

Lois de probabilité

Anita Burgun

Problème posé

- Le problème posé en statistique:
 - On s'intéresse à une population
 - On extrait un échantillon
 - On se demande quelle sera la composition de l'échantillon (pourcentage théorique vs pourcentage observé)
 - Calcul des probabilités : lois

Contenu des cours

- Loi binomiale
- Loi hypergéométrique
- Loi de Poisson
- Loi normale
- Loi du χ^2 (définition)
- Loi T de Student (définition)

Contenu des cours

- Loi binomiale
- *Loi hypergéométrique*
- *Loi de Poisson*
- *Loi normale*
- *Loi du χ^2 (définition)*
- *Loi T de Student (définition)*

Loi binomiale

- Loi née du jeu: gagner aux dés en sortant des as
- On utilise la loi binomiale dans 2 situations: lorsqu'on désire connaître
 - 1) la probabilité de k succès au bout de n tentatives sachant la probabilité P de gagner à chacune des tentatives. C'est la situation de jeu de hasard:
quelle est la probabilité de tirer 3 as en lançant 10 fois un dé?
 - 2) La probabilité d'observer k individus possédant une caractéristique c donnée dans un échantillon de n individus tirés d'une population où la proportion P de la caractéristique c est connue. *Quelle est la probabilité d'observer 3 malades dans un échantillon de 10 sujets choisis au hasard dans une population où la fréquence de la maladie est de 17%?*

Rappel

- Une v.a. X est discrète si l'ensemble des réalisations possibles x_1, x_2, \dots, x_n pour cette variable est fini ou dénombrable.
- Famille de 4 enfants dont les deux parents sont porteurs d'un gène d'une maladie héréditaire. Nombre d'enfants atteints. $X=0,1,2,3,4$
- Nombre d'étudiants ayant les yeux bleus en PCEM1. $X=0,1,2,\dots,n$

Définition

- A chacune des réalisations x_i de la v.a. X est associée une probabilité $P(X=x_i) = p_i$

$$\forall i, 0 \leq p_i \leq 1$$

- L'ensemble des couples (x_i, p_i) forme une loi de probabilité.

$$\sum_{i=0}^n p_i = 1$$

Loi de Bernoulli

- Soit une expérience aléatoire ayant deux résultats possibles :
 - le succès (probabilité p)
 - ou l'échec ($q=1-p$)
- Famille de 4 enfants dont les deux parents sont porteurs d'un gène d'une maladie héréditaire m (récessive).
- La v.a. « avoir la maladie m » suit une loi de Bernoulli de paramètre p égal à 0.25
- $E(X) = p$ $V(X) = p(1-p)$

Loi binomiale

- On répète n fois dans des conditions identiques une expérience aléatoire dont l'issue se traduit par l'apparition ou la non apparition d'un événement A de probabilité p , le résultat de chaque expérience étant indépendant des résultats précédents.
- Soit X le nombre d'apparitions de l'événement A parmi ces n expériences ($0 \leq X \leq n$).
- X suit une loi binomiale de paramètres n et p .

$$\mathcal{B}(n; p)$$

Loi binomiale

- *On répète n fois dans des conditions identiques une expérience aléatoire dont l'issue se traduit par l'apparition ou la non apparition d'un événement A de probabilité p , le résultat de chaque expérience étant indépendant des résultats précédents.*
- *Soit X le nombre d'apparitions de l'événement A parmi ces n expériences ($0 \leq X \leq n$).*
- *X suit une loi binomiale de paramètres n et p .*
- A chaque expérience on peut associer une variable de Bernoulli

Loi binomiale (def 2)

- La somme de n v.a. de Bernoulli X_i indépendantes et de même paramètre p est une v.a. discrète qui suit une loi binomiale de paramètres n et p

$$X = \sum_{i=0}^n X_i$$

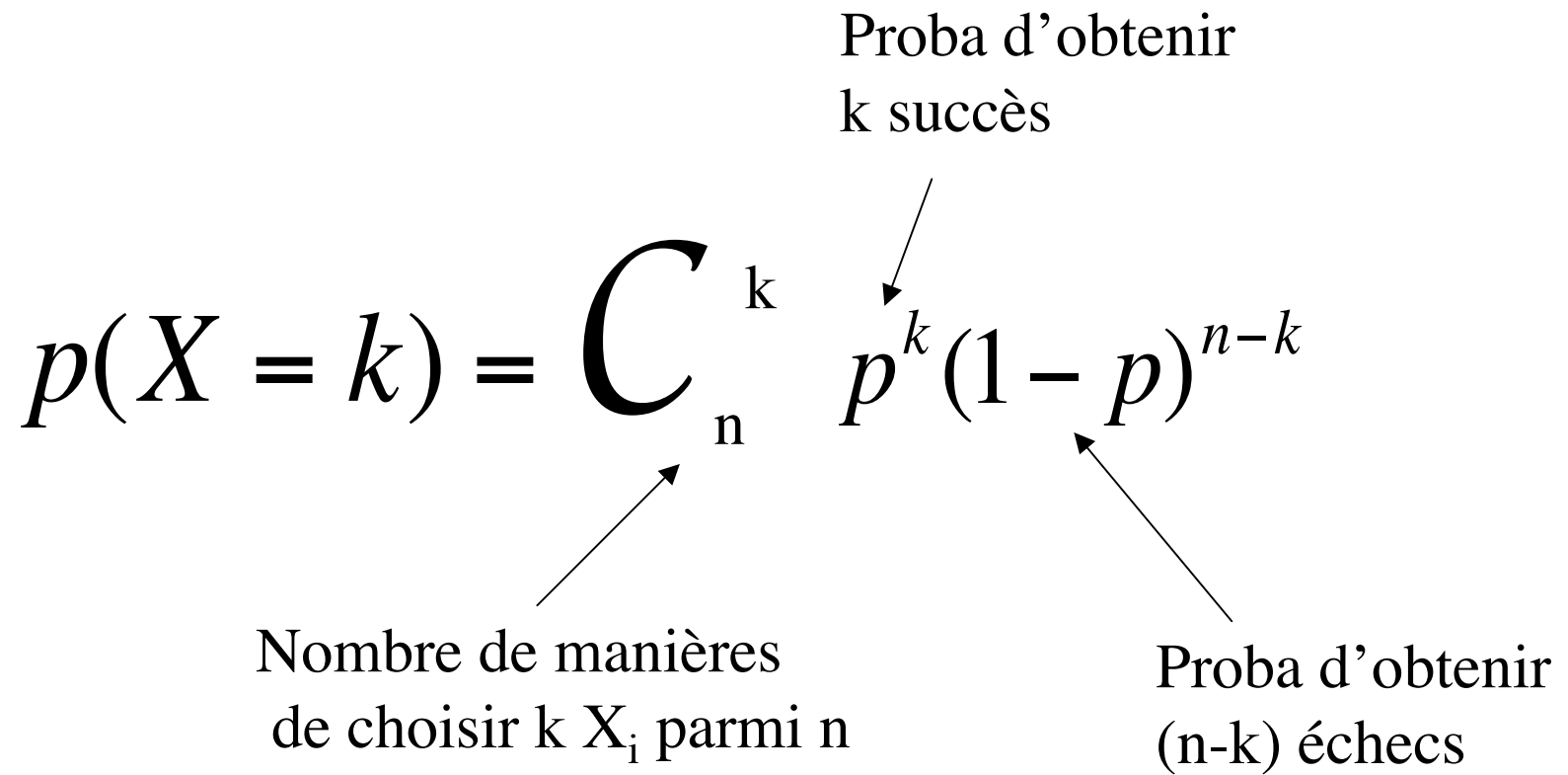
Loi binomiale

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

Proba d'obtenir
k succès

Nombre de manières
de choisir k X_i parmi n

Proba d'obtenir
(n-k) échecs

The diagram illustrates the components of the binomial probability formula. The formula is written as $p(X = k) = C_n^k p^k (1 - p)^{n-k}$. Three arrows point from descriptive text to parts of the formula: one from 'Proba d'obtenir k succès' to p^k , one from 'Nombre de manières de choisir k X_i parmi n' to C_n^k , and one from 'Proba d'obtenir (n-k) échecs' to $(1 - p)^{n-k}$.

Loi binomiale

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

$$P_k = p(X = k) = \frac{n!}{k! (n-k)!} p^k (1 - p)^{n-k}$$

Loi BINOMIALE

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

Loi BINOMIALE parce que ses différents termes sont les termes du développement de $(p+(1-p))^n$ selon la formule du binôme de Newton

$$(a + b)^n = \sum_{k=0}^n C_n^k a^k b^{n-k}$$

$$(p + (1 - p))^n = \sum p_k = 1$$

$$P_0 = (1 - p)^n$$

$$P_{k+1} = \frac{n!}{(k+1)!(n-(k+1))!} p^{k+1} (1-p)^{n-k-1}$$

$$P_{k+1} = \frac{n-k}{k+1} \frac{n!}{k!(n-k)!} \frac{p}{(1-p)} p^k (1-p)^{n-k}$$

$$P_{k+1} = \frac{p}{(1-p)} \frac{n-k}{k+1} P_k$$

Pour aider dans les calculs

$$(k + 1)! = (k + 1) \times k!$$

$$k! = \frac{(k + 1)!}{k + 1}$$

$$\frac{1}{k!} = \frac{k + 1}{(k + 1)!}$$

Espérance, variance

- La somme de n v.a. de Bernoulli X_i indépendantes et de même paramètre p est une v.a. discrète qui suit une loi binomiale de paramètres n et p

$$X = \sum_{i=0}^n X_i$$

Espérance

$$E(X_1 + X_2) = E(X_1) + E(X_2)$$

$$X \sim \mathcal{B}(n; p)$$

$$E(X) = \sum E(X_i)$$

$$E(X) = np$$

Variance

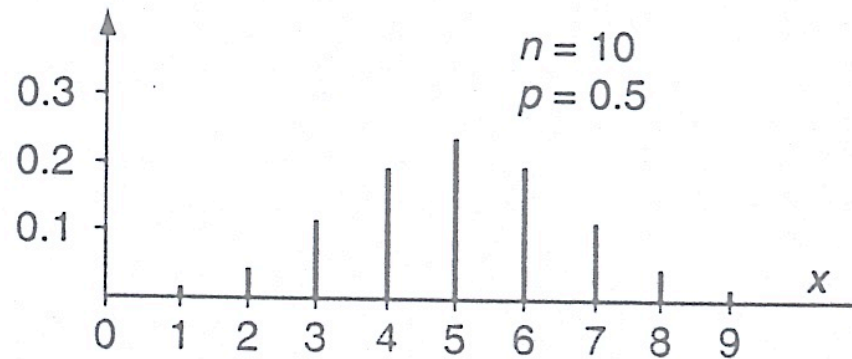
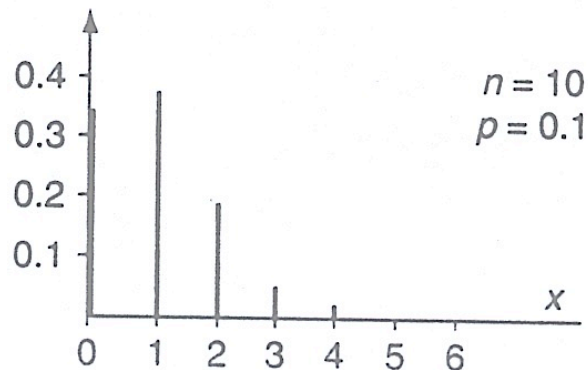
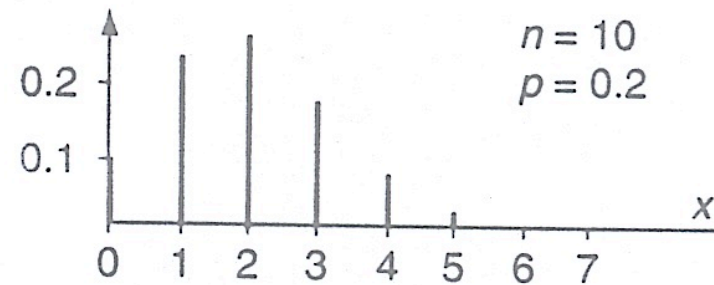
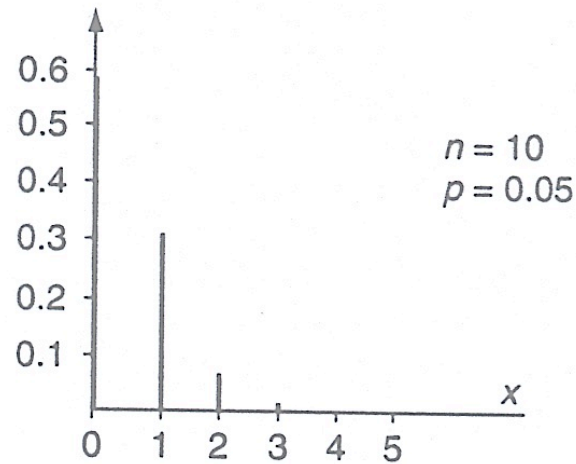
$$V(X_1 + X_2) = V(X_1) + V(X_2)$$

$$X \sim \mathcal{B}(n; p)$$

$$V(X) = \sum V(X_i)$$

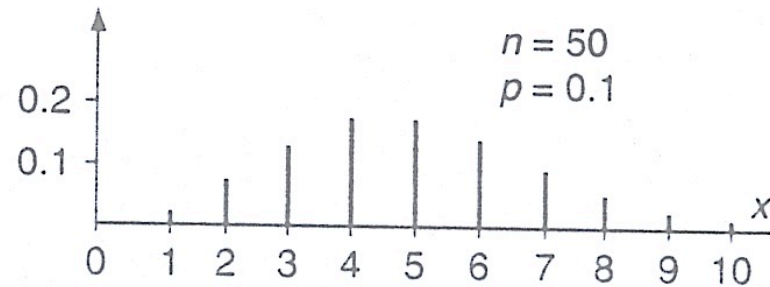
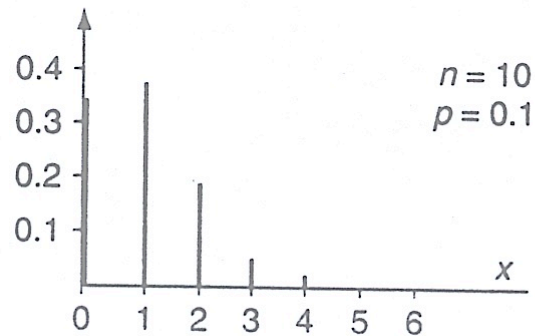
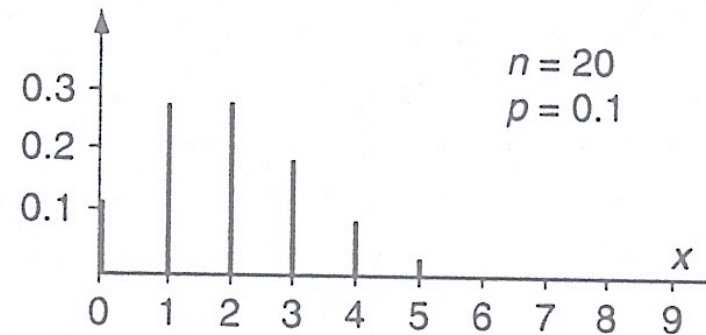
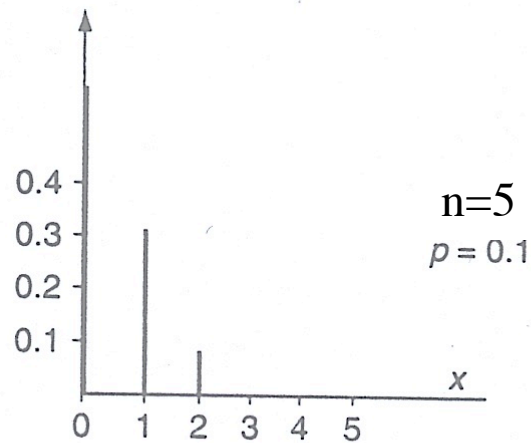
$$V(X) = n p (1-p)$$

Représentation graphique (1)



Tiré de G Saporta, Probabilités, analyse de données et statistique Technip

Représentation graphique (2)



Tiré de G Saporta, Probabilités, analyse de données et statistique Technip

Remarques utiles

- Dans certaines circonstances, la loi binomiale peut être approximée soit par une loi de Poisson (p petit) soit par une loi normale.
- La somme de 2 v.a. binomiales indépendantes et de même paramètre p est une v.a. binomiale

$$X_1 + X_2 \sim \mathcal{B}(n_1 + n_2; p)$$

Illustration (1/3)

- La probabilité d'avoir un garçon lors d'une naissance est de 50%, le sexe des enfants est indépendant d'une naissance à l'autre. Madame X a eu 3 garçons et elle est enceinte pour la quatrième fois. Quelles sont les propositions vraies?
- A. La probabilité que Mme X soit enceinte d'un garçon est de 0,0625
- B. Si l'enfant à naître est un garçon, la famille de Mme X fera partie des 12,5% de familles de 4 enfants du même sexe
- C. Si l'enfant à naître est une fille, la famille de Mme X fera partie des 50% de familles de 4 enfants qui ont 3 enfants du même sexe
- D. La probabilité que mme X soit enceinte d'une fille est de 0,5.
- E. Aucune de ces propositions n'est exacte

Illustration (1/3)

La probabilité d'avoir un garçon lors d'une naissance est de 50%, le sexe des enfants est indépendant d'une naissance à l'autre.

« avoir un garçon »

Variable de Bernouilli de probabilité $p=0,5$

A. La probabilité que Mme X soit enceinte d'un garçon est de 0,0625
FAUX

D. La probabilité que Mme X soit enceinte d'une fille est de 0,5.
VRAI

On s'intéresse au nb de garçons dans une famille de 4 enfants.

Loi binomiale de paramètres $n=4$ et $p=0,5$

- E. Aucune de ces propositions n'est exacte

Illustration 1/3

- Quelles sont les propositions vraies?
- B. Si l'enfant à naître est un garçon, la famille de Mme X fera partie des 12,5% de familles de 4 enfants du même sexe
- C. Si l'enfant à naître est une fille, la famille de Mme X fera partie des 50% de familles de 4 enfants qui ont 3 enfants du même sexe

Illustration (1/3)

- Nombre de garçons loi binomiale $n=4$, $p=0,5$
- B. Si l'enfant à naître est un garçon, la famille de Mme X fera partie des 12,5% de familles de 4 enfants du même sexe VRAI
- $P(X=4 \text{ garçons}) = 0,5^4 = 0,0625$
- Idem pour les filles 0,0625
- donc 12,5% de familles de 4 enfants du même sexe
- C. Si l'enfant à naître est une fille, la famille de Mme X fera partie des 50% de familles de 4 enfants qui ont 3 enfants du même sexe

Illustration (1/3)

- Nombre de garçons loi binomiale $n=4$, $p=0,5$
- C. Si l'enfant à naître est une fille, la famille de Mme X fera partie des 50% de familles de 4 enfants qui ont 3 enfants du même sexe
- VRAI

$$p(3\text{garçons}) = \frac{4!}{3!(4-3)!} (0.5)^3 (0.5)^1 = 0.25$$

Illustration (1/3)

- La probabilité d'avoir un garçon lors d'une naissance est de 50%, le sexe des enfants est indépendant d'une naissance à l'autre. Madame X a eu 3 garçons et elle est enceinte pour la quatrième fois. Quelles sont les propositions vraies?
- A. La probabilité que Mme X soit enceinte d'un garçon est de 0,0625
- B. Si l'enfant à naître est un garçon, la famille de Mme X fera partie des 12,5% de familles de 4 enfants du même sexe
- C. Si l'enfant à naître est une fille, la famille de Mme X fera partie des 50% de familles de 4 enfants qui ont 3 enfants du même sexe
- D. La probabilité que Mme X soit enceinte d'une fille est de 0,5.
- E. Aucune de ces propositions n'est exacte
- Réponse B,C,D

Illustration (2/3)

- Quelle est la probabilité de tirer 3 as en jetant 10 fois un dé?
- On pose $k = \text{nb d'as à sortir} = 3$
 $n = \text{nb de tentatives} = 10$
 $p = \text{probabilité de sortir un as} = 1/6 = 0.17$

$$p(3as) = \frac{10!}{3!(10-3)!} (0.17)^3 (0.83)^7 = 0.16$$

Illustration (3/3)

- Quelle est la probabilité d'observer 3 malades dans un échantillon de 10 sujets choisis au hasard dans une population où la fréquence de la maladie m est 17%?
- On pose $k = \text{nb de malades} = 3$
 $n = \text{taille de l'échantillon} = 10$
 $p = \text{fréquence de la maladie dans la population} = 0.17$

$$p(3\text{malades}) = \frac{10!}{3!(10-3)!} (0.17)^3 (0.83)^7 = 0.16$$

Illustration (3/3)

- Quelle est la probabilité d'observer moins de 4 malades dans un échantillon de 10 sujets choisis au hasard dans une population où la fréquence de la maladie m est 17%?

$$p(X < 4) = p(0) + p(1) + p(2) + p(3)$$

$$p(X < 4) = 0.155 + 0.318 + 0.293 + 0.160 = 0.926$$

Références

- Régis Beuscart et al. Biostatistique, Omniscience
 - Pour le programme PAES
- Gilbert Saporta Probabilités, analyses des données et statistiques. Technip
 - Pour connaître les démonstrations et approfondir les aspects mathématiques des biostat
- Daniel Schwartz Méthodes statistiques à l'usage des médecins et des biologistes. Flammarion Médecine Science
 - Pour les applications médicales