# Probabilities

## Common Discrete Probability Distributions

**Mohamad GHASSANY**

EFREI PARIS

Loi Uniforme Discrète $\mathcal{U}(n)$

Bernoulli distribution $\mathcal{B}(p)$

Binomial distribution $\mathcal{B}(n, p)$

Poisson distribution $\mathcal{P}(\lambda)$

Geometric or Pascal distribution $\mathcal{G}(p)$

Negative Binomial distribution $\mathcal{BN}(r, p)$

# Loi Uniforme Discrète $\mathcal{U}(n)$

## Definition

*A random variable* $X$ *has a* ***discrete uniform distribution*** *if each of the* $n$ *values in its range, say,* $x_1, x_2, \ldots, x_n$*, has equal probability. Then:*

$$P(X = x_i) = \frac{1}{n} \qquad \forall i \in \{1, \ldots, n\}$$

*We say* $X \sim \mathcal{U}(n)$.

## Example

The distribution of the numbers obtained at the throw of the dice (if it is fair) follows a uniform distribution whose probability distribution is the following:

| $x_i$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $P(X = x_i)$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ |

# Moments of uniform distribution

## Particular case

In the particular case of a discrete uniform distribution where each value of the random variable $X$ corresponds to its rank, i.e. $x_i = i \; \forall i \in \{1, \dots, n\}$, we have:

$$E(X) = \frac{n+1}{2} \quad \text{et} \quad V(X) = \frac{n^2-1}{12}$$

## Demonstration

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2} \quad \text{et} \quad \sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}.$$

## Example

The example of the throw of the dice: we can directly calculate the moments of $X$:

$$E(X) = \frac{6+1}{2} = 3.5 \quad \text{et} \quad V(X) = \frac{6^2-1}{12} = \frac{35}{12} \simeq 2.92.$$

# Bernoulli distribution $\mathcal{B}(p)$

## Indicator random variable

Let $A$ an event; the indicator random variable of $A$, defined by $X = \mathbb{1}_A$, is:

$$X(\omega) = \mathbb{1}_A(\omega) = \begin{cases} 0 & \text{si } \omega \in \bar{A} \\ 1 & \text{si } \omega \in A \end{cases}$$

So $X(\Omega) = \{0, 1\}$ with:

$$\begin{aligned} &P(X = 1) = P\{\omega \in \Omega / X(\omega) = 1\} = P(A) = p \\ &P(X = 0) = P\{\omega \in \Omega / X(\omega) = 0\} = P(\bar{A}) = 1 - P(A) = q \\ &\text{avec } p + q = 1 \end{aligned}$$

## Definition

*We say $X$ follows a Bernoulli distribution of parameter $p = P(A)$, we write $X \sim \mathcal{B}(p)$. A Bernoulli distribution is associated to "**Bernoulli's event**", which is a random experience having two possibilities: **success** $(X = 1)$ or **fail** $(X = 0)$.*

## Bernoulli's Distribution function

$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1-p & \text{si } 0 \leqslant x < 1 \\ 1 & \text{si } x \geqslant 1. \end{cases}$$

## Expected value

$$E(X) = 1 \times P(A) + 0 \times P(\bar{A}) = P(A) = p$$

## Variance

$$V(X) = E(X^2) - E^2(X) = p - p^2 = p(1-p) = pq$$

because

$$E(X^2) = 1^2 \times P(A) + 0^2 \times P(\bar{A}) = P(A) = p$$

# Binomial distribution $\mathcal{B}(n, p)$

# Binomial distribution $\mathcal{B}(n, p)$

- Described for the 1st time by *Isaac Newton* in 1676 and demonstrated for the 1st time by the swiss mathematician *Jacob Bernoulli* in 1713.
- Binomial distribution is one most frequently used probability distributions in applied statistics.
- $n$ **independant** Bernoulli events.
- Each has $p$ as probability of success and $1 - p$ probability of fail.

$$A \quad A \quad \bar{A} \quad A \quad \bar{A} \quad \dots \quad \bar{A} \quad A \quad A$$

$$S \quad S \quad E \quad S \quad E \quad \dots \quad E \quad S \quad S$$

- $X =$ **the number of success** on all $n$ events.
- $X$ depends of two parameters $n$ and $p$.

$$S \quad S \quad E \quad S \quad E \quad \dots \quad E \quad S \quad S$$

- $X =$ **the number of success** on all $n$ events.
- $X(\Omega) = \{0, 1, \dots, n\}$

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k} \qquad 0 \leqslant k \leqslant n$$

- $\binom{n}{k}$ is the number of all samples of size $n$ containing exactly $k$ successes, of probability $p^k$, order is not counted, and $n - k$ fails, of probability $(1 - p)^{n-k}$.
- We write $X \sim \mathcal{B}(n, p)$.

**Remark**

A Birnoulli random variable is a Binomal variable of parameters $(1, p)$.

$$X \sim \mathcal{B}(p) \iff X \sim \mathcal{B}(1, p)$$

## Pascal's triangle

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k} \quad \forall\, n \geqslant 1 \,\text{et}\, 1 \leqslant k \leqslant n-1$$

$$\begin{array}{lccccccccc}
n=0: & & & & & 1 & & & & \\
n=1: & & & & 1 & & 1 & & & \\
n=2: & & & 1 & & 2 & & 1 & & \\
n=3: & & 1 & & 3 & & 3 & & 1 & \\
n=4: & 1 & & 4 & & 6 & & 4 & & 1
\end{array}$$

## Binomial theorm

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k$$

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

### Example

We flip five coins. The results are supposed to be independent. What is the probability distribution of $X$ who is thenumber of heads.

### Solution

- $X =$ nombre de piles (*succès*).
- $n = 5$.
- $p = 1/2$.
- $X \sim \mathcal{B}(5, \frac{1}{2})$.
- $X(\Omega) = \{0, 1, \dots, 5\}$
- $P(X = 0) = \binom{5}{0}\left(\frac{1}{2}\right)^0\left(1 - \frac{1}{2}\right)^{5-0} = \frac{1}{32}$
- $P(X = 1) = \binom{5}{1}\left(\frac{1}{2}\right)^1\left(1 - \frac{1}{2}\right)^{4} = \frac{5}{32}$
- $P(X = 2) = \binom{5}{2}\left(\frac{1}{2}\right)^2\left(1 - \frac{1}{2}\right)^{3} = \frac{10}{32}$
- $P(X = 3) = \binom{5}{3}\left(\frac{1}{2}\right)^3\left(1 - \frac{1}{2}\right)^{2} = \frac{10}{32}$
- $P(X = 4) = \binom{5}{4}\left(\frac{1}{2}\right)^4\left(1 - \frac{1}{2}\right)^{1} = \frac{5}{32}$
- $P(X = 5) = \binom{5}{5}\left(\frac{1}{2}\right)^5\left(1 - \frac{1}{2}\right)^{0} = \frac{1}{32}$

If $X \sim \mathcal{B}(n,p)$ so $E(X) = np$ and $V(X) = np(1-p)$

**demonstration**

***1st method:*** We assigne to each $i$, $1 \leqslant i \leqslant n$, a Bernoulli random variable

$$\mathbb{1}_A = X_i = \begin{cases} 1 & \text{if } A \text{ is realized} \\ 0 & \text{if not} \end{cases}$$

So we write: $X = \sum_{i=1}^{n} X_i = X_1 + X_2 + \ldots + X_n$

Then

$$E(X) = E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i) = np$$

et

$$V(X) = V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V(X_i) = np(1-p)$$

because $X_i$ are indenpendent.

*2nd method:* Direct calculation.

- $E(X) = \sum_{k=0}^{n} k\binom{n}{k}p^k(1-p)^{n-k} = \ldots = np$
- $V(X) = E(X^2) - E^2(X)$
- To get $E(X^2)$ we go through $E[X(X-1)]$.
- $V(X) = E(X^2) - E^2(X) = E[X(X-1)] + E(X) - E(X^2)$
- $E[X(X-1)] = \sum_{k=0}^{n} k(k-1)\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k} = \ldots = n(n-1)p^2$
- $V(X) = n(n-1)p^2 + np - (np)^2 = np(1-p)$

**Example**

The number of heads after $n$ coin flips follows a Binomial distribution $\mathcal{B}(n, 1/2)$:

$$P(X = k) = \binom{n}{k} \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{n-k} = \frac{\binom{n}{k}}{2^n}, \quad 0 \leqslant k \leqslant n$$

with $E(X) = n/2$ and $V(X) = n/4$.

**Example**

The number $N$ of red balls appearing in $n$ draws with replacement from an urn containing two red, three green and one black follows a Binomial distribution $\mathcal{B}(n, 1/3)$:

$$P(N = k) = \binom{n}{k} \left(\frac{1}{3}\right)^k \left(\frac{2}{3}\right)^{n-k} = \binom{n}{k} \frac{2^{n-k}}{3^n}, \quad 0 \leqslant k \leqslant n$$

with $E(X) = n/3$ and $V(X) = 2n/9$.

**Remark**

If $X_1 \sim \mathcal{B}(n_1, p)$ et $X_2 \sim \mathcal{B}(n_2, p)$, $X_1$ and $X_2$ being **indenpendent**, so $X_1 + X_2 \sim \mathcal{B}(n_1 + n_2, p)$. This results from the definition of a Binomial distribution since we sum up here the result of $n_1 + n_2$ independent events.

# Poisson distribution $\mathcal{P}(\lambda)$

**Definition**

*A random variable $X$ follows a Poisson distribution of parameter $\lambda > 0$ if it is defined on $\mathbb{N}$ and*

$$P(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}, \quad k \in \mathbb{N}$$

*This distribution depends on only one real positive parameter $\lambda$, we write $X \sim \mathcal{P}(\lambda)$.*

**Remark**

$$e^x = \sum_{i=0}^{+\infty}\frac{x^i}{i!}$$

So

$$\sum_{k=0}^{\infty} P(X = k) = \sum_{k=0}^{\infty} e^{-\lambda}\frac{\lambda^k}{k!} = e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!} = e^{-\lambda}e^{\lambda} = 1$$

# Moments of Poisson distribution

If $X \sim \mathcal{P}(\lambda)$ So $E(X) = \lambda$ and $V(X) = \lambda$

## Expected value

$$\begin{aligned} E(X) &= \sum_{k=0}^{\infty} kP(X = k) \\ &= \dots \\ &= \lambda. \end{aligned}$$

## Variance

- First we calculate $E(X^2) = \sum_{k=0}^{\infty} k^2 P(X = k) = \dots = \lambda(\lambda + 1)$.

Then

$$V(X) = \lambda(\lambda + 1) - \lambda^2 = \lambda$$

# Poisson distribution $\mathcal{P}(\lambda)$

## Example

- $X =$ number of laptops sold by day in a shop.
- Suppose that $X \sim \mathcal{P}(5)$.
- The probability of solding 5 laptops by day is

$$P(X = 5) = e^{-5}\frac{5^5}{5!} = e^{-5} \simeq 0.1755$$

- The probability of solding at least 2 laptops is

$$P(X \geqslant 2) = 1 - \left(e^{-5}\frac{5^0}{0!} + e^{-5}\frac{5^1}{1!}\right) \simeq 0.9596$$

- The aveage number of laptops sold by day is 5 since $E(X) = \lambda = 5$.

## Proprieties

If $X$ and $Y$ are two **indenpendent** random variable of Poisson distribution, $X \sim \mathcal{P}(\lambda)$ and $Y \sim \mathcal{P}(\mu)$, Then their sum is also Poisson: $X + Y \sim \mathcal{P}(\lambda + \mu)$.

$$\text{If } n \to \infty \text{ and } p \to 0 \text{ alors } X : \mathcal{B}(n, p) \sim \mathcal{P}(\lambda)$$

**Remark**

A good approximationis obtained if $n \geqslant 50$ and $np \leqslant 5$.

In this context, the Poisson distribution is often used to model the number of successes when an experiment with a very low chance of success is repeated a very large number of times.

**Applications of Poisson distribution**

- The number of persons over 100 years in a community.
- The number of fake phone numbers dialed in one day.
- The number of customers entering a given post office in one day.
- The number of $\alpha$ particles emitted by a radioactive material during a certain period of time.

The variables in these examples are approximately Poisson.

# Geometric or Pascal distribution $\mathcal{G}(p)$

- $\varepsilon$ : "Repeat a Bernoulli event until the first success".
- Example:

$$\bar{A} \quad \bar{A} \quad \bar{A} \quad \bar{A} \quad \bar{A} \quad \dots \quad \bar{A} \quad \bar{A} \quad A$$

$$E \quad E \quad E \quad E \quad E \quad \dots \quad E \quad E \quad S$$

- Each trial has $p$ as probability of success and $1-p$ as probability of fail.
- $X =$"number of events".

$$\underbrace{E \quad E \quad E \quad E \quad E \quad \dots \quad E \quad E}_{k-1} \quad S$$

- $X(\Omega) = \mathbb{N}^* = \{1, 2, 3, \dots\}$. We say $X \sim \mathcal{G}(p)$.
- $\forall\, k \in \mathbb{N}^* \quad P(X = k) = (1-p)^{k-1}p$
- Attention: Sometimes $X =$ "number of events until having the first success". In this case $X(\Omega) = \mathbb{N}$. We say $X \sim \mathcal{G}(p)$ on $\mathbb{N}$.
- This distribution is often used to model lifetimes, or waiting time, when the time is measured in discrete way (number of days for example).
- Série entière : $\sum_{k=0}^{\infty} x^k = 1/(1-x)$ pour $|x| < 1$
- $\sum_{k=1}^{\infty} P(X=k) = \sum_{k=1}^{\infty}(1-p)^{k-1}p = p\sum_{j=0}^{\infty}(1-p)^j \sum_{k=1}^{\infty}(1-p)^{k-1}p = p\sum_{j=0}^{\infty}(1-p)^j = p\frac{1}{1-(1-p)} = 1$

## Expected value

- $E(X) = \sum_{k=1}^{\infty} kP(X=k) = \sum_{k=1}^{\infty} kp(1-p)^{k-1} = p\sum_{k=1}^{\infty} k(1-p)^{k-1}$
- Power series: $\sum_{k=0}^{\infty} x^k = 1/(1-x) \quad \text{pour} \quad |x| < 1$
- 1st derivative: $\sum_{k=1}^{\infty} kx^{k-1} = 1/(1-x)^2$
- So $E(X) = \frac{p}{[1-(1-p)]^2} = \frac{1}{p}$

In other words, if independent trials with probability $p$ of success are performed until the first success occurs, the expected number of trials needed is equal to $1/p$. For example, the expected number of throws of a balanced die that it takes to get the value 1 is 6.

## Variance of Geometric distribution

- $V(X) = E(X^2) - E^2(X) = E[X(X-1)] + E(X) - E^2(X)$. While,

$$E[X(X-1)] = \sum_{k=2}^{\infty} k(k-1)p(1-p)^{k-1}$$
$$= p(1-p)\sum_{k=2}^{\infty} k(k-1)(1-p)^{k-2}$$

- 1st derivative of Power series: $\sum_{k=1}^{\infty} kx^{k-1} = 1/(1-x)^2$

- 2nd derivative of Power series: $\sum_{k=2}^{\infty} k(k-1)x^{k-2} = 2/(1-x)^3$

- So $E[X(X-1)] = \frac{2p(1-p)}{[1-(1-p)]^3} = \frac{2(1-p)}{p^2}$

- Then $V(X) = E[X(X-1)] + E(X) - E^2(X) = \frac{1-p}{p^2}$.

# Negative Binomial distribution

$\mathcal{BN}(r, p)$

- $\varepsilon$ : "Repeat a Bernoulli event until $r$ successes".
- Example with $r = 3$:

$$\bar{A} \quad A \quad \bar{A} \quad \bar{A} \quad \bar{A} \quad A \quad \bar{A} \quad \bar{A} \quad A$$

$$E \quad S \quad E \quad E \quad E \quad S \quad E \quad E \quad S$$

- But we can obtain $r$ successes in other ways:

$$S \quad E \quad E \quad E \quad E \quad E \quad S \quad E \quad S$$

$$E \quad E \quad E \quad E \quad S \quad E \quad S \quad E \quad S$$

- Each trial has $p$ as probability of success and $1 - p$ as probability of fail.
- Let $X =$"number of trails to obtain this result".

$$\underbrace{\overbrace{E \quad S \quad E \quad E \quad E \quad S \quad E \quad E}^{r-1 \text{ succès and } k-r \text{ échecs}} \quad S}_{X=k}$$

- $X(\Omega) = \{r, r+1, r+2, \ldots\}$. We say $X \sim \mathcal{BN}(r, p)$.
- $\forall\, k \in X(\Omega)$,

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

# $\mathcal{G}(p) = \mathcal{BN}(1, p)$

- $\varepsilon$ : "Repeat a Bernoulli event until $r$ successes".
- Let,

$$E \quad \dots \quad E \quad S \quad E \quad \dots \quad E \quad S \dots \quad E \dots \quad E \quad S$$

- Let, $Y_1$ the number of trials until the first success, $Y_2$ the number of supplementary trials until the 2nd success, $Y_3$ until the 3rd success and so on.
- Which means,

$$\underbrace{E \quad \dots \quad E \quad S}_{Y_1} \; \underbrace{E \quad \dots \quad E \quad S}_{Y_2} \; \underbrace{\dots}_{\dots} \; \underbrace{E \quad \dots \quad E \quad S}_{Y_r}$$

- The draws being indenpendent and always having the same probability of success, each of the variables $Y_1, Y_2, \dots, Y_r$ is Geometric $\mathcal{G}(p)$.
- $X =$"number of trials until $r$ successes"$= Y_1 + Y_2 + \dots + Y_r$.
- So,

$$E(X) = E(Y_1) + E(Y_2) + \dots + E(Y_r) = \sum_{i=1}^{r} \frac{1}{p} = \frac{r}{p}$$

and

$$V(X) = \sum_{i=1}^{r} V(Y_i) = \frac{r(1-p)}{p^2}$$

since $Y_i$ are indenpendent.