example (basketball, cont'd)

① $P(\text{win}) = p$

② $P(\text{win}) = p^3 + 3p^2(1-p)$

Compare $p^3 + 3p^2(1-p)$ to p

$p^2 + 3p(1-p)$ to 1

easy calc .... $p > p^3 + 3p^2(1-p)$ iff $p \in (0, \frac{1}{2})$

$\Rightarrow$ option 2 is for "better" players. //

<u>example</u> Monte Hall Game

You are on a game show facing 3 doors. Behind one of the doors is a car, behind 2 others are goats.

You choose door # 1.

Before opening door # 1, the host opens door # 2, there is a goat.

The host offers you a choice to change to # 3 or not.

① ② ③

? "goat" ?

Should you switch?

Let C_1, C_2, C_3 = events that car is behind door # 1, 2, or 3.

E = host opens door # 2, and a goat behind it

$$P(C_3 \mid E) =$$

$$= \frac{P(E \mid C_3)\, P(C_3)}{\sum_{i=1}^{3} P(E \mid C_i)\, P(C_i)} =$$

$$= \frac{1 \times \frac{1}{3}}{\frac{1}{2} \times \frac{1}{3} + 0 \times \frac{1}{3} + 1 \times \frac{1}{3}} = 2/3$$

$\Rightarrow$ <u>switch</u> ! //

<u>A</u> q<u>uick</u> <u>note</u> <u>on</u> <u>conditional</u> <u>probabilities</u>.

Let F s.t. $P(F) > 0$ then

axiom 1   1) $0 \le P(E \mid F) \le 1$;

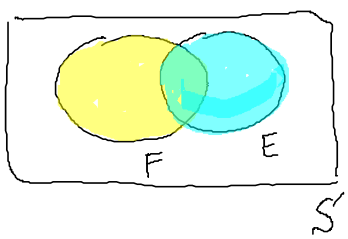
axiom 2   2) $P(S \mid F) = 1$;

axiom 2 2) $P(S|F) = 1$;

axiom 3 3) If E_i, $i = 1, 2, \ldots$, are mutially exclusive events then

$$P\left(\bigcup_{i=1}^{\infty} E_i \,|\, F\right) = \sum_{i=1}^{\infty} P(E_i|F).$$

Conditional probabilities "act" similarly to unconditional probabilities.



$S \longrightarrow \tilde{S} = F$

Now, all probabilities are relative frequencies of E within F.

Random Variables and their Distributions.

ex) Toss 2 fair dice,

May be interested in their sum rather than individual values on each die.

Def Real-valued functions defined on the sample space are random variables,

so

so

$$X : S \longrightarrow \mathbb{R}$$

random variable.

<u>example</u> 1) $S = \{1, 2, 3\}$

* X is the identity, so

$X(i) = i \Rightarrow$ random variable.

* $X(2) = X(3) = 5$

$X(1) = 1 \Rightarrow$ random variable.

2) 2 die example, sum

$$X(\{i, j\}) = i + j$$

$1 \le i \le 6,\ 1 \le j \le 6$

$\Rightarrow$ random variable.

<u>Important</u>: r.v. is a <u>function</u>!

For any $w \in S$ there exists only <u>one</u> $x \in \mathbb{R}$ s.t. $X(w) = x$.

<u>Convention</u>: We will use capitol letters for random variables, and script letters for specific values.

<u>example</u> 1) Suppose we toss 3 fair coins

$$S = \{HHH, HHT, \ldots\}$$

8 outcomes $= 2^3$.

Let $Y : S \to R$

$= \#$ of heads in the outcome.

Values of Y : 0, 1, 2, 3

$P(Y=0) = \frac{1}{8} = P(Y=3)$

$P(Y=1) = \frac{3}{8} = P(Y=2)$

$\underbrace{\sum_{i=0}^{3} P(Y=i) = 1}$

a way to check your computation!

HHH	$\to$ 3
HHT, HTH, THH	$\to$ 2
HTT, THT, TTH	$\to$ 1
TTT	$\to$ 0

---

2) Independendent tosses of a coin until a "head" occurs.

$P(H) = p \in (0,1)$

$X = \#$ of trials until first "head".

Values of X : $1, 2, 3, \ldots\ldots$, i.e. $\mathbb{N}$

$\underbrace{H\ldots}_{\downarrow\, 1}$, $\underbrace{TH\ldots}_{\downarrow\, 2}$, $\ldots\ldots$, $\underbrace{\underbrace{TTT\ldots T}_{k-1}H\ldots}_{\downarrow\, k}$, ...

$P(X=1) = p$

$P(X=2) = (1-p)\cdot p$

$P(X=3) = (1-p)^2 \cdot p$

$$P(X=3) = (1-p)^2 \cdot p$$

$\vdots$

$$P(X=k) = (1-p)^{k-1} \cdot p$$

$\vdots$

$$\sum_{k=1}^{\infty} P(X=k) = \sum_{k=1}^{\infty} (1-p)^{k-1} \cdot p$$

$$= p \cdot \frac{1}{1-(1-p)} = 1 \quad \checkmark$$

---

(3) Urn contains 11 balls:
3 white, 3 red, 5 black

A player wins \$1 each time a ball selected is white & looses \$1 if red ball is selected.

Suppose 3 balls are chosen w/o replacement.

Let X = winnings of the player.

Possible values of X : $0, \pm 1, \pm 2, \pm 3$

$$P(X=0) = P(\text{all black}) + P(R, W, B)$$

$$= \frac{\binom{5}{3}}{\binom{11}{3}} + \frac{\binom{3}{1}\binom{3}{1}\binom{5}{1}}{\binom{11}{3}}$$

$$= \frac{55}{165}$$

$$P(X=1) = P(2W, 1R) + P(1W, 2B)$$

$$= \frac{\binom{3}{2}\binom{3}{1}}{\binom{11}{3}} + \frac{\binom{3}{1}\binom{5}{2}}{\binom{11}{3}} =$$

$$= \frac{39}{165} = P(X = -1)$$

$$P(X = 2) = P(X = -2) = P(2W, 1B)$$

$$= \frac{\binom{3}{2}\binom{5}{1}}{\binom{11}{3}} = 15/165$$

$$P(X = 3) = P(X = -3) = P(3W) = \frac{1}{\binom{11}{3}} =$$

$$= 1/165$$

$$\sum_{x=-3}^{3} P(X = x) = \frac{55 + (39 + 15 + 1) \times 2}{165} = 1$$

//

Discrete Random Variables

Def A random variable that can take on at most countably many values is called discrete r.v. ("random variable");

we say r.v. has discrete distribution

(i.e. list of all possible values & their probabilities).

For a discrete random variable X

we define a probability mass function

(pmf) $p(x) = P(X = x)$ $(p : \mathbb{R} \to [0,1])$

Clearly, 1) $0 \leq p(x) \leq 1$;

2) $\displaystyle\sum_{x:\text{ all values of } X} p(x) = 1$;

3) If x is not a possible value of $X \Rightarrow p(x) = 0$.

---

examples: 1) Uniform distribution on $\{1, 2, \ldots, N\}$

All outcomes (or values $1, 2, \ldots, N$) are equally likely.

X is a r.v. takes values $1, 2, \ldots, N$

for any $1 \leq x \leq N$ (x integer)

$$P(X = x) = \frac{1}{N}.$$

Shorthand, $X \sim \text{Unif}\{N\}$

"distributed like" "parameter" (# of values here)

---

2) Binomial Distribution

Shorthand $X \sim \text{Bin}(N, p)$ parameters

X = number of successes in N trials, where the following holds:

1) Fixed (known in advance) # of trials, N;
2) Trials are independent;
3) Each trial results in "success" with probability p.

examples

a) A fair coin is tossed 10 times

X = # of heads among 10 tosses.

Yes: $N = 10$, $p = \frac{1}{2}$.

$\Rightarrow \quad X \sim \text{Bin}(10, \frac{1}{2})$.

b) # of phone calls received in one hour by a phone operator? No

c) Household with 5 adults

X = # among them that will vote for a democratic candidate in upcoming election.

No, most likely independence is not satisfied. //

X = # successes among N independent trials, $P(\text{"success"}) = p$.

Values of X : $\{0, 1, 2, \ldots, N\} \ni x$

pmf $\quad p(x) = P(X = x) = \binom{N}{x} p^x (1-p)^{N-x}$

$$\begin{array}{ccccc} 1 & 0 & 1 & \cdots & 1 \\ 1 & 2 & 3 & \cdots & N \\ \uparrow & \uparrow & \uparrow & & \uparrow \\ p & \times (1-p) & \times p & \cdots & \times p \end{array} = p^x \cdot (1-p)^{N-x}$$

- probability of a specific sequence with x successes.

- $\binom{N}{x}$ such sequences.

$$\sum_{x=0}^{N} P(X = x) = \sum_{x=0}^{N} \binom{N}{x} p^x (1-p)^{N-x}$$

$$\boxed{\text{Binomial formula: } (a+b)^N = \sum_{k=0}^{N} \binom{N}{k} a^k b^{N-k}}$$

$$= (p + (1-p))^N = 1$$

More important distributions later...

We first study common features useful for describing distributions.

Def. The *cumulative distribution function* (cdf) of a random variable X is

$F : \mathbb{R} \to [0, 1]$ such that

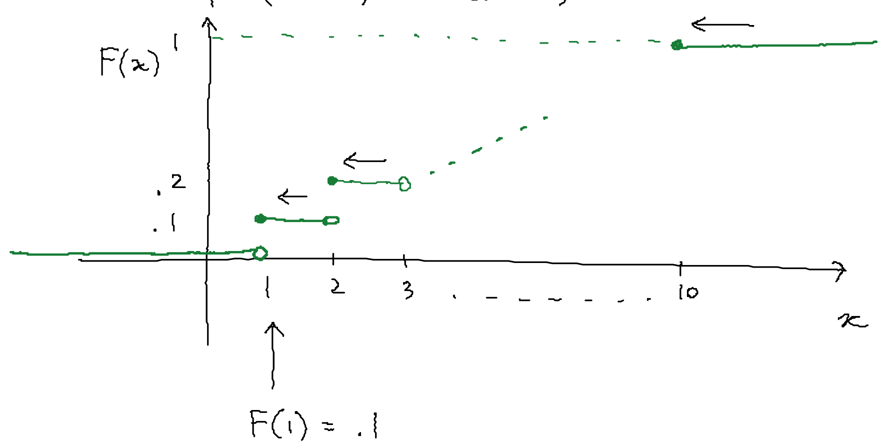
$$F(x) = P(X \le x)$$

<u>example</u> Consider $X \sim \text{Uniform}\{10\}$

$$F(x) = P(X \le x) = \sum_{y \le x} p(y)$$ pmf

$F(-1) = 0$, $F(2) = \frac{1}{5} = p(1) + p(2)$

$F(5.5) = \frac{1}{2}$, $F(10) = 1$



---

1) Cdfs completely describe distributions (unique desciption);

2) $F(x)$, cdf, is a nondecreasing function;

Let $x \ge y$ $\quad \{X \le x\} \supset \{X \le y\}$

$$F(x) = P(X \le x) \ge P(X \le y) = F(y).$$

3) $F(x)$ is a right continuous function

$$\lim_{y \downarrow x} F(y) = F(x) \quad \text{and}$$

$$\lim_{x \to -\infty} F(x) = 0 \;;\; \lim_{x \to +\infty} F(x) = 1.$$

Note : pmfs are defined for discrete r.v.'s only while cdfs are for all real-valued r.v.'s.

Continuous Distributions

Generally, random variables that can assume values in an (entire) interval, bounded or not, have continuous distributions; such random variables are continuous r.v.'s.

Def X has continuous distribution if there exists a non-negative function $f : \mathbb{R} \to [0, \infty)$ s.t. for any $A \subset \mathbb{R}$ (measurable)

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

(e.g. $A = [a, b)$; $P(X \in A) = P(a \leq X < b)$

(e.g. $A = [a,b)$; $P(X \in A) = P(a \le X < b)$

$$= \int_{[a,b)} f(x)\,dx = \int_a^b f(x)\,dx \ )$$

Such function f is called

probability density function (pdf)

(density)

1) $f(x) \ge 0$, all $x \in \mathbb{R}$

2) $\int_{\mathbb{R}} f(x)\,dx = \int_{-\infty}^{\infty} f(x)\,dx = 1$

---

example Uniform distribution on interval $[a,b]$

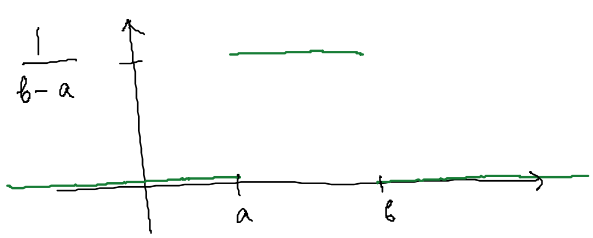
$X \sim \text{Unif}\,[a,b]$

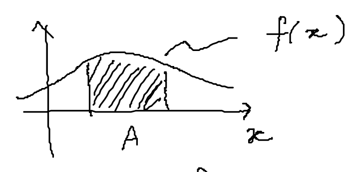
pdf, density $f(x) = \begin{cases} \text{const}, & a \le x \le b \\ 0, & \text{o.w.} \end{cases}$

("otherwize")

$1 = \int_a^b (\text{const})\,dx = \text{const}\,(b-a) \Rightarrow \text{const} = \frac{1}{b-a}$

$\Rightarrow \quad f(x) = \frac{1}{b-a}, \quad a \le x \le b.$





area $= \int_A f(x)\,dx$

Say $a = 0$, $b = 1$ (standard uniform)

$P\left(X \in \left(\frac{1}{2}, \frac{3}{4}\right)\right) = \int_{\frac{1}{2}}^{\frac{3}{4}} f(x)\,dx = \int_{\frac{1}{2}}^{3/4} 1\,dx = \frac{1}{4}$

$A = (.2, 1.3]$

$$P(X \in A) = P(.2 < X \le 1.3)$$

$$= \int_{.2}^{1.3} f(x)\,dx = \int_{.2}^{1} 1\,dx + \int_{1}^{1.3} 0\,dx$$

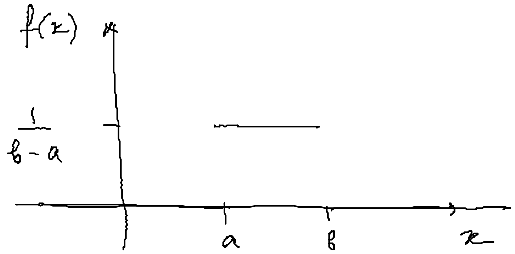
$$= .8$$

Fix any $x \in \mathbb{R}$, assume that X is a continuous r.v. then

$$P(X = \{x\}) = \int_{\{x\}} f(u)\,du$$

$$= \int_{x}^{x} f(u)\,du = 0, \quad x \in \mathbb{R}.$$

$X \sim Unif\ [a, b]$



cdf F?

$x < a, \quad F(x) = 0$

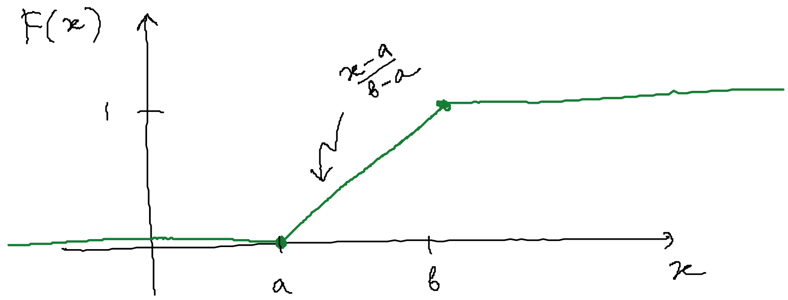
$x > b, \quad F(x) = 1$

$a \le x \le b$

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(u)\,du$$

$$= \int_{-\infty}^{a} 0\,du + \int_{a}^{x} \frac{1}{b-a}\,du$$

$$= 0 + \frac{u}{b-a}\Big|_a^x = \frac{x-a}{b-a}, \quad a \le x \le b.$$



Several Useful Identities

1) For any value $x \in \mathbb{R}$

$$P(X > x) = 1 - P(X \le x) = 1 - F(x)$$

2) For any $x_1 < x_2$

$$P(x_1 < X \le x_2) = P(X \le x_2) - P(X \le x_1)$$

$$= F(x_2) - F(x_1)$$

Unless it is clearly stated that the distribution is continuous, we need to be careful about the endpoints.

<u>example</u> $X \sim \text{Unif}\{5\}$ discrete

$$p(x) = P(X = x) = \underbrace{\frac{1}{5}, \quad x = 1, 2, \ldots, 5}$$

(0, o.w. → will not write it but assume it)

$$F(x) = \begin{cases} 0, & x < 1 \\ k/5, & k \le x < k+1, \quad k = 1, 2, 3, 4 \end{cases}$$

$$F(x) = \begin{cases} 0, & x < 1 \\ k/5, & k \leq x < k+1, \quad k = 1, 2, 3, 4 \\ 1, & x \geq 5 \end{cases}$$

$$p(x) = P(X = x) = F(x) - \lim_{y \to x-} F(y)$$

$$p(2) = \frac{2}{5} - \frac{1}{5} = \frac{1}{5}$$

$$p(3.5) = \frac{3}{5} - \frac{3}{5} = 0$$

//
