Probability Theory

Textbook: *Introduction to Probability* by Blitzstein and Hwang

---

Previous Lecture

- Indep of Events (conditional indep)
- Conditioning as Problem-Solving
  Wishful Thinking/First Step Analysis
- Coherency of Bayes' Rule
- Pitfalls and Paradoxes

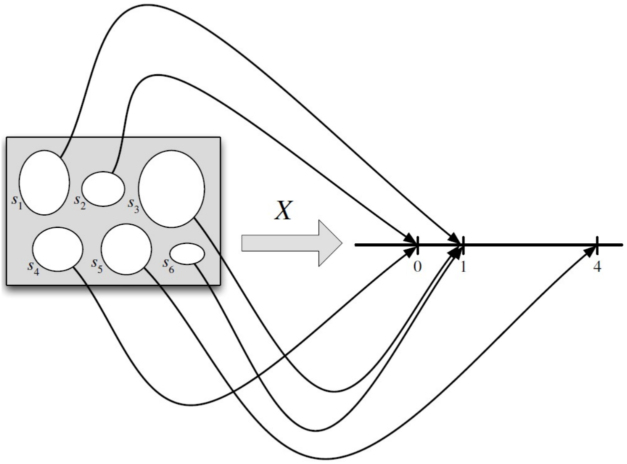
---

§3.1 - Random Variables

To simplify notation and many calculations, we introduce the idea of a random variable.

"Random variables" are not really random, but rather associated with a random process (or "experiment"). (random just expresses our inability to predict something). They're also not really variables, but rather functions.

Def (**Random Var**). Given an experiment with sample space S, a random var (rv) X is a function from S to $\mathbb{R}$.



Ex (**Coin Tosses**). Imagine an experiment where we toss a fair coin twice. The sample space is $S = \{HH, HT, TH, TT\}$.

- Let X be # of Heads. Then X is a rv w/values $0, 1,$ and 2.
  As a function, X assigns 2 to the outcome HH,
  1 to HT and TH, and 0 to TT.
  That is, $X(HH) = 2,\ X(HT) = X(TH) = 1,$ and $X(TT) = 0$.

- Let Y be # of Tails. In terms of X:

$$Y = 2 - X. \text{ In other words, } Y(s) = 2 - X(s) \text{ for all } s \in S.$$

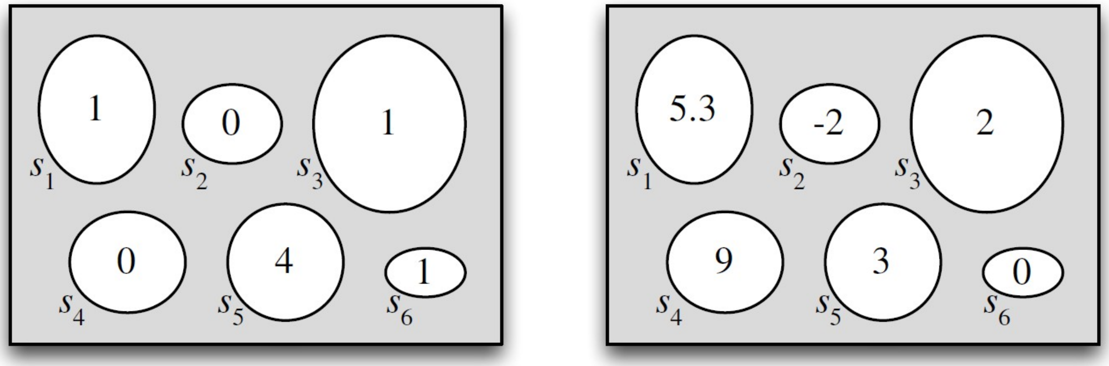
- Let I be 1 if first toss is Heads, and 0 otherwise.

  Then I assigns 1 to HH and HT, and 0 to TH and TT.

  This is an *indicator* rv since it indicates whether the first toss is Heads. 1 means "yes," 0 means "no."

We can also encode the sample space as $S_2 = \{(1,1),(1,0),(0,1),(0,0)\}$, where 1 is Heads & 0 is Tails.
Then we can give explicit formulas for X, Y, I :

$$X(s_1,s_2) = s_1 + s_2, \qquad Y(s_1,s_2) = 2 - s_1 - s_2, \qquad I(s_1,s_2) = s_1, \text{ where } (s_1,s_2) \in S_2.$$



Rv assigns #s to outcomes in S

The sample space S can be multidimensional, and the outcomes $s \in S$ may be non-numeric (color, labels, etc.).

So, rvs provide **numerical summaries** of an experiment.

Plinko Random Process

The path the coins take as they're dropped is too complicated: random.

Sample space is the 5 outcomes (slots) at the bottom $\{s_1, s_2, s_3, s_4, s_5\}$.

$X(s)$ then assigns a number to each outcome $1,2,3,4,5$.

If we send many coins down (repeat X many times), we can discover X's distr. It's the shape below the plinko board. It looks mound shaped (normal).

Harvard Video (2nd half): youtube.com/watch?v=PNrqCdslGi4&list=PL2SOU6wwxB0uwwH80KTQ6ht66KWxbzTIo&index=8

---

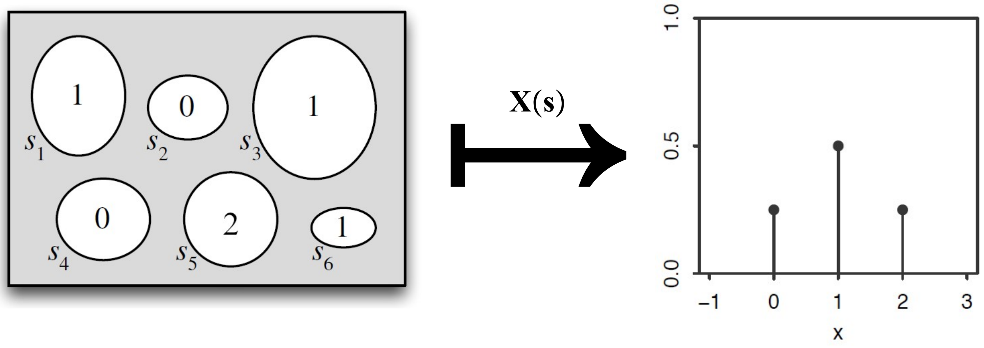
§3.2 - Distr's & Prob Mass Functions (PMF/PF)

One can classify rvs as being discrete or continuous. In this section we'll focus on discrete rvs.

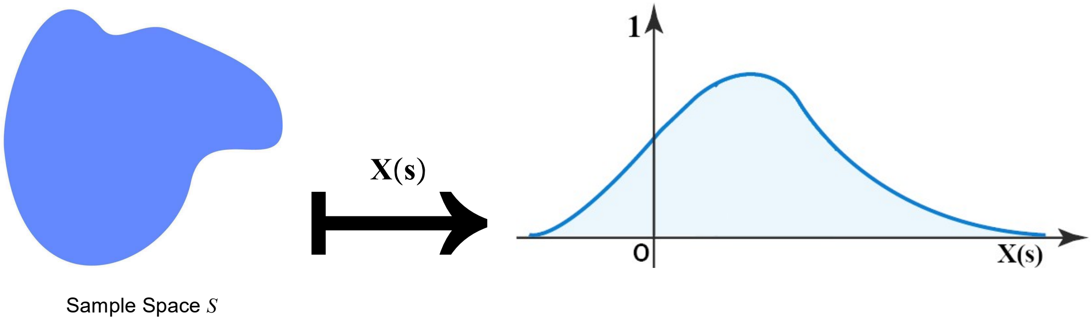
Def (**Discrete Rv**): X is **discrete** if there is a finite list of values $a_1, a_2, \ldots, a_n$ or an infinite list of values

$a_1, a_2, \ldots$ such that $P(X = b) = 0$ for any b not in the list of values. So the rv only take on the values in the discrete list a_i.

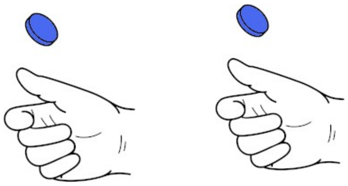
If X is discrete, then the set of values x such that $P(X = x) > 0$ is called the **support of** X.



In chpt 5, we'll look at continuous rvs that can take on any value in $\mathbb{R}$.



Def (**Prob Mass Function, PF**): The PF of a discrete rv X is given by $p_X(x) = P(X = x)$. Note this is positive if x is in support of X, and 0 otherwise.



Ex (**Coin Tosses Cont'd**). We'll find PFs of all rvs in the "Coin Tosses" example in §3.1 above, (involving two fair coin tosses).

- X, the # of Heads.

Since $X = 0$ if TT occurs, $X = 1$ if HT or TH occurs, and $X = 2$ if HH occurs, the PF of X is:

$$p_X(0) = P(X = 0) = \tfrac{1}{4},$$
$$p_X(1) = P(X = 1) = \tfrac{1}{2},$$
$$p_X(2) = P(X = 2) = \tfrac{1}{4},$$

and $p_X(x) = 0$ otherwise.

- $Y = 2 - X$, the # of Tails.

Reasoning as above or using the fact that $P(Y = y) = P(2 - X = y) = P(X = 2 - y) = p_X(2 - y)$, the PF of Y is:

$$p_Y(0) = P(Y = 0) = \tfrac{1}{4},$$
$$p_Y(1) = P(Y = 1) = \tfrac{1}{2},$$

$p_Y(2) = P(Y = 2) = \frac{1}{4}$,

and $p_Y(y) = 0$ otherwise.

Note that X and Y have the same PF (p_X and p_Y are the same function) even though X and Y are not the same rv (X and Y are different functions from $\{HH, HT, TH, TT\} \to \mathbb{R}$).

❗ Rvs concern themselves with what #s go w/different outcomes.

PFs concern themselves w/the frequency with which those #s occur.

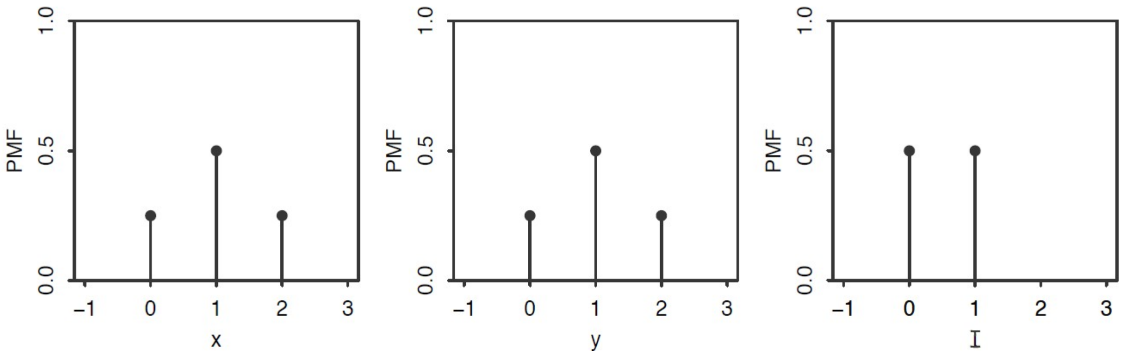
- I, the indicator of the first toss landing Heads.

Since $I = 0$ if TH or TT occurs, and 1 if HH or HT occurs, the PF of I is:

$p_I(0) = P(I = 0) = \frac{1}{2}$,

$p_I(1) = P(I = 1) = \frac{1}{2}$,

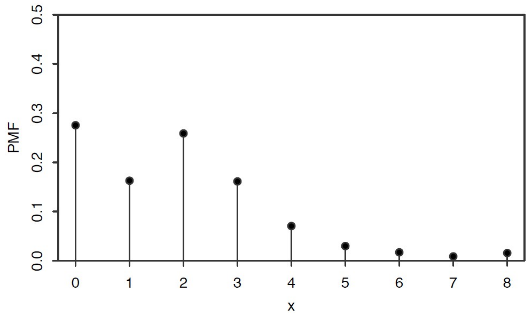
and $p_I(i) = 0$ otherwise.



PMFs for X, Y, I

Thm (Valid PFs). Let X be discrete w/support $x_1, x_2, \ldots$. The PF p_X must satisfy:

- Nonnegative: $p_X(x) > 0$ if $x = x_j$ for some x_j in the support, and $p_X(x) = 0$ otherwise,
- Sums to 1: $\sum_{j=1}^{\infty} p_X(x_j) = 1$.



Proof. First criterion is true since prob is always nonnegative.

Second is true since X must take on some value, and the events $\{X = x_j\}$ are disjoint, so:

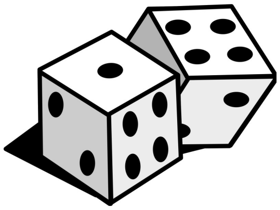
$\sum_{j=1}^{\infty} P(X = x_j) = P(\cup_{j=1}^{\infty} \{X = x_j\}) = P(X = x_1 \text{ or } X = x_2 \text{ or} \ldots) = 1.$ ■

Conversely, if distinct values $x_1, x_2, \ldots$ are specified and we have a function satisfying the two criteria above, then this function is the PF of *some* rv, we will show how to construct such a rv in Chapter 5.

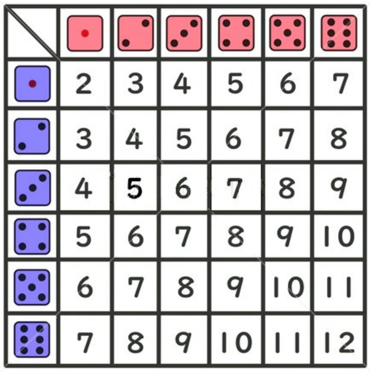
The PF is one way of expressing the distribution of a discrete rv. In particular:

Prob X is in a Set: Given a discrete X and a set $B \in \mathbb{R}$, if we know the PF of X we can find $P(X \in B)$, by summing up the heights of the vertical bars in B in the plot of the PF. (You'll do this in HW 4!)

Ex (**PF of Two Dice**): Let T be the sum of two fair die rolls. Assume we've already calculated the PF of T as:



$$P(T=2) = P(T=12) = \tfrac{1}{36},$$
$$P(T=3) = P(T=11) = \tfrac{2}{36},$$
$$P(T=4) = P(T=10) = \tfrac{3}{36},$$
$$P(T=5) = P(T=9) = \tfrac{4}{36},$$
$$P(T=6) = P(T=8) = \tfrac{5}{36},$$
$$P(T=7) = \tfrac{6}{36}.$$



	⚀	⚁	⚂	⚃	⚄	⚅
⚀	2	3	4	5	6	7
⚁	3	4	5	6	7	8
⚂	4	5	6	7	8	9
⚃	5	6	7	8	9	10
⚄	6	7	8	9	10	11
⚅	7	8	9	10	11	12

Sample Space

What's the prob that T is in the interval $[1,4]$?

So, T has support on $\{2,3,4\}$. We know the prob of these values from above, so:

$$P(1 \le T \le 4) = P(T=2) + P(T=3) + P(T=4) = \tfrac{6}{36}. \qquad \square$$

Harvard Video: youtube.com/watch?v=k2BB0p8byGA&list=PL2SOU6wwxB0uwwH80KTQ6ht66KWxbzTIo&index=9

What did we learn?

- Random Vars
- Discrete Rvs
- Prob Mass Functions, PFs
- Valid PFs