Midterm [20 pts]

March 16th, 2022 in class

PUBH 8442: Bayes Decision Theory and Data Analysis

Give your final answers in simplified, closed form wherever possible. However, partial credit will be awarded for incomplete solutions. Good luck!

1. *Regression hypotheses* [9 pts]
   Consider the simple linear regression model $\mathbf{y} = \mathbf{x}\beta + \epsilon$, where $\mathbf{y} : n \times 1$, $\mathbf{x} : n \times 1$ and $\epsilon : n \times 1$ with $\epsilon \sim N(\mathbf{0}, \sigma^2 I)$. Assume the residual variance σ^2 is fixed, and there are two models for β:

   - M_0 is the null model: $\beta = 0$.
   - M_1 is a model with a normal prior: $\beta \sim N(0, \sigma^2\tau^2)$ with τ^2 fixed.

   Assume each model has prior probability 1/2, $P(M_0) = P(M_1) = 1/2$.

   (a) (3 points) What is the marginal distribution for $\mathbf{y}$ under M_1, $p(\mathbf{y} \mid M_1)$?

   Under M_1 $\mathbf{y}$ is multivariate normal, with mean

   $$E(\mathbf{y}) = \mathbf{x}E(\beta) + E(\epsilon) = 0 + 0 = 0$$

   and variance matrix

   $$\text{Var}(\mathbf{y}) = \mathbf{x}\text{Var}(\beta)\mathbf{x}^T + \text{Var}(\epsilon) = \sigma^2\tau^2\mathbf{x}\mathbf{x}^T + \sigma^2\mathbf{I}$$

   . That is, $p(\mathbf{y} \mid M_1) = N(\mathbf{0}, \sigma^2(\tau^2\mathbf{x}\mathbf{x}^T + \mathbf{I}))$.

   (b) (3 points) What is the posterior probability of M_1, $p(M_1 \mid \mathbf{y})$?
   Using the given prior probabilities $P(M_0) = P(M_1) = 1/2$,

   $$p(M_1 \mid \mathbf{y}) = \frac{P(M_1)p(\mathbf{y} \mid M_1)}{P(M_0)p(\mathbf{y} \mid M_0) + P(M_1)P(\mathbf{y} \mid M_1)} = \frac{p(\mathbf{y} \mid M_1)}{p(\mathbf{y} \mid M_0) + p(\mathbf{y} \mid M_1)}$$

   where $p(\mathbf{y} \mid M_1)$ is given in part (a) and $p(\mathbf{y} \mid M_1) = N(\mathbf{0}, \sigma^2 I)$.

   (c) (3 points) What is the posterior expectation of β, marginalizing over M_0 and M_1, $E(\beta \mid \mathbf{y})$?

   Note that

   $$p(\beta \mid \mathbf{y}) = P(M_0 \mid \mathbf{y})p(\beta \mid M_0, \mathbf{y}) + P(M_1 \mid \mathbf{y})p(\beta \mid M_1, \mathbf{y}),$$

   which gives

   $$\begin{aligned} E(\beta \mid \mathbf{y}) &= P(M_0 \mid \mathbf{y})E(\beta \mid M_0, \mathbf{y}) + P(M_1 \mid \mathbf{y})E(\beta \mid M_1, \mathbf{y}) \\ &= P(M_1 \mid \mathbf{y})E(\beta \mid M_1, \mathbf{y}), \end{aligned}$$

   where $P(M_1 \mid \mathbf{y})$ is found in part (b), and $E(\beta \mid M_1, \mathbf{y}) = (\mathbf{x}^T\mathbf{x} + (\tau^2)^{-1})^{-1}\mathbf{x}^T\mathbf{y}$ (see slide 17 of Bayesian Linear Models notes).

2. *Telemedicine call-ins* [7 pts]
Let Y be the number of call-ins to a new telemedicine service within the first hour, which has a Poisson distribution with rate parameter λ: $Y \sim \text{Poisson}(\lambda)$. Assume λ has a Gamma$(1,1)$ prior distribution.

(a) (3 pts) What is the marginal probability that $Y = 0$, $P(Y = 0)$?

By the Poisson-Gamma marginal distribution (see lecture notes for "Model Comparison") with $y = 0$ and $\alpha = \beta = 1$:

$$P(Y = 0) = \frac{1^1\Gamma(0+1)}{\Gamma(1)\cdot 1 \cdot (1+1)^{0+1}} = \frac{1}{2}.$$

(b) (4 pts) Consider the wait time (in hrs) until the next call-in after the first hour, X, where $X \sim \text{Exponential}(\lambda)$. What is the conditional density of X, given Y, $p(X \mid Y)$? (Assume X and Y are independent given λ.)

By the Poisson-Gamma model, $p(\lambda \mid Y) \sim \text{Gamma}(Y+1, 2)$. So,

$$\begin{aligned}
p(X \mid Y) &= \int_0^\infty p(X \mid \lambda)p(\lambda \mid Y)d\lambda \\
&= \int \lambda e^{-\lambda X}\frac{2^{(Y+1)}}{\Gamma(Y+1)}\lambda^Y e^{-2\lambda}d\lambda \\
&= \frac{2^{(Y+1)}}{\Gamma(1+Y)}\int \lambda^{(Y+1)}e^{-(X+2)\lambda} \\
&= \frac{2^{(Y+1)}}{\Gamma(1+Y)}\frac{\Gamma(Y+2)}{(X+2)^{(Y+2)}} \\
&= \frac{(Y+1)2^{(Y+1)}}{(X+2)^{(Y+2)}}
\end{aligned}$$

for $X \geq 0$.

3. *Bayes estimator* [4 pts]
Consider a Bayesian model with likelihood $p(\mathbf{y} \mid \theta)$ and prior $p(\theta)$ defined for $\theta > 0$. We would like to estimate θ using the loss function

$$l(\theta, d(\mathbf{y})) = \frac{(\theta - d(\mathbf{y}))^2}{d(\mathbf{y})} \quad \text{for } d(\mathbf{y}) > 0.$$

What is the Bayes decision rule, $d(\mathbf{y}) = \hat{\theta}_{\mathbf{y}}$, under this loss function? Write your answer in terms of the posterior mean $\text{E}(\theta \mid \mathbf{y})$ and variance $\text{V}(\theta \mid \mathbf{y})$.

Write $E = E_{\theta|\mathbf{y}}$, $V = \text{Var}_{\theta|\mathbf{y}}$, and $d = d(\mathbf{y})$ for short. The posterior risk is

$$\begin{aligned}
E\frac{(\theta - d)^2}{d} &= E\left(\frac{\theta^2 - 2\theta d + d^2}{d}\right) \\
&= \frac{1}{d}E(\theta^2) - 2E(\theta) + d.
\end{aligned}$$

Which is minimized where the derivative is 0:

$$\frac{d}{dd}E\frac{(\theta-d)^2}{d} = \frac{-E(\theta^2)}{d^2} + 1 = 0$$
$$\rightarrow d = \sqrt{E(\theta^2)} = \sqrt{V(\theta) + E(\theta)^2}$$

Reference: Density and probability mass functions

Normal$(\boldsymbol{\mu}, \boldsymbol{\Sigma}) : p(\mathbf{x}) = \frac{e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})}}{\det(2\pi\Sigma)^{1/2}}$ Exponential$(\lambda) : p(x) = \lambda e^{-\lambda x}$

Gamma$(\alpha, \beta) : p(x) = \frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}$ Poisson$(\lambda) : P(x) = \frac{\lambda^x e^{-\lambda}}{x!}$