Averages.

$1, 2, 3, 4, 5, 6$ Mean $= \dfrac{1+2+3+4+5+6}{6}$

$= 3.5$

$1, 1, 1, 1, 3, 4, 6, 6$ $\dfrac{1+1+1+1+3+4+6+6}{8} = \dfrac{23}{8}$

OR $\dfrac{4}{8} \times 1 + \dfrac{1}{8} \times 3 + \dfrac{1}{8} 4 + \dfrac{2}{8} \times 6$

Mean as a weighted sum.

weights are relative frequencies of each value.

$$E(X) = \sum_x x\, P(X = x)$$

"Expectation of X"
"Expected value"

$X \sim \text{Bernoulli}(p)$

$$E(X) = 1 \times P(X=1) + 0 \times P(X=0)$$
$$= 1 \times p + 0$$
$$= p$$

Indicator RV.

$$X = \begin{cases} 1 & \text{if event } A \text{ occurs} \\ 0 & \text{otherwise} \end{cases}$$

X is an indicator variable for the event A.

$$E(X) = P(A)$$

the probability of the event is equivalent to the expected value of a suitably chosen indicator RV

$X \sim \text{Binomial}(n, p)$.

$$E(X) = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k}$$

$$= \sum_{k=0}^{n} n \binom{n-1}{k-1} p^k (1-p)^{n-k}$$

[choosing a committee of size k with 1 person designated as chair.

$$= np \sum_{k=0}^{n} \binom{n-1}{k-1} p^{k-1} q^{n-k}.$$

$$= np \sum_{k=1}^{n} \binom{n-1}{k-1} p^{k-1} q^{n-k}.$$

[$\binom{n}{k} = 0$ for $k < 0$

let $j = k-1$

$$= np \underbrace{\sum_{j=0}^{n-1} \binom{n-1}{j} p^j q^{n-1-j}}_{\text{Binomial Theorem } (p+q)^{n-1} = 1}$$

$$E(X) = np$$

<u>Linearity of Expectation.</u>

$$E(X+Y) = E(X) + E(Y)$$

even if X and Y are not independent.

$$E(cX) = c\,E(X)$$

$X \sim \text{Binomial}(n, p)$

Sum of n Bernoulli(p) RV

$$X = X_1 + X_2 + \ldots + X_n$$

$$E(X) = E(X_1 + X_2 + \ldots + X_n)$$
$$= E(X_1) + E(X_2) + \ldots + E(X_n)$$
$$= p + p + \ldots + p$$
$$= np$$

Hypergeometric

$$E(x) = \sum_{k=0}^{t} k \frac{\binom{t}{k}\binom{d}{n-k}}{\binom{t+d}{n}}$$

5 cards $\quad k$ # aces.

X_j $\quad j^{th}$ card is an ace. $\quad \leftarrow$ indicator variable.

$$X = \sum_{j=1}^{5} X_j$$

$$\begin{aligned} E(x) &= E(X_1 + X_2 + X_3 + X_4 + X_5) \\ &= E(X_1) + E(X_2) + \cdots E(X_5) \\ &= 5E(X_1) \\ &= 5 \times \frac{4}{52} \end{aligned}$$

$$E(x) = \frac{5}{13}$$

In this case the X_j's are <u>not</u> independent

However, <u>before</u> we look at any of the cards, we have no reason to think that the 1st card has a different distribution than any of the others.

Generalize

Expected value of a Hypergeometric $= np$

Geometric Distribution.

$X \sim \text{Geometric}(p)$

Series of independent Bernoulli (p) trials.

count the number of failures before the 1st success.

(don't include the success)

PMF: $P(X = k)$.

T T T T T H $\qquad q^5 p$

$$P(X = k) = q^k p \qquad k \in \{0, 1, 2, 3, \dots\}$$

Is this a valid PMF?

1) $P(X = k) \geq 0$

2) $\sum_k P(X = k) = 1$.

$$\sum_{k=0}^{\infty} q^k p = p \underbrace{\sum_{k=0}^{\infty} q^k}_{\text{sum of a Geometric series}} = \frac{p}{1-q} = \frac{p}{p} = 1$$

Expected Value.

$$E(X) = \sum_{k=0}^{\infty} k \times p q^k$$

$$= p \sum_{k=0}^{\infty} k q^k$$

Start from $\displaystyle\sum_{k=0}^{\infty} q^k = (1-q)^{-1}$

Take derivatives wrt q.

$$\sum_{k=0}^{\infty} k q^{k-1} = +(1-q)^{-2}.$$

multiply by q.

$$\sum_{k=0}^{\infty} k q^k = \frac{q}{(1-q)^2}$$

$$E(x) = p \sum_{k=0}^{\infty} k q^k$$

$$= \frac{pq}{(1-q)^2} = \frac{pq}{p^2} = \frac{q}{p}$$

Alternatively

consider the 1st flip

if it's a H, then we have no failures.
$X = 0$ with prob p

if it's a T, then we have 1 failure, and we restart the problem. This happens with prob q.

$$E(X) = p \times 0 + q(1 + E(x)).$$

$$E(x) = q(1 + E(x)).$$

$$E(x) = q + q\,E(x).$$

$$E(x)(1 - q) = q$$

$$E(x) = \frac{q}{1-q} = \frac{q}{p}$$

Linearity of Expectation.

$$T = X + Y$$

$$E[T] = E[X] + E[Y]$$

even if X and Y are not independent

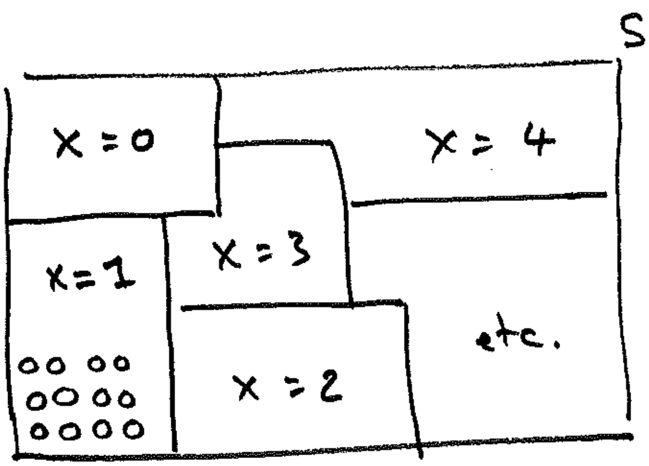
1, 1, 1, 1, 2, 2, 4, 5

$$= \frac{1+1+1+1+2+2+4+5}{8}$$

$$= \frac{4}{8} \times 1 + \frac{2}{8} \times 2 + \frac{1}{8} \times 4 + \frac{1}{8} \times 5$$

Grouped and ungrouped sums

sum over individual elements of the state space.



$$E[X] = \sum_x x P(X = x)$$

$$= \sum_s X(s) P(s).$$

sum ~~of~~ over individual elements of the state space.

$$E[T] = \sum_s (X+Y)(s)\, P(s)$$

$$= \sum_s (X(s) + Y(s))\, P(s)$$

$$= \sum_s X(s) P(s) + \sum_s Y(s) P(s)$$

$$= E(X) + E(Y)$$

Hence.

$$E(X+Y) = E(X) + E(Y)$$

$$E(cX) = c\,E(X)$$

Extreme case of dependence. $X = Y$

$$E(X+Y) = E(2X) = E(X) + E(Y)$$

So linearity of ~~independence~~ expectation holds even when X, Y are not independent.

Negative Binomial Distribution

Generalization of the Geometric Distribution.

Series of independent Bernoulli(p) trials.

failures before the r^{th} success.

$r = 5$

0001 01 0000101001 $\quad r = 5$

$n = 11$ failures.

$$P(X=n) = \underbrace{\binom{n+r-1}{r-1} p^{r-1} q^{n}}_{\substack{r-1 \text{ successes in } 1^{st} \\ n+r-1 \text{ trials}}} \times \underbrace{p}_{\substack{\text{success in} \\ n+r^{th} \text{ trial.}}}$$

$$= \binom{n+r-1}{r-1} p^{r} q^{n}$$

$E[X]$

$r = 1 \rightarrow$ Geometric (p) $\qquad E(\,) = \dfrac{q}{p}$

$r = 2. \rightarrow$ wait for the 1^{st} success,

then wait for the 2^{nd} success.

$$X = X_1 + X_2 + X_3 + \dots X_r$$

$$E(X) = E(X_1 + X_2 + \dots X_r).$$
$$= E(X_1) + E(X_2) + \dots E(X_r)$$

Each $X_i \sim$ Geometric (p)

$$E(X) = r\,E(X_1)$$
$$= r\frac{q}{p}$$

<u>1st success distribution</u>.

trials to the 1st success, including the success.

$$X \sim FS(p).$$

$$Y = X - 1.$$

$$Y \sim \text{Geometric}(p)$$

$$E[X] = E[Y] + 1$$
$$= \frac{q}{p} + 1 = \frac{1}{p}$$

Game.

Toss a coin repeatedly, wait for 1st H.

H on 1st trial	win \$2.
2nd	\$ 4
3rd	\$ 8
⋮	
xth	\$ 2^x

where x is # flips to 1st head, including the 1st head.

How much are you willing to pay to play this game?

Fair: what price would make the expected return zero?

$$Y = 2^x \qquad E[Y]$$

$$E[Y] = \sum_{k=1}^{\infty} 2^k \times \frac{1}{2^k} = \sum_{k=1}^{\infty} 1 = \infty$$

you should be willing to pay me an infinite amount of money.

What if I don't have an infinite amount of money to give you?

Let's say I have \$1T $\approx 2^{40}$

so if $x \geq 40$, I give you all the \$1T

$$E[Y] = \sum_{k=1}^{40} 2^k \times \frac{1}{2^k} + \sum_{k=41}^{\infty} \frac{2^{40}}{2^k}$$

$$= 41$$