2.5 Jointly cont. r. vectors

Def. (X,Y) is a (jointly) continuous r. vector if there is $f \geq 0$ so that the joint df of (X,Y) is

$$F(x,y) = \int_{-\infty}^{x} \left(\int_{-\infty}^{y} f(u,v)\, dv \right) du, \quad -\infty < x, y < \infty.$$

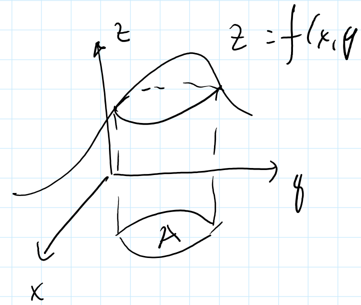
$f(x,y)$ is called joint pdf of (X,Y).

Note if (X,Y) is continuous, then

(i) $F(x,y)$ is continuous;

(ii) joint pdf

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



(iii) For $A \subset \mathbb{R}^2$,

$$P((X,Y) \in A) = \iint_A f(x,y)\, dx\, dy, \quad A \in \mathcal{B}.$$

Geometrically, it is the volume under $z = f(x,y)$ above A.

Ex1. Let (X,Y) be continuous. Find

a) $P(X = Y) = 0$; b) $P(Y = h(X)) = 0$

Answer. a) $P(X=Y) = P((X,Y) \in \Delta) = \iint_\Delta f(x,y)\, dx\, dy = 0$

$\Delta = \{(x,y) : x = y\}$ is the line: there is no volume above Δ:

b) Similarly, $A = \{(x,y) : y = h(x)\}$ is a curve ... .

Also, $P((X,Y) = (a,b)) = 0$.

Ex 2. Let (X,Y) be continuous with joint df $F(x,y)$ and joint pdf $f(x,y)$. Then df and pdf of X are

$$F_X(x) = \int_{-\infty}^{x}\left(\int_{-\infty}^{\infty} f(u,y)\,dy\right)du, \quad f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy, \ -\infty < x < \infty.$$

Answer. For any x, y

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

We found that $F_X(x) = \lim_{y\to\infty} F(x,y) = \int_{-\infty}^{x}\left(\int_{-\infty}^{\infty} f(u,v)\,dv\right)du$,

$f_X(x) = F_X'(x) = \int_{-\infty}^{\infty} f(x,v)\,dv, \ -\infty < x < \infty$

Similarly, $f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,dx, \quad -\infty < y < \infty$.

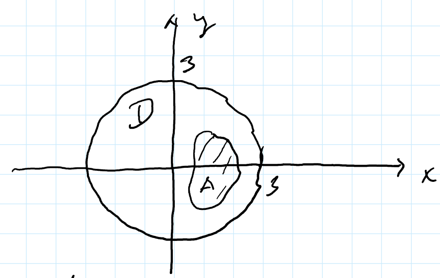
In general, if $(X_1, \dots, X_d)$ is cont. r. vector, then any $(X_{i_1}, \dots, X_{i_k})$, $k = 1, \dots, d$, is cont. r. vector.

Ex 3. Dart is flung at $D = \{(x,y) : x^2 + y^2 < 3^2\}$.

Let (X,Y) be coordinates of the hit.

"Any point in D is equally likely to be hit" is modelled by continuous r. vector (X,Y) with constant joint pdf

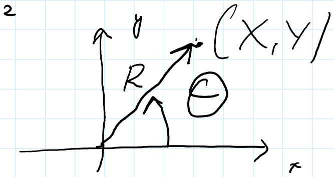
$$f(x,y) = \begin{cases} c, & (x,y) \in D \\ 0, & \text{otherwise} \end{cases}$$



$\iint f(x,y)\,dx\,dy = \iint c\,dx\,dy = c|A|$

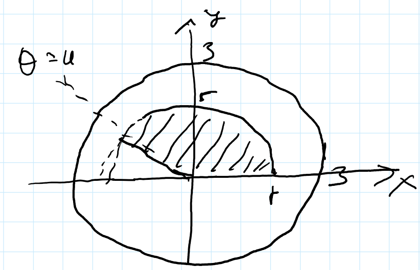
$$P((X,Y)\in A) = \iint_A f(x,y)\,dx\,dy = \iint_A c\,dx\,dy = c|A|.$$

with $A = D$, $1 = c|D|$, $c = \frac{1}{|D|} = \frac{1}{\pi\cdot 3^2}$.



Find joint pdf of $R = \sqrt{X^2+Y^2}$, Θ

Note $R > 0$, $0 \le \Theta < 2\pi$



Answer. For $r > 0$, $0 < u < 2\pi$,

$$P(R \le r, \Theta \le u) = \frac{\frac{ur^2}{2}}{\pi\cdot 3^2} = \frac{u}{2\pi}\cdot\frac{r^2}{3^2}.$$

2.5 Discrete r.v.

Def. (X,Y) is discrete if it takes finite or countable number of values.

Def. Joint pmf of (X,Y) is the function

$$f(x,y) = P(X=x, Y=y), \quad -\infty < x, y < \infty.$$

Ex1. Let (X,Y) be discrete. Find pmf of X.

Answer. The pmf

$$f_X(x) = P(X=x) = \sum_y P(X=x, Y=y) = \sum_y f(x,y).$$

Similarly, $f_Y(y) = \sum_x f(x,y)$.

3.2. Independence and pmf.

Def. Discrete r.v. X, Y are independent if

$\{X=x\}, \{Y=y\}$ are independent for all x, y:

$P(X=x, Y=y) = P(X=x)\, P(Y=y)$ for all x, y, equivalently,

joint pmf $f(x,y) = f_X(x)\, f_Y(y)$ for all x, y.

<u>Def.</u> $X_1, \dots, X_n$ are independent if $\{X_1 = x_1\}, \dots, \{X_n = x_n\}$ is independent family of events for any $x_1, \dots, x_n$.

<u>Thm.</u> If X, Y are independent, then $h(X)$ and $g(Y)$ are independent for any functions h, g.

<u>3.5.</u> <u>Examples of discrete r.v.</u>

1. $X \sim$ binomial (n,p): $X = \#$ of "successes" in n independent trials with $P(\text{success}) = p$, $P(\text{failure}) = q = 1-p$,

$$P(X=k) = \binom{n}{k} p^k q^{n-k}, \quad k = 0, 1, \dots, n.$$

Consider $X_i = \begin{cases} 1 & \text{if success in } i\text{th trial} \\ 0 & \text{if failure in } i\text{th trial}. \end{cases}$

Then $X_1, \dots, X_n$ are independent Bernoulli(p) and $X = X_1 + \dots + X_n$.

2. (hypergeometric) There are m red and w white balls in the box. $X = \#$ of red balls among n selected" is called <u>hypergeometric r.v.</u> Range of $X = \{0, 1, \dots, n\}$, denoting $N = m + w$,

$$\boxed{(1) \quad P(X=k) = \frac{\binom{m}{k}\binom{N-m}{n-k}}{\binom{N}{n}}, \quad k = 0, 1, \dots, n}$$

X is the number of successes but trials are not independent.

Rem 4. If N, m are large and $\frac{m}{N} \approx p$, then in (1):

$$P(X=k) \approx \binom{n}{k} p^k q^{n-k}, \quad k=0,1,\dots,n, \quad q=1-p.$$

3. X is Poisson(λ): it models

(i) number of earthquakes
(ii) number of customer arrivals } in a given time interval.

(iii) number of typos on a printed page.

. . . .

$$\boxed{\begin{array}{l} \text{Range of } X = \{0,1,2,\dots\} \\ P(X=k) = e^{-\lambda}\dfrac{\lambda^k}{k!}, \quad k=0,1,\dots \end{array}}$$

Note for $\lambda=1$, $P(X\geq 6) = \sum_{k=6}^{\infty} P(X=k) = 0.0006$

Where Poisson(λ) comes from?

1. Poisson X arises as approximation of binomial(n,p) - r.v. Y, when n is "large", p is "small", $\lambda = n\cdot p$ is "moderate":

$$P(X=k) = \binom{n}{k} p^k q^{n-k} \approx e^{-\lambda}\frac{\lambda^k}{k!} \text{ with } \lambda = n\cdot p,\ k=0,1,\dots \tag{2}$$

It can be shown,

$$\boxed{|P(X\in A) - P(Y\in A)| \leq np\cdot p = np^2}$$

Why (2)? $P(Y=k) = \dfrac{n!}{(n-k)!}\cdot\dfrac{1}{k!}\, p^k \dfrac{(1-p)^n}{(1-p)^k}$. $\quad = \dfrac{n\cdot(n-1)\dots(n-k+1)}{n^k}$.

Why (2)? $P(Y=k) = \frac{n!}{(n-k)!} \cdot \frac{1}{k!} p^k \frac{(1-p)^n}{(1-p)^k}$ $= \frac{n\cdot(n-1)\cdots(n-k+1)}{n^k}$ $\cdot$

$(1-p)^k \approx 1$

$\cdot \frac{(np)^k}{k!} \left(1-\frac{np}{n}\right)^n \underset{\lambda = np}{\approx} \frac{\lambda^k}{k!}\left(1-\frac{\lambda}{n}\right)^n \approx \frac{\lambda^k}{k!} e^{-\lambda}.$