S = set of outcomes of an experiment

Event is a set of outcomes, i.e any subset of S

$P : S \to [0,1]$

$$0 \le P(s) \le 1$$

$$P(S) = 1$$

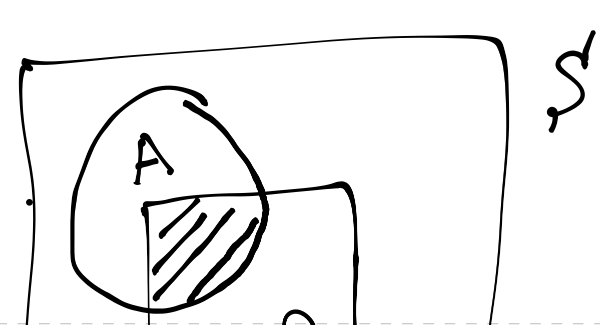
$E_1, \dots, E_n$ mutually

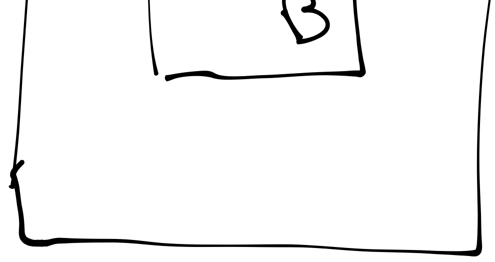
$$P\left(\bigcup_{i=1}^{n} E_i\right) = \sum P(E_i)$$

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

A, B, $P(A \cup B) = P(A) + P(B) - \underline{P(A \cap B)}$

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





$E_1, \dots, E_n$

$$P(A) = \sum_{i=1}^{n} P(A \mid E_i)\, P(E_i)$$

$$= \sum_{i=1}^{n} P(A \cap E_i)$$

$$P(\text{disease} \mid \text{positive}) = \frac{P(\text{positive} \mid \text{disease})\, P(\text{disease})}{P(\text{positive} \mid \text{disease}) \cdot P(\text{disease}) + P(\text{positive} \mid \text{disease}^c) \cdot P(\text{disease}^c)}$$

$\Longrightarrow$ population prevalence is very low relative to even the false positive rate, many of the positive results will actually

---

Def Two events A and B are independent if $P(A \cap B) = P(A) \cdot P(B)$.

$\hookrightarrow P(B|A) = P(B)$

$P(A|B) = P(A)$

not hard to see

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{\cancel{P(A)} \cdot P(B)}{\cancel{P(A)}}$$

→ if we have information about A or B occuring, we don't gain any information about the other occuring.

Ex say I roll two dice → should these outcomes be independent?

→ should be independent

Ex Generic event A. Is A independent of A^c?

$\Rightarrow$ A and A^c are mutually exclusive events

$\Rightarrow P(A \cap A^c) = 0$

unless $A = \emptyset$ or S

A & A^c are not independent.

Ex Let X_1 = { event that dice 1 is a 1 }

E_7 = { event that the sum is 7 }

$$P(E_7 \cap X_1) = P(\{(1,6)\}) = 1/36$$

$$P(E_7) = P(\{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\})$$

$$= 6/36 = 1/6$$

$$P(X_1) = 1/6$$

$$P(E_7 \cap X_1) = P(E_7) \cdot P(X_1)$$

$$E_7 \perp\!\!\!\perp X_1$$

X_2 = { dice is 1 is a 2 }

E_5 = { sum of the dice is 5 }

$$P(X_2 \cap E_5) = P(\{(2,3)\}) = 1/36$$

$$P(X_2) = 1/6$$

$$P(E_5) = P(\{(1,4),(2,3),(3,2),(4,1)\})$$

$$= 1/9$$

$\Rightarrow \frac{1}{30} \neq \frac{1}{6} \cdot \frac{1}{9}$ X_2 & E_5 are not independent

Ex E_7 as above

X_4 1st dice is 4

Y_3 2nd dice is 3

$X_4 \perp\!\!\!\perp Y_3$ are clearly independent

$X_4 \perp\!\!\!\perp E_7$

$Y_3 \perp\!\!\!\perp E_7$

$(X_4 \cap Y_3)$ is it independent of E_7?

$$P(X_4 \cap Y_3) = 1/36 = P(\{(4,3)\})$$

$$P(X_4 \cap Y_3 \cap E_7) = \frac{1}{36}$$

$$\underset{\frac{1}{36}}{P(X_4 \cap Y_3 \cap E_7)} \neq \underset{\frac{1}{36}}{P(X_4 \cap Y_3)} \cdot \underset{\frac{1}{6}}{P(E_7)}$$

Three events here that are pairwise independent but their insersection were not independent.

Def events $A, B, C \Rightarrow$ mutually independent

if
$$P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$$
$$P(A \cap B) = P(A) \cdot P(B)$$
$$P(B \cap C) = P(B) \cdot P(C)$$
$$P(A \cap C) = P(A) \cdot P(C)$$

$\Rightarrow$ more than 3 we basically need all possible intersections

---

Def Given a sample space S, a random Variable X is a function that maps outcome in $S \to \mathbb{R}$

$$X : S \to \mathbb{R}.$$

Ex Let $S = \{$ outcomes from rolling a pair of dice $\}$
$$= \{ (i, j) : i = 1, \dots, 6,\ j = 1, \dots, 6 \}$$

X = the sum of the two dice

$= i + j$

$\underline{Ex}$ Let X = # of times I get tails before tossing a heads

$S = \{ \{H\}, \{TH\}, \{TTH\}, \{TTTH\}, \dots \}$

$X(H) = 0$

$X(TH) = 1$

$X(TTH) = 2$

$\underline{Ex\ 3}$ Randomly select a person in the class. X: the person's height

$\underline{Ex\ 4}$ $X = \begin{cases} 1, & \text{voted in the 2016 US elections} \\ 0, & \text{otherwise} \end{cases}$

$\underline{Ex\ 5}$ I pick a lightbulb out at the store,

X = # of hours of being on until it breaks

Since the vals of a random variable is determined by the corresponding outcome of the underlying experiment, we can talk about probabilities.

If $X =$ sum of two dice

$P(X=2) = 1/36$

$P(X=3) = 2/36$

$P(X=4) = 3/36$, etc

$P(X=12) = 1/36$

In general there are two kinds of random variables:

- Discrete $\to$ X takes a finite # of vals or a countable # of values
- Continuous $\to$ X can take all of the values in some interval $[a,b]$ (or a union of such intervals)

countable $\to$ mapping between the $\mathbb{Z}^+$ and the items in the set

Def
For a discrete random, there is a fx. called probability mass fx (p.m.f.)

$p(x) = P(X = x)$

$x_1, x_2, \ldots, x_n$ $\quad p(x_i) > 0$

$x_1, x_2, \ldots,$ $\quad p(x) = 0, \; x \notin \{x_1, x_2 \ldots\}$

$$\sum_{x_i} p(x_i) = 1$$

Def For a continuos random varieble, there exists some function f called the probebility density function (pdf or density) s/t for any set $A \in \mathbb{R}$

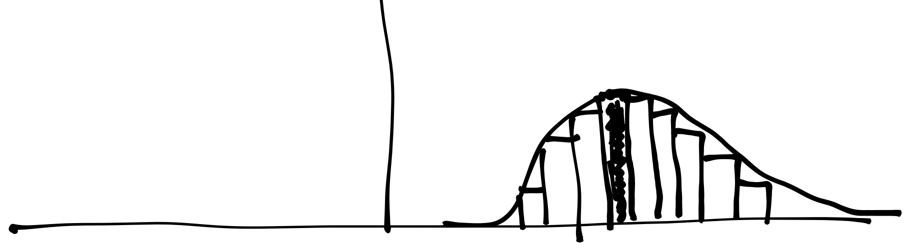
$$P(X \in A) = \int_A f(x)\,dx$$

$\Rightarrow$ any functn s/t $f(x) \geq 0$

$\int_{\mathbb{R}} f(x)\,dx = 1$ is the density of some random varieble.

integral is th area under

the curve



$$P(X \in (x-\underline{d}x, x+\underline{dx}))$$

$$\approx f(x)\cdot(2dx)$$

For any random variable, there exists a function called the cumulative distribution function (cdf or distribution)

$$F(x) = P(X \leq x)$$

discrete:

$x_1, \dots, x_n$

or $x_1, x_2, \dots$

$$F(x) = \sum_{x_i \leq x} p(x_i)$$

cont:

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

$$\left(f(x) = \frac{d}{dx} F(x) \right)$$

<u>Ex</u> $X : S \to \{1,2,3\}$

$p(1) = 0.5$

$p(2) = 1/3$

$p(3) = 1/6$

$$F(x) = \begin{cases} 0, & x < 1 \\ 0.5, & 1 \leq x < 2 \\ 5/6, & 2 \leq x < 3 \\ 1, & x > 3 \end{cases}$$

Properties of the cdf

$F(x)$ is an increasing fx

$F(x) \geq F(y)$ if $x > y$

$$\lim_{x \to -\infty} F(x) = 0$$

$$\lim_{x \to \infty} F(x) = 1$$

Ex Toss coin until H comes up

H, TH, TTH

0, 1, 2

$p(0) = \frac{1}{2}$ $p(1) = (\frac{1}{4})$ $p(2) = (\frac{1}{8})$

$$F(x) = \begin{cases} 0, & x < 0 \\ \sum_{j=0}^{i} \left(\frac{1}{2}\right)^j, & i \le x < i+1 \end{cases}$$

geometric series

Ex $f(x) = \begin{cases} 2x, & 0 < x < 1 \\ 0, & o/w \end{cases}$

$X \to$ takes on vals in the interval $(0,1)$
and more likely to take
on vals closer to 1

$$F(x) = \begin{cases} 0, & x < 0 \\ x^2, & 0 \le x < 1 \\ 1, & x \ge 1 \end{cases}$$

Note $P(a < X \le b) = F(b) - F(a)$

$P(\{X \le b\}) = P(\{X \le a\}) + P(\{a < X \le b\})$
$\cup$

$\shortparallel$

$\parallel$ $F(n)$ $\parallel$ $F(\cdot)$

---

pmf of a discrete random variable $p(i)$

tells us exactly the probability

$P(\underline{X}=i)$

continuous random variables have densities

so if I want to know the probability

that $\underline{X}$ is $\in [a,b]$

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

$\underline{Ex}$ Toss coin until H comes up

H, TH, TTH, TTTH, TTTTTH

$\underline{0}$ $\underline{1}$, $\underline{2}$

magic coin and it comes up heads with prob p

$P(X=0) = P(\text{heads on first toss})$

$= p$

$P(X=1) = P(\text{tails on my first toss, heads on my second toss})$

$= P(\text{tails on 1})P(\text{heads on 2})$

$= (1-p) \cdot p$

$P(X=2) = P(\text{tails on 1 and 2, heads on 3})$

$= P(\text{tails 1})P(\text{tails 2})P(\text{H 3})$

$= (1-p) \cdot (1-p) \cdot p$

$= (1-p)^2 p$

$P(X=i) = (1-p)^i \cdot p$

$$p(x) = \begin{cases} (1-p)^x \cdot p, & \text{for } x = 0, 1, 2, \dots \\ 0, & \text{o/w} \end{cases}$$

$F(x) = P(X \leq x)$

$$= \sum_{0 \le i \le x} p(i) \quad , i = 0, 1, 2, \ldots$$

$$F(-1) = 0$$

$$F(0.5) = p(0)$$

$$F(2) = p(0) + p(1) + p(2)$$

$$F(2.1) = p(0) + p(1) + p(2)$$

$$F(1000) = \sum_{i=0}^{1000} p(i) = \sum_{i=0}^{1000} (1-p)^i p$$

$$p = 1/2 \qquad p(i) = \left(\frac{1}{2}\right)^{i+1}$$

$F(0) = \frac{1}{2}$

$F(1) = 3/4$

$F(2) = 7/8$

$F(x)$