3Y03 - 3J04 PROBABILITY & STATISTICS FOR (CIVIL) ENGINEERING

C01 - Lecture 7

Last time Random Variables $X, Y, \ldots$

→ represent numerical outcomes of an experiment

↳ discrete (range finite or countably infinite)
or ↳ continuous (range contains an interval)

Example Play video game 5 times; games independent. Stop after losing. $P(\text{win}) = 0.7$. $X = \#$ wins. Find $P(X=x)$
$[P(\text{loss}) = 0.3]$ for each $x = 0, 1, 2, 3, 4, 5$.

$P(L) = P(X=0) = 0.3$

$P(WL) = P(X=1) = 0.7 \times 0.3 = 0.21$

$P(WWL) = P(X=2) = (0.7)^2 \times 0.3 = 0.14$

$P(WWWL) = P(X=3) = (0.7)^3 \times 0.3 = 0.1029$

$P(WWWWL) = P(X=4) = (0.7)^4 \times 0.3 = 0.07203$

$P(WWWWW) = P(X=5) = (0.7)^5 = 0.16807$

Add up to 1

We can sometimes describe the prob. distribution with a function e.g. here

$$P(X=x) = \begin{cases} (0.7)^x \cdot 0.3 & x \neq 5 \\ (0.7)^5 & x = 5 \end{cases}$$

[3.2]

Def^n (Probability Mass Function)

For a discrete random variable X which can take values in the range of X: $R_X = \{x_1, x_2, \ldots, x_n\}$ (or

$R_X = \{x_1, x_2, x_3, \dots\}$) countably infinite ↖ finite

a **probability mass function** $f(x)$ is a function

such that

(1) $f(x_i) \geq 0$

(2) $\sum_{i=1}^{n} f(x_i) = 1$ $\left(\text{or } \sum_{i=1}^{\infty} f(x_i) = 1\right)$

(3) $f(x_i) = P(X = x_i)$

Notice

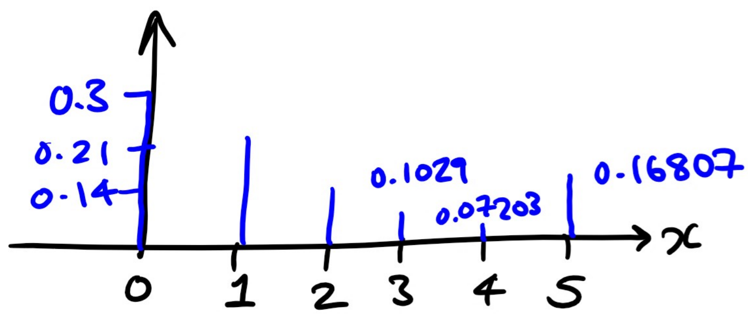
By Prob. Axiom (3) this gives us $P(X \in A) = \sum_{x_i \in A} f(x_i)$

A: set of x_is

(the events $X = x_1$, $X = x_2$, etc. are mutually exclusive) X cannot equal 2 values at the same time

We can visualize p.m.f.

e.g. Example above:



[3.3] Cumulative Distribution Function

Similar idea (above: $P(X = x)$). Now we have a function describing $P(X \leq x)$.

We call this a cumulative distribution function (c.d.f.) denoted by $F(x) = P(X \leq x)$

$$= \sum_{x_i \leq x} f(x_i)$$

↑ p.m.f.

Example Suppose you know that the prob. mass function $f(x)$ of a discrete random variable X is given by:

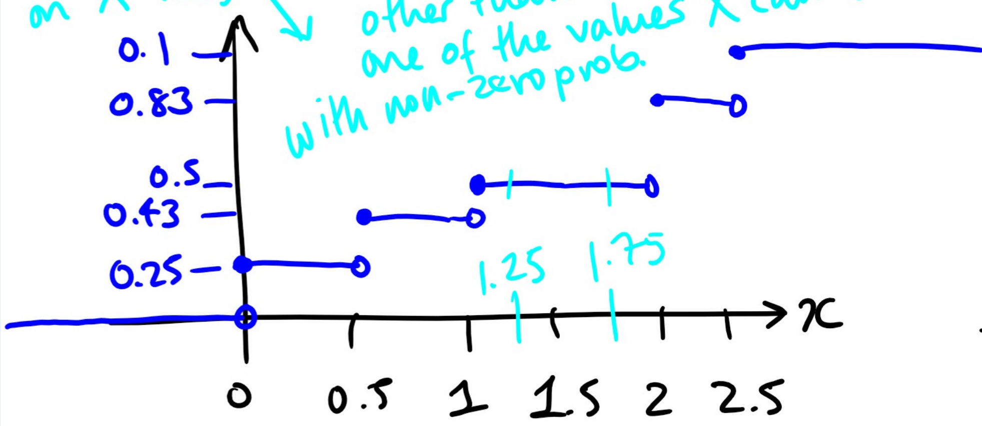
x	0	0.5	1	1.5	2	2.5	
$f(x)$	0.25	0.18	0.07	0	0.33	0.17	← add to 1

Find $F(x)$, the cumulative distribution function.

Solution

$$P(X \leq x) = F(x) = \begin{cases} 0 & \text{if } x < 0 \\ 0.25 & \cancel{\text{if } x = 0} \text{ if } 0 \leq x < 0.5 \\ 0.43 & \text{if } 0.5 \leq x < 1 \\ 0.5 & \text{if } 1 \leq x < \cancel{1.5}\, 2 \\ 0.83 & \text{if } 2 \leq x < 2.5 \\ 1 & \text{if } 2.5 \leq x \end{cases}$$

Notice that $P(X \leq 1.75) = P(X \leq 1.25)$, say.
Here the value of little x (circled) of interest as a bound on X might be something other than one of the values X can take with non-zero prob.



Notice

- $F(x)$ starts at 0, ends at 1
- Non-decreasing ($x \leq y \Rightarrow F(x) \leq F(y)$)
- Right-continuous everywhere
- Jumps at x are equal to p.m.f. at x

So for any $F(x)$ with these properties we can describe a corresponding p.m.f. (just calculate the value of the jump at x to get $f(x)$)

[3.4] Mean & Variance of a Discrete Random Variable

- ways to summarise a prob. dist. function

- Mean : "centre" / "balance point" (where prob. masses would balance)
  ↳ weighted average (weights ~ probabilities)

Denoted by $\underbrace{\mu = E(X)}_{\text{"expected value" of } X} = \sum_x x f(x)$

- long-run average of X after repeated observations (Law of Large Numbers)

Notice If $f(x)$ is same for all x i.e. $f(x) = \frac{1}{N}$ then μ is just average of x-values.

Variance "dispersion", "spread"

$$\sigma^2 = V(X) = E((X-\mu)^2) = \sum_x (x-\mu)^2 f(x)$$

σ standard deviation

expected value of square of distance from mean μ.