Homework 2 Part 1

This is an individual assignment.

Solve all problems by hand. You may type your answers in markdown cells or `push` a PDF file with your handwritten answers.

Problem 1 (15 points)

Consider a training set containing positive natural numbers - including zero - ($x \in \mathbb{N}_0$) for 2 classes, C_0 and C_1. The training set has 100 samples for class C_0 and 50 for C_1.

Suppose that you have reason to believe that samples belonging from C_0 are drawn from a Poisson random variable with parameter $\lambda > 0$, and samples belonging to C_1 are drawn from a Binomial random variable with parameters $n \in \mathbb{N}_0$ and $p \in [0,1]$. In other words:

$$p(x|C_0) = \frac{\lambda^x e^{-\lambda}}{x!}$$
$$p(x|C_1) = \binom{n}{x} p^x (1-p)^{n-x}$$

where $\binom{n}{x} = \frac{n!}{(n-x)!}$ is the binomial coefficient of "n choose x".

Consider $\lambda = 2$, $n = 10$ and $p = 0.5$. Consider the test point $x = 3$. Use the Bayesian classifier to assign $x = 3$ to C_0 or C_1? Show your work.

Q1. $P(C_0) = \frac{\cancel{100}}{\cancel{150}} = \frac{2}{3}$, $P(C_1) = \frac{50}{150} = \frac{1}{3}$

$P(x|C_0) = \frac{\lambda^x e^{-\lambda}}{x!}$, $P(x|C_1) = \binom{n}{x} p^x (1-p)^{n-x}$

$\lambda = 2$

$x = 3$

$n = 10$, $p = 0.5$

$x = 3$

$\therefore P(3|C_0) = \frac{2^3 e^{-2}}{3!}$

$\therefore P(3|C_1) = \binom{10}{3}(0.5)^3(0.5)^7$

$= \frac{10(9)(8)\cancel{(7!)}}{\cancel{(7!)}\,3(2)}(0.5)^{10}$

$\therefore \frac{P(C_0|x)}{P(C_1|x)} = \frac{P(x|C_0)\,P(C_0)}{P(x|C_1)\,P(C_1)} = \frac{\frac{2^3 e^{-2}}{3!}\cdot\left(\frac{2}{3}\right)}{120\,(0.5)^{10}\left(\frac{1}{3}\right)} = \frac{0.361}{0.1172} = \boxed{3.08}$

$\therefore \frac{P(C_0|3)}{P(C_1|3)} > 1 \iff P(C_0|3) > P(C_1|3)$

$\therefore$ Assign $x=3$ to class C_0.

Problem 2 (30 points)

Suppose you have a training set with N data points $\{x_i\}_{i=1}^N$, where $x_i \in \mathbb{Z}_0^+$ (set of nonnegative integers - $0, 1, 2, 3, \ldots$). Assume the samples are independent and identically distributed (i.i.d.), and each sample is drawn from a Geometric random variable with probability mass function:

$$P(x|\rho) = \rho(1-\rho)^x$$

Moreover, consider the Beta density function as the prior probability on the success probability, ρ,

$$P(\rho|\alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \rho^{\alpha-1}(1-\rho)^{\beta-1}$$

Answer the following questions:

1. (8 points) **Derive the maximum likelihood estimate (MLE) for the rate parameter ρ. Show your work.**
2. (10 points) **Derive the maximum a posteriori (MAP) estimate for the rate parameter ρ. show your work.**
3. (6 points) **Is the Beta distribution a conjugate prior for the success probability, ρ, of the Geometric distribution? Why or why not?**
4. (6 points) **Suppose you would like to update the Beta prior distribution and the MAP point estimation in an online fashion, as you obtain more data. Write the pseudo-code for the online update of the prior parameters. In your answer, specify the new values for the parameters of the prior.**

Q2 $P(x|\rho) = \rho(1-\rho)^x$

$$P(\rho|\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\,\Gamma(\beta)} \rho^{\alpha-1} (1-\rho)^{\beta-1}$$

(1) → MLE:

$\cancel{\mathcal{L} = \sum_{i=1}^{N} P(x|\rho)}$ $\mathcal{L}^0 = \prod_{i=1}^{N} P(x_i|\rho)$

$$= \prod_{i=1}^{N} \rho(1-\rho)^{x_i} = \rho^N (1-\rho)^{\sum_{i=1}^{N} x_i}$$

$$\mathcal{L} = \ln(\mathcal{L}^0) = \ln\left(\prod_{i=1}^{N} \rho(1-\rho)^{x_i}\right)$$

$$\cancel{\sum_{i=1}^{N} \ln\left[\rho(1-\rho)^{x_i}\right]} = \ln\left(\rho^N (1-\rho)^{\sum_{i=1}^{N} x_i}\right)$$

$$= \ln(\rho^N) + \ln(1-\rho)^{\sum_{i=1}^{N} x_i}$$

$$= N\ln(\rho) + \ln(1-\rho)\left(\sum_{i=1}^{N} x_i\right)$$

$$\therefore \frac{\partial \mathcal{L}}{\partial \rho} = 0 \Leftrightarrow N\left(\frac{1}{\rho}\right) + \left(\frac{-1}{1-\rho}\right)\left(\sum_{i=1}^{N} x_i\right) = 0$$

$$\therefore N - N\rho - \rho \sum_{i=1}^{N} x_i = 0 \Leftrightarrow \boxed{\rho_{MLE} = \frac{N}{N + \sum_{i=1}^{N} x_i}}$$

(2) MAP :

$$\mathcal{L}^{0} = \left[\prod_{i=1}^{N} P(x_i|\rho) \right] P(\rho|\alpha,\beta)$$

$$= \left[\prod_{i=1}^{N} \rho(1-\rho)^{x_i} \right] \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} (\rho^{\alpha-1}) (1-\rho)^{\beta-1}$$

$$= \left[\rho^{N} (1-\rho)^{\sum_{i=1}^{N} x_i} \right] \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} (\rho^{\alpha-1}) (1-\rho)^{\beta-1}$$

$$= (\rho^{N+\alpha-1}) (1-\rho)^{\sum_{i=1}^{N} x_i+\beta-1} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}$$

$$\mathcal{L}^{0} \propto (\rho^{N+\alpha-1}) (1-\rho)^{\sum_{i=1}^{N} x_i+\beta-1}$$

$$\mathcal{L} = \ln(\mathcal{L}^{0}) = (N+\alpha-1)\ln(\rho) + \left(\sum_{i=1}^{N} x_i+\beta-1\right)\ln(1-\rho)$$

$$\frac{\partial \mathcal{L}}{\partial \rho} = 0 \Leftrightarrow (N+\alpha-1)\left(\frac{1}{\rho}\right) + \left(\sum_{i=1}^{N} x_i+\beta-1\right)\left(\frac{-1}{1-\rho}\right) = 0$$

$$\therefore (N+\alpha-1)(1-\rho) + \left(\sum_{i=1}^{N} x_i+\beta-1\right)(-\rho) = 0$$

$$\therefore N+\alpha-1 = \left(\sum_{i=1}^{N} x_i + \beta - 1 + N + \alpha - 1\right)\rho$$

$$\therefore \rho_{MAP} = \left(\frac{N+\alpha-1}{\sum_{i=1}^{N} x_i + \alpha + \beta - 2 + N}\right)$$

(3) Yes, it is a conjugate prior because it brings the posterior in the same parameters as prior.

$$\alpha \leftarrow N + \alpha$$

$$\beta \leftarrow \sum_{i=1}^{N} x_i + \beta$$

(4) pseudo-code for online update:

(1) start iteration $t=0$. Initialize the prior parameters $\alpha^{(t)}$ and $\beta^{(t)}$

(2) Compute the parameter estimation for the current prior probability:

$$\rho_{MAP} = \frac{N+\alpha-1}{\sum_{i=1}^{N} x_i + N + \alpha + \beta - 2}$$

(3) Update the prior parameters.

$$\alpha \leftarrow N + \alpha$$
$$\beta \leftarrow \sum_{i=1}^{N} x_i + \beta$$

(4) Increment iteration counter

$$t \leftarrow t + 1$$

On-Time (5 points)

Submit your assignment before the deadline.

Submit Your Solution

Confirm that you've successfully completed the assignment.

Along with the Notebook, include a PDF of the notebook with your solutions.

`add` and `commit` the final version of your work, and `push` your code to your GitHub repository.

Submit the URL of your GitHub Repository as your assignment submission on Canvas.
