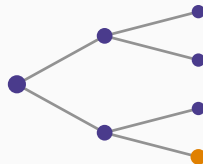


Probabilités

Chapitre 13 : Bernoulli et la loi binomiale



L'épreuve élémentaire

Répéter l'épreuve

La binomiale au travail

Ce qu'il faut retenir

L'épreuve élémentaire

Les mêmes situations reviennent sans cesse

Un crédit fait défaut ou non. Une pièce est conforme ou non. Un client achète ou non. Un candidat réussit ou non.

Ces situations ont la même structure. Plutôt que de refaire chaque fois le calcul, on la nomme une fois, on établit ses formules, et on les réutilise.

Loi de Bernoulli

X suit une loi de **Bernoulli** de paramètre p si elle ne prend que deux valeurs :

$$P(X = 1) = p \quad P(X = 0) = 1 - p$$

On note $X \sim \mathcal{B}(p)$. La valeur 1 s'appelle *succès* — sans jugement de valeur : pour un banquier, le « succès » est un défaut.

Espérance et variance

$$E(X) = p \qquad V(X) = p(1 - p)$$

La démonstration, en deux lignes

$$E(X) = 0(1 - p) + 1(p) = p$$

$$E(X^2) = 0^2(1 - p) + 1^2(p) = p \quad \text{donc} \quad V(X) = p - p^2 = p(1 - p)$$

Le point clé : $X^2 = X$ puisque $0^2 = 0$ et $1^2 = 1$.

Un résultat qu'il faut savoir lire

$V(X) = p(1 - p)$ vaut 0 en $p = 0$ et en $p = 1$, et son maximum est en $p = 0,5$, où elle vaut 0,25.

Le maximum d'incertitude est à égalité de chances. Un événement quasi certain ou quasi impossible ne comporte presque aucun risque : on sait ce qui va se passer.

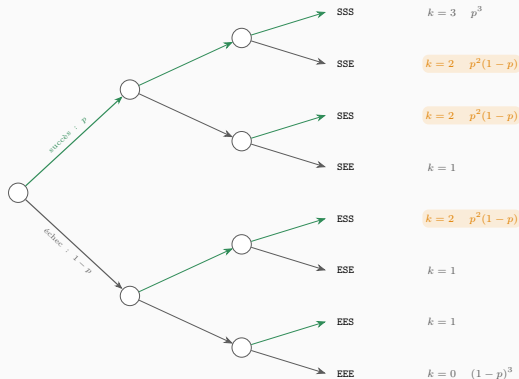
La conséquence pour un modèle de score

Un modèle qui attribue à un dossier une probabilité de défaut de 0,5 n'apporte aucune information : c'est le cas le plus incertain de tous.

Un bon modèle **repousse les dossiers vers les extrêmes**. C'est exactement ce que mesurent les indicateurs de discrimination utilisés en scoring.

Répéter l'épreuve

De l'arbre à la formule



Trois chemins mènent à $k = 2$, et chacun pèse exactement $p^2(1-p)$.

Trois, c'est $\binom{3}{2}$: le nombre de façons de choisir les 2 places du succès parmi 3.

Le dénombrement compte les chemins, la probabilité pèse chaque chemin :

$$P(X = k) = \underbrace{\binom{n}{k}}_{\text{combien de chemins}} \times \underbrace{p^k(1-p)^{n-k}}_{\text{poids d'un chemin}}$$

Toute la loi binomiale tient dans cette phrase — et c'est pourquoi le dénombrement venait *avant*.

Loi binomiale

X suit une loi **binomiale** de paramètres n et p , notée $X \sim \mathcal{B}(n; p)$, si elle compte le nombre de succès sur n épreuves de Bernoulli **indépendantes** et **de même paramètre** p .

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}, \quad k = 0, 1, \dots, n$$

La formule se lit en deux morceaux

$p^k (1 - p)^{n-k}$ est la probabilité **d'un** chemin précis comportant k succès : on multiplie le long du chemin, ce qui n'est licite que grâce à l'indépendance.

$\binom{n}{k}$ compte **combien** de chemins conviennent : c'est le dénombrement du chapitre 5, et l'ordre n'y compte pas puisque seul le total importe.

Les trois conditions à vérifier

Sans elles, la formule est fausse

1. Un nombre **fixé** d'épreuves : n est connu à l'avance.
2. Deux issues seulement par épreuve.
3. Épreuves **indépendantes** et de probabilité p **constante**.

Trois énoncés à trier

Situation	Binomiale ?
200 crédits indépendants, $p = 0,04$	oui
Tirage de 5 boules sans remise dans une urne	non : p change
On lance jusqu'au premier six	non : n n'est pas fixé

Espérance et variance

$$E(X) = np \qquad V(X) = np(1 - p) \qquad \sigma(X) = \sqrt{np(1 - p)}$$

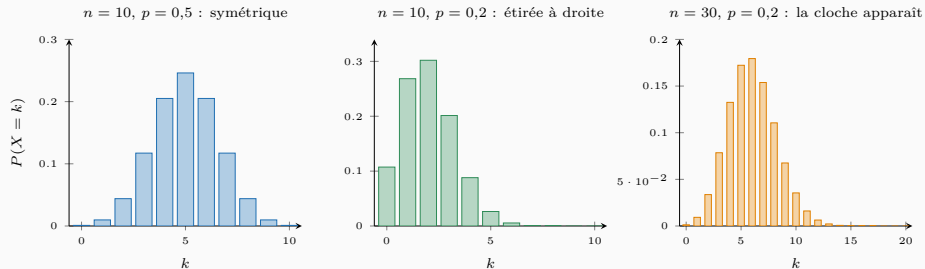
Obtenues sans aucun calcul

$X = X_1 + \dots + X_n$, somme de n indicatrices de paramètre p .

L'espérance est **toujours** additive : $E(X) = np$. Les variances s'additionnent parce que les épreuves sont **indépendantes** : $V(X) = np(1 - p)$.

C'est tout le chapitre 12 qui travaille ici — et c'est la démonstration la plus courte du cours.

Trois allures selon les paramètres



$$E(X) = np \quad \text{et} \quad V(X) = np(1-p). \quad \text{Ici : } 5 \text{ et } 2,5 \ ; \ 2 \text{ et } 1,6 \ ; \ 6 \text{ et } 4,8.$$

Quand n grandit, la forme se régularise et devient une cloche : c'est déjà le théorème central limite qui pointe (chapitre 21).

Le rôle de chaque paramètre

p commande la **forme** : symétrique si $p = 0,5$, étirée à droite si p est petit, à gauche s'il est grand.

n commande la **régularité** : plus n est grand, plus la loi ressemble à une cloche.

Ce que la troisième vignette annonce

À $n = 30$, la binomiale a déjà l'allure d'une loi normale. Ce n'est pas une coïncidence graphique.

Une binomiale **est** une somme de n variables indépendantes. Le théorème central limite (chapitre 21) s'applique donc directement, et l'approximation normale de la binomiale en sera le premier usage.

La binomiale au travail

Une chaîne de production

Une machine produit 3 % de pièces défectueuses. On prélève 20 pièces.

$$P(X = 0) = (0,97)^{20} = 0,544$$

$$P(X = 1) = \binom{20}{1} (0,03)(0,97)^{19} = 20 \times 0,03 \times 0,561 = 0,336$$

$$P(X \geq 2) = 1 - 0,544 - 0,336 = 0,120$$

La règle de décision

Si l'on refuse le lot dès deux défauts, on rejette à tort 12 % des lots pourtant conformes.

C'est le risque du fournisseur, et il se calcule avant de rédiger le contrat, non après le premier litige.

Le mode

Le mode d'une binomiale est l'entier situé juste en dessous de $(n + 1)p$.

Pour $n = 200$ et $p = 0,04$: $(n + 1)p = 8,04$, donc le mode est 8.

Mode, moyenne, médiane

Ici les trois coïncident presque : 8. Quand p est petit et n modeste, ils s'écartent — la loi est dissymétrique.

C'est la même lecture qu'en statistique descriptive : **l'écart entre moyenne et mode mesure l'asymétrie.**

Ce qu'il faut retenir

Six points

1. **Bernoulli** : deux issues, $E(X) = p$, $V(X) = p(1 - p)$, maximale en $p = 0,5$.
2. **Binomiale** : n épreuves indépendantes, $P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$.
3. La formule se lit : *combien de chemins* $\times$ *poids d'un chemin*.
4. Trois conditions : n fixé, deux issues, indépendance avec p constant.
5. $E(X) = np$ et $V(X) = np(1 - p)$, obtenues par décomposition en indicatrices.
6. Pour n grand, la binomiale prend l'allure d'une **cloche**.

La suite

Deux questions échappent encore à la binomiale.

Combien de temps avant le premier succès? — le nombre d'épreuves n'est pas fixé. *Combien d'événements rares dans une période?* — n est immense et p minuscule.

Ce sont la loi **géométrique** et la loi de **Poisson**, au chapitre 14.