CSE 312: Foundations of Computing II **Summer 2020**

Lecture 5.5: Convolution

Lecturer: Alex Tsun *Monday, July 27th* *Scribe: Pemi Nguyen*

5.5.1 Law of Total Probability for Random Variables

Definition 5.5.1.1: Law of Total Probability for Random Variables

Discrete version: If X, Y are discrete random variables:

$$p_X(x) = \sum_y p_{X,Y}(x,y) = \sum_y p_{X|Y}(x|y)p_Y(y)$$

Continuous version: If X, Y are continuous random variables:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy = \int_{-\infty}^{\infty} f_{X|Y}(x|y)f_Y(y)dy$$

Examples

Let $X, Y \sim Unif(1,4)$ be independent rolls of a fair 4-sided die. What is the PMF of $Z = X + Y$? Well we know that for the range of Z we have the following, since it is the sum of two values each in the range $\{1,2,3,4\}$:

$$\Omega_Z = \{2,3,4,5,6,7,8\}$$

To solve for the probability mass function we can consider the probability of getting each pair of X and Y that sum up to some value, but this would be complicated to consider all the cases of, so instead we can rewrite this in terms of probabilities of X and Y (since $Y = Z - X$), which we know both of the probability mass functions for. Then because X and Y are independent, we can separate this. So we have the following:

$$\begin{aligned} p_Z(z) &= \mathbb{P}(Z = z) \\ &= \sum_{x \in \Omega_x} \mathbb{P}(X = x, Y = z - x) \\ &= \sum_{x \in \Omega_x} \mathbb{P}(X = x)\,\mathbb{P}(Y = z - x) \\ &= \sum_{x \in \Omega_x} p_X(x)p_Y(z - x) \end{aligned}$$

Examples

Now let's consider the continuous case. What if X and Y are continuous RVs with respective probability and cummulative density functions f_X, F_X, f_Y and F_Y. If we define $Z = X + Y$, then how can we solve for the probability density function for Z, $f_Z(z)$? Well, we can work with the cumulative density functions to relate these to probabilities we can rewrite as follows:

$$\begin{aligned}
F_Z(z) &= \mathbb{P}(Z \leq z) \\
&= \mathbb{P}(X + Y \leq z) \\
&= \int_{-\infty}^{\infty} \mathbb{P}(X + Y \leq z \mid X = x) f_X(x) dx \\
&= \int_{-\infty}^{\infty} \mathbb{P}(x + Y \leq z \mid X = x) f_X(x) dx \\
&= \int_{-\infty}^{\infty} \mathbb{P}(Y \leq z - x) f_X(x) dx \\
&= \int_{-\infty}^{\infty} F_Y(z - x) f_X(x) dx
\end{aligned}$$

Then, we can solve for the probability density function by differentiating:

$$\begin{aligned}
f_Z(z) &= \frac{d}{dz} F_Z(z) \\
&= \int_{-\infty}^{\infty} f_Y(z - x) f_X(x) dx
\end{aligned}$$

5.5.2 Convolution

Convolution is a mathematical operation that allows to derive the distribution of a sum of two independent random variables.

For example, suppose the amount of gold a company can mine is X tons per year in country A, and the amount of gold the company can mine is Y tons per year in country B, independently. You have some distribution to model each. What is the distribution of the total amount of gold you mine, $Z = X + Y$?

Definition 5.5.2.1: Convolution

Let X, Y be independent random variables, and $Z = X + Y$.
Discrete version: If X, Y are discrete:

$$p_Z(z) = \sum_{x \in \Omega_X} p_X(x) p_Y(z - x)$$

Continuous version: If X, Y are continuous:

$$F_Z(z) = \int_{-\infty}^{\infty} f_X(x) F_Y(z - x) dx$$

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x) dx$$

Note: You can swap the roles of X and Y. Note the similarity between the cases!

Proof of Convolution.:

- Discrete case:

$$\begin{aligned}
p_Z(z) &= \mathbb{P}(Z = z) \\
&= \sum_{x \in \Omega_X} \mathbb{P}(X = x, Z = z) && \text{[LTP/marginal]} \\
&= \sum_{x \in \Omega_X} \mathbb{P}(X = x, Y = z - x) && [(X = x, Z = z) \text{ equivalent to } (X = x, Y = z - x)] \\
&= \sum_{x \in \Omega_X} \mathbb{P}(X = x) \mathbb{P}(Y = z - x) && [X \text{ and } Y \text{ are independent}] \\
&= \sum_{x \in \Omega_X} p_X(x) p_Y(z - x)
\end{aligned}$$

- Continuous case:

$$\begin{aligned}
F_Z(z) &= \mathbb{P}(Z \le z) \\
&= \mathbb{P}(X + Y \le z) && [\text{def of } Z] \\
&= \int_{-\infty}^{\infty} \mathbb{P}(X + Y \le z | X = x) f_X(x) dx) && \text{[LTP]} \\
&= \int_{-\infty}^{\infty} \mathbb{P}(x + Y \le z | X = x) f_X(x) dx) && [\text{Given } X = x] \\
&= \int_{-\infty}^{\infty} \mathbb{P}(Y \le z - x | X = x) f_X(x) dx) && \text{[algebra]} \\
&= \int_{-\infty}^{\infty} \mathbb{P}(Y \le z - x) f_X(x) dx) && [X \text{ and } Y \text{ are independent}] \\
&= \int_{-\infty}^{\infty} F_Y(z - x) f_X(x) dx && \text{[def of CDF]}
\end{aligned}$$

Now we can take the derivative to get the density with respect to z to get the pdf:

$$f_Z(z) = \frac{d}{dz} F_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z - x) dx$$

□

Examples

Suppose X and Y are two independent random variables such that $X \sim \text{Poi}(\lambda_1)$ and $Y \sim \text{Poi}(\lambda_2)$, and let $Z = X + Y$. Prove that $Z \sim \text{Poi}(\lambda_1 + \lambda_2)$.

The range of X, Y are $\Omega_X = \Omega_Y = \{0, 1, 2, \dots\}$, and so $\Omega_Z = \{0, 1, 2, \dots\}$ as well. For $n \in \Omega_Z$:

$$\begin{aligned}
p_Z(n) &= \sum_{k=0}^{n} p_X(k) p_Y(n-k) && \text{[convolution formula]} \\
&= \sum_{k=0}^{n} e^{-\lambda_1} \frac{\lambda_1^k}{k!} \cdot e^{-\lambda_2} \frac{\lambda_2^{n-k}}{(n-k)!} && \text{[plug in Poisson PMFs]} \\
&= e^{-(\lambda_1+\lambda_2)} \sum_{k=0}^{n} \frac{1}{k!(n-k)!} \lambda_1^k (1-\lambda_2)^{n-k} \\
&= e^{-(\lambda_1+\lambda_2)} \frac{1}{n!} \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} \lambda_1^k (1-\lambda_2)^{n-k} && \text{[multiply and divide by } n!] \\
&= e^{-(\lambda_1+\lambda_2)} \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} \lambda_1^k (1-\lambda_2)^{n-k} && \left[\binom{n}{k} = \frac{n!}{k!(n-k)!}\right] \\
&= e^{-(\lambda_1+\lambda_2)} \frac{(\lambda_1+\lambda_2)^n}{n!} && \text{[binomial theorem]}
\end{aligned}$$

Thus, $Z \sim \text{Poi}(\lambda_1 + \lambda_2)$, as its PMF matches that of a Poisson distribution.

Examples

Suppose X, Y are independent and identically distributed (iid) continuous $\text{Unif}(0, 1)$ random variables. Let $Z = X + Y$. What is $f_Z(z)$?

We always begin by calculating the range: we have $\Omega_Z = [0, 2]$.

For a $U \sim Unif(0, 1)$ random variable, we know $\Omega_U = [0, 1]$, and that

$$f_U(u) = \begin{cases} 1 & 0 \le u \le 1 \\ 0 & \text{otherwise} \end{cases}$$

The convolution formula tells us that

$$f_Z(z) = \int_0^1 f_X(x) f_Y(z-x) dx = \int_0^1 f_Y(z-x) dx$$

where the second formula holds since $f_X(x) = 1$ for all $0 \le x \le 1$ as we saw above.

For $f_Y(z-x) > 0$, we need $0 \le z - x \le 1$ (otherwise it will be $= 0$). We'll split into two cases depending on whether $z \in [0, 1]$ or $z \in [0, 2]$, which compose its range $\Omega_Z = [0, 2]$.

- If $z \in [0, 1]$, we already have $z - x \le 1$ since $z \le 1$. We also need $z - x \ge 0$ for the density to be nonzero: $x \le z$. Hence, our integral becomes:

$$\begin{aligned}
f_Z(z) &= \int_0^z f_Y(z-x) dx + \int_z^1 f_Y(z-x) dx \\
&= \int_0^z 1 dx + 0 = [x]_0^z = z
\end{aligned}$$

- If $z \in [1, 2]$, we already have $z - x \geq 0$ since $z \geq 1$. We now need the other condition $z - x \leq 1$ for the density to be nonzero: $x \geq z - 1$. Hence, our integral becomes:

$$\begin{aligned} f_Z(z) &= \int_0^{z-1} f_Y(z-x)dx + \int_{z-1}^1 f_Y(z-x)dx \\ &= 0 + \int_{z-1}^1 1dx = [x]_{z-1}^1 = 2 - z \end{aligned}$$

Thus, putting these two cases together gives:

$$f_Z(z) = \begin{cases} z & 0 \leq z \leq 1 \\ 2 - z & 1 \leq z \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

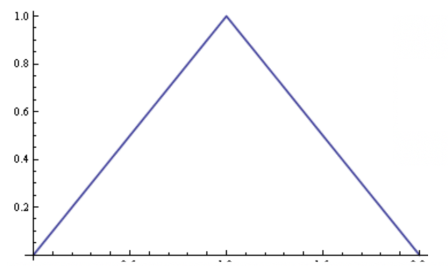

Figure 5.5.1: Triangular distribution for Z

This makes sense because there are "more ways" to get a value of 1 for example than any other point. Whereas to get a value of 2, there's only one way - we need both X, Y to be equal to 1.