UC Berkeley
Department of Electrical Engineering and Computer Sciences

EECS 126: PROBABILITY AND RANDOM PROCESSES

Discussion 3
Spring 2021

1. Covariance Matrix and Independence

For random variables $X_1, X_2, \ldots, X_n$, define the covariance matrix as a matrix Σ with entries $\Sigma_{ij} = \text{cov}(X_i, X_j)$ for all $i, j \in \{1, \ldots, n\}$. For this question, let $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T$, and assume $\mathbb{E}[\mathbf{X}] = \mathbf{0}$.

(a) Show that Σ is symmetric and positive semi-definite (PSD). Recall that a square matrix $A \in \mathbb{R}^{n \times n}$ is symmetric iff $A = A^T$ and PSD iff $\mathbf{u}^T A \mathbf{u} \geq 0$ for all $\mathbf{u} \in \mathbb{R}^n$.
Hint: For any vector of random variables, $\Sigma = \mathbb{E}[(\mathbf{X} - \mathbb{E}[\mathbf{X}])(\mathbf{X} - \mathbb{E}[\mathbf{X}])^T]$, can you see for yourself why this is true?

(b) Show that if the X_i's are pairwise independent, then Σ is a diagonal matrix.

(c) Give an example of two random variables X_1 and X_2 with a diagonal covariance matrix, but X_1 and X_2 are not independent.

Solution:

(a) Σ is symmetric because $\Sigma_{ij} = \text{cov}(X_i, X_j) = \text{cov}(X_j, X_i) = \Sigma_{ji}$.
Σ is PSD because for any $\mathbf{u} \in \mathbb{R}^n$,

$$\begin{aligned}
\mathbf{u}^T \Sigma \mathbf{u} &= \mathbf{u}^T\, \mathbb{E}[\mathbf{X}\mathbf{X}^T]\mathbf{u} \\
&= \mathbb{E}[\mathbf{u}^T \mathbf{X}\mathbf{X}^T \mathbf{u}] \qquad \text{by linearity of expectation} \\
&= \mathbb{E}[(\mathbf{u}^T\mathbf{X})(\mathbf{u}^T\mathbf{X})^T] \\
&= \mathbb{E}[S^2] \qquad \text{where } S = \mathbf{u}^T\mathbf{X} = \sum_{i=1}^{n} u_i X_i \text{ is a zero-mean scalar r.v.} \\
&= \text{var}(S) \\
&\geq 0
\end{aligned}$$

Note: This result holds even if $\mathbf{X}$ is not zero-mean, since we can replace $\mathbf{X}$ with the zero-mean r.v. zero-mean $\dot{\mathbf{X}} = \mathbf{X} - \mathbb{E}[\mathbf{X}]$ and do the same calculations.

(b) Σ is a diagonal matrix if $\Sigma_{ij} = 0$ for all $i \neq j$. We know that if two r.v. X_i and X_j are independent, then $\mathbb{E}[X_i X_j] = \mathbb{E}[X_i]\,\mathbb{E}[X_j]$, so

$$\begin{aligned}
\Sigma_{ij} = \text{cov}(X_i, X_j) &= \mathbb{E}[X_i X_j] - \mathbb{E}[X_i]\,\mathbb{E}[X_j] \\
&= \mathbb{E}[X_i]\,\mathbb{E}[X_j] - \mathbb{E}[X_i]\,\mathbb{E}[X_j] \\
&= 0, \text{for } i \neq j
\end{aligned}$$

(c) Consider $X_1 \sim \mathcal{N}(0, 1)$ and $X_2 = ZX_1$, where $Z \in \{1, -1\}$ with probability $\{1/2, 1/2\}$ (Z is called a Rademacher random variable). Now since $E[X_1] = 0$, we get $\mathbb{E}[X_1]\,\mathbb{E}[X_2] = 0$.

Also, $\mathbb{E}[X_1 X_2] = \mathbb{E}[Z X_1^2] = \mathbb{E}[Z]\,\mathbb{E}[X_1^2] = 0$. Therefore, $\mathrm{cov}(X_1, X_2) = \mathrm{cov}(X_2, X_1) = \mathbb{E}[X_1 X_2] - \mathbb{E}[X_1]\,\mathbb{E}[X_2] = 0$. So, X_1 and Y_2 are uncorrelated (they have diagonal covariance matrix). But observe that X_2 is generated based on X_1, so X_2 cannot be independent of X_1.

2. Poisson Merging

(a) Let $X \sim \text{Poisson}(\lambda)$ and $Y \sim \text{Poisson}(\mu)$ be independent Poisson random variables. Prove that $X + Y \sim \text{Poisson}(\lambda + \mu)$. (This property known as **Poisson merging** will be extensively used when we discuss Poisson processes.)

Note that it is **not** sufficient to use linearity of expectation to say that $X + Y$ has mean $\lambda + \mu$. You are asked to prove that the *distribution* of $X + Y$ is Poisson.

Hint: You may find the binomial theorm helpful, which states $(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$.

(b) Suppose that you and your friend are folding paper cranes. However, you both are very slow at it. The number of paper cranes you and your friend can fold in an hour follows a Poisson distribution with mean 3 and mean 2 respectively (A Poisson is appropriate here since finishing a paper crane is a rare event).

What is the distributions of the total number of paper cranes you and your friend can fold together in an hour?

Solution:

(a) For $z \in \mathbb{N}$,

$$\begin{aligned}
\Pr(X + Y = z) &= \sum_{j=0}^{z} \Pr(X = j, Y = z - j) = \sum_{j=0}^{z} \frac{\mathrm{e}^{-\lambda}\lambda^j}{j!} \frac{\mathrm{e}^{-\mu}\mu^{z-j}}{(z-j)!} \\
&= \frac{\mathrm{e}^{-(\lambda+\mu)}}{z!} \sum_{j=0}^{z} \frac{z!}{j!(z-j)!} \lambda^j \mu^{z-j} \\
&= \frac{\mathrm{e}^{-(\lambda+\mu)}}{z!} \sum_{j=0}^{z} \binom{z}{j} \lambda^j \mu^{z-j} = \frac{\mathrm{e}^{-(\lambda+\mu)}(\lambda+\mu)^z}{z!}.
\end{aligned}$$

This is an incredible result; not a lot of distributions have this property! Here is some intuition for why Poisson merging holds. If we are interested in the number of customers entering a store in the next hour, we can discretize the hour into n time intervals, where n is a positive integer. In each time interval, independently of other time intervals, the probability that a female customer enters the store is λ/n and the probability that a male customer enters the store is μ/n. Since the two types of customers are assumed to be independent, the probability that a customer, disregarding gender, enters the store is $\lambda/n + \mu/n - \lambda\mu/n^2$. As $n \to \infty$, the number of customers who enter the store in the hour is Poisson with mean $\lim_{n\to\infty} n[\lambda/n + \mu/n - \lambda\mu/n^2] = \lambda + \mu$.

We will be able to give a much easier proof of this result after we introduce transforms of random variables.

(b) Using Poisson merging, the total number of paper cranes follows a Poisson distribution as well, with mean $3 + 2 = 5$.

3. **Triangle Density**

Consider random variables X and Y which have a joint PDF uniform on the triangle with vertices at $(0,0),(1,0),(0,1)$.

(a) Find the joint PDF of X and Y.

(b) Find the marginal PDF of Y.

(c) Find the conditional PDF of X given Y.

(d) Find $\mathbb{E}[X]$ in terms of $\mathbb{E}[Y]$.

(e) Find $\mathbb{E}[X]$.

Solution:

(a) Note that the joint PDF is uniform on the triangle, which has area $1/2$, so for all valid x, y, $f_{X,Y}(x,y) = 2$.

(b) In order to find the marginal PDF, we integrate out:

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}x = \int_0^{1-y} 2\,\mathrm{d}x = 2(1-y)$$

where $0 \le y \le 1$.

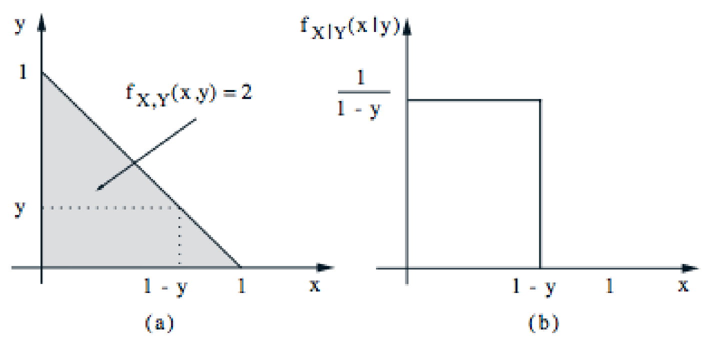

Figure 1: Joint density of (X, Y) (a) and the conditional density $X \mid Y$ (b). Image taken from Bertsekas and Tsitsiklis.

(c) The conditional density is given by, for $0 \le y \le 1$,

$$f_{X|Y}(x \mid y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{2}{2(1-y)} = \frac{1}{1-y}, \qquad 0 \le x \le 1-y.$$

This should agree with your intuition that given $Y = y$, X should be uniform.

(d) We use the tower property: $\mathbb{E}[\mathbb{E}(X \mid Y)] = \mathbb{E}[X]$. Note that for $0 \le y \le 1$,

$$\begin{aligned}\mathbb{E}[X \mid Y = y] &= \int_0^{1-y} x f_{X|Y}(x \mid y)\,\mathrm{d}x = \int_0^{1-y} x \frac{1}{1-y}\,\mathrm{d}x \\ &= \frac{1}{1-y}\left[\frac{(1-y)^2}{2}\right] = \frac{1-y}{2}.\end{aligned}$$

Thus, we have:

$$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}(X \mid Y)] = \int_0^1 \mathbb{E}[X \mid Y = y] f_Y(y)\,\mathrm{d}y.$$

Note that we are simply trying to find $\mathbb{E}[X]$ in terms of $\mathbb{E}[Y]$, so there is no need to expand out $f_Y(y)$, so we have:

$$\begin{aligned}\mathbb{E}[X] &= \int_0^1 \mathbb{E}[X \mid Y = y] f_Y(y)\,\mathrm{d}y = \int_0^1 \frac{1-y}{2} f_Y(y)\,\mathrm{d}y \\ &= \frac{1}{2} - \frac{1}{2}\int_0^1 y f_Y(y)\,\mathrm{d}y = \frac{1-\mathbb{E}[Y]}{2}.\end{aligned}$$

(e) Finally, we note that by symmetry, $\mathbb{E}[X]$ should be equal to $\mathbb{E}[Y]$, so we have

$$\mathbb{E}[X] = \frac{1-\mathbb{E}[X]}{2},$$

and

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