4.1.1 Functions of Two Random Variables

Ex: Let X and Y be both uniformly distributed in $[0, 1]$ and independent. Let $Z = XY$. Find the PDF of Z.

Ex: Let X and Y be two independent discrete random variables. Express the PMF of $Z = X + Y$ in terms of the PMFs $p_X(x)$ and $p_Y(y)$ of X and Y. Do you recognize this expression?

Ex: Let X and Y be two independent continuous random variables. Show that, similarly to the discrete case, the PDF of $Z = X + Y$ is given by the "convolution" of the PDFs $f_X(x)$ and $f_Y(y)$ of X and Y.

Ex: Let X_1, X_2, X_3, $\ldots$, be a sequence of IID (independent, identically distributed) random variables, whose distribution is uniform in $[0, 1]$. Using convolution, compute and sketch the PDF of $X_1 + X_2$. As exercise, also compute and sketch the PDF of $X_1 + X_2 + \ldots + X_n$ for $n = 3, 4$, and observe the trend. As we add more and more random variables, the pdf of the sum is getting smoother and smoother. It turns out that, in the limit the shape of the density around the center will be converge to the Gaussian PDF. (It turns out that the Gaussian PDF is a fixed point for convolution: convolving two Gaussian PDFs results in another Gaussian PDF.)

4.2 Covariance and Correlation

Covariance of two random variables, X and Y, is defined as:

$$cov(X,Y) = E[(X - E(X))(Y - E(Y))]$$

It is a quantitative measure of the relationship between the two random variables: the magnitude reflects the strength of the relationship, the sign conveys the direction of the relationship. When $cov(X,Y) = 0$, the two random variables are said to be "uncorrelated". A positive correlation implies, roughly speaking, that they tend to increase or decrease together. A negative correlation, on the other hand, implies that when X increases, Y tends to decrease, and vice versa.

Ex: Show that $cov(X,Y) = E(XY) - E(X)E(Y)$ (as exercise.)

It follows from the alternative definition proven in the above exercise that independence implies uncorrelatedness.

Ex: Exhibit a counterexample showing that uncorrelatedness does not necessarily imply independence. (Consider X and Y uniformly distributed in an area shaped like a rhombus.)

The correlation coefficient is a normalized form of covariance:

$$\rho(X,Y) = \frac{cov(X,Y)}{\sqrt{var(X)var(Y)}}$$

What is the possible range of values for $\rho(X,Y)$? The positive and negative extremes of this range correspond to full positive and full negative correlation, respectively. They are only attained when $X = aY$, where a

is a positive, or negative scalar, respectively.

Ex: Show that, for any two random variables X and Y

$$var(X+Y) = varX + varY + 2cov(X,Y)$$