Lecture 9

Def: A <u>random variable</u> is a function $X : S \to \mathbb{R}$, where (S, P) is a probability space.

↳ $\boxed{\text{Ex}}$: Summing the result of two dice rolls $X : S \to \mathbb{R}$

$(n, m) \quad n, m \in \{1, 2, \ldots, 6\}$

$X[(1,5)] = 6$

↳ $\boxed{\text{Ex}}$ The lifetime of a lightbulb is a <u>continuous</u> random variable.

Convention: Random variables are to be denoted by capital letters: $A, B, C, \ldots, X, Y, Z$

Def: Let X be a random variable. The <u>cumulative distribution function</u> (cdf) of X is the function $F_X : \mathbb{R} \to [0,1]$, $F_X(x) = P(X \leq x)$

$\boxed{\text{Ex}}$ Flip a fair coin once. Let X be the number of heads.

↳ $S = \{H, T\} \qquad X : S \to \{0, 1\}$

$X(H) = 1 \qquad P(X=1) = \frac{1}{2}$

$X(T) = 0 \qquad P(X=0) = \frac{1}{2}$

$P(X = \sqrt{2}) = 0$

$$F_X(x) = \begin{cases} 0, & \text{if } x < 0 \\ \frac{1}{2}, & \text{if } 0 \leq x < 1 \\ 1, & \text{if } x \geq 1 \end{cases}$$

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

$\boxed{\text{Ex}}$ Roll a pair of fair dice, let X be the sum of two numbers:

Find $F_X(1)$, $F_X(7)$, $F_X(15)$

$F_X(1) = P(X \le 1) = 0$

$F_X(15) = P(X \le 15) = 1$

$F_X(7) = P(X \le 7) = \frac{\text{favorable events}}{\text{total outcomes}}$

$$\text{favorable events: } \begin{cases} (1,1),\ (1,2),\ \dots\ (1,6) \\ (2,1),\ (2,2),\ \dots\ (2,5) \\ \vdots \\ (6,1) \end{cases} = 21$$

total outcomes: $6 \times 6 = 36$

So $F_X(7) = \frac{21}{36}$

<u>Def</u>: (Quantiles) Suppose that X is a random variable with cumulative distribution function F_X. For any number $p \in [0,1]$ the <u>p-quantile</u> of X, denoted by $Q_X(p)$ is the smallest number x_0 such that:

$$F_X(x_0) = P(x \le x_0) \ge p$$

- Thus, x_0 is the p-quantile of X $[x_0 = Q_X(p)]$

  if

  - $P(x \le x_0) \ge p$ <u>and</u>
  - $P(x \le x) < p$ for any $x < x_0$

↳ $\boxed{\text{Ex}}$ In this cdf:

$$F_X(x) = \begin{cases} 0, & \text{if } x < 0 \\ \frac{1}{2}, & \text{if } 0 \le x < 1 \\ 1, & \text{if } x \ge 1 \end{cases}$$

- the $\frac{1}{4}$-quantile $= 0$

• The median of random variable X is the $\frac{1}{2}$-quantile

Ex :

score x	50	60	70	80	90
$P(X=x)$	0.1	0.2	0.3	0.25	0.15

← Assume these are the only scores possible

cdf

$$F_X(x) = P(X \le x) = \begin{cases} 0.1 & x=50 \\ 0.3 & x=60 \\ 0.6 & x=70 \\ 0.85 & x=80 \\ 1 & x=90 \end{cases}$$

0.75 - quantile is 80 :

$F_X(70) = 0.6$ which is smaller than 0.75 NOT Good

$F_X(80) = 0.85$ which is greater than 0.75 Good

• This means 75% of students scored at least an 80.

• Quantiles convert probabilities into meaningful cutoff values and give robust, interpretable summaries of distributions

- especially for skewed data and risk analysis.

• **Ex** Let S be the number of steps a person walks in a day, measured across a large group participating in a welness program.

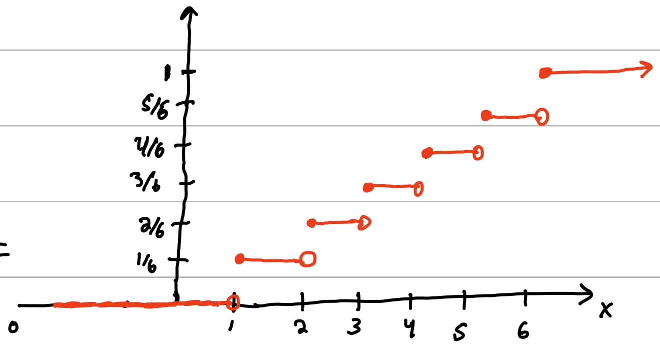
↳ the 0.7 quantile is 10000 steps:

↳ means that 70% of participants walk 10000 steps or fewer, and less than 70% walk 9999 steps or fewer.

$\boxed{\text{Ex}}$ Roll a dice. X is the outcome. Find the median of X.

1) cdf:

$F_X(x) =$



$P(X \le 3) = \frac{1}{2}$ So $P(X \le 3) \ge \frac{1}{2}$ ✓

$P(X < 3) = \frac{2}{6}$ So $P(X < 3) < \frac{1}{2}$ ✓

$\boxed{3 \text{ is the median}}$

Def: Independence

↳ Assume (S, P) is a probability space, $X_i : S \to \mathbb{R}$

i) The random variables $X_1, X_2, \dots, X_n$ are called independent if for any numbers $x_1, \dots, x_n \in \mathbb{R} \cup \{+\infty\}$ the events $\{X_1 \le x_1\}, \dots, \{X_n \le x_n\}$ are independent.

↳ $P(X_1 \le x_1, \dots, X_n \le x_n) = P(X_1 \le x_1) \cdot \ldots \cdot P(X_n \le x_n)$

ii) An infinite sequence of regular variables $X_n : S \to \mathbb{R}$ is called independent if $X_1, \dots, X_n$ are independent for any n.

Notation: $X \perp\!\!\!\perp Y$, X, Y are regular variables.

• The sequence is called i.i.d. (independent, identically distributed) if it is independent and the regular variable have the same cdf.

$P(X_i \le x) = P(X_j \le x) \quad \forall\, i, j$