

MA2202: PROBABILITY I

Random vectors

Spring 2021

Satvik Saha
19MS154

*Indian Institute of Science Education and Research, Kolkata,
Mohanpur, West Bengal, 741246, India.*

Definition 4.1 (Random vector). A random vector $\mathbf{X}: \Omega \rightarrow \mathbb{R}^n$ is a tuple of random variables $X_i: \Omega \rightarrow \mathbb{R}$.

Definition 4.2 (Joint cumulative distribution function). The joint cumulative distribution function of a random vector \mathbf{X} is the map $F_{\mathbf{X}}: \mathbb{R}^n \rightarrow [0, 1]$, given as

$$F_{\mathbf{X}}(\mathbf{s}) = P(X_1 \leq s_1, \dots, X_n \leq s_n).$$

Definition 4.3 (Joint probability mass function). If X_i are discrete random variables, their joint probability mass function is the map $p_{\mathbf{X}}: \mathbb{R}^n \rightarrow [0, 1]$,

$$p_{\mathbf{X}}(\mathbf{s}) = P(X_1 = s_1, \dots, X_n = s_n).$$

Definition 4.4 (Joint probability density function). Suppose that

$$F_{\mathbf{X}}(\mathbf{s}) = \int_{-\infty}^{s_n} \cdots \int_{-\infty}^{s_1} f_{\mathbf{X}}(t_1, \dots, t_n) dt_1 \cdots dt_n,$$

then $f_{\mathbf{X}}: \mathbb{R}^n \rightarrow [0, 1]$ is the probability density function corresponding to the joint cumulative distribution function $F_{\mathbf{X}}$.

Remark. If $f_{\mathbf{X}}$ is continuous, then

$$f_{\mathbf{X}} = \frac{\partial F_{\mathbf{X}}(t_1, \dots, t_n)}{\partial t_1 \cdots \partial t_n}.$$

Definition 4.5 (Joint moment generating function). Let \mathbf{X} be a random vector. Then, its joint moment generating function is defined as

$$M_{\mathbf{X}}(\mathbf{t}) = E \left[e^{\mathbf{t}^\top \mathbf{X}} \right] = E \left[e^{t_1 X_1 + \dots + t_n X_n} \right].$$

Remark. If X_1, \dots, X_n are independent, then

$$M_{\mathbf{X}}(\mathbf{t}) = \prod M_{X_i}(t_i).$$

Theorem 4.1. If X and Y are independent continuous random variables, then the probability density function of their sum is the convolution $f_{X+Y} = f_X * f_Y$,

$$f_{X+Y}(x) = \int_{\mathbb{R}} f_X(x-t) f_Y(t) dt.$$

Example. When X and Y are identical and uniform on $[0, 1]$, then

$$f_{X+Y}(x) = \int_0^1 f(x-t) dt = \begin{cases} x, & \text{if } x \in [0, 1], \\ 2-x, & \text{if } x \in [1, 2], \\ 0, & \text{otherwise.} \end{cases}$$

Also,

$$M_{X+Y}(t) = (M(t))^2 = \frac{1}{t^2}(e^t - 1)^2.$$

Definition 4.6 (Conditional distribution). Let X and Y be two discrete random variables. We write

$$P(X = s | Y = t) = \frac{P(X = s, Y = t)}{P(Y = t)}$$

for $P(Y = t) > 0$. We also have

$$P(X \leq s | Y = t) = \sum_{r \leq s} P(X = r | Y = t).$$

If X and Y are continuous random variables, then the conditional distribution of X given $Y = t$ is described as

$$F_{X|Y=t}(r) = \int_{-\infty}^r \frac{f_{X,Y}(r, t)}{f_Y(t)} dr.$$

Example. Consider two continuous random variables X and Y which have a joint probability mass function

$$f_{X,Y}(s, t) = \begin{cases} \alpha t, & \text{if } 0 < t < s < 1, \\ 0, & \text{otherwise.} \end{cases}$$

First normalize, by demanding

$$\iint_{\mathbb{R}^2} f_{X,Y}(s, t) dt ds = \int_0^1 \int_0^s \alpha t dt ds = 1,$$

whence $\alpha = 6$. Thus,

$$E[Y | X = s] = \int_{\mathbb{R}} t \cdot \frac{f_{X,Y}(s, t)}{f_X(s)} dt.$$

Now,

$$f_X(s) = \int_{\mathbb{R}} f_{X,Y}(s, t) dt = \int_0^s 6t \alpha dt = 3s^2$$

for $0 < s < 1$, and simply 3 for $s \geq 1$. Thus,

$$E[Y | X = s] = \int_0^s t \cdot \frac{6t}{3s^2} dt = \frac{2}{3}s,$$

in the region $0 < s < 1$. For $s \geq 1$, the expectation becomes $2/3$. Also,

$$\text{Var}[Y | X = s] = E[Y^2 | X = s] - E[Y | X = s]^2.$$

The first term is

$$E[Y^2 | X = s] = \int_0^s t^2 \cdot \frac{6t}{3s^2} dt = \frac{1}{2}s^2.$$

Thus,

$$\text{Var}[Y | X = s] = \frac{1}{2}s^2 - \frac{4}{9}s^2 = \frac{1}{18}s^2.$$

Note that

$$f_Y(t) = \int_{\mathbb{R}} f_{X,Y}(s, t) ds = \int_t^1 6t \alpha ds = 6t(1 - t)$$

in the region $0 < t < 1$. Thus,

$$F_Y(t) = \int_0^t 6t'(1 - t') dt' = t^2(3 - 2t)$$

for $0 < t < 1$. $F_Y(t) = 1$ for $t \geq 1$.

Theorem 4.2. For discrete or continuous random variables X and Y ,

$$E[E[X | Y]] = E[X].$$

Proof.

$$E[E[X | Y]] = \sum_n E[X | Y = n] P(Y = n) = \sum_{nm} m P(X = m, Y = n).$$

Reordering the summations, we get

$$\sum_m m \sum_n P(X = m, Y = n) = \sum_m m P(X = m) = E[X].$$

The proof for discrete random variables is analogous, switching the sums for integrals. \square

Theorem 4.3. *For random variables X and Y ,*

$$\text{Var}[X] = \text{Var}[E[X | Y]] + E[\text{Var}[X | Y]].$$

Proof. Using the previous theorem,

$$\text{Var}[E[X | Y]] = E[E[X | Y]^2] - E[E[X | Y]]^2 = E[E[X | Y]^2] - E[X]^2,$$

and

$$E[\text{Var}[X | Y]] = E[E[X^2 | Y] - E[X | Y]^2] = E[X^2] - E[E[X | Y]^2].$$

Adding the above gives the desired result. \square