Math 180A: Intro to Probability (for Data Science)

www.math.ucsd.edu/~tkemp/180A

Today: § 3.1-3.2

Next: § 2.4-2.5

HW2 due TONIGHT!

Lab 2.2 due Monday, 10/14.

* Regrade requests for HW1: window Tuesday 10/15 8am – 11pm
  - ↳ separate request for each problem
  - ↳ detailed, polite responses please.
* Numerical answers for HW in rubric on Gradescope
* Lab 1 solutions posted on datahub.ucsd.edu.

Random Variables

3.1

Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ a random variable is a function

$$X : \Omega \to \mathbb{R}$$

This is a bad, old-fashioned name. Would be better to call it a random function or random measurement.

Eg. Toss a fair coin 4 times. Let X = number of tails.

Eg. Shoot an arrow at a circular target. Y = distance from center.

Eg. Your car is in a minor accident; the damage repair cost is a random number between \$100 and \$1500. Your insurace deductible is \$500. Z = your out of pocket expenses.

In all these examples, think about what you observe. You can't really see a formula for the function. By repeating the experiment over and over, all you can learn is the probability distribution.

Probability Distribution

3.2

Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a random variable $X : \Omega \to \mathbb{R}$

the probability distribution or law of X is a probability measure μ_X on $\mathbb{R}$.

$$A \subseteq \mathbb{R} \quad \rightsquigarrow \quad \mu_X(A) = \mathbb{P}(\{X \in A\})$$

[Caution: for this to make sense, we need to have a designated set of allowed "events" in $\mathbb{R}$; call this collection $\mathcal{B}(\mathbb{R})$

Then, we must have

$$\boxed{\text{For each } A \in \mathcal{B}(\mathbb{R}),\ \{X \in A\} \in \mathcal{F}.}$$

This is a condition on X; we call such functions **measurable**.

We will ignore these technicalities in this course; all our random variables are indeed measurable.]

E.g. $\underbrace{\text{Toss a fair coin 4 times}}$. Let X = number of tails.

$\Omega = \{ (x_1, x_2, x_3, x_4) \in \{H, T\}^4 \}$

$\mathbb{P}$ = uniform on Ω;

$\mathbb{P}\{(x_1, x_2, x_3, x_4)\} = \frac{1}{2^4} = \frac{1}{16}$

$\{X=2\} = \{ (T,T,H,H), (T,H,T,H), \dots \}$

$\#\{X=2\} = \binom{4}{2}$.

$\mathbb{P}(X=2) = \frac{\binom{4}{2}}{16} = \frac{3}{8}$

$\mathbb{P}(X=k) = \frac{\binom{4}{k}}{16}$

$\downarrow$

$X \in \{0, 1, 2, 3, 4\} \subset \mathbb{R}$

if $A \subseteq \mathbb{R}$ does not contain one of these numbers, $\mu_X(A) = 0$.

By additivity of probability measures, to understand μ_X, just need to know $\mu_X(\{k\}) = ?$ $0 \le k \le 4$

$$\mu_X(\{k\}) = \mathbb{P}(X=k) = p_X(k)$$

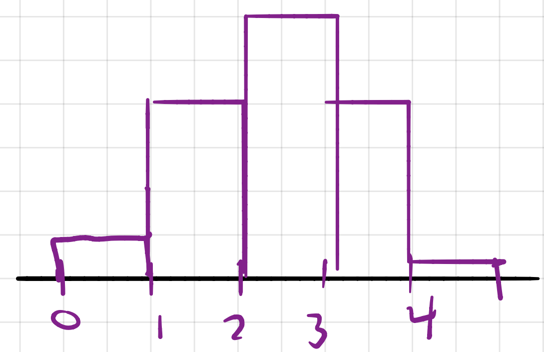
k	0	1	2	3	4
$p_X(k)$	1/16	1/4	3/8	1/4	1/16

X = # tails in n coin tosses,

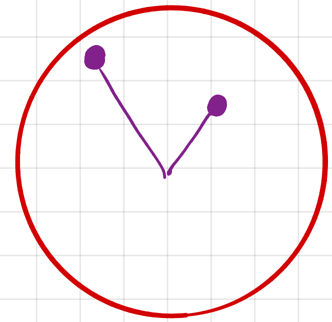
$$p_X(k) = \mathbb{P}(X=k) = \frac{1}{2^n}\binom{n}{k} \quad 0 \le k \le n$$

Binomial

$\mathrm{Bin}(n, \frac{1}{2})$



Eg. Shoot an arrow at a circular target of radius 1.



Y = distance from center.

Last lecture, you calculated that:

- If $r \le 0$, $\mathbb{P}(Y \le r) = 0$
- If $r \ge 1$, $\mathbb{P}(Y \le r) = 1$.
- If $r \in [0,1]$, $\mathbb{P}(Y \le r) = \dfrac{\text{Area}(\mathbb{D}_r)}{\text{Area}(\mathbb{D}_1)} = r^2$.

$$\mathbb{P}(Y \le r) = \mathbb{P}(\{Y \in (-\infty, r]\})$$

← Can get information about other sets $A \subseteq \mathbb{R}$ using the properties of $\mathbb{P}$.

$$\mathbb{P}(Y = 0.4) \le \mathbb{P}(Y \in (0.4-\varepsilon, 0.4]) \to 0 \text{ as } \varepsilon \to 0.$$

as $\varepsilon \to 0$, this $\to 0$.

What can we say about $\mathbb{P}(Y = 0.4) \overset{?}{=} 0$.

$(0.3, 0.4]$

$$\mathbb{P}(Y \in (0.3, 0.4]) = 0.4^2 - 0.3^2 = 0.07$$

$$\mu_Y(-\infty, 0.4] = \mu_Y(-\infty, 0.3] + \mu_Y(0.3, 0.4]$$

0.3 0.4

$$(-\infty, 0.4] = (-\infty, 0.3] \cup (0.3, 0.4]$$

$$\mathbb{P}(Y \in (0.4-\varepsilon, 0.4]) = 0.4^2 - (0.4-\varepsilon)^2$$

We will focus mostly on two kinds of random variables:

discrete: There are finitely (or countably) many possible values $\{k_1, k_2, k_3, \dots\}$ for X.

↳ μ_X is described by the probability mass function $p_X(k) = \mathbb{P}(X=k)$, $k \in \{k_1, k_2, k_3, \dots\}$

In this case, by the laws of probability,

$$p_X(k) \geq 0 \text{ for each } k, \quad \sum_{j=1}^{n} p_X(j) = 1.$$

continuous: For any real number $t \in \mathbb{R}$, $\mathbb{P}(X=t) = 0$.

↳ μ_X is captured by understanding $\mathbb{P}(X \leq r)$ as a function of r.

Eg. $\mathbb{P}(X \in [a,b]) = \mathbb{P}(\{X=a\} \cup \{X \in (a,b]\})$

$$= \underbrace{\mathbb{P}(X=a)}_{0} + \mathbb{P}(X \in (a,b])$$

$$= \mathbb{P}(X \leq b) - \mathbb{P}(X \leq a)$$

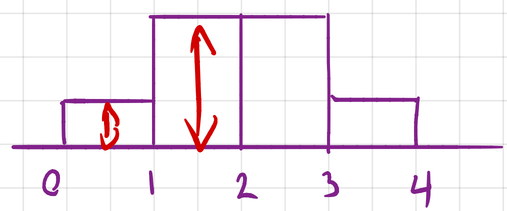
Cumulative Distribution Function (CDF)

For any random variable X, $F_X(r) = \mathbb{P}(X \le r)$.

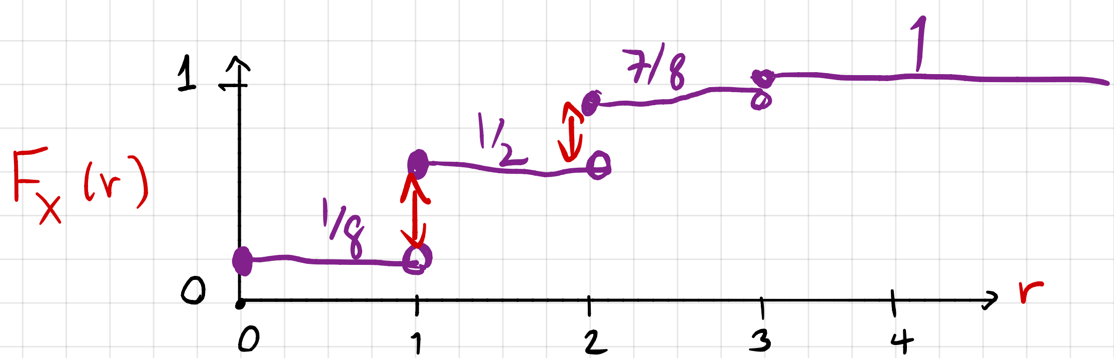
Actually works in all cases – including discrete.

Eg. $\mu_X = \mathrm{Bin}(3, \frac{1}{2})$

k	0	1	2	3
$p_X(k)$	$1/8$	$3/8$	$3/8$	$1/8$



- $r < 0$, $\{X \le r\} = \emptyset$ $\quad \mathbb{P}(X \le r) = 0$
- $0 \le r < 1$, $\{X \le r\} = \{X = 0\}$ $\quad \mathbb{P}(X \le r) = \mathbb{P}(X = 0) = 1/8$
- $1 \le r < 2$, $\{X \le r\} = \{X = 0, 1\}$ $\quad \mathbb{P}(X \le r) = \mathbb{P}(X = 0) + \mathbb{P}(X = 1) = \frac{1}{8} + \frac{3}{8} = \frac{1}{2}$
- $2 \le r < 3$, $\{X \le r\} =$ $\vdots$
- $r \ge 3$, $\{X \le r\} =$ $\vdots$



Properties of the CDF $F_X(r) = \mathbb{P}(X \leq r)$

(1) Monotone increasing: $s \leq t \Rightarrow F_X(s) \leq F_X(t)$

(2) $\lim_{r \to -\infty} F_X(r) = 0$, $\lim_{r \to +\infty} F_X(r) = 1$.

(3) The function F_X is right-continuous: $\lim_{t \to r+} F_X(t) = F_X(r)$.

Corollary: If X is a continuous random variable, F_X is a continuous function.

Densities

Some continuous random variables have probability densities. This is an infinitesimal version of a probability mass function.

X discrete, $\in \{k_1, k_2, k_3, \ldots\}$

$p_X(k) = \mathbb{P}(X = k)$

$$\mathbb{P}(X \in A) = \sum_{k \in A} \mathbb{P}(X = k)$$

$$= \sum_{k \in A} p_X(k)$$

$p_X(k) \geq 0$, $\sum_k p_X(k) = 1$.

X continuous

$\mathbb{P}(X = t) = 0$ for all $t \in \mathbb{R}$.

$$\mathbb{P}(X \in A) = \int f_X(t)\, dt$$

Eg. Shoot an arrow at a circular target of radius 1.

Y = distance from center.

$$\int_{-\infty}^{r} f(t)\,dt \overset{?}{=} \mathbb{P}(Y \in (-\infty, r]) = F_Y(r) = \begin{cases} 0, & r \leq 0 \\ r^2, & 0 \leq r \leq 1 \\ 1, & r \geq 1 \end{cases}$$