AI 501 Machine Learning for AI, Spring 2021

Homework assignment 2

1. (15 points) Let $\mathbb{P}$ be a probability measure. Using the definition of probability measures, show that

   1. For any events $A \subset B$, $\mathbb{P}(A) \leq \mathbb{P}(B)$.
   2. For any countable set $\{A_i\}_{i=1}^n$, $\mathbb{P}(\bigcup_{i=1}^n A_i) \leq \sum_{i=1}^n \mathbb{P}(A_i)$.

2. (15 points) You have written a computer program that compiles with probability π and does not with probability $1 - \pi$ (code does not change). You decide to model the number of successful compiles x in a sequence of m independent experiments using a Binomial distribution:

$$p(x|m, \pi) = \binom{m}{x} \pi^x (1 - \pi)^{m-x}, \ x \in \{0, 1, \ldots, m\}. \tag{1}$$

   Choose a conjugate prior for the parameter π of the Binomial likelihood and compute the posterior distribution $p(\pi|x, m)$.

3. (20 points) There are two bags. The first bag contains four mangos and two apples; the second bag contains four mangos and four apples. We also have a biased coin, which shows "heads" with probability 0.7 and "tails" with probability 0.3. If the coin shows "heads", we pick a fruit at random from bag 1; otherwise we pick a fruit at random from bag 2.

   (a) (5 points) Your friend flips the coin (you cannot see the result) and picks a fruit at random from the corresponding bag. What is the probability that the fruit is a mango.

   (b) (5 points) If your friend presented you a mango, what is the probability that the mango was picked from bag 2?

   (c) (10 points) This time, your friend picks 3 fruits one by one. The picked fruit is put back into its corresponding bag before picking the next fruit. The first fruit is picked in the same way as in (a) and (b). For the second fruit, your friend flips the coin. If the coin lands "heads", a fruit is picked at random from the same bag that was previously picked from; otherwise a fruit is picked at random from the other bag. Your friend repeats this process in the third choice. If your friend picked three apples, what is the probability that those apples were only picked from bag 1?

4. (16 points) Consider a two dimensional multivariate normal distribution,

$$p(x, y) = \mathcal{N}\left(\begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}, \begin{bmatrix} \sigma_{xx}^2 & \sigma_{xy}^2 \\ \sigma_{yx}^2 & \sigma_{yy}^2 \end{bmatrix} \right). \tag{2}$$

   Furthermore, we have

$$\begin{bmatrix} \mathrm{z} \\ \mathrm{w} \end{bmatrix} = \boldsymbol{A} \begin{bmatrix} \mathrm{x} \\ \mathrm{y} \end{bmatrix} + \mathbf{b} \tag{3}$$

   where $\boldsymbol{A} \in \mathbb{R}^{2\times 2}$ and $\mathbf{b} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{Q})$ is independent Gaussian noise. "Independent" implies that (x, y) and $\mathbf{b}$ are independent random variables and that $\boldsymbol{Q}$ is a diagonal matrix.

(a) (8 points) Compute the likelihood $p(z, w|x, y)$.

(b) (8 points) Compute the marginal likelihood $p(z, w)$.

5. (14 points) Compute *entropy*,

$$h[\mathbf{x}] \stackrel{\text{def}}{=} -\int p(\boldsymbol{x}) \log p(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}, \tag{4}$$

of an n-dimensional multivariate normal distribution with Probability Density Function (PDF),

$$p(\boldsymbol{x}) = \exp\bigg(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu}) - \frac{1}{2}\log|\boldsymbol{\Sigma}| - \frac{n}{2}\log 2\pi \bigg). \tag{5}$$

Hint: $(\boldsymbol{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu}) = \mathrm{tr}(\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})(\boldsymbol{x}-\boldsymbol{\mu})^T)$ and $\mathbb{E}[\mathrm{tr}(A)] = \mathrm{tr}(\mathbb{E}[A])$.

6. (20 points) Let $\mathrm{x}_i \sim \mathrm{Gamma}(\alpha_i, 1)$ for $i = 1, \dots, n$ with PDF

$$p(x) = \frac{x^{\alpha_i - 1} e^{-x}}{\Gamma(\alpha_i)}, \tag{6}$$

where

$$\Gamma(z) := \int_0^\infty t^{z-1} e^{-t} \mathrm{d}t, \tag{7}$$

denotes the gamma function. Show that

$$\mathbf{y} := \left(\frac{\mathrm{x}_1}{\sum_{i=1}^n \mathrm{x}_i}, \dots, \frac{\mathrm{x}_n}{\sum_{i=1}^n \mathrm{x}_i} \right), \tag{8}$$

is a Dirichlet random variable with PDF

$$f(\boldsymbol{y}) = \frac{\Gamma(\sum_{i=1}^n \alpha_i)}{\prod_{i=1}^n \Gamma(\alpha_i)} \prod_{i=1}^n y_i^{\alpha_i - 1}. \tag{9}$$

Hint: use the change of variable

$$(\mathrm{x}_1, \dots, \mathrm{x}_n) \to (\mathrm{y}_1, \dots, \mathrm{y}_{n-1}, \mathrm{z}), \tag{10}$$

where

$$\mathrm{y}_i = \frac{\mathrm{x}_i}{\mathrm{z}} \text{ for } i = 1, \dots, n-1, \quad \mathrm{z} = \sum_{i=1}^n \mathrm{x}_i. \tag{11}$$