Uniform
Assume that $Pr[X = i] = 1/n$ for $i \in \{1, \ldots, n\}$. Then

$E[X] = \sum_{i=1}^{n} i \times Pr[X = i] = \frac{1}{n} \sum_{i=1}^{n} i$

$= \frac{1}{n} \cdot \frac{n(n+1)}{2} = \frac{n+1}{2}$

Also,

$E[X^2] = \sum_{i=1}^{n} i^2 \times Pr[X = i] = \frac{1}{n} \sum_{i=1}^{n} i^2$

$= \frac{1}{6} (1 + 3n + 2n^2)$, as you can verify.

This gives

$var(X) = \frac{1}{6} (1 + 3n + 2n^2) - \left(\frac{n+1}{2}\right)^2 = \frac{n^2}{3} - 1$.

Variance: binomial.

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

$= \frac{n}{n} \binom{n}{1} p^1 (1-p)^{n-1} + \frac{1}{n} \binom{n}{2} p^2 (1-p)^{n-2}$

$= \frac{n+1}{n} \binom{n}{2} p^2 (1-p)^{n-2}$

$E[X^2] = \sum_{i=0}^{n} i^2 \cdot \binom{n}{i} p^i (1-p)^{n-i}$

$= \frac{n+1}{n} \binom{n}{2} p^2 (1-p)^{n-2}$

$E[X^2] = E[X]$ when $X = 1 \times Pr[X = 1] = 1 + 0 \times Pr["\text{anything else}"]$

$= \frac{n+1}{n} \binom{n}{2} p^2 (1-p)^{n-2}$

$\Rightarrow \sigma(X) \neq E[|X - E[X]|]$.

Exercise: How big can you make $\frac{E[X]}{E[|X - E[X]|]}$?

Fixed points.

Number of fixed points in a random permutation of $n$ items.

$X =$ Number of student that get assignment back.

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

where $X_i$ is indicator variable for $i$th student getting assignment back.

$E[X] = \sum_{i=0}^{n} E[X_i] = \sum_{i=0}^{n} \frac{1}{n} \times \frac{1}{n}$

$= 1/n + (n-1)/n \times 1/(n-1) + \cdots + 1/n \times 1/n = 1 + 1 = 2.$

$E[X^2] = E[X] = 1/n$

$E[X|X_i] = 1 \times Pr[X = 1] = 1 + 0 \times Pr["\text{anything else}"]$

$= \frac{1+1}{(n-1)}$.

$var[X] = E[X^2] - E[X]^2 = 2 - 1 = 1.$

Slide animation by Satish Rao
Properties of variance.

1. \( \text{var}[cX] = c^2 \text{var}[X] \), where \( c \) is a constant.
   Scales by \( c^2 \).
2. \( \text{var}[X + c] = \text{var}[X] \), where \( c \) is a constant.
   Shifts center.

Proof:

\[ \text{var}[cX] = E[(cX)^2] - (E[cX])^2 = c^2E[X^2] - c^2(E[X])^2 = c^2(\text{var}[X]) \]

\[ \text{var}[X + c] = E[(X + c)^2] - (E[X + c])^2 = E[X^2] + 2cE[X] + c^2 - (E[X] + c)^2 = \text{var}[X] \]

Variance of sum of independent random variables

**Theorem:**

If \( X \) and \( Y \) are independent, then

\[ \text{var}[X + Y] = \text{var}[X] + \text{var}[Y] \]

**Proof:**

Since shifting the random variables does not change their variance, let us subtract their means.

That is, we assume that \( E[X] = 0 \) and \( E[Y] = 0 \).

Then, by independence,

\[ E[XY] = E[X]E[Y] = 0. \]

Hence,

\[ \text{var}[X + Y] = E[(X + Y)^2] = E[X^2 + 2XY + Y^2] \]


Variance of geometric distribution.

\( X \) is a geometrically distributed RV with parameter \( p \).

Thus, \( P[X = n] = (1 - p)^{n-1}p \) for \( n \geq 1 \). Recall \( E[X] = 1/p \).

\[ E[X^2] = p + 4p(1-p) + 9p(1-p)^2 + ... \]

\[ = \frac{p}{(1-p)^2} \]

\[ pE[X^2] = p + 3p(1-p) + 3p(1-p)^2 + ... \]

\[ = \frac{1}{1-p} \]

\[ pE[X^2] = \frac{1}{1-p} \]

\[ \Rightarrow E[X] = \frac{2}{p} - 1 = \frac{2-p}{p} \]

\( \Rightarrow \text{var}[X] = E[X^2] - E[X]^2 = \frac{2-p}{p} - \frac{2}{p} = \frac{1-p}{p} \).

\( \sigma(X) = \sqrt{\frac{1-p}{p}} \approx E[X] \) when \( p \) is small(ish).

Variance of Binomial Distribution.

Flip coin with heads probability \( p \).

**X** - how many heads?

\[ X_i = \begin{cases} 1 & \text{if } \text{ith flip is heads} \\ 0 & \text{otherwise} \end{cases} \]

**Note:** \( p = 0 \Rightarrow \text{var}[X] = 0 \). Also, \( p = 1 \Rightarrow \text{var}[X] = 0 \)

Now,

\[ X = X_1 + X_2 + ... + X_n \]

\( X_i \) and \( X_j \) are independent. Hence,

\[ \text{var}[X] = \text{var}[X_1] + ... + \text{var}[X_n] = n \times \text{var}[X_1] = np(1-p). \]

Coupon Collectors Problem.

**Experiment:** Get coupons at random from \( n \) until collect all \( n \) coupons.

**Outcomes:** \{12345...n\} \( n \times \)

**Random Variable:** \( X \) - length of outcome.

Before: \( P[X \geq n \ln 2n] \leq \frac{1}{2} \).

Today: \( E[X] \)?

\[ E[X^2] = E[(1 + ZY)^2] = E[1 + 2ZY + Z^2Y^2] \]

\[ = 1 + 2qE[Y] + qE[Y^2] = 1 + 2q/p + qE[X^2] \]

Thus, \( pE[X^2] = 1 + 2q/p = (p + 2q)/p = (2-p)/p \), so that \( E[X^2] = (2-p)/p^2 \).

Hence,

\[ \text{var}[X] = E[X^2] - E[X]^2 = E[X^2] - 1/p^2 = (1-p)/p^2. \]
Time to collect coupons

$X$-time to get $n$ coupons.
$X_1$ - time to get first coupon. Note: $X_1 = 1$. $E(X_1) = 1$.
$X_2$ - time to get second coupon after getting first.

$Pr[\text{\textquoteleft get second coupon\textquoteright } | \text{\textquoteleft got milk first coupon\textquoteright } ] = \frac{n-1}{n}$
$E(X_2)$? Geometric ! ! ! $E(X_2) = \frac{1}{p} = \frac{1}{\frac{n-1}{n}}$.

$Pr[\text{\textquoteleft get i-th coupon\textquoteright } | \text{\textquoteleft got i-1st coupons\textquoteright } ] = \frac{n-(i-1)}{n}$

$E(X_i) = \frac{1}{p} = \frac{n}{n-i+1}, i = 1,2,\ldots,n.$

$E[X] = E[X_1] + \cdots + E[X_n] = \frac{n}{n} + \frac{n}{n-1} + \frac{n}{n-2} + \cdots + \frac{n}{1}$

$= n(1 + \frac{1}{2} + \cdots + \frac{1}{n}) = nH(n) \approx n(\ln(n) + \gamma)$

Review: Harmonic sum

$H(n) = 1 + \frac{1}{2} + \cdots + \frac{1}{n} \approx \int_1^n \frac{1}{x} \, dx = \ln(n)$.

A good approximation is

$H(n) \approx \ln(n) + \gamma$ where $\gamma \approx 0.58$ (Euler-Mascheroni constant).

Harmonic sum: Paradox

Consider this stack of cards (no glue!):

If each card has length 2, the stack can extend $H(n)$ to the right of the table. As $n$ increases, you can go as far as you want!

Paradox

par\-a\-do\-x

/ para daks/

noun

a statement or proposition that, despite sound (or apparently sound) reasoning from acceptable premises, leads to a conclusion that seems sensible, logically unacceptable, or self-contradictory.

"a potentially serious conflict between quantum mechanics and the general theory of relativity known as the information paradox"

- a seemingly absurd or self-contradictory statement or proposition that when investigated or explained may prove to be well founded or true.

"in a paradox, he has discovered that stepping back from his job has increased the rewards he gains from it"

synonyms: contradiction, contradiction in terms, self-contradiction, inconsistency.

- a situation, person, or thing that combines contradictory features or qualities.

"the mingling of deciduous trees with elements of desert flora forms a fascinating ecological paradox"

Stacking

The cards have width 2. Induction shows that the center of gravity after $n$ cards is $H(n)$ away from the right-most edge.

Simeon Poisson

The Poisson distribution is named after:

Siméon Denis Poisson (1781–1840)
Equal Time: B. Geometric

The geometric distribution is named after:

![B. Geometric](image)

I could not find a picture of D. Binomial, sorry.

Poisson Distribution: Definition and Mean

Definition: Poisson Distribution with parameter \( \lambda > 0 \)

\[
X = P(\lambda) \iff Pr\{X = m\} = \frac{\lambda^m}{m!} e^{-\lambda}, m \geq 0.
\]

Fact: \( E[X] = \lambda \).

Proof:

\[
E[X] = \sum_{m=1}^{\infty} m \cdot \frac{\lambda^m}{m!} e^{-\lambda} = e^{-\lambda} \sum_{m=1}^{\infty} \frac{\lambda^m}{(m-1)!}
\]

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

Poisson Distribution: Variance

Recall: \( X = P(\lambda) \iff Pr\{X = n\} = \frac{\lambda^n}{n!} e^{-\lambda}, n \geq 0. \)

Fact: \( \text{var}[X] = \lambda. \) One finds

\[
E[X(X - 1)] = \sum_{n=0}^{\infty} n(n-1) \frac{\lambda^n}{n!} e^{-\lambda} = \lambda e^{-\lambda} \sum_{n=2}^{\infty} \frac{\lambda^{n-2}}{(n-2)!} = \lambda^2 e^{-\lambda} \sum_{m=0}^{\infty} \frac{\lambda^m}{m!} = \lambda^2.
\]

Hence,

\[
E[X^2] - E[X] = \lambda^2.
\]

Thus,

\[
E[X^2] = \lambda^2 + E[X] = \lambda^2 + \lambda.
\]

Consequently,

\[
\text{var}[X] = E[X^2] - E[X]^2 = E[X^2] - \lambda^2 = \lambda.
\]

Review: Distributions

Experiment: flip a coin \( n \) times. The coin is such that \( Pr[H] = \frac{\lambda}{n}. \)

Random Variable: \( X \) - number of heads. Thus, \( X = B(n, \frac{\lambda}{n}). \)

Poisson Distribution is distribution of \( X \) “for large \( n. \)”

\[
Pr[X = m] = \left(\frac{n}{m}\right) p^m(1-p)^{n-m}, \text{ with } p = \frac{\lambda}{n}
\]

\[
= \frac{n(n-1)\cdots(n-m+1)}{m!} \frac{\lambda^m}{m!} \left(1 - \frac{\lambda}{n}\right)^{n-m}
\]

\[
= \frac{\lambda^m}{m!} \left(1 - \frac{\lambda}{n}\right)^n \approx \frac{\lambda^m}{m!} \left(1 - \frac{\lambda}{n}\right)^n \approx \frac{\lambda^m}{m!} e^{-\lambda}.
\]

We used \((1-a/n)^n \approx e^{-a}.\)
Summary

Variance, Geometric, time to collect coupons; Poisson

- \( \text{var}[a + bX] = b^2 \text{var}[X] \);
- \( X, Y \) independent \( \Rightarrow \text{var}[X + Y] = \text{var}[X] + \text{var}[Y] \);
- Time to collect all of \( n \) coupons: \( nH(n) \approx n(\ln(n) + 0.58) \);
- Intuition: Last coupon takes \( 1/n \) to collect!
- Remember: \( E[X_1 + \cdots + X_n] = E[X_1] + \cdots + E[X_n] \).
- Distributions:
  - \( G(p) : E[X] = 1/p, \text{var}[X] = (1 - p)/p^2 \);
  - \( B(n,p) : E[X] = np, \text{var}[X] = np(1 - p) \);
  - \( P(\lambda) : E[X] = \lambda, \text{var}[X] = \lambda \).