Recitation 4 Solution

1. (Coin Toss) In a large collection of coins, the probability X that a head will be obtained when a coin is tossed varies from one coin to another, and the distribution of X in the collection is specified by the following pdf:

$$\begin{cases} f(x) = 6x(1-x) & \text{for } 0 < x < 1, \\ 0 & \text{otherwise.} \end{cases}$$

Suppose that a coin is selected at random from the collection and tossed once. Let Y be the event that a head is obtained.

(a) What is the conditional pmf of $Y|X = x$?

(b) What is the joint distribution of X and Y?

(c) Suppose the outcome of the coin flip is heads. What is the conditional distribution of X given that the outcome is heads?

(a) The conditional pmf of $Y|X = x$ is

$$p(Y|X = x) = \begin{cases} x & \text{for } y = 1 \\ 1-x & \text{for } y = 0 \\ 0 & \text{otherwise.} \end{cases}$$

(b)

$$f_{X,Y}(x,y) = \begin{cases} 6x^2(1-x) & \text{for } 0 < x < 1 \text{ and } y = 1 \\ 6x(1-x)^2 & \text{for } 0 < x < 1 \text{ and } y = 0 \\ 0 & \text{otherwise.} \end{cases}$$

(c) To find the conditional distribution of $X|Y = 1$, we first need to find the marginal distribution of Y.

$$\begin{aligned} P(Y = 1) &= \int_0^1 6x^2(1-x)dx \\ &= \frac{1}{2}. \end{aligned}$$

Then we have

$$\begin{aligned} f_{X|Y}(X|Y = 1) &= \frac{f_{X,Y}(x,1)}{P(Y = 1)} \\ &= 12x^2(1-x). \end{aligned}$$

2. (Sum of geometric random variables)

(a) X_1 and X_2 are independent, geometric random variables with the same parameter, p. What is their joint distribution?

(b) What is the distribution of their sum, $Y = X_1 + X_2$?

(c) Which probability is greater: $P(X_1 = 13, X_2 = 3)$ or $P(X_1 = 8, X_2 = 8)$?

(d) If $X_1, ...X_n$ are geometric random variables with the same parameter, p, what is the distribution of their sum, $Z = \sum_{i=1}^n X_i$?

solution

(a) The joint distribution is:

$$p_{X_1,X_2}(x_1, x_2) = (1-p)^{x_1-1}p * (1-p)^{x_2-1}p = (1-p)^{x_1+x_2-2}p^2, \text{ where } x_1, x_2 \geq 1 \tag{1}$$

(b) Intuitively, the sum of two geometric random variables is the number of events before two successes. So the distribution of the sum can be written:

$$p_Y(y) = \binom{x_1 + x_2 - 1}{1} p_{X_1+X_2}(x_1, x_2) = \binom{x_1 + x_2 - 1}{1} (1-p)^{x_1+x_2-2}p^2 = \tag{2}$$

$$= \binom{y-1}{1} (1-p)^{y-2}p^2, \text{ where } y = x_1 + x_2 \text{ and } y \geq 2 \tag{3}$$

(c) Since both random variables have the same parameter, the "failed events" can be distributed in any way over X_1 and X_2 and their probability is the same.

(d) We can extend the intuition of part b to the sum of n geometric random variables:

$$p_Z(z) = p_{X_1+...+X_n}(x_1, ..., x_n) = \binom{(\sum_{i=1}^n x_i) - 1}{n-1} (1-p)^{x_1+...+x_n-n}p^n = \tag{4}$$

$$= \binom{z-1}{n-1} (1-p)^{z-n}p^n, \text{ where } z = \sum_{i=1}^n x_i \text{ and } z \geq n \tag{5}$$

3. (Markov chain)

(a)

$$T = \begin{bmatrix} 0.5 & 0.25 & 0 \\ 0.5 & 0.5 & 0.5 \\ 0 & 0.25 & 0.5 \end{bmatrix}$$

(b) Eigenvalues:

$$\lambda_1 = 0, \lambda_2 = \frac{1}{2}, \lambda_3 = 1$$

Corresponding eigenvectors:

$$v_1 = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \tag{6}$$

$$v_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \tag{7}$$

$$v_3 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \tag{8}$$

Let v denote our initial state vector. Then by Lemma 4.3 we know:

$$p_{\bar{x}[n]} = T^{n-1}v$$

Our initial vector v is a probability distribution, so it must sum to 1. Since our eigenvectors v_1, v_2, v_3 are a basis for $\mathbb{R}^3$, we can write our initial vector v as a unique linear combination of the eigenvectors:

$$v = av_1 + bv_2 + cv_3$$

Since, the elements in both v_1 and v_2 sum to zero, but we know the elements of the initial vector v must sum to 1, then we must have $c > 0$. In fact, we must have $c = 0.25$ in order for the elements of v to sum to 1.

To see how our Markov chain converges over time, we look at what happens to $p_{\bar{x}[n]} = T^{n-1}v$ over many steps, i.e. as n tends to infinity.

$$T^{n-1}v = T^{n-2}(Tv) = T^{n-2}(\lambda_1 av_1 + \lambda_2 bv_2 + \lambda_3 cv_3)$$

Then $a\lambda_1^n$ tends to 0, $b\lambda_2^n$ tends to 0, and so our state vector tends to cv_3 where $c = 0.25$.

Therefore the Markov chain converges to the following fixed state vector, also called the stationary distribution:

$$\begin{bmatrix} 0.25 \\ 0.5 \\ 0.25 \end{bmatrix}$$