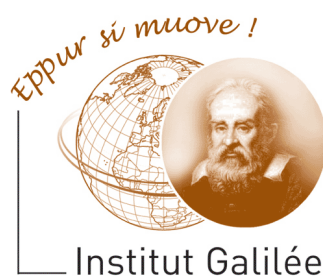


Cours Commun Scientifique
de
Probabilités & Statistiques

Résumé de cours
Suivi des fiches d'exercices

Responsable du cours : Laurent Tournier



Ce document, ainsi éventuellement que d'autres à venir, peut être trouvé au format PDF à l'adresse suivante :

<http://www.math.univ-paris13.fr/~tourner/enseignement>

1 Espaces de probabilités. Loi binomiale

1.1 Définitions

Définition

Un **espace de probabilité** (Ω, P) est constitué de

- Ω , un ensemble (l'ensemble des résultats d'une expérience aléatoire)
- P , une probabilité sur Ω .

Un élément $\omega \in \Omega$ est appelé une **réalisation**, c'est un résultat possible d'une expérience aléatoire.

Un sous-ensemble $A \subset \Omega$ est appelé un **événement**. C'est un ensemble de réalisations (par exemple, celles qui vérifient une certaine condition). L'ensemble des événements est donc l'ensemble $\mathcal{P}(\Omega)$ des parties (ou sous-ensembles) de Ω .

Définition

Une **probabilité** sur Ω est une application $P : \mathcal{P}(\Omega) \rightarrow [0,1]$, définie sur les événements, telle que

1. $P(\Omega) = 1$

2. pour toute suite $(A_n)_n$ d'événements disjoints deux à deux, $P\left(\bigcup_n A_n\right) = \sum_n P(A_n)$.

Si un événement A vérifie $P(A) = 0$, on dit que A est **négligeable**; et si $P(A) = 1$, on dit que A est **presque sûr**, ou que A a lieu presque sûrement, abrégé « p.s. ».

Propriétés

(i) $P(\emptyset) = 0$

(ii) Pour tout événement A , $P(A^c) = 1 - P(A)$

(iii) Si $A \subset B$, alors $P(A) \leq P(B)$

(iv) Pour tous événements A et B , $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

(v) Pour toute suite croissante $(A_n)_n$ d'événements (c'est-à-dire $A_n \subset A_{n+1}$ pour tout n),

$$P\left(\bigcup_n A_n\right) = \lim_{n \uparrow} P(A_n)$$

(vi) Pour toute suite décroissante $(A_n)_n$ d'événements (c'est-à-dire $A_{n+1} \subset A_n$ pour tout n),

$$P\left(\bigcap_n A_n\right) = \lim_{n \downarrow} P(A_n)$$

Pour simplifier, on suppose dans ce cours que tout ensemble de réalisations est un événement. En réalité, ceci n'est plus possible dans le cas par exemple où $\Omega = [0,1]$ est muni de la probabilité uniforme car on ne peut pas définir l'intégrale sur n'importe quel ensemble mais seulement sur des ensembles « mesurables ». En pratique, ceci n'est pas une limitation car tous les ensembles que l'on utilise sont mesurables. Néanmoins, pour définir *rigoureusement* la théorie des probabilités, on appellerait espace de probabilités un *triplet* (Ω, \mathcal{A}, P) où \mathcal{A} est l'ensemble des événements, qui doit être une **tribu** sur Ω , c'est-à-dire un ensemble de parties de Ω tel que a) $\Omega \in \mathcal{A}$, b) si $A \in \mathcal{A}$ alors $A^c \in \mathcal{A}$, et c) si $A_n \in \mathcal{A}$ pour tout $n \in \mathbb{N}$, alors $\bigcup_n A_n \in \mathcal{A}$; et P est uniquement définie sur \mathcal{A} . Puis, sur \mathbb{R} , on définirait la **tribu borélienne**, qui est la plus petite tribu contenant les intervalles.

1.2 Cas élémentaire : équiprobabilité

On suppose que $\text{Card } \Omega = n$, avec $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$.

Définition

La **probabilité uniforme** sur Ω (ou distribution équiprobable) est la probabilité P définie par : pour tout $A = \{\omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_k}\} \subset \Omega$,

$$P(A) = \frac{k}{n} = \frac{\text{Card } A}{\text{Card } \Omega}.$$

Autrement dit, $P(\text{événement}) = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}}$.

Rappels de dénombrement :

Proposition

Soit E un ensemble fini.

– Une **permutation** de E est une façon d'ordonner les éléments de E .

Le nombre de permutations d'un ensemble à n éléments est

$$n! = 1 \times 2 \times 3 \times \dots \times (n-2) \times (n-1) \times n.$$

– Un **arrangement** de k éléments de E est une façon de choisir et d'**ordonner** k éléments de E : c'est une suite de k éléments de E distincts 2 à 2.

Le nombre d'arrangements de k éléments parmi n éléments (où $0 \leq k \leq n$) est

$$A_n^k = n(n-1) \dots (n-k+1) = \frac{n!}{(n-k)!}.$$

– Une **combinaison** de k éléments de E est une façon de choisir k éléments de E , **sans spécifier d'ordre** : c'est un sous-ensemble de E à k éléments.

Le nombre de combinaisons de k éléments parmi n éléments (où $0 \leq k \leq n$) est

$$\binom{n}{k} = C_n^k = \frac{n(n-1) \dots (n-k+1)}{k!} = \frac{n!}{k!(n-k)!}.$$

On peut dire aussi qu'un arrangement correspond à un tirage de k éléments un par un (et sans remise) en mémorisant l'ordre de tirage, tandis qu'une combinaison correspond à un tirage de k éléments simultanément.

Un arrangement de n éléments parmi n est une permutation, donc $A_n^n = n!$.

1.3 Probabilités conditionnelles

Définition

Soit B un événement tel que $P(B) > 0$. L'application $P(\cdot|B) : \mathcal{P}(\Omega) \rightarrow [0,1]$ définie par

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

est une probabilité sur Ω .

$P(A|B)$ est appelée la **probabilité conditionnelle de A sachant B** .

Proposition

On a

$$P(A \cap B) = P(A|B)P(B) = P(B|A)P(A).$$

Définition

Un **système complet d'événements** est une partition $(A_n)_n$ de Ω , c'est-à-dire une suite finie ou infinie $(A_n)_n$ d'événements disjoints, dont la réunion est Ω :

$$\text{pour tous } i \neq j, A_i \cap A_j = \emptyset, \quad \text{et} \quad \Omega = \bigcup_n A_n.$$

Par exemple, pour tout événement B , le couple (B, B^c) est un système complet d'événements.

Théorème (Théorème des probabilités totales)

Soit $(A_n)_n$ un système complet d'événements. Pour tout événement A ,

$$P(A) = \sum_n P(A|A_n)P(A_n).$$

En particulier, pour tous événements A et B ,

$$P(A) = P(A|B)P(B) + P(A|B^c)P(B^c).$$

Théorème (Formule de Bayes)

Soit $(A_n)_n$ un système complet d'événements. Pour tout événement A , et tout événement A_i du système,

$$P(A_i|A) = \frac{P(A|A_i)P(A_i)}{\sum_n P(A|A_n)P(A_n)}.$$

En particulier, pour tous événements A et B ,

$$P(B^c|A) = \frac{P(A|B^c)P(B^c)}{P(A|B)P(B) + P(A|B^c)P(B^c)}.$$

1.4 Événements indépendants

Définition

Deux événements A et B sont **indépendants** si

$$P(A \cap B) = P(A)P(B).$$

On a alors $P(A|B) = P(A)$ et $P(B|A) = P(B)$ (si $P(A) \neq 0$ et $P(B) \neq 0$).

Proposition

Si deux événements A et B sont indépendants, alors A^c et B^c le sont aussi.

Définition

Une famille $(A_i)_i$ d'événements est indépendante si pour toute sous-famille finie on a

$$P(A_{i_1} \cap A_{i_2} \cdots \cap A_{i_k}) = P(A_{i_1})P(A_{i_2}) \cdots P(A_{i_k}).$$

En particulier, des événements A , B et C sont indépendants si

$$P(A \cap B) = P(A)P(B), \quad P(B \cap C) = P(B)P(C), \quad P(A \cap C) = P(A)P(C) \\ \text{et} \quad P(A \cap B \cap C) = P(A)P(B)P(C).$$

1.5 Comptage d'événements indépendants réalisés : loi binomiale

On considère n tirages à Pile-ou-Face avec la même pièce biaisée, qui tombe sur Pile avec probabilité p et donc sur Face avec probabilité $1 - p$ (où $p \in]0,1[$ est fixé).

En notant 1 pour Pile et 0 pour Face, chaque réalisation ω se représente par une suite de 0 et de 1 de longueur n : par exemple, $\omega = (1,0,1,1, \dots)$ si le premier tirage est Pile, le deuxième Face, le troisième Pile, etc. Vu que les tirages sont indépendants, on a par exemple (ici, $n = 4$)

$$P(\{(1,0,1,1)\}) = p \times (1 - p) \times p \times p$$

et on voit qu'en général, si la séquence ω comporte k fois 1 (et donc $n - k$ fois 0),

$$P(\{\omega\}) = p^k(1 - p)^{n-k}.$$

Soit $0 \leq k \leq n$. On cherche la probabilité de l'événement

$$A_k = \{\text{Exactement } k \text{ pièces tombent sur Pile}\}.$$

On vient de voir que, pour toute séquence $\omega \in A_k$, $P(\{\omega\}) = p^k(1 - p)^{n-k}$. Par ailleurs, le nombre de telles séquences est $\text{Card } A_k = \binom{n}{k}$. On en déduit que

$$P(A_k) = \binom{n}{k} p^k (1 - p)^{n-k}.$$

En notant X le nombre de fois où Pile est apparu, on dira plus tard que X est une variable aléatoire qui suit la **loi binomiale** $\mathcal{B}(n, p)$.

Plus généralement, on peut bien sûr appliquer ce qui précède pour évaluer le nombre d'événements réalisés parmi une suite de n événements **indépendants** B_1, \dots, B_n ayant tous la **même probabilité** $P(B_i) = p$: on pose $A_k = \{k \text{ événements exactement parmi } B_1, \dots, B_n \text{ se réalisent}\}$. Ci-dessus, on avait $B_i = \{\text{le } i\text{-ième tirage est Pile}\}$.

2 Variables aléatoires. Généralités

Soit (Ω, P) un espace de probabilité.

Définition

Une **variable aléatoire** est une application $X : \Omega \rightarrow \mathbb{R}$.

Définition

Soit X une variable aléatoire. La **loi** de X est la probabilité P_X sur \mathbb{R} définie par :

$$\text{pour tout } B \subset \mathbb{R}, \quad P_X(B) = P(\{\omega \in \Omega \mid X(\omega) \in B\}) \\ = P(X \in B).$$

P_X peut aussi être vue comme une probabilité sur $X(\Omega)$, ensemble des valeurs prises par X , aussi appelé **support** de P_X . On note parfois $X \sim P_X$ pour indiquer que X suit la loi P_X .

La seconde égalité est une nouvelle notation : on note $\{X \in B\}$ l'événement formé des éventualités ω pour lesquelles $X(\omega) \in B$, et on abrège $P(\{X \in B\}) = P(X \in B)$.

Définition

Si A est un événement, on introduit la variable aléatoire **fonction indicatrice de A** , notée $\mathbf{1}_A$, qui indique si l'événement A est réalisé :

$$\text{pour tout } \omega \in \Omega, \quad \mathbf{1}_A(\omega) = \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{si } \omega \notin A. \end{cases}$$

2.1 Loys discrètes

Une variable aléatoire X est dite **discrète** si l'ensemble $X(\Omega)$ des valeurs qu'elle prend est *dénombrable* (c'est-à-dire que l'on peut trouver une suite qui énumère tous les éléments de $X(\omega)$: c'est le cas notamment si $X(\Omega)$ est un ensemble fini, \mathbb{N} , \mathbb{Z} ou \mathbb{Q} , mais pas l'intervalle $[0,1]$ ni \mathbb{R}). On dit aussi que la *loi* de X est discrète.

Si X est discrète, alors, pour tout $B \subset \mathbb{R}$, on peut calculer

$$P_X(B) = P(X \in B) = \sum_{x \in B \cap X(\Omega)} P(X = x) = \sum_{x \in B} P(X = x).$$

Pour caractériser une loi discrète, il suffit donc de se donner les **probabilités élémentaires**

$$p_X(x) = P(X = x) \quad \text{pour tout } x \in X(\Omega)$$

Définition-Proposition

Soit $E \subset \mathbb{R}$. Une famille $(p(x))_{x \in E} \in \mathbb{R}^E$ est une **famille de probabilités élémentaires** si

1. pour tout $x \in E$, $p(x) \geq 0$
2. $\sum_{x \in E} p(x) = 1$.

Dans ce cas, il existe une variable aléatoire X (sur un espace de probabilité (Ω, P)), à valeurs dans E , de probabilités élémentaires $p_X = p$, c'est-à-dire

$$\text{pour tout } x \in E, \quad P(X = x) = p(x).$$

2.2 Lois continues

Une variable aléatoire X est dite **continue** ou **à densité** s'il existe une fonction $f_X : \mathbb{R} \rightarrow \mathbb{R}$ telle que, pour tout $B \subset \mathbb{R}$,

$$P_X(B) = P(X \in B) = \int_B f_X(x) dx.$$

La fonction f_X (unique) est appelée la **densité** de X .

Définition-Proposition

Une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ est une **fonction de densité de probabilité** si

1. pour tout $x \in \mathbb{R}$, $f(x) \geq 0$

2. $\int_{\mathbb{R}} f(x) dx = 1$.

Dans ce cas, il existe une variable aléatoire X (sur un espace de probabilité (Ω, P)) de densité $f_X = f$.

Si X a une densité alors, pour tout $x \in \mathbb{R}$, $P(X = x) = \int_x^x f_X(t) dt$ donc

$$P(X = x) = 0.$$

De plus, si $f_X(x) = 0$ pour tout $x \in B$, alors $P(X \in B) = \int_B f_X(x) dx = 0$.

On en déduit que X est à valeurs dans $\{x \in \mathbb{R} \mid f_X(x) > 0\}$.

2.3 Fonction de répartition

Définition

Soit X une variable aléatoire. La **fonction de répartition** de X est la fonction $F_X : \mathbb{R} \rightarrow \mathbb{R}$ définie par

$$\text{pour tout } x \in \mathbb{R}, \quad F_X(x) = P(X \leq x).$$

Proposition

a) Soit X une variable aléatoire. Sa fonction de répartition F_X est une fonction croissante,

$$\lim_{x \rightarrow -\infty} F_X(x) = 0 \quad \text{et} \quad \lim_{x \rightarrow +\infty} F_X(x) = 1.$$

b) Si X et Y sont deux variables aléatoires telles que $F_X(t) = F_Y(t)$ pour tout $t \in \mathbb{R}$, alors X et Y ont même loi.

c) Si X est une variable aléatoire discrète, F_X est une fonction constante par morceaux, dont les sauts se situent aux points de $X(\Omega)$, et le saut en $x \in X(\Omega)$ a pour hauteur $P(X = x)$.

d) Si X est une variable aléatoire de densité f_X , on a

$$\text{pour tout } x \in \mathbb{R}, \quad F_X(x) = \int_{-\infty}^x f_X(t) dt$$

et on a la dérivée $(F_X)'(x) = f_X(x)$ (en tout point x où f_X est continue).

2.4 Espérance d'une variable aléatoire

Définition

L'**espérance** d'une variable aléatoire X , notée $E[X]$, est la moyenne de ses valeurs, pondérées par leurs probabilités.

Si X est discrète,

$$E[X] = \sum_{x \in X(\Omega)} xP(X = x).$$

Si X est continue, de densité f ,

$$E[X] = \int_{\mathbb{R}} xf(x)dx.$$

Attention. L'espérance n'est pas toujours définie. Il faut pour cela que la série ou l'intégrale ci-dessus converge absolument.

Propriétés

(i) Si X est constante, égale à $c \in \mathbb{R}$ (pour tout $\omega \in \Omega$, $X(\omega) = c$), alors $E[X] = E[c] = c$.

(ii) Pour tout événement $A \subset \Omega$, $E[\mathbf{1}_A] = P(A)$.

(iii) L'espérance est linéaire : pour toutes variables aléatoires X et Y , et tout réel a ,

$$E[aX] = aE[X] \quad \text{et} \quad E[X + Y] = E[X] + E[Y].$$

(iv) L'espérance est croissante : si $X \leq Y$ p.s., alors $E[X] \leq E[Y]$.

Proposition

Soit X une variable aléatoire, et $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ une fonction.

– Si X est discrète, alors

$$E[\varphi(X)] = \sum_{x \in X(\Omega)} \varphi(x)P(X = x).$$

– Si X est continue, alors

$$E[\varphi(X)] = \int_{\mathbb{R}} \varphi(x)f_X(x)dx.$$

(À condition que la série et l'intégrale soient bien définies)

2.5 Variance d'une variable aléatoire

Définition

Soit X une variable aléatoire. La **variance** de X est l'espérance des carrés des écarts de X à sa moyenne :

$$\text{Var}(X) = E\left[(X - E[X])^2\right] \geq 0.$$

L'**écart type** de X est $\sigma(X) = \sqrt{\text{Var}(X)}$.

Attention. La variance n'est pas toujours définie. Il faut que l'espérance $E[X]$ soit définie et que l'espérance ci-dessus converge. Ceci revient à demander à ce que $E[X^2]$ converge.

NB. À la différence de la variance, l'écart type $\sigma(X)$ est *homogène* à X : si par exemple X est une distance, alors $\sigma(X)$ est une distance aussi. Ceci justifie l'intérêt de l'écart type.

Propriétés

Pour toutes variables aléatoires X et Y et toute constante a ,

1. $\text{Var}(X) = E[X^2] - E[X]^2$
2. $\text{Var}(aX) = a^2 \text{Var}(X)$
3. $\text{Var}(X + a) = \text{Var}(X)$
4. $\text{Var}(X + Y) = \text{Var}(X) + 2 \text{Cov}(X, Y) + \text{Var}(Y)$, où la **covariance** est définie par

$$\text{Cov}(X, Y) = E\left[(X - E[X])(Y - E[Y])\right] = E[XY] - E[X]E[Y].$$

Proposition

Pour toute variable aléatoire X possédant une variance, la variable aléatoire $Y = \frac{X - E[X]}{\sigma(X)}$ est centrée ($E[Y] = 0$) et réduite ($\text{Var}(Y) = 1$).

Plus généralement, pour $r > 0$, on définit (s'il existe) le **moment d'ordre r** :

$$m_r(X) = E[X^r],$$

et le **moment centré d'ordre r** :

$$\mu_r(X) = E\left[(X - E[X])^r\right].$$

Proposition (*Inégalité de Markov*)

Soit X une variable aléatoire. Pour tout $a > 0$,

$$P(|X| \geq a) \leq \frac{E[|X|]}{a}.$$

Plus généralement, pour tout $a > 0$ et $r > 0$,

$$P(|X| \geq a) \leq \frac{E[|X|^r]}{a^r}.$$

Démonstration : On a toujours $|X|^r \geq 0$ et, si $|X| \geq a$, alors $|X|^r \geq a^r$. D'où $a^r \mathbf{1}_{\{|X| \geq a\}} \leq |X|^r$, ce qui donne, en prenant l'espérance de chaque membre :

$$E[a^r \mathbf{1}_{\{|X| \geq a\}}] \leq E[|X|^r].$$

Et $E[a^r \mathbf{1}_{\{|X| \geq a\}}] = a^r P(|X| \geq a)$, d'où résulte l'inégalité annoncée ($r = 1$ donne la première). ■

Proposition (*Inégalité de Bienaymé-Tchebychev*)

Soit X une variable aléatoire. Pour tout $a > 0$,

$$P\left(|X - E[X]| \geq a\right) \leq \frac{\text{Var}(X)}{a^2}.$$

Démonstration : Prendre $r = 2$ et remplacer X par $X - E[X]$ dans l'inégalité de Markov. ■

2.6 Indépendance de variables aléatoires

Définition

Des variables aléatoires X_1, \dots, X_n sont **indépendantes** si, pour tous $B_1, \dots, B_n \subset \mathbb{R}$,

$$P(X_1 \in B_1, \dots, X_n \in B_n) = P(X_1 \in B_1) \cdots P(X_n \in B_n).$$

(où les virgules se lisent « et » : $P(X_1 \in B_1, \dots, X_n \in B_n) = P(\{X_1 \in B_1\} \cap \dots \cap \{X_n \in B_n\})$)

Par exemple, deux variables aléatoires X et Y sont indépendantes si les événements qui ne dépendent que de X sont indépendants des événements qui ne dépendent que de Y .

Proposition

1. Si X_1, \dots, X_n sont indépendantes, alors les variables aléatoires $f_1(X_1), \dots, f_n(X_n)$ sont indépendantes, quelles que soient les fonctions f_1, \dots, f_n .
2. « **Indépendance par paquets** ». Si X_1, \dots, X_n sont indépendantes alors, par exemple, les variables aléatoires $f_{1,2}(X_1, X_2)$, $f_4(X_4)$, $f_{3,5,6}(X_3, X_5, X_6), \dots$ sont indépendantes : les fonctions de « paquets disjoints » de variables sont indépendantes.
3. Si des événements A_1, \dots, A_n sont indépendants alors leurs fonctions indicatrices $\mathbf{1}_{A_1}, \dots, \mathbf{1}_{A_n}$ sont des variables aléatoires indépendantes ; et réciproquement.

Proposition

Si X_1, \dots, X_n sont des variables aléatoires indépendantes, alors

1. si leurs espérances sont bien définies,

$$E[X_1 \cdots X_n] = E[X_1] \cdots E[X_n]$$

2. si leurs variances sont bien définies, alors on a $\text{Cov}(X_i, X_j) = 0$ pour tous $i \neq j$, d'où

$$\text{Var}(X_1 + \cdots + X_n) = \text{Var}(X_1) + \cdots + \text{Var}(X_n).$$

NB. Les réciproques sont fausses !

Par le 1. de la proposition précédente on déduit, si X_1, \dots, X_n sont indépendantes,

$$E[f_1(X_1) \cdots f_n(X_n)] = E[f_1(X_1)] \cdots E[f_n(X_n)].$$

On rappelle que E est toujours linéaire : même si X_1, \dots, X_n ne sont pas indépendantes,

$$E[X_1 + \cdots + X_n] = E[X_1] + \cdots + E[X_n].$$

2.7 Loi faible des grands nombres

Théorème

Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires indépendantes, et de même loi, d'espérance m et de variance σ^2 . On définit la variable aléatoires \bar{X}_n , appelée **moyenne empirique**, par

$$\bar{X}_n = \frac{X_1 + \cdots + X_n}{n}.$$

On a :

$$\text{pour tout } \varepsilon > 0, \quad P\left(|\bar{X}_n - m| < \varepsilon\right) \xrightarrow[n \rightarrow \infty]{} 1.$$

Démonstration : Par linéarité de l'espérance, $E[\bar{X}_n] = \frac{1}{n}(E[X_1] + \dots + E[X_n]) = m$, et les variables sont indépendantes donc

$$\text{Var}(\bar{X}_n) = \frac{1}{n^2} \sum_{i=1}^n \text{Var}(X_i) = \frac{\sigma^2}{n}.$$

L'inégalité de Bienaymé-Tchebychev pour \bar{X}_n s'écrit donc, pour tout $\varepsilon > 0$,

$$P(|\bar{X}_n - m| \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2},$$

ce qui donne, en passant au complémentaire,

$$P(|\bar{X}_n - m| < \varepsilon) \geq 1 - \frac{\sigma^2}{n\varepsilon^2}.$$

Comme le terme de droite converge vers 1 quand $n \rightarrow \infty$, et que le terme de gauche est ≤ 1 , on obtient l'énoncé. ■

NB. Si $(A_n)_{n \geq 1}$ est une suite d'événements *indépendants* et qui ont *même probabilité* p (par exemple, dans une suite de tirages à Pile-ou-Face, $A_n = \{\text{le } n\text{-ième tirage est Pile}\}$, et $p = \frac{1}{2}$), alors en posant $X_i = \mathbf{1}_{A_i}$, on a

$$\bar{X}_n = \frac{\mathbf{1}_{A_1} + \dots + \mathbf{1}_{A_n}}{n} = \frac{\text{nombre d'événements réalisés parmi } A_1, \dots, A_n}{n}$$

donc \bar{X}_n est la **fréquence de réalisation** des événements A_1, \dots, A_n , et la loi des grands nombres montre que, si n est grand, cette fréquence a de grandes chances d'être proche de $E[X_1] = p$, qui est la probabilité commune des événements A_n . Ainsi, la fréquence d'apparition de Pile dans une suite de tirages à Pile-ou-Face indépendants converge vers $\frac{1}{2}$. Ou, si la pièce est biaisée, vers la probabilité d'obtenir Pile à un tirage.

3 Lois usuelles

3.1 Loi de Bernoulli de paramètre p , $\mathcal{B}(p)$

C'est la loi d'une variable aléatoire X qui ne peut prendre que 2 valeurs, notées 1 et 0, et $p \in [0,1]$ est la probabilité de la valeur 1 :

$$P(X = 1) = p \quad \text{et} \quad P(X = 0) = 1 - p.$$

C'est donc la loi de la fonction indicatrice $\mathbf{1}_A$ d'un événement A tel que $P(A) = p$.

On a

$$E[X] = p \quad \text{et} \quad \text{Var}(X) = p(1 - p).$$

3.2 Loi binomiale de paramètres n et p , $\mathcal{B}(n,p)$

Soit n variables aléatoires X_1, X_2, \dots, X_n , indépendantes et de même loi $\mathcal{B}(p)$.

La loi binomiale $\mathcal{B}(n,p)$ est la loi de la variable aléatoire $S_n = X_1 + X_2 + \dots + X_n$.

C'est donc **la loi du nombre d'événements** parmi A_1, \dots, A_n qui sont réalisés, si A_1, \dots, A_n sont indépendants et de même probabilité p . (Ci-dessus, $X_n = \mathbf{1}_{A_n}$)

S_n est à valeurs dans $\{0, 1, \dots, n\}$ et on a (cf. chapitre 1)

$$\text{pour } k = 0, 1, \dots, n, \quad P(S_n = k) = \binom{n}{k} p^k (1 - p)^{n-k}.$$

De plus

$$E[X] = \sum_{i=1}^n E[X_i] = np$$

et comme les variables aléatoires sont indépendantes,

$$\text{Var}(X) = \sum_{i=1}^n \text{Var}(X_i) = np(1 - p).$$

3.3 Loi de Poisson de paramètre λ , $\mathcal{P}(\lambda)$

Soit $\lambda > 0$. Une variable aléatoire X suit la loi de Poisson de paramètre λ si

$$X(\Omega) = \mathbb{N}$$

et

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

On a

$$E[X] = \lambda \quad \text{et} \quad \text{Var}(X) = \lambda.$$

C'est la « loi des petites probabilités » car la loi limite de la loi binomiale $\mathcal{B}(n,p)$, avec $np \sim \lambda$:

Proposition

Si, pour tout n , S_n suit la loi $\mathcal{B}(n,p_n)$, et $np_n \xrightarrow[n \rightarrow \infty]{} \lambda$, alors

$$\text{pour tout } k \in \mathbb{N}, \quad P(S_n = k) \xrightarrow[n \rightarrow \infty]{} e^{-\lambda} \frac{\lambda^k}{k!}.$$

Dans la pratique, on peut approcher la loi binomiale par une loi de Poisson lorsque

$$n \geq 50, \quad p \leq 0,1 \quad \text{et} \quad \lambda = np \leq 15.$$

3.4 Loi géométrique de paramètre p , $\mathcal{G}(p)$

Soit $p \in]0,1[$. Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires indépendantes et de même loi $\mathcal{B}(p)$. La loi géométrique de paramètre p est la loi de la variable aléatoire

$$N = \min\{n \geq 1 | X_n = 1\}.$$

C'est donc **la loi du premier instant** n tel que A_n est réalisé, si $(A_n)_{n \geq 1}$ est une suite d'événements indépendants et de même probabilité p . (Ci-dessus, $X_n = \mathbf{1}_{A_n}$)

N est à valeurs dans \mathbb{N}^* et on a

$$\text{pour } k \in \mathbb{N}^*, \quad P(N = k) = P(X_1 = 0, X_2 = 0, \dots, X_{k-1} = 0, X_k = 1) = (1-p)^{k-1}p.$$

On a

$$E[N] = \frac{1}{p} \quad \text{et} \quad \text{Var}(N) = \frac{1-p}{p^2}.$$

3.5 Loi uniforme sur $[a,b]$, $\mathcal{U}([a,b])$

Soit $a < b$. La loi uniforme sur $[a,b]$ est la loi de densité

$$f(x) = \frac{1}{b-a} \mathbf{1}_{[a,b]}(x) = \begin{cases} (b-a)^{-1} & \text{si } a \leq x \leq b \\ 0 & \text{si } x \notin [a,b]. \end{cases}$$

Une variable aléatoire X de loi $\mathcal{U}([a,b])$ est donc à valeurs dans $[a,b]$ et sa fonction de répartition est donnée par :

$$\begin{aligned} \text{pour tout } x \in [a,b], \quad F_X(x) &= P(X \leq x) = \frac{1}{b-a} \int_a^x dt = \frac{x-a}{b-a}, \\ \text{pour tout } x < a, \quad F_X(x) &= 0 \quad \text{et, pour tout } x > b, \quad F_X(x) = 1. \end{aligned}$$

On a

$$E[X] = \frac{a+b}{2} \quad \text{et} \quad \text{Var}(X) = \frac{1}{12}(b-a)^2.$$

3.6 Loi exponentielle de paramètre λ , $\mathcal{E}(\lambda)$

Soit $\lambda > 0$. La loi exponentielle de paramètre λ est la loi de densité

$$f(x) = \lambda e^{-\lambda x} \mathbf{1}_{\mathbb{R}_+}(x).$$

Une variable aléatoire X de loi $\mathcal{E}(\lambda)$ est donc à valeurs dans \mathbb{R}_+ et sa fonction de répartition est donnée par :

$$\begin{aligned} \text{pour tout } x \geq 0, \quad F_X(x) &= P(X \leq x) = \int_0^x \lambda e^{-\lambda t} dt = 1 - e^{-\lambda x} \\ \text{et, pour tout } x < 0, \quad F_X(x) &= 0. \end{aligned}$$

On a

$$\begin{aligned} \text{pour tout } x \geq 0, \quad P(X \geq x) &= e^{-\lambda x}, \\ E[X] &= \frac{1}{\lambda} \quad \text{et} \quad \text{Var}(X) = \frac{1}{\lambda^2}. \end{aligned}$$

La loi exponentielle est une loi « sans mémoire ». En effet, pour tous $s, t \geq 0$,

$$P(X \geq s+t | X > s) = \frac{P(\{X \geq s+t\} \cap \{X \geq s\})}{P(X \geq s)} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} = e^{-\lambda t} = P(X \geq t),$$

en utilisant le fait que l'événement $\{X \geq s+t\}$ est inclus dans l'événement $\{X \geq s\}$.

3.7 La loi normale de moyenne m et variance σ^2 , $\mathcal{N}(m, \sigma^2)$

La loi normale centrée ($m = 0$) réduite ($\sigma = 1$), notée $\mathcal{N}(0,1)$, est la loi de densité

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

La courbe représentative de f est la « courbe en cloche ». Si X suit la loi $\mathcal{N}(0,1)$, on a

$$P(-2 \leq X \leq 2) \simeq 0,954 \quad \text{et} \quad P(-3 \leq X \leq 3) \simeq 0,997.$$

De la loi normale centrée réduite se déduisent les autres lois normales :

Si X suit la loi $\mathcal{N}(0,1)$, alors $m + \sigma X$ suit la loi $\mathcal{N}(m, \sigma^2)$. La densité de la loi $\mathcal{N}(m, \sigma^2)$ est

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}.$$

Lorsque X suit une loi normale, on dit que X est une variable aléatoire **gaussienne**.

Proposition

Toute combinaison linéaire de variables aléatoires gaussiennes indépendantes est une variable aléatoire gaussienne.

Plus précisément, si X_1, \dots, X_n sont indépendantes et $X_i \sim \mathcal{N}(m_i, \sigma_i^2)$ alors, pour tous $a_1, \dots, a_n \in \mathbb{R}$, la variable aléatoire $X = a_1 X_1 + \dots + a_n X_n$ suit la loi $\mathcal{N}(M, \Sigma^2)$, où

$$M = E[X] = \sum_{i=1}^n a_i m_i \quad \text{et} \quad \Sigma^2 = \text{Var}(X) = \sum_{i=1}^n a_i^2 \sigma_i^2.$$

Les lois normales interviennent dans le théorème central limite.

4 Couples de variables aléatoires

4.1 Loi du couple, loi marginale

Définition

Soit X, Y deux variables aléatoires. La **loi du couple** (X, Y) est la probabilité $P_{(X, Y)}$ sur \mathbb{R}^2 qui vérifie :

$$\text{pour tous } A, B \subset \mathbb{R}, \quad P_{(X, Y)}(A \times B) = P(X \in A, Y \in B).$$

Inversement, les lois de X et de Y sont appelées les **lois marginales** de $P_{(X, Y)}$.

(la virgule se lit « et » : $P(X \in A, Y \in B) = P(\{X \in A\} \cap \{Y \in B\})$)

Si X et Y sont indépendantes, la loi du couple est fournie par les lois de X et de Y :

$$P_{(X, Y)}(A \times B) = P_X(A)P_Y(B).$$

La loi du couple contient davantage d'information que P_X et P_Y : elle indique aussi la façon dont les variables dépendent l'une de l'autre (connaître X peut renseigner sur Y).

Cas de deux variables discrètes. Si X et Y sont discrètes alors la loi de (X, Y) est donnée par les probabilités élémentaires :

$$p_{(X, Y)}(x, y) = P(X = x, Y = y) \quad \text{pour tous } x \in X(\Omega), y \in Y(\Omega).$$

Elles vérifient $p_{(X, Y)}(x, y) \in [0, 1]$ pour tous x, y , et

$$\sum_{x \in X(\Omega)} \sum_{y \in Y(\Omega)} p_{(X, Y)}(x, y) = 1.$$

Inversement, les lois marginales se déduisent des $(p_{(X, Y)}(x, y))$:

$$\text{pour tout } x \in X(\Omega), \quad p_X(x) = P(X = x) = \sum_{y \in Y(\Omega)} P(X = x, Y = y) = \sum_{y \in Y(\Omega)} p_{(X, Y)}(x, y),$$

$$\text{pour tout } y \in Y(\Omega), \quad p_Y(y) = P(Y = y) = \sum_{x \in X(\Omega)} P(X = x, Y = y) = \sum_{x \in X(\Omega)} p_{(X, Y)}(x, y).$$

NB. X et Y sont indépendantes si, et seulement si $p_{(X, Y)}(x, y) = p_X(x)p_Y(y)$ pour tous x, y .

Cas où $P_{(X, Y)}$ a une densité. On dit que le couple (X, Y) a une densité s'il y a une fonction $f_{(X, Y)} : \mathbb{R}^2 \rightarrow \mathbb{R}$ telle que

$$\text{pour tout } D \subset \mathbb{R}^2, \quad P_{(X, Y)}(D) = \iint_D f_{(X, Y)}(x, y) dx dy.$$

$f_{(X, Y)}$ est appelée la **densité** du couple (X, Y) . Alors $f_{(X, Y)}(x, y) \geq 0$ pour tous $x, y \in \mathbb{R}$, et

$$\int_{\mathbb{R}} \int_{\mathbb{R}} f_{(X, Y)}(x, y) dx dy = 1.$$

En particulier,

$$\text{pour tous } A, B \subset \mathbb{R}, \quad P(X \in A, Y \in B) = \int_A \left(\int_B f_{(X, Y)}(x, y) dy \right) dx$$

On déduit les lois marginales de la loi du couple et, *dans le cas indépendant*, on déduit la loi du couple des lois marginales :

Proposition

1. Si (X,Y) a pour densité $f_{(X,Y)}$, alors X et Y ont des densités f_X et f_Y données par

$$f_X(x) = \int_{\mathbb{R}} f_{(X,Y)}(x,y)dy \quad \text{et} \quad f_Y(y) = \int_{\mathbb{R}} f_{(X,Y)}(x,y)dx.$$

2. Si X et Y ont des densités f_X et f_Y et sont indépendantes, alors (X,Y) a pour densité

$$f_{(X,Y)}(x,y) = f_X(x)f_Y(y).$$

Réciproquement, si $f_{(X,Y)}(x,y) = f(x)g(y)$ pour deux fonctions f et g , alors X et Y sont indépendantes, et les densités de X et Y sont proportionnelles à f et g .

4.2 Lois conditionnelles

Si X,Y sont des variables aléatoires, connaître la valeur de X peut influencer sur la loi de Y .

Cas où X et Y sont discrètes. Étant donné $x \in X(\Omega)$, la **loi conditionnelle de Y sachant $X = x$** est la probabilité sur $Y(\Omega)$ donnée par les probabilités élémentaires suivantes :

$$\text{pour tout } y \in Y(\Omega), \quad p_{Y|X=x}(y) = P(Y = y|X = x) = \frac{P(X = x, Y = y)}{P(X = x)} = \frac{p_{(X,Y)}(x,y)}{p_X(x)}.$$

Cas où $P_{(X,Y)}$ a une densité. Étant donné $x \in \mathbb{R}$ tel que $f_X(x) > 0$, la **loi conditionnelle de Y sachant $X = x$** est la probabilité sur \mathbb{R} donnée similairement par la densité suivante :

$$f_{Y|X=x}(y) = \frac{f_{(X,Y)}(x,y)}{f_X(x)}.$$

4.3 Loi de $(U,V) = \varphi(X,Y)$

Soit un couple (X,Y) qui admet une densité $f_{(X,Y)}$. On suppose que (X,Y) est à valeurs dans $A \subset \mathbb{R}^2$ (ceci revient à dire que $f_{(X,Y)}(x,y) = 0$ si (x,y) est hors de A).

Soit φ un \mathcal{C}^1 -difféomorphisme de A sur $B \subset \mathbb{R}^2$, c'est-à-dire que φ est différentiable sur A , que φ est une bijection de A sur B , et que φ^{-1} est différentiable sur B .

On considère le couple de variables aléatoires

$$(U,V) = \varphi(X,Y).$$

La loi du (U,V) se calcule par changement de variable : en posant $(u,v) = \varphi(x,y)$, on aura $(x,y) = \varphi^{-1}(u,v)$ et donc

$$dx dy = |\text{Jac}_{\varphi^{-1}}(u,v)| du dv$$

où Jac est le déterminant jacobien : si $\psi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ s'écrit $\psi(u,v) = (\psi_1(u,v), \psi_2(u,v))$,

$$\text{Jac}_{\psi}(u,v) = \begin{vmatrix} \frac{\partial \psi_1}{\partial u} & \frac{\partial \psi_1}{\partial v} \\ \frac{\partial \psi_2}{\partial u} & \frac{\partial \psi_2}{\partial v} \end{vmatrix}.$$

Ainsi, si $C \subset B$,

$$\begin{aligned} P((U,V) \in C) &= P((X,Y) \in \varphi^{-1}(C)) = \iint_{\varphi^{-1}(C)} f_{(X,Y)}(x,y) dx dy \\ &= \iint_C f_{(X,Y)}(\varphi^{-1}(u,v)) |\text{Jac}_{\varphi^{-1}}(u,v)| du dv. \end{aligned}$$

Ceci montre que le couple (U,V) a pour densité

$$f_{(U,V)}(u,v) = f_{(X,Y)}(\varphi^{-1}(u,v)) |\text{Jac}_{\varphi^{-1}}(u,v)|.$$

Il est parfois plus simple, pour effectuer le changement de variable, de calculer plutôt

$$du dv = |\text{Jac}_{\varphi}(x,y)| dx dy$$

de sorte à obtenir que la densité de (U,V) est

$$f_{(U,V)}(u,v) = f_{(X,Y)}(x,y) |\text{Jac}_{\varphi}(x,y)|^{-1}$$

où il reste à remplacer (x,y) en fonction de (u,v) (on a $(x,y) = \varphi^{-1}(u,v)$).

4.4 Conséquence : loi de $Z = h(X,Y)$

Soit un couple (X,Y) qui admet une densité $f_{(X,Y)}$. Soit h une fonction $\mathbb{R}^2 \rightarrow \mathbb{R}$. On suppose que, pour tout $x \in \mathbb{R}$, $y \mapsto h(x,y)$ est strictement croissante, ou strictement décroissante (ou que $\frac{\partial h}{\partial y} \neq 0$).

On considère la variable aléatoire

$$Z = h(X,Y).$$

Pour obtenir la loi de Z , on calcule la densité de $(X,Z) = \varphi(X,Y)$ par la formule précédente, où $\varphi(x,y) = (x, h(x,y))$,

$$f_{(X,Z)}(x,z) = f_{(X,Y)}(x,y) \left| \begin{array}{cc} 1 & 0 \\ \frac{\partial h}{\partial x} & \frac{\partial h}{\partial y} \end{array} \right|^{-1} = \frac{f_{(X,Y)}(x,y)}{\left| \frac{\partial h}{\partial y} \right|},$$

on remplace y par son expression en fonction de z (il faut $h(x,y) = z$), et on en déduit la loi de la marginale Z par la formule usuelle :

$$f_Z(z) = \int_{\mathbb{R}} f_{(X,Z)}(x,z) dx.$$

Exemple : somme de variables aléatoires indépendantes. On suppose X et Y indépendantes, de densités f_X et f_Y et on considère $Z = X + Y = h(x,y)$ où $h(x,y) = x + y$.

On a vu que (X,Y) a pour densité $f_X(x)f_Y(y)$, et on calcule $f_{(X,Z)}(x,z) = f_{(X,Y)}(x,y) = f_{(X,Y)}(x,z-x) = f_X(x)f_Y(z-x)$, d'où

$$f_{X+Y}(z) = \int_{\mathbb{R}} f_X(x)f_Y(z-x) dx = (f_X * f_Y)(z).$$

f_{X+Y} est le **produit de convolution** de f_X et f_Y .

5 Estimation

5.1 Principe, statistiques classiques

Définition

Soit X une variable aléatoire. Un **échantillon de taille n de X** est une famille X_1, \dots, X_n de n variables aléatoires indépendantes et de même loi que X .

On souhaite étudier la loi de X . Par exemple, X est la taille en centimètres d'un individu choisi uniformément dans la population adulte française. Son espérance est donc la taille moyenne d'un Français adulte, que l'on peut vouloir estimer.

On ne dispose pour cela que d'une réalisation d'un **échantillon de taille n** : une réalisation $(x_1, \dots, x_n) \in \mathbb{R}^n$ de n variables aléatoires indépendantes (X_1, \dots, X_n) qui ont la même loi que X . À défaut de pouvoir mesurer toute la population, ce qui serait long, coûteux et compliqué, on se contente de mesurer la taille de n personnes choisies au hasard parmi les Français adultes.

L'objectif de l'estimation statistique consiste donc à déduire certaines propriétés de la loi de X (son espérance, sa variance, ses paramètres...) à partir d'un échantillon de valeurs X_1, \dots, X_n .

Statistiques simples. Les quantités les plus classiques pour décrire un échantillon sont

– la **moyenne empirique** :

$$\bar{X}_n = \frac{X_1 + \dots + X_n}{n} = \frac{1}{n} \sum_{i=1}^n X_i$$

– la **variance empirique** :

$$S_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2$$

– la **variance empirique modifiée** :

$$\Sigma_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2.$$

Proposition

Si X a pour espérance m et pour écart type σ , alors

$$E[\bar{X}_n] = m, \quad \text{Var}(\bar{X}_n) = \frac{\sigma^2}{n},$$

$$E[S_n^2] = \frac{n-1}{n} \sigma^2, \quad E[\Sigma_n^2] = \sigma^2.$$

5.2 Estimateurs

On suppose que X suit une loi P_θ qui dépend d'un paramètre $\theta \in \Theta$, où $\Theta \subset \mathbb{R}$ est l'ensemble des valeurs a priori possibles du paramètre. On ignore la valeur de θ , et on souhaite l'estimer. NB. Dans ce qui suit, on pourrait noter E_θ au lieu de E pour indiquer que la loi de X et donc de X_1, \dots, X_n dépend de θ ; on ne le fait pas, afin d'alléger la notation, mais il faut le garder à l'esprit.

Définition

Un **estimateur** de θ est une variable aléatoire $T_n = f(X_1, \dots, X_n)$ qui dépend d'un échantillon X_1, \dots, X_n de X . On utilise souvent la notation $\hat{\theta}$ pour un estimateur de θ .

Une **estimation** de θ est la valeur réelle $t_n = f(x_1, \dots, x_n)$ prise par une réalisation particulière de l'échantillon.

NB. La définition d'estimateur peut paraître curieuse à plusieurs titres. On note que n n'apparaît pas dans l'appellation « estimateur de θ » mais apparaît dans la définition. En fait, un estimateur peut être vu comme une suite $(T_n)_n$ de variables aléatoires où T_n dépend de X_1, \dots, X_n , et on utilisera la variable T_n adaptée à la taille de l'échantillon dont on dispose. De plus, θ n'apparaît pas dans la définition : n'importe quelle fonction de X_1, \dots, X_n est donc un estimateur de θ . En revanche θ intervient pour mesurer la qualité de l'estimateur :

Définition

Soit T_n un estimateur de θ .

Le **biais** de T_n est la différence $E[T_n] - \theta$.

On dit que T_n est **sans biais** si $E[T_n] = \theta$, quel que soit $\theta \in \Theta$.

On dit que T_n est **asymptotiquement sans biais** si $E[T_n] \xrightarrow[n \rightarrow \infty]{} \theta$, quel que soit $\theta \in \Theta$.

On dit que T_n est **convergent** si, quel que soit $\theta \in \Theta$,

$$\text{pour tout } \alpha > 0, \quad P(|T_n - \theta| > \alpha) \xrightarrow[n \rightarrow \infty]{} 0.$$

Proposition

Tout estimateur asymptotiquement sans biais dont la variance tend vers 0 est convergent.

Définition

Le **risque quadratique** d'un estimateur T_n de θ est

$$R_{T_n}(\theta) = E[(T_n - \theta)^2].$$

On dit que l'estimateur S_n est **meilleur** que T_n si, quel que soit θ ,

$$R_{S_n}(\theta) \leq R_{T_n}(\theta).$$

Par l'inégalité de Markov, un estimateur dont le risque quadratique tend vers 0 (quel que soit θ) est convergent.

NB. Si T_n est sans biais, alors $R_{T_n}(\theta) = \text{Var}(T_n)$.

5.3 Construction d'estimateurs

Méthode des moments Le principe est d'utiliser la loi des grands nombres pour estimer les moments, et d'utiliser ensuite ces estimateurs des moments pour estimer θ .

Par la loi des grands nombres, on a :

Proposition

Soit X une variable aléatoire d'espérance m et de variance σ^2 .

1. La moyenne empirique est un estimateur sans biais et convergent de m .
2. La variance empirique est un estimateur asymptotiquement sans biais et convergent de σ^2 , et la variance empirique modifiée est un estimateur sans biais et convergent de σ^2 .
3. Pour tout $r > 0$, le moment empirique d'ordre r ,

$$\widehat{m}_r = \frac{1}{n} \sum_{i=1}^n (X_i)^r$$

est un estimateur sans biais et convergent de $m_r = E[X^r]$. (Si m_r est bien défini)

On en déduit la méthode des moments : exprimer (si possible) θ à l'aide des moments $(m_r)_{r>0}$, puis remplacer dans cette expression les moments par les moments empiriques. Ceci fournit un estimateur convergent de θ . L'expression peut aussi faire intervenir σ^2 , que l'on remplace par S_n^2 ou Σ_n^2 .

En pratique, on calcule $E[X]$, $E[X^2]$, etc., jusqu'à obtenir une expression faisant intervenir θ (souvent, $E[X]$ suffit), et on inverse pour obtenir θ en fonction de $E[X]$, $E[X^2]$, etc. Il ne reste plus qu'à remplacer $m_1 = E[X]$ par \bar{X}_n , $m_2 = E[X^2]$ par \widehat{m}_2 , etc.

Méthode du maximum de vraisemblance Le principe est d'estimer θ par la valeur qui maximise la densité du vecteur (X_1, \dots, X_n) .

Définition

La **vraisemblance** de l'échantillon (X_1, \dots, X_n) est la fonction L définie par :

– si X est discrète, de probabilité élémentaire P_θ , pour tous x_1, \dots, x_n ,

$$L(x_1, \dots, x_n; \theta) = \prod_{i=1}^n P_\theta(x_i)$$

– si X est continue, de densité f_θ , pour tous x_1, \dots, x_n ,

$$L(x_1, \dots, x_n; \theta) = \prod_{i=1}^n f_\theta(x_i).$$

Un **estimateur du maximum de vraisemblance** (ou EMV) pour θ est un estimateur $h(X_1, \dots, X_n)$ tel que, pour toutes les valeurs $x_1, \dots, x_n \in X(\Omega)$,

$$L(x_1, x_2, \dots, x_n; h(x_1, \dots, x_n)) = \max_{\theta \in \Theta} L(x_1, \dots, x_n; \theta).$$

NB. Pour tous x_1, \dots, x_n , $h(x_1, \dots, x_n)$ est la, ou l'une des valeurs de θ où $L(x_1, \dots, x_n; \theta)$ est maximum. Ceci définit l'estimation (pratique). L'*estimateur* est la variable aléatoire $h(X_1, \dots, X_n)$. Sous des hypothèses assez générales, on montre que ceci définit un bon estimateur convergent.

Pour le calcul, on est amené à maximiser $L(x_1, \dots, x_n; \theta)$ selon θ . Vu que le logarithme est strictement croissant, c'est équivalent à maximiser la **log-vraisemblance** $\ln(L(x_1, \dots, x_n; \theta))$, souvent plus pratique. La dérivée étant nulle au maximum, ceci mène à chercher θ tel que

$$\frac{\partial}{\partial \theta} \ln L(x_1, \dots, x_n; \theta) = 0.$$

C'est l'équation de la vraisemblance.

5.4 Intervalles de confiance

Définition

Un **intervalle de confiance de niveau** $1 - \alpha$ est un intervalle $IC_{1-\alpha}$, qui dépend de X_1, \dots, X_n , contenant la valeur θ avec probabilité $\geq 1 - \alpha$.

Intervalle de confiance pour la moyenne m . Soit X une variable aléatoire de moyenne m et de variance σ^2 .

Soit (X_1, X_2, \dots, X_n) un échantillon de X . On cherche un intervalle de confiance à partir des estimateurs sans biais \bar{X}_n et Σ_n^2 .

Théorème (*Théorème central limite*)

Soit $(X_n)_n$ une suite de variables aléatoires indépendantes, de même loi de moyenne m et d'écart type σ . Soit Z_n la variable aléatoire définie par

$$Z_n = \frac{\sqrt{n}(\bar{X}_n - m)}{\sigma}.$$

Lorsque $n \rightarrow +\infty$, Z_n converge en loi vers une variable $Z \sim \mathcal{N}(0,1)$. On a donc

$$\text{pour tout intervalle } I \subset \mathbb{R}, \quad P(Z_n \in I) \xrightarrow{n \rightarrow \infty} P(Z \in I) = \frac{1}{\sqrt{2\pi}} \int_I e^{-\frac{x^2}{2}} dx.$$

Dans la pratique on applique ce résultat dès que n est suffisamment grand ($n \geq 30$).

Soit $Z \sim \mathcal{N}(0,1)$ et a et α définis par

$$P(-a \leq Z \leq a) = 1 - \alpha$$

Par exemple, on sait que (voir la table)

$$P(|Z| \leq 1,96) = 95\% \quad \text{et} \quad P(|Z| \leq 2,576) = 99\%.$$

Si $n \geq 30$, le théorème de la limite centrale permet d'écrire

$$P(-a \leq \frac{\sqrt{n}(\bar{X}_n - m)}{\sigma} \leq a) \simeq 1 - \alpha.$$

On obtient, vu que Σ_n^2 est un estimateur convergent de σ^2 ,

$$P(-a \leq \frac{\sqrt{n}(\bar{X}_n - m)}{\Sigma_n} \leq a) \simeq 1 - \alpha,$$

ce qui se réécrit

$$P\left(\bar{X}_n - a \frac{\Sigma_n}{\sqrt{n}} \leq m \leq \bar{X}_n + a \frac{\Sigma_n}{\sqrt{n}}\right) \simeq 1 - \alpha.$$

Si \bar{x}_n est la moyenne observée et σ_n^2 la variance corrigée observée, en posant $\pi_1 = \bar{x}_n - a \frac{\sigma_n}{\sqrt{n}}$ et $\pi_2 = \bar{x}_n + a \frac{\sigma_n}{\sqrt{n}}$, on en déduit que $[\pi_1, \pi_2]$ est un intervalle de confiance de m de niveau $1 - \alpha$.

Intervalle de confiance pour une proportion p . Ici X est une variable aléatoire de Bernoulli de paramètre p , d'où $E[X] = p$ et $\text{Var}(X) = p(1 - p)$.

On a, pour n grand,

$$\frac{\sqrt{n}(\bar{X}_n - p)}{\sqrt{p(1 - p)}} \sim \mathcal{N}(0,1)$$

et l'intervalle de confiance pour p de niveau $1 - \alpha$ précédent est donc

$$I = \left[\bar{x}_n - a \frac{\sqrt{(\bar{x}_n)(1 - \bar{x}_n)}}{\sqrt{n}}, \bar{x}_n + a \frac{\sqrt{(\bar{x}_n)(1 - \bar{x}_n)}}{\sqrt{n}} \right].$$

Comme, pour tout p tel que $0 < p < 1$, on a $p(1 - p) \leq \frac{1}{4}$, on déduit

$$I \subset J = \left[\bar{x}_n - \frac{a}{2\sqrt{n}}, \bar{x}_n + \frac{a}{2\sqrt{n}} \right].$$

On utilise généralement plutôt J (plus simple) comme intervalle de confiance de niveau $1 - \alpha$.

Remarques.

1. Attention : pour que ces approximations soient justifiées, les valeurs de $n\pi_1$, $n\pi_2$, $n(1 - \pi_1)$ et $n(1 - \pi_2)$ doivent être toutes les quatre supérieures ou égales à 5.
2. Les intervalles de confiance donnés ci-dessus permettent aussi de déterminer la grandeur n de l'échantillon nécessaire pour avoir une précision donnée pour l'estimation d'une proportion.

5.5 Application aux tests de différence

On souhaite, à partir de l'échantillon observé x_1, \dots, x_n , savoir si l'on peut raisonnablement conclure qu'une certaine hypothèse sur la loi de X est fautive (en vue de prendre une décision). L'hypothèse est appelée **hypothèse nulle**, et notée \mathcal{H}_0 . C'est une hypothèse que l'on veut avoir « peu de chance » de rejeter si elle est vraie (erreur de première espèce).

Le **seuil de risque** du test est

$$\alpha = P(\text{rejet de } \mathcal{H}_0) \text{ quand } \mathcal{H}_0 \text{ est vraie.}$$

α doit être petit, en général on souhaite $\alpha = 5\%$.

Exemple 1 : Test sur une proportion

Soit p la proportion observée d'un caractère C sur un échantillon de taille n . On cherche à savoir si cette proportion paraît conforme à une fréquence théorique annoncée, π .

Hypothèse \mathcal{H}_0 : la fréquence théorique est π .

Soit X une variable de Bernoulli de paramètre π . Si n est suffisamment grand,

$$Z_n = \frac{\sqrt{n}(\bar{X}_n - \pi)}{\sqrt{\pi(1 - \pi)}} \sim \mathcal{N}(0,1),$$

d'où

$$P(-1,96 \leq \frac{\sqrt{n}(\bar{X}_n - \pi)}{\sqrt{\pi(1 - \pi)}} \leq 1,96) = 95\%.$$

Comme p est l'estimation associée à \bar{X}_n , ceci conduit au test suivant, qui a un seuil de risque de 5% : si

$$\left| \frac{\sqrt{n}(p - \pi)}{\sqrt{\pi(1 - \pi)}} \right| > 1,96$$

alors on rejette l'hypothèse \mathcal{H}_0 .

NB : Si le test est faux, on ne peut pas conclure que la proportion *est* π , mais simplement que l'échantillon *ne permet pas d'exclure* que la proportion est π . On utilise ce test pour détecter (avec grande probabilité) les cas où la proportion *n'est pas* conforme à π . Par exemple pour vérifier si un fabricant fournit bien des pièces qui ont une précision donnée, ou des médicaments qui ont une certaine efficacité, etc.

Exemple 2 : Test sur une moyenne

Soit m la moyenne observée sur un échantillon de taille n . On veut savoir si cette moyenne est conforme à la moyenne théorique annoncée μ .

Hypothèse \mathcal{H}_0 : la moyenne théorique est μ .

Soit σ^2 la variance corrigée observée sur l'échantillon. Le test est le suivant : si

$$\left| \frac{\sqrt{n}(m - \mu)}{\sigma} \right| > 1,96$$

alors on rejette l'hypothèse \mathcal{H}_0 . C'est un test dont le seuil de risque est 5%.

Si l'hypothèse est rejetée, on peut donc conclure que la moyenne n'est pas μ , et on se trompe dans seulement 5% des cas.

Cours Commun Scientifique
de
Probabilités & Statistiques
—
Fiches d'exercices

FICHE 1 – CALCULS ÉLÉMENTAIRES,
PROBABILITÉS CONDITIONNELLES, LOI BINOMIALE

Exercice 1.

Combien de nombres peut-on former avec les chiffres 0, 1, 2, 3, 4, 5, 6 et 7, en tenant compte de l'ensemble des contraintes suivantes :

- chaque nombre est composé de chiffres différents ;
 - chaque nombre commence par un 7 ;
 - chaque nombre est divisible par 5,
1. si les nombres sont de 8 chiffres ?
 2. si les nombres sont de 6 chiffres ?

Exercice 2.

Un ensemble de délégués, composé de 6 étudiants d'*informatique*, de 7 étudiants d'*énergétique* et de 5 étudiants de *télécom*, doit élire un comité représentatif formé de 2 étudiants d'*informatique*, 2 de *télécom* et de 2 d'*énergétique* . Quel est le nombre de comités possibles si :

1. les 18 élèves sont éligibles ?
2. un étudiant d'*énergétique* est élu d'office ?
3. 2 étudiants d'*informatique* ne sont pas éligibles ?

Exercice 3.

Bruno et Alain jouent avec deux dés. La règle du jeu est la suivante :

- Bruno joue le premier, et gagne s'il fait un double ;
- sinon, Alain joue, et gagne si la somme des chiffres tirés est impaire ;
- si aucun d'eux n'a gagné, on joue un deuxième tour dans les mêmes conditions, et ainsi de suite.

Calculer la probabilité de succès des deux joueurs pour chacun des tours successifs, puis globalement. Discuter les résultats obtenus.

Exercice 4.

Les trois fils de Ken, au cours d'une partie de chasse, aperçoivent un superbe faisan huppé se percher sur la ligne électrique alimentant leur ferme. Ils tirent tous les trois en même temps. Les tirs de chasseurs différents sont supposés indépendants, et chacun d'eux tire une fois et une seule. On admet que chaque chasseur a une chance sur deux de tuer le volatile, et une chance sur cinq de couper le fil, ces deux événements étant également supposés indépendants.

1. Quelle est la probabilité pour chaque chasseur
 - de tuer le faisan et de couper le fil ?
 - de tuer le faisan ou de couper le fil ? (« ou » inclusif)
 - de tuer le faisan ou de couper le fil, mais pas les deux ? (« ou » exclusif)
2. Calculer la probabilité de ne pas être obligé de dîner le soir aux chandelles (en d'autres termes : quelle est la probabilité que le fil ne soit pas coupé ?).
3. Quelle est la probabilité pour le faisan ait reçu :
 - aucune balle ?
 - exactement deux balles ?
 - au moins deux balles ?

- au plus deux balles ?
 - exactement une balle ?
4. Montrer que les deux événements « le fil est coupé » et « le faisan est mort » sont indépendants.
- Suggestion* : on pourra montrer que les événements complémentaires le sont.
5. En déduire la probabilité de déguster le soir le faisan aux chandelles, après l'avoir fait rôtir dans la cheminée puisque le four électrique ne fonctionne plus.

Exercice 5.

Une population est composée de 40% d'hommes et de 60% de femmes. Dans cette population, 50% des femmes et 30% des hommes lisent plus de 10 romans par an. Quelle est la probabilité pour qu'un lecteur de plus de 10 romans par an, choisi au hasard, soit un homme ?

Exercice 6.

Lors d'une brève conversation vous apprenez d'un homme qu'il a deux enfants dont au moins une fille. Quelle est la probabilité pour qu'il ait deux filles ?

Exercice 7.

Deux usines fabriquent les mêmes pièces. La première en produit 70% de bonnes et la deuxième 90%. Les deux usines fabriquent la même quantité de pièces.

1. Quel est le pourcentage de bonnes pièces sur l'ensemble du marché, supposé alimenté par les deux usines ?
2. On achète une pièce, elle est bonne ; quelle est la probabilité pour qu'elle provienne de la deuxième usine ?
3. Mêmes questions lorsque la première usine produit 2,5 fois plus que la deuxième.

Exercice 8.

Un antiquaire possède trois coffres C_1, C_2, C_3 . Ces coffres ont chacun deux tiroirs et dans chacun de ces 6 tiroirs, il y a une pièce de monnaie.

Dans le coffre C_1 , il y a 2 pièces d'or,
 dans le coffre C_2 , il y a 2 pièces d'argent,
 dans le coffre C_3 , il y a 1 pièce d'or et 1 pièce d'argent.

1. On ouvre un des six tiroirs au hasard et on trouve une pièce en argent. Quelle est la probabilité pour que cette pièce vienne du coffre C_2 ?
2. On recommence l'expérience et on trouve une deuxième fois une pièce en argent. Quelle est la probabilité pour que l'on ait ouvert deux fois le même coffre ?

Exercice 9.

Le directeur d'un centre de loisirs ouvert aux adultes et adolescents de plus de 15 ans s'interroge sur la nécessité de recruter un moniteur de planche à voile. Il décide de le faire si la proportion des personnes choisissant cette activité est supérieure à 15%.

Une étude de la fréquentation du centre sur l'année précédente a montré que les activités choisies par les adultes ($\text{âge} \geq 18$ ans) fréquentant ce centre se répartissent entre : 50% pour le tennis, 40% pour les sports nautiques et 10% pour l'équitation. Parmi les adultes ayant choisi les sports nautiques, un quart fait de la planche à voile.

1. Déterminer quelle est la proportion d'adultes ayant choisi la planche à voile.
 Le groupe des adolescents ($15 \leq \text{âge} < 18$) a pour sa part des activités qui se répartissent entre : 55% pour le tennis, 40% pour les sports nautiques (dont la moitié pour la planche à voile), 5% pour l'équitation.
2. Déterminer parmi les adolescents la proportion de ceux qui font de la planche à voile.
3. Sachant qu'une personne sur 5 fréquentant ce centre est un adolescent, déterminer la proportion π_0 des usagers (adultes et adolescents confondus) faisant de la planche à voile. Le directeur du centre va-t-il recruter un moniteur spécialisé ?

Exercice 10.

Un modèle simplifié d'évolution du prix d'une action suppose que chaque jour le prix de l'action grimpe de 1 euro avec une probabilité p ($0 < p < 1$) et descend de 1 euro avec une probabilité $1 - p$. L'évolution d'un jour donné est indépendante de l'évolution des jours passés.

1. Quelle est la probabilité qu'après 2 jours le prix de l'action soit celui du premier jour ?
2. Quelle est la probabilité qu'après 3 jours le prix de l'action ait grimpé d'un euro ?
3. Sachant qu'après 3 jours le prix de l'action a grimpé d'un euro, quelle est la probabilité que le mouvement du premier jour ait été une hausse ?

Exercice 11.

Maryse joue à un jeu télévisé. Elle a, face à elle, trois portes (A, B et C) identiques. Derrière l'une d'elles se trouvent 5000 euros et derrière les deux autres rien du tout.

Maryse choisit une des portes (la porte A par exemple). L'animateur, qui connaît la porte gagnante, ouvre une autre porte (disons la C) et lui montre qu'il n'y a rien derrière. Il demande alors à Maryse si elle maintient son choix ou si elle préfère la porte B. Quel choix donne à Maryse la plus grande probabilité de gagner ? Justifiez votre réponse.

Exercice 12.

On considère un avion de 50 places. La probabilité pour qu'un voyageur ayant réservé ne se présente pas à l'embarquement est de 20%. Un jour la compagnie a enregistré 52 réservations. Quelle est la probabilité pour qu'elle se trouve dans une situation embarrassante ?

Exercice 13.

En tant qu'ingénieur, on sollicite votre avis sur un projet de lancement d'une fusée. Le moteur et l'équipement électronique se composent de 1 500 000 pièces distinctes. Chaque pièce a une chance sur dix millions de se révéler défectueuse. On précise que la défectuosité d'une seule pièce suffit à mettre le projet en échec, et que les causes de défectuosité de chacune des pièces sont statistiquement indépendantes. Qu'en pensez-vous ?

Exercice 14.

Un nouveau test de dépistage d'une maladie rare, touchant environ une personne sur 100 000, vient d'être mis au point. Pour tester sa validité, on a effectué un test statistique : sur 534 sujets sains, le test a été positif 1 seule fois, et, sur 17 sujets malades, il a été positif 16 fois.

Une personne effectue ce test ; le résultat est positif. Quelle est la probabilité pour qu'elle soit atteinte par cette maladie ?

Au vu de ces résultats, peut-on commercialiser le test ?

Exercice 15.

La probabilité pour que l'injection d'un vaccin à un individu choisi au hasard provoque une réaction allergique est de 0,1%. Quelle est la probabilité pour que, sur 900 individus vaccinés, on observe l'allergie dans :

1. exactement trois cas ?
2. au plus trois cas ?
3. au moins trois cas ?

Exercice 16.

Un système de communication est constitué de n composants qui fonctionnent indépendamment les uns des autres avec une probabilité pour chacun d'entre eux de $p \in]0,1[$. Le système est opérationnel dès qu'au moins la moitié de ses composants fonctionnent. Pour quelles valeurs de p un système à 5 composants a une plus forte probabilité d'être opérationnel qu'un système à 3 composants ?

FICHE 2 – VARIABLES ALÉATOIRES (I)

Exercice 1.

Le trousseau de clés d'un gardien de nuit comporte dix clés, dont une seule ouvre la porte du poste de garde. Pour qu'il y pénètre, il y a deux scénarios possibles :

- Cas A : il prend une clé au hasard, l'essaie, la met de côté si elle n'ouvre pas, et ainsi de suite.
- Cas B : il prend une clé au hasard, l'essaie, mais la laisse sur le trousseau si elle n'ouvre pas, et ainsi de suite.

On désigne respectivement par X_A et X_B les variables aléatoires égales aux nombres d'essais (y compris le bon) avant le succès, dans le premier et le second scénarios.

Déterminer la loi de probabilité et la fonction de répartition de X_A et de X_B .

Calculer $E[X_A]$ et $E[X_B]$.

Le gardien utilise la méthode B un jour sur trois. Un jour, après avoir essayé 8 clés, il n'a toujours pas ouvert la porte. Quelle est la probabilité pour qu'il ait utilisé la méthode B ?

Exercice 2.

On considère une fonction F définie par

$$F(x) = \begin{cases} a & \text{si } x \leq -1, \\ bx + c & \text{si } x \in]-1, 1[\\ d & \text{si } x \geq 1. \end{cases}$$

1. Déterminer les valeurs de a , b , c et d pour que F soit une fonction de répartition d'une variable aléatoire à densité.
2. Représenter graphiquement F .
3. Déterminer la densité de probabilité associée à F .

Exercice 3.

Soit f la fonction de \mathbb{R} dans \mathbb{R}^+ définie par

$$\text{pour tout } x \in [2, 4], \quad f(x) = \lambda \frac{1}{(1-x)^2} \quad \text{et, si } x \notin [2, 4], \quad f(x) = 0.$$

1. Déterminer $\lambda \in \mathbb{R}$ pour que f soit une densité de probabilité.
2. Soit X une variable aléatoire de densité f . Déterminer sa fonction de répartition F .
3. Calculer l'espérance $E[X]$ et la variance $\text{Var}(X)$ de la variable aléatoire X .

On pourra utiliser les égalités

$$x = (x-1) + 1, \quad x^2 = (x-1)^2 + 2(x-1) + 1.$$

Exercice 4.

Dans l'urne du bureau de vote du quartier Saint-Roman à Roquebrune (06), les électeurs ont déposé, à l'occasion de l'élection présidentielle, N bulletins de vote, dont S pour le candidat des verts. Le président du bureau extrait successivement, *sans les remettre*, n bulletins de l'urne. On désigne par X la variable aléatoire égale au nombre de bulletins verts sur les n .

Remarque : pour que le calcul qui suit ait un sens, on est amené à supposer que la taille n de l'échantillon extrait de l'ensemble des bulletins est suffisamment petite pour que $n < S$.

1. Montrer que X suit la loi hypergéométrique :

$$\text{pour } 0 \leq k \leq n, \quad P(X = k) = \frac{C_S^k C_{N-S}^{n-k}}{C_N^n}.$$

2. Vérifier la normalisation de cette loi. Pour cela, on pourra développer les deux membres de l'identité :

$$(1+x)^N = (1+x)^S (1+x)^{N-S}$$

et comparer les coefficients.

3. Calculer la moyenne $E[X]$ et la variance $\text{Var}(X)$. On montrera que :

$$E[X] = np \quad \text{et} \quad \text{Var}(X) = np(1-p) \frac{N-n}{N-1}$$

où p désigne la proportion de bulletins verts : $S = pN$.

Exercice 5.

On considère une variable aléatoire réelle X dont la fonction de répartition $F_X(x)$ est donnée par

$$F_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - (1 + \frac{x}{2})e^{-\frac{x}{2}} & \text{si } x \geq 0. \end{cases}$$

1. $F_X(x)$ est-elle continue sur \mathbb{R} ?
2. Déterminer $\lim_{x \rightarrow +\infty} F_X(x)$. Interprétation ?
3. Calculer la densité de probabilité $f_X(x)$. Quel est le mode de X ? (C'est-à-dire la valeur x où $f_X(x)$ est maximale)
4. Calculer l'espérance et l'écart-type de X .
5. Déduire de ce qui précède les variations et la courbe représentative de $F_X(x)$.
6. Calculer $P(1 \leq X < 2)$.

Exercice 6.

Soit T une variable aléatoire, de densité de probabilité f_T donnée par :

$$f_T(t) = \lambda(1-t^2)\mathbf{1}_{[-1,1]}(t).$$

1. Calculer λ et représenter graphiquement $f_T(t)$.
2. Déterminer la fonction de répartition $F_T(t)$ et tracer sa courbe représentative.
3. Calculer la probabilité de l'événement $\{|T| \geq \frac{1}{2}\}$. Représenter cette probabilité sur chacun des deux graphiques précédents.
4. Calculer l'espérance et la variance de T . Quelle est sa médiane ? (C'est-à-dire la valeur t telle que $P(T \leq t) = \frac{1}{2}$)

Exercice 7.

Soit U une variable aléatoire de loi uniforme sur $[0,1]$ (sa densité de probabilité, f_U est définie par $f_U(x) = \mathbf{1}_{[0,1]}(x)$). On pose

$$X = -\frac{1}{\lambda} \ln(U).$$

On note F_X la fonction de répartition de X et f_X sa densité de probabilité.

1. Rappeler la fonction de répartition $F_U(u) = P(U \leq u)$, pour $u \in \mathbb{R}$, de la variable aléatoire U .
2. Déterminer la fonction de répartition F_X et la densité de probabilité de la variable aléatoire X .

3. Quelle est la loi de X ? Donner les valeurs de $E[X]$ et $\text{Var}(X)$. On rappelle que, pour tout $n \in \mathbb{N}$,

$$\int_0^{+\infty} x^n e^{-x} dx = n! .$$

Exercice 8.

On cherche à comparer l'efficacité de deux méthodes pour réaliser dans un délai très court un contrôle médical sur 1000 personnes.

Ce contrôle utilise un prélèvement sanguin et doit permettre de déterminer la présence ou l'absence d'un virus dont on sait qu'il atteint un individu donné avec la probabilité 0,01.

Méthode A : On teste séparément les 1000 personnes ; ce qui conduit à réaliser 1000 analyses.

Méthode B : On répartit les 1000 individus en n groupes de r personnes (avec $nr = 1000$). Pour chaque groupe, on mélange ensemble les prélèvements sanguins effectués sur les r personnes du groupe et on analyse le mélange. Si le résultat est positif pour un groupe, alors on analyse séparément le sang des r personnes qui composent ce groupe. Donc si y est le nombre de groupes positifs, la méthode B conduit à faire $n + ry$ analyses.

1. Étude de la méthode B.

a) Quelle est la probabilité q pour qu'un groupe de r personnes soit négatif? En déduire la probabilité p pour qu'un groupe de r personnes soit positif. Donner en fonction de p et q la loi de probabilité et l'espérance de la variable aléatoire

$$Y = \text{nombre de groupes positifs} .$$

b) Calculer en fonction de r l'espérance de la variable aléatoire

$$X = \text{nombre d'analyses dans la méthode B.}$$

Montrer qu'une valeur approchée de $E[X]$ est

$$E[X] \simeq 10 \left(r + \frac{100}{r} \right) .$$

2. Comparaison des deux méthodes.

Montrer que l'on peut trouver des valeurs de r pour lesquelles la méthode B est meilleure que la méthode A. Pour quelle valeur de r la méthode B est-elle optimale? Pour cette valeur, quel est le nombre moyen d'analyses effectuées?

Exercice 9.

Chaque page du site web de l'institut Galilée comporte un grand nombre de caractères, et un gros effort a été fait pour éviter coquilles et autres erreurs. On constate cependant qu'on y trouve en moyenne trois fautes toutes les deux pages. Déterminez la probabilité pour que la page que vous êtes en train de consulter présente au moins trois fautes.

Exercice 10.

Le nombre de connexions par jour au site web de l'institut Galilée est une variable poissonnienne de paramètre c . Chaque connexion peut provenir d'un ordinateur ou d'un smartphone. On note p la probabilité pour qu'elle provienne d'un ordinateur. Déterminer la loi suivie par le nombre de connexions par jour établies à partir d'un ordinateur.

Exercice 11.

Un avion peut transporter 100 passagers et leurs bagages. Il pèse sans les passagers mais avec l'équipage et le carburant 120 tonnes. Les consignes de sécurité interdisent le décollage si le poids de l'appareil dépasse 129,42 tonnes.

Les 100 places ont été occupées. Le poids d'un voyageur suit une loi d'espérance 70 kg et d'écart type 10 kg. Le poids de ses bagages suit une loi d'espérance 20 kg et d'écart type 10 kg. Toutes ces variables sont supposées indépendantes.

1. Calculer l'espérance du poids de l'avion au décollage. Est-elle conforme aux normes de sécurité ?
2. Calculer l'écart type du poids total de l'appareil.
3. En utilisant l'inégalité de Bienaymé-Tchebychev, trouver un majorant de la probabilité pour que le poids réel de l'appareil au décollage dépasse 129,42 tonnes.

Exercice 12.

1. Soient X et Y deux variables aléatoires indépendantes de lois de Poisson respectives $\mathcal{P}(\lambda)$ et $\mathcal{P}(\mu)$.

a) Soit n un entier strictement positif. En remarquant que

$$\{X + Y = n\} = \bigcup_{k=0}^n (\{X = k\} \cap \{Y = n - k\}),$$

montrer que $X + Y$ suit une loi de Poisson de paramètre $\lambda + \mu$.

b) Calculer, pour $k = 0, 1, \dots, n$,

$$P(X = k | X + Y = n).$$

En déduire que la loi de probabilité conditionnelle de X sachant $X + Y = n$ est la loi binomiale

$$\mathcal{B}\left(n, \frac{\lambda}{\lambda + \mu}\right).$$

2. Application :

Soit deux échantillons d'articles produits en série. On sait que les nombres X et Y d'articles défectueux dans chacun des deux échantillons suivent respectivement des lois de Poisson $\mathcal{P}(2)$ et $\mathcal{P}(3)$. On réunit les deux échantillons. Sachant que le nombre total d'articles défectueux est 3, quelle est la probabilité de l'événement $\{X \leq 1\}$?

Exercice 13.

Soit Y une variable aléatoire de loi $\mathcal{E}(1/2)$. On appelle g sa densité (la rappeler) et G sa fonction de répartition.

On étudie la loi d'une variable aléatoire X telle que $Y = X^2$ et dont la densité f est paire. On appelle F sa fonction de répartition.

1. En écrivant que pour tout $a > 0$ on a

$$P(0 \leq Y \leq a) = P(-\sqrt{a} \leq X \leq \sqrt{a}),$$

montrer que

$$G(a) - G(0) = F(\sqrt{a}) - F(-\sqrt{a}).$$

2. En déduire la valeur de $f(\sqrt{a})$ pour tout $a > 0$, puis que

$$\text{pour tout } x \in \mathbb{R}, \quad f(x) = \frac{|x|}{2} e^{-\frac{1}{2}x^2}.$$

3. Déterminer $E[X]$ et $\text{Var}(X)$.

4. Comment pourrait-on définir X à partir de Y ? Introduire une variable aléatoire ε indépendante de Y et de loi donnée par $P(\varepsilon = 1) = \frac{1}{2} = P(\varepsilon = -1)$ et considérer $X = \varepsilon\sqrt{Y}$.

Exercice 14.

Une urne contient des boules, dont une proportion p ($0 < p < 1$) de boules blanches, les autres boules étant rouges. On tire les boules une à une avec remise.

1. Soit X_n la variable aléatoire égale au nombre de boules blanches obtenues au cours des n premiers tirages. Quelle est la loi de X_n ?

2. Soit Y_1 la variable aléatoire égale au rang d'apparition de la première boule blanche. Quelle est la loi de Y_1 ? Retrouver pour tout entier $k \geq 1$ la probabilité $P(Y_1 = k)$ et calculer l'espérance de Y_1 .

On pourra dériver terme à terme l'égalité suivante :

$$(E) \quad \text{pour tout } q \text{ tel que } |q| < 1, \quad \sum_{k=0}^{+\infty} q^k = \frac{1}{1-q}.$$

3. Soit $n > 1$. On note Y_n la variable aléatoire égale au rang d'apparition de la n -ième boule blanche.

a) Quelles sont les valeurs prises par Y_n ?

b) Après avoir remarqué que, pour tout $k \geq n$,

$$\{Y_n = k\} = \{X_{k-1} = n-1\} \cap \{\text{la } k\text{-ième boule est blanche}\},$$

donner la loi de Y_n .

c) Calculer l'espérance de Y_n . On pourra dériver n fois terme à terme l'égalité (E).

Exercice 15.

On pose pour tout $a > 0$,

$$\Gamma(a) = \int_0^{+\infty} x^{a-1} e^{-x} dx.$$

On rappelle que pour tout $n \in \mathbb{N}^*$, $\Gamma(n) = (n-1)!$.

Une variable aléatoire Z suit une loi Gamma de paramètres a et θ ($a > 0$ et $\theta > 0$), notée $\Gamma(a, \theta)$, si sa densité de probabilité est

$$f(x) = \frac{1}{\Gamma(a)\theta^a} x^{a-1} e^{-\frac{x}{\theta}} \quad \text{si } x > 0, \quad f(x) = 0 \quad \text{sinon.}$$

1. Vérifier que f est bien une densité de probabilité. Quelle est la moyenne de la variable aléatoire $Z \sim \Gamma(a, \theta)$?

2. Le nombre d'appels reçus par un standard téléphonique pendant une durée t ($t > 0$) est une variable aléatoire X_t suivant une loi de Poisson de paramètre λt où λ est un réel strictement positif.

Soit Y_n le temps d'arrivée du n -ième appel (à partir de $t = 0$).

a) Remarquer que $P(Y_1 > t) = P(X_t = 0)$.

En déduire que Y_1 suit une loi exponentielle $\mathcal{E}(\lambda)$.

b) Plus généralement, calculer $P(Y_n > t)$ en utilisant la variable aléatoire X_t .

En déduire que la fonction de répartition de Y_n est :

$$F_n(t) = 1 - e^{-\lambda t} \sum_{k=0}^{n-1} \frac{(\lambda t)^k}{k!} \quad \text{si } t \geq 0, \quad F_n(t) = 0 \quad \text{sinon,}$$

puis que Y_n suit une loi Gamma dont on donnera les paramètres.

Exercice 16.

On cherche à caractériser deux sources photoniques par la valeur moyenne λ du nombre de photons émis par unité de temps. Le nombre de photons émis par unité de temps suit une loi de Poisson.

On sait que les sources appartiennent à deux familles distinctes (F_0, F_1) associées aux paramètres λ_0 et λ_1 . Par ailleurs on sait qu'une source sur 10 est une source F_1 .

1. Donner les probabilités $P(k | F_0)$ et $P(k | F_1)$ d'observer k photons sachant que la source provient de la famille F_0 ou F_1 respectivement.

2. Calculer en fonction de λ_0 et λ_1 la probabilité $P(k)$ d'observer k photons en prenant une source au hasard.

3. On a observé k photons. Calculer la probabilité $P(F_0 | k)$ que la source observée soit une source F_0 . De même calculer $P(F_1 | k)$.

On décide que la source observée appartient à la famille F_1 si $P(F_1 | k) > P(F_0 | k)$.

4. Montrer que cette règle de décision conduit à choisir F_1 si

$$\frac{P(k | F_1)}{P(k | F_0)} > 9.$$

5. En déduire une règle de décision permettant de choisir directement F_0 ou F_1 suivant le nombre k de photons observés.

Application numérique : $\lambda_0 = 1$, $\lambda_1 = 4$.

Exercice 17.

Le conseil d'administration d'une banque décide d'organiser sa gestion de manière à ce qu'il y ait 999 chances sur 1000 de toujours pouvoir faire face aux demandes de retrait de ses clients. La banque a 1000 clients, le dépôt de chaque client est de 1000 euros. La probabilité pour qu'un client retire son argent un jour donné est 0,001. Dans ces conditions, combien la banque doit-elle conserver de liquidités journalières pour suivre le principe de gestion qui a été posé ? (On pourra utiliser une loi de Poisson.)

Exercice 18.

On lance une pièce de monnaie n fois.

1. Soit X la variable aléatoire égale au nombre de piles obtenus. Quelle est la loi de probabilité suivie par X ? Préciser son espérance m et sa variance σ^2 en fonction de n .

2. Soit $Y = \frac{X}{n}$ la proportion de piles obtenus. Déterminer $E[Y]$ et $\text{Var}(Y)$.

3. En utilisant l'inégalité de Bienaymé-Tchebychev, déterminer un nombre de lancers n tel que

$$P(0,4 < Y < 0,6) \geq 90\%.$$

Exercice 19.

Une variable aléatoire X suit une loi normale $\mathcal{N}(m(\lambda), 1)$ où λ est un paramètre prenant les valeurs 0, 1, 2 avec les probabilités

$$P(\lambda = 0) = a \quad P(\lambda = 1) = 1 - 2a \quad P(\lambda = 2) = a,$$

et la moyenne de X est donnée par $m(\lambda) = \lambda - 1$.

1. Quelles sont les conditions sur a pour que le support de la loi de λ soit $\{0, 1, 2\}$?

2. Soit Z une variable aléatoire de loi normale centrée réduite. Soit $x \in \mathbb{R}$. On pose

$$\beta(x) = P(Z > x).$$

Déterminer en fonction de $\beta(x)$ les probabilités

$$P(Z < x), \quad P(Z < -x), \quad P(|Z| < x), \quad P(0 < Z < x).$$

3. La table de la loi gaussienne centrée réduite donne les valeurs de $\Phi(x) = P(Z < x)$. Déduire de la table les valeurs $\beta(0)$, $\beta(1)$ et $\beta(2)$.

4. Calculer, en fonction de a , la probabilité $P(-1 < X < 1)$.

FICHE 3 – VARIABLES ALÉATOIRES (II),
COUPLES DE VARIABLES ALÉATOIRES

Exercice 1.

On considère deux variables aléatoires indépendantes X et Y , définies sur \mathbb{R} . Ces deux variables suivent chacune la loi normale centrée réduite $\mathcal{N}(0,1)$.

1. Soit Z la variable aléatoire définie par $Z = X + Y$ et T la variable aléatoire définie par $T = aX$ où a est un réel.

Déterminer les espérances $E[Z]$ et $E[T]$ puis les variances $\text{Var}(Z)$ et $\text{Var}(T)$ des variables aléatoires Z et T .

Sachant que toute combinaison linéaire de variables aléatoires gaussiennes indépendantes est une variable aléatoire gaussienne, en déduire les lois et les densités de probabilité des variables aléatoires Z et T .

2. a) Soit f une fonction paire et $x \in \mathbb{R}_+$. Montrer que

$$\int_{-\sqrt{x}}^{\sqrt{x}} f(t) dt = \int_0^x \frac{f(\sqrt{u})}{\sqrt{u}} du.$$

b) On appelle U la variable aléatoire définie par $U = X^2$. Déterminer pour tout réel x la valeur de la probabilité

$$P(U \leq x).$$

En déduire la densité de probabilité de la variable aléatoire U .

Exercice 2.

On dit qu'une variable aléatoire Z suit une loi $\Gamma(\lambda, n)$, où $n \in \mathbb{N}$ et $\lambda > 0$, si Z a pour densité de probabilité la fonction :

$$f_Z(z) = \frac{\lambda^n}{(n-1)!} z^{n-1} e^{-\lambda z} \mathbf{1}_{\mathbb{R}_+}(z)$$

1. Vérifier que la loi $\Gamma(\lambda, n)$ est bien définie, puis calculer $E[Z]$, $E[Z^2]$ et $\text{Var}(Z)$.

2. Soient X et Y deux variables aléatoires indépendantes de lois respectives $\Gamma(\lambda, m)$ et $\Gamma(\lambda, n)$. Déterminer la loi jointe du couple (U, V) défini par :

$$U = X + Y, \quad V = \frac{X}{X + Y}$$

3. Déterminer les lois marginales de U et V . Ces variables sont-elles indépendantes ?

On donne :

$$\int_0^1 t^{m-1} (1-t)^{n-1} dt = \frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)}$$

où Γ est la fonction d'Euler :

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt$$

et on rappelle que pour $n \in \mathbb{N}$, $\Gamma(n) = (n-1)!$.

Exercice 3.

Un atelier fonctionne avec deux équipes d'ouvriers, une du matin et l'autre du soir. Chaque jour on enregistre le nombre d'ouvriers absents. On note X (respectivement Y), le nombre d'absences dans l'équipe de jour (respectivement de nuit). La loi jointe P de (X, Y) est donnée par

$$\begin{aligned} P(0,0) = c & \quad P(0,1) = 2c & \quad P(0,2) = 0 & \quad P(0,3) = 3c \\ P(1,0) = 0 & \quad P(1,1) = 3c & \quad P(1,2) = 4c & \quad P(1,3) = 0 \\ P(2,0) = 0 & \quad P(2,1) = 4c & \quad P(2,2) = 2c & \quad P(2,3) = c \end{aligned}$$

1. Déterminer la constante c .
2. Donner les lois marginales de X et Y ainsi que leurs espérances.
3. Une absence coûte 30 euros à l'usine. Quelle est la perte journalière moyenne due aux absences ?

Exercice 4.

1. Soit $u \in [0,1]$, $v \in [0,1]$, et le changement de variables

$$x = (-2 \ln u)^{\frac{1}{2}} \cos(2\pi v)$$

$$y = (-2 \ln u)^{\frac{1}{2}} \sin(2\pi v)$$

Déterminer à quel ensemble appartiennent x et y et le jacobien J associé à ce changement de variables. Écrire J en fonction de u puis en fonction de x et y .

2. Soient U et V deux variables aléatoires indépendantes de loi uniforme sur $[0,1]$.

Soient X et Y les variables aléatoires définies par

$$X = (-2 \ln U)^{\frac{1}{2}} \cos(2\pi V)$$

$$Y = (-2 \ln U)^{\frac{1}{2}} \sin(2\pi V)$$

- a) Préciser la densité de probabilité du couple (U, V) .
- b) Déterminer la densité du couple (X, Y) et montrer que les variables aléatoires X et Y sont indépendantes et suivent une loi normale centrée réduite.

Exercice 5.

On considère deux variables aléatoires indépendantes X_1 et X_2 , obéissant à des lois définies respectivement par les fonctions de répartition F_{X_1} et F_{X_2} . On pose $U = \max(X_1, X_2)$ et $V = \min(X_1, X_2)$.

1. Déterminer les lois de U et de V . Que deviennent-elles si X_1 et X_2 suivent la même loi, de fonction de répartition F ?
2. On suppose à présent X_1 et X_2 équiréparties sur l'intervalle $[0,1]$. Déterminer la probabilité conditionnelle $P(U > \frac{3}{4} \mid V < \frac{1}{3})$.

Exercice 6.

On considère un couple (X_1, X_2) de variables aléatoires de loi uniforme sur le carré $[0,1] \times [0,1]$.

1. Montrer que X_1 et X_2 sont deux variables indépendantes.
2. On pose $Y_1 = \frac{X_1}{X_2}$ et $Y_2 = X_1 + X_2$. Déterminer la loi jointe de ces deux nouvelles variables.
3. Déterminer les lois marginales de Y_1 et Y_2 . Ces deux variables sont-elles indépendantes ?

Exercice 7.

Soit (X,Y) un couple de variables aléatoires de densité

$$f(x,y) = \frac{1}{2x} \quad \text{si } (x,y) \in D, \quad f(x,y) = 0 \quad \text{sinon}$$

avec

$$D = \left\{ (x,y) \mid 0 < y \leq x \quad \text{et} \quad 0 < y \leq \frac{1}{x} \right\} .$$

1. Donner une représentation graphique de D .
2. Déterminer les densités des variables aléatoires X et Y .
Les variables aléatoires X et Y sont-elles indépendantes ?
3. On pose $S = \sqrt{XY}$ et $T = \sqrt{\frac{Y}{X}}$.
 - a) Montrer que l'ensemble des valeurs prises par S et T est l'intervalle $]0,1]$.
 - b) Soit $(s,t) \in]0,1]^2$ et $(x,y) \in D$ tels que $s = \sqrt{xy}$ et $t = \sqrt{\frac{y}{x}}$.
Calculer x et y en fonction de s et t .
 - c) Déterminer la densité $h(s,t)$ du couple (S,T) .

Exercice 8.

1. Soit D le domaine du plan

$$D = \{(x,y) \mid x > 0, y > 0, y^2 > x\}.$$

Dessiner D .

2. Soit (X,Y) un couple de variables aléatoires à valeurs réelles de densité de probabilité

$$f(x,y) = \begin{cases} \frac{1}{2\sqrt{x}} e^{-y} & \text{si } (x,y) \in D \\ 0 & \text{ailleurs.} \end{cases}$$

Déterminer les lois de X et Y . Les variables aléatoires X et Y sont-elles indépendantes ?

Exercice 9.

Lors d'un salon de l'artisanat, un artisan a un stand où il ne vend que deux produits.

Le produit A au prix unitaire de 8 euros, le produit B au prix unitaire de 12 euros.

On suppose dans la suite de l'exercice,

- que les quantités achetées par les différents clients sont indépendantes,
- que le nombre de produits achetés par une personne s'arrêtant au stand est :
 - pour le produit A , une variable aléatoire de Poisson de paramètre 2 notée X ,
 - pour le produit B , une variable aléatoire de Poisson de paramètre 1 notée Y ,
- que les variables aléatoires X et Y sont indépendantes.

1. Calculer les probabilités, pour une personne s'arrêtant au stand :

- qu'elle n'achète rien
- qu'elle achète uniquement des produits A
- qu'elle achète uniquement des produits B
- qu'elle achète des produits A et des produits B
- qu'elle achète des produits A ou des produits B (ou inclusif).

2. On note $Z = 8X + 12Y$ la variable aléatoire égale au montant en euros de l'achat d'une personne s'arrêtant au stand. Calculer $E[Z]$ et $\text{Var}(Z)$.

3. On a remarqué que 350 personnes en moyenne s'arrêtaient au stand chaque jour. Soit W la variable aléatoire égale au chiffre d'affaire sur une journée de l'artisan.

- a) Déterminer la moyenne et la variance de W .
- b) En utilisant l'inégalité de Bienaymé-Tchebychev, trouver un minorant de la probabilité $P(9000 < W < 10600)$.

c) On suppose que la loi de W peut être approchée par une loi normale de moyenne $m = 9800$ et d'écart type $\sigma = 309$.

- Quelle est la loi suivie par la variable aléatoire $W^* = \frac{W-9800}{309}$?
- Avec la table de la loi gaussienne centrée réduite, trouver, si $Z \sim \mathcal{N}(0,1)$,

$$P(Z \leq 0,16) \quad \text{et} \quad P(Z \leq 0,49).$$

- En déduire la probabilité pour que le chiffre d'affaire W d'une journée soit compris entre 9750 et 9951 euros.

Exercice 10.

1. Soit \mathcal{C} le carré de sommets $A (0,0)$, $B (0,1)$, $C (1,1)$, $D (1,0)$.

Déterminer, pour tout $u \in [0,1]$, l'aire $D(u)$ du domaine $\mathcal{D}(u)$ défini par

$$\mathcal{D}(u) = \mathcal{C} \cap \{(x,y) \mid -u \leq x - y \leq u\} .$$

2. Deux personnes se donnent rendez-vous entre 19h et 20h. On suppose que l'on peut associer aux instants d'arrivée de ces deux personnes des variables aléatoires X et Y continues, indépendantes et uniformes sur $[0,1]$. Soit $U = |X - Y|$ la variable aléatoire égale au temps d'attente de la première personne. Montrer que la fonction de répartition de U est donnée par

$$F(u) = \begin{cases} 0 & \text{si } u < 0 \\ D(u) & \text{si } u \in [0,1] \\ 1 & \text{si } u > 1. \end{cases}$$

3. Calculer la densité de U et son espérance $E[U]$.

4. Ces deux personnes conviennent que la première arrivée s'en ira après un temps d'attente égal à $2E[U]$. Quelle est la probabilité pour que le rendez-vous ait lieu ?

FICHE 4 – LOIS LIMITES, ESTIMATION

Exercice 1.

Trois machines, A , B , C fournissent respectivement 50%, 30%, 20% de la production d'une usine. Les pourcentages de pièces défectueuses produites par ces machines sont respectivement 3%, 4%, 5%.

1. Quelle est la probabilité pour qu'une pièce prise au hasard dans la production soit défectueuse ?
2. On prélève un échantillon de 100 pièces sur la production. Quelle est la loi suivie par le nombre de pièces défectueuses ? Par quelle loi peut-on l'approcher ?
3. Quelle est la probabilité pour qu'une pièce défectueuse prise au hasard provienne de A , de B , de C ?

Exercice 2.

Une compagnie d'assurance assure 500 navires pour une somme de 5 millions d'euros chacun. Chaque navire a chaque année une probabilité égale à 0,1% de subir un sinistre majeur couvert par l'assurance. Soit X le nombre de navires perdus en une année. Donner la loi de X , son espérance et sa variance.

On cherche à estimer quelles réserves doit posséder la compagnie d'assurance pour être sûre de pouvoir payer les indemnités avec une probabilité égale à 99,9% à la fin de chaque année.

1. Pourquoi peut-on estimer que la variable aléatoire X suit une loi de Poisson ? Quelle est le paramètre de cette loi ?

La table de la loi de Poisson pour le paramètre $\lambda = 0,5$ donne

$$P(X \leq 3) = 0,9982 \quad \text{et} \quad P(X \leq 4) = 0,9998.$$

Conclure.

2. Une seconde compagnie d'assurance assure également 500 navires dans les mêmes conditions que la précédente. Les compagnies ont-elles intérêt à fusionner ?

La table de la loi de Poisson pour le paramètre $\lambda = 1$ donne

$$P(X \leq 4) = 0,9963 \quad \text{et} \quad P(X \leq 5) = 0,9994.$$

Exercice 3.

Montrer que tout estimateur T_n asymptotiquement sans biais dont la variance tend vers 0 est convergent. *Utiliser l'inégalité de Markov.*

Exercice 4.

Soit (X_1, X_2) un échantillon d'une variable aléatoire X admettant pour densité :

$$f(x) = \begin{cases} \frac{3x^2}{\theta^3} & \text{si } 0 < x < \theta \\ 0 & \text{ailleurs.} \end{cases}$$

1. Calculer $E[X]$, en déduire un estimateur T_1 de θ par la méthode des moments.
2. Soit $Z = \max(X_1, X_2)$. Calculer la fonction de répartition de Z et en déduire sa densité g . Calculer $E[Z]$ et en déduire un estimateur T_2 de θ par la méthode des moments.
3. Montrer que T_1 et T_2 sont sans biais. Quel est le meilleur ?

Exercice 5.

Soit X une variable aléatoire admettant pour densité

$$f(x) = \theta x^{\theta-1} \mathbf{1}_{]0,1]}(x),$$

où $\theta > 0$ est un paramètre.

1. Calculer $E[X]$ et $\text{Var}(X)$.

2. Soit $Y = \ln X$.

a) Préciser l'ensemble des valeurs prises par Y . Déterminer la fonction de répartition et la densité de Y .

b) Calculer

$$\int_{-\infty}^0 x \theta e^{x\theta} dx$$

et en déduire $E[Y]$.

3. Soit (X_1, X_2, \dots, X_n) un échantillon de la variable aléatoire X .

a) On pose $\alpha = \frac{1}{\theta}$.

Écrire la vraisemblance $L(x_1, x_2, \dots, x_n; \alpha)$ de l'échantillon pour α . En déduire un estimateur $\hat{\alpha}$ de α par la méthode du maximum de vraisemblance. Montrer que cet estimateur est sans biais.

b) On pose $\beta = \frac{\theta}{\theta+1}$. Donner un estimateur sans biais de β par la méthode des moments.

Exercice 6.

Soit X une variable aléatoire de loi uniforme sur $[0, \theta]$, où $\theta > 0$ est un paramètre. Soit (X_1, X_2, \dots, X_n) un échantillon de X . Montrer que :

1. $T_1 = 2\overline{X}_n$ est un estimateur sans biais et convergent de θ ;

2. $T_2 = \max(X_1, X_2, \dots, X_n)$ est un estimateur asymptotiquement sans biais et convergent de θ .

Exercice 7.

Soit X une variable aléatoire dont la loi dépend de deux paramètres p_1 et p_2 par :

$$P(X = 0) = 1 - p_1 - p_2, \quad P(X = 1) = p_1 \quad \text{et} \quad P(X = 2) = p_2.$$

1. Trouver les conditions à vérifier par p_1 et p_2 pour que le support de la loi de X soit égal à $\{0, 1, 2\}$. Calculer $E[X]$, $E[X^2]$, $\text{Var}(X)$.

2. Soit (X_1, X_2, \dots, X_n) un échantillon de X . Déterminer des estimateurs L_1 et L_2 de p_1 et p_2 par la méthode des moments. Montrer qu'ils sont sans biais.

3. Pour $j = 0, 1, 2$, on désigne par N_j le nombre de X_k égaux à j . Écrire la vraisemblance de l'échantillon en fonction de p_1 , p_2 , N_0 , N_1 , N_2 . Déterminer les estimateurs Z_1 et Z_2 de p_1 et p_2 par la méthode du maximum de vraisemblance.

4. Montrer que $L_1 = Z_1$ et $L_2 = Z_2$.

5. Un échantillon de taille $n = 100$ de X a donné les observations suivantes :

$$N_0 = 20, \quad N_1 = 50, \quad N_2 = 30 .$$

À quelles estimations de p_1 et p_2 conduisent les estimateurs L_1 et L_2 ?

Exercice 8.

Soit X une variable aléatoire admettant pour densité de probabilité :

$$f(x) = \begin{cases} a & 0 \leq x < \frac{1}{2} \\ b & \frac{1}{2} \leq x \leq 1 \\ 0 & \text{sinon.} \end{cases}$$

1. Indiquer les relations que doivent satisfaire a et b pour que f soit une densité de probabilité de support $[0,1]$. Exprimer b en fonction de a , paramètre que l'on cherche à estimer.
2. Calculer $E[X]$, $E[X^2]$ et $\text{Var}(X)$ en fonction de a .
3. Soit (X_1, X_2, \dots, X_n) un échantillon de X . Déterminer un estimateur L_n de a par la méthode des moments. Montrer que cet estimateur est sans biais et que sa variance tend vers 0 quand $n \rightarrow +\infty$; en déduire qu'il est convergent.
4. On désigne par N_0 le nombre de variables aléatoires X_i appartenant à l'intervalle $[0, \frac{1}{2}[$. Montrer que l'estimateur Z_n de a par la méthode du maximum de vraisemblance peut s'écrire

$$Z_n = \frac{2}{n} \sum_{i=1}^n Y_i$$

où $Y_i = \mathbf{1}_{\{X_i < \frac{1}{2}\}}$ est la variable aléatoire égale à 1 si $X_i \in [0, \frac{1}{2}[$, et égale à 0 si $X_i \in [\frac{1}{2}, 1[$.

5. Quel est le meilleur des estimateurs Z_n et L_n ?

Exercice 9.

Soit X une variable aléatoire suivant une loi de Poisson de paramètre $\lambda > 0$. On considère un échantillon (X_1, X_2, \dots, X_n) de X .

1. On pose $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$. Que valent $E[\bar{X}_n]$ et $\text{Var}(\bar{X}_n)$?

2. a) Ecrire la vraisemblance $L(k_1, k_2, \dots, k_n; \lambda)$ de l'échantillon (X_1, X_2, \dots, X_n) .

b) En déduire que l'équation de la vraisemblance est

$$\sum_{i=1}^n \left(-1 + \frac{k_i}{\lambda} \right) = 0,$$

et que l'estimateur T_n de λ par maximum de vraisemblance est donc égal à \bar{X}_n .

Exercice 10.

Soient (X_1, X_2, \dots, X_n) , n variables aléatoires indépendantes et de même loi de densité de probabilité

$$f(x) = \frac{1}{a} e^{-\frac{x}{a}} \mathbf{1}_{]0, +\infty[}(x),$$

où $a > 0$ est un paramètre.

1. Calculer $E[X_i]$ et en déduire un estimateur de a par la méthode des moments.
2. Écrire la vraisemblance $L(x_1, x_2, \dots, x_n; a)$ de l'échantillon (X_1, X_2, \dots, X_n) . En déduire l'estimateur du maximum de vraisemblance de a .
3. On définit la variable aléatoire

$$I = \min(X_1, X_2, \dots, X_n).$$

- a) Déterminer, pour tout $x \in \mathbb{R}$, la probabilité $P(X_i \geq x)$, puis la probabilité $P(I \geq x)$.
- b) En déduire la fonction de répartition de la variable aléatoire I , puis sa densité de probabilité.
- c) Calculer $E[I]$ et en déduire un estimateur sans biais de a de la forme kI .

Exercice 11.

Un directeur de centre de loisirs souhaite embaucher un moniteur de planche à voile si la proportion de personnes choisissant cette activité est supérieure à 15%.

Il a observé que $\pi_0 = 12\%$ des personnes fréquentant le centre de loisirs avaient choisi la planche à voile l'année dernière. Il souhaite aujourd'hui réactualiser sa décision. Pour cela, il dispose de plus du pourcentage $\pi_1 = 16\%$ des personnes ayant choisi la planche à voile parmi les n premières venues en juillet. Il souhaite estimer avec un risque de 5% si π_1 est conforme aux observations de l'année précédente.

On définit n variables aléatoires indépendantes et de même loi $(X_i)_{1 \leq i \leq n}$ par :

$$X_i = \begin{cases} 1 & \text{si la } i\text{-ième personne choisit la planche à voile} \\ 0 & \text{si la } i\text{-ième personne choisit un autre sport.} \end{cases}$$

1. Soit $\pi_0 = E[X_i]$, que vaut $\text{Var}(X_i)$?
2. On note $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$. Que valent $E[\bar{X}_n]$ et $\sigma(\bar{X}_n)$?
3. Déterminer en fonction de n un intervalle I_n tel que

$$P(\bar{X}_n \in I_n) = 95\%.$$

On rappelle que si $Z \sim \mathcal{N}(0,1)$, $P(-1,96 \leq Z \leq 1,96) = 95\%$.

4. Que vaut I_n lorsque $n = 30$? Le pourcentage π_1 est-t-il significativement différent, au risque de 5%, du pourcentage π_0 obtenu l'année dernière ?
5. Sur quel nombre minimum de personnes aurait-il dû observer la proportion π_1 pour en conclure qu'il devait peut-être changer d'avis ?

Exercice 12.

Soit Y une variable aléatoire de Bernoulli de paramètre p et \bar{Y}_n la variable aléatoire

$$\bar{Y}_n = \frac{1}{n}(Y_1 + Y_2 + \dots + Y_n),$$

les Y_i étant des variables aléatoires indépendantes et de même loi que Y .

1. Quel théorème permet de dire que si $n \geq 30$, la variable aléatoire

$$\sqrt{n} \left(\frac{\bar{Y}_n - p}{\sqrt{p(1-p)}} \right)$$

suit approximativement la loi gaussienne centrée réduite $\mathcal{N}(0,1)$? Énoncer ce théorème dans le cas général où Y est une variable aléatoire de moyenne m et d'écart type σ .

2. On pose $n = 100$ et $p = 0,1$. Sur la table de la loi gaussienne, on voit que $P(Z < 1,96) \simeq 0,975$ si $Z \sim \mathcal{N}(0,1)$. En déduire le plus petit réel A tel que

$$P(\bar{Y}_{100} > A) \leq 0,025 .$$

3. Application. Un acheteur veut tester les livraisons qu'il reçoit et savoir si elles sont conformes à ce que dit le fabricant. Le fabricant assure que moins de 10% des vis livrées sont défectueuses. L'acheteur prélève 100 vis à chaque livraison et relève la proportion π de vis défectueuses. Pour quelles valeurs de π , l'acheteur peut-il légitimement refuser la livraison au seuil de risque de 2,5% ?

Exercice 13.

On cherche à comparer les pourcentages de satisfaction de deux groupes de personnes utilisant un même produit cosmétique.

Chacune des personnes du groupe A a une probabilité π_0 d'être satisfaite.

Chacune des personnes du groupe B a une probabilité π_1 d'être satisfaite.

On dispose d'un échantillon de 50 personnes dans le groupe A , dont 15 se déclarent satisfaites.

Pour le groupe B , sur un échantillon de 100 personnes, 60 sont satisfaites.

1. Quels sont les pourcentages p_0 et p_1 de satisfaction observés sur ces deux échantillons ?

2. Soit une variable aléatoire X de loi normale de moyenne m et d'écart type σ : $X \sim \mathcal{N}(m, \sigma)$.

Rappeler quelle est la loi suivie par la variable aléatoire

$$Y = \frac{X - m}{\sigma} .$$

3. Soit \bar{X} une variable aléatoire de Bernoulli de paramètre π_0 . Montrer que la moyenne empirique \bar{X}_{50} suit une loi normale de moyenne π_0 et d'écart type

$$\sqrt{\frac{\pi_0(1 - \pi_0)}{50}} .$$

De la même façon, donner la loi suivie par \bar{Y}_{100} , où Y est une variable aléatoire de Bernoulli de paramètre π_1 .

4. Soit D la variable aléatoire $D = \bar{X}_{50} - \bar{Y}_{100}$. Quelle est la loi suivie par D ?

5. Soit \mathcal{H}_0 l'hypothèse « $\pi_0 = \pi_1$ », que l'on souhaite tester. Sous cette hypothèse,

$$D \sim \mathcal{N}\left(0, p(1 - p)\left(\frac{1}{50} + \frac{1}{100}\right)\right)$$

où

$$p = \frac{50p_0 + 100p_1}{150}$$

est la moyenne des pourcentages observés sur les deux échantillons. Préciser la valeur de p .

a) Déterminer un intervalle I tel que sous l'hypothèse \mathcal{H}_0

$$P(|D| \in I) = 95\% .$$

(On rappelle que si $Z \sim \mathcal{N}(0,1)$, $P(|Z| \leq 1,96) = 95\%$)

b) Quelle est la valeur observée pour D ? Conclure : les pourcentages de satisfaction des deux populations sont-ils significativement différents au seuil de 5% ?

Exercice 14.

La durée de fonctionnement d'une ampoule suit une loi normale $\mathcal{N}(m, \sigma^2)$. On examine 50 ampoules et on enregistre les résultats suivants :

$$\bar{X}_{50}(\omega) = 58 \text{ jours} \quad \text{et} \quad \frac{1}{50} \sum_{i=1}^{50} (X_i(\omega) - \bar{X}_{50}(\omega))^2 = 99 \text{ jours}^2 .$$

Déterminer un intervalle de confiance de niveau 95% de la moyenne m .

Exercice 15.

Sur un échantillon de 300 pièces fabriquées par une machine, 60 sont défectueuses.

Trouver un intervalle de confiance I de niveau 99% de la proportion réelle de pièces défectueuses fabriquées par la machine.

Exercice 16.

On a observé que l'âge X des étudiants assidus en première année d'université suivait une loi normale $\mathcal{N}(18,1)$ et que l'âge Y des étudiants non assidus suivait la loi normale $\mathcal{N}(20,1)$.

1. Calculer la probabilité pour qu'un étudiant assidu ait plus de 20 ans.

2. Dans un premier temps, pour organiser les groupes de TD, il est décidé d'un test basé sur l'âge de l'étudiant pour évaluer son assiduité.

Soit \mathcal{H}_0 l'hypothèse : « l'étudiant est assidu » et \mathcal{H}_1 l'hypothèse alternative : « l'étudiant n'est pas assidu ».

On décide de rejeter l'hypothèse \mathcal{H}_0 si l'âge de l'étudiant est supérieur à $A = 20$ ans.

Soit

$$\alpha = P(\text{rejeter } \mathcal{H}_0 \mid \mathcal{H}_0 \text{ vraie})$$

le risque de première espèce et

$$\beta = P(\text{ ne pas rejeter } \mathcal{H}_0 \mid \mathcal{H}_0 \text{ fausse}) = P(\text{ ne pas rejeter } \mathcal{H}_0 \mid \mathcal{H}_1 \text{ vraie})$$

le risque de deuxième espèce.

a) Remarquer que le risque de première espèce, α , est égal à $P(X > 20)$. Que vaut β ?

b) On admet un risque de première espèce égal à 5%. Quelle valeur peut-on prendre pour A ? Que vaut alors β ?

Exercice 17.

1. La table statistique donnée en annexe donne la probabilité $\Phi(t)$ pour qu'une variable aléatoire normale Z centrée réduite ($Z \sim \mathcal{N}(0,1)$) soit inférieure ou égale à une valeur donnée t .

Soit X une variable normale de moyenne m et d'écart type σ . Rappeler comment se calcule, en fonction de Z , la probabilité $P(X < A)$.

2. Un grossiste reçoit des pelotes de laine provenant de deux usines différentes. Toutes les pelotes d'une même usine n'ont pas exactement le même poids, mais le poids (en grammes) d'une pelote prise au hasard dans un lot de même provenance suit une loi normale $\mathcal{N}(m, \sigma^2)$. Les usines A et B annoncent respectivement

$$m_A = 100\text{g}, \quad \sigma_A = 8\text{g} \quad ; \quad m_B = 96\text{g}, \quad \sigma_B = 4\text{g}.$$

a) Un premier client refuse toute pelote dont le poids est inférieur à 90g. De quelle usine provient le lot donnant le plus faible pourcentage de rebut ?

b) Pour un autre client, le poids minimum exigé est inconnu ; mais on constate qu'il est tel que le choix entre les deux provenances est indifférent. Déterminer ce poids minimum exigé ainsi que le pourcentage de rebut (commun au deux lots).

3. Le grossiste de son côté souhaite vérifier si les lots reçus sont conformes aux qualifications annoncées. Pour cela il prélève un échantillon de 100 pelotes dans le lot provenant de l'usine A et détermine le poids moyen m d'une pelote de cet échantillon.

a) Soient $(X_1, X_2, \dots, X_{100})$, 100 variables aléatoires indépendantes de même loi $\mathcal{N}(m_A, \sigma_A^2)$. On note \bar{X}_{100} la moyenne de ces variables aléatoires :

$$\bar{X}_{100} = \frac{1}{100} \sum_{i=1}^{100} X_i.$$

Quelle loi peut approcher la loi suivie par la variable aléatoire

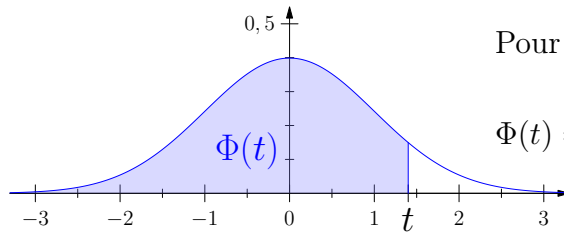
$$Y = 10 \frac{\bar{X}_{100} - m_A}{\sigma_A} ?$$

Énoncer le théorème utilisé.

b) Quelles sont les valeurs de m pour lesquelles il pourra rejeter, au seuil de risque 5%, l'hypothèse $H_0 : « m = m_A »$?

On rappelle que si $Z \sim \mathcal{N}(0,1)$, $P(|Z| \leq 1,96) = 95\%$.

Table de la fonction de répartition de la loi normale centrée réduite $\mathcal{N}(0,1)$



Pour Z de loi $\mathcal{N}(0,1)$,

$$\Phi(t) = F_Z(t) = P(Z \leq t) = \int_{-\infty}^t e^{-\frac{x^2}{2}} \frac{dx}{\sqrt{2\pi}}$$

t	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,1	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,4	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998
3,5	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998
3,6	0,9998	0,9998	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
3,7	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
3,8	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
3,9	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000

Exemples : $\Phi(0,25) \simeq 0,5987$, $\Phi(-0,32) = 1 - \Phi(0,32) \simeq 1 - 0,6255 = 0,3745$

Memento des lois usuelles

Lois discrètes

Nom	Paramètres	Support	Définition : $P(A) = \sum_{a \in A} p(a)$	Espérance	Variance
Loi de Dirac δ_a	$a \in \mathbb{R}$	$\{a\}$	$p(a) = 1$	a	0
Loi de Bernoulli $\mathcal{B}(p)$	$p \in [0,1]$	$\{0,1\}$	$p(0) = 1 - p, p(1) = p$	p	$p(1 - p)$
Loi binomiale $\mathcal{B}(n,p)$	$n \in \mathbb{N}, p \in [0,1]$	$\{0, \dots, n\}$	$p(k) = \binom{n}{k} p^k (1 - p)^{n-k}$	np	$np(1 - p)$
Loi géométrique $\mathcal{G}(p)$	$p \in]0,1]$	\mathbb{N}^*	$p(k) = (1 - p)^{k-1} p$	$\frac{1}{p}$	$\frac{1 - p}{p^2}$
Loi de Poisson $\mathcal{P}(\lambda)$	$\lambda \in]0, +\infty[$	\mathbb{N}	$p(k) = e^{-\lambda} \frac{\lambda^k}{k!}$	λ	λ

Lois continues

Nom	Paramètres	Support	Définition : $P(A) = \int_A f(x) dx$	Espérance	Variance
Loi uniforme $\mathcal{U}([a,b])$	$a < b$	$[a,b]$	$f(x) = \frac{1}{b - a} \mathbf{1}_{[a,b]}(x)$	$\frac{a + b}{2}$	$\frac{(b - a)^2}{12}$
Loi exponentielle $\mathcal{E}(\lambda)$	$\lambda \in]0, \infty[$	$]0, +\infty[$	$f(x) = \lambda e^{-\lambda x} \mathbf{1}_{]0, +\infty[}(x)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Loi normale/gaussienne $\mathcal{N}(m, \sigma^2)$	$m \in \mathbb{R}, \sigma^2 \in]0, +\infty[$	\mathbb{R}	$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - m)^2}{2\sigma^2}\right)$	m	σ^2