1. Week 11 Supplementary material

Remark 1.1. Let $X = (X_1, \ldots, X_p)$ be a p-dimensional discrete/continuous random vector with joint p.m.f./p.d.f. f_X. We are interested in the distribution of $Y = h(X)$ for functions $h : \mathbb{R}^p \to \mathbb{R}^q$. Here, $Y = (Y_1, \ldots, Y_q)$ is a q-dimensional random vector with $Y_j = h_j(X_1, \ldots, X_p)$, where $h_j : \mathbb{R}^p \to \mathbb{R}, j = 1, 2, \cdots, q$ denotes the component functions of h. The distribution of Y is uniquely determined as soon as we are able to compute the joint DF F_Y of Y. Note that

$$F_Y(y_1, \cdots, y_q) = \mathbb{P}(Y_1 \le y_1, \cdots, Y_q \le y_q) = \mathbb{P}(h_1(X) \le y_1, \cdots, h_q(X) \le y_q), \forall (y_1, \cdots, y_q) \in \mathbb{R}^q.$$

Once the joint DF F_Y is known, the joint p.m.f./p.d.f. of Y can then be deduced by standard techniques.

Example 1.2. Let $X_1 \sim Uniform(0,1)$ and $X_2 \sim Uniform(0,1)$ be independent RVs. Suppose we are interested in the distribution of $Y = X_1 + X_2$. By independence of X_1 and X_2, the joint p.d.f. (X_1, X_2) is given by

$$\begin{aligned} f_{X_1,X_2}(x_1, x_2) &= f_{X_1}(x_1) f_{X_2}(x_2) \\ &= \begin{cases} 1, & \text{if } x_1, x_2 \in (0,1) \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

Consider the function $h : \mathbb{R}^2 \to \mathbb{R}$ defined by $h(x_1, x_2) := x_1 + x_2, \forall (x_1, x_2) \in \mathbb{R}^2$. Then $Y = h(X_1, X_2)$. Now, for $y \in \mathbb{R}$

$$\begin{aligned} F_Y(y) &= \mathbb{P}(Y \le y) \\ &= \mathbb{P}(h(X_1, X_2) \le y) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} 1_{(-\infty,y]}(h(x_1, x_2)) f_{X_1,X_2}(x_1, x_2)\, dx_1 dx_2 \\ &= \int_0^1 \int_0^1 1_{(-\infty,y]}(x_1 + x_2)\, dx_1 dx_2 \end{aligned}$$

$$= \begin{cases} 0, \text{ if } y < 0, \\ \int_{x_1=0}^{y} \int_{x_2=0}^{y-x_1} \, dx_2 dx_1, \text{ if } 0 \le y < 1, \\ 1 - \frac{1}{2} \times (2-y) \times (2-y), \text{ if } 1 \le y < 2, \\ 1, \text{ if } y \ge 2 \end{cases}$$

$$= \begin{cases} 0, \text{ if } y < 0, \\ \frac{y^2}{2}, \text{ if } 0 \le y < 1, \\ \frac{4y-y^2-2}{2}, \text{ if } 1 \le y < 2, \\ 1, \text{ if } y \ge 2 \end{cases}$$

Here, F_Y is differentiable everywhere except possibly at the points $0, 1, 2$ and

$$F_Y'(y) = \begin{cases} y, \text{ if } y \in (0,1), \\ 2-y, \text{ if } y \in (1,2), \\ 0, \text{ otherwise.} \end{cases}$$

Observe that $\int_{-\infty}^{\infty} F_Y'(y)\, dy = 1$ and the derivative is non-negative. Hence, Y is a continuous RV with the p.d.f. given by F_Y'.

Example 1.3. Fix $p \in (0,1)$ and let $n_1, \cdots, n_q$ be positive integers. Let $X_1, \cdots, X_q$ be independent RVs with $X_i \sim Binomial(n_i, p), i = 1, \cdots, q$. Here, the with p.m.f.s are given by

$$f_{X_i}(x_i) = \begin{cases} \binom{n_i}{x} p^x (1-p)^{n_i - x}, \forall x \in \{0, 1, \cdots, n_i\}, \\ 0, \text{ otherwise} \end{cases}$$

for $i = 1, \cdots, q$. Using independence, the joint p.m.f. is given by

$$f_X(x_1, \cdots, x_q) = \begin{cases} \prod_{i=1}^{q} \binom{n_i}{x_i} p^{\sum_{i=1}^{q} x_i} (1-p)^{n - \sum_{i=1}^{q} x_i}, \forall (x_1, \cdots, x_q) \in \prod_{i=1}^{q} \{0, 1, \cdots, n_i\}, \\ 0, \text{ otherwise} \end{cases}$$

where $n = n_1 + \cdots + n_q$. Consider $Y = X_1 + \cdots + X_q$. Now, if $y \notin \{0, 1, \cdots, n\}$, $f_Y(y) = \mathbb{P}(X_1 + \cdots + X_q = y) = 0$ and if $y \in \{0, 1, \cdots, n\}$, then

$$\begin{aligned}
f_Y(y) &= \mathbb{P}(X_1 + \cdots + X_q = y) \\
&= \sum_{\substack{(x_1, \cdots, x_q) \in \prod_{i=1}^q \{0,1,\cdots,n_i\} \\ x_1 + \cdots + x_q = y}} f_X(x_1, \cdots, x_q) \\
&= p^y (1-p)^{n-y} \sum_{\substack{(x_1, \cdots, x_q) \in \prod_{i=1}^q \{0,1,\cdots,n_i\} \\ x_1 + \cdots + x_q = y}} \prod_{i=1}^q \binom{n_i}{x_i} \\
&= \binom{n}{y} p^y (1-p)^{n-y}.
\end{aligned}$$

Therefore, $Y = X_1 + \cdots + X_q \sim Binomial(n, p)$ with $n = n_1 + \cdots + n_q$.

Remark 1.4. We had earlier mentioned that $Bernoulli(p)$ distribution is the same as $Binomial(1, p)$ distribution. Using the above computation, we can identify a $Binomial(n, p)$ RV as a sum of n independent RVs each having distribution $Bernoulli(p)$.

Remark 1.5 (Correlation and Independence). If X and Y are independent RVs defined on the same probability space, then $Cov(X, Y) = 0$ and hence X and Y are uncorrelated. However, the converse is not true. We illustrate this problem with examples.

(a) Let $X = (X_1, X_2)$ be a bivariate discrete random vector, i.e. a 2-dimensional discrete random vector with joint p.m.f. given by

$$f_X(x_1, x_2) = \begin{cases} \frac{1}{2}, & \text{if } (x_1, x_2) = (0, 0), \\ \frac{1}{4}, & \text{if } (x_1, x_2) = (1, 1) \text{ or } (1, -1), \\ 0, & \text{otherwise.} \end{cases}$$

The marginal p.m.fs are

$$f_{X_1}(x_1) = \begin{cases} \frac{1}{2}, \text{ if } x_1 \in \{0,1\} \\ 0, \text{ otherwise} \end{cases} \quad , \quad f_{X_2}(x_2) = \begin{cases} \frac{1}{2}, \text{ if } x_2 = 0 \\ \frac{1}{4}, \text{ if } x_2 \in \{1,-1\} \\ 0, \text{ otherwise} \end{cases}$$

We have $f_{X_1,X_2}(0,0) = \frac{1}{2} \neq \frac{1}{4} = f_{X_1}(0) f_{X_2}(0)$ and hence X_1 and X_2 are not independent. But, $\mathbb{E}X_1 = \frac{1}{2}, \mathbb{E}X_2 = 0, \mathbb{E}(X_1X_2) = 0, Var(X_1) > 0$ and $Var(X_2) > 0$. Therefore $Cov(X_1, X_2) = 0$ and hence X_1 and X_2 are uncorrelated.

(b) Let $X = (X_1, X_2)$ be a bivariate continuous random vector, i.e. a 2-dimensional continuous random vector with joint p.d.f. given by

$$f_X(x_1, x_2) = \begin{cases} 1, & \text{if } 0 < |x_2| \leq x_1 < 1 \\ 0, & \text{otherwise.} \end{cases}$$

Then,

$$\mathbb{E}(X_1X_2) = \int_0^1 \int_{-x_1}^{x_1} x_1x_2 \, \mathrm{d}x_2 \, \mathrm{d}x_1 = 0,$$

and

$$\mathbb{E}(X_1) = \int_0^1 \int_{-x_M1}^{x_1} x_1 \, \mathrm{d}x_2 \, \mathrm{d}x_1 = \frac{2}{3}, \quad \mathbb{E}(X_2) = \int_0^1 \int_{-x_1}^{x_1} x_2 \, \mathrm{d}x_2 \, \mathrm{d}x_1 = 0.$$

Hence, $\mathbb{E}(X_1X_2) = (\mathbb{E}X_1)(\mathbb{E}X_2)$, which implies $Cov(X_1, X_2) = 0$. A similar computation shows $Var(X_1)$ and $Var(X_2)$ exists and are non-zero. Hence, X_1 and X_2 are uncorrelated. Now, by computing the marginal p.d.f.s f_{X_1} and f_{X_2}, it is immediate that the equality

$$f_{X_1,X_2}(x_1, x_2) = f_{X_1}(x_1) f_{X_2}(x_2)$$

does not hold for all $x = (x_1, x_2) \in \mathbb{R}^2$. Here, X_1 and X_2 are not independent. The verification with the marginal p.d.f.s is left as an exercise in practice problem set 10.