SUMS OF RANDOM VARIABLES: A KEY PGF RESULT

Suppose that X_1 and X_2 are discrete random variables with ranges $\mathbb{X}_1$ and $\mathbb{X}_2$, probability mass functions f_{X_1} and f_{X_2} and probability generating functions G_{X_1} and G_{X_2} respectively. Suppose also that X_1 and X_2 are independent. Define discrete random variable Y by

$$Y = X_1 + X_2$$

Using the Theorem of Total Probability, by partitioning on the different possible values of X_1, for y in an appropriate range $\mathbf{Y}$

$$\begin{aligned} f_Y(y) = \mathrm{P}[\ Y = y\ ] &= \sum_x \mathrm{P}[\ (X_1 = x) \cap (X_2 = y - x)\ ] \\ &= \sum_x \mathrm{P}[\ X_1 = x\ ]\mathrm{P}[\ X_2 = y - x\ ] \\ &= \sum_x f_{X_1}(x) f_{X_2}(y - x) \end{aligned}$$

where the second line follows from the independence of X_1 and X_2, and all summations are over $x \in \mathbb{X}_1$. Hence we now have an expression for the pmf for the new variable Y.

Now,

$$\begin{aligned} G_Y(t) &= \sum_y f_Y(y) t^y = \sum_y \left\{ \sum_x f_{X_1}(x) f_{X_2}(y - x) \right\} t^y \\ &= \sum_{x_1} \sum_{x_2} f_{X_1}(x_1) f_{X_2}(x_2) t^{x_1 + x_2} \qquad \text{changing variables to } x_1 = x \quad x_2 = y - x \\ &= \left\{ \sum_{x_1} f_{X_1}(x_1) t^{x_1} \right\} \left\{ \sum_{x_2} f_{X_2}(x_2) t^{x_2} \right\} \\ &= G_{X_1}(t)\ G_{X_2}(t) \end{aligned}$$

so therefore

$$G_Y(t) = G_{X_1}(t)\ G_{X_2}(t)$$

EXTENSION

If $X_1, X_2, ..., X_n$ are independent discrete random variables with pgfs $G_{X_1}, G_{X_2}, ..., G_{X_n}$ .respectively, and discrete random variable Y is defined by

$$Y = \sum_{i=1}^{n} X_i$$

then by induction on n using the above result we have that

$$G_Y(t) = \prod_{i=1}^{n} G_{X_i}(t)$$

so that if $X_1, X_2, ..., X_n$ are also identically distributed with pgf G_X then

$$G_Y(t) = \{G_X(t)\}^n$$

(see, for example, the Bernoulli/Binomial pgfs or the Geometric/Negative Binomial pgfs)