Lecture 19: Joint, Conditional and Marginal Distributions

Joint CDF: $F(x,y) = P(X \leq x, Y \leq y)$

Continuous case: Joint PDF ⟶ not a prob., that's a density, we integrate

$$f(x,y) = \frac{\partial^2}{\partial x \partial y} F(x,y)$$

In uniform case, prob. is proportional to area.

$$P((X,Y) \in A) = \iint_A f(x,y)\,dx\,dy$$

↳ could be any region in a plane

– Joint PDF is what we integrate to get the probability of (X,Y) being in any particular set.

Marginal PDF of X: $\displaystyle\int_{-\infty}^{\infty} f(x,y)\,dy$ (we marginalized out the y, marginalization)

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy = 1$$

Conditional PDF of $Y|X$ is

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{\text{Joint density}}{\text{marginal density of } X}$$

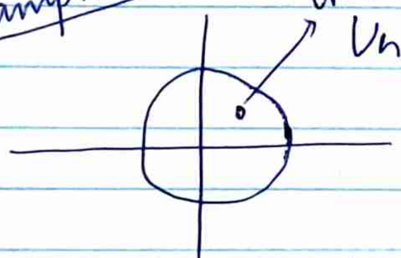
$$= \frac{f_{X|Y}(x|y)\, f_Y(y)}{f_X(x)} \quad \left(\begin{array}{c}\text{Analogous to}\\ \text{Bayes' rule}\end{array}\right)$$

X, Y independent if $f_{X,Y}(x,y) = f_X(x)\, f_Y(y)$ for all x, y.

Example

picking a uniformly random point

Unif in disk $x^2+y^2 \leq 1$



$$f(x,y) = \begin{cases} 1/\pi, & x^2+y^2 \leq 1 \\ 0, & \text{outside} \end{cases}$$

Uniform means that probability of some region is proportional to area.

When we have problems that involve a uniform distribution on some region in the plane, we can actually think of probability in terms of area, or atleast it's proportional to area.

X and Y are dependent here.

Marginal pdf

$$f_X(x) = \int_{-\sqrt{1-x^2}}^{\sqrt{1+x^2}} \frac{1}{\pi}\, dy \qquad x^2+y^2 \leq 1, \quad y^2 \leq 1-x^2$$

$$= \frac{2\sqrt{1-x^2}}{\pi}, \quad -1 \leq x \leq 1$$

not uniform

$$f_{Y|X}(y|x) = \frac{1/\pi}{2/\pi\sqrt{1-x^2}} = \frac{1}{2\sqrt{1-x^2}}, \text{ if } -\sqrt{1-x^2} \leq y \leq \sqrt{1-x^2}$$

joint pdf

marginal pdf

we are treating X as a constant here.

and we are explicitly calling that constant x

[which means $Y|X=x \sim \text{Unif}(-\sqrt{1-x^2}, \sqrt{1-x^2})$]

$Y|X \sim \text{Unif}(-\sqrt{1-X^2}, \sqrt{1-X^2})$

given that we get to know what X is, here's the distribution of Y

Since $f_{Y|X}(y|x)$ is a constant for each fixed x.

⇐ X and Y are not independent, since
$$f_{X,Y}(x,y) \neq f_X(x) f_Y(y)$$

⇐ Another way to say X and Y are not independent is that the conditional distribution of Y given X is not same as the unconditional distribution of Y, i.e., learning X gives us information.

2-D LOTUS (Continuous case) [Discrete case LOTUS is also possible]
Let (X,Y) have joint PDF $f(x,y)$, and let $g(x,y)$ be a real-valued function of x, y. Then,
$$E g(X,Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f(x,y)\, dx\, dy$$

Theorem: If X, Y are independent, then $E(XY) = E(X)E(Y)$
("Independence implies uncorrelated.")

Proof (Continuous case)

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

Since X, Y are independent, the joint PDF is just the product of the marginal PDFs

(from 2-D LOTUS)

$$= \int_{-\infty}^{\infty} y f_Y(y) \underbrace{\int_{-\infty}^{\infty} x f_X(x)\, dx}_{\text{Constant} = E(X)}\, dy$$

$$= E(X) \underbrace{\int_{-\infty}^{\infty} y f_Y(y)\, dy}_{E(Y)}$$

$$= E(X)\, E(Y)$$

Example $X, Y \overset{i.i.d.}{\sim} \text{Unif}(0,1)$, find $E|X-Y|$.

LOTUS: $\displaystyle\int_0^1\int_0^1 |X-Y|\,dx\,dy$ [Since, they are i.i.d. Uniformed, the PDF is just 1.]

$$= \iint_{x>y} (x-y)\,dx\,dy + \iint_{x\leq y} (y-x)\,dx\,dy$$

$$= 2\iint_{x>y} (x-y)\,dx\,dy$$

[problem is completely symmetrical because ~~of the i.i.d.~~ X and Y are i.i.ds and $|X-Y|$ is a symmetrical function.]

$$= 2\int_0^1\int_y^1 (x-y)\,dx\,dy$$

$$= 2\int_0^1 \left(\frac{x^2}{2} - yx\right)\Big|_y^1 dy$$

$$= 2\int_0^1 \left(\frac{1}{2} - y - \frac{y^2}{2} + y^2\right) dy$$

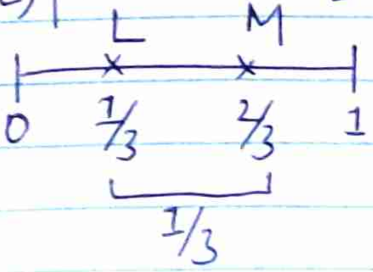
$$= 2\int_0^1 \left(y^2/2 - y + 1/2\right) dy = 2\left[\frac{y^3}{6} - \frac{y^2}{2} + \frac{1}{2}y\right]_0^1$$

$$= 2\left(\frac{1}{6} - \cancel{1/2} + \cancel{1/2}\right) = 1/3$$

So, the average distance between two uniforms is $1/3$.

Let $M = \max(X, Y)$.
$L = \min(X, Y)$
$|X-Y| = M - L$
$E(M-L) = 1/3$
$E(M) - E(L) = 1/3$

$X+Y$
$E(M+L) = E(M) + E(L)$ | $E(X) = E(Y) = \frac{1}{2}$
$= 1$



$E(M) = 2/3$
$E(L) = 1/3$

Example (discrete case) chicken-egg problem

~~10 eggs~~ $N \sim Pois(\lambda)$ eggs, each hatches with probability p independently.
Let X = # that hatch, so $X|N \sim Bin(N,p)$
Let Y = # that don't hatch.

pretend N is a known constant i.e., we know the no. of eggs (Actually N is Poisson but just pretend)

So, $\boxed{X+Y=N}$. Find joint PMF of X, Y. Are X and Y independent?

Soln: $P(X=i, Y=j) = \sum_{n=0}^{\infty} P(X=i, Y=j | N=n) P(N=n)$ (LOTP)

$P(X=3, Y=5 | N=10) = 0$

$P(X=3, Y=5 | N=2) = 0$

Redundant information

$= P(X=i, \cancel{Y=j} | N=i+j) P(N=i+j)$

If we know there is $(i+j)$ eggs and i hatched, then we already know that j hatched.

$= P(X=i | N=i+j) \underbrace{P(N=i+j)}_{\text{Poisson PMF}}$,

$= {}^{(i+j)}C_i \, p^i (1-p)^j \cdot \dfrac{e^{-\lambda} \lambda^{i+j}}{(i+j)!}$

$= \left(e^{-\lambda p} \dfrac{(\lambda p)^i}{i!} \right) \left(e^{-\lambda q} \dfrac{(\lambda q)^j}{j!} \right)$, $q = (1-p)$

$\Rightarrow X, Y$ are independent

$X \sim Pois(\lambda p)$

$Y \sim Pois(\lambda q)$

If ~~you~~ we change Poisson to anything else ($N \sim Pois(\lambda)$), X and Y become dependent.