Vectors of Random Variables

Guy Lebanon

January 6, 2006

So far we treated random variables $X : \Omega \to \mathbb{R}$ and probabilities associated with them $P(X \in A)$. Much of the remaining material in the course will involve multiple RV and their probabilities. We will consider for the next several weeks what happens with a finite number of RV $X_1, \ldots, X_n$, where each X_i is a function $X_i : \Omega \to \mathbb{R}$. Later on in the course, we will discuss random processes which are an infinite number of random variables.

Definition 1. *Let Ω be a sample space, P a probability distribution $P : \Omega \to \mathbb{R}$ on it, and $(X_1, \ldots, X_n)$ be n random variables $X_i : \Omega \to \mathbb{R}$. We then also say that $\vec{\boldsymbol{X}} = (X_1, \ldots, X_n)$ is a vector of random variables or a random vector. Alternatively, we can think of $\vec{\boldsymbol{X}}$ as a function $\vec{\boldsymbol{X}} : \Omega \to \mathbb{R}^n$ as follows $\vec{\boldsymbol{X}}(\omega) = (X_1(\omega), \ldots, X_n(\omega)) \in \mathbb{R}^n$. Assuming that $\boldsymbol{A} \subset \mathbb{R}^n$ we write $\vec{\boldsymbol{X}} \in \boldsymbol{A}$ for the event $\{\omega \in \Omega : (X_1(\omega), \ldots, X_n(\omega)) \in \boldsymbol{A}\}$ and $P(\vec{\boldsymbol{X}} \in \boldsymbol{A})$ for its probability $P(\vec{\boldsymbol{X}} \in \boldsymbol{A}) = P(\{\omega \in \Omega : \vec{\boldsymbol{X}}(\omega) \in \boldsymbol{A})$.*

Random vectors are discrete, continuous or neither. The definition is very similar to the one dimensional case: $\vec{\boldsymbol{X}}$ is discrete if there exists a finite or countable set K s.t. $P(\vec{\boldsymbol{X}} \in K) = 1$ and $\vec{\boldsymbol{X}}$ is continuous if $P(\vec{\boldsymbol{X}} = (x_1, \ldots, x_n)) = 0$ for all $(x_1, \ldots, x_n)$. For example, the voltage measurement at several locations of a circuit is a continuous random vector. The measurement of color and sex of a particular species of animals is a discrete random vector (note that in this case we have to map the color and sex to real numbers e.g. brown=0, white=1, black=2 and male=0, female=1). The measurements of height, weight and gender of a person is neither discrete nor continuous.

Often (but not always), the events $\vec{\boldsymbol{X}} \in \boldsymbol{A}$ will take a simpler form:

$$\begin{aligned}\{X_1 \in A_1, \ldots, X_n \in A_n\} &= \{\omega \in \Omega : X_1(\omega) \in A_1, \ldots, X_n(\omega) \in A_n\} \\ &= \{X_1 \in A_1\} \cap \cdots \cap \{X_n \in A_n\} \subset \Omega\end{aligned}$$

and $P(X_1 \in A_1, \ldots, X_n \in A_n)$ denotes the probability of that event. Such an event is called a factored event (you may visualize it as an n dimensional rectangle of sides $A_1, \ldots, A_n$).

Definition 2. *The random variables $X_1, \ldots, X_n$ are independent if for all $A_1, \ldots, A_n$*

$$P(X_1 \in A_1, \ldots, X_n \in A_n) = P(X_1 \in A_1) \cdots P(X_n \in A_n)$$

For example, let Ω describes the result of throwing a pair of dice, X_1 is a RV measuring the sum of the two dice and X_2 is a RV measuring the difference of the two dice. We have $P(X_1 = 2, X_2 = 2) = P(X_1 \in \{2\}, X_2 \in \{2\}) = \emptyset, P(X_1 = 6, X_2 = 2) = P(\{(4, 2)\}) = 1/36$ and

$$P(X_1 > 4, X_2 < 0) = P(\{(1,4),(1,5),(1,6),(2,3),(2,4),(2,5),(2,6),(3,4),(3,5),(3,6),(4,5),(4,6),(5,6)\} = 13/36.$$

We also have $P(X_1 > 4) = 1 - P(\{(1,1),(1,2),(2,1),(1,3),(3,1),(2,2)\}) = 30/36$ and $P(X_2 < 0) = 15/36$. Since $P(X_1 > 4, X_2 < 0) \neq P(X_1 > 4)P(X_2 < 0)$, X_1, X_2 are not independent.

As with one dimensional RV we define cdf, pmf and pdf to help us calculate probabilities. The cdf of $\vec{\boldsymbol{X}}$ is defined as

$$F_{\vec{\boldsymbol{X}}}(x_1, \ldots, x_n) = P(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_n \leq x_n)$$

the pdf is defined as its multiple derivative

$$f_{\vec{\boldsymbol{X}}(x_1,\ldots,x_n)}(x_1, \ldots, x_n) = \frac{\partial^n}{\partial x_1 \cdots \partial x_n} F_{\vec{\boldsymbol{X}}}(x_1, \ldots, x_n)$$

if it exists and 0 otherwise and the pmf is defined as

$$p_{\vec{\boldsymbol{X}}}(x_1, \ldots, x_n) = P(X_1 = x_1, \ldots, X_n = x_n).$$

We will sometime abbreviate the above formulas by denoting $\vec{\boldsymbol{x}} = (x_1, \ldots, x_n)$ e.g. $p_{\vec{\boldsymbol{X}}}(\vec{\boldsymbol{x}}) = P(\vec{\boldsymbol{X}} = \vec{\boldsymbol{x}})$. As in the one dimensional case, the pdf (for continuous $\vec{\boldsymbol{X}}$) has to be non-negative and integrate to 1 $\int \cdots \int f_{\vec{\boldsymbol{X}}}(\vec{\boldsymbol{x}}) d\vec{\boldsymbol{x}} = 1$, and the pmf (for discrete $\vec{\boldsymbol{X}}$) has to be non-negative and sum to one $\sum_{x_1} \cdots \sum_{x_2} p_{\vec{\boldsymbol{X}}}(\vec{\boldsymbol{x}}) = 1$.

From the fundamental theorem of calculus, we have that for continuous RV

$$F_{\vec{\boldsymbol{X}}}(x_1, x_2, \ldots, x_n) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \cdots \int_{-\infty}^{x_n} f_{\vec{\boldsymbol{X}}}(t_1, t_2, \ldots, t_n)\ dt_1 dt_2 \cdots dt_n.$$

The formulas for computing $P(\vec{\boldsymbol{X}} \in \boldsymbol{A}), \boldsymbol{A} \subset \mathbb{R}^n$ are similar to the one-dimensional case, and their proofs carry over without much problems

$$P(\vec{\boldsymbol{X}} \in \boldsymbol{A}) = \begin{cases} \sum_{\vec{\boldsymbol{x}} \in \boldsymbol{A}} p_{\vec{\boldsymbol{X}}}(\vec{\boldsymbol{x}}) & \vec{\boldsymbol{X}} \text{ is a discrete random vector} \\ \int_{\vec{\boldsymbol{X}} \in \boldsymbol{A}} f_{\vec{\boldsymbol{X}}}(\vec{\boldsymbol{x}})\, d\vec{\boldsymbol{x}} & \vec{\boldsymbol{X}} \text{ is a continuous random vector} \end{cases}.$$

The marginal pmf, pdf and cdf are obtained as follows:

$$p_{X_i}(x_i) = \sum_{x_1} \cdots \sum_{x_{i-1}} \sum_{x_{i+1}} \cdots \sum_{x_n} p_{\vec{\boldsymbol{X}}}(x_1, \ldots, x_n) \quad \text{e.g.} \quad p_{X_1}(x_1) = \sum_{x_2 \in \mathbb{R}} p_{X_1, X_2}(x_1, x_2)$$

$$f_{X_i}(x_i) = \int \cdots \int f_{\vec{\boldsymbol{X}}}(x_1, \ldots, x_n)\, dx_1 \cdots dx_{i-1} dx_{i+1} dx_n \quad \text{e.g.} \quad f_{X_1}(x_1) = \int_{-\infty}^{\infty} f_{X_1, X_2}(x_1, x_2)\, dx_2$$

$$F_{X_i}(x_i) = F_{\vec{\boldsymbol{X}}}(\infty, \cdots, \infty, x_i, \infty, \cdots, \infty) \quad \text{e.g.} \quad F_{X_1}(x_1) = F_{X_1, X_2}(x_1, \infty)$$

The following result is useful when the variables are independent and we want to compute $P(\vec{\boldsymbol{X}} \in \boldsymbol{A})$ or if we want to check whether some random variables are indeed independent.

Theorem 1. *The vector $\vec{\boldsymbol{X}}$ are independent if and only if the cdf factors $F_{\vec{\boldsymbol{X}}}(x_1, \ldots, x_n) = F_{X_1}(x_1) \cdots F_{X_n}(x_n)$. $\vec{\boldsymbol{X}}$ is also independent if and only if the pmf factors (for discrete random vectors) $p_{\vec{\boldsymbol{X}}}(x_1, \ldots, x_n) = p_{X_1}(x_1) \cdots p_{X_n}(x_n)$ or the pdf factors (for continuous random vectors) $f_{\vec{\boldsymbol{X}}}(x_1, \ldots, x_n) = f_{X_1}(x_1) \cdots f_{X_n}(x_n)$.*

Example: Suppose that the vector (X, Y) has the joint-normal pdf $f_{X,Y}(x, y) = \frac{1}{2\pi} e^{-(x^2+y^2)/2}$. Since the pdf factors we know that $f_{X,Y}(x, y) = f_X(x) f_Y(x)$ where X, Y are independent normal distribution with $\mu = 0, \sigma = 1$ $f_X(x) = \frac{1}{2\pi} e^{-x^2/2}$. Because of this factorization of the pdf we can compute probabilities of factored sets (rectangles) more easily:

$$P(3 < X < 6, Y < 9) = \int_3^6 \int_{\infty}^9 f_{X,Y}(x, y)\, dxdy = \int_3^6 \int_{\infty}^9 f_X(x) f_Y(y)\, dxdy = \int_3^6 f_X(x) dx \int_{\infty}^9 f_Y(y) dy.$$

Example: Let (X, Y) be a random vector where X, Y are independent exponential random variables with parameters λ_1, λ_2. Then

$$f_{X,Y}(x, y) = f_X(x) f_Y(y) = \lambda_1 \lambda_2 e^{-\lambda_1 x} e^{-\lambda_2 y} = \lambda_1 \lambda_2 e^{-\lambda_1 x - \lambda_2 y}$$

and the joint cdf is

$$F_{X,Y}(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} \lambda_1 \lambda_2 e^{-\lambda_1 x - \lambda_2 y} dxdy = \int_{-\infty}^{x} \lambda_1 e^{-\lambda_1 x} \int_{-\infty}^{y} \lambda_2 e^{-\lambda_2 y} = (1 - e^{-\lambda_1 x})(1 - e^{-\lambda_1 x}) = F_X(x) F_Y(y).$$