Statistiques et probabilités
Cours n°3

Guillaume Postic

Université Paris-Saclay, Univ. Evry
Département informatique

Master 1 MIAGE - 2023/2024

Variable aléatoire discrète 🆚 continue

- VA discrète prends un ensemble fini de valeurs distincts
    - Exemple {1, 2, 3, 4, 5, 6} pour un dé à 6 faces

- VA continue peut prendre une infinité de valeurs possibles dans l'intervalle de définition
    - Exemples : longueur, temps, température

- VA discrètes et continues sont des **VA quantitatives**
    - Par opposition aux VA **qualitatives**…

Variable aléatoire discrète

Soit une variable aléatoire X qui assigne un nombre a pour la réalisation d'un évènement ω :

- La fonction de masse de X est donnée par

$$p_X(a) \text{ ou } p(a) = P(X = a)$$

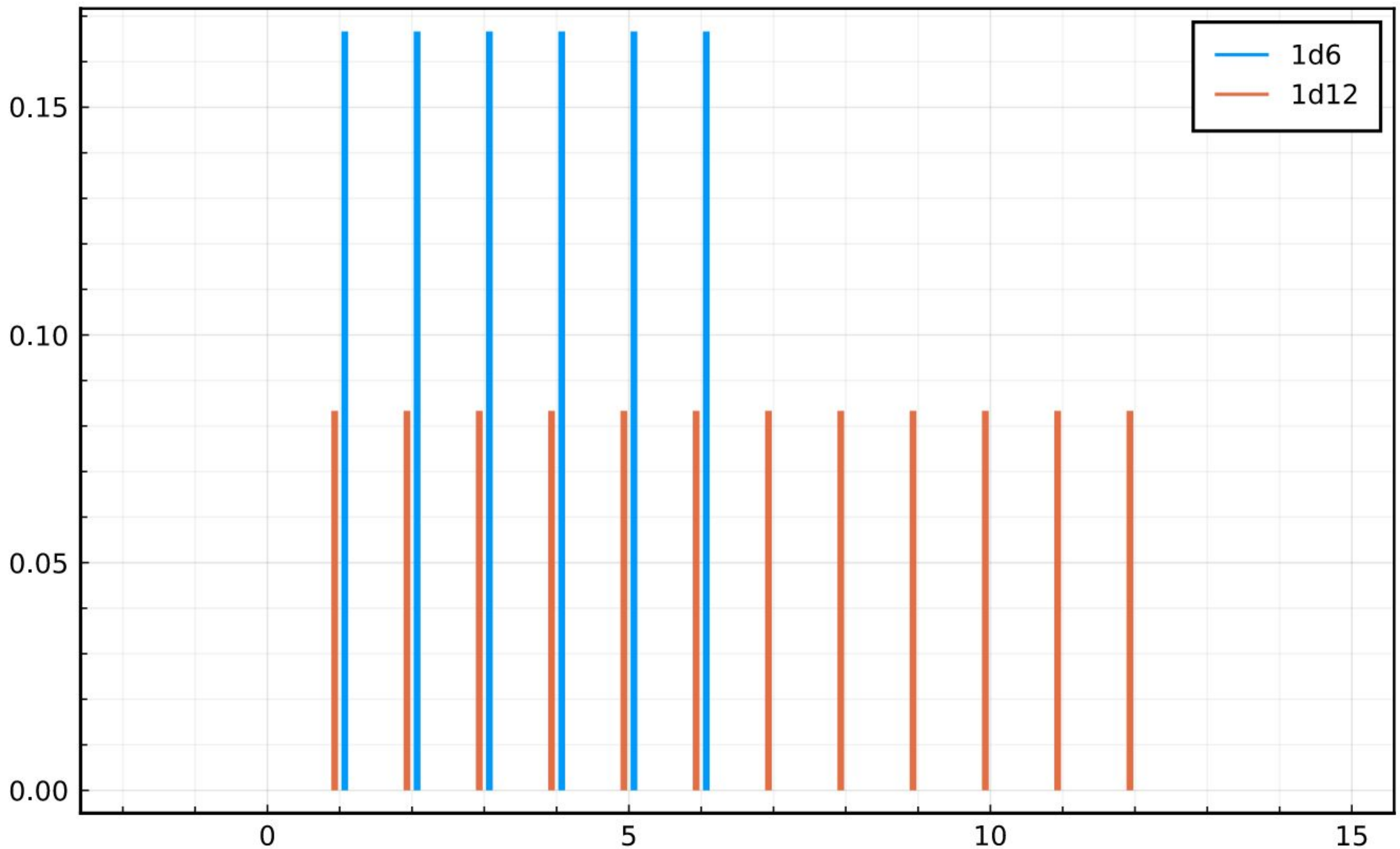
Note : elle défini la loi de probabilité discrète suivie par X

- La fonction de répartition (fonction de distribution cumulative) est donnée par

$$F_X(a) \text{ ou } F(a) = P(X \leq a)$$

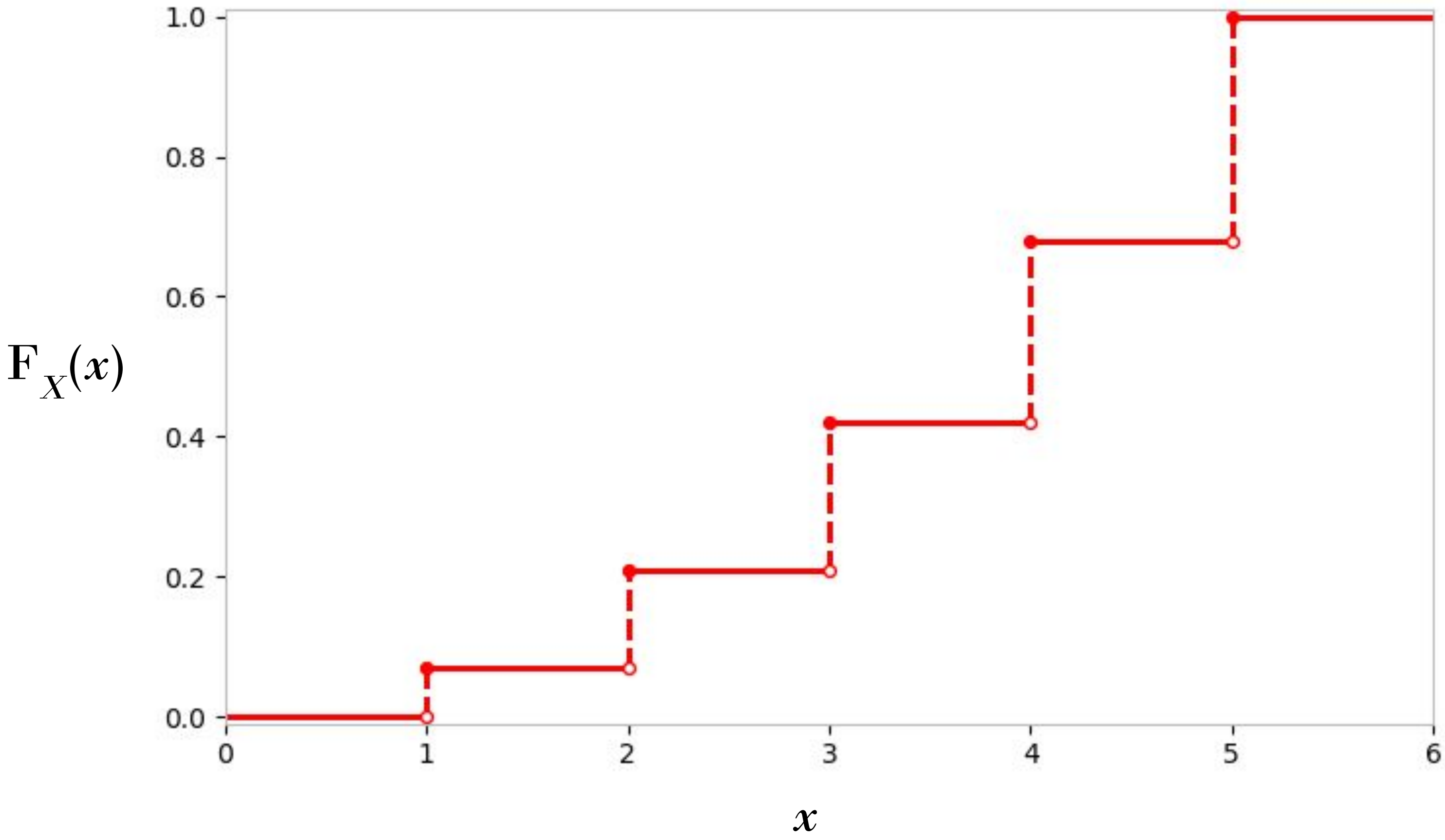
Variable aléatoire discrète

Exemple : fonction de masse du résultat d’un jet de dé à [**six**|**douze**] faces



Variable aléatoire discrète

Exemple : fonction de répartition du résultat d'un jet de dé à six faces

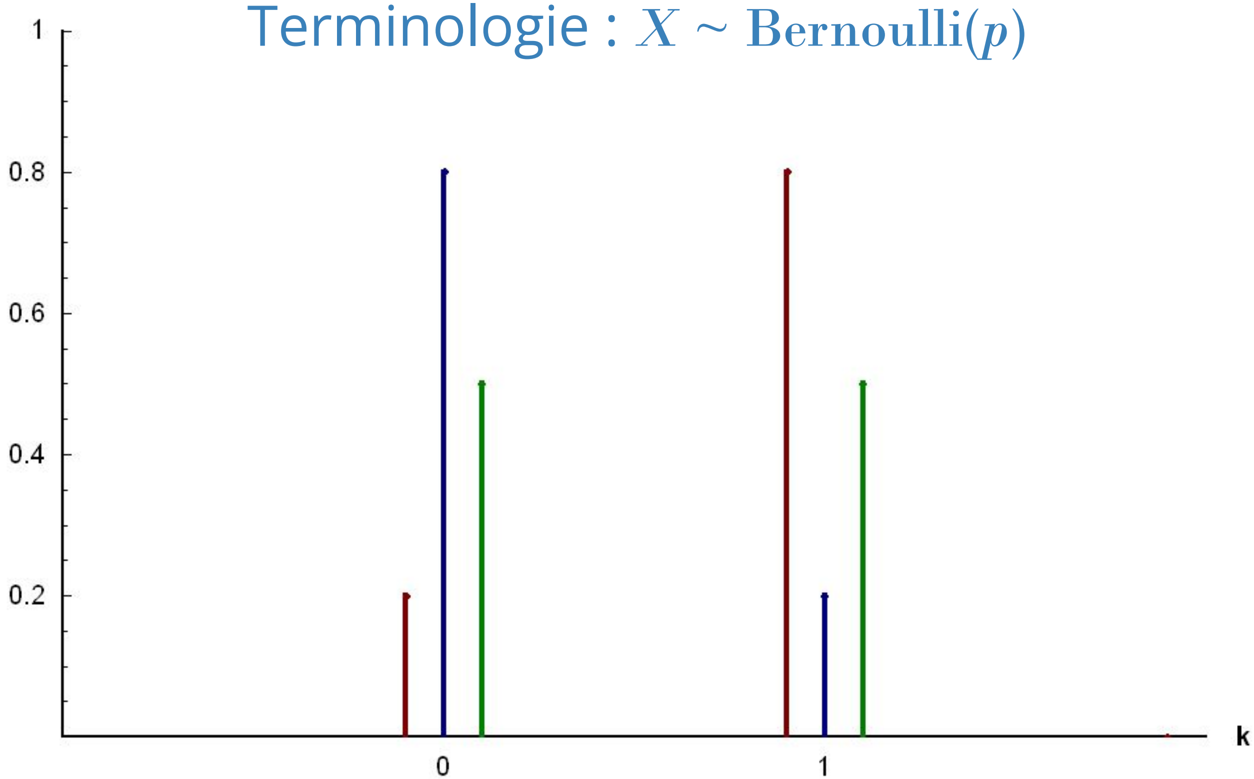


Distributions discrètes (1)

Loi de Bernoulli

Loi de probabilité d'une variable aléatoire discrète qui prend la valeur 1 avec la probabilité *p* et 0 avec la probabilité *q* = 1 – *p*.

Terminologie : $X \sim \mathbf{Bernoulli}(p)$



Distributions discrètes (2)

Loi binomiale

Soit $X_1, X_2, \ldots, X_n$ *n* variables aléatoires de Bernoulli de paramètre *p*, indépendantes et identiquement distribuées, alors leur somme *N* est une variable aléatoire, qui suit la loi binomiale :

$$N = \sum_{k=1}^{n} X_k \sim \mathcal{B}(n, p)$$

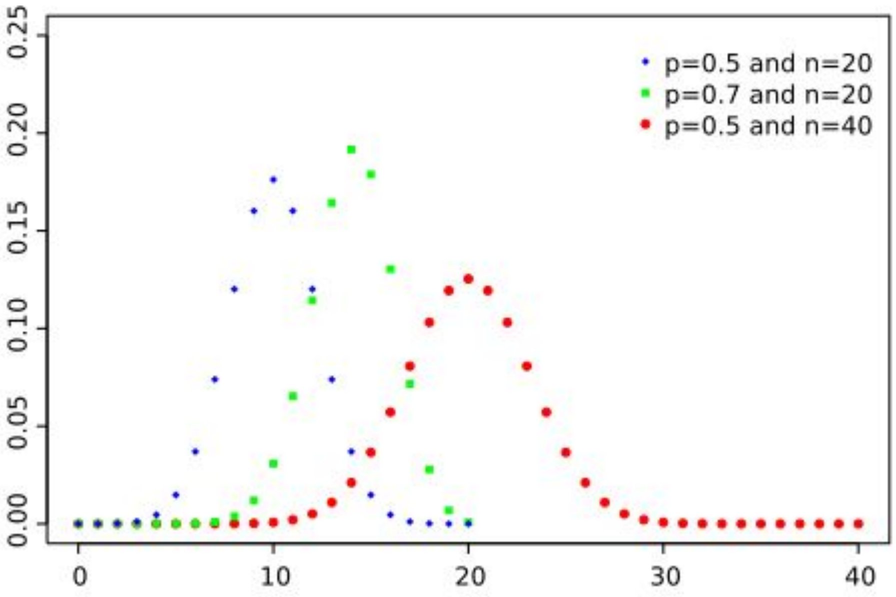
Sa fonction de masse donne la probabilité d'obtenir *k* succès après *n* épreuves de Bernoulli :

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

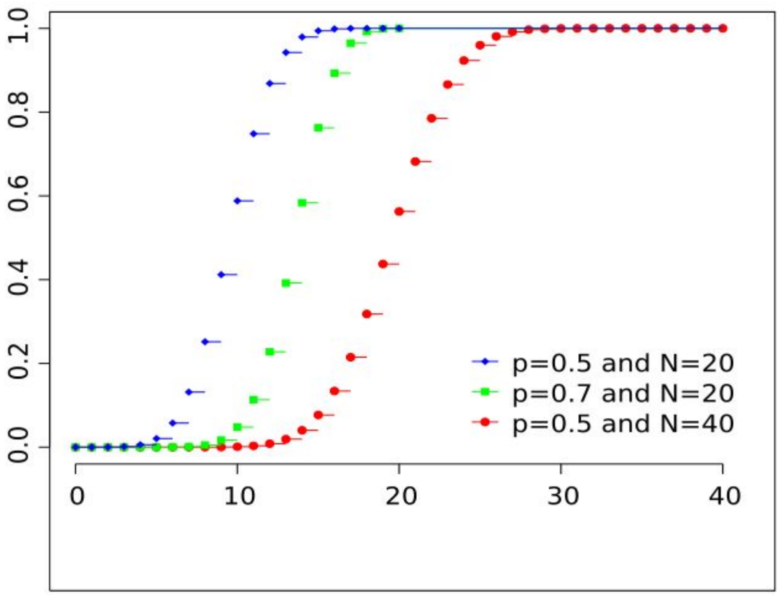
université PARIS-SACLAY

Distributions discrètes (2)

Loi binomiale



Fonction de masse



Fonction de répartition

$$E(X) = np$$

$$V(X) = npq$$

Distributions discrètes (3)

Loi géométrique

Selon la convention choisie :

- la loi du nombre X d'épreuves de Bernoulli indépendantes de probabilité de succès $p \in]0,1[$ (ou $q = 1 - p$ d'échec) nécessaire pour obtenir le premier succès. X est la variable aléatoire donnant le rang du premier succès. Le support de la loi est alors {1, 2, 3, ...}.

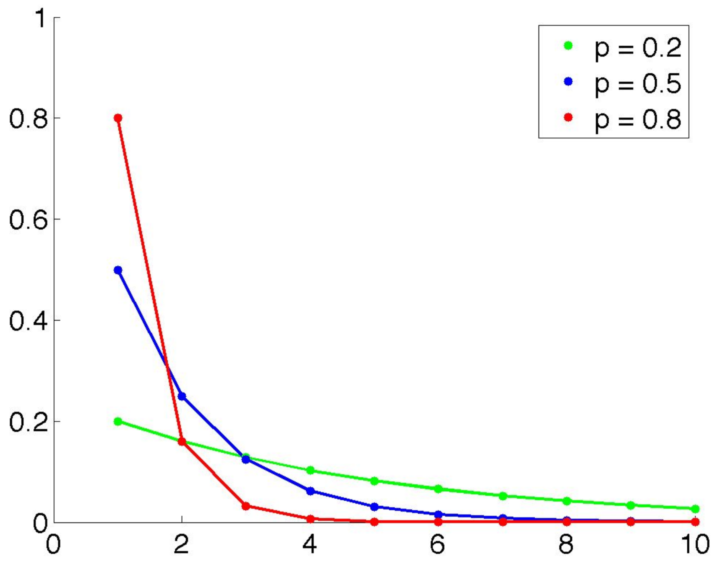
$$\mathbb{P}(X = k) = q^{k-1}p.$$

- La loi du nombre $Y = X - 1$ d'échecs avant le premier succès. Le support de la loi est alors {0, 1, 2, 3, ...}.

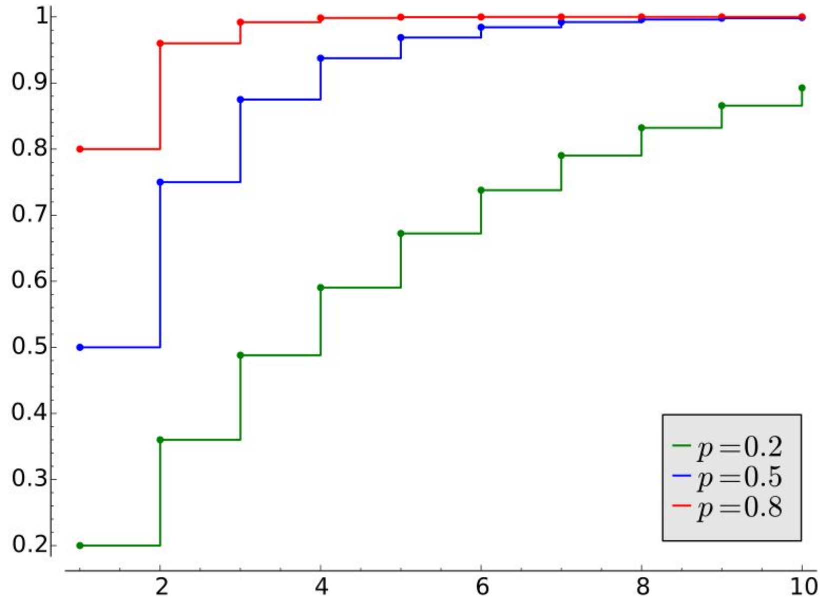
$$\mathbb{P}(Y = k) = q^{k}p.$$

Distributions discrètes (3)

Loi géométrique



Fonction de masse



Fonction de répartition

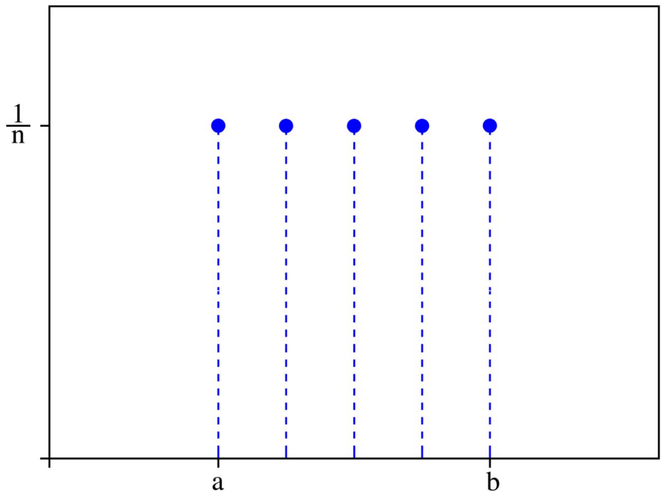
$$E(X) = 1 / p$$

$$V(X) = q / p^2$$

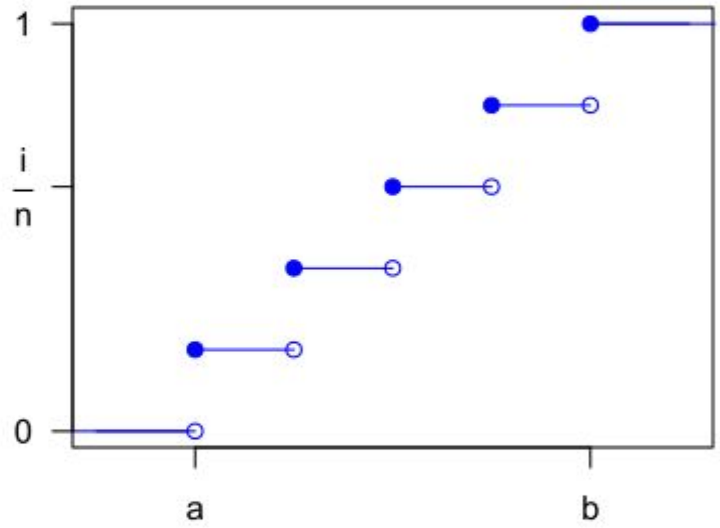
Distributions discrètes (4)

Loi uniforme

Loi de probabilité discrète indiquant une probabilité de se réaliser identique (équiprobabilité) à chaque valeur d'un ensemble fini de valeurs possibles.



Fonction de masse



Fonction de répartition

$$E(X) = (a + b)\ /\ 2$$

$$V(X) = (n^2 - 1)\ /\ 12$$

Convolution (1)

Soit X et Y, deux VA **indépendantes**.

La loi de probabilité de la **somme** $Z = X + Y$ est le *produit de convolution* de leurs lois de probabilités individuelles.

Dans le cas discret, la fonction de masse de Z est :

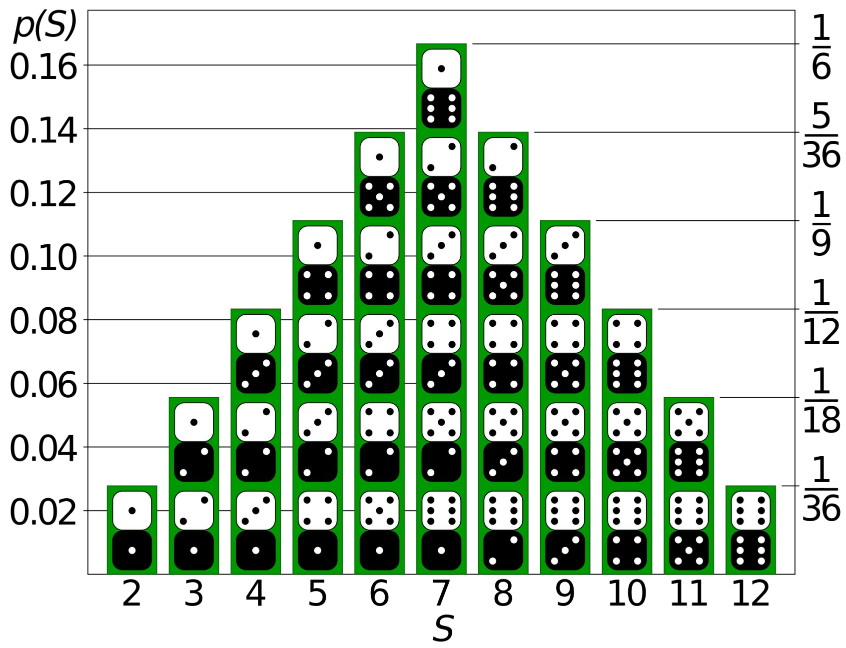
$$p_Z(z) = (p_X * p_Y)(z) = \sum_x p_X(x) \cdot p_Y(z - x)$$

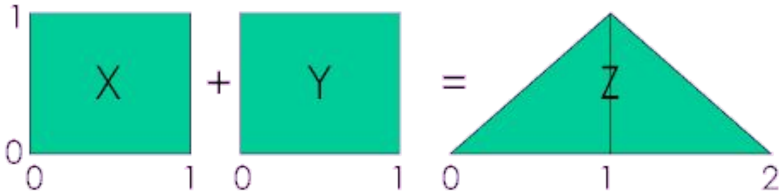
Une loi binomiale de paramètres (n, p) est donc la convolution de n lois de Bernoulli de paramètres p.

Convolution (2)

Exemple :

La somme de deux dés à six faces équilibrés suit une loi « triangle » : convolution de deux lois uniformes.





Convolution (2)

Distribution de Z, somme de dé n°1 et dé n°2, **équilibrés** :

$P(z = 2) = 1/36$

$P(z = 3) = 2/36 = 1/18$

$\boxed{P(z = 4) = 3/36} = 1/12$

$P(z = 5) = 4/36 = 1/9$

$P(z = 6) = 5/36$

$P(z = 7) = 6/36 = 1/6$

$\boxed{P(z = 8) = 5/36}$

$P(z = 9) = 4/36 = 1/9$

$P(z = 10) = 3/36 = 1/12$

$P(z = 11) = 2/36 = 1/18$

$P(z = 12) = 1/36$

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Convolution (2)

Exemple :

La somme de deux dés à six faces équilibrés suit une loi « triangle » : convolution de deux lois uniformes.

Pour $z = 2$

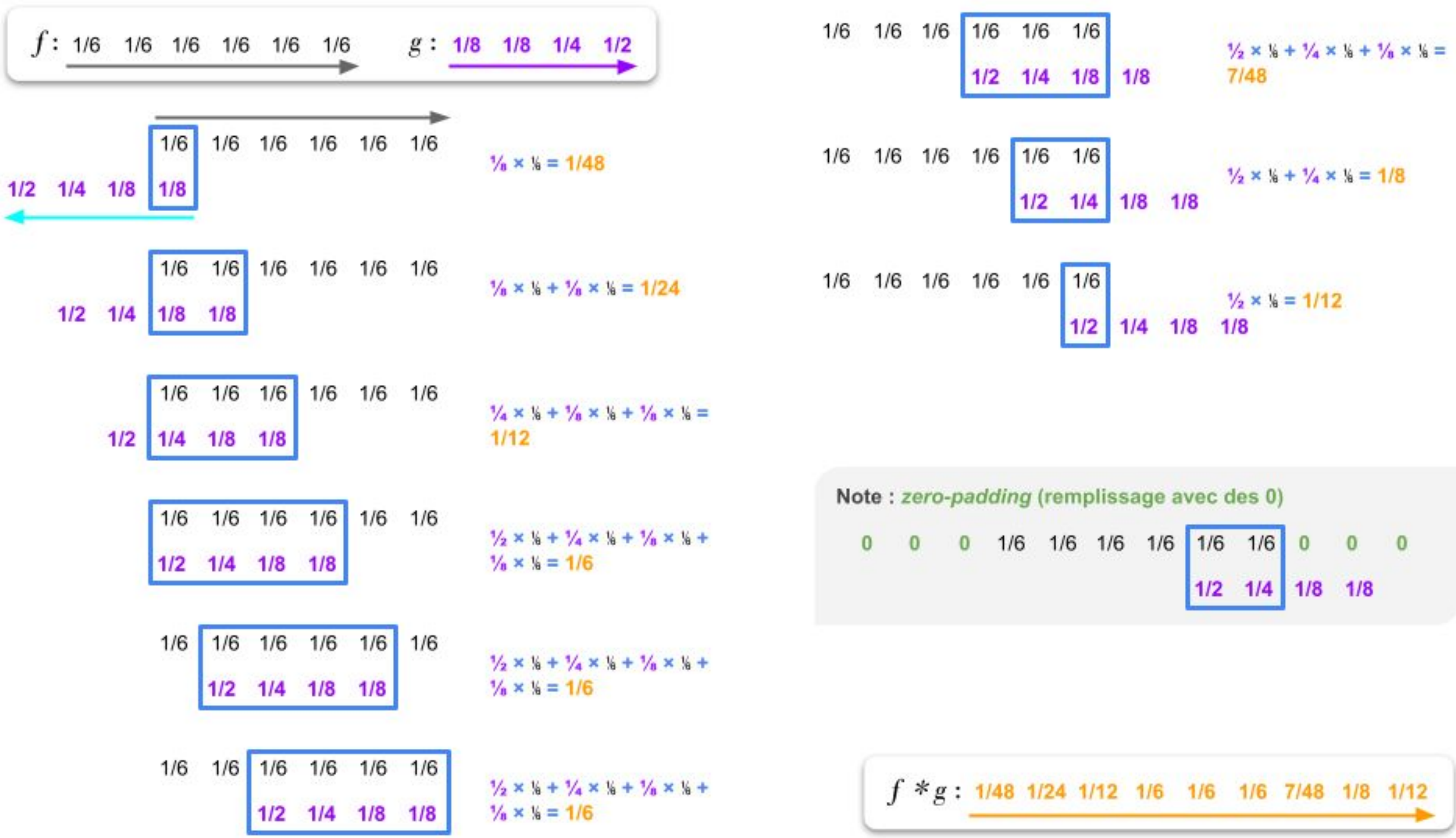
$$p_Z(2) = \sum_{x=1}^{6} p_X(x) \cdot p_Y(2 - x) = p_X(1) \cdot p_Y(1) = ⅙ \cdot ⅙ = 1/36$$

Pour $z = 3$

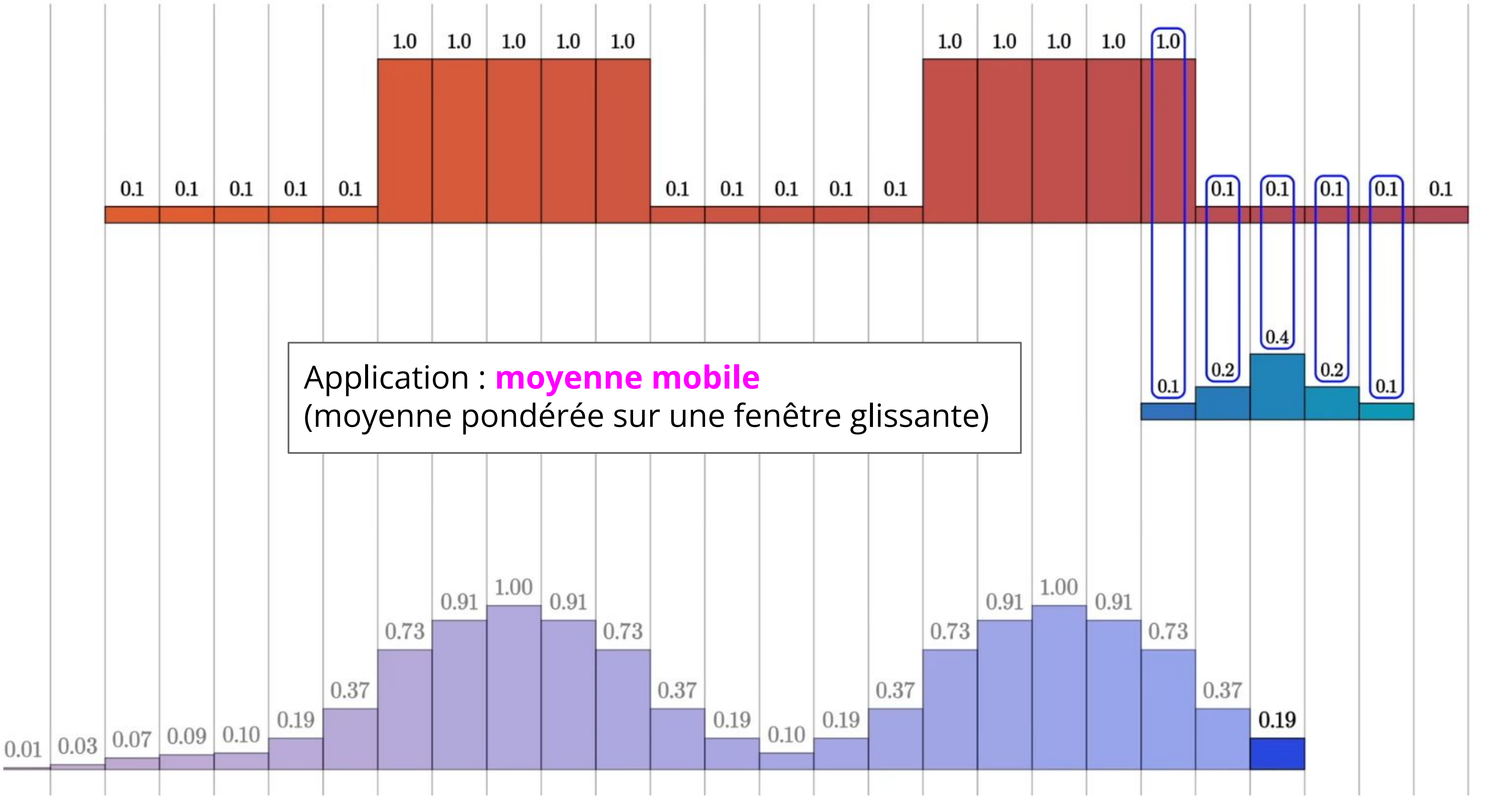
$$p_Z(3) = \sum_{x=1}^{6} p_X(x) \cdot p_Y(3 - x) = p_X(1) \cdot p_Y(2) + p_X(2) \cdot p_Y(1) = 1/18$$

…

Convolution (3)



Convolution (4)



Application : **moyenne mobile**
(moyenne pondérée sur une fenêtre glissante)

Effet de **lissage**
Exemple : flou gaussien en analyse d'images

Distributions discrètes (5)

Loi de Poisson

Rappel de la loi binomiale :

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

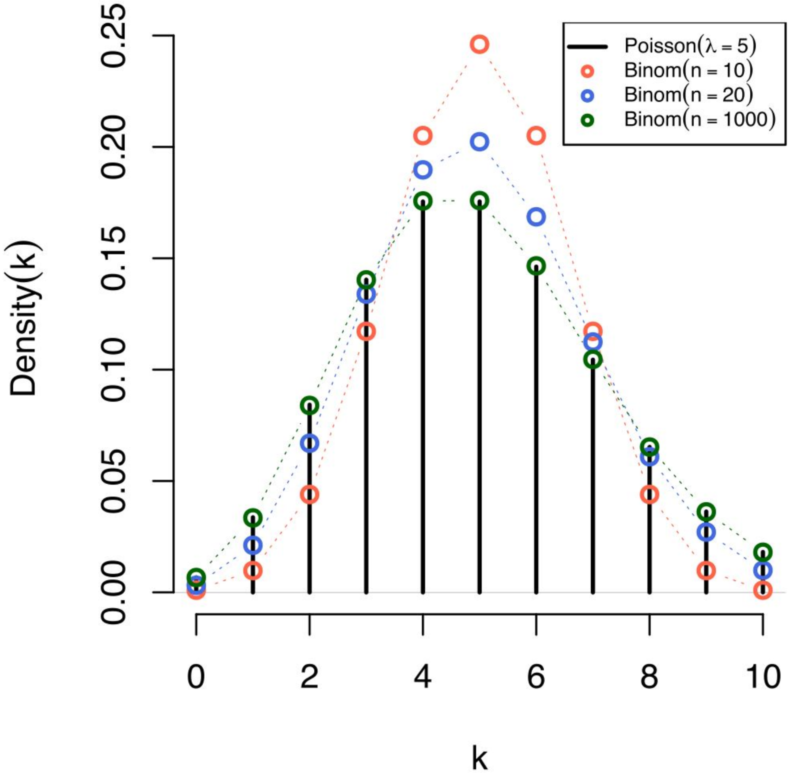
Limite de la loi binomiale quand $n \to +\infty$, $p \to 0$ (« événements rares ») et $\lambda = pn$ constant :

$$p(k) = \mathbb{P}(X = k) = \frac{\lambda^k}{k!} \mathrm{e}^{-\lambda}$$

Distributions discrètes (5)

Loi de Poisson

Pour des valeurs de n grandes et p faibles, le calcul de la loi binomiale devient lourd. La loi de Poisson offre une bonne **approximation**.



Distributions discrètes (5)

Loi de Poisson

Loi de probabilité discrète qui décrit le comportement du **nombre d'événements se produisant dans un intervalle de temps (ou d'espace) fixé**, si ces événements se produisent avec une fréquence moyenne ou espérance connue, et indépen damment du temps écoulé depuis l'événement précédent.

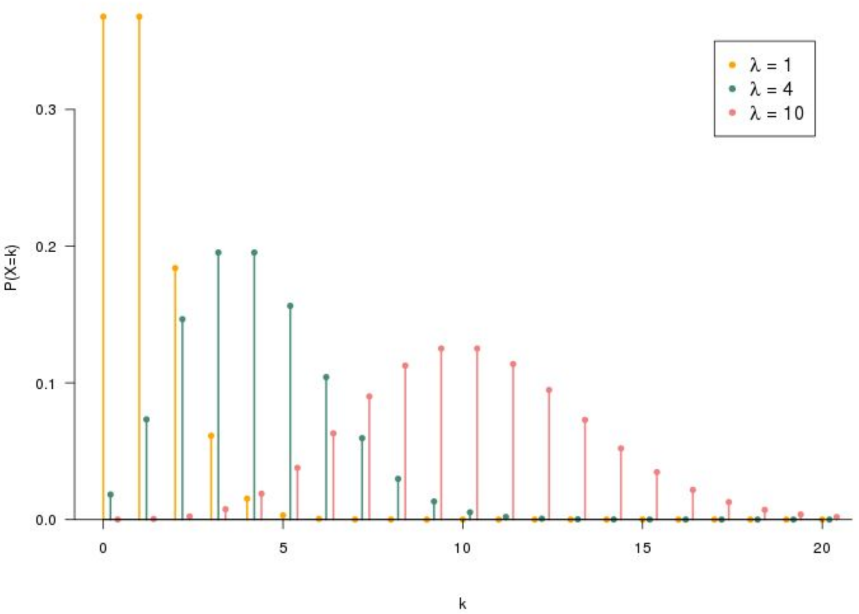
Si le nombre moyen d'occurrences dans un intervalle de temps fixé est λ, alors la probabilité qu'il existe exactement *k* occurrences (*k* étant un entier naturel, *k* = 0, 1, 2…) est :

$$p(k) = \mathbb{P}(X = k) = \frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}$$

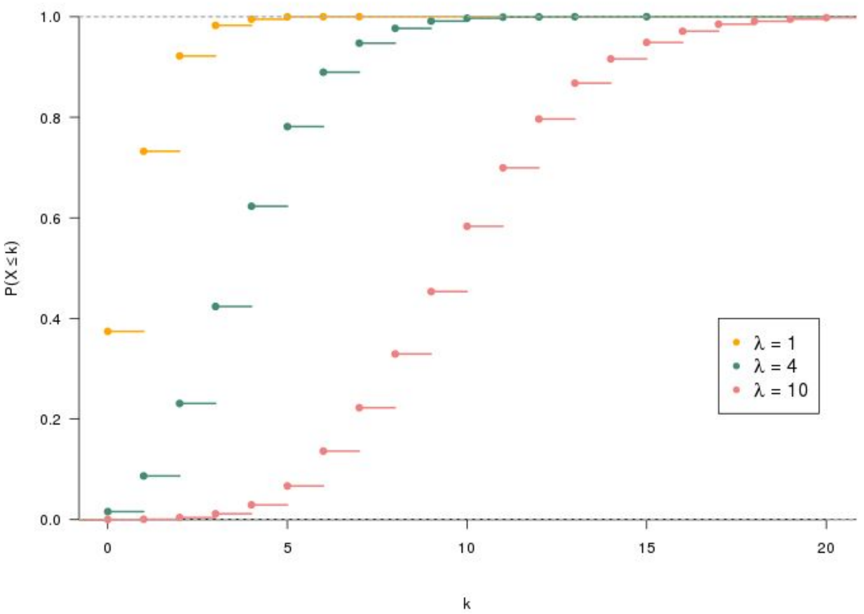
Distributions discrètes (5)

Loi de Poisson

$$X \sim \text{Pois}(\lambda)$$



Fonction de masse



Fonction de répartition

$$\mathrm{E(X)} = \lambda$$

$$\mathrm{V(X)} = \lambda$$

Distributions discrètes (5)

Loi de Poisson

Si l'intervalle de temps n'est pas fixé, alors

- **λ** est la **fréquence** à laquelle l'évènement survient
- le nombre N_t d'occurrences dans un intervalle de longueur t suit une loi de Poisson « d'intensité » λt :

$$\mathbb{P}(N_t = k) = \mathrm{e}^{-\lambda t} \frac{(\lambda t)^k}{k!}$$