Variance, Independent Random Variables, Begin Zoo

CSE 312 Spring 26

Lecture 11

Review Expected Value of a Random Variable

Definition. Given a discrete RV $X: \Omega \to \mathbb{R}$, the **expectation** or **expected value** or **mean** of X is

$$\mathbb{E}[X] = \sum_{x \in \Omega_X} x \cdot P(X = x) \qquad = \sum_{x \in \Omega_X} x \cdot p_X(x)$$

Intuition: "Weighted average" of the possible outcomes (weighted by probability)

Recap Linearity of Expectation

Theorem. For any two random variables X and Y

$$\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y].$$

Or, more generally: For any random variables $X_1, \dots, X_n$, and constants $a_1, a_2, \dots, a_n, b$,

$$\mathbb{E}[a_1 X_1 + \cdots + a_n X_n + b] = a_1 \mathbb{E}[X_1] + \cdots + an\mathbb{E}[X_n] + b.$$

Recap Using LOE to compute complicated expectations

Often boils down to the following three steps:

- Decompose: Finding the right way to decompose the random variable into sum of simple random variables

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

- LOE: Apply linearity of expectation.

$$\mathbb{E}[X] = \mathbb{E}[X_1] + \cdots + \mathbb{E}[X_n].$$

- Conquer: Compute the expectation of each X_i

Often, X_i are **indicator** (0/1) random variables.

Recap Expected Value of $g(X)$ -- LOTUS

Definition. Given a discrete RV $X: \Omega \to \mathbb{R}$, the **expectation** or **expected value** or **mean** of $g(X)$ is

$$\mathbb{E}[g(X)] = \sum_{x \in \Omega_X} g(x) \cdot P(X = x) \quad = \sum_{x \in \Omega_X} g(x) \cdot p_X(x)$$

$Y = g(X)$

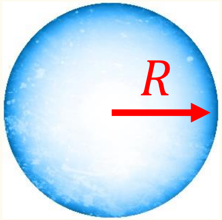
Also known as **LOTUS:** “Law of the unconscious statistician

LOTUS Example

Definition. Given a discrete RV $X: \Omega \to \mathbb{R}$, the **expectation** or **expected value** or **mean** of $g(X)$ is

$$\mathbb{E}[g(X)] = \sum_{x \in \Omega_X} g(x) \cdot P(X = x) \quad = \sum_{x \in \Omega_X} g(x) \cdot p_X(x)$$

Consider a sphere, whose radius is a random variable R:



$$R = \begin{cases} 1 & \text{w.p. } \frac{3}{8} \\ 2 & \text{w.p. } \frac{1}{4} \\ 3 & \text{w.p. } \frac{3}{8} \end{cases}$$

Q: What is the expected radius of the sphere? $E(R) = 2$

Q: What is the expected volume of the sphere?

LOTUS example (2)

$\text{Vol} = g(R) = \frac{4}{3}\pi R^3$

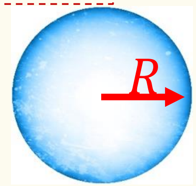
$E(aX) = aE(X)$

Definition. Given a discrete RV $X: \Omega \to \mathbb{R}$, the **expectation** or **expected value** or **mean** of $g(X)$ is

$$\mathbb{E}[g(X)] = \sum_{x \in \Omega_X} g(x) \cdot P(X = x) \quad = \sum_{x \in \Omega_X} g(x) \cdot p_X(x)$$

Consider a sphere, whose radius is a random variable R:

$$R = \begin{cases} 1 & \text{w.p. } \frac{3}{8} \\ 2 & \text{w.p. } \frac{1}{4} \\ 3 & \text{w.p. } \frac{3}{8} \end{cases}$$



$$\mathbb{E}[\text{Volume}] = \mathbb{E}\left[\frac{4}{3}\pi R^3\right]$$

$$= \frac{4}{3}\pi E(R^3)$$

$$= \frac{4}{3}\pi \left[1^3 \cdot \frac{3}{8} + 2^3 \cdot \frac{1}{4} + 3^3 \cdot \frac{3}{8}\right]$$

$g(x) = x^3$

Q: Is $\mathbb{E}[R^3] \neq (\mathbb{E}[R]^3)$?

inside parenthes $E[R^3]$ $2^3 = 8$

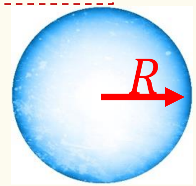
$$E[g(X)] \neq g(E(X))$$

LOTUS example (3)

Definition. Given a discrete RV $X\colon \Omega \to \mathbb{R}$, the **expectation** or **expected value** or **mean** of $g(X)$ is

$$\mathbb{E}[g(X)] = \sum_{x \in \Omega_X} g(x) \cdot P(X = x) \quad = \sum_{x \in \Omega_X} g(x) \cdot p_X(x)$$

Consider a sphere, whose radius is a random variable R:



$$R = \begin{cases} 1 & \text{w.p. } \frac{3}{8} \\ 2 & \text{w.p. } \frac{1}{4} \\ 3 & \text{w.p. } \frac{3}{8} \end{cases}$$

$$\mathbb{E}[\text{Volume}] = \mathbb{E}\left[\frac{4}{3}\pi R^3\right]$$

$$= \frac{4}{3}\pi \cdot 1^3 \cdot \frac{3}{8} + \frac{4}{3}\pi \cdot 2^3 \cdot \frac{1}{4} + \frac{4}{3}\pi \cdot 3^3 \cdot \frac{3}{8}$$

$$= 7\frac{2}{3}\pi$$

Variance – Properties

$$\mathbb{E}[X] = \sum_{x \in \Omega_X} x \cdot P(X = x)$$

$$\mathbb{E}[g(X)] = \sum_{x \in \Omega_X} g(x) \cdot P(X = x)$$

Definition. The **variance** of a (discrete) RV X is

$$\mathrm{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \sum_{x \in \Omega_X} p_X(x) \cdot (x - \mathbb{E}[X])^2$$

$g(x) = (x - E(X))^2$

Intuition: Variance (or standard deviation) is a quantity that measures, in expectation, how "far" the random variable is from its expectation.

Variance - summary

Definition. The **variance** of a (discrete) RV X is

$$\mathrm{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \sum_{x \in \Omega X} p_X(x) \cdot (x - \mathbb{E}[X])^2$$

Standard deviation: $\sigma(X) = \sqrt{\mathrm{Var}(X)}$

Recall $\mathbb{E}[X]$ is a **constant**, not a random variable itself.

Intuition: Variance (or standard deviation) is a quantity that measures, in expectation, how "far" the random variable is from its expectation.

Agenda

- Recap
- Properties of Variance
- Independent Random Variables
- Properties of Independent Random Variables

Variance – Properties (1)

$E(aX) = aE(X)$

Definition. The **variance** of a (discrete) RV X is

$$\mathrm{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \sum_x p_X(x) \cdot (x - \mathbb{E}[X])^2$$

Theorem. $\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$

$(a-b)^2 = a^2 - 2ab + b^2$

$$E[(X - E(X))^2]$$

$$= E[X^2 - 2XE(X) + (E(X))^2]$$

$$= E[X^2] - \underbrace{E[2X E(X)]}_{2E(X)E(X)} + \underbrace{E(E(X)^2)}_{(E(X))^2} = E(X^2) - [E(X)]^2$$

Variance – Properties (2)

Theorem. $\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$

Proof: $\mathrm{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2]$

Recall $\mathbb{E}[X]$ is a **constant**

$$= \mathbb{E}[X^2 - 2\mathbb{E}[X] \cdot X + \mathbb{E}[X]^2]$$

$$= \mathbb{E}(X^2) - 2\mathbb{E}[X]\mathbb{E}[X] + \mathbb{E}[X]^2$$

(linearity of expectation!)

$$= \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

$\mathbb{E}[X^2]$ and $\mathbb{E}[X]^2$ are different !

Variance – Properties (3)

$$\mathbb{E}[X] = \sum_{x \in \Omega_X} x \cdot P(X = x)$$

$$\mathbb{E}[g(X)] = \sum_{x \in \Omega_X} g(x) \cdot P(X = x)$$

Definition. The **variance** of a (discrete) RV X is

$$\mathrm{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \sum_{x \in \Omega_X} p_X(x) \cdot (x - \mathbb{E}[X])^2$$

Theorem. $\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$

Theorem. For any $a, b \in \mathbb{R}$, $\mathrm{Var}(a \cdot X + b) = a^2 \cdot \mathrm{Var}(X)$

$Y = aX + b$

$\mathrm{Var}(aX+b) = \mathrm{Var}(aX)$

$\mathrm{Var}(X+b) = \mathrm{Var}(X)$

$\mathrm{Var}(aX) = E((aX)^2) - [E(aX)]^2 = a^2(E(X^2) - E(X))^2)$

$E[a^2X^2]$ $(aE[X])^2$ $= a^2 Var(X)$

Variance of Indicator Random Variables

Suppose that X_A is an indicator RV for event A with $P(A) = p$ so

$$X_A = \begin{cases} 1 & \text{w/prob } p \\ 0 & 1-p. \end{cases}$$

$$\mathbb{E}[X_A] = P(A) = p$$

$$\mathrm{Var}(X_A) = \mathbb{E}[X_A^2] - \mathbb{E}[X_A]^2 = p - p^2 = p(1-p)$$

$$E[X_A^2] = 1^2 \cdot p + 0^2(1-p) = p$$

Variance of Indicator Random Variables - calculation

Suppose that X_A is an indicator RV for event A with $P(A) = p$ so

$$\mathbb{E}[X_A] = P(A) = p$$

Since X_A only takes on values 0 and 1, we always have $X_A^2 = X_A$ so

$$\mathrm{Var}(X_A) = \mathbb{E}[X_A^2] - \mathbb{E}[X_A]^2 = \mathbb{E}[X_A] - \mathbb{E}[X_A]^2 = p - p^2 = p(1-p)$$

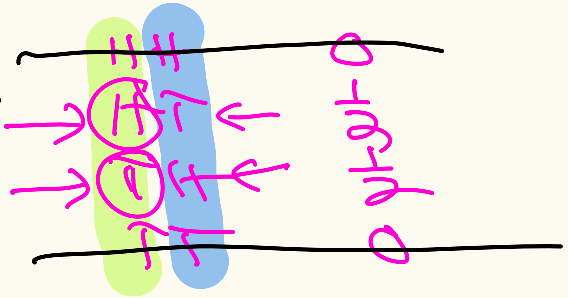
Another example: $\mathrm{Var}(X+Y) \neq \mathrm{Var}(X) + \mathrm{Var}(Y)$

$$X_1 = \begin{cases} 1 & \frac{1}{2} \\ 0 & \frac{1}{2} \end{cases}$$

Proof by counter-example:

Recall glued coins

- Let X_1 be a r.v. that indicates if the first coin comes up heads.
- Let X_2 be a r.v. that indicates if the second coin comes up heads.



$$E(X_1) = \frac{1}{2} = E(X_2)$$

$$\mathrm{Var}(X_1) = \frac{1}{4} = \mathrm{Var}(X_2)$$

$$\mathrm{Var}(X_1) + \mathrm{Var}(X_2) = \frac{1}{2}$$

$$\mathrm{Var}(\underbrace{X_1 + X_2}_{1 \text{ w/prob } 1}) = 0$$

Another example: $\mathrm{Var}(X+Y) \neq \mathrm{Var}(X) + \mathrm{Var}(Y)$ (Explanation)

Proof by counter-example:

Recall glued coins

- Let X_1 be a r.v. that indicates if the first coin comes up heads.
- Let X_2 be a r.v. that indicates if the second coin comes up heads.

- Outcomes are HT and TH, each with probability 0.5
- Therefore, X_1 and X_2 are indicator random variables with probability 0.5 of being 1.
- Therefore, they both have expectation 0.5 and variance 0.25.
- Thus $\mathrm{Var}(X_1) + \mathrm{Var}(X_2) = 0.5$
- On the other hand, $X_1 + X_2$ counts the number of heads in the outcome, which is always 1. Therefore $\mathrm{Var}(X_1 + X_2) = 0$

Question 1

The **variance** of a (discrete) RV X is

$$\mathrm{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \sum_x p_X(x) \cdot (x - \mathbb{E}[X])^2.$$

Can the variance of a random variable be negative?

Poll:
https://pollev.com/ annakarlin185

A. Yes
B. No
C. It will take me too long to figure out.
D. I don't know how/where to start.

Question 2

The **variance** of a (discrete) RV X is

$$\mathrm{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \sum_x p_X(x) \cdot (x - \mathbb{E}[X])^2.$$

Is Var(X + 5) = Var (X) + 5?

$\mathrm{Var}(X+5) = \mathrm{Var}(X)$

Poll:
https://pollev.com/ annakarlin185

A. Yes
B. No
C. It will take me too long to figure out.
D. I don't know how/where to start.

Question 3

The **variance** of a (discrete) RV X is

$$\mathrm{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \sum_x p_X(x) \cdot (x - \mathbb{E}[X])^2.$$

Is it true that if Var(X) = 0, then X is a constant?

Poll:

https://pollev.com/ annakarlin185

A. Yes
B. No
C. It will take me too long to figure out.
D. I don't know how/where to start.

Question 4

The **variance** of a (discrete) RV X is

$$\mathrm{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 \geq 0$$

What is the relationship between $E(X^2)$ and $[E(X)]^2$?

$$E(X^2) = \mathrm{Var}(X) + [E(X)]^2$$

<u>Poll:</u>
https://pollev.com/ annakarlin185

A. $\mathbb{E}[X^2] \geq \mathbb{E}[X]^2$ for all X
B. $\mathbb{E}[X^2] \leq \mathbb{E}[X]^2$ for all X
C. $\mathbb{E}[X^2] = \mathbb{E}[X]^2$ for all X
D. None of the above.

Agenda (2)

- Variance
- Properties of Variance
- Independent Random Variables
- Properties of Independent Random Variables

Random Variables and Independence

Comma is shorthand for AND

Definition. Two random variables X, Y are **(mutually) independent** if for all x, y,

$$P(X = x, Y = y) = P(X = x) \cdot P(Y = y)$$

Intuition: Knowing X doesn't help you guess Y and vice versa

$\forall x, y \in \mathbb{R}$ $\quad \{X = x\}, \{Y = y\}$ are indep

Multiple Random Variables and Independence

Comma is shorthand for AND

Definition. Two random variables X, Y are **(mutually) independent** if for all x, y,

$$P(X = x, Y = y) = P(X = x) \cdot P(Y = y)$$

Intuition: Knowing X doesn't help you guess Y and vice versa

Definition. The random variables $X_1, \dots, X_n$ are **(mutually) independent** if for all $x_1, \dots, x_n$,

$$P(X_1 = x_1, \dots, X_n = x_n) = P(X_1 = x_1) \cdots P(X_n = x_n)$$

Note: No need to check for all subsets, but need to check <u>for all outcomes</u>!

These random variables are not independent

Recall glued coins

- Let X_1 be a r.v. that indicates if the first coin comes up heads.
- Let X_2 be a r.v. that indicates if the second coin comes up heads.

$$\underbrace{P(X_1=1, X_2=1)}_{0} \neq \underbrace{P(X_1=1)}_{\frac{1}{2}} \underbrace{P(X_2=1)}_{\frac{1}{2}}$$

Example

$p = \frac{1}{2}$

Let X be the number of heads in n *independent* coin flips of the same coin. Let $Y = X \bmod 2$ be the parity (even/odd) of X.

Are X and Y independent?

even odd

$\Omega_Y = \{0, 1\}$

$\Omega_X = \{0, 1, 2, \ldots, n\}$

$P(Y=0, X=1) = 0$

$P(Y=0) = \frac{1}{2}$

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

Poll:

https://pollev.com/ annakarlin185

A. Yes
B. No
C. It will take me too long to figure out.
D. I don't know how to start.

Example 2

Make $2n$ *independent* coin flips of the same coin.

Let X be the number of heads in the first n flips and Y be the number of heads in the last n flips.

Are X and Y independent?

$P(X=k, Y=j)$

Poll:

https://pollev.com/ annakarlin185

A. Yes
B. No
C. It will take me too long to figure out.
D. I don't know how to start.

Agenda (3)

- Variance
- Properties of Variance
- Independent Random Variables
- Properties of Independent Random Variables

Important Facts about Independent Random Variables

Theorem. If X, Y independent, $\mathbb{E}[X \cdot Y] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$

Theorem. If X, Y independent, $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$

Corollary. If $X_1, X_2, \ldots, X_n$ mutually independent,

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i}^{n} \mathrm{Var}(X_i)$$

(Not Covered) Proof of $\mathbb{E}[X \cdot Y] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$

Theorem. If X, Y independent, $\mathbb{E}[X \cdot Y] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$

Proof

Let $x_i, y_i, i = 1, 2, \ldots$ be the possible values of X, Y.

$$\mathbb{E}[X \cdot Y] = \sum_i \sum_j x_i \cdot y_j \cdot P(X = x_i \wedge Y = y_j)$$

independence

$$= \sum_i \sum_j x_i \cdot y_i \cdot P(X = x_i) \cdot P(Y = y_j)$$

$$= \sum_i x_i \cdot P(X = x_i) \cdot \left(\sum_j y_j \cdot P(Y = y_j) \right)$$

$$= \mathbb{E}[X] \cdot \mathbb{E}[Y]$$

Note: *NOT* true in general; see earlier example $\mathbb{E}[X^2] \neq \mathbb{E}[X]^2$

(Not Covered) Proof of $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$

Theorem. If X, Y independent, $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$

Proof

$$
\begin{aligned}
&Var(X+Y) \\
&= \mathbb{E}[(X+Y)^2] - (\mathbb{E}[X+Y])^2 \\
&= \mathbb{E}[X^2 + 2XY + Y^2] - (\mathbb{E}[X] + \mathbb{E}[Y])^2 \\
&= \mathbb{E}[X^2] + 2\,\mathbb{E}[XY] + \mathbb{E}[Y^2] - (\mathbb{E}[X]^2 + 2\,\mathbb{E}[X]\,\mathbb{E}[Y] + \mathbb{E}[Y]^2) \\
&= \mathbb{E}[X^2] - \mathbb{E}[X]^2 + \mathbb{E}[Y^2] - \mathbb{E}[Y]^2 + 2\,\mathbb{E}[XY] - 2\,\mathbb{E}[X]\,\mathbb{E}[Y] \\
&= Var(X) + Var(Y) + 2\,\mathbb{E}[XY] - 2\,\mathbb{E}[X]\,\mathbb{E}[Y] \\
&= Var(X) + Var(Y)
\end{aligned}
$$

linearity

equal by independence

Recap Random Variables and Independence

Comma is shorthand for AND

Definition. Two random variables X, Y are **(mutually) independent** if for all x, y,

$$P(X = x, Y = y) = P(X = x) \cdot P(Y = y)$$

Intuition: Knowing X doesn't help you guess Y and vice versa

Definition. The random variables $X_1, \ldots, X_n$ are **(mutually) independent** if for all $x_1, \ldots, x_n$,

$$P(X_1 = x_1, \ldots, X_n = x_n) = P(X_1 = x_1) \cdots P(X_n = x_n)$$

Note: No need to check for all subsets, but need to check for all values!

Example – Coin Tosses

We flip n *independent* coins, each one heads with probability p

- $X_i = \begin{cases} 1, & i^{\text{th}} \text{ outcome is heads} \\ 0, & i^{\text{th}} \text{ outcome is tails.} \end{cases}$
- $Z =$ number of heads

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

Fact. $Z = \sum_{i=1}^{n} X_i$

$P(X_i = 1) = p$
$P(X_i = 0) = 1 - p$
$\mathbb{E}(X_i) = p$

By LOE $\mathbb{E}[Z] = \sum_{i=1}^{n} \mathbb{E}(X_i) = np$

$\cancel{P(Z = k) =}$

$\text{Var}(X_i) = p(1-p)$

$\text{Var}(Z) = \text{Var}(X_1 + X_2 + \cdots + X_n) \underset{\text{indep.}}{=} \overbrace{\text{Var}(X_1)}^{p(1-p)} + \text{Var}(X_2) + \cdots + \text{Var}(X_n) = np(1-p)$

Example – Coin Tosses (2)

We flip n *independent* coins, each one heads with probability p

- $X_i = \begin{cases} 1, & i^{\text{th}} \text{ outcome is heads} \\ 0, & i^{\text{th}} \text{ outcome is tails.} \end{cases}$
- $Z =$ number of heads

Fact. $Z = \sum_{i=1}^{n} X_i$

$P(X_i = 1) = p$
$P(X_i = 0) = 1 - p$
$\mathbb{E}(X_i) = p$

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

By LOE $\mathbb{E}[Z] = \sum_{i=1}^{n} \mathbb{E}(X_i) = np$

Note: $X_1, \dots, X_n$ are mutually independent!

$$\mathrm{Var}(Z) = \sum_{i=1}^{n} \mathrm{Var}(X_i) = n \cdot p(1-p)$$

Note $\mathrm{Var}(X_i) = p(1-p)$

Agenda (if time)

- Zoo of Discrete RVs, Part I
  - Uniform Random Variables
  - Bernoulli Random Variables
  - Binomial Random Variables
  - Geometric Random Variables

Motivation for "Named" Random Variables

Random Variables that show up all over the place.

- Easily solve a problem by recognizing it's a special case of one of these random variables.

Each RV introduced today will show:

- A general situation it models
- Its name and parameters
- Its PMF, Expectation, and Variance
- Example scenarios you can use it

Welcome to the Zoo! (Preview) 🦍🐘🦁🐅🦓🐪🦒

Bernoulli = Indicator

$X \sim \mathrm{Unif}(a, b)$

$$P(X = k) = \frac{1}{b - a + 1}$$

$$\mathbb{E}[X] = \frac{a + b}{2}$$

$$\mathrm{Var}(X) = \frac{(b - a)(b - a + 2)}{12}$$

$X \sim \mathrm{Ber}(p)$

$$P(X = 1) = p, P(X = 0) = 1 - p$$

$$\mathbb{E}[X] = p$$

$$\mathrm{Var}(X) = p(1 - p)$$

$X \sim \mathrm{Bin}(n, p)$

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

$$\mathbb{E}[X] = np$$

$$\mathrm{Var}(X) = np(1 - p)$$

$X \sim \mathrm{Geo}(p)$

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

$$\mathbb{E}[X] = \frac{1}{p}$$

$$\mathrm{Var}(X) = \frac{1 - p}{p^2}$$

$X \sim \mathrm{NegBin}(r, p)$

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

$$\mathbb{E}[X] = \frac{r}{p}$$

$$\mathrm{Var}(X) = \frac{r(1 - p)}{p^2}$$

$X \sim \mathrm{HypGeo}(N, K, n)$

$$P(X = k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$$

$$\mathbb{E}[X] = n\frac{K}{N}$$

$$\mathrm{Var}(X) = n\frac{K(N - K)(N - n)}{N^2(N - 1)}$$