não brigam e há congestionamento, a probabilidade de perda de paciência é 0.5. Se a viagem não tiver congestionamento e não tiver briga, o pai dirige sem perder a paciência com probabilidade 1. Determine a probabilidade de:

a - não ter havido congestionamento dado que o pai não perdeu a paciência com os filhos.

b - ter havido briga dado que o pai perdeu a paciência no trajeto.

14. Em um programa de TV o participante deve selecionar uma de três portas. Atrás de uma das portas existe um prêmio. Atrás das outras duas portas não há nada. Depois que o participante escolheu sua porta, o apresentador do programa abre uma das portas não escolhidas. Todos na plateia constatam que não havia prêmio na porta que foi aberta. O apresentador então pergunta ao participante se ele gostaria de mudar sua escolha para a outra porta fechada. Probabilisticamente, é vantajoso mudar de porta ou permanecer com a escolha inicial? Responda usando o Teorema de Bayes.

15. Em um aeroporto todos os passageiros são submetidos a uma verificação cuidadosa. Seja $T \in \{0, 1\}$ uma variável aleatória indicando se um indivíduo é terrorista ($T = 1$) ou não ($T = 0$). Considere também que $X \in \{0, 1\}$ é uma variável aleatória indicando prisão ($X = 1$ significa prisão, $X = 0$ indica sem prisão). De acordo com dados históricos, um terrorista é preso com probabilidade $P(X = 1|T = 1) = 0.98$. Um não terrorista é preso incorretamente com probabilidade $P(X = 1|T = 0) = 0.001$. Considere que 1 em cada 100 mil passageiros circulando pelo aeroporto é terrorista. Uma pessoa foi presa no aeroporto no dia em que você foi viajar, qual é a probabilidade dessa pessoa ser realmente um terrorista?

16. Uma empresa vende um sistema (câmera e software) de reconhecimento facial que controla a abertura de uma porta. Se uma pessoa deseja passar pela porta, o sistema irá tirar uma foto do rosto e comparar com imagens salvas na memória. O software calcula um valor X (escore) no intervalo de 0 a 1 que mede a semelhança entre as imagens. Após alguns testes envolvendo diversas pessoas, chegou-se em formulações para representar a distribuição condicional de X dado uma variável indicadora C. Admita que: $C = 1$ indica o comparativo "foto vs. imagem" de uma mesma pessoa e $C = 0$ indica o comparativo "foto vs. imagem" de pessoas diferentes. As distribuições são as seguintes:

$$f_{X|\alpha_1,\lambda_1,C=1}(x) = \alpha_1 \exp\{\lambda_1 x\} \quad \text{e} \quad f_{X|\alpha_2,\lambda_2,C=2}(x) = \alpha_2 \exp\{-\lambda_2 x\} \quad \text{para } \; 0 < x < 1.$$

O termo λ_j é o parâmetro do modelo probabilístico. O termo α_j é uma constante normalizadora que pode ser escrita em função de λ_j.

a - Escreva as constantes normalizadoras α_1 e α_2 em função dos parâmetros λ_1 e λ_2, respectivamente.

b - Calcule as constantes normalizadoras para $\lambda_1 = \ln(1000)$ e $\lambda_2 = \ln(0.001)$.

17. Nos casos listados abaixo, calcule $E(X)$, $Var(X)$, mediana de X e os valores $\{q_1, q_2\}$ tais que $P(X \leq q_1) = 0.025$ e $P(X \leq q_2) = 0.975$.

a - $X \sim \text{Beta}(3, 5)$.

b - $X \sim \text{Beta}(12, 4)$.

c - $X \sim \text{Gama}(12, 4)$.

d - $X \sim \text{Gama}(26, 5)$.

e - $X \sim \text{Normal}(120, 64)$.

f - $X \sim \text{Normal}(860, 576)$.

18. Uma cidade está considerando construir uma nova praça. Um jornal local decidiu investigar o nível de apoio da população para este projeto. Uma pesquisa foi conduzida entrevistando 120 moradores da cidade e, deste total de participantes, 74 são favoráveis à construção da praça (o restante é contra).

    a - Qual é a distribuição de probabilidade da variável Y = número de pessoas na entrevista que apoiam a criação da praça.

    b - Admita uma distribuição *a priori* uniforme para a proporção p de moradores que apoiam a praça. Qual é a distribuição *a posteriori* de p.

19. Sofia é editora de um jornal estudantil e ela pretende realizar uma pesquisa de opinião para determinar o nível de apoio que o atual presidente da associação de estudantes possui. Ela vai especificar uma distribuição *a priori* para p que é a proporção de estudantes apoiando o atual presidente. Sofia acredita que o valor de p está em torno de 0.5 e o desvio padrão, representando a incerteza *a priori*, vale 0.15.

    a - Estabeleça uma distribuição *a priori* Beta(a, b) que satisfaça a informação inicial de Sofia.

    b - Sofia entrevistou 68 estudantes e 21 deles declararam apoio ao presidente (o restante é contra). Determine a distribuição *a posteriori* que Sofia vai obter para p.

    c - Calcule a esperança, o desvio padrão e a mediana da distribuição *a posteriori* obtida em (b).

20. Selecionamos uma amostra aleatória de moradores de uma cidade para estimar a proporção p de pessoas que praticam alguma atividade física regular. Antes de analisar os dados, precisamos especificar uma distribuição *a priori* que expressa nosso conhecimento sobre p. Você acredita que a proporção p deve esta ao redor de 0.4 e sua incerteza sobre isso pode ser representada pelo desvio padrão 0.1.

    a - Estabeleça uma distribuição *a priori* Beta(a, b) que satisfaça a sua informação inicial.

    b - A pesquisa foi respondida por 100 moradores. Deste total, 21 disseram praticar atividade física regular. Determine a distribuição *a posteriori* de p.

    c - Calcule a esperança, o desvio padrão e a mediana da distribuição *a posteriori* obtida em (b).

21. Um empresa de seguros está medindo o número de reclamações que recebe por semana. Seja Y_i o número de reclamações na i-ésima semana. Adote $Y_i|\mu \sim \text{Poisson}(\mu)$ e assuma independência condicional dos Y_i's dado μ. O gerente da empresa escolheu a distribuição *a priori* Gama(0.1,0.1) para expressar a informação inicial dele sobre a taxa média μ de reclamações semanais. Responda:

    a - Calcule a esperança, a variância e a mediana da distribuição *a priori* sugerida pelo gerente.

    b - Em um período de 10 semanas, observou-se os seguintes valores de Y_i: 5, 8, 4, 6, 11, 6, 6, 5, 6 e 4. Determine a distribuição *a posteriori* de μ.

c - Calcule a esperança, a variância e a mediana da distribuição *a posteriori* obtida em (b).

22. Seja Y o número de sucessos em $n = 10$ ensaios Bernoulli independentes com probabilidade de sucesso igual a p em todos os ensaios. Suponha que apenas quatro valores são possíveis para p, eles são: 0.2, 0.4, 0.6 e 0.8. Na análise Bayesiana que será desenvolvida, não desejamos favorecer qualquer destes quatro valores, então adota-se uma distribuição *a priori* uniforme discreta para p. Admita que observou-se $Y = 7$. Preencha toda a tabela abaixo para obter a distribuição *a posteriori* de p.

p	priori	verossimilhança	priori $\times$ verossimilhança	posteriori
0.2				
0.4				
0.6				
0.8				
Marginal $P(Y = 7)$				