LECTURE 6 - 3/2/2021

Today: Chapters 3 and 6 (joint distributions).
↑ discrete ↑ continuous

① discrete r.v.
- multivariate pmf, marginals
- expectation
- independence (time permitting).

② Continuous r.v.
- multivariate densities (pdfs), marginals
- expectation
- independence (time permitting)

Multivariate discrete distributions

ex: $X, Y \sim \text{Bernoulli}(p)$

$$X = \begin{cases} 0 & \text{prob } 1-p \\ 1 & \text{prob } p \end{cases}, \quad \Omega_X = \{0,1\}$$

$$Y = \begin{cases} 0 & \text{prob } 1-p \\ 1 & \text{prob } p \end{cases}, \quad \Omega_Y = \{0,1\}$$

$p = \frac{1}{2}$.

Ω_X and Ω_Y are different sample spaces.

$E[X \cdot Y] \leftarrow$ nonsense $\rightarrow$ we can't follow the formula below because there isn't the Ω to sum over.

$$E[Z] = \sum_{w \in \Omega} P(Z=w)\, w$$

If we want to deal with X, Y together we need the joint sample space $\Omega = \{(0,0), (0,1)(1,0), (1,1)\}$
(x ↑ ↑ y)

we also need a probability (we can choose ...).

Ex 1

$$(X,Y) = \begin{cases} (0,0) & \text{prob } 1/2 \\ (0,1) & \text{prob } 0 \\ (1,0) & \text{prob } 0 \\ (1,1) & \text{prob } 1/2 \end{cases}$$

$(0,0), (0,1)$: $= \frac{1}{2} = P(X=0)$.

$(1,0), (1,1)$: $= \frac{1}{2} = P(X=1)$

$\Rightarrow$ X and Y are completely linked

Ex 2

$$(X,Y) = \begin{cases} (0,0) & \text{prob } 1/4 \\ (0,1) & \text{prob } 1/4 \\ (1,0) & \text{prob } 1/4 \\ (1,1) & \text{prob } 1/4 \end{cases}$$

$(0,0), (0,1)$: $= 1/2 = P(X=0)$

$(1,0), (1,1)$: $= \frac{1}{2} = P(X=1)$

$\Rightarrow$ X and Y are independent. (more later)

in these above we are given the marginal pmfs.

(Non-valid) Ex 3

$$(X,Y) = \begin{cases} (0,0) & \text{prob } 3/4 \\ (0,1) & \text{prob } 0 \\ (1,0) & \text{prob } 0 \\ (1,1) & \text{prob } 1/4 \end{cases}$$

$(0,0), (0,1)$: $\neq \frac{1}{2} = P(X=0)$

$(1,0), (1,1)$: $\neq \frac{1}{2} = P(X=0)$

so far [Given the pmf of X and Y $\longrightarrow$ define a joint distribution

later [determine the pdf of X, Y $\longleftarrow$ Given ~~define~~ a joint distribution

Def: Given a joint distribution/~~pdf~~ pmf of (X,Y), the <u>marginal</u> distributions/pmf ~~of~~ are the pmfs of X and Y individually.

$$\overset{\text{given}}{\frac{1}{2}} = P(X=0) = P(X=0, Y=0) + P(X=0, Y=1)$$

this is in the joint r.v. / the marginal probability of X.

This is because $\{X=0\} = \{X=0, Y=0\} \cup \{X=0, Y=1\}$

$$\overset{\text{given}}{\frac{1}{2}} = P(X=1) = P(X=1, Y=0) + P(X=1, Y=1).$$

what are all possible joint distributions of (X, Y) that are consistent with given marginals of X, Y?

Ex 1

	$Y=0$	$Y=1$
$X=0$	$1/2$	0
$X=1$	0	$1/2$

$\xrightarrow[\text{general}]{\text{in}}$

	$Y=0$	$Y=1$	
$X=0$	a	b	$\frac{1}{2}$
$X=1$	c	d	$\frac{1}{2}$
	$1/2$	$1/2$	

$$\frac{1}{2} = P(X=0) = P(X=0, Y=0) + P(X=0, Y=1) = a + b$$

$$\Rightarrow a + b = \tfrac{1}{2}$$

all the sums of each row and each column here must be $\frac{1}{2}$.

Additionally, $a+b+c+d = 1$.

we must solve this:

$$a+b+c+d = 1$$

$$a+b = 1/2$$
$$c+d = 1/2$$
$$a+c = 1/2$$
$$b+d = 1/2$$

Assume $X \sim \text{Bernoulli}(p)$, $Y \sim \text{Bernoulli}(q)$.

	$Y=0$	$Y=1$	
$X=0$	a	b	$1-p$
$X=1$	c	d	p
	$1-q$	q	

$$\begin{cases} a+b+c+d = 1 \\ a+b = 1-p \\ c+d = p \\ a+c = 1-q \\ b+d = q \end{cases}$$

But $c+d = p$ and $b+d = q$ are redundant.

So really, we have $\begin{cases} a+b+c+d = 1 \\ a+b = 1-p \\ a+c = 1-q \end{cases}$.

3 eq, 4 vars.

we could do algebra and solve for everything, or we can use 1 degree of freedom.

Pick $0 \le a \le 1-p, 1-q$. } pick a and everything else is determined.

Ex 1

	Y=0	Y=1
X=0	$a=\frac{1}{2}$	$b=0$
X=0	$c=0$	$d=\frac{1}{2}$

$a+c=\frac{1}{2}$ $\quad a+b=\frac{1}{2}$

$a+b+c+d=1$

Given Marginal X, Marginal Y $\rightarrow$ Build joint distribution on joint sample space Ω.

THERE IS NOT A UNIQUE WAY TO DO THIS.

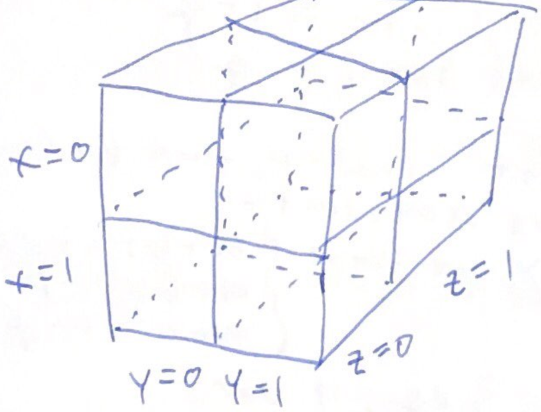
Book focuses on joint dist of two r.v. X Y, called a bivariate dist.

But you could have n r.v. $\underbrace{X_1, X_2, \ldots, X_n}_{n \text{ r.v.}}$, called a multivariate dist, which we will cover.

As n increases, specifying the joint pmf becomes more involved/parameterized.

Ex: Consider ~~X, Y, Z ~ Bernoulli~~
$X \sim$ Bernoulli(p), $Y =$ Bernoulli(q), $Z =$ Bernoulli(r).

we need joint space $\Omega = \{(0,0,0), (0,0,1)(0,1,0), \ldots (1,1,1)\}$.



$\Rightarrow a_1 + a_2 + \cdots + a_8 = 1$

$\vdots$

This is difficult to set out.

Q: $X_1, X_2, \ldots, X_n \sim \text{Bernoulli}(p)$.

$\Omega = \{ \underbrace{(0,0,\ldots,0)}_{n}, \underbrace{(1,1,\ldots,1)}_{n}, (0,1,0,\ldots) \ldots \}$ $\quad |\Omega| = 2^n$.

For large n, this specification is too large.
How do we define joint distributions for $X_1, X_2, \ldots, X_n$?

If we can extend the number of vars, we can also extend such that each var has its own distribution

Ex: $X \sim \text{Poisson}(\mu)$
$Y \sim \text{Bernoulli}(p)$.

$P(X=k) = \dfrac{e^{-\mu}\mu^k}{k!}$

	$Y=0$	$Y=1$	
$X=0$	a_0	b_0	$e^{-\mu}$
$X=1$	a_1	b_1	$e^{-\mu}\mu$
$X=2$	a_2	b_2	$\dfrac{e^{-\mu}\mu^2}{2!}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$X=n$	a_n	b_n	$\dfrac{e^{-\mu}\mu^n}{n!}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
	$1-p$	p	

$\Omega = \{ (0,0), (1,0), (2,0), (3,0), \ldots$
$(0,1), (1,1), (2,1) (3,1) \ldots \}$.

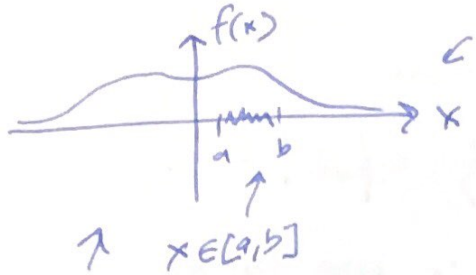
Now we move to continuous r.v. and consider joint distributions similarly to the above.

Let X, Y be continuous r.v.

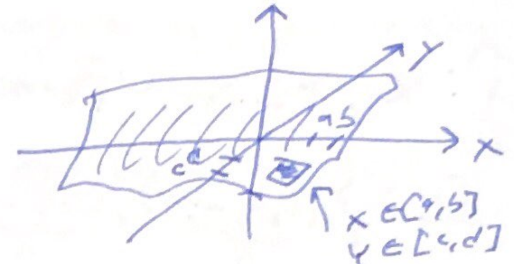
$X \sim f_X(x)$ pdf $\quad Y \sim f_Y(x)$ pdf $\Big\}$ these are the marginals.

$\Omega_X = \mathbb{R} \qquad \Omega_Y = \mathbb{R}$.

so joint space $\Omega = \mathbb{R}^2$ (cartesian plane).
(X, Y) is a random vector in $\mathbb{R}^2$.



← 1 dim



joint pdf $f(x,y) : \mathbb{R}^2 \to \mathbb{R}$.

a surface that lives above $\mathbb{R}^2$.

For a univarite pdf

$$P(X \in [a,b]) = \int_a^b f_X(x)\,dx$$

$$P(Y \in (c,d)) = \int_c^d f_Y(y)\,dy$$

For a bivariate pdf

$$P((X,Y) \in [a,b] \times [c,d])$$

$$= P(X \in [a,b], Y \in [c,d])$$

$$= \int_a^b \int_c^d f(x,y)\, \cancel{dx\,dy}\; dy\,dx$$

$$= \int_a^b dx \int_c^d dy\, f(x,y) \quad \leftarrow \text{alt notation}$$

we will now look at getting the marginals given the joint distribution (the other direction that we already did is hard in the continuous case).

$\underline{\text{Ex}}$: 6.26 $f(x,y) = \begin{cases} e^{-x-y}, & x,y \geq 0 \\ 0 & \text{o/w} \end{cases}$.

$$P(X \leq 1, Y \geq 2) = \int_2^\infty \int_0^1 e^{-x-y}\, dx\, dy$$

$$= \int_2^\infty e^{-y} dy \left(\int_0^1 e^{-x} dx \right)$$

$$= \int_2^\infty e^{-y} dy \left(-e^{-x} \Big|_0^1 \right)$$

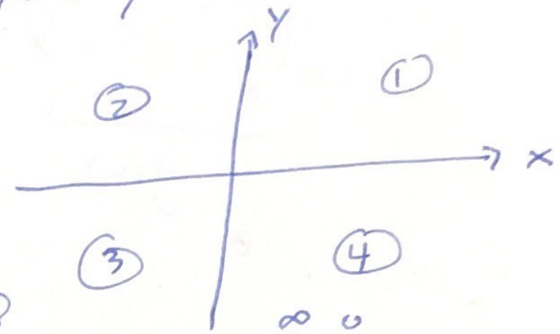
$$= \int_2^\infty e^{-y}dy\,(-e^{-1}+1) = (1-e^{-1})\int_2^\infty e^{-y}dy = (1-e^{-1})(-e^{-y})\Big|_2^\infty$$

$$= (1-e^{-1})e^{-2}.$$

For univariate case, $\int_{-\infty}^{\infty} f(x)dx = 1$. is required.

For multivariate case, and joint distributions $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy = 1$ is required.

In the preceding example, we can break into these quadrant regions.



$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy = \overset{1}{\int_0^{\infty}\int_0^{\infty}} f(x,y)\,dx\,dy + \int_0^{\infty}\int_{-\infty}^{0} f(x,y)\,dx\,dy$$

$$+ \int_{-\infty}^{0}\int_{\infty}^{0} f(x,y)\,dx\,dy + \int_{-\infty}^{0}\int_{0}^{\infty} f(x,y)\,dx\,dy$$

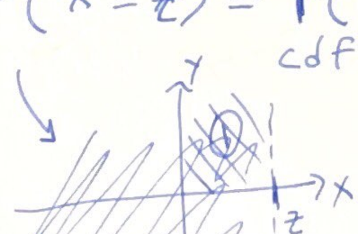
$f(x,y) = 0$ in regions 2, 3, 4.

For this example then, $\int_0^{\infty}\int_0^{\infty} e^{-x-y}\,dx\,dy = 1$.

What is the marginal distribution of X?

we want pdf of X.

$$P(X \le z) = \underset{\text{cdf of X}}{F(z)} = \int_{-\infty}^{\infty}\int_{-\infty}^{z} f(x,y)\,dx\,dy$$



$$= \int_0^{\infty}\int_0^{z} f(x,y)\,dx\,dy$$

$$f_X(z) = F'(z) = \frac{d}{dz}\left(\int_0^\infty \int_0^z e^{-x-y}\,dx\,dy\right)$$

<u>Aside</u>: <u>Recall</u>: $\frac{d}{dz}\int_0^z g(w)\,dw = g(z)$.

$$f_X(z) = \frac{d}{dz}\left[\int_0^z \underbrace{\left(\int_0^\infty e^{-x-y}\,dy\right)}_{g(w)} dz\right] = \int_0^\infty e^{-z-y}\,dy = e^{-z}.$$

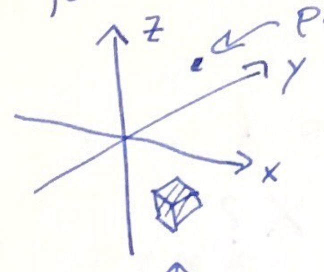
$$\Rightarrow f_X(z) = e^{-z}.$$

we could do the double integral and then take the deriv, but this trick is <u>good</u>.

Additionally, we can extend beyond bivariate case. We can look at X, Y, Z, continuous r.v.

joint distribution (X, Y, Z). $\Omega = \mathbb{R}^3$.

prob of singleton event.



$$f(x,y,z) : \mathbb{R}^3 \to \mathbb{R}.$$

$$P(X \in (a,b), Y \in (c,d), Z \in (e,f))$$

$$= \int_a^b \int_c^d \int_e^f f(x,y,z)\,dz\,dy\,dz.$$

prob that we are in some box in $\mathbb{R}^3$.

In general multivariate case, $X_1, X_2, \ldots, X_n$.

$\Omega = \mathbb{R}^n$. $f(x_1, \ldots, x_n) : \mathbb{R}^n \to \mathbb{R}$.

$$P(X_1 \in (a_1,b_1), \ldots, X_n \in (a_n,b_n)) = \int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} f(x_1, \ldots, x_n)\,dx_n \ldots dx_1.$$

we rarely see or work with multidimensional cdf, but we will define it here:

Def: The joint cumulative distribution function (cdf) of X, Y is given by

$$F(x,y) = P(X \le x, Y \le y)$$

$$= \int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)\,dv\,du$$