Joint, Conditional, and Marginal PDFs

Lec. 23

July 30, 2020

Joint PDFs

Let X and Y be two continuous random variables. Then the joint density function $f_{X,Y} : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ satisfies:

$$\mathbb{P}(a \le X \le b, c \le Y \le d) = \int_a^b \int_c^d f_{X,Y}(x,y)dydx$$

and

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y)dydx = 1$$

and

$$f_{X,Y}(x,y) \ge 0 \qquad \forall x, y \in \mathbb{R}$$

Joint CDFs

(applies to continuous & discrete in general)

Let X and Y be two random variables. Then the joint cumulative distribution function $F_{X,Y} : \underbrace{\mathbb{R} \times \mathbb{R}} \to \mathbb{R}$ satisfies:

$$F_{X,Y}(x, y) = \underbrace{\mathbb{P}(X \leq x, Y \leq y)}$$

In the continuous case, assuming the cdf is continuous.

$$f_{X,Y}(x, y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x, y)$$

$$\frac{\partial^2}{\partial x \partial y} = \frac{\partial}{\partial x} \cdot \frac{\partial}{\partial y}$$

$\frac{\partial^2}{\partial y \partial x}$ is also fine.

Conditional PDFs

Just like $P(A|B) = \frac{P(A \cap B)}{P(B)}$

Let X and Y be two continuous random variables with joint density function $f_{X,Y}$.
For any y with $f_y(y) > 0$, the conditional distribution of X given $Y = y$ is defined as:

$$f_{x|y}(x|y) := \frac{f_{x,y}(x,y)}{f_y(y)}$$

When Y is continuous, even though $P(Y = y) = 0$, if $f_y(y) > 0$, then:

$$P(a \leq X \leq b | Y = y) = \int_a^b f_{x|y}(x|y)dx$$

Independence

Let X and Y be two continuous random variables. X and Y are independent if:

$$f_{x,y}(x,y) = f_x(x) f_y(y)$$

for all x, y.

Since $\underbrace{f_{x,y}(x,y) = f_{x|y}(x|y) f_y(y)}$, this implies $\underbrace{f_{x|y}(x|y) = f_x(x)}$.

Just like $\mathbb{P}(A \cap B) = \mathbb{P}(A|B) \cdot \mathbb{P}(B)$ in discrete setting

in order for X and Y to be independent

Marginalization

To recover the individual pdfs from the joint pdf:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dx$$

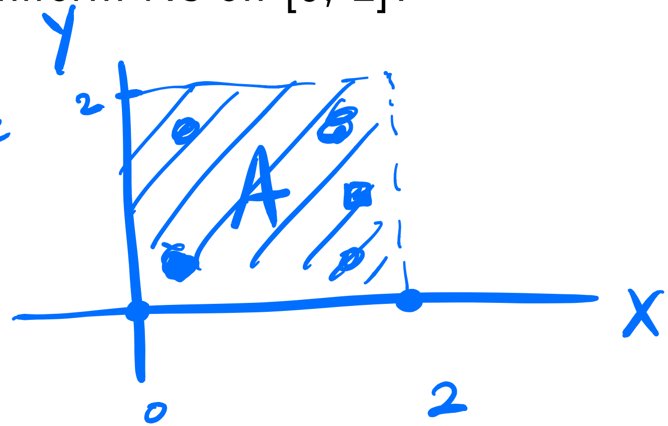
for $f_X(x)$
to be a valid pdf,
$\int_{-\infty}^{\infty} f_X(x)\,dx = 1$ ✓

$f_X(x)$ is a marginal pdf
of $f_{X,Y}(x,y)$

Example: $X, Y \overset{i.i.d}{\sim} Unif(0, 2)$

independent and identically distributed.

What is $f_{X,Y}(x, y)$ for two uniform rvs on [0, 2]?



Uniform on 2x2 square → Constant

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

$$\int_0^2 \int_0^2 c\, dy\, dx = 1$$

$$c \int_0^2 \int_0^2 1 \cdot dy\, dx = 1$$

$$c \cdot 4 = 1$$

$$c = \frac{1}{4}$$

$$f_{x,y}(x,y) = \begin{cases} \frac{1}{4} & \text{if } (x,y) \in A \\ 0 & \text{o.w.} \end{cases}$$

Example: $X, Y \overset{i.i.d}{\sim} Unif(0, 2)$

What is $f_{X,Y}(x, y)$ for two uniform rvs on [0, 2]?
We know that it must be a nonzero constant c on the two-by-two square, since all x, y pairs are equally likely. Again, we can use the constraint:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x, y)dydx = 1$$

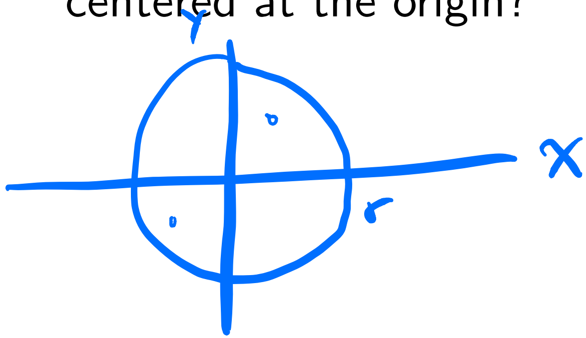
$$\int_{0}^{2}\int_{0}^{2} cdydx = 1$$

$$4 \cdot c = 1$$

$$\rightarrow c = \frac{1}{4}$$

Uniform Density Over a Disk: Joint

What is $f_{X,Y}(x,y)$ for a uniform density over a disk of radius r centered at the origin?



$$f_{X,Y}(x,y) = \begin{cases} c & \text{if } x^2+y^2 \le r^2 \\ 0 & \text{o.w} \end{cases}$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx\,dy = \iint_{x^2+y^2\le r^2} c\,dx\,dy$$

$$= c \iint_{x^2+y^2\le r^2} 1\,dx\,dy$$

$$= c \cdot (\text{area of disk})$$

$$= c \cdot \pi r^2$$

We know pdf must integrate to 1, so $c\pi r^2 = 1 \Rightarrow c = \frac{1}{\pi r^2}$

Uniform Density Over a Disk: Joint

What is $f_{X,Y}(x, y)$ for a uniform density over a disk of radius r centered at the origin?

$$f(x,y) = \begin{cases} c, & \text{if } x^2 + y^2 \le r^2 \\ 0, & \text{otherwise} \end{cases}$$

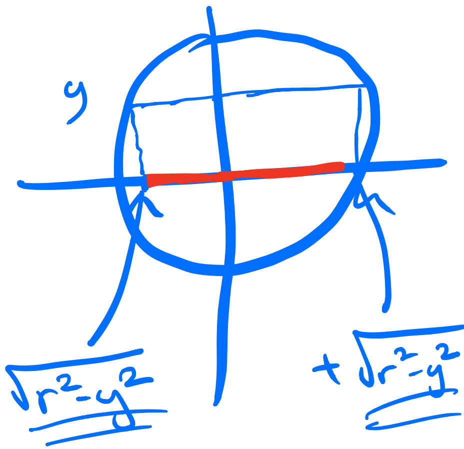
$$\begin{aligned} \int\int_{x^2+y^2\le r^2} c \cdot dx \cdot dy &= c \cdot \int\int_{x^2+y^2\le r^2} 1 \cdot dx \cdot dy \\ &= c \cdot [\text{area of disk}] \\ &= c \cdot \pi \cdot r^2 \end{aligned}$$

By definition, the joint must integrate to 1, so $1 = c \cdot \pi \cdot r^2 \Rightarrow c = \frac{1}{\pi \cdot r^2}$

Uniform Density Over a Disk: Marginals

$x^2+y^2 \le r^2$

What is $f_Y(y)$, and $f_X(x)$, now that we know the joint over the disk?



$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx$$

$$= \int_{-\sqrt{r^2-y^2}}^{\sqrt{r^2-y^2}} \frac{1}{\pi r^2}\,dx$$

$$= \frac{1}{\pi r^2}\int_{-\sqrt{r^2-y^2}}^{\sqrt{r^2-y^2}} 1\,dx$$

$$= \frac{2\cdot\sqrt{r^2-y^2}}{\pi r^2} \quad (\text{for } -r \le y \le r)$$
$$0 \quad \text{o.w.}$$

$$f_X(x) = \frac{2\cdot\sqrt{r^2-x^2}}{\pi r^2} \quad (\text{for } -r \le x \le r)$$
$$0 \quad \text{o.w.}$$

Uniform Density Over a Disk: Marginals

What is $f_y(y)$, and $f_x(x)$, now that we know the joint over the disk?

$$\begin{aligned} f_y(y) &= \int_{-\sqrt{r^2-y^2}}^{\sqrt{r^2-y^2}} \frac{1}{\pi \cdot r^2} \cdot dx \\ &= \frac{1}{\pi \cdot r^2} \int_{-\sqrt{r^2-y^2}}^{\sqrt{r^2-y^2}} \cdot dx \\ &= \frac{2\sqrt{r^2 - y^2}}{\pi \cdot r^2} \end{aligned}$$

for $-r \le y \le r$
By symmetry,

$$f_x(x) = \frac{2\sqrt{r^2 - y^2}}{\pi \cdot r^2}$$

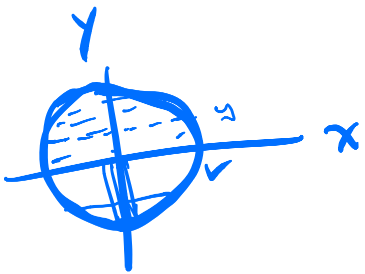
Uniform Density Over a Disk: Conditional PDFs

What is $f_{x|y}(x|y)$?

By def, $f_{X|Y}(x|y) = \dfrac{f_{X,Y}(x,y)}{f_Y(y)}$

$= \dfrac{\frac{1}{\pi r^2}}{\frac{2\sqrt{r^2-y^2}}{\pi r^2}}$

$= \dfrac{1}{2\sqrt{r^2-y^2}}$ (for $x^2+y^2 \le r^2$, also $|y| < r$)



Note: $f_{X|Y}(x|y) \neq \dfrac{2\sqrt{r^2-x^2}}{\pi r^2} = f_X(x)$

$\Rightarrow$ X and Y are not independent

What about $\text{Cov}(X,Y)$?

$\mathbb{E}[X] = 0$

$\mathbb{E}[Y] = 0$ by symmetry

$\mathbb{E}[X|Y=y] = 0$

$\mathbb{E}[X \cdot Y | Y=y] = y \cdot \mathbb{E}[X|Y=y]$

$= y \cdot 0 = 0$

$\Rightarrow \mathbb{E}[X \cdot Y] = 0$

$\text{Cov}(X,Y) = \mathbb{E}[X \cdot Y] - \mathbb{E}[X]\mathbb{E}[Y]$

$= 0 - 0 = 0$

X, Y are uncorrelated, but are dependent

Note:

Independent $\Rightarrow \mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$

(converse is not true)

(the converse is not true) as shown in this example.) independent $\Rightarrow$ uncorrelated (but not the other way around)

Uniform Density Over a Disk: Conditional PDFs

What is $f_{x|y}(x|y)$? If $x^2 + y^2 \leq r^2$ and $|y| < r$,

$$f_{x|y}(x|y) = \frac{f_{x,y}(x,y)}{f_y(y)} = \frac{\frac{1}{\pi \cdot r^2}}{\frac{2\sqrt{r^2 - y^2}}{\pi \cdot r^2}} = \frac{1}{2\sqrt{r^2 - y^2}}$$

$\Updownarrow$

$f_{x|y}(x)$ $\quad$ $f_{x|y}(x,y)$

Since this differs from $f_x(x)$, we know X and Y are dependent. Furthermore,

$$\mathbb{E}[X] = \mathbb{E}[Y] = 0 \text{ by symmetry}$$
$$\mathbb{E}[X|Y = y] = 0 \text{ by symmetry}$$
$$\mathbb{E}[XY|Y = y] = y \cdot \mathbb{E}[X|Y = y] = y \cdot 0 = 0 \Rightarrow \mathbb{E}[XY] = 0$$
$$Cov(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = 0$$

X, Y are uncorrelated but dependent!!

2D LOTUS

2 min break

Law of the unconcious statistician.

Let X, Y be two random variables with joint PDF $f_{X,Y}(x, y)$, and let $g(x, y)$ be a real-valued function of x, y. Then

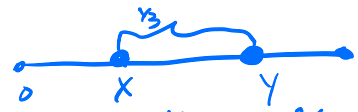
$$\mathbb{E}[g(X, Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x, y) f_{X,Y}(x, y)dxdy$$

use the known joint pdf of x and Y

You don't need to find the joint pdf of $g(X,Y)$ in order to calculate $\mathbb{E}[g(X,Y)]$

Expected Distance Between Two Points

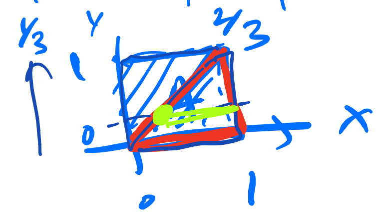
Let $X, Y \overset{i.i.d}{\sim} Unif(0,1)$. What is $\mathbb{E}[|X - Y|]$?



$\mathbb{E}[|X-Y|] = \int_0^1 \int_0^1 |x-y| \cdot f_{X,Y}(x,y)\,dx\,dy$

$= \int_0^1 \int_0^1 |x-y| \cdot 1 \; dx\,dy$

$= \underset{x>y}{\iint} (x-y)\cdot dx\,dy + \underset{y>x}{\iint} (y-x)\,dx\,dy$



$f_{XY}(x,y) = \begin{cases} 1 & \text{if } x,y \in A \\ 0 & \text{o.w.} \end{cases}$

$= 2 \cdot \underset{x>y}{\iint} (x-y)\,dx\,dy$

$\int_y^1 (x-y)\,dx$

$\frac{x^2}{2} - yx \Big|_y^1$

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

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

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

$= \frac{1}{3}$

Observe: $|X-Y| = \max(X,Y) - \min(X,Y)$

$M = \max(X,Y)$

$L = \min(X,Y)$

$\mathbb{E}[|X-Y|] = \mathbb{E}[M-L] = \frac{1}{3}$

$= \mathbb{E}[M] - \mathbb{E}[L] = \frac{1}{3}$

Also,

$\mathbb{E}[M+L] = \mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$

$\mathbb{E}[M] + \mathbb{E}[L] = \frac{1}{2} + \frac{1}{2} = 1$

solve the two equations:

$\mathbb{E}[M] = \frac{2}{3} \quad \mathbb{E}[L] = \frac{1}{3}$

Expected Distance Between Two Points

Let $X, Y \overset{i.i.d}{\sim} Unif(0,1)$. What is $\mathbb{E}[|X-Y|]$?

$$\begin{align}
\mathbb{E}[|X-Y|] &= \int_0^1 \int_0^1 |x-y| f_{X,Y}(x,y) dxdy \\
&= \int_0^1 \int_0^1 |x-y| \cdot 1 dxdy \\
&= \int \int_{X>Y} (x-y) dxdy + \int \int_{Y>X} (y-x) dxdy \\
&= 2 \cdot \int_0^1 \int_y^1 (x-y) dxdy \\
&= 2 \cdot \int_0^1 \left(\frac{x^2}{2} - yx\right)\Big|_{x=y}^1 dy \\
&= \frac{1}{3} \\
&
\end{align}$$

Total Probability Theorem

$$\mathbb{P}(A \cap B) = \mathbb{P}(A \mid B) \cdot \mathbb{P}(B)$$

Discrete Partition.

Continuous Partition

we know

Partition

dx

probability

	Dis.	Cont.
Dis.	$P(A) = \sum_1^n P(B_i) \cdot P(A \mid B_i)$	$\mathrm{P}(\mathrm{A}) = \int_{-\infty}^{\infty} f_x(x) \cdot P(A \mid X = x)d$
Cont.	$\mathrm{f}_x(x) = \sum_1^n P(B_i) \cdot f_{X\mid B_i}(x)$	$\mathrm{f}_x(x) = \int_{-\infty}^{\infty} f_y(y) \cdot f_{x\mid y}(x \mid y)dy$

we know (analogous)

$f_x(x) \rightarrow$ continuous

$\mathbb{P}(X = x) \rightarrow$ disc.

Total Probability Theorem Examples

Top right

X ~ cont.
Y discrete.

Discrete/Continuous:
$P(Y > X) = \int_{-\infty}^{\infty} f_x(x) \cdot P(Y > X | X = x) dx$

Continuous/Discrete: Bottom left.
Flip a fair coin. If its heads, then $X \sim Exp(\lambda_1)$, otherwise $X \sim Exp(\lambda_2)$. Then, for $x > 0$,

$$f_x(x) = \frac{1}{2} \cdot \lambda_1 e^{-\lambda_1 x} + \frac{1}{2} \cdot \lambda_2 e^{-\lambda_2 x}$$

Bayes' Rule

	Dis.	Cont.
Dis.	$P(A_i\|B) = \frac{P(A_i)\cdot P(B\|A_i)}{\sum_{j=1}^{n} P(A_j)\cdot P(B\|A_j)}$	$P(A_i\|X = x) = \frac{P(A_i)\cdot f_{X\|A_i}(x)}{\sum_j^n P(A_j)\cdot f_{X\|A_j}(x)}$
Cont.	$f_{X\|A}(x) = \frac{f_X(x)\cdot P(A\|X=x)}{\int_{-\infty}^{\infty} f_X(t)P(A\|X=t)dt}$	$f_{x\|y}(x\|y) = \frac{f_x(x)\cdot f_{y\|x}(y\|x)}{\int_{-\infty}^{\infty} f_X(t)f_{Y\|X}(Y\|t)dt}$

Discrete/Continuous: top right.

$$
\begin{aligned}
P(A|X = x) &= P(A|X \in [x, x+\delta]) \\
&= \frac{P(A)\cdot P(X \in [x, x+\delta]|A)}{P(A)\cdot P(X \in [x, x+\delta]|A) + P(A^c)\cdot P(X \in [x, x+\delta]|A^c} \\
&= \frac{P(A)\cdot f_{X|A}(x)\cdot \delta}{P(A)\cdot f_{X|A}(x)\cdot\delta + P(A^c)\cdot f_{X|A^c}(x)\cdot\delta} \\
&= \frac{P(A)\cdot f_{X|A}(x)}{P(A)\cdot f_{X|A}(x) + P(A^c)\cdot f_{X|A^c}(x)}
\end{aligned}
$$