Joint Distribution Review. Week 1

Prob. space (Ω, F, P) → probability $P(A) \in [0,1]$

Ω ↑ sample space, F ↑ collection of events

If $A_i \cap A_j = \emptyset$ for any $i \neq j$, the $P(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n)$.

$P(A^c) = 1 - P(A)$.

e.g. Flip a coin twice. $P(H) = p$, $P(T) = 1-p$

$\Omega = \{HH, TH, HT, TT\}$

$F = \{$all possible subsets of $\Omega\}$

P : $P(HH) = p^2$

$P(TH) = (1-p)(p)$

☆ Joint prob. mass function.

$p_{X,Z}(x,z) = P(X=x \text{ and } Z=z)$

$p_X(x) = P(X=x)$

$= P(X=x \text{ and } Z=z \text{ for all possible } z)$

$= \sum_z p_{X,Z}(x,z)$

e.g. Pick a person at random from a group of people (e.g. UW undergrade)

$\Omega = \{$ all individuals of N people$\}$.

$X(w)$ = heart rate of $w \in \Omega$

$Y(w)$ = blood pressure of $w \in \Omega$

$Z(w)$ = temperature of w.

Joint distribution function for (X,Y):

$F_{X,Y}(x,y) = P(X \leq x, Y \leq y)$ for $x, y \in \mathbb{R}$.

$$F_X(x) = P(X \le x) = \lim_{y\to\infty} P(X \le x, Y \le y)$$ marginal distribution

$$F_Y(y) = \lim_{x\to\infty} F_{X,Y}(x,y)$$

☆ Joint prob. density function.

$f(x,y) \ge 0$ for cont. random variables X & Y.

1) $\iint f(x,y)\,dx\,dy = 1$

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

$$\begin{aligned} F_{X,Y}(x,y) &= P(X \le x,\ Y \le y) \\ &= P((X,Y) \in (-\infty, x] \times (-\infty, y]) \\ &= \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(u,v)\,du\,dv \end{aligned}$$

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

e.g. (X,Y) is of uniform distribution over $B_r = \{(x,y) : x^2+y^2 \le r^2\}$.

take a point randomly from B_r.

$$f_{X,Y}(x,y) = \begin{cases} c & \text{if } (x,y)\in B_r \\ 0 & \text{elsewhere} \end{cases}$$

$\iint f(x,y)\,dx\,dy = 1 \Rightarrow \iint_{B_r} c\,dx\,dy = 1$

$\qquad\qquad\qquad\qquad\qquad\qquad\quad \parallel$

$c\,\text{Area}(B_r) = c\pi r^2$

$c = \frac{1}{\pi r^2}$

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{\pi r^2}, & (x,y)\in B_r \\ 0, & \text{else} \end{cases}$$

Independence

Definition: Random variables $X_1, X_2, \dots, X_n$ are independent if

$$P(X_1 \in A_1, X_2 \in A_2, \dots, X_n \in A_n) = P(X_1 \in A_1) P(X_2 \in A_2) \cdots P(X_n \in A_n)$$

$$P(X_2 \in A_2 \mid X_1 \in A_1) \overset{\text{def}}{=} \frac{P(X_2 \in A_2 \text{ and } X_1 \in A_1)}{P(X_1 \in A_1)}$$

$$\overset{\text{indep.}}{=} \frac{P(X_2 \in A_2) P(X_1 \in A_1)}{P(X_1 \in A_1)}$$

$$= P(X_2 \in A_2)$$

For discrete r.v. X_1 & X_2, they're independent iff,

$$P(X_1 = k_1 \text{ and } X_2 = k_2) = P(X_1 = k_1) P(X_2 = k_2)$$

Fact: For cont. r.vector $(X_1, \dots, X_n)$, they have joint density function $f(x_1, \dots, x_n)$

1) $X_1, X_2, \dots, X_n$ are independent

$$\Leftrightarrow f(x_1, x_2, \dots, x_n) = f_{X_1}(x_1) f_{X_2}(x_2) \cdots f_{X_n}(x_n)$$

2) If $f(x_1, \dots, x_n) = f_1(x_1) f_2(x_2) \cdots f_n(x_n)$, where $f_1 \geq 0, f_2 \geq 0, \dots, f_n \geq 0$, then $X_1, \dots, X_n$ are independent.

e.g. Randomly sample a point (X, Y) from a disc with radius $r > 0$.

(X, Y) has density $f(x,y) = \begin{cases} \frac{1}{\pi r^2}, & x^2 + y^2 < r^2 \\ 0, & \text{else} \end{cases}$

We know $f_X(x) = \begin{cases} \frac{2\sqrt{r^2 - x^2}}{\pi r^2} & \text{for } -r < x < r \\ 0 & \text{else.} \end{cases}$

$f_Y(y) = \begin{cases} \frac{2\sqrt{r^2 - y^2}}{\pi r^2} & \text{for } -r < y < r \\ 0 & \text{else.} \end{cases}$

As $f_X(x) f_Y(y) \neq f(x,y)$ as functions of (x,y), X & Y are not indep.

e.g. Let $U_1, U_2, \dots, U_n$ be indep. identically distributed r.v. having uniform distribution over $[0,1]$.

Let X = largest
$= \max\{U_1, \dots, U_n\}$

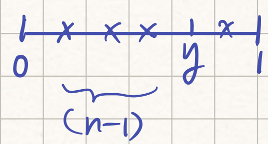
Y = second largest of $U_1, U_2, \dots, U_n$

Find joint density func. of (X,Y).

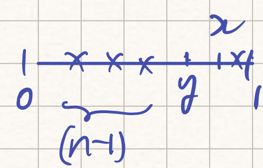
Note $X > Y$. $0 \le X \le 1$, $0 \le Y \le 1$. Take $1 \geq x > y \geq 0$.

$$F(X \le x \text{ and } Y \le y) = P(Y \le y) - P(Y \le y \text{ and } X > x) \qquad (*)$$

$$\begin{aligned} P(Y \le y) &= P(\text{either all of } U_j\text{'s are} \le y \text{ or exactly } (n-1) \text{ of } U_j\text{'s} \le y) \\ &= P(\text{all of } U_j\text{'s} \le y) + P(\text{exactly } (n-1) \text{ of } U_j\text{'s} \le y) \\ &= \underbrace{y \cdot y \cdot \dots \cdot y}_{n} + \binom{n}{1}(1-y) \cdot y^{n-1} \\ &= y^n + n(1-y)y^{n-1} \end{aligned}$$



$$\begin{aligned} P(Y \le y \text{ and } X < x) &= \binom{n}{1}(1-x)y^{n-1} \\ &= n(1-x)y^{n-1} \end{aligned}$$



So
$$\begin{aligned} F(x,y) &= P(X \le x \text{ and } Y \le y) \\ &\overset{(*)}{=} y^n + n(1-y)y^{n-1} - n(1-x)y^{n-1} \\ &= nxy^{n-1} + (n-1)y^n \end{aligned}$$

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \, \partial y} = \begin{cases} n(n-1)y^{n-1} & \text{for } 0 < y < x < 1 \\ 0 & \text{else.} \end{cases}$$