Lecture 26.

Proposition: Random vars X and Y are independent iff $f_{XY}(x,y) = g(x)\,h(y)$ $\quad -\infty < x < \infty,\ -\infty < y < \infty$ for some functions h and g. Here (f_{XY} is either the prob. mass function or prob. density function).

Proof (continuous case): We already know that X and Y are indep iff $f_{XY} = f_X f_Y$, where f_X and f_Y are the marginal probabilities of X and Y resp. Therefore, we only need to show that any factorization $f_{XY}(x,y) = g(x)h(y)$ works. Now suppose $f_{XY} = g(x)h(y)$. Then

$$1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)\,dx\,dy = \underbrace{\left(\int_{-\infty}^{\infty} g(x)\,dx\right)}_{c_1}\underbrace{\left(\int_{-\infty}^{\infty} h(y)\,dy\right)}_{c_2}$$

Also
$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy = c_2\, g(x)$$

and
$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dx = c_1\, h(y)$$

But then

$$f_{XY}(x,y) = g(x)\,h(y) = c_1 c_2\, g(x)\,h(y)$$
$$= (c_2 g(x))(c_1 h(y)) = f_X(x)\,f_Y(y)$$

So X and Y are independent.

Ex: Suppose X, Y have a joint density given by

$$f(x,y) = \begin{cases} 6e^{-2x-3y} & , \ 0<x<\infty, \ 0<y<\infty \\ 0 & \text{otherwise.} \end{cases}$$

Are X and Y independent?

Yes, since $6e^{-2x-3y} = (6e^{-2x})(e^{-3y})$

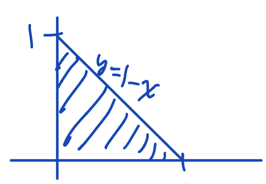
and so $f(x,y) = g(x)\,h(y)$

$$g(x) = \begin{cases} 6e^{-2x} & 0<x<\infty \\ 0 & \text{otherwise} \end{cases} \qquad h(x) = \begin{cases} e^{-3y} & 0<y<\infty \\ 0 & \text{otherwise.} \end{cases}$$

Ex: Suppose that X, Y have joint density given by

$$f(x,y) = \begin{cases} 24xy & 0<x<1, \ 0<y<1, \ 0<x+y<1 \\ 0 & \text{otherwise.} \end{cases}$$

So $f(x,y)$ is non zero on the domain



Are X and Y indep? **NO:** why?

Set $\quad I(x,y) = \begin{cases} 1 & \text{if } 0<x<1,\ 0<y<1,\ 0<x+y<1. \\ 0 & \text{otherwise.} \end{cases}$

Then $\quad f(x,y) = 24xy\,I(x,y)$. But $I(x,y)$ does not factor as a product $g(x)h(y)$ on $-\infty < x, y < \infty$, so neither does $f(x,y)$.

* i.e. If X, Y are independent, then $f_{X,Y}(x,y)$ is non-zero on a rectangular domain!

↳ the converse is <u>not</u> true!

Independence of many variables:

A set of random vars $X_1, \dots, X_n$ is independent, iff, for any sets $A_1, \dots, A_n \subseteq \mathbb{R}$,

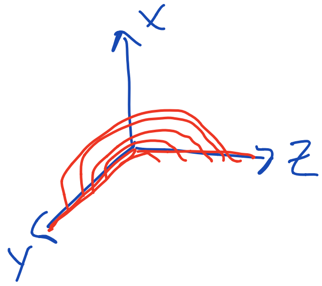
$$P(X_1 \in A_1, X_2 \in A_2, \dots, X_n \in A_n) = \prod_{i=1}^{n} P(X_i \in A_i).$$

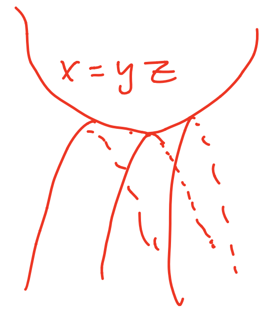
Equivalently, $X_1, \dots, X_n$ are independent iff for all $a_1, \dots, a_n$,

$$P(X_1 \le a_1, \dots, X_n \le a_n) = \prod_{i=1}^{n} P(X_i \le a_i).$$

More generally, given <u>any</u> collection $\{X_i : i \in I\}$, I an arbitrary index set, we say the collection is independent iff for any finite subset $J \subseteq I$, $\{X_i : i \in J\}$ is independent

Ex: Suppose X, Y, Z are independent RV's, each being uniformly distributed over $(0,1)$. What is $P(X \geq YZ)$?





By independence:

$$f_{XYZ}(x,y,z) = f_X(x) f_Y(y) f_Z(z) = \begin{cases} 1 & 0 \leq x, y, z \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

So $$P(X \geq YZ) = \iiint_{\{x \geq yz\}} f_{XYZ} \, dx\,dy\,dz$$

$$= \int_0^1 \int_0^1 \int_{yz}^1 dx\,dy\,dz = \int_0^1 \int_0^1 (1 - yz)\,dy\,dz$$

$$= \int_0^1 y - \frac{y^2 z}{2} \Big|_0^1 dz = \int_0^1 1 - \frac{z}{2}\,dz = z - \frac{z^2}{4}\Big|_0^1$$

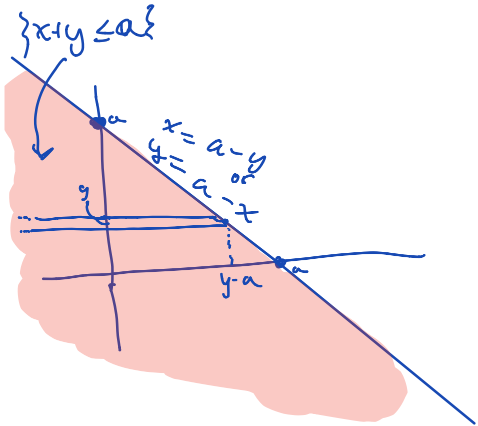
$$= 1 - \frac{1}{4} = \frac{3}{4}.$$

Sums of independent RV's.

- it is often useful to be able to determine the distribution of a sum of RV's.
- suppose that X and Y are independent continuous RVs with pdf's f_X and f_Y respectively.

Then:

$$F_{X+Y}(a) = P(X+Y \le a) = \iint_{x+y\le a} \underbrace{f_X(x)\, f_Y(y)}_{=f_{XY}(x,y)}\, dx\, dy \quad \text{(this is by independence)}$$



$$= \int_{-\infty}^{\infty} \int_{-\infty}^{a-y} f_X(x)\, f_Y(y)\, dx\, dy.$$

$$= \int_{-\infty}^{\infty} f_Y(y) \int_{-\infty}^{a-y} f_X(x)\, dx\, dy$$

$$= \underbrace{\int_{-\infty}^{\infty} f_Y(y)\, F_X(a-y)\, dy}_{\text{called the convolution of } F_X \text{ and } F_Y}.$$

called the convolution of F_X and F_Y.

To get the corresponding density, we differentiate:

$$\frac{d}{da}\left(F_{X+Y}(a)\right) = \frac{d}{da} \int_{-\infty}^{\infty} f_Y(y)\, F_X(a-y)\, dy$$

$$= \int_{-\infty}^{\infty} f_Y(y)\, \frac{d}{da} F_X(a-y)\, dy$$

this is a rule called "Leibniz's rule"

$$= \int_{-\infty}^{\infty} f_Y(y)\, f_X(a-y)\, dy$$

Rmk: In general the convolution of two functions $f(x)$ and $g(y)$ is defined

$$(f * g)(z) = \int_{-\infty}^{\infty} f(z-y)\, g(y)\, dy = \int_{-\infty}^{\infty} g(z-x) f(x)\, dx.$$

Ex: Sum of two uniform RV's on $(0,1)$. X, Y indep, uniform on $(0,1)$, $Z = X+Y$.

Then

$$f_X = f_Y = \begin{cases} 1 & \text{on } (0,1) \\ 0 & \text{otherwise.} \end{cases}$$

So $$f_Z = \int_{-\infty}^{\infty} f_X(z-y) f_Y(y)\,dy$$

$$= \int_0^1 f_X(z-y)\,dy$$

Now, $$f_X(z-y) = \begin{cases} 1 & \text{for } 0 < z-y < 1 \\ 0 & \text{otherwise.} \end{cases}$$

So if $0 < z \leq 1$ (so $0 < z-y < 1$ iff $0 < y < z$)

$$f_Z(z) = \int_0^z dy = z.$$

If $1 < z < 2$, (so $0 < z-y < 1$ iff $z-1 < y <$)

then $$f_Z(z) = \int_{z-1}^1 dy = 1 - (z-1) = 2 - z.$$

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