The probability integral transform

Jacob Fine

Dec 10th, 2024

A general property of any probability density function (PDF) is that, using its cumulative density function (CDF), we can convert each value of the random variable into the value of a standard uniform distribution. Another way of phrasing this is that for any PDF, if we think of the values output by its CDF as a random variable (the ranks), then that random variable always follows a standard uniform distribution.

Formally, for some CDF

$$P(X \leq x) = F_X(x)$$

for an original random variable X, where r is the rank associated with $X = x$, the distribution of $P(R \leq r)$ must be a standard uniform distribution. That is, the ranks follow a standard uniform distribution for any PDF, i.e.,

$$R \sim U(0, 1)$$

To show this, let us consider the inverse CDF, often called the 'quantile function', which gives the lowest real value $x \in \mathbb{R}$ such that $r \leq F_X(x)$. Simply put, the quantile function outputs the value of x from the original distribution of X associated with the rank of interest r. So, while the CDF outputs the rank of each x, the inverse CDF outputs the x (called a quantile) of each rank.

$$F_X^{-1}(r) = \inf \{x \in \mathbb{R} : r \leq F_X(x)\}$$

Consider the CDF of R, $F_R(r)$. We have

$$F_R(r) = P(R \leq r)$$

$$\implies F_R(r) = P(F_X(x) \leq r)$$

$$\implies F_R(r) = P(X \leq F_X^{-1}(r))$$

This is just the CDF of X bounded by $F_X^{-1}(r)$

$$\implies F_R(r) = F_X(F_X^{-1}(r))$$

$$\implies F_R(r) = r$$

This is the definition of the standard uniform distribution, i.e., the value of its CDF is just the current value of the random varaible. Therefore

$$R \sim U(0, 1)$$

This result is especially useful for generating samples to 'match' the distribution of any random variable, through 'inverse integral transform sampling'. Using this method, we can take this standard uniform distribution of ranks 'feed it' into the quantile function of our choice to output matching quantiles.