Jointly distributed Random Variables Lecture 2

Independent Random Variables

X and Y are independent if $P(X \in A, Y \in B) =$
$= P(X \in A)\,P(Y \in B)$.

It can be shown that X and Y are independent iff

$$F(a,b) = P(X \leq a, Y \leq b) = P(X \leq a)\,P(Y \leq b)$$
$$= F_X(a)\,F_Y(b).$$

When X and Y are discrete, the condition of independence is equivalent to $p(x,y) = p_X(x)\,p_Y(y)$.

One direction is obvious. For the other direction, observe that
$$P(X \in A, Y \in B) = \sum_{y \in B} \sum_{x \in A} p(x,y)$$
$$= \sum_{y \in B} \sum_{x \in A} p_X(x)\,p_Y(y) = \sum_{y \in B} p_Y(y) \sum_{x \in A} p_X(x)$$
$$= P(Y \in B)\,P(X \in A).$$

In the jointly continuous case, the condition of independence is equivalent to $f(x,y) = f_X(x) f_Y(y)$.

Ex. $n+m$ independent trials are performed. $P(\text{success}) = p$. Let $X = \#$ of successes in the first n trials and $Y = \#$ of successes in the last m trials.

Then $P(X=k, Y=j) = \binom{n}{k} p^k (1-p)^{n-k} \binom{m}{j} p^j (1-p)^{m-j}$

$= P(X=k) P(Y=j)$

Independence of several random variables $X_1, X_2, \ldots, X_n$ may be defined similarly:

Def: $X_1, X_2, \ldots, X_n$ are said to be independent if, for all sets of real numbers $A_1, A_2, \ldots, A_n$

$$P(X_1 \in A_1, X_2 \in A_2, \ldots, X_n \in A_n) = \prod_{k=1}^{n} P(X_k \in A_k).$$

Ex. Each Calculus I student joins the video lecture independently with a beta distribution $\sim B(3,1)$

If there are 12 students (the rest don't bother to show up at all), compute

(a) The probability that 10 or more join in within the first half of the lesson.

(b) The students log on in alphabetic order.

Solution:

(a) Let $X_1, X_2, \ldots, X_{12}$ be the times when each student joins the lesson, where the index $1 \le k \le 12$ lists students in alphabetic order.

The distribution for each X_k is given by

$$f_k(x) = f(x) = 3x^2 \quad 0 \le x \le 1$$

Then P(student k joins before the half mark)

$$= P\left(X_k \le \tfrac{1}{2}\right) = \int_0^{\frac{1}{2}} 3x^2\,dx = x^3 \Big|_0^{\frac{1}{2}} = \frac{1}{8}$$

The likelihood that 10 or more join in before half of the lesson is over is

$$\sum_{k=10}^{12} \binom{12}{k} \left(\frac{1}{8}\right)^k \left(\frac{7}{8}\right)^{12-k} \approx 4.8 \times 10^{-8}$$

or essentially never!

(b) $P(X_1 < X_2 < X_3 < \ldots < X_{12}) =$

$$= \int_0^1 \int_{x_1}^1 \int_{x_2}^1 \cdots \int_{x_{11}}^1 3^{12} x_1^2 x_2^2 x_3^2 \ldots x_{12}^2 \, dx_{12} \, dx_{11} \ldots dx_1$$

It is easier to note by symmetry that all orders are equally likely.

Thus $P(X_1 < X_2 < X_3 < \ldots < X_{12}) = \frac{1}{12!}$

<u>Ex.</u> A man and a woman decide to meet at a certain location. If each of them independently arrives at a time uniformly distributed between 12 noon and 1 P.M. Find the probability that the first to arrive has to wait longer than 10 minutes.

<u>Solution:</u> Let X and Y be the times of arrival for the woman and the man respectively. Clearly X and Y are uniformly distributed over $(0, 60)$.

We wish to calculate $P(X+10 < Y \text{ or } Y+10 < X)$

$$= P(X+10 < Y) + P(Y+10 < X)$$

By symmetry, this is $2P(Y+10 < X) =$

$$= 2 \iint\limits_{Y+10<X} f(x,y)\,dy\,dx = 2 \iint\limits_{Y+10<X} f_X(x) f_Y(y)\,dy\,dx$$

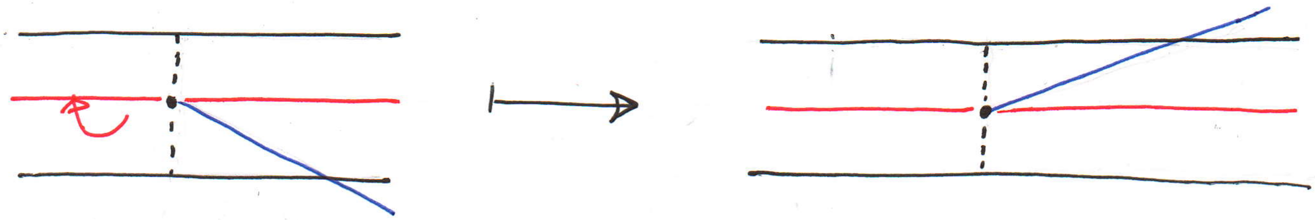
$$= 2 \int_{10}^{60} \int_{0}^{x-10} \left(\frac{1}{60}\right)^2 dy\,dx = \frac{2}{60^2} \int_{10}^{60} (x-10)\,dx$$

$$= \frac{25}{36}.$$

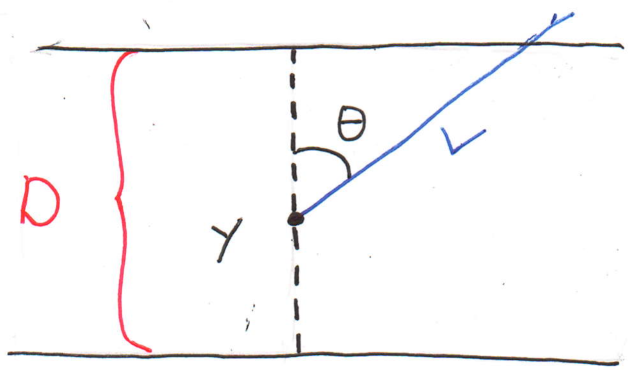
<u>Ex.</u> (Buffon's needle problem) A table is ruled with equidistant parallel lines a distance D apart. A needle of length L, where $L \le D$, is randomly thrown on the table. What is the probability that the needle will intersect one of the lines?

<u>Solution:</u>

When needle is dropped draw a vertical line through the leftmost point on the needle. If the slope of the needle is negative, reflect through a horizontal line.



Thus, without loss of generality we may assume that the leftmost point of the needle falls on Y uniformly distributed over $[0, D]$ and makes an angle θ uniformly distributed over $[0, \frac{\pi}{2}]$:



The joint density is $f(\theta, Y) = \frac{2}{\pi D}$; $0 \leq \theta \leq \frac{\pi}{2}$, $0 \leq Y \leq D$.

The probability that the needle intersects the line is thus $P(Y + L\cos\theta \geq D) = P(Y \geq D - L\cos\theta)$

$$= \left(\int_0^{\frac{\pi}{2}} \int_{D - L\cos\theta}^{D} dy\, d\theta \right) \frac{2}{\pi D} = \frac{2}{\pi D}$$

$$= \frac{2}{\pi D} \int_0^{\frac{\pi}{2}} L \cos\theta \, d\theta = \frac{2L}{\pi D}.$$

<u>Ex.</u> Let X, Y, Z be independent and uniformly distributed over $(0,1)$. Compute $P(X \geq YZ)$.

<u>Solution:</u>

Clearly $f(x,y,z) = f_X(x) f_Y(y) f_Z(z) = 1$ for $x, y, z \in [0,1]$.

Thus $P(X \geq YZ) = \int_0^1 \int_0^1 \int_{yz}^1 dx\,dy\,dz =$

$$= \int_0^1 \int_0^1 (1 - yz)\,dy\,dz = 1 - \int_0^1 \int_0^1 yz\,dy\,dz = 1 - \left(\int_0^1 y\,dy \right)^2$$

$$= 1 - \left(\frac{1}{2}\right)^2 = \frac{3}{4}.$$