Averages.

\[1, 2, 3, 4, 5, 6\]

Mean \[= \frac{1+2+3+4+5+6}{6} = 3.5\]

\[1, 1, 1, 1, 3, 4, 6, 6\]

Mean as a weighted sum.

Weights are relative frequencies of each value.

\[E(x) = \sum_{x} x \cdot P(x=x)\]

"Expectation of X"

"Expected value"

\[X \sim \text{Bernoulli}(p)\]

\[E(x) = 1 \cdot P(x=1) + 0 \cdot P(x=0) = 1 \cdot p + 0 = p\]
Indicator RV.

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

\[ X \text{ is an indicator variable for the event } A. \]

\[ E(X) = P(A) \quad \text{the probability of the event is equivalent to the expected value of a suitably chosen indicator RV} \]

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

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

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

\[ = np \sum_{k=0}^{n} \binom{n-1}{k-1} p^{k-1} q^{n-k} \]

\[ = np \sum_{k=1}^{n} \binom{n-1}{k-1} p^{k-1} q^{n-k} \]

\[ \left\{ \begin{array}{ll} \binom{n}{k} = 0 & \text{for } k < 0 \\ \end{array} \right. \]
Let $j = k - 1$

$$E(x) = np$$

---

**Linearity of Expectation**

$$E(x + y) = E(x) + E(y)$$

$$E(cX) = cE(X)$$

---

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

*Sum of $n$ Bernoulli $(p)$ R.V.*

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

$$E(x) = E(X_1 + X_2 + \ldots + X_n)$$

$$= E(X_1) + E(X_2) + \ldots + E(X_n)$$

$$= np + np + \ldots + np$$

$$= np$$
Hypergeometric

\[ E(X) = \sum_{k=0}^{t} \binom{t}{k} \binom{d}{n-k} \frac{(t-k)}{(t+d)} \]

5 cards \( k \neq \text{aces} \).

\( X_j \) \( j^{\text{th}} \) card is an ace. \( \leftarrow \) indicator variable.

\[ X = \sum_{j=1}^{5} X_j \]

\[ E(X) = E(X_1 + X_2 + X_3 + X_4 + X_5) \]
\[ = E(X_1) + E(X_2) + \cdots + E(X_5) \]
\[ = 5E(X_1) \]
\[ = 5 \times \frac{4}{13} \]

\[ E(X) = \frac{5}{13} \]

In this case the \( X_j \)'s are not independent.

However, before 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 is \( np \)

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)

\[
\begin{align*}
\text{PMF:} & \quad P(X = k) = q^k p \\
\text{T T T T T T H} & \quad q^5 p \\
\text{P}(X = k) & = q^k p \\
& \quad k \in \{0, 1, 2, 3, \ldots \}
\end{align*}
\]

Is this a valid PMF?

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

2) \( \sum_{k} P(X = k) = 1 \)
\[
\sum_{k=0}^{\infty} q^k p = p \sum_{k=0}^{\infty} q^k = \frac{p}{1-q} = \frac{1}{p} = 1
\]

*Sum of a geometric series*

**Expected Value.**

\[
E(X) = \sum_{k=0}^{\infty} k \cdot p \cdot q^k
\]

\[
= p \sum_{k=0}^{\infty} k \cdot q^k
\]

*Start from*

\[
\sum_{k=0}^{\infty} q^k = \frac{1}{1-q}
\]

*Take derivatives wrt q.*

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

*Multiply by q.*

\[
\sum_{k=0}^{\infty} kq^k = \frac{q}{(1-q)^2}
\]
\[ E(x) = p \sum_{k=0}^{\infty} kq^k \]
\[ = \frac{pq}{(1-q)^2} = \frac{pq}{\rho^2} = \frac{q}{\rho} \]

Alternatively, consider the 1st flip:

If it's a H, then we have no failures.

\[ x = 0 \quad \text{with prob} \quad p \]

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

\[ E(x) = px \cdot 0 + q(1 + E(x)) \]
\[ E(x) = q(1 + E(x)) \]
\[ E(x) = q + qE(x) \]
\[ E(x)(1-q) = q \]
\[ E(x) = \frac{q}{1-q} = \frac{q}{p} \]
Linearity of Expectation.

\[ T = X + Y \]

Even if \( X \) and \( Y \) are not independent

\[ E[X+Y] = E[X] + E[Y] \]

\[ 1, 1, 1, 1, 2, 2, 4, 5 \]

\[ = \frac{1+1+1+1+2+2+4+5}{8} \]

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

\[ \uparrow \]

Grouped and ungrouped sums

Sum over individual elements of the state space.

\[ E[X] = \sum_{x} x P(x=x) \]

\[ = \sum_{s} x(s) P(s) \]

\[ \uparrow \]

Sum of over individual elements of the state space.
\[
E[T] = \sum_s (X(s) + Y(s)) P(s)
\]
\[
= \sum_s X(s) P(s) + \sum_s Y(s) P(s)
\]
\[
= E(X) + E(Y).
\]

Hence,
\[
E(X + Y) = E(X) + E(Y)
\]
\[
E(cX) = c E(X)
\]

Extremal case of dependence. \(X = Y\)
\[
E(X + Y) = E(2X) = E(X) + E(Y).
\]

So linearity of 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 0000 10 10 01

$n = 11$ failures.

$P(X = n) = \binom{n+r-1}{r-1} p^{r-1} q^{n} \times p$

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

\[\text{r-1 successes in 1st} \]
\[\text{n+r-1 trials}\]

\[\text{Success in (n+r)th trial.}\]

$E[X]$

$r = 1 \rightarrow \text{Geometric (p)} \quad E() = \frac{q}{p}$

$r = 2 \rightarrow \text{wait for the 1st success, then wait for the 2nd success.}$
\[ X = X_1 + X_2 + X_3 + \ldots + X_r \]

\[ E(x) = E(X_1 + X_2 + \ldots + X_r) = E(X_1) + E(X_2) + \ldots + E(X_r) \]

Each \( X_i \sim \text{Geometric}(\rho) \)

\[ E(X) = r E(X_1) = \frac{rq}{\rho} \]

1st success distribution.

\# trials to the 1st success, including the success.

\( X \sim FS(\rho) \).

\( Y = X - 1 \).

\( Y \sim \text{Geometric}(\rho) \)

\[ E[X] = E[Y] + 1 \]

\[ = \frac{q}{\rho} + 1 = \frac{1}{\rho} \]
Game.

Toss a coin repeatedly, wait for 1st H.

1st trial win $2.
2nd $4
3rd $8

$2^x$ 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?

$Y = 2^x$ \hspace{1cm} 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 can't have an infinite amount of money to give you?

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

so if $x = 2^{40}$, I give you all the $1T$

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

$$= 41$$