Formulaire de probabilités

Damien Robert

13–24 Janvier 2020

En ligne : https://www.math.u-bordeaux.fr/~damienrobert/capes2020
Légende :

☆ Pas explicitement au programme du Capès, mais utile ;

★ Complètement hors programme, inclus dans ce document uniquement par soucis de complétude.

1 Rappels de combinatoire

Definition 1.1.

- Nombre possibles de tirage de r objets parmi n avec remise : n^r (=les fonctions de $\{1, \dots, r\}$ dans $\{1, \dots, n\}$.
- Arrangements : $A_k^n = \frac{n}{(n-k)!}$: nombre de façons distinctes de choisir *dans l'ordre* k objets parmi n (tirage sans remise) ;
- Binomial : $\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n(n-1)\dots(n-k+1)}{k(k-1)\dots 1}$: nombre de façons distinctes de choisir k objets parmi n (tirage non ordonné sans remise) ;
- Partitionnement : répartir $N = r_1 + \dots + r_k$ objets dans k familles de telle sorte que la i-ième famille soit de cardinal r_i : $\frac{N!}{r_1! \dots r_k!}$. C'est aussi le nombre d'anagramme d'un mot de longueur N dont les lettres ont des répétitions $r_1, \dots, r_k$;
- Si on a N_1 boules blanches et N_2 boules noires, nombre de possibilités de tirer n boules sachant qu'il y en a k de blanches : $\frac{\binom{N_1}{k}\binom{N_2}{n-k}}{\binom{N_1+N_2}{n}}$.

Propriétés 1.2.

- $\binom{n}{k} = \binom{n}{n-k}$ *(interprétation : le choix de k objets parmi n est déterminé par son complémentaire) ;*
- *Pascal : $\binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1}$ (interprétation : ensembles de $k+1$ éléments= ensembles contenant un élément fixé x_0 $\coprod$ ensembles ne contenant pas x_0). Donc $\binom{n+1}{k+1} = \sum_{i=k}^{n} \binom{i}{k}$.*
- $(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$. *Donc* $2^n = \sum_{k=0}^{n} \binom{n}{k}$ *et* $0 = \sum_{k=0}^{n} (-1)^k \binom{n}{k}$.
- $k\binom{n}{k} = n\binom{n-1}{k-1}$ *(interprétation : compter le nombre de couples (a, A) où A est une partie de $\{1, \dots, n\}$ de cardinal k et $a \in A$, d'abord en fixant A puis en fixant a).*
- $\sum_{k=0}^{n} k\binom{n}{k} = n2^{n-1}$;
- *Van Der Monde : $\binom{m+n}{k} = \sum_{i=0}^{k} \binom{m}{i}\binom{n}{k-i}$ (k éléments parmi $n+m$ = i élément parmis m et $k-i$ éléments parmi n pour $0 \le i \le k$).*

2 Rappels sur les familles sommables

Ici I dénote un ensemble dénombrable, et $(a_i)_{i \in I}$ une famille dénombrable de nombres réels, indexée par I.
Cas positif, $a_i \geq 0$ **:**

- La famille est sommable si $\{\sum_{i \in J} a_i\}$ est majoré pour tout $J \subset I$ fini. La somme $\sum_{i \in I} a_i$ est alors la limite supérieure de cet ensemble.
- Si l'on se fixe une bijection $\mathbb{N} \to I$, $(a_i)_{i \in I}$ est sommable $\Leftrightarrow$ $\sum_{i=0}^{\infty} a_i$ est une série convergente.

- Donc pour tester la sommabilité, on peut décomposer et permuter I comme on veut (on peut aussi faire des blocs de blocs).

Cas général, $a_i \in \mathbb{R}$ **:**

- $(a_i)_{i\in \mathrm{I}}$ est sommable de somme $a \Leftrightarrow (|a_i|)_{i\in\mathrm{I}}$ est sommable $\Leftrightarrow a_i^+$ et a_i^- sont sommables où $a_i^+ = max(a_i, 0)$, $a_i^- = -min(a_i, 0)$ (donc $a_i = a_i^+ - a_i^-$) et $a = \sum_{i=1}^{\infty} a_i = \sum_{i\in\mathrm{I}} a_i^+ - \sum_{i\in\mathrm{I}} a_i^-$ $\Leftrightarrow$ la série $\sum_{i=1}^{\infty} a_i$ est absolument convergente et de limite a *pour une* bijection $\mathbb{N} \to \mathrm{I}$ $\Leftrightarrow$ la série $\sum_{i=1}^{\infty} a_i$ converge vers la même limite a *pour toute* bijection $\mathbb{N} \to \mathrm{I}$.

Propriétés 2.1.

- *Si* $(x_i)_{i\in\mathrm{I}}$ *et* $(y_i)_{i\in\mathrm{I}}$ *sont sommables, alors* $(ax_i + by_i)_{i\in\mathrm{I}}$ *est sommable de somme*
$$\sum_{i\in\mathrm{I}} ax_i + by_i = a\sum_{i\in\mathrm{I}} x_i + b\sum_{i\in\mathrm{I}} y_i$$
- *Si* $0 \leq x_i \leq y_i$ *et* $(y_i)_{i\in\mathrm{I}}$ *est sommable alors* $(x_i)_{i\in\mathrm{I}}$ *aussi et* $\sum_{i\in\mathrm{I}} x_i \leq \sum_{i\in\mathrm{I}} y_i$ *;*
- Une fois *montré la sommabilité (en regardant la famille* $(|a_i|)_{i\in\mathrm{I}}$*), on peut calculer la limite* a *en décomposant* I *comme l'on veut :*
- *Sommation par bloc : si* $\mathrm{I} = \coprod_{j\in\mathrm{J}} \mathrm{I}_j$ *alors pour tout* $j \in \mathrm{J}$ *chaque famille* $(a_i)_{i\in\mathrm{I}_j}$ *est sommable et on a*
$$\sum_{i\in\mathrm{I}} a_i = \sum_{j\in\mathrm{J}}\sum_{i\in\mathrm{I}_j} a_i.$$
- *Fubbini : si* $\mathrm{I} = \mathrm{I}_1 \times \mathrm{I}_2$,
$$\sum_{(i_1,i_2)\in\mathrm{I}} a_{(i_1,i_2)} = \sum_{i_1\in\mathrm{I}_1}\sum_{i_2\in\mathrm{I}_2} a_{(i_1,i_2)}.$$

3 Espace de probabilités

3.1 Cas d'un univers fini

Definition 3.1.

- Ω désigne un univers fini (univers des résultats possibles du tirage aléatoire) ;
- Un événement est un sous-ensemble $\mathrm{A} \subset \Omega$;
- Une probabilité est une fonction $\mathrm{P} : \mathsf{P}(\Omega) \to [0, 1]$ qui satisfait
  - $\mathrm{P}(\Omega) = 1$;
  - $\mathrm{P}(\mathrm{A} \coprod \mathrm{B}) = \mathrm{P}(\mathrm{A}) + \mathrm{P}(\mathrm{B})$;

Dictionnaire entre probabilités et ensembles :

Événement	Ensemble
Univers	Ω
A	$\mathrm{A} \subset \Omega$
Événement élémentaire	Singleton $\{w\}$
A ou B	$\mathrm{A} \cup \mathrm{B}$
A et B	$\mathrm{A} \cap \mathrm{B}$
non A	$\overline{\mathrm{A}} = \Omega - \mathrm{A}$
A implique B	$\mathrm{A} \subset \mathrm{B}$
A et B incompatible	$\mathrm{A} \cap \mathrm{B} = \emptyset$
w réalise A	$w \in \mathrm{A}$
Événement certain	$\mathrm{A} = \Omega$
Événement impossible	$\mathrm{A} = \emptyset$
Probabilité de A	$\mathrm{P}(\mathrm{A})$

Exemple 3.2. Pour l'expérience de lancer de dé, on peut prendre $\Omega = \{1, 2, 3, 4, 5, 6\}$, et pour le lancer de deux dés on peut prendre $\Omega = \{1, 2, 3, 4, 5, 6\}^2$. L'événement « la somme des deux dés est supérieure à 11 » est représenté par l'ensemble $\mathrm{A} = \{(5, 6), (6, 5), (6, 6)\}$.

Exemple 3.3 (Probabilité uniforme)**.** Si Ω est fini de taille N, P est uniforme si $\mathrm{P}(\{x\}) = 1/\mathrm{N}$. Alors pour tout événement : $\mathrm{P}(\mathrm{A}) = \#\mathrm{A}/\#\Omega$.

Ainsi une probabilité peut se voir comme une généralisation de la notion de cardinal : si $\Omega = \{x_1, \dots, x_n\}$, P est entièrement déterminé par $p_i = \mathrm{P}(\{x_i\})$. Inversement, si on a n réels $p_i \in [0; 1]$ tels que $\sum_{i=1}^{n} p_i = 1$, alors on peut associer une probabilité $\mathrm{P}(\mathrm{A}) = \sum_{x_i\in\mathrm{A}} p_i$.

3.2 ☆ Cas d'un univers général

Definition 3.4. ★

- Ω désigne un univers (univers des tirages aléatoires)
- $\mathcal{T}$ une famille d'événements (tribu). Un événement $A \in \mathcal{T}$ est un sous-ensemble de Ω, et
  - Ω (l'événement certain) et $\emptyset$ (l'événement impossible) sont des événements.
  - Si A est un événement, $\overline{A} := \Omega - A$ aussi,
  - Si $(A_i)_{i\in\mathbb{N}}$ sont des événements, $\cup A_i$ et $\cap A_i$ aussi.

Ce sont des définitions techniques (à oublier !) pour formaliser la notion d'expérience aléatoire. En pratique si Ω est fini ou dénombrable, si on veut que $\{x\}$ soit un événement possible pour tout $x \in \Omega$ on va prendre $\mathcal{T} = \mathsf{P}(\Omega)$. Le problème se pose pour définir sur $[0;1]$ une probabilité telle que $P([a;b]) = b - a$. On ne peut la définir que sur un sous-ensemble de parties : la tribu des Boréliens, qui est engendrée par ces intervalles.

Definition 3.5. Une probabilité sur les événements $\mathcal{T}$ dans Ω est alors une fonction $P : \mathcal{T} \to \mathbb{R}$ qui satisfait

- $P(\Omega) = 1$, et pour tout $A \in \mathcal{T}$, $P(A) \geq 0$;
- Si $(A_i)_{i\in\mathbb{N}}$ est une suite d'événements disjoints, $P(\coprod A_i) = \sum P(A_i)$.

Propriétés 3.6.

- *Pour tout événement* A, $P(A) \in [0;1]$, *donc* $P : \Omega \to [0;1]$, *;*
- A *événement presque sûr :* $P(A) = 1$ *;*
- A *événement presque impossible :* $P(A) = 0$ *;*
- $P(A) = 1 - P(\overline{A})$ *;* $P(\emptyset) = 0$ *;*
- *Si* $A \subset B$, $P(A) \leq P(B)$ *et* $P(B - A) = P(B) - P(A)$ *;*
- *Si* $A_1 \subset A_2 \subset \dots$, $P(A_n)$ *est croissante et* $P(\cup A_n) = \lim_{n\to\infty} P(A_n)$ *;*
- *Si* $A_1 \supset A_2 \supset \dots$, $P(A_n)$ *est décroissante et* $P(\cap A_n) = \lim_{n\to\infty} P(A_n)$ *;*
- $P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$ *;*
- $P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3) - P(A_1 \cap A_2) - P(A_1 \cap A_3) - P(A_2 \cap A_3) + P(A_1 \cap A_2 \cap A_3)$ *;*
- *Si on a une probabilité* P_1 *sur* Ω_1 *et une probabilité* P_2 *sur* Ω_2, *on peut définir une probabilité produit* P *sur* $\Omega = \Omega_1 \times \Omega_2$ *par* $P(A \times B) = P_1(A)P_2(B)$.[1]

Si $\Omega = \{x_i\}_{i\in I}$ est dénombrable, P est entièrement déterminé par $p_i = P(\{x_i\})$, qui forment une famille sommable de somme 1. Inversement, si on a des réels $p_i \in [0;1]$ sommables de somme 1, on peut associer une probabilité sur $\mathcal{T} = \mathsf{P}(\Omega)$ par $P(A) = \sum_{x_i \in A} p_i$.

3.3 Probabilité conditionnelle

Definition 3.7 (Probabilité conditionnelle)**.** Si B est un événement de probabilité non nulle, $P_B(A) = \frac{P(A\cap B)}{P(B)}$ (autre notation : $P(A \mid B)$). On obtient ainsi une probabilité sur l'univers B.

Propriétés 3.8.

- $P(A \cap B) = P(A \mid B)P(B) = P(B \mid A)P(A)$ *;*
- $P(A_1 \cap \dots \cap A_n) = P(A_1)P(A_2 \mid A_1)P \dots (A_{n-1} \mid A_1 \cap A_2 \cap \dots \cap A_{n-2})P(A_n \mid A_1 \cap \dots \cap A_{n-1})$ *;*
- *Formule de Bayes :* $P(B \mid A) = \frac{P(A|B)P(B)}{P(A)}$.
- *Formule des probabilités totales : Si* $\Omega = \coprod_{i\in I} A_i$, $P(A) = \sum_{i\in I} P(A \mid A_i)P(A_i) \Rightarrow$ *arbre de probabilité*
- *Si* $n \in I$, $P(A_n \mid A) = P(A \mid A_n)P(A_n) / \sum_{i\in I} P(A \mid A_i)P(A_i)$.

[1] ★ En prenant la tribu $\mathcal{T}$ sur Ω engendrée par $\mathcal{T}_1 \times \mathcal{T}_2$, et en étendant P par linéarité. Réciproquement pour tester si P est une probabilité produit, il

Exemple 3.9 (Arbre de probabilité)**.** Si $\Omega = \coprod_{i\in \mathrm{I}} \mathrm{A}_i$, $\mathrm{A}_i = \coprod_j \mathrm{A}_{ij}$, $\mathrm{A}_{ij} = \coprod_k \mathrm{A}_{ijk}$, alors $\Omega = \coprod_{i,j,k} \mathrm{A}_{ijk}$. On peut alors calculer

$$\mathrm{P}(\mathrm{A}) = \sum_{i,j,k} \mathrm{P}(\mathrm{A} \mid \mathrm{A}_{ijk})\mathrm{P}(\mathrm{A}_{ijk}) = \sum_{i,j,k} \mathrm{P}(\mathrm{A} \mid \mathrm{A}_{ijk})\mathrm{P}(\mathrm{A}_{ijk} \mid \mathrm{A}_{ij})\mathrm{P}(\mathrm{A}_{ij} \mid \mathrm{A}_i)\mathrm{P}(\mathrm{A}_i)$$

Note : plus généralement, faire les produits à partir d'un noeud de l'arbre (plutôt que la racine) permet de retrouver les probabilités conditionnelles à ce noeud : $\mathrm{P}(\mathrm{A}_{ijk} \mid \mathrm{A}_i) = \mathrm{P}(\mathrm{A}_{ijk} \mid \mathrm{A}_{ij})\mathrm{P}(\mathrm{A}_{ij} \mid \mathrm{A}_i)$.

3.4 Indépendence

Definition 3.10 (Indépendence)**.**

- A est indépendant de B $\Leftrightarrow \mathrm{P}(\mathrm{A} \cap \mathrm{B}) = \mathrm{P}(\mathrm{A})\mathrm{P}(\mathrm{B}) \Leftrightarrow \mathrm{P}(\mathrm{A} \mid \mathrm{B}) = \mathrm{P}(\mathrm{A})$ (ou $\mathrm{P}(\mathrm{B}) = 0$).
- $\mathrm{A}_1, \dots, \mathrm{A}_n$ sont indépendantes si $\mathrm{P}(\mathrm{A}_{i_1} \cap \dots \cap \mathrm{A}_{i_k}) = \mathrm{P}(\mathrm{A}_{i_1}) \cdot \dots \cdot \mathrm{P}(\mathrm{A}_{i_k})$ pour tout $\{i_1, \dots, i_k\} \subset \{1, \dots, n\}$.

Propriétés 3.11.

- *A et B sont indépendants* $\Leftrightarrow \overline{\mathrm{A}}$ *et* B *le sont* $\Leftrightarrow$ A *et* $\overline{\mathrm{B}}$ *le sont* $\Leftrightarrow \overline{\mathrm{A}}$ *et* $\overline{\mathrm{B}}$ *le sont ;*
- *Si* $\mathrm{A}_1, \dots, \mathrm{A}_n$ *sont indépendants, ils le sont deux à deux (attention, la réciproque est fausse) ;*
- *Attention : on peut avoir* A *indépendant de* B *et* B *indépendant de* C *mais* A *non indépendant de* C.

4 Variable aléatoire

Definition 4.1.

- Une variable aléatoire réelle est une fonction $\mathrm{X} : \Omega \to \mathbb{R}$.[2]
- X induit une probabilité sur $\mathbb{R}$: $\mathrm{P}(\mathrm{X} \in \mathrm{A}) := \mathrm{P}(\mathrm{X}^{-1}(\mathrm{A})$ (aussi noté $\mathrm{P}_\mathrm{X}(\mathrm{A})$).[3]
- On dit que X a une loi $\mathcal{L}$ si $\mathrm{P}_\mathrm{X} = \mathcal{L}$. Notation : $\mathrm{X} \sim \mathcal{L}$;
- Un vecteur aléatoire réel est un n-uplet[4] $\mathrm{X} = (\mathrm{X}_1, \dots, \mathrm{X}_n)$ de fonctions : $\mathrm{X} \to \mathbb{R}^n$.

Exemple 4.2. Pour un jeu de pile ou face, on peut prendre la variable aléatoire X représentant les gains du jeu : X(pile) $= 10$, X(face) $= -10$.

4.1 Indépendance

Definition 4.3.

- X et Y sont indépendantes ssi $\forall a, b \in \mathbb{R}$,

$$\mathrm{P}(\mathrm{X} \in [a_1; b_1] \cap \mathrm{Y} \in [a_2, b]_2) = \mathrm{P}(\mathrm{X} \in [a_1; b_1])\mathrm{P}(\mathrm{Y} \in [a_2, b]_2);$$

- $\mathrm{X}_1, \dots, \mathrm{X}_n$ sont des variables aléatoires indépendantes ssi

$$\mathrm{P}(\cap \mathrm{X}_i \in [a_i, b_i]) = \prod_{i=1}^{n} \mathrm{P}(\mathrm{X}_i \in [a_i, b_i]).$$

- Si $\mathrm{X}_1, \dots, \mathrm{X}_n$ sont indépendantes, alors elles sont indépendantes deux à deux (attention, la réciproque est fausse).

Exemple 4.4. Deux événements A et B sont indépendants $\Leftrightarrow$ les variables aléatoires $\mathbb{1}_\mathrm{A}$ et $\mathbb{1}_\mathrm{B}$ le sont.

Propriétés 4.5.

- $\mathrm{X}_1, \dots, \mathrm{X}_n$ *sont indépendantes* $\Leftrightarrow \mathrm{P}_{(\mathrm{X}_1, \dots, \mathrm{X}_n)}$ *est la probabilité produit* $\mathrm{P}_{\mathrm{X}_1} \times \dots \times \mathrm{P}_{\mathrm{X}_n}$.
- *Attention : la loi de* $\mathrm{X}_1, \dots, \mathrm{X}_n$ *ne détermine pas en général la loi de* $(\mathrm{X}_1, \dots, \mathrm{X}_n)$. *Mais si les* X_i *sont indépendantes, la loi de* $(\mathrm{X}_1, \dots, \mathrm{X}_n)$ *est déterminée par la loi produit des* X_i.

Proposition 4.6. *Si* $\mathrm{X}_1, \dots, \mathrm{X}_n$ *sont indépendantes, toute fonction* $f(\mathrm{X}_1, \dots, \mathrm{X}_p)$ *est indépendante de* $g(\mathrm{X}_{p+1}, \dots, \mathrm{X}_n)$.

---

suffit de vérifier $\mathrm{P}(\mathrm{A} \times \mathrm{B}) = \mathrm{P}_1(\mathrm{A})\mathrm{P}_2(\mathrm{B})$ sur des générateurs de $\mathcal{T}_1$ et $\mathcal{T}_2$.

[2]★ qui est mesurable, ie satisfait $\mathrm{X}^{-1}([a, b]) \in \mathcal{T}$ pour tout $a, b \in \mathbb{R}$ [3]★ la tribu sur $\mathbb{R}$ est la tribu des Boréliens [4]★ mesurable

4.2 ☆ Fonction de répartition

- $F : \mathbb{R} \to [0;1], x \mapsto P(X \leq x)$.
- F est croissante, continue à droite, tend vers 0 en $-\infty$ et 1 en $+\infty$ et caractérise P_X car $P(X \in]a;b]) = F(b) - F(a)$.
- Réciproquement, toute fonction $F : \mathbb{R} \to [0;1]$ satisfaisant ces propriétés vient d'une variable aléatoire X.
- F a une limite à gauche en tout point : $F(x) - \lim_{t \to x^-} F(t) = P(X = x)$.

5 Espérance et variance

5.1 Espérance

Definition 5.1.

- Si X est une variable aléatoire finie, $X(\Omega) = \{x_1, \dots, x_n\}$, $E(X) = \sum_{i=1}^n x_i P(X = x_i)$.
- Si X est une variable aléatoire discrète, $X(\Omega) = \{x_1, \dots, x_n, \dots\}$, $E(X) = \sum_{n=1}^\infty x_i P(X = x_i)$. **Attention** : l'espérance n'a de sens que si la série est absolument convergente.
- Si X est à densité f, $E(X) = \int_{-\infty}^\infty t f(t) dt$. **Attention** : l'espérance n'a de sens que si cette intégrale converge absolument.

Propriétés 5.2.

- *Linéarité : si* X *et* Y *ont une espérance, alors* $\alpha X + \beta Y$ *aussi et* $E(\alpha X + \beta Y) = \alpha E(X) + \beta E(Y)$.
- *Si* $X = \sum_{i \in I} x_i \mathbb{1}_{A_i}$ *(avec les* A_i *disjoints*[5]*) et que l'espérance de* X *existe, alors* $E(X) = \sum x_i P(A_i)$
- *Si on a un arbre de probabilité,* $E(X) = \sum_{ijk} E(X \mid A_{ijk}) P(A_{ijk}) = \sum_{ijk} E(X \mid A_{ijk}) P(A_{ijk} \mid A_{ij}) P(A_{ij} \mid A_i) P(A_i)$.
- *Si* X *et* Y *sont indépendantes et ont une espérance, alors* XY *aussi et* $E(XY) = E(X)E(Y)$.
- *Transfert, cas* X *fini ou discret : si* $g : \mathbb{R} \to \mathbb{R}$ *est une fonction continue par morceaux,* $E(g(X)) = \sum_{i \in I} g(x_i) P(X = x_i)$.
- *Transfert, cas* X *à densité : si* $g : \mathbb{R} \to \mathbb{R}$ *est une fonction continue par morceaux,* $E(g(X)) = \int_{-\infty}^\infty g(t) f(t) dt$.

5.2 Variance

Definition 5.3.

- Variance : $V(X) = E((X - E(X))^2)$. Attention : la variance n'a de sens que si $E(X)$ existe et si la variable aléatoire $(X - E(X))^2$ a une espérance ;
- Variance, cas X fini ou discret : $V(X) = \sum_{i \in I} (x_i - E(X))^2 P(X = x_i)$;
- Variance, cas X à densité : $V(X) = \int_{-\infty}^\infty (t - E(X))^2 f(t) dt$;
- Écart type : $\sigma = \sqrt{V(X)}$.

Propriétés 5.4.

- $V(X) \geq 0$ *;* $V(X) = 0 \Leftrightarrow X$ *est constante ;*
- *Si* X *et* Y *sont indépendantes,* $V(X + Y) = V(X) + V(Y)$ *;*
- *Koenig-Huygens :* $V(X) = E(X^2) - (E(X))^2$.
- $V(aX + b) = a^2 V(X)$.
- *Variable centrée réduite :* $E(X) = 0$ *et* $V(X) = 1$. *Exemple :* $\frac{1}{\sigma(X)}(X - E(X))$ *est toujours centrée réduite.*

5.3 ☆ Moments

Definition 5.5.

[5] par exemple si X est une va discrète, $A_i = X^{-1}(x_i)$

- Moment d'ordre $k \in \mathbb{N}$ = espérance de X^k (si elle existe).
- Cas X fini ou discret : $E(X^k) = \sum_{n=1}^{\infty} x_i^k P(X = x_i)$.
- Cas X à densité : $E(X^k) = \int_{-\infty}^{\infty} t^k f(t)dt$.

Propriétés 5.6 (Relations entre les moments)**.**

- *Si* X *a un moment d'ordre k, il a un moment d'ordre* k' *pour tout* $k' \leq k$ *(car* $E(|X^{k'}|)^{1/k'} \leq E(|X^k|)^{1/k}$*) ;*
- X *a un moment d'ordre* $2 \Leftrightarrow$ *la variance de* X *existe.*
- ★ *Hölder : si* $1/p + 1/q = 1$, $E(|XY|) \leq E(|X|^p)^{1/p} E(|Y|^q)^{1/q}$
- ★ *Minkowski :* $E(|X+Y|^p)^{1/p} \leq E(|X^p|)^{1/p} + E(|Y^p|)^{1/p}$

5.4 ☆ Covariance

Definition 5.7.

- Covariance : $Cov(X, Y) = E(XY) - E(X)E(Y)$
- Coefficient de corrélation : $\rho(X, Y) = \frac{Cov(X,Y)}{\sigma(X)\sigma(Y)} \leq 1$.

Propriétés 5.8.

- $Cov(X, Y) = E((X - E(X))(Y - E(Y)))$ *;*
- Cov *est bilinéaire, symétrique* ($Cov(X, Y) = Cov(Y, X)$*) et positive* ($Cov(X, X) = V(X) \geq 0$*) ;*
- $V(\sum_{i=1}^{n} X_i) = \sum_{1 \leq i,j \leq n} Cov(X_i, X_j) = \sum_{i=1}^{n} V(X_i) + \sum_{1 \leq i < j \leq n} Cov(X_i, X_j)$.
- *Si les* X_i *sont indépendantes deux à deux :* $E(X_i X_j) = E(X_i)E(X_j)$ *donc* $Cov(X_i, X_j) = 0$ *et* $V(\sum X_i) = \sum V(X_i)$.

Exemple 5.9 (Droite de régression linéaire)**.**

- Coefficient de corrélation : $\rho(X, Y) = \frac{cov(X,Y)}{\sigma_X \sigma_Y}$. $|\rho(X, Y)| \leq 1$ et on a égalité $\Leftrightarrow$ il existe une droite $aX + bY + c = 0$ telle que $P(aX + bY + c = 0) = 1$.
- Si X et Y variables aléatoires réelles avec des moments d'ordre 2. La droite $\Delta : Y = aX + b$ de régression linéaire minimisant les moindres carrés par rapport à Y (c'est à dire minimisant $E((Y - aX - b)^2)$) est donnée par $a = Cov(X, Y)/V(X)$ et $b = EY - aEX$, autrement dit $(Y - EY) = \rho(X, Y)\frac{\sigma_Y}{\sigma_X}(X - EX)$. Le minimum est donné par $V(Y)(1 - \frac{\rho(X,Y)}{V(X)V(Y)})$.
- On peut aussi définir la droite $\Delta' : X = a'Y + b'$ minimisant les moindres carrés par rapport à X. Alors $aa' = \rho(X, Y)^2$, donc $\Delta = \Delta' \Leftrightarrow aa' = 1 \Leftrightarrow$ les points sont alignés.
- Application : régression linéaire sur un nuage de n points $P_i = (x_i, y_i)$. On prend une loi $P((X, Y) = P_i$ proportionnelle à la fréquence de répétition de P_i.

6 Variable aléatoire finie ou discrète

Definition 6.1. X est une variable aléatoire finie (resp. discrète) si $X(\Omega)$ est fini (resp. dénombrable) : $X(\Omega) = \{x_i\}_{i \in I}$. P_X est alors complètement déterminé par $P(X = x_i)$ pour $x_i \in X(\Omega)$.

Propriétés 6.2.

- *Si* X *et* Y *sont discrètes (ou finies), le vecteur* (X, Y) *est déterminé par* $p_{ij} := P(X = x_i \cap Y = y_j)$ *;*
- X *et* Y *sont indépendantes* $\Leftrightarrow P(X = x_i \cap Y = y_j) = P(X = x_i)P(Y = y_j)$ *;*
- *La loi de* X *peut se retrouver à partir de la loi de* (X, Y) *:* $P(X = x_i) = \sum_{j \in J} P(X = x_i \cap Y = y_j) = \sum_{j \in J} p_{ij}$ *;*
- *Loi conditionnelle :* $P(Y = y_j \mid X = x_i) = \frac{p_{ij}}{\sum_{j' \in J} p_{ij'}}$.
- *Si* X *et* Y *sont indépendantes,* $X + Y$ *a pour loi* $P(X + Y = z) = \sum_{x+y=z} P(X = x)P(Y = y)$.

6.1 Loi uniforme $\mathrm{U}([1;n])$

- $\mathrm{X}(\Omega) = \{1, \dots, n\}, \mathrm{P}(\mathrm{X} = i) = 1/n$;
- $\mathrm{E}(\mathrm{X}) = \frac{n+1}{2}, \mathrm{V}(\mathrm{X}) = \frac{n^2-1}{12}$.

Exemple type : pile ou face, lancé de dé.

6.2 Loi de Bernoulli $\mathrm{B}(p)$

- $\mathrm{X}(\Omega) = \{0, 1\}, \mathrm{P}(\mathrm{X} = 0) = 1 - p, \mathrm{P}(\mathrm{X} = 1) = p$;
- $\mathrm{E}(\mathrm{X}) = p, \mathrm{V}(\mathrm{X}) = p(1 - p)$.

Interprétation : $\mathrm{X} = \mathbb{1}_\mathrm{A}$ où $\mathrm{A} \in \mathcal{T}$ est un évènement, $p = \mathrm{P}(\mathrm{A})$ (probabilité de tomber dans A).
Exemple type : un lancé de pile ou face biaisé.

6.3 Loi binomiale $\mathrm{B}(n;p)$

- $\mathrm{X}(\Omega) = \{0, \dots, n\}, \mathrm{P}(\mathrm{X} = k) = \binom{n}{k} p^k (1-p)^{n-k}$;
- $\mathrm{E}(\mathrm{X}) = np, \mathrm{V}(\mathrm{X}) = np(1 - p)$;
- ☆ Loi multinomiale : $\mathrm{P}(\mathrm{X} = (j_1, \dots, j_k)) = \frac{n!}{j_1! \dots j_k!} p_1^{j_1} \dots p_k^{j_k}$, où $n = j_1 + \dots + j_k$.

Interprétation : $\mathrm{X} = \mathrm{X}_1 + \dots + \mathrm{X}_n$ somme de variables aléatoires indépendantes de loi de Bernouilli de paramètre p (schéma de Bernouilli). Plus généralement la loi binomiale est stable par addition : si X_i sont indépendantes de loi $\mathrm{B}(n_i; p)$, $\mathrm{X} = \sum \mathrm{X}_i$ est de loi $\mathrm{B}(\sum n_i, p)$.

Exemple type : n lancés de pile ou face biaisé. (Si $n = 1$ on retrouve une loi de Bernouili). Pour la loi multinomiale tirage avec remise de boules de k couleurs différentes en proportions $p_1, \dots, p_k$).

Exemple 6.3. Si on a n boules, avec n_1 blanches et n_2 noires, X la variable aléatoire donnant le nombre de boules blanches quand on tire r boules avec remises, alors $\mathrm{X} \simeq \mathcal{B}(r, n_1/n)$.

6.4 ☆ Loi hypergéométrique $\mathcal{H}(\mathrm{N}, n, p)$

- $\mathrm{P}(\mathrm{X} = k) = \frac{\binom{\mathrm{N}p}{k}\binom{\mathrm{N}(1-p)}{n-k}}{\binom{\mathrm{N}}{n}}$;
- $\mathrm{E}(\mathrm{X}) = np, \mathrm{V}(\mathrm{X}) = \frac{\mathrm{N}-n}{\mathrm{N}-1} np(1 - p)$.

Exemple type : on a N boules dans une urne, avec une proportion p de boules rouges. On tire n boules sans remises, X est la variable aléatoire égale au nombre de boules rouges obtenues, $\mathrm{X} \simeq \mathcal{H}(\mathrm{N}, n, p)$.

6.5 ☆ Loi géométrique $\mathcal{G}(p)$

- $\mathrm{X}(\Omega) = \mathbb{N}^*, \mathrm{P}(\mathrm{X} = k) = p(1-p)^{k-1}$;
- $\mathrm{E}(\mathrm{X}) = \frac{1}{p}, \mathrm{V}(\mathrm{X}) = \frac{1-p}{p^2}$.

Interprétation : rang d'apparition du premier succès d'une loi de Bernoulli $\mathrm{B}(p)$ itérée infiniment.
Exemple type : lancé de dé jusqu'à obtenir un 6.

6.6 Loi de Poisson $\mathcal{P}(\lambda)$

- $\mathrm{X}(\Omega) = \mathbb{N}, \mathrm{P}(\mathrm{X} = k) = e^{-\lambda} \frac{\lambda^k}{k!}$;
- $\mathrm{E}(\mathrm{X}) = \lambda, \mathrm{V}(\mathrm{X}) = \lambda$.

Interprétation : modélise une file d'attente sans mémoire, ou plus généralement le comportement du nombre d'événements se produisant dans un intervalle de temps fixé, si ces événements se produisent avec une fréquence moyenne ou espérance connue, et indépendamment du temps écoulé depuis l'événement précédent.

Exemple type : la somme S d'un grand nombre de variables de Bernouilli indépendantes de petit paramètres suit (approximativement) une loi de Poisson.

Stabilité par la somme : $\mathrm{X} = \sum_i \mathrm{X}_i$ somme de variables aléatoires indépendantes de Poisson de paramètre λ_i ($\mathrm{X}_i \sim \mathcal{P}(\lambda_i)$, $\mathrm{X} \sim \mathcal{P}(\sum \lambda_i)$.

Exemple 6.4. A un carrefour routier, où il y a deux directions A et B. Un véhicule suit la direction A avec probabilité p. Il y a N véhicules par heure, où N suit une loi de Poisson de paramètre λ. Alors le nombre de véhicule suivant la direction A suit une loi de Poisson de paramètre λp. La loi de n sachant qu'il y a k véhicule ayant pris la direction A est une loi de Poisson de paramètre $\lambda(1-p)$ translatée de k.

6.7 ★ Fonctions génératrices

- Fonction génératrice : $\mathrm{G} = \mathrm{E}(z^{\mathrm{X}}) = \sum_{k=0}^{\infty} p_k z^k$ où $p_k = \mathrm{P}(\mathrm{X} = k)$.
- X et Y sont indépendantes $\Leftrightarrow \mathrm{E}(z_1^{\mathrm{X}} z_2^{\mathrm{Y}}) = \mathrm{E}(z_1^{\mathrm{X}})\mathrm{E}(z_2^{\mathrm{Y}})$;
- Si X et Y sont indépendantes, $\mathrm{G}_{\mathrm{X}+\mathrm{Y}} = \mathrm{G}_{\mathrm{X}}\mathrm{G}_{\mathrm{Y}}$;
- X admet une espérance ssi G' a une limite à gauche en 1 : $\mathrm{E}(\mathrm{X}) = \lim_{z\to 1^-} \mathrm{G}'(z)$.
- X admet un moment d'ordre 2 ssi G" a une limite à gauche en 1 : $\mathrm{E}(\mathrm{X}^2) - \mathrm{E}(\mathrm{X}) = \lim_{z\to 1^-} \mathrm{G}"(z)$.
- X admet un moment d'ordre k ssi $\mathrm{G}^{(k)}$ a une limite à gauche en 1 : $\mathrm{E}(\mathrm{X}(\mathrm{X}-1)\dots(\mathrm{X}-k+1)) = \lim_{z\to 1^-} \mathrm{G}^{(k)}(z)$.
- Rappel : pour une série entière G à coefficients positifs de rayon $\mathrm{R} > 0$, G a une limite à gauche en R $\Leftrightarrow \sum a_n \mathrm{R}^n$ converge.
- Si $\mathrm{S} = \sum_{i=1}^{\mathrm{T}} \mathrm{X}_i$ où les X_i sont indépendantes de même loi, et T est une variable aléatoire à valeur dans $\mathbb{N}^*$, $\mathrm{G}_{\mathrm{S}} = \mathrm{G}_{\mathrm{T}} \circ \mathrm{G}_{\mathrm{X}_1}$. Et (Wald) $\mathrm{E}(\mathrm{S}) = \mathrm{E}(\mathrm{X}_1)\mathrm{E}(\mathrm{T})$ et $\mathrm{V}ar(\mathrm{S}) = \mathrm{V}ar(\mathrm{X}_1)\mathrm{E}(\mathrm{T}) + \mathrm{V}ar(\mathrm{T})\mathrm{E}(\mathrm{X}_1)^2$.
- Si $\mathrm{X} \sim \mathcal{B}(p)$, $\mathrm{G}(t) = 1 - p + pt$;
- Si $\mathrm{X} \sim \mathcal{B}(n, p)$, $\mathrm{G}(t) = (1 - p + pt)^n$ (plus généralement la fonction génératrice de $\sum \mathrm{X}_i$ où les X_i sont indépendantes sera $\mathrm{G} = \prod \mathrm{G}_i$) ;
- Si $\mathrm{X} \sim \mathcal{G}(p)$, $\mathrm{G}(t) = \frac{pt}{1-(1-p)t}$;
- Si $\mathrm{X} \sim \mathcal{P}(\lambda)$, $\mathrm{G}(t) = e^{\lambda(t-1)}$.

7 Variable aléatoire à densité

Definition 7.1. $\mathrm{X} : \Omega \to \mathbb{R}$ est à densité f (une fonction continue par morceaux) si $\mathrm{P}(\mathrm{X} \in [a, b]) = \int_{t=a}^{b} f(t)dt$.
Dans ce cas f est positive[6] et satisfait $\int_{-\infty}^{\infty} f(t)dt = 1$.

Propriétés 7.2.

- $\mathrm{E}(\mathrm{X}) = \int_{-\infty}^{\infty} tf(t)dt$ *(si cette intégrale converge absolument).*
- *Fonction de répartition :* $\mathrm{F}(x) = \int_{-\infty}^{x} f(t)dt$. *Elle est continue et satisfait* $\mathrm{F}'(x_0) = f(x_0)$ *en les points de continuité de* f *;*
- $\mathrm{P}(\mathrm{X} \in]a, b[) = \mathrm{P}(\mathrm{X} \in [a, b]) = \mathrm{F}(b) - \mathrm{F}(a)$ *;*

7.1 ☆ Loi uniforme sur $[a; b]$

- Densité : $\frac{1}{b-a}\mathbb{1}_{[a;b]}$.
- Fonction de répartition : $\mathrm{F}(x) = 0$ si $x < a$, $\mathrm{F}(x) = \frac{1}{b-a}(x-a)$ si $a \le x \le b$, $\mathrm{F}(x) = 1$ si $x > b$.
- $\mathrm{E}(\mathrm{X}) = \frac{a+b}{2}$, $\mathrm{V}(\mathrm{X}) = \frac{(b-a)^2}{12}$.

Exemple 7.3. On peut aussi considérer des lois uniformes sur des parties du plan, par exemple si on joue aux fléchettes, on peut regarder la loi uniforme sur la cible (la probabilité de tomber dans une zone de la cible étant proportionnelle à son aire).

7.2 Loi exponentielle

- Densité : $f(x) = \lambda e^{-\lambda x}$ si $x \le 0$ (et $f(x) = 0$ si $x < 0$).
- Fonction de répartition : $\mathrm{F}(x) = 0$ si $x < 0$ et $\mathrm{F}(x) = 1 - e^{-\lambda x}$ sinon.
- $\mathrm{E}(\mathrm{X}) = \frac{1}{\lambda}$, $\mathrm{V}(\mathrm{X}) = \frac{1}{\lambda^2}$.
- Loi sans mémoire : $\forall s \in \mathbb{R}, t \ge 0, \mathrm{P}(\mathrm{X} > s + t \mid \mathrm{X} > t) = \mathrm{P}(\mathrm{X} > s)$ (et réciproquement une variable aléatoire sans mémoire suit une loi exponentielle).

[6]★ quitte à la changer presque partout

7.3 ☆ Loi normale (ou de Laplace-Gauss)

- Densité dans le cas centré réduite ($X \sim \mathcal{N}(0,1)$) : $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$;
- Fonction de répartition : donnée par une table ;
- Intervalle de confiance : pour tout $\alpha \in]0,1[$, il existe un unique $u_\alpha > 0 telque P(-u_\alpha < X < u_\alpha) = 1 - \alpha$;
- Cas général d'une loi normale $X \sim \mathcal{N}(m,\sigma)$ d'espérance m et d'écart type σ (donc de variance σ^2) : $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2}}$;
- $X \sim \mathcal{N}(m,\sigma) \Leftrightarrow \frac{X-m}{\sigma} \sim \mathcal{N}(0,1)$.

7.4 ★ Vecteurs aléatoires à densité

Propriétés 7.4.

- *Si (X,Y) a pour densité $f(x,y)$, alors X a pour densité $f_Y : x \mapsto \int_{-\infty}^{\infty} f(x,y)dy$;*
- *Loi conditionnelle :* $P(Y \in B \mid X \in A) = \frac{\iint_{A \times B} f(x,y)dxdy}{\int_{x \in A} f_Y(x)dx}$.
- *Si X de densité f est indépendante de Y de densité g, $X+Y$ a pour densité la convolution $f \star g : t \mapsto \int_{-\infty}^{\infty} f(x)g(t-x)dt$;*
- *Transfert : si $(X_1, \dots, X_p)$ a pour densité f, alors $E(g(X)) = \iint g(x_1,\dots,x_p) f(x_1,\dots,x_p) dx_1 \dots dx_p$.*
- *Si (X_i) sont indépendantes de densité f_i, alors $X = (X_1,\dots,X_p)$ a pour densité $f = f_1 \dots f_p$.*
- *Réciproquement si X a pour densité $f_1(x_1)\dots f_p(x_p)$, alors les X_i sont indépendantes et $f_{X_i} = \lambda_i f_i$ (pour des $\lambda_i \in \mathbb{R}_+^*$ tels que $\prod \lambda_i = 1$).*

Definition 7.5 (Loi normale centrée vectorielle)**.**

- Loi normale vectorielle centrée : densité $f(x) = \frac{1}{\alpha} e^{-1/2Q(x)}$ où Q forme quadratique définie positive.
- Si $\Sigma_{ij} = Cov(X_i, X_j)$ (matrice de variances-covariances), alors $f(x) = \frac{1}{\sqrt{2\pi}^p} \frac{1}{\det \Sigma} e^{-1/2 t_x \Sigma^{-1} x}$.
- Les composantes X_i suivent une loi normale, et elles sont indépendantes ssi Σ est diagonale ;
- X est normale $\Leftrightarrow$ toute combinaison linéaire $Y = \sum a_i X_i$ est normale.

8 ★ Convergence de variables aléatoires

Definition 8.1.

- Convergence en probabilité $X_n \to X$: $\forall \alpha > 0, P(|X_n - X| \geq \alpha) \xrightarrow[n\to\infty]{} 0$;
- Convergence en loi $X_n \to X$: si F_n est la fonction de répartition de X_n et F celle de X alors $F_n(x) \to F(x)$ en tout point de continuité de F ;

Exemple 8.2 (Convergence de sommes de variables aléatoires indépendantes)**.**

- Approximation de la loi hypergéométrique par la loi binomiale : si $X_N \sim \mathcal{H}(N,n,p)$, X_N converge en loi vers $X \sim B(n,p)$
- Approximation de la loi binomiale par la loi de Poisson : si $X_n \sim \mathcal{B}(n, \lambda/n)$[7], X_n converge en loi vers $X \sim \mathcal{P}(\lambda)$.[8]
- Approximation de la loi géométrique par la loi exponentielle : si $X_n \sim \mathcal{G}(p_n)$ où $\lim_n p_n = 0$ et $\lim_n p_n/a_n = \lambda > 0$ alors $a_n X_n$ converge en loi vers la loi exponentielle de paramètre λ.
- Inversement, si X suit une loi exponentielle de paramètre λ, l'échantillon discret $Y = \lceil \theta X \rceil$ ($\theta > 0$) suit une loi géométrique de paramètre $p = 1 - e^{-\lambda/\theta}$.

Propriétés 8.3.

- *Inégalité de Markov : Si $|X|$ a une espérance m, $\forall a > 0, P(|X| \geq a) \leq m/a$;*

[7] ou plus généralement X_n de paramètre p_n tel que $np_n \to \lambda$

[8] En pratique on remplace la loi binomiale par la loi de Poisson dès que $n > 30$ et $np < 5$ ou dès que $n > 50$ et $p < 0.1$. Et on la remplace par

- *Inégalité de Bienaymé-Techbychev : si* X *a une espérance* m *et une variance* σ^2*, alors* $\forall a > 0, \mathrm{P}(|\mathrm{X} - m| \geq a) \leq \frac{\sigma^2}{a^2}$ *;*
- *Loi faible des grands nombres : si* X_n *indépendantes de même espérance* m *et variance* σ*,* $\frac{1}{n}\sum_{i=1}^n \mathrm{X}_i \to m$ *en probabilité ;*
- *Loi forte des grands nombres : si* X_n *indépendantes de même loi, admettant une variance finie et une espérance* m*, alors* $\frac{1}{n}\sum_{i=1}^n \mathrm{X}_i \to m$ *presque sûrement.*
- *Théorème central limite (TCL) : si* X_n *suite de variable aléatoires indépendantes de même loi, d'espérance* m *et de variance* σ^2*, alors la réduction centrée réduite de* $\sum_{i=1}^n \mathrm{X}_i$ *converge en loi vers* $\mathcal{N}(0,1)$ *:*
$$\mathrm{P}\left(\frac{\sum_{i=1}^n \mathrm{X}_i - nm}{\sqrt{n}\sigma} \leq x\right) \to \int_{-\infty}^{x} \frac{1}{2\sqrt{\pi}} e^{-\frac{t^2}{2}} dt.$$
- *Convergence presque sûre* $\Rightarrow$ *convergence en probabilité* $\Rightarrow$ *convergence en loi ;*
- *Convergence en moment d'ordre* k *(nome* $\mathcal{L}^k$*)* $\Rightarrow$ *convergence en moment d'ordre* k' *pour* $k' \leq k \Rightarrow$ *convergence en probabilité* $\Rightarrow$ *convergence en loi.*

Exemple 8.4 (Échantillons)**.** Échantillon : si X est de moyenne inconnue m de d'écart type σ, on estime m par $\overline{\mathrm{X}}_n = \frac{1}{n}\sum_{i=1}^n \mathrm{X}_i$. Pour n grand, $\mathrm{P}(|\overline{\mathrm{X}}_n - m| \geq \alpha) \simeq 1 - (\Phi(\sqrt{n}\alpha/\sigma) - \Phi(-\sqrt{n}\alpha/\sigma)) = 2(1 - \Phi(\sqrt{n}\alpha/\sigma))$ où $\Phi \sim \mathcal{N}(0,1)$. Donc le plus petit entier tel que $\mathrm{P}(|\overline{\mathrm{X}}_n - m| \geq \alpha) \leq \beta$ est $n = \lceil \frac{\sigma^2}{\alpha^2}\Phi^{-1}(1 - \beta/2)^2 \rceil$.

Taux de confiance usuels : $\Phi(1.64) - \Phi(-1.64) = 0.9$ $0.9 : \Phi^{-1}(0.95) = 1.64)$, $\Phi(1.96) - \Phi(-1.96) = 0.95$, $\Phi(3.09) - \Phi(-3.09) = 0.99$.

Application numérique : si $\sigma/\alpha = 10$, pour avoir $\mathrm{P}(|\overline{\mathrm{X}}_n - m| \geq \alpha) \leq 0.05$ il faut $n \geq 10^2\Phi^{-1}(0.975)^2 = 100 * 1.96^2 = 384.16$.

9 Références

9.1 Livres

- *Probabilités et statistiques pour le capes et l'agrégation interne*, Jérôme Escoffier. (Avec exercices corrigés).
- *Probabilités 1, Licence capes*, Jean-Yves Ouvrard. (Avec exercices corrigés).

9.2 Polycopiés

- https://www.lpsm.paris/pageperso/roux/enseignements/1516/capes_probas/poly_probas_avec_solutions.pdf. *Probabilités et statistique pour le CAPES*, Beatrice de Tilièere et Frederique Petit, 149p., 2016. (Avec exercices corrigés).
- https://www-fourier.ujf-grenoble.fr/~dpiau/capes08/fprobas08.pdf. Liste de leçons de probabilités et des exercices non corrigés.
- http://math.univ-lyon1.fr/~duheille/cours-capes.pdf, *Préparation au Capes de mathématiques, Probabilités*, F. Bienvenue-Duheille, 46p. (2007-2008). Attention : niveau largement au-dessus du programme.

$\mathcal{N}(np, np(1-p))$ dès que $n \geq 30$, $np \geq 15$ et $np(1-p) \geq 5$.