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

**Linearity of Expectation**

\[ T = X + Y \]

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

True even if \( X, Y \) are not independent.

\[ 1, 1, 1, 2, 2, 4, 5 \quad \Rightarrow \quad \frac{1 + 1 + 1 + 2 + 2 + 4 + 5}{7}. \]

\[ \Rightarrow \quad \frac{3}{7} \times 1 + \frac{2}{7} \times 2 + \frac{1}{7} \times 4 + \frac{1}{7} \times 5 \]

**Proof** is by summing over individual elements of the state-space.

\[ E[X] = \sum x P(X = x) \]

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

**Grouped**

\[ E[T] = \sum_{s} (X(s) + Y(s)) P(s) \]

\[ = \sum_{s} (X(s) + Y(s)) P(s) \]
\[ = \sum_{s} X(s) P(s) + \sum_{s} Y(s) P(s) \]

\[ = E(X) + E(Y). \]

\[ E(x+y) = E(x) + E(y). \] (for discrete case)

Similarly \[ E(cx) = c E(x) \]

Extreme case of dependence \[ X = Y \]

\[ E(x+y) = E(2x) = E(x) + E(y). \]

---

**Negative Binomial Distribution.**

Generalization of geometric distribution.

Independent Bernoulli(p) trials.

# failures before the r-th success.

\[ \begin{array}{cccccccc}
1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\
\end{array} \]

\[ r = 5 \]

\[ n = 9 \] failures.

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

\[ = \binom{n+r-1}{r-1} p^{r} (1-p)^n. \]
1st \( n+r-1 \) trials must have
\( n \) failures
\( r-1 \) successes
in any order.

Last trial must be a success.

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

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

\( r = 2 \) \quad \text{wait for 1st success}
then wait for 2nd success.

\[
X = X_1 + X_2 + \ldots + X_r
\]

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

\( X_j \sim \text{Geometric (p)} \)

\[
E(X) = r \cdot E(X_1)
\]

\[= \frac{rq}{p}\]

\[\text{\# trials (i.e. including failures)}
\]

\[\text{Here - only counting successes}
\]

\[\text{sometimes N.B. if}
\]

\[\text{defined over}
\]

\[\text{\# trials (i.e. including successes)
\]

\[\text{the failures}.
\]
$X \sim FS(p)$ - $1^{st}$ success distribution

$= \text{trials to the 1}^{st} \text{ success including the success.}$

$Y = X - 1$

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

$E(X) = E(Y) + 1$

$= \frac{q}{p} + 1 = \frac{1}{p}$

- thus in reasonable

  intuitive

  eg 2 tosses before to

  get 1st H on

  fair coin.

---

**St. Petersburg Paradox**

Flip fair coin repeatedly until 1st H.

If H on 1st trial - win $2.

2nd

3rd

\[= x^{n} \]

\[= \$2^{x} \]

\[\text{where } x \text{ is #flips to 1st H, including the 1st H.} \]

\[Q. \text{ How much would you be willing to pay to play this game?} \]
Fari: what price would make the expected return zero?

\[ Y = 2^x \quad \text{Find } E(Y) \]

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

\( (k-1) \text{ tails followed by } H. \)

What happens if the person you are playing against only has \$1T

\[ \$10^{12} < 2^{40} \]

If \( x \geq 40 \), I give you \$1T

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

\[ = 41 \]
Poisson Distribution.

\( X \sim \text{Pois} (\lambda) \quad \lambda \) "rate parameter"

\( \lambda > 0 \)

\[ P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!} \quad \text{P.M.F.} \]

Is this a valid PMF?

\[ P(X = k) > 0 \]

\[ \sum_{k=0}^{\infty} \frac{e^{-\lambda} \lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} e^\lambda = 1 \]

Taylor series for \( e^\lambda \)

\[ E[X] = \sum_{k=0}^{\infty} k \frac{e^{-\lambda} \lambda^k}{k!} \]

\[ = e^{-\lambda} \sum_{k=0}^{\infty} \frac{k \lambda^k}{k!} \]

\[ = e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^k}{k!} \]

\[ = e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \]
\[ = e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} \]

let \( j = k - 1 \)

\[ = e^{-\lambda} \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} \]

\[ = e^{-\lambda} \lambda e^\lambda = \lambda \]

\[ E[X] = \lambda. \]

Why is the Poisson Useful?

- Applications where we're counting the number of "successes" where there are a large number of trials and each trial has a small probability of success.
eg. # emails received in an hour.
    - lots of people could send you an email
    - prob. that a specific person will send you an email is small.

# memory errors on a memory chip on a spacecraft.
    - lots of sub-atomic particles hit the memory
    - chance that any one particle will cause an error is small.

# chocolate chips in a cookie.
    - lots of bits of dough
    - each has a small chance to be a chocolate chip.

In each case the Poisson distribution is a model.
    - it is not exact.

↑

important concept when applying probability to the real world.
**Poisson Approximation**

Events $A_1, A_2, \ldots, A_n$ with $P(A_j) = p_j$ for $j = 1, 2, \ldots, n$, where $n$ is large and $p_j$ is small.

Events are independent or "weakly dependent" if $A_j$'s occur in approximately $\text{Poisson}(\lambda)$,

$$\lambda = \sum_{j=1}^{n} p_j.$$

If $A_j$'s are exactly independent, and $p_j = p$ for all $j$,

$$X = \text{Bin}(n, p) \quad \text{let } n \to \infty,$$

$$p \to 0,$$

$$\lambda = np \text{ stays constant}.$$

$$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

$$= \frac{n(n-1)(n-2)\ldots(n-k+1)\left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k}}{k!}.$$

$\lambda \to p = \frac{\lambda}{n}$. 

\text{Poisson} even if $A_j$'s are not all same or even if $A_j$'s are not completely independent.
\[ \text{take limit } \lim_{n \to \infty} \]

\[ n \cdot (n-1) \cdot (n-2) \ldots (n-k+1) \to n^k \]

\[ (1 - \frac{\lambda}{n})^{n-k} = (1 - \frac{\lambda}{n})^n \cdot (1 - \frac{\lambda}{n})^{-k} \]

\[ (1 - \frac{\lambda}{n})^{-k} \to 1 \]

\[ (1 - \frac{\lambda}{n})^n \to e^{-\lambda} \]

\[ (1 + \frac{\lambda}{n})^n \to e^\lambda \]

\[ P(X = k) = \frac{n^k \cdot \lambda^k}{k!} \cdot \frac{e^{-\lambda}}{n^k} \to \frac{e^{-\lambda} \cdot \lambda^k}{k!} \]

\[ P(X = k) = \frac{e^{-\lambda} \cdot \lambda^k}{k!} \]

Poisson distribution is the limit of Binomial as \( n \to \infty \)

\( p \to 0 \)

in such a way that the expected number of successes stays fixed (\( = \lambda \)).
$10^6$ particles incident on the chip.

A parity-check so that $b_i$th is even.

<table>
<thead>
<tr>
<th>0</th>
<th>0</th>
<th>1</th>
<th>0</th>
<th>1</th>
<th>0</th>
<th>1</th>
<th>1</th>
<th>0</th>
</tr>
</thead>
</table>

Flipping 1 0 1 0 1 0 1 1 0 $\leftarrow$ odd # of 1's.

Error correcting codes.

Shannon's theorem.

---

Binomial $\left(10^6, \frac{1}{10^{-10}}\right) \approx$ Poisson $(10^{-2})$

Makes practical problems easier (numerically).