classmate

Date

Page

CS660 Homework 5

① Exercise 6.1

Consider the following bivariate distribution (x, y) of two discrete random variables X and Y.

		x_1	x_2	x_3	x_4	x_5
Y	y_1	0.01	0.02	0.03	0.1	0.1
	y_2	0.05	0.1	0.05	0.07	0.2
	y_3	0.1	0.05	0.03	0.05	0.04

X

Compute:

(a) The marginal distributions $p(x)$ and $p(y)$.

→ The marginal probability of $X = x_i$, $p(x_i)$ is the sum of the joint probabilities over all values of Y for each x_i:

$$p(x_i) = \sum_y p(x_i, y)$$

Similarly, the marginal probability of $Y = y_i$, $p(y_i)$

$$p(y_i) = \sum_x p(x, y_j)$$

Calculating $p(x)$.

$p(x_1) = 0.01 + 0.05 + 0.1 = 0.16$

$p(x_2) = 0.02 + 0.1 + 0.05 = 0.17$

$p(x_3) = 0.03 + 0.05 + 0.03 = 0.11$

$p(x_4) = 0.1 + 0.07 + 0.05 = 0.22$

$p(x_5) = 0.1 + 0.2 + 0.04 = 0.34$.

∴ the marginal distribution $p(x)$ is:

$$p(x) = \{0.16, 0.17, 0.11, 0.22, 0.34\}.$$

Calculating $p(y)$

$p(y_1) = 0.01 + 0.02 + 0.03 + 0.1 + 0.1 = 0.26$

$p(y_2) = 0.05 + 0.01 + 0.05 + 0.07 + 0.2 = 0.47$

$p(y_3) = 0.1 + 0.05 + 0.03 + 0.05 + 0.04 = 0.27$

$\therefore$ the marginal distribution $p(y)$ is:

$p(y) = \{0.26, 0.47, 0.27\}$.

(b) The conditional distributions $p(x \mid Y = y_1)$ and $p(y \mid X = x_3)$.

→ The conditional probability $p(x \mid Y = y_1)$ is given by:

$$p(x \mid Y = y_1) = \frac{p(x, y_1)}{p(y_1)}$$

As $p(y_1) = 0.26$; we calculate values as follows:

$$p(x_1 \mid Y = y_1) = \frac{0.01}{0.26} \approx 0.0385$$

$$p(x_2 \mid Y = y_1) = \frac{0.02}{0.26} \approx 0.0769$$

$$p(x_3 \mid Y = y_1) = \frac{0.03}{0.26} \approx 0.1154$$

$$p(x_4 \mid Y = y_1) = \frac{0.1}{0.26} \approx 0.3846$$

$$p(x_5 \mid Y = y_1) = \frac{0.1}{0.26} \approx 0.3846$$

$\therefore$ the Conditional probability $p(x \mid Y = y_1)$ is:

$p(x \mid Y = y_1) \approx \{0.0385, 0.0769, 0.1154, 0.3846, 0.3846\}$

The conditional probability $p(y|X=x_3)$ is:

$$p(y|X=x_3) = \frac{p(x_3, y)}{p(x_3)}$$

As $p(x_3) = 0.11$, we calculate values as follows:

$$p(y_1|X=x_3) = \frac{0.03}{0.11} \approx 0.2727$$

$$p(y_2|X=x_3) = \frac{0.05}{0.11} \approx 0.4545$$

$$p(y_3|X=x_3) = \frac{0.03}{0.11} \approx 0.2727$$

$\therefore$ the conditional distribution $p(y|X=x_3)$ is:

$$p(y|X=x_3) \approx \{0.2727, 0.4545, 0.2727\}$$

② You have written a computer program that sometimes compiles and sometimes not (code does not change). You decide to model the apparent stochasticity (success vs no success) x of the compiler using a Bernoulli distribution with parameter μ:

$$p(x|\mu) = \mu^x (1-\mu)^{1-x}, \quad x \in \{0, 1\}.$$

Choose a conjugate prior for the Bernoulli likelihood and compute the posterior distribution $p(\mu | x_1, \ldots, x_N)$

→ ① The Model and likelihood:

$$p(x|\mu) = \mu^x (1-\mu)^{1-x}, \quad x \in \{0, 1\}$$

$x=1$ indicates success,

$x=0$ indicates failure, μ is probability of success.

(ii) Choosing a Conjugate Prior.
The conjugate prior for the Bernoulli likelihood is the Beta distribution. The Beta distribution is parameterized by two shape parameters α and β:

$$p(\mu) = \text{Beta}(\mu|\alpha,\beta) = \frac{\mu^{\alpha-1}(1-\mu)^{\beta-1}}{B(\alpha,\beta)}$$

$B(\alpha,\beta)$ is Beta function which normalizes the distribution.

(iii) Likelihood of Observing the Data.
Suppose we observe a set of N independent compilation results $x_1, x_2, x_3, \ldots x_N$, where each x_i is either 0 or 1.

$$p(x_1, \ldots\ldots x_N|\mu) = \prod_{i=1}^{N} p(x_i|\mu) =$$

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

let $S = \sum_{i=1}^{N} x_i$ be the total number of successful compilations. Then the likelihood simplifies to:

$$p(x_1, \ldots, x_N|\mu) = \mu^S(1-\mu)^{N-S}.$$

(iv) Computing the Posterior Distribution.
Applying Baye's theorem,

$$p(\mu|x_1, \ldots, x_N) \propto p(x_1, \ldots, x_N|\mu) \cdot p(\mu)$$

Using our prior $p(\mu) = \text{Beta}(\mu|\alpha,\beta) = \frac{\mu^{\alpha-1}(1-\mu)^{\beta-1}}{B(\alpha,\beta)}$;

$$p(\mu|x_1, \ldots\ldots, x_N) \propto \mu^S(1-\mu)^{N-S} \cdot \mu^{\alpha-1}(1-\mu)^{\beta-1}$$

Combining the power of μ and $1-\mu$, we have:

$$p(\mu|x_1,\ldots,x_N) \propto \mu^{S+\alpha-1}(1-\mu)^{N-S+\beta-1}$$

$\therefore$ The posterior distribution for μ is:

$$p(\mu|x_1,\ldots,x_N) = \text{Beta}(\mu|\alpha+S, \beta+N-S).$$

(3) Excercise 6.4.

→ Event B_1: The fruit was picked from Bag 1.

Event B_2: The fruit was picked from Bag 2.

Event M: The fruit picked was a mango.

To find - probability that the mango was picked from Bag 2; $P(B_2|M)$.

Baye's Theorem states:

$$P(B_2|M) = \frac{P(M|B_2)\cdot P(B_2)}{P(M)}.$$

(i) Finding $P(B_1)$ and $P(B_2)$.

with probability 0.6 the coin lands heads and picked from bag 1.

$$P(B_1) = 0.6.$$

with probability 0.4 the coin lands ~~heads~~ tails and picked from Bag 2.

$$P(B_2) = 0.4.$$

(ii) Finding $P(M|B_1)$ and $P(M|B_2)$.

Bag 1 contains 4 mangoes and 2 apples;

$$P(M|B_1) = \frac{\text{Number of mangoes in Bag 1}}{\text{Total fruits in Bag 1.}}$$

$$= \frac{\cancel{4}\,2}{\cancel{6}\,3} = \frac{2}{3}.$$

$$P(M|B_2) = \frac{\text{No. of mangoes in Bag 2}}{\text{Total fruits in Bag 2}}$$

$$= \frac{\cancel{4}^{1}}{\cancel{8}_{2}} = \frac{1}{2}.$$

(iii) $P(M) = P(M|B_1)\cdot P(B_1) + P(M|B_2)\cdot P(B_2)$

$$\therefore P(M) = \frac{2}{\cancel{3}_{1}} \times \cancel{0.6}^{\,0.2} + \frac{1}{\cancel{2}_{1}} \times \cancel{0.4}^{\,0.2}$$

$$P(M) = 0.4 + 0.2$$

$$P(M) = 0.6.$$

Applying Baye's Theorem;

$$P(B_2|M) = \frac{P(M|B_2)\cdot P(B_2)}{P(M)}$$

$$= \frac{\frac{1}{\cancel{2}_{1}} \times \cancel{0.4}^{\,0.2}}{0.6}$$

$$= \frac{0.2}{0.6} = \frac{1}{3}$$

$\therefore$ The probability that the mango was picked from Bag 2 is $\frac{1}{3}$.

(4) Exercise 6.11

→ The expectation of x with respect to its joint distribution $p(x, y)$ is:

$$E[x] = \sum_x \sum_y x p(x, y);$$

if x and y are discrete, or as:

$$E[x] = \int_x \int_y x p(x, y)\, dy\, dx;$$

if x and y are continuous.

We can express the joint distribution of (x, y) as:

$$p(x, y) = p(x|y) p(y)$$

$$\therefore E[x] = \sum_x \sum_y x p(x|y) p(y)$$

Or, in the continuous case:

$$E[x] = \int_x \int_y x p(x|y)\, p(y)\, dy\, dx$$

Since we are summing/integrating over both x and y, we can change the order of summation/integration:

$$E[x] = \sum_y p(y) \sum_x x p(x|y)$$

or, in the continuous case:

$$E[x] = \int_y p(y) \int_x x p(x|y)\, dx\, dy.$$

We can see that the inner sum/integral $\sum_x x p(x|y)$ or $\int_x x p(x|y) dx$ is the definition of the conditional expectation of x given y, denoted $E[x|y]$.

$$E[x|y] = \sum_x p(x|y)$$

or,

$$E(x|y) = \int_x x\,p(x|y)\,dx$$

$$\therefore \quad E[x] = \sum_y p(y)\,E(x|y]$$

or,

$$E[x] = \int_y p(y)\,E[x|y]\,dy$$

The expression $\sum_y p(y) E[x|y]$ or $\int_y p(y) E[x|y] dy$ is simply the expectation of $E[x|y]$ with respect to y.

$$\therefore E[x] = E_y[E[x|y]]$$

(5) Exercise 6.13.

Given:

Let X be a continuous random variable with cumulative distribution function $F_X(x)$.

$$Y := F_X(x).$$

We want to prove that Y is uniformly distributed on $[0, 1]$. That is, we want to show that $Y \sim \text{Uniform}(0, 1)$.

(1) Find the CDF of Y:

To show that Y is uniformly distributed, we need to find its CDF, $F_Y(y) = P(Y \leq y)$.

By definition of Y, we have:-

$$F_Y(y) = P(Y \leq y) = P(F_X(X) \leq y).$$

(ii) Rewrite in terms of X:
Since $F_X(x)$ is a continuous and strictly increasing function (as X has strictly monotonic CDF), we can take the inverse of F_X on both sides of the inequality $F_X(X) \leq y$ to rewrite this probability in terms of X:

$$P(F_X(X) \leq y = P(X \leq F_X^{-1}(y)).$$

(iii) By the definition of the CDF $F_X(x)$, we know:

$$P(X \leq F_X^{-1}(y)) = F_X(F_X^{-1}(y))$$

Therefore,

$$F_Y(y) = P(Y \leq y) = y$$

The CDF of Y is $F_Y(y) = y$ for $y \in [0,1]$. This is exactly the CDF of a uniform random variable on $[0,1]$, since for $Y \sim$ Uniform $(0,1)$, we have $F_Y(y) = y$ for $y \in [0,1]$

$\therefore$ CDF of $Y := F_X(X)$ is $F_Y(y) = y$ for $y \in [0,1]$.
$\therefore Y := F_X(X)$ is uniformly distributed on $[0,1]$.

(6) Let A be the event that a customer invests in tax-free bonds.
Let B be the event that a customer invests in mutual funds.

Given:

$P(A) = 0.6$ (tax-free bonds)

$P(B) = 0.3$ (mutual funds)

$P(A \cap B) = 0.15$ (the probability that the customer invests in both tax-free bonds and mutual funds).

(a) Probability that a customer will invest in either tax-free bonds or mutual funds:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
$$= 0.6 + 0.3 + 0.15$$
$$= 0.9 - 0.15$$
$$P(A \cup B) = 0.75$$

(b) Probability that a customer will invest in neither tax-free bonds nor mutual funds:

$$P(\text{neither } A \text{ nor } B) = 1 - P(A \cup B)$$
$$= 1.0.75$$

$$\therefore P(\text{neither } A \text{ nor } B) = 0.25$$

(7) Let G represent the event that the suspect is guilty.
Let I represent the event that the suspect is innocent,
Let T represent the event that the truth serum test indicates the suspect is guilty.

Given:

$P(G) = 0.05$ (guilty)

$P(I) = 1 - P(G) = 0.95$ (95% of suspects are innocent)

The test (truth serum) has following properties

- For guilty suspects: The test correctly identifies them as guilty 90% of the time

$$P(T|G) = 0.9$$

This implies

$$P(\text{Not } T | G) = 0.1.$$

- For innocent suspects: The test incorrectly identifies them as guilty 1% of the time

$$P(T|I) = 0.01$$

This implies

$$P(\text{Not } T | I) = 0.99.$$

We want to find the probability that the suspect is innocent given that the test indicates they are guilty ($P(I|T$

Apply Baye's Theorem,

$$P(I|T) = \frac{P(T|I) \times P(I)}{P(I)}.$$

We can find $P(T)$ by considering both ways in which the test could indicate guilt

$$P(T) = P(T|G) \cdot P(G) + P(T|I) \cdot P(I)$$

classmate

Date
Page

(i) Calculating $P(T)$:

$$P(T) = (0.9) \cdot (0.05) + (0.01) \cdot (0.95)$$
$$= 0.045 + 0.0095$$
$$= 0.0545.$$

(ii) Calculating $P(I|T)$:

$$P(I|T) = \frac{P(T|I) \cdot P(I)}{P(T)}.$$

$$= \frac{(0.01) \cdot (0.95)}{0.0545}$$

$$= \frac{0.0095}{0.0545}$$

$$\therefore P(I|T) \approx 0.1743$$

So, there is about a 17.43% chance that the suspect is innocent even though the truth serum indicates they are guilty

(8) Given:

$$f(x) = 2(1-x), \quad 0 < x \leq 1$$
$$\text{and } f(x) = 0, \quad \text{elsewhere.}$$

The expected value $E[X]$ of X is defined by:

$$E[X] = \int_{-\infty}^{\infty} x f(x)\,dx$$

Since $f(x) = 0$ outside the interval $(0, 1)$, this simplifies to:

$$E[X] = \int_0^1 x \cdot 2(1-x)\,dx$$

Substitute $f(x) = 2(1-x)$ into the formula:

$$E[x] = \int_0^1 x \cdot 2(1-x)\, dx$$

$$= 2\int_0^1 x(1-x)\, dx$$

$$E[x] = 2\int_0^1 (x - x^2)\, dx$$

$$= 2\left[\int_0^1 x\, dx - \int_0^1 x^2\, dx\right.$$

$$= 2\left[\left(\frac{x^2}{2}\right)_0^1 - \left(\frac{x^3}{3}\right)_0^1\right]$$

$$= 2\left(\frac{1}{2} - \frac{1}{3}\right)$$

$$= \cancel{2}\left(\frac{1}{\cancel{6}_3}\right)$$

$$\therefore E[X] = \frac{1}{3}$$

Since each unit represents \$5000, the average profit in dollars is:

$$\frac{1}{3} \times 5000$$

$$= \frac{5000}{3} \approx 1666.67.$$

$\therefore$ The average profit per automobile is approximately 1666.67 dollars.

classmate

Date
Page

(9) Given:

$\alpha = 3$ and $\beta = 2$.

$X \sim Beta(3,2)$

we want to find the probability that at least 80% of the new models will require service during the first year of operation. Mathematically, we need to calculate:

$$P(X \geq 0.8).$$

(i) The PDF Beta (α, β) is given by:

$$f(x; \alpha, \beta) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}, \quad 0 \leq x \leq 1$$

As $\alpha = 3$ and $\beta = 2$.

$$\therefore f(x; 3, 2) = \frac{x^2 (1-x)^1}{B(3,2)}, \quad 0 \leq x \leq 1.$$

(ii) Calculating $P(X \leq 0.8)$ using Beta CDF.

using beta distribution table, the $P(X \leq 0.8) \approx 0.942$

$$P(X \geq 0.8) = 1 - P(X \leq 0.8)$$

$$P(X \geq 0.8) = 1 - 0.942$$
$$= 0.058$$

This means there is about a 5.8% chance that at least 80% of these television sets will require service in their first year.

(10) Given:

- Service calls arrive at a maintenance center according to a Poisson process with rate λ (calls per minute).
- We have a data set of 20 one-minute periods, yielding an average of 1.8 calls per minute.
- The prior distribution for λ is an exponential distribution with mean 2.

Let:

- Y be the total number of calls observed over the 20 one minutes periods.
- $Y \sim$ Poisson (20λ).

(i) The probability mass function of Y for a Poisson random variable is:

$$P(Y|\lambda) = \frac{(20\lambda)^{Y} e^{-20\lambda}}{Y!}$$

The total number of calls observed over 20 minute is:

$$Y = 20 \times 1.8 = 36$$

we have $Y = 36$ as our observed value.

(ii) An exponential distribution with mean 2 has a rate parameter $\theta = \frac{1}{2}$, so:

$$p(\lambda) = \theta e^{-\theta\lambda} = \frac{1}{2} e^{-\lambda/2}$$

Substituting likelihood and prior expressions:

$$p(\lambda|Y) \propto \frac{(20\lambda)^{36} e^{-20\lambda}}{36!} \times \frac{1}{2} e^{-\lambda/2}$$

Ignoring Constant terms; we get.

$$P(\lambda | y) \propto \lambda^{36} e^{-(20+\frac{1}{2}\lambda)} = \lambda^{36} e^{-20.5\lambda}$$

(iii) The posterior distribution $p(\lambda|y)$ has the form:

$$p(\lambda | y) \propto \lambda^{36} e^{-20.5\lambda}$$

The Gamma distribution with shape parameter α and rate parameter β has the form:

$$\text{Gamma}(\lambda | \alpha, \beta) \propto \lambda^{\alpha-1} e^{-\beta\lambda}$$

Shape parameter $\alpha = 36+1 = 37$

Rate parameter $\beta = 20.5$.

$\therefore$ The posterior distribution of λ is:

$$\lambda | Y \sim \text{Gamma}(37, 20.5)$$