Recall

Independence of n random variables

For $n \geq 2$, let $X_1, \ldots, X_n$ be n random variables. The *joint cumulative distribution function* (joint CDF) of $X_1, \ldots, X_n$ is

$$F(a_1, \ldots, a_n) \coloneqq P(X_1 \leq a_1, \ldots, X_n \leq a_n), \quad \forall\, a_1, \ldots, a_n \in \mathbb{R}.$$

- $X_1, \ldots, X_n$ are *joint continuous* if there exists $f \colon \mathbb{R}^n \to [0, \infty)$ such that

  $$P\{(X_1, \ldots, X_n) \in C\} = \int \cdots \int_C f(x_1, \ldots, x_n) dx_1 \cdots dx_n$$

  for all "measurable" sets $C \subset \mathbb{R}^n$.

- $X_1, \ldots, X_n$ are *independent* if

  $$P(X_1 \in A_1, \ldots, X_n \in A_n) = P(X_1 \in A_1) \cdots P(X_n \in A_n), \quad \forall\, A_1, \ldots, A_n \subset \mathbb{R},$$

  which is equivalently characterized by the joint CDF

  $$F(a_1, \ldots, a_n) = F_{X_1}(a_1) \cdots F_{X_n}(a_n), \quad \forall\, a_1, \ldots, a_n \in \mathbb{R}.$$

Example 1 (Pairwise independence $\not\Longrightarrow$ independence)**.** Let $X \sim Bern(\frac{1}{2})$ and $Y \sim Bern(\frac{1}{2})$. Suppose that X, Y are independent. Define $Z = \begin{cases} 1 & \text{if } X \neq Y \\ 0 & \text{if } X = Y. \end{cases}$

Then the joint PMF of X, Y, Z is

$$p(x, y, z) = \begin{cases} \frac{1}{4} & (x, y, z) = (0, 0, 0) \\ \frac{1}{4} & (x, y, z) = (0, 1, 1) \\ \frac{1}{4} & (x, y, z) = (1, 0, 1) \\ \frac{1}{4} & (x, y, z) = (1, 1, 0) \\ 0 & \text{otherwise.} \end{cases}$$

Hence

$$p_X(a) = p_Y(a) = p_Z(a) = \begin{cases} \frac{1}{2} & a = 1 \\ \frac{1}{2} & a = 0 \\ 0 & \text{otherwise.} \end{cases}$$

and

$$p_{X,Y}(a, b) = p_{X,Z}(a, b) = p_{Y,Z}(a, b) = \begin{cases} \frac{1}{4} & (a, b) = (0, 0) \\ \frac{1}{4} & (a, b) = (0, 1) \\ \frac{1}{4} & (a, b) = (1, 0) \\ \frac{1}{4} & (a, b) = (1, 1) \\ 0 & \text{otherwise.} \end{cases}$$

Thus X, Y, Z are pairwise independent. However, since

$$p(0,0,0) = \frac{1}{4} \quad \text{while} \quad p_X(0)p_Y(0)p_Z(0) = \frac{1}{8},$$

we have X, Y, Z are NOT independent.

Example 2. Review [Tutorial 8, Example 3 & Example 4].

Convolution formula for sum of independent random variables

Let X, Y be **independent** random variables. Then

$$\begin{cases} \text{if } X,Y \text{ joint continuous:} & f_{X+Y} = f_X * f_Y, \text{ that is } \forall\, z \in \mathbb{R},\ f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(z-y) f_Y(y)dy. \\ \text{if } X,Y \text{ discrete:} & p_{X+Y} = p_X * p_Y, \text{ that is } \forall\, z \in \mathbb{R},\ p_{X+Y}(z) = \sum_y p_X(z-y)p_Y(y). \end{cases}$$

In particular,

- If $X, Y \sim U(0,1)$ independent, then $f_{X+Y}(z) = \begin{cases} z & 0 \le z \le 1 \\ 2-z & 1 < z \le 2 \\ 0 & \text{otherwise.} \end{cases}$
- If $X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$ independent, then $X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$. Combining with the previous result, we have for $a, b, c \in \mathbb{R}$, $aX + bY + c \sim N(a\mu_1 + b\mu_2 + c, a^2\sigma_1^2 + b^2\sigma_2^2)$.
- If $X \sim Poisson(\lambda_1), Y \sim Poisson(\lambda_2)$ independent, then $X + Y \sim Poisson(\lambda_1 + \lambda_2)$.

Example 3. For $m, n \in \mathbb{N}$ and $0 \le p \le 1$, let $X \sim Bin(n,p)$ and $Y \sim Bin(m,p)$ be independent. Then $X + Y \sim Bin(n+m, p)$.

Proof. Let $z \in \{0, \dots, n+m\}$. With convention $\binom{n}{k} = 0$ if $k > n$ and by the independence,

$$\begin{aligned} p_{X+Y}(z) &= p_X * p_Y(z) \\ &= \sum_{y=0}^{z} \binom{n}{z-y} p^{z-y}(1-p)^{n-(z-y)} \cdot \binom{m}{y} p^y (1-p)^{m-y} \\ &= \sum_{y=0}^{z} \binom{n}{z-y}\binom{m}{y} p^z (1-p)^{m+n-z} \\ &= \binom{m+n}{z} p^z (1-p)^{m+n-z} \end{aligned}$$

where the third equality is by $\sum_{y=0}^{z} \binom{n}{z-y}\binom{m}{y} = \binom{m+n}{z}$. When $z \notin \{0, \dots, n+m\}$, we have $p_{X+Y}(z) = 0$. Hence $X + Y \sim Bin(n+m, p)$. □