Probabilités et statistiques

Denis Vekemans [*]

Table des matières

[*]Laboratoire de mathématiques pures et appliquées Joseph Liouville ; 50, rue Ferdinand Buisson BP 699 ; 62 228 Calais cedex ; France

1 Combinatoire

1.1 Analyse combinatoire sans répétition

1.1.1 Cardinal de l'ensemble des applications de E dans F

Soit E un ensemble de cardinal p et F un ensemble de cardinal n. On note $\mathcal{F}(E,F)$ l'ensemble des applications de E dans F. On a

$$\#\mathcal{F}(E,F) = n^p.$$

1.1.2 Cardinal de l'ensemble des injections de E dans F

Soit E un ensemble de cardinal p et F un ensemble de cardinal n. On note $\mathcal{I}(E,F)$ l'ensemble des injections de E dans F. On a

$$\#\mathcal{I}(E,F) = A_n^p = n(n-1)\dots(n-p+1) = \frac{n!}{(n-p)!}.$$

On appelle p-arrangement de F (ensemble de cardinal n), toute application injective f de $[[1,p]]$ dans F. En identifiant f au p-uplet $\{f(1), f(2), \dots, f(p)\}$, un p-arrangement apparaît comme une suite ordonnée de p éléments distincts de F. Le nombre des p-arrangements de F est donc A_n^p.

1.1.3 Cardinal de l'ensemble des bijections de E dans F

Soit E un ensemble de cardinal n et F un ensemble de cardinal n. On note $\mathcal{B}(E,F)$ l'ensemble des bijections de E dans F. On a

$$\#\mathcal{B}(E,F) = A_n^n = n!.$$

On appelle permutation de E toute bijection de E dans E. Le nombre de permutations de E est $n!$.

1.1.4 Cardinal de l'ensemble des parties de E

Soit E un ensemble de cardinal n et soit p un entier tel que $0 \leq p \leq n$.

Le nombre de parties de E de cardinal p est

$$C_n^p = \begin{pmatrix} n \\ p \end{pmatrix} = \frac{n!}{p!(n-p)!}.$$

On appelle p-combinaison de E (ensemble de cardinal n), toute partie de E de cardinal p. Le nombre des p-combinaisons de E est donc C_n^p.

Propriétés de C_n^p :

- $C_n^p = C_n^{n-p}$;
- $C_n^p = C_{n-1}^p + C_{n-1}^{p-1}$.

Formule du binôme de Newton : étant donnés deux réels a et b et un entier naturel n, on a

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

1.2 Analyse combinatoire avec répétition

Soit E un ensemble de cardinal n et soient $n_1, n_2, \ldots, n_r$ r entiers naturels tels que $\sum_{k=0}^{r} n_k = n$.

On appelle $(n_1, n_2, \ldots, n_r)$-permutation (avec répétition) de E toute application de E dans $[[1, r]]$ telle que $\# f^{-1}(k) = n_k$.

Le nombre des $(n_1, n_2, \ldots, n_r)$-permutation (avec répétition) de E est

$$P_{n_1, n_2, \ldots, n_r} = \frac{n!}{n_1! n_2! \ldots n_r!}.$$

Formule du multinôme : étant donnés r réels a_1, a_2, …, a_r et un entier naturel n, on a

$$(a_1 + a_2 + \ldots + a_r)^n = \sum_{n_1+n_2+\ldots n_r = n} P_{n_1, n_2, \ldots, n_r} a_1^{n_1} a_2^{n_2} \ldots a_r^{n_r}.$$

Exercice 1 [9, exercice 1, p. 7] Soit E un ensemble de cardinal n, montrer que $\#\mathcal{P}(E) = 2^n$.

Exercice 2 [9, exercice 2, p. 7] Soit f une application d'un ensemble E de cardinal n dans lui même. Montrer que

$$f \text{ est injective} \iff f \text{ est surjective} \iff f \text{ est bijective.}$$

Exercice 3 [9, exercice 4, p. 8] Soit $A_n = \sum_{p=0}^{n} A_n^p$ (où $A_n^0 = 1$) le nombre total des arrangements d'un ensemble de cardinal n. Montrer que A_n est la partie entière de $n!e$.

Exercice 4 [9, exercice 5, p. 8] Soient 4 individus A, B, C et D. De combien de manières peut-on les classer

– en supposant qu'il n'y ait pas d'*ex aequo* ?
– en supposant qu'il puisse y avoir 0, 1 ou plusieurs *ex aequo* ?

Exercice 5 [9, exercice 8, p. 10] Soit E un ensemble à n éléments. Quel est le nombre de couples de parties (A, B) de E telles que $A \subset B$?

Généralisation. Soit E un ensemble à n éléments. Quel est le nombre de p-uplets de parties $(A_1, A_2, \ldots, A_p)$ de E telles que $A_1 \subset A_2 \subset \ldots \subset A_p$?

Exercice 6 [9, exercice 10, p. 11] Démontrer les identités suivantes

1. $C_n^0 + C_n^2 + C_n^4 + \ldots = C_n^1 + C_n^3 + C_n^5 + \ldots = 2^{n-1}$, $\forall n \in [[1, \infty[[$.
2. $0C_n^0 + 2C_n^2 + 4C_n^4 + \ldots = 1C_n^1 + 3C_n^3 + 5C_n^5 + \ldots$, $\forall n \in [[2, \infty[[$.
3. $C_n^0 - \frac{1}{2}C_n^1 + \frac{1}{3}C_n^2 - \frac{1}{4}C_n^3 + \ldots = \frac{1}{n+1}$, $\forall n \in [[1, \infty[[$.

Exercice 7 [9, exercice 12, p. 12] Soit $S_r(n) = 1^r + 2^r + \ldots + n^r$. Montrer que

$$\sum_{k=0}^{r} C_{r+1}^k S_k(n) = (n+1)^{r+1} - 1.$$

Calculer $S_0(n)$, $S_1(n)$, $S_2(n)$, $S_3(n)$.

Exercice 8 [9, exercice 13, p. 12]Calculer les 3 sommes

1. $S_1 = C_n^0 + C_n^3 + C_n^6 + \ldots$;

2. $S_2 = C_n^1 + C_n^4 + C_n^7 + \ldots$;
3. $S_3 = C_n^2 + C_n^5 + C_n^8 + \ldots$.

$\boxed{\textbf{Exercice 9}}$ [9, exercice 14, p. 13] Soit $E = [[1, n]]$.

1. Quel est le nombre de parties de E de cardinal k dont deux éléments quelconques ne sont pas consécutifs ?
2. Quel est le nombre de *mots* de longueur n formé des deux lettres a et b et dans lequel deux lettres a ne se suivent pas ?

$\boxed{\textbf{Exercice 10}}$ [9, exercice 15, p. 14]

1. Étant donnés des entiers naturels a, b et n, montrer que
$$C_{a+b}^n = \sum_{k=0}^{n} C_a^k C_b^{n-k}.$$
2. En déduire les identités
$$\begin{aligned} C_{2n}^n &= \sum_{k=0}^{n} \left(C_n^k\right)^2. \\ \frac{n}{2} C_{2n}^n &= \sum_{k=0}^{n} k \left(C_n^k\right)^2. \end{aligned}$$

$\boxed{\textbf{Exercice 11}}$ [9, exercice 18, p. 16] De combien de manières peut-on ranger p billes indiscernables dans n boîtes ?

$\boxed{\textbf{Exercice 12}}$ [9, exercice 19, p. 16] Soient n et p des entiers supérieurs ou égaux à 1. Quel est le nombre de vecteurs $(x_1, x_2, \ldots, x_n) \in \mathbb{N}^n$ solutions de l'équation

1. $$x_1 + x_2 + \ldots + x_n = p \ ?$$
2. $$x_1 + x_2 + \ldots + x_n \leq p \ ?$$

$\boxed{\textbf{Exercice 13}}$ [9, exercice 20, p. 16] Pour tout entier $n \in [[1, \infty[[$, on note $S(n)$ la somme des chiffres de l'écriture de n dans le système de numération décimale.

1. Étant donnés $k \in \mathbb{N}$ et $p \in [[1, 9]]$, combien y a-t-il de nombres entiers n tels que
$$0 \leq n < 10^k \text{ et } S(n) = p \ ?$$
2. Étant donnés $k \in \mathbb{N}$, combien y a-t-il de nombres entiers n tels que
$$1 \leq n \leq 10^k \text{ et } S(n) = 10 \ ?$$

$\boxed{\textbf{Exercice 14}}$ [9, exercice 21, p. 17] Soit $E = [[1, m]]$ et $F = [[1, n]]$.

1. Quel est le nombre des applications strictement croissantes de E dans F ?
2. Quel est le nombre des applications croissantes au sens large de E dans F ?

Exercice 15 [9, problème 24, p. 18] Soit $E = \{e_1, e_2, \ldots, e_n\}$.

1. (a) Soit Γ une permutation de E. Montrer que les propositions suivantes sont équivalentes :

   i. Γ^k est sans point fixe pour $k \in [[1, n-1]]$ et $\Gamma^n = Id$.

   ii. $\forall i \in [[1, n]],\ E = \left\{\Gamma^k(e_i) | k \in [[0, n-1]]\right\}$.

   iii. $\exists i \in [[1, n]],\ E = \left\{\Gamma^k(e_i) | k \in [[0, n-1]]\right\}$.

   iv. Il existe une permutation P de E telle que

   $$\Gamma \circ P(e_i) = P(e_{i+1}) \text{ pour } i \in [[1, n-1]] \text{ et } \Gamma \circ P(e_n) = P(e_1).$$

   Lorsque l'une de ces propositions est vraie, on dit que Γ est une *permutation circulaire* de E.

   (b) Montrer que le nombre de permutations circulaires de E est $(n-1)!$.

2. Soit D_n le nombre de permutations sans point fixe (ou dérangements) de E.

   (a) Montrer que pour $n \in [[3, \infty[[,\ D_n = (n-1)(D_{n-1} + D_{n-2})$ avec $D_1 = 0$ et $D_2 = 1$.

   (b) Montrer par récurrence que
   $$D_n = \sum_{k=0}^{n} (-1)^k \frac{n!}{k!}.$$
   En déduire que lorsque n tend vers l'infini, D_n est équivalent à $\frac{n!}{e}$.

2 Statistique descriptive

2.1 Série statistique

- $X = (x_1, \ldots, x_p)$. Les x_i sont appelées les *observations*.
- $\{(x_i, n_i),\ i \in [[1, p]]\}$. Les n_i sont appelées les *effectifs* et si $n = \sum_{i=1}^{p} n_i$, les $f_i = \frac{n_i}{n}$ sont appelées les *fréquences*.

Fonction de répartition.

$F : \mathbb{R} \to \mathbb{R}; x \mapsto \sum_{i \text{ tel que } x_i \leq x} f_i$ est une fonction en escalier, croissante de 0 à 1, continue à droite.

2.2 Série statistique classée

C'est le cas lorsque $x_i = [a_i, a_{i+1}[$ et x_i est appelée une *classe*. Les a_i forment une suite croissante. Il est une hypothèse usuelle de répartition homogène à l'intérieur de chaque classe.

Ensuite, $h(x) = h_i = \frac{\phi_i}{\alpha_i} = \frac{\nu_i}{n\alpha_i}$ si $x \in [a_i, a_{i+1}[$; $h(x) = 0$ si $x < a_1$; $h(x) = 0$ si $x \geq a_{p+1}$

- ν_i est l'*effectif de la classe* $[a_i, a_{i+1}[$;
- $\phi_i = \frac{\nu_i}{n}$ est la *fréquence de la classe* $[a_i, a_{i+1}[$;
- $\alpha_i = a_{i+1} - a_i$ est l'*amplitude* de la classe $[a_i, a_{i+1}[$.

Histogramme.

On appelle histogramme la représentation graphique de $h : \mathbb{R} \to \mathbb{R}; x \mapsto h(x)$.

Fonction de répartition.

$\Phi : \mathbb{R} \to \mathbb{R}; x \mapsto \int_{-\infty}^{x} h(t)dt$ est une fonction croissante de 0 à 1, continue, affine par morceaux, telle que $\Phi(a_i) = \sum_{j<i} \phi_j$.

2.3 Caractéristiques des séries statistiques

Quelques éléments caractéristiques de position.

- **Le mode** : toute valeur (respectivement classe) telle que n_i (respectivement h_i) soit maximum.
- **La médiane** Me. Dans le cas de série non classée : toute valeur μ telle que $\sum_{x_i\leq\mu} n_i \geq \frac{n}{2}$ et $\sum_{x_i\geq\mu} n_i \geq \frac{n}{2}$; dans le cas de série classée : l'ensemble des valeurs μ telles que $\Phi(\mu) = 0,5$.
- **Le premier quartile** Q_1. Dans le cas de série non classée : toute valeur μ telle que $\sum_{x_i\leq\mu} n_i \geq \frac{n}{4}$ et $\sum_{x_i\geq\mu} n_i \geq \frac{n}{4}$; dans le cas de série classée : l'ensemble des valeurs μ telles que $\Phi(\mu) = 0,25$.
- **Le troisième quartile** Q_3. Dans le cas de série non classée : toute valeur μ telle que $\sum_{x_i\leq\mu} n_i \geq \frac{3n}{4}$ et $\sum_{x_i\geq\mu} n_i \geq \frac{3n}{4}$; dans le cas de série classée : l'ensemble des valeurs μ telles que $\Phi(\mu) = 0,75$.
- On définirait de façon analogue les déciles (D_1, …, D_9).
- **La moyenne** $\overline{x}$. Dans le cas de série non classée : $\sum_{i=1}^{p} f_i x_i$; dans le cas de série classée : $\int_{-\infty}^{\infty} th(t)dt$.

Quelques éléments caractéristiques de dispersion.

- **L'étendue**. Dans le cas de série non classée : $\max_i x_i - \min_i x_i$; dans le cas de série classée : $a_{p+1} - a_p$.
- **L'intervalle interquartile** : $Q_3 - Q_1$ (si plusieurs Q_i correspondent, on prend la moyenne de ceux qui correspondent).
- **L'intervalle interdécile** : $D_9 - D_1$ (si plusieurs D_i correspondent, on prend la moyenne de ceux qui correspondent).
- **L'écart moyen** E_m. Dans le cas de série non classée : $\sum_{i=1}^{p} f_i|x_i - \overline{x}|$; dans le cas de série classée : $\int_{-\infty}^{\infty} |t - \overline{x}|h(t)dt$.
- **La variance** V_x **et l'écart-type** σ_x. Dans le cas de série non classée : $V_x = \sum_{i=1}^{p} f_i(x_i - \overline{x})^2$; dans le cas de série classée : $V_x = \int_{-\infty}^{\infty}(t - \overline{x})^2 h(t)dt$. Dans tous les cas, $\sigma_x = \sqrt{V_x}$.

Quelques éléments caractéristiques de forme.

Soit m_k le moment d'ordre k. Dans le cas de série non classée : $m_k = \sum_{i=1}^{p} f_i x_i^k$; dans le cas de série classée : $m_k = \int_{-\infty}^{\infty} t^k h(t)dt$.

Soit μ_k le moment centré d'ordre k. Dans le cas de série non classée : $\mu_k = \sum_{i=1}^{p} f_i|x_i - \overline{x}|^k$; dans le cas de série classée : $\mu_k = \int_{-\infty}^{\infty}(t - \overline{x})^k h(t)dt$.

- **Coefficient d'asymétrie** : $\frac{\mu_3}{\sigma_x^3}$.
- **Coefficient d'applatissement** : $\frac{\mu_4}{\sigma_x^4}$.

2.4 Série statistique double

C'est le cas lorsque $x_i = (y_i, z_i)$ et x_i est appelée un couple.

Quelques éléments propres aux séries statistiques doubles. On restreint ici au cas non classé.

- **La covariance** : $Cov_{(y,z)} = \sum_{i=1}^{p} f_i(y_i - \overline{y})(z_i - \overline{z})$.
- **L'ajustement affine** : $Z = aY + b$ avec $a = \frac{Cov_{(y,z)}}{V_y}$, $b = \overline{z} - a\overline{y}$; et $r_{(y,z)}$, le coefficient de corrélation $r_{(y,z)} = \frac{Cov_{(y,z)}}{\sigma_y \sigma_z}$.

  Cet ajustement est basé sur la minimisation de la somme des carrés des écarts en ordonnée entre le nuage de point et la droite.

Pour fixer les idées sur la statistique descriptive, on détaillera les quatre exercices suivants en rappelant s'il y a lieu les définitions principales.

Exercice 16 Cas d'une entrée simple discrète.

Les résultats obtenus à partir du lancer d'un dé à 6 faces sur 100 expériences sont donnés dans le tableau suivant :

Résultats	1	2	3	4	5	6
Nombre d'observations	15	17	14	18	20	16

1. Donner le mode (i.e. le résultat le plus fréquent).
2. Tracer la fonction cumulative croissante des fréquences relatives.
3. Donner la médiane et l'écart interquartile.
4. Donner la moyenne des observations.
5. Donner l'écart-type des observations.

Solution :

Résultats	1	2	3	4	5	6
Nombre d'observations	15	17	14	18	20	16
Fréquences relatives	$0,15$	$0,17$	$0,14$	$0,18$	$0,20$	$0,16$
Fréquences relatives cumulées	$0,15$	$0,32$	$0,46$	$0,64$	$0,84$	$1,00$

1. Mode : 5.
2. Tracer.
3. $Q_1 = 2$.
   $Q_2 = 4$.
   $Q_3 = 5$.
   $Q_3 - Q_1 = 3$ (écart interquartile).
4. Moyenne : $3,59$.
5. Ecart-type : $\simeq 1,69$.

Exercice 17 Cas d'une entrée simple continue.

Le taux de triglycérides est observé chez 250 hommes de 20 à 30 ans.

On relève les résultats suivants :

Triglycérides (en g/l)	$[0.0; 0.6[$	$[0.6; 0.8[$	$[0.8; 1.0[$	$[1.0; 1.2[$	$[1.2; 1.4[$	$[1.4; \infty[$
Nombre d'observations	5	32	86	89	32	6

1. Donner la classe modale (i.e. la classe la plus fréquente).
2. Tracer la fonction cumulative croissante des fréquences relatives.
3. Donner la médiane et l'écart interquartile.
   Aide : pour avoir des classes de même longueur, on remplacera dans le calcul de la moyenne et de l'écart-type, la première par $[0.4; 0.6[$ et la dernière par $[1.4; 1.6[$.
4. Donner la moyenne des observations avec 4 décimales.
5. Donner l'écart-type des observations avec 4 décimales.

Solution :

Classes	$[0.0;0.6[$	$[0.6;0.8[$	$[0.8;1.0[$	$[1.0;1.2[$	$[1.2;1.4[$	$[1.4;\infty[$
Centres des classes	$0,5$	$0,7$	$0,9$	$1,1$	$1,3$	$1,5$
Nombre d'observations	5	32	86	89	32	6
Fréquences relatives	$0,020$	$0,128$	$0,344$	$0,356$	$0,128$	$0,024$
Fréquences relatives cumulées	$0,020$	$0,148$	$0,492$	$0,848$	$0,976$	$1,000$

1. Classe modale : $[1.0;1.2[$.
2. Tracer.
3. $\frac{Q_1-1}{0,25-0,492} = \frac{0,8-1}{0,148-0,492} \simeq 0,581 \Longrightarrow Q_1 \simeq 0,859$.
   $\frac{Q_2-1}{0,5-0,492} = \frac{1,2-1}{0,848-0,492} \simeq 0,561 \Longrightarrow Q_2 \simeq 1,004$ (médiane).
   $\frac{Q_3-1}{0,75-0,492} = \frac{1,2-1}{0,848-0,492} \simeq 0,561 \Longrightarrow Q_3 \simeq 1,144$.
   $Q_3 - Q_1 \simeq 0,285$ (écart interquartile).
4. Moyenne : $\simeq 1,0032$.
5. Ecart-type : $\simeq 0,2025$.

Exercice 18 Cas d'une entrée double discrète.

On donne un tableau statistique à double entrée X et Y. Pour l'observation i, X prend la valeur x_i, Y prend la valeur y_i, et ceci dénombré n_i fois.

x_i	2	2	2	3	3	3	4	4	5
y_i	4	5	6	6	7	8	9	10	11
n_i	2	5	3	1	6	2	3	1	1

1. Donner les moyennes à 10^{-3} près.
2. Donner les variances à 10^{-3} près.
3. Donner la covariance à 10^{-3} près.
4. Donner la droite de régression de Y en fonction de X.
5. Donner le coefficient de corrélation et conclure quant à la qualité de cet ajustement.
6. Donner une estimation de Y, lorsque X prend la valeur 1, à 10^{-2} près.

<u>Solution</u> :

1. $E(X) = \frac{68}{24} \simeq 2,833$.
   $E(Y) = \frac{163}{24} \simeq 6,791$.
2. $V(X) = \frac{210}{24} - \frac{68^2}{24^2} = \frac{416}{24^2} \simeq 0,722$.
   $V(Y) = \frac{1187}{24} - \frac{163^2}{24^2} = \frac{1919}{24^2} \simeq 3,331$.
3. $Cov(X,Y) = \frac{497}{24} - \frac{68 \cdot 163}{24^2} = \frac{844}{24^2} \simeq 1,465$.
4. $y = \frac{844}{416}x + \frac{163}{24} - \frac{844 \cdot 68}{416 \cdot 24} \simeq 1,043 + 2,028x$.
5. $\rho(X,Y) = \frac{844}{\sqrt{416 \cdot 1919}} \simeq 0,94$.
   Très bon ajustement.
6. Si $\hat{x} = 1$, alors $\hat{y} = \frac{844}{416} + \frac{163}{24} - \frac{844 \cdot 68}{416 \cdot 24} = \frac{1278}{416} \simeq 3,07$.

Exercice 19 Cas d'une entrée double discrète.

On donne un tableau statistique à double entrée X et Y. Pour l'observation i, X prend la valeur x_i, Y prend la valeur y_i, et ceci dénombré 1 fois. On prend i variant de 0 à $n-1$ (avec $n \geq 3$), $x_i = \cos(\frac{2i\pi}{n})$ et $y_i = \sin(\frac{2i\pi}{n})$.

1. Donner les moyennes.
2. Donner les variances.
3. Donner la covariance.
4. Donner la droite de régression de Y en fonction de X.
5. Donner le coefficient de corrélation et conclure quant à la qualité de cet ajustement.

Solution :

1.

$$\begin{aligned}
nE(X) &= \sum_{i=0}^{n-1} \cos(\frac{2i\pi}{n}) \\
&= \underbrace{(\frac{1}{2}\left(1+e^{\frac{2\pi I}{n}}+e^{\frac{4\pi I}{n}}+\ldots+e^{\frac{2(n-1)\pi I}{n}}\right))}_{=0} + \underbrace{(\frac{1}{2}\left(1+e^{-\frac{2\pi I}{n}}+e^{-\frac{4\pi I}{n}}+\ldots+e^{-\frac{2(n-1)\pi I}{n}}\right))}_{=0} \\
&= 0;
\end{aligned}$$

$$\begin{aligned}
nE(Y) &= \sum_{i=0}^{n-1} \sin(\frac{2i\pi}{n}) \\
&= \underbrace{(\frac{1}{2I}\left(1+e^{\frac{2\pi I}{n}}+e^{\frac{4\pi I}{n}}+\ldots+e^{\frac{2(n-1)\pi I}{n}}\right))}_{=0} - \underbrace{(\frac{1}{2I}\left(1+e^{-\frac{2\pi I}{n}}+e^{-\frac{4\pi I}{n}}+\ldots+e^{-\frac{2(n-1)\pi I}{n}}\right))}_{=0} \\
&= 0.
\end{aligned}$$

2.

$$\begin{aligned}
nV(X) &= \sum_{i=0}^{n-1} \cos^2(\frac{2i\pi}{n}) \\
&= \underbrace{(\frac{1}{4}\left(1+e^{\frac{4\pi I}{n}}+e^{\frac{8\pi I}{n}}+\ldots+e^{\frac{4(n-1)\pi I}{n}}\right))}_{=0} + \underbrace{\frac{1}{2}(1+1+1+\ldots+1)}_{=\frac{n}{2}} \\
&\quad + \underbrace{(\frac{1}{4}\left(1+e^{-\frac{4\pi I}{n}}+e^{-\frac{8\pi I}{n}}+\ldots+e^{-\frac{4(n-1)\pi I}{n}}\right))}_{=0} \\
&= \frac{n}{2};
\end{aligned}$$

$$\begin{aligned}
nV(Y) &= \sum_{i=0}^{n-1} \sin^2(\frac{2i\pi}{n}) \\
&= \underbrace{(-\frac{1}{4}\left(1+e^{\frac{4\pi I}{n}}+e^{\frac{8\pi I}{n}}+\ldots+e^{\frac{4(n-1)\pi I}{n}}\right))}_{=0} + \underbrace{\frac{1}{2}(1+1+1+\ldots+1)}_{=\frac{n}{2}} \\
&\quad - \underbrace{(\frac{1}{4}\left(1+e^{-\frac{4\pi I}{n}}+e^{-\frac{8\pi I}{n}}+\ldots+e^{-\frac{4(n-1)\pi I}{n}}\right))}_{=0} \\
&= \frac{n}{2}.
\end{aligned}$$

3.

$$\begin{aligned}
nCov(X,Y) &= \sum_{i=0}^{n-1} \cos(\frac{2i\pi}{n})\sin(\frac{2i\pi}{n}) \\
&= \underbrace{(\frac{1}{4I}\left(1+e^{\frac{4\pi I}{n}}+e^{\frac{8\pi I}{n}}+\ldots+e^{\frac{4(n-1)\pi I}{n}}\right))}_{=0} - \underbrace{(\frac{1}{4I}\left(1+e^{-\frac{4\pi I}{n}}+e^{-\frac{8\pi I}{n}}+\ldots+e^{-\frac{4(n-1)\pi I}{n}}\right))}_{=0} \\
&= 0.
\end{aligned}$$

4. $y=0$.
5. $\rho(X,Y)=0$.
   On ne peut imaginer pire ajustement.

Exercice 20 [9, problème 14, p. 39] Soit X une série statistique ordonnée par valeurs croissantes $X=(x_1,x_2,\ldots,x_n)$ avec $x_1 \leq x_2 \leq \ldots \leq x_n$.

1. Déterminer l'ensemble des médianes en distinguant le cas où n est impair et le cas où n est pair.
2. Sous quelles conditions la médiane est-elle unique ?
3. Étant donné un réel x_0, on considère l'écart moyen autour de x_0, défini par :
   $$e_{x_0} = \frac{1}{n}\sum_{i=1}^{n}|x_i - x_0|.$$
   Montrer que e_{x_0} est minimum quand x_0 est une médiane.
   Examiner d'abord le cas où $n=2$ et en déduire l'étude du cas général en distinguant le cas où n est impair et le cas où n est pair.
4. Soit μ une médiane quelconque. Montrer que, lorsque la médiane n'est pas unique, e_μ ne dépend pas de μ. On appellera écart médian la valeur de e_μ.
5. Montrer que l'écart médian est inférieur ou égal à l'écart moyen.

Exercice 21 [9, problème 16, p. 41] Soit (x_i,n_i), $\forall i \in [[1,p]]$, une distribution statistique telle que $0 < x_1 < x_2 < \ldots < x_p$. On suppose $p \in [[2,\infty[[$.

On pose $n=\sum_{i=1}^{p} n_i$ et $f_i = \frac{n_i}{n}$, $\forall i \in [[1,p]]$ et on note m la moyenne de cette distribution.

Soit ϕ une application de $\mathbb{R}_+^*$ dans $\mathbb{R}$ continue et strictement monotone.

1. Montrer qu'il existe un réel M_ϕ et un seul quel que $\phi(M_\phi)=\sum_{i=1}^{p} f_i\phi(x_i)$.
   M_ϕ est appelé la ϕ-moyenne de la distribution statistique.
2. On suppose que ϕ admet sur $\mathbb{R}_+^*$ une dérivée seconde strictement positive. Montrer que $\phi(M_\phi) > \phi(m)$ (on étudiera d'abord le cas où $p=2$ et on généralisera par récurrence).
   En déduire que
   – si ϕ est croissante, alors $M_\phi > m$;
   – si ϕ est décroissante, alors $M_\phi < m$.
3. Comparer M_ϕ et m dans le cas où ϕ admet sur $\mathbb{R}_+^*$ une dérivée seconde strictement négative.
4. Soit q la moyenne quadratique définie par $q=\sqrt{\sum_{i=1}^{p} f_i x_i^2}$. Montrer que $m<q$. En déduire que pour une distribution statistique quelconque ayant au moins deux valeurs distinctes, l'écart moyen est strictement inférieur à l'écart type.

5. Soit g la moyenne géométrique définie par $g = \prod_{i=1}^{p} x_i^{f_i}$. Montrer que $g < m$.
6. Soit h la moyenne harmonique définie par $\frac{1}{h} = \sum_{i=1}^{p} \frac{f_i}{x_i}$. Montrer que $h < g$.
7. Soit $\alpha \in \mathbb{R}^*$ et a_α la moyenne d'ordre α définie par

$$a_\alpha = \left(\sum_{i=1}^{p} f_i x_i^\alpha \right)^{\frac{1}{\alpha}}.$$

   (a) Montrer que $\lim_{\alpha\to\infty} a_\alpha = x_p$; $\lim_{\alpha\to-\infty} a_\alpha = x_1$; et $\lim_{\alpha\to 0} a_\alpha = g$.

   (b) On pose $a_0 = g$. Montrer que l'application $a \ : \ \mathbb{R} \longrightarrow \mathbb{R} \ ; \ \alpha \longmapsto a_\alpha$ est strictement croissante sur $\mathbb{R}$. En déduire à nouveau les inégalités

$$x_1 < h < g < m < q < x_p.$$

$\boxed{\textbf{Exercice 22}}$ [9, exercice 8, p. 60] Pour étudier la liaison entre deux caractères qualitatifs A et B, on considère les variables x et y définies comme suit. Pour chaque individu, on pose $x = 1$ s'il possède la caractère A, $x = 0$ dans l'autre cas ; $y = 1$ s'il possède la caractère B, $y = 0$ dans l'autre cas.

Sur un ensemble de n individus, on a obtenu la distributionn statistique

$X \quad Y$	1	0
1	a	b
0	c	d

avec $a + b + c + d = n$.

1. Montrer que le coefficient de corrélation linéaire entre x et y est le suivant

$$r = \frac{ad - bc}{\sqrt{(a+b)(c+d)(a+c)(b+d)}}.$$

2. Que peut-on dire des caractères A et B lorsque $r = 0$, $r = 1$ et $r = -1$?

$\boxed{\textbf{Exercice 23}}$ [9, problème 19, p. 70] Soit $X_1, X_2, \ldots, X_p$ et Y des vecteurs colonne de $\mathbb{R}^n$. On se propose d'ajuster sur ces données une relation linéaire $y = a_1x_1 + a_2x_2 + \ldots + a_px_p$.

Pour cela, on adopte le critère des moindres carrés, c'est-à-dire que l'on cherche les réels $a_1, a_2, \ldots, a_p$ qui rendent minimum

$$S = \sum_{k=1}^{n} (y_k - a_1x_{1,k} - a_2x_{2,k} - \ldots - a_px_{p,k})^2.$$

On notera X la matrice $(X_1, X_2, \ldots, X_p)$ à p colonnes et n lignes, X^t sa transposée et A le vecteur colonne de $\mathbb{R}^p$ de composantes $a_1, a_2, \ldots, a_p$.

On suppose que les vecteurs $X_1, X_2, \ldots, X_p$ sont linéairement indépendants.

1. Montrer que la matrice X^tX est inversible.
2. Montrer que le critère des moindres carrés revient à déterminer le vecteur A de $\mathbb{R}^p$ qui minimise $||Y - XA||_2$.
3. En déduire qu'il existe une solution et une seule définie par $A = (X^tX)^{-1}X^tY$.
4. Montrer que l'on a, après ajustement,

$$||Y||_2^2 = ||XA||_2^2 + ||Y - XA||_2^2.$$

   En déduire que $q = \frac{||XA||_2^2}{||Y||_2^2}$ est un indice de qualité de l'ajustement compris entre 0 et 1. Quelle est la signification de $q = 1$?

5. Supposons que $p = 2$ et que $x_{2,k} = 1, \ \forall k \in [[1, n]]$.

   Retrouver en calculant a_1 et a_2 les résultats concernant l'ajustement d'une relation affine $y = a_1x_1 + a_2$.

3 Espaces probabilisés

Au XVIII$^{\text{ème}}$ siècle, le calcul des probabilités est issu des jeux de hasard (dés, cartes, ...) et se veut être un moyen de modélisation. Aujourd'hui, il est introduit dans toutes les branches de l'activité scientifique (en économétrie, en génétique pour les lois de Mendel, en physique et plus particulièrement en physique nucléaire, en informatique comme outil de modélisation, ...).

On peut certainement lier le vocabulaire propre aux statistiques à celui des probabilités comme ceci :

$$\begin{aligned} POPULATION &\longleftrightarrow REFERENTIEL \\ FREQUENCE\ RELATIVE &\longleftrightarrow PROBABILITE \\ CARACTERE &\longleftrightarrow EVENEMENT \end{aligned}$$

3.1 Expérience aléatoire

Dès qu'une expérience dépend du hasard (i.e. peut prétendre à plusieurs résultats observables différents), on est en présence d'une expérience aléatoire.

Comme exemple d'expérience aléatoire, on pourra citer le lancer d'une pièce de monnaie ou le tirage d'une boule dans un sac, ...

3.2 Evénements, événements élémentaires

Un événement est un résultat observable à l'issue d'une certaine expérience.

Dans le cadre de l'expérience aléatoire consistant à lancer une pièce de monnaie, un événement pourra être, par exemple, l'obtention de face, ...

On appelle événement élémentaire un événement qui ne peut être réalisé de plusieurs façons distinctes.

Dans le cas d'une expérience aléatoire consistant à lancer un dé à 6 faces, l'obtention d'un nombre pair n'est pas un événement élémentaire car cet événement peut être réalisé par l'obtention de 2, de 4 ou encore de 6, alors que l'obtention de 6 est un événement élémentaire.

3.3 Ensemble fondamental, référentiel

On appelle ensemble fondamental, l'ensemble formé de tous les événements élémentaires possibles à l'issue de l'expérience en question. Le référentiel est, lui, l'ensemble formé de tous les événements possibles à l'issue de l'expérience en question.

Pour une expérience aléatoire consistant à lancer un dé à 6 faces, l'ensemble fondamental est constitué de l'obtention de 1, de l'obtention de 2, de l'obtention de 3, de l'obtention de 4, de l'obtention de 5 et de l'obtention de 6.

3.4 Algèbre des événements

On considère deux événements A et B faisant partie du référentiel E.

3.4.1 Définitions

1. L'inclusion

   $A \subset B$ signifie que la réalisation de A entraîne celle de B.

2. L'événement certain

   Cet événement est noté E, il est toujours réalisé.

3. L'événement impossible

   Cet événement est noté Ø, il n'est jamais réalisé.

4. La réunion

   $A \bigcup B$ signifie que A ou B est réalisé.

5. L'intersection

   $A \bigcap B$ signifie que A et B sont réalisés.

6. La compatibilité

   Deux événements sont compatibles, i.e. peuvent être réalisés en même temps, si $A \bigcap B \neq \emptyset$.

7. L'incompatibilité

   Deux événements sont incompatibles, i.e. ne peuvent être réalisés en même temps, si $A \bigcap B = \emptyset$.

8. L'absorption

   Si $B \subset A$, alors $A \bigcap B = B$.

9. La complémentarité

   On appelle complémentaire de A l'événement contraire de A, noté $\overline{A}$, tel que $A \bigcup \overline{A} = E$ et que $A \bigcap \overline{A} = \emptyset$.

3.4.2 Propriétés diverses

On considère trois événements A, B et C du référentiel E.

1. $\overline{\overline{A}} = A$.
2. $A \bigcup B = B \bigcup A$, commutativité.
3. $A \bigcup (B \bigcup C) = (A \bigcup B) \bigcup C$, associativité.
4. $A \bigcup (B \bigcap C) = (A \bigcup B) \bigcap (A \bigcup C)$, distributivité.
5. $A \bigcap B = B \bigcap A$, commutativité.
6. $A \bigcap (B \bigcap C) = (A \bigcap B) \bigcap C$, associativité.
7. $A \bigcap (B \bigcup C) = (A \bigcap B) \bigcup (A \bigcap C)$, distributivité.
8. $\overline{A \bigcup B} = \overline{A} \bigcap \overline{B}$.
9. $\overline{A \bigcap B} = \overline{A} \bigcup \overline{B}$.

Exercice 24 [13] On se donne A, B et C trois événements.

Exprimer

1. A seul se produit (Réponse : $A \bigcap \overline{B} \bigcap \overline{C}$)
2. A et B sont réalisés, mais pas C (Réponse : $A \bigcap B \bigcap \overline{C}$)
3. Les trois se produisent en même temps (Réponse : $A \bigcap B \bigcap C$)
4. Au moins l'un des trois se produit (Réponse : $A \bigcup B \bigcup C$)
5. Aucun des trois ne se produit (Réponse : $\overline{A \bigcup B \bigcup C} = \overline{A} \bigcap \overline{B} \bigcap \overline{C}$)
6. Au moins deux se produisent (Réponse : $(A \bigcap B) \bigcup (A \bigcap C) \bigcup (B \bigcap C))$
7. Un et un seul se produit (Réponse : $(A \bigcap \overline{B} \bigcap \overline{C}) \bigcup (B \bigcap \overline{A} \bigcap \overline{C}) \bigcup (C \bigcap \overline{A} \bigcap \overline{B}))$
8. A ou B est réalisé, mais pas C (Réponse : $(A \bigcap \overline{C}) \bigcup (B \bigcap \overline{C}))$
9. A ou B seulement (ou exclusif) est réalisé, en même temps que C (Réponse : $(A \bigcap \overline{B} \bigcap C) \bigcup (B \bigcap \overline{A} \bigcap C))$

3.5 Probabilité

3.5.1 Le schéma d'urne

Ce schéma est un exemple introductif, il ne permet en aucun cas de définir le mot *probabilité*.

On parle de schéma d'urne pour une expérience consistant à tirer itérativement une boule dans une urne après remise de celle-ci dans l'urne (on abrégera par tirage avec remise), cette urne contenant initialement n boules blanches et p boules noires (indiscernables). On définit alors la probabilité de tirer une boule blanche dans l'urne par la valeur $n/(n+p)$, et celle de tirer une boule noire par la valeur $p/(n+p)$. Ces valeurs représentent en fait la proportion de boules du type donné dans l'urne.

3.5.2 La loi des grands nombres

On ne décrit pas ici la loi des grands nombres, il s'agit simplement d'en donner une interprétation en se basant sur l'exemple introductif du scéma de type urne.

Si dans une épreuve de type schéma d'urne, la probabilité d'un événement est p, et si l'on répète l'expérience un grand nombre de fois dans des conditions identiques, alors la fréquence f de l'événement tend à se rapprocher de plus en plus de p.

Une chose essentielle découle de cette loi : si on considère une urne dont on ne connaît pas la nature de la population, il est possible de la déduire à partir d'un grand nombre d'expériences suivant le schéma d'urne.

3.5.3 Calcul de la probabilité d'un événement

L'exemple introductif illustre aussi la formule de Laplace ...

Découlant directement des axiomes du calcul des probabilités que nous verrons ultérieurement, on a :

$$P(A) = \frac{\text{nombre de cas favorables à } A}{\text{nombre de cas possibles}}$$

Cette relation est appelée **formule de Laplace** (elle est supposée connue) et s'applique dans le cas où les événements élémentaires sont équiprobables (i.e. de même probabilité).

Les événements sont équiprobables lorsque, par exemple, le dé est non pipé, la pièce est équilibrée, les cartes sont indiscernables, ...

Exercice 25 [13] On tire au hasard une carte dans un jeu de 52 cartes.

Donner les probabilités des événements suivants :

1. A On a tiré une carte rouge (Réponse : $P(A) = \frac{1}{2}$)
2. B On a tiré une carte de carreau (Réponse : $P(B) = \frac{1}{4}$)
3. C On a tiré la dame de pique (Réponse : $P(C) = \frac{1}{52}$)
4. D On a tiré un as (Réponse : $P(D) = \frac{1}{13}$)
5. E On a tiré un valet ou un roi (Réponse : $P(E) = \frac{2}{13}$)

Ainsi que celles de $A\bigcap B$, $A\bigcap C$, $A\bigcap D$, $A\bigcap E$, puis celles de $A\bigcup B$, $A\bigcup C$, $A\bigcup D$ et enfin celle de $A\bigcup E$.

(Réponses : $P(A\bigcap B) = \frac{1}{4}, P(A\bigcap C) = 0, P(A\bigcap D) = \frac{1}{26}, P(A\bigcap E) = \frac{1}{13}, P(A\bigcup B) = \frac{1}{2}, P(A\bigcup C) = \frac{27}{52}, P(A\bigcup D) = \frac{7}{13}, P(A\bigcup E) = \frac{15}{26}$)

Exercice 26 On tire 4 cartes sans remise et consécutivement dans un jeu de 32 cartes. On note A, B, C, D et E les événements suivants :

1. A La première est rouge, les deux dernières ne sont pas des dames.

   Réponse :

$première\ carte$	$seconde\ carte$	$troisième\ carte$	$quatrième\ carte$	$cas\ favorables$
$D,\ R$	D	$non\ D$	$non\ D$	$2 \times 3 \times 28 \times 27$
$D,\ R$	$non\ D$	$non\ D$	$non\ D$	$2 \times 28 \times 27 \times 26$
$non\ D,\ R$	D	$non\ D$	$non\ D$	$14 \times 4 \times 27 \times 26$
$non\ D,\ R$	$non\ D$	$non\ D$	$non\ D$	$14 \times 27 \times 26 \times 25$

   $P(A) = \frac{328860}{863040}$.

2. B Les trois premières sont des as et la dernière est un carreau.

   Réponse :

$première\ carte$	$seconde\ carte$	$troisième\ carte$	$quatrième\ carte$	$cas\ favorables$
$1,\ \diamondsuit$	$1,\ non\ \diamondsuit$	$1,\ non\ \diamondsuit$	$\diamondsuit$	$1 \times 3 \times 2 \times 7$
$1,\ non\ \diamondsuit$	$1,\ \diamondsuit$	$1,\ non\ \diamondsuit$	$\diamondsuit$	$3 \times 1 \times 2 \times 7$
$1,\ non\ \diamondsuit$	$1,\ non\ \diamondsuit$	$1,\ \diamondsuit$	$\diamondsuit$	$3 \times 2 \times 1 \times 7$
$1,\ non\ \diamondsuit$	$1,\ non\ \diamondsuit$	$1,\ non\ \diamondsuit$	$\diamondsuit$	$3 \times 2 \times 1 \times 8$

   $P(B) = \frac{174}{863040}$.

3. C L'une au moins des deux premières est de pique, aucune n'est de coeur.

   Réponse :

$première\ carte$	$seconde\ carte$	$troisième\ carte$	$quatrième\ carte$	$cas\ favorables$
$\spadesuit$	$non\ \heartsuit$	$non\ \heartsuit$	$non\ \heartsuit$	$8 \times 23 \times 22 \times 21$
$\diamondsuit\ ou\ \clubsuit$	$\spadesuit$	$non\ \heartsuit$	$non\ \heartsuit$	$16 \times 8 \times 22 \times 21$

   $P(C) = \frac{144144}{863040}$.

4. D Une seule des trois premières est un trêfle, la dernière est un as de trêfle.

   Réponse :

$quatrième\ carte$	$première\ carte$	$seconde\ carte$	$troisième\ carte$	$cas\ favorables$
$1,\ \clubsuit$	$\clubsuit$	$non\ \clubsuit$	$non\ \clubsuit$	$1 \times 7 \times 24 \times 23$
$1,\ \clubsuit$	$non\ \clubsuit$	$\clubsuit$	$non\ \clubsuit$	$1 \times 24 \times 7 \times 23$
$1,\ \clubsuit$	$non\ \clubsuit$	$non\ \clubsuit$	$\clubsuit$	$1 \times 24 \times 23 \times 7$

   $P(D) = \frac{11592}{863040}$.

5. E Une seule des trois premières est un trêfle, la dernière est un as.

   Réponse :

$quatrième\ carte$	$première\ carte$	$seconde\ carte$	$troisième\ carte$	$cas\ favorables$
$1,\ \clubsuit$	$\clubsuit$	$non\ \clubsuit$	$non\ \clubsuit$	$1 \times 7 \times 24 \times 23$
$1,\ \clubsuit$	$non\ \clubsuit$	$\clubsuit$	$non\ \clubsuit$	$1 \times 24 \times 7 \times 23$
$1,\ \clubsuit$	$non\ \clubsuit$	$non\ \clubsuit$	$\clubsuit$	$1 \times 24 \times 23 \times 7$
$1,\ non\ \clubsuit$	$\clubsuit$	$non\ \clubsuit$	$non\ \clubsuit$	$3 \times 8 \times 23 \times 22$
$1,\ non\ \clubsuit$	$non\ \clubsuit$	$\clubsuit$	$non\ \clubsuit$	$3 \times 23 \times 8 \times 22$
$1,\ non\ \clubsuit$	$non\ \clubsuit$	$non\ \clubsuit$	$\clubsuit$	$3 \times 23 \times 22 \times 8$

$P(D) = \frac{48024}{863040}$.

Exercice 27 On tire 4 cartes sans remise dans un jeu de 32 cartes.

On note A, B, et C les événements suivants :

1. A Les 4 cartes sont rouges

   (Réponse :

première carte	*seconde carte*	*troisième carte*	*quatrième carte*	*cas favorables*
R	R	R	R	$16 \times 15 \times 14 \times 13$

   $P(A) = \frac{16\times15\times14\times13}{32\times31\times30\times29} = \frac{43680}{863040}$).

2. B Une seule de ces 4 cartes est un as

   (Réponse :

première carte	*seconde carte*	*troisième carte*	*quatrième carte*	*cas favorables*
1	*non* 1	*non* 1	*non* 1	$4 \times 28 \times 27 \times 26$
non 1	1	*non* 1	*non* 1	$28 \times 4 \times 27 \times 26$
non 1	*non* 1	1	*non* 1	$28 \times 27 \times 4 \times 26$
non 1	*non* 1	*non* 1	1	$28 \times 27 \times 26 \times 4$

   $P(B) = \frac{4\times(4\times28\times27\times26)}{32\times31\times30\times29} = \frac{314496}{863040}$).

3. C L'une de ces 4 cartes au moins est un carreau

   (Réponse :

première carte	*seconde carte*	*troisième carte*	*quatrième carte*	*cas favorables*
$\diamondsuit$	?	?	?	$8 \times 31 \times 30 \times 29$
non $\diamondsuit$	$\diamondsuit$	?	?	$24 \times 8 \times 30 \times 29$
non $\diamondsuit$	*non* $\diamondsuit$	$\diamondsuit$	?	$24 \times 23 \times 8 \times 29$
non $\diamondsuit$	*non* $\diamondsuit$	*non* $\diamondsuit$	$\diamondsuit$	$24 \times 23 \times 22 \times 8$

   $P(C) = \frac{608016}{863040}$).

4. Calculer $P(A \bigcap C)$.

   Réponse :

première carte	*seconde carte*	*troisième carte*	*quatrième carte*	*cas favorables*
$\diamondsuit$	R	R	R	$8 \times 15 \times 14 \times 13$
$\heartsuit$	$\diamondsuit$	R	R	$8 \times 8 \times 14 \times 13$
$\heartsuit$	$\heartsuit$	$\diamondsuit$	R	$8 \times 7 \times 8 \times 13$
$\heartsuit$	$\heartsuit$	$\heartsuit$	$\diamondsuit$	$8 \times 7 \times 6 \times 8$

   Donc $P(A \bigcap C) = \frac{42000}{863040}$.

5. Calculer $P(A \bigcup C)$.

   Réponse :

première carte	*seconde carte*	*troisième carte*	*quatrième carte*	*cas favorables*
$\diamondsuit$	?	?	?	$8 \times 31 \times 30 \times 29$
non $\diamondsuit$	$\diamondsuit$	?	?	$24 \times 8 \times 30 \times 29$
non $\diamondsuit$	*non* $\diamondsuit$	$\diamondsuit$	?	$24 \times 23 \times 8 \times 29$
non $\diamondsuit$	*non* $\diamondsuit$	*non* $\diamondsuit$	$\diamondsuit$	$24 \times 23 \times 22 \times 8$
$\heartsuit$	$\heartsuit$	$\heartsuit$	$\heartsuit$	$8 \times 7 \times 6 \times 5$

Donc $P(A \bigcup C) = \frac{609696}{863040}$.

6. Calculer $P(A \bigcap B)$.

Réponse :

première carte	*seconde carte*	*troisième carte*	*quatrième carte*	*cas favorables*
$1,\ R$	*non* $1,\ R$	*non* $1,\ R$	*non* $1,\ R$	$2 \times 14 \times 13 \times 12$
non $1,\ R$	$1,\ R$	*non* $1,\ R$	*non* $1,\ R$	$14 \times 2 \times 13 \times 12$
non $1,\ R$	*non* $1,\ R$	$1,\ R$	*non* $1,\ R$	$14 \times 13 \times 2 \times 12$
non $1,\ R$	*non* $1,\ R$	*non* $1,\ R$	$1,\ R$	$14 \times 13 \times 12 \times 2$

Donc $P(A \bigcap B) = \frac{17472}{863040}$.

7. Calculer $P(A \bigcup B)$.

Réponse :

première carte	*seconde carte*	*troisième carte*	*quatrième carte*	*cas favorables*
$1,\ R$	$1,\ R$	*non* $1,\ R$	*non* $1,\ R$	$2 \times 1 \times 14 \times 13$
$1,\ R$	*non* $1,\ R$	$1,\ R$	*non* $1,\ R$	$2 \times 14 \times 1 \times 13$
$1,\ R$	*non* $1,\ R$	*non* $1,\ R$	$1,\ R$	$2 \times 14 \times 13 \times 1$
non $1,\ R$	$1,\ R$	$1,\ R$	*non* $1,\ R$	$14 \times 2 \times 1 \times 13$
non $1,\ R$	$1,\ R$	*non* $1,\ R$	$1,\ R$	$14 \times 2 \times 13 \times 1$
non $1,\ R$	*non* $1,\ R$	$1,\ R$	$1,\ R$	$14 \times 13 \times 2 \times 1$
$1,\ R$	*non* 1	*non* 1	*non* 1	$2 \times 28 \times 27 \times 26$
non 1	$1,\ R$	*non* 1	*non* 1	$28 \times 2 \times 27 \times 26$
non 1	*non* 1	$1,\ R$	*non* 1	$28 \times 27 \times 2 \times 26$
non 1	*non* 1	*non* 1	$1,\ R$	$28 \times 27 \times 26 \times 2$
non $1,\ R$	*non* $1,\ R$	*non* $1,\ R$	*non* $1,\ R$	$14 \times 13 \times 12 \times 11$
$1,\ N$	*non* 1	*non* 1	*non* 1	$2 \times 28 \times 27 \times 26$
non 1	$1,\ N$	*non* 1	*non* 1	$28 \times 2 \times 27 \times 26$
non 1	*non* 1	$1,\ N$	*non* 1	$28 \times 27 \times 2 \times 26$
non 1	*non* 1	*non* 1	$1,\ N$	$28 \times 27 \times 26 \times 2$

Donc $P(A \bigcup B) = \frac{340704}{863040}$.

8. Calculer $P(B \bigcap C)$.

Réponse :

première carte	*seconde carte*	*troisième carte*	*quatrième carte*	*cas favorables*
$1,\ \diamondsuit$	$non\ 1$	$non\ 1$	$non\ 1$	$1 \times 28 \times 27 \times 26$
$1,\ non\ \diamondsuit$	$non\ 1,\ \diamondsuit$	$non\ 1$	$non\ 1$	$3 \times 7 \times 27 \times 26$
$1,\ non\ \diamondsuit$	$non\ 1,\ non\ \diamondsuit$	$non\ 1,\ \diamondsuit$	$non\ 1$	$3 \times 21 \times 7 \times 26$
$1,\ non\ \diamondsuit$	$non\ 1,\ non\ \diamondsuit$	$non\ 1,\ non\ \diamondsuit$	$non\ 1,\ \diamondsuit$	$3 \times 21 \times 20 \times 7$
$non\ 1,\ \diamondsuit$	1	$non\ 1$	$non\ 1$	$7 \times 4 \times 27 \times 26$
$non\ 1,\ \diamondsuit$	$non\ 1$	1	$non\ 1$	$7 \times 27 \times 4 \times 26$
$non\ 1,\ \diamondsuit$	$non\ 1$	$non\ 1$	1	$7 \times 27 \times 26 \times 4$
$non\ 1,\ non\ \diamondsuit$	$1,\ \diamondsuit$	$non\ 1$	$non\ 1$	$21 \times 1 \times 27 \times 26$
$non\ 1,\ non\ \diamondsuit$	$1,\ non\ \diamondsuit$	$non\ 1,\ \diamondsuit$	$non\ 1$	$21 \times 3 \times 7 \times 26$
$non\ 1,\ non\ \diamondsuit$	$1,\ non\ \diamondsuit$	$non\ 1,\ non\ \diamondsuit$	$non\ 1,\ \diamondsuit$	$21 \times 3 \times 20 \times 7$
$non\ 1,\ non\ \diamondsuit$	$non\ 1,\ \diamondsuit$	1	$non\ 1$	$21 \times 7 \times 4 \times 26$
$non\ 1,\ non\ \diamondsuit$	$non\ 1,\ \diamondsuit$	$non\ 1$	1	$21 \times 7 \times 26 \times 4$
$non\ 1,\ non\ \diamondsuit$	$non\ 1,\ non\ \diamondsuit$	$1,\ \diamondsuit$	$non\ 1$	$21 \times 20 \times 1 \times 6$
$non\ 1,\ non\ \diamondsuit$	$non\ 1,\ non\ \diamondsuit$	$1,\ non\ \diamondsuit$	$non\ 1,\ \diamondsuit$	$21 \times 20 \times 3 \times 7$
$non\ 1,\ non\ \diamondsuit$	$non\ 1,\ non\ \diamondsuit$	$non\ 1,\ \diamondsuit$	1	$21 \times 20 \times 7 \times 4$
$non\ 1,\ non\ \diamondsuit$	$non\ 1,\ non\ \diamondsuit$	$non\ 1,\ non\ \diamondsuit$	$1,\ \diamondsuit$	$21 \times 20 \times 19 \times 1$

Donc $P(B \bigcap C) = \frac{218736}{863040}$.

9. Calculer $P(B \bigcup C)$.

   Réponse :

première carte	*seconde carte*	*troisième carte*	*quatrième carte*	*cas favorables*
$1,\ \diamondsuit$	?	?	?	$1 \times 31 \times 30 \times 29$
$1,\ non\ \diamondsuit$	$1,\ \diamondsuit$	?	?	$3 \times 1 \times 30 \times 29$
$1,\ non\ \diamondsuit$	$1,\ non\ \diamondsuit$	$\diamondsuit$	?	$3 \times 2 \times 8 \times 29$
$1,\ non\ \diamondsuit$	$1,\ non\ \diamondsuit$	$non\ \diamondsuit$	$\diamondsuit$	$3 \times 2 \times 22 \times 8$
$1,\ non\ \diamondsuit$	$non\ 1,\ \diamondsuit$	?	?	$3 \times 7 \times 30 \times 29$
$1,\ non\ \diamondsuit$	$non\ 1,\ non\ \diamondsuit$	$1,\ \diamondsuit$	?	$3 \times 21 \times 1 \times 29$
$1,\ non\ \diamondsuit$	$non\ 1,\ non\ \diamondsuit$	$1,\ non\ \diamondsuit$	$\diamondsuit$	$3 \times 21 \times 2 \times 8$
$1,\ non\ \diamondsuit$	$non\ 1,\ non\ \diamondsuit$	$non\ 1,\ \diamondsuit$	?	$3 \times 21 \times 7 \times 29$
$1,\ non\ \diamondsuit$	$non\ 1,\ non\ \diamondsuit$	$non\ 1,\ non\ \diamondsuit$	$1,\ \diamondsuit$	$3 \times 21 \times 20 \times 1$
$1,\ non\ \diamondsuit$	$non\ 1,\ non\ \diamondsuit$	$non\ 1,\ non\ \diamondsuit$	$non\ 1,\ \diamondsuit$	$3 \times 21 \times 20 \times 7$
$1,\ non\ \diamondsuit$	$non\ 1,\ non\ \diamondsuit$	$non\ 1,\ non\ \diamondsuit$	$non\ 1,\ non\ \diamondsuit$	$3 \times 21 \times 20 \times 19$
$non\ 1,\ \diamondsuit$	?	?	?	$21 \times 31 \times 30 \times 29$
$non\ 1,\ non\ \diamondsuit$	$1,\ \diamondsuit$	?	?	$21 \times 1 \times 30 \times 29$
$non\ 1,\ non\ \diamondsuit$	$1,\ non\ \diamondsuit$	$1,\ \diamondsuit$	?	$21 \times 3 \times 1 \times 29$
$non\ 1,\ non\ \diamondsuit$	$1,\ non\ \diamondsuit$	$1,\ non\ \diamondsuit$	$\diamondsuit$	$21 \times 3 \times 2 \times 8$
$non\ 1,\ non\ \diamondsuit$	$1,\ non\ \diamondsuit$	$non\ 1,\ \diamondsuit$	?	$21 \times 3 \times 7 \times 29$
$non\ 1,\ non\ \diamondsuit$	$1,\ non\ \diamondsuit$	$non\ 1,\ non\ \diamondsuit$	$1,\ \diamondsuit$	$21 \times 3 \times 20 \times 1$
$non\ 1,\ non\ \diamondsuit$	$1,\ non\ \diamondsuit$	$non\ 1,\ non\ \diamondsuit$	$non\ 1,\ \diamondsuit$	$21 \times 3 \times 20 \times 7$
$non\ 1,\ non\ \diamondsuit$	$1,\ non\ \diamondsuit$	$non\ 1,\ non\ \diamondsuit$	$non\ 1,\ non\ \diamondsuit$	$21 \times 3 \times 20 \times 19$
$non\ 1,\ non\ \diamondsuit$	$non\ 1,\ \diamondsuit$	?	?	$21 \times 7 \times 30 \times 29$
$non\ 1,\ non\ \diamondsuit$	$non\ 1,\ non\ \diamondsuit$	$1,\ \diamondsuit$	?	$21 \times 20 \times 1 \times 29$
$non\ 1,\ non\ \diamondsuit$	$non\ 1,\ non\ \diamondsuit$	$1,\ non\ \diamondsuit$	$1,\ \diamondsuit$	$21 \times 20 \times 3 \times 1$
$non\ 1,\ non\ \diamondsuit$	$non\ 1,\ non\ \diamondsuit$	$1,\ non\ \diamondsuit$	$non\ 1,\ \diamondsuit$	$21 \times 20 \times 3 \times 7$
$non\ 1,\ non\ \diamondsuit$	$non\ 1,\ non\ \diamondsuit$	$1,\ non\ \diamondsuit$	$non\ 1,\ non\ \diamondsuit$	$21 \times 20 \times 3 \times 19$
$non\ 1,\ non\ \diamondsuit$	$non\ 1,\ non\ \diamondsuit$	$non\ 1,\ \diamondsuit$	?	$21 \times 20 \times 7 \times 29$
$non\ 1,\ non\ \diamondsuit$	$non\ 1,\ non\ \diamondsuit$	$non\ 1,\ non\ \diamondsuit$	$1,\ \diamondsuit$	$21 \times 20 \times 19 \times 1$
$non\ 1,\ non\ \diamondsuit$	$non\ 1,\ non\ \diamondsuit$	$non\ 1,\ non\ \diamondsuit$	$1,\ non\ \diamondsuit$	$21 \times 20 \times 19 \times 3$
$non\ 1,\ non\ \diamondsuit$	$non\ 1,\ non\ \diamondsuit$	$non\ 1,\ non\ \diamondsuit$	$non\ 1,\ \diamondsuit$	$21 \times 20 \times 19 \times 7$

Donc $P(B \bigcup C) = \frac{703776}{863040}$.

3.6 Les axiomes du calcul de probabilités

3.6.1 Enoncé

Soit E un ensemble fondamental composé d'un ensemble dénombrable d'événements élémentaires. Les événements élémentaires possibles sont $E_1, E_2, \ldots, E_n, \ldots$.
Soit $\mathcal{E}$ le référentiel associé à l'ensemble fondamental E (i.e. les événements $A_1, A_2, \ldots, A_n, \ldots$ de $\mathcal{E}$ peuvent s'écrire comme réunion dénombrable d'événements élémentaires $E_1, E_2, \ldots, E_n, \ldots$ de E).
On définit la fonction $P\ :\ \mathcal{E} \longrightarrow \mathbb{R}$ telle que $P(A_i)$ désigne la probabilité de l'événement A_i si les axiomes suivants sont vérifiés :

- $\forall A_i,\ 0 \leq P(A_i) \leq 1$.
- $P(E) = 1$, $E = \bigcup_i E_i$ est appelé événement certain.
- $\forall A_i,\ \forall A_j$ tels que $A_i \bigcap A_j = \emptyset$ (événements incompatibles), on a $P(A_i \bigcup A_j) = P(A_i) + P(A_j)$.

3.6.2 Conséquences et propriétés

1. Si $\emptyset$ désigne l'événement impossible, on a $P(\emptyset) = 0$.
   Démonstration $A\bigcap\emptyset{=}\emptyset{\Longrightarrow}\ P(A\bigcup\emptyset) = P(A) + P(\emptyset) = P(A)$
   $\Longrightarrow P(\emptyset) = 0$.
2. Si $\overline{A}$ est l'événement contraire de A, alors $P(\overline{A}) = 1 - P(A)$.
   Démonstration $A\bigcup\overline{A} = E$ et $A\bigcap\overline{A}=\emptyset$,
   donc $P(A\bigcup\overline{A}) = P(A) + P(\overline{A}) = P(E) = 1$, et donc $P(\overline{A}) = 1 - P(A)$.
3. Soient A et B quelconques, alors $P(A) = P(A\bigcap B) + P(A\bigcap\overline{B})$.
   Démonstration $(A\bigcap B)\bigcup(A\bigcap\overline{B}) = A$ et $(A\bigcap B)\bigcap(A\bigcap\overline{B})=\emptyset$,
   donc $P((A\bigcap B)\bigcup(A\bigcap\overline{B})) = P(A\bigcap B) + P(A\bigcap\overline{B}) = P(A)$.
4. Si $A \subset B$, on a $P(A) \leq P(B)$.
   Démonstration $A\bigcup(B\bigcap\overline{A}) = B$ car $A \subset B$ et $A\bigcap(B\bigcap\overline{A})=\emptyset$,
   donc $P(B) = P(A) + P(B\bigcap\overline{A})$ et puis $P(A) \leq P(B)$.

Exercice 28 [9, exercice 8, p. 90] Soit $(A_n)_{n\in[[1,\infty[[}$ une suite d'événements.
Montrer que

$$P(\bigcup_{n\geq 1} A_n) \leq \sum_{n\geq 1} P(A_n).$$

Exercice 29 [9, exercice 9, p. 90] Soient $A_1, A_2, \ldots, A_n$ n événements.
On considère les 2^n événements de la forme $E_1\bigcup E_2\bigcup\ldots\bigcup E_n$, où $E_i \in \{A_i, \overline{A_i}\}$.
Montrer que la somme de leurs probabilités vaut $2^n - 1$.

3.6.3 Lien entre la probabilité de l'union et celle de l'intersection de deux événements

Énoncé du théorème

$$\forall A,\ \forall B,\ P(A\bigcup B) = P(A) + P(B) - P(A\bigcap B)$$

Démonstration $A\bigcup(B\bigcap\overline{A}) = A\bigcup B$ et $A\bigcap(B\bigcap\overline{A})=\emptyset$,
donc $P(A\bigcup B) = P(A) + P(B\bigcap\overline{A})$.
Or $P(B) = P(B\bigcap A) + P(B\bigcap\overline{A})$, donc, directement,
$P(A\bigcup B) = P(A) + P(B) - P(A\bigcap B)$.

Remarque On pourra constater cette propriété sur les exercices 3.5.3.

Exercice 30 [13] Au cours d'un sondage, on obtient les informations suivantes : 35 pourcents des gens vont au cinéma C, 12 pourcents au musée M et 6 pourcents aux deux. Exprimer le pourcentage de gens

1. allant au cinéma ou au musée (Réponse : $P(C\bigcup M) = P(C) + P(M) - P(C\bigcap M) = 0,35 + 0,12 - 0,06$ $= 0,41$)
2. n'allant pas au cinéma (Réponse : $P(\overline{C}) = 1 - P(C) = 1 - 0,35 = 0,65$)
3. n'allant ni au cinéma ni au musée (Réponse : $P(\overline{C}\bigcap\overline{M}) = P(\overline{C\bigcup M} = 1 - P(C\bigcup M) = 1 - 0,41 = 0,59$)
4. allant au cinéma mais pas au musée (Réponse : $P(C\bigcap\overline{M}) = P(C) - P(C\bigcap M) = 0,35 - 0,06 = 0,29$)

3.6.4 Formule du crible ou formule de Poincaré

La formule du crible ou de Poincaré est une généralisation d'un théorème déjà établi. Elle lie les probabilités de l'union et celle des différentes intersections possibles dans le cas de n événements. Elle s'écrit :

$$P(\bigcup_{i=1}^{n} A_i) = \sum_i P(A_i) - \sum_{i<j} P(A_i \bigcap A_j) + \sum_{i<j<k} P(A_i \bigcap A_j \bigcap A_k) + \ldots + (-1)^{n-1} P(\bigcap_{i=1}^{n} A_i).$$

Démonstration

Cette formule se démontre par récurrence sur n.

- La propriété est vraie pour $n = 1$ et $n = 2$.
- Supposons la formule vraie pour $n-1$. On pose $A' = \bigcup_{i=1}^{n-1} A_i$.

  On a

$$P(\bigcup_{i=1}^{n} A_i) = P(A' \bigcup A_n) = P(A') + P(A_n) - P(A' \bigcap A_n).$$

  Notations : soit $I_k(n)$ une partie quelconque de $[[1,n]]$ de cardinal k, $S_0 = 1$ et

$$S_k = \sum_{I_k(n)} P(\bigcap_{i\in I_k(n)} A_i).$$

  Avec ces notations, on veut montrer que

$$P(\bigcup_{i=1}^{n} A_i) = \sum_{k=1}^{n} (-1)^{k+1} S_k.$$

  On a, par hypothèse de récurrence

$$P(A') = \sum_{k=1}^{n-1} (-1)^{k+1} \sum_{I_k(n-1)} P(\bigcap_{i\in I_k(n-1)} A_i).$$

  D'autre part

$$\begin{aligned} P(A' \bigcap A_n) &= P(\bigcup_{i=1}^{n-1} (A_i \bigcap A_n)) \\ &= \sum_{k=1}^{n-1} (-1)^{k+1} \sum_{I_k(n-1)} P((\bigcap_{i\in I_k(n-1)} A_i) \bigcap A_n) \\ &= \sum_{K=2}^{n} (-1)^{K} \sum_{I_{K-1}(n-1)} P(\bigcap_{i\in I_{K-1}(n-1)} (A_i \bigcap A_n)) \text{ en posant } K = k+1 \end{aligned}$$

  Donc, en considérant que $I_k(n)$ est une partie à k éléments de $[[1,n]]$ qui soit contient l'élément n, soit ne le contient pas,

$$\begin{aligned} P(\bigcup_{i=1}^{n} A_i) &= \sum_{i=1}^{n} P(A_i) + \sum_{k=2}^{n-1} (-1)^{k+1} \sum_{I_k(n)} P(\bigcap_{i\in I_k(n)} A_i) + (-1)^n P(\bigcap_{i=1}^{n} A_i) \\ &= \sum_{k=1}^{n} (-1)^{k+1} S_k. \end{aligned}$$

Exercice 31 [9, problème 18, p. 94] Soit $\mathcal{A}$ un référentiel. Pour A et B dans $\mathcal{A}$, on définit $A\Delta B = (A \bigcap \overline{B}) \bigcup (B \bigcap \overline{A})$.

1. Montrer que $(\mathcal{A}, \Delta, \bigcap)$ est un anneau commutatif.

2. A chaque A de $\mathcal{A}$, on associe sa fonction indicatrive ϕ_A définie par $\phi_A(\omega) = 1$ si $\omega \in A$ et $\phi_A(\omega) = 0$ si $\omega \notin A$.
Montrer que
$$\phi_{A_1 \Delta A_2 \Delta \ldots \Delta A_n} = \sum_{i=1}^{n} \phi_{A_i} \ mod\ 2.$$
3. Étant donnés deux événements A et B, montrer que
$$P(A \Delta B) = P(A) + P(B) - 2P(A \bigcap B).$$
4. Étant donnés n événements $A_1, A_2, \ldots, A_n$, on pose pour tout entier $k \in [[0, n]]$, $I_k(n)$ une partie quelconque de $[[1, n]]$ de cardinal k, $S_0 = 1$ et
$$S_k = \sum_{I_k(n)} P(\bigcap_{i \in I_k(n)} A_i).$$
Montrer par récurrence que
$$P(A_1 \Delta A_2 \Delta \ldots \Delta A_n) = \sum_{k=1}^{n} (-2)^{k+1} S_k.$$

3.7 Probabilité conditionnelle, indépendance

3.7.1 Probabilité conditionnelle

Définition Soient A et B deux événements compatibles et B de probabilité non nulle. La probabilité que A soit réalisé sachant que B est réalisé est appelée probabilité de A par rapport à B. Elle est notée $P(A/B)$ et vaut :
$$P(A/B) = \frac{P(A \bigcap B)}{P(B)}.$$

3.7.2 Propriété

Une probabilité conditionnelle est une probabilité.

Pour le montrer, il suffit de vérifier les trois axiomes définissant une probabilité.

Exercice 32 Soit $(A_k)_{k \in \{1,\ldots,n\}}$ une suite de n événements tels que $A_1 \bigcap A_2 \bigcap \ldots \bigcap A_n$ soit de probabilité non nulle.

Alors,
$$P(A_1 \bigcap A_2 \bigcap \ldots \bigcap A_n) = P(A_1) \times P(A_2/A_1) \times \ldots \times P(A_n/A_1 \bigcap A_2 \bigcap \ldots \bigcap A_{n-1}).$$
Formule des probabilités composées.

Exercice 33 [13] On note G l'événement naissance d'un garçon, F l'événement naissance d'une fille et L l'événement présenter une luxation congénitale de la hanche à la naissance.
On donne $P(G) = 0,52$, $P(F) = 0,48$, $P(L/G) = 0,01$ et $P(L/F) = 0,02$.

- Quelle est la probabilité qu'un nouveau-né présente cette luxation congénitale de la hanche ?
  Réponse : $(L \bigcap G) \bigcup (L \bigcap F) = L$ et $(L \bigcap G) \bigcap (L \bigcap F) = \emptyset$,
  donc $P(L) = P(L \bigcap G) + P(L \bigcap F)$
  $= P(G)P(L/G) + P(F)P(L/F) = 0,52 \times 0,01 + 0,48 \times 0,02 = 0,0148.$
- Quelle est la probabilité qu'un nouveau-né présentant cette luxation congénitale de la hanche soit une fille ?
  Réponse : $P(F \bigcap L) = P(L)P(F/L) = P(F)P(L/F)$, donc
  $P(F/L) = \dfrac{P(F)P(L/F)}{P(L)} = \dfrac{0,48 \times 0,02}{0,0148} = 0,6486.$

Exercice 34 [9, exercice 3, p. 105] Deux urnes A et B contiennent des boules numérotées de 1 à n. On tire au hasard une boule dans chaque urne. Quelle est la probabilité que $\frac{a}{b}$ soit un entier si a est le numéro de la boule tirée dans l'urne A et si b est le numéro de la boule tirée dans l'urne B.

Encadrer cette probabilité et en donner un équivalent quand n tend vers l'infini.

Exercice 35 [9, exercice 5, p. 106] Une urne contient a boules blanches et b boules noires. On tire successivement k boules, sans remise.

Quelle est la probabilité que la $k$ième boule tirée soit noire ?

Exercice 36 [9, exercice 6, p. 107] On considère n urnes numérotées de 1 à n. L'urne k contient n boules noires et k boules blanches.

On choisit au hasard une urne et dans cette urne, une boule.

Quelle est la probabilité d'obtenir une boule noire ?

Que vaut approximativement cette probabilité quand n est grand ?

Exercice 37 [5, exercice 4, p. 108–110] **Le tournoi**.

A, B et C jouent à pile ou face (pièce équilibrée non truquée).

– A l'étape 1, A et B jouent.
– Le gagnant de l'étape $n-1$ joue à l'étape n contre le non-joueur à l'étape $n-1$.

Dès que quelqu'un gagne deux parties de suite, il est déclaré vainqueur et la partie s'arrête.

Quelles sont les probabilités de chacun de gagner ?

Exercice 38 [9, exercice 22.1, p. 119] U contient a boules blanches et b boules noires. V contient b boules blanches et a boules noires.

On effecctue des tirages avec remises. On désigne par U_n l'évenement "le $n$ième tirage s'effectue dans U", par V_n l'évenement "le $n$ième tirage s'effectue dans V" et par B_n l'évenement "la $n$ième boule tirée est blanche".

Ainsi $P(B_n/U_n) = \frac{a}{a+b}$ et $P(B_n/V_n) = \frac{b}{a+b}$.

Règle : si à l'étape n on a tiré une boule blanche, le $n+1$ième tirage s'effectue dans U ; si à l'étape n on a tiré une boule noire, le $n+1$ième tirage s'effectue dans V.

1. Donner une relation entre $P(B_{n+1})$ et $P(B_n)$.
2. En déduire $P(B_n)$ en fonction de $P(B_1)$.
3. Quelle est la limite de $P(B_n)$ lorsque n tend vers ∞ ?

3.7.3 Indépendance en probabilité

On dit que A et B sont indépendants si $P(A/B) = P(A)$.

On pourra remarquer que $P(A/B) = P(A) \Longleftrightarrow P(A \bigcap B) = P(A)P(B)$.

Propriété

Soient A et B deux événements indépendants. Alors :

– A et $\overline{B}$ sont indépendants

Démonstration $(A \bigcap B) \bigcup (A \bigcap \overline{B}) = A$ et $(A \bigcap B) \bigcap (A \bigcap \overline{B}) = \emptyset$.

Donc, $P(A) = P(A \bigcap B) + P(A \bigcap \overline{B}) = P(A)P(B) + P(A \bigcap \overline{B})$.

Donc $P(A\bigcap\overline{B}) = P(A) - P(A)P(B) = P(A)(1-P(B)) = P(A)P(\overline{B})$, ce qui montre que A et $\overline{B}$ sont indépendants.

– B et $\overline{A}$ sont indépendants

Démonstration La démonstration est la même que précédemment en interchangeant les rôles de A et B.

– $\overline{A}$ et $\overline{B}$ sont indépendants

Démonstration On a montré que A et $\overline{B}$ sont indépendants, d'après le premier point, donc, d'après le second point, $\overline{A}$ et $\overline{B}$ sont indépendants.

3.7.4 Le théorème de Bayes

Soit un référentiel E, et soient $E_1, E_2, \dots, E_n$ formant une partition de E.

– $\forall i$ tel que $E_i \subset E$, $E_i \neq \emptyset$.
– $\forall (i,j)$ tel que $i \neq j$, $E_i \bigcap E_j = \emptyset$ (incompatibles)
– $\bigcup_{i=1}^n E_i = E$

Alors, $\forall A \subset E,\ P(A) = \sum_{j=1}^n P(E_j)P(A/E_j)$ (formule dite des probabilités totales),

et $\forall i, P(E_i/A) = \dfrac{P(E_i)P(A/E_i)}{\sum_{j=1}^n P(E_j)P(A/E_j)}$.

Démonstration $\forall A \subset E,\ A = A\bigcap E = A\bigcap(\bigcup_{i=1}^n E_i) = \bigcup_{i=1}^n (A\bigcap E_i)$.

Si $i \neq j$, $A\bigcap E_i$ et $A\bigcap E_j$ sont incompatibles car E_i et E_j le sont.

Donc, $P(A) = \sum_{j=1}^n P(A\bigcap E_j) = \sum_{j=1}^n P(E_j)P(A/E_j)$.

De plus $P(E_i/A) = \dfrac{P(E_i\bigcap A)}{P(A)} = \dfrac{P(E_i\bigcap A)}{\sum_{j=1}^n P(E_j)P(A/E_j)} = \dfrac{P(E_i)P(A/E_i)}{\sum_{j=1}^n P(E_j)P(A/E_j)}$.

Exercice 39 [13] Soit une voiture M de marque X. Cette voiture démarre mal. Cependant, le garagiste sait que parmi les véhicules qui démarrent mal, la probabilité pour que la cause en soit le démarreur est $P(A) = 0,5$, que la cause en soit la batterie est $P(B) = 0,4$, puis que la cause en soit les bougies est $P(C) = 0,1$. De plus, sur le nombre total de voitures démarrant mal à cause du démarreur, 10 pourcents sont de marque X, sur le nombre total de voitures démarrant mal à cause de la batterie, 20 pourcents sont de marque X, et sur le nombre total de voitures démarrant mal à cause des bougies, 5 pourcents sont de marque X. Sachant cela aussi, le garagiste se penchera-t-il d'abord sur le démarreur, la batterie ou les bougies ?

Solution : X est l'événement être de marque X.

$P(X/A) = 0,1; P(X/B) = 0,2; P(X/C) = 0,05$.

Calculons $P(A/X), P(B/X)$ et $P(C/X)$.

$$P(A/X) = \frac{0,5\times 0,1}{0,5\times 0,1 + 0,4\times 0,2 + 0,1\times 0,05} \approx 0,37.$$
$$P(B/X) = \frac{0,4\times 0,2}{0,5\times 0,1 + 0,4\times 0,2 + 0,1\times 0,05} \approx 0,59.$$
$$P(C/X) = \frac{0,1\times 0,05}{0,5\times 0,1 + 0,4\times 0,2 + 0,1\times 0,05} \approx 0,04.$$

Le garagiste vérifiera donc d'abord la batterie.

Exercice 40 [9, exercice 17, p. 112] On considère deux événements indépendants A et B.

Montrer que
$$\max(P(A\bigcap B), P(\overline{A}\bigcap\overline{B}), P(A\Delta B)) \geq \frac{4}{9}.$$

3.7.5 Indépendance mutuelle d'un nombre fini d'événements

Soient $E_1, E_2, \ldots, E_n$, n événements. On dit que tous ceux-ci sont mutuellement indépendants si

$$P(\bigcap_{i=1}^{n} E_i) = \prod_{i=1}^{n} P(E_i).$$

3.7.6 Indépendance deux à deux d'un nombre fini d'événements

Soient $E_1, E_2, \ldots, E_n$, n événements. On dit qu'ils sont indépendants deux à deux si $\forall (i,j)$ tel que $i \neq j$, on ait $P(E_i \bigcap E_j) = P(E_i)P(E_j)$.

3.7.7 Pas d'implication entre les notions d'indépendance mutuelle et deux à deux d'un nombre fini d'événements

Exemple 1 On lance un dé équilibré à huit faces.
On note A_1 l'événement obtenir 1,2,7 ou 8, A_2 l'événement obtenir 2,3,6 ou 8 et A_3 l'événement obtenir 3,4,5 ou 8.
Ainsi, $P(A_i) = 1/2$, $P(\bigcap_{i=1}^{3} A_i) = \prod_{i=1}^{3} P(A_i) = 1/8$.
On a :
$P(A_1 \bigcap A_2) = 1/4 = P(A_1)P(A_2)$,
$P(A_2 \bigcap A_3) = 1/4 = P(A_2)P(A_3)$,
mais,
$P(A_1 \bigcap A_3) = 1/8 \neq P(A_1)P(A_3) = 1/4$.

Exemple 2 On lance successivement 2 dés à 6 faces.
On note A_1 l'événement obtenir un nombre pair avec la premier dé, A_2 l'événement obtenir un nombre impair avec le second dé et A_3 l'événement obtenir deux dés de même parité.
Ainsi, $P(A_1) = 1/2$, $P(A_2) = 1/2$ et $P(A_3) = 1/2$.
On a :
$P(A_1 \bigcap A_2) = 1/4 = P(A_1)P(A_2)$,
$P(A_2 \bigcap A_3) = 1/4 = P(A_2)P(A_3)$,
et $P(A_1 \bigcap A_3) = 1/4 = P(A_1)P(A_3)$.
Mais, $P(\bigcap_{i=1}^{3} A_i) = P(\emptyset) = 0 \neq \prod_{i=1}^{3} P(A_i) = 1/8$.

Conclusion
Il n'y a aucune implication entre indépendance deux à deux et indépendance mutuelle.

3.7.8 Indépendance totale d'un nombre fini d'événements

Soient $E_1, E_2, \ldots, E_n$, n événements. On dit qu'ils sont totalement indépendants si $\forall I \subset \{1, 2, \ldots, n\}$, on a $P(\bigcap_{i \in I} E_i) = \prod_{i \in I} P(E_i)$.

3.8 Epreuves répétées

3.8.1 Epreuves répétées non exhaustives

Soit E un référentiel partitionné en deux événements notés A et $\overline{A}$. L'événement A est celui de tirer une boule blanche dans une urne parmi des blanches et des noires. Les probabilités de ces deux événements sont données par $P(A) = p$ et $P(\overline{A}) = q = 1 - p$.

On répète l'épreuve un certain nombre de fois sans modifier le référentiel. Pendant n épreuves, $P(A)$ est toujours égal à p. C'est le cas d'un tirage avec remise, ...

La probabilité que A soit réalsé exactement k fois, à l'issue de n épreuves, dans un ordre précis est

$$p^k q^{n-k}.$$

La probabilité que A soit réalsé exactement k fois, à l'issue de n épreuves, sans tenir compte de l'ordre est

$$C_n^k p^k q^{n-k}.$$

Nous aurons l'occasion de retrouver ce dernier résultat lorsque nous étudierons la loi binomiale.

3.8.2 Epreuves répétées exhaustives

Soit E un référentiel partitionné en deux événements notés A et $\overline{A}$. L'événement A est celui de tirer une boule blanche dans une urne parmi des blanches et des noires. Les probabilités de ces deux événements sont données par $P(A) = p$ et $P(\overline{A}) = q = 1 - p$.

L'urne contient au début N boules dont pN blanches et qN noires. On répète l'épreuve un certain nombre de fois en modifiant le référentiel. Pendant les n épreuves, $P(A)$ va varier. C'est le cas d'un tirage sans remise, ...

La probabilité que A soit réalsé exactement k fois, à l'issue de n épreuves, sans tenir compte de l'ordre est

$$\frac{C_{Np}^k C_{Nq}^{n-k}}{C_N^n}.$$

Nous aurons l'occasion de retrouver ce résultat lorsque nous étudierons la loi hypergéométrique.

Exercice 41 On considère 30 sujets de CAPES. 20 exactement contiennent de l'analyse numérique (AN) et 4 exactement, des probabilités (P). L'un des sujets contient à la fois de l'analyse numérique et des probabilités.

Dans ce lot de sujets,

1. on prend au hasard l'un d'eux. Quelle est la probabilité pour qu'il contienne soit de l'analyse numérique, soit des probabilités.
(Réponse : $P(AN \bigcup P) = P(AN) + P(P) - P(AN \bigcap P) = \frac{20+4-1}{30} = \frac{23}{30}$).
2. on prend au hasard 3 d'entre eux avec remise (A). Quelle est la probabilité pour qu'ils contiennent tous soit de l'analyse numérique, soit des probabilités.
(Réponse : $P(A) = P(AN \bigcup P)^3 = (\frac{23}{30})^3$).
3. on prend au hasard 3 d'entre eux sans remise (B). Quelle est la probabilité pour qu'ils contiennent tous soit de l'analyse numérique, soit des probabilités.
(Réponse : $P(B) = \frac{23\times22\times21}{30\times29\times28} = \frac{C_{23}^3 C_7^0}{C_{30}^3}$).

Exercice 42 [9, exercice 18, p. 115] On lance un dé n fois de suite.

Quelle est la probabilité que le produit des points obtenus soit un nombre pair ?

Exercice 43 On considère une urne contenant 100 jetons indiscernables numérotés de 1 à 100. N personnes (N est supposé supérieur ou égal à 2) viennent tirer un jeton au hasard et le remettent dans l'urne.

1. Quelle est la probabilité pour que 2 personnes au moins (parmi les N personnes) aient pris un numéro identique ?
2. À partir de quelle valeur de N, cette probabilité est-elle supérieure à $\frac{1}{2}$?

Exercice 44 Une urne contient 6 boules dont 3 blanches, 2 rouges et 1 noire. On tire 3 boules dans cette urne. On définit les événements suivants :

1. A : "Tirer une boule blanche, une rouge et une noire".
2. B : "Tirer au moins deux boules blanches".
3. C : "Tirer une boule blanche, puis une rouge et puis une noire".
4. D : "Tirer deux boules blanches, puis une rouge".
5. E : "Ne pas tirer de boule noire".
6. F : "Ne pas tirer plus d'une boule blanche".

On note P_1 la probabilité de cet événement lorsqu'il est effectué avec remise et P_2 lorsqu'il est effectué sans remise. Calculer $P_1(A), P_1(B), P_1(C), P_1(D), P_1(E), P_1(F), P_2(A), P_2(B), P_2(C), P_2(D), P_2(E)$ et $P_2(F)$.

Réponses :
$P_1(A) = 6 \times (\frac{1}{2} \times \frac{1}{3} \times \frac{1}{6}) = \frac{1}{6}$
$P_1(B) = \frac{1}{2}$
En effet,

première boule	*seconde boule*	*troisième boule*	*cas favorables*
blanche	*blanche*	?	$3 \times 3 \times 6$
blanche	*non blanche*	*blanche*	$3 \times 3 \times 3$
non blanche	*blanche*	*blanche*	$3 \times 3 \times 3$

et le nombre de cas possibles est 6^3.
$P_1(C) = \frac{1}{2} \times \frac{1}{3} \times \frac{1}{6} = \frac{1}{36}$
$P_1(D) = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{3} = \frac{1}{12}$
$P_1(E) = (\frac{5}{6})^3 = \frac{125}{216}$
$P_1(F) = (\frac{1}{2})^3 + 3 \times (\frac{1}{2})^3 = \frac{1}{2}$
$P_2(A) = 6 \times (\frac{3}{6} \times \frac{2}{5} \times \frac{1}{4}) = \frac{3}{10}$
$P_2(B) = \frac{1}{2}$
En effet,

première boule	*seconde boule*	*troisième boule*	*cas favorables*
blanche	*blanche*	?	$3 \times 2 \times 4$
blanche	*non blanche*	*blanche*	$3 \times 3 \times 2$
non blanche	*blanche*	*blanche*	$3 \times 3 \times 2$

et le nombre de cas possibles est $6 \times 5 \times 4$.
$P_2(C) = \frac{3}{6} \times \frac{2}{5} \times \frac{1}{4} = \frac{1}{20}$
$P_2(D) = \frac{3}{6} \times \frac{2}{5} \times \frac{2}{4} = \frac{1}{10}$
$P_2(E) = \frac{5}{6} \times \frac{4}{5} \times \frac{3}{4} = \frac{1}{2}$
$P_2(F) = \frac{3}{6} \times \frac{2}{5} \times \frac{1}{4} + 3 \times (\frac{3}{6} \times \frac{3}{5} \times \frac{2}{4}) = \frac{1}{2}$.

4 Variables aléatoires réelles

4.1 Notion de variable aléatoire réelle

Dès que l'on mesure une grandeur dont les valeurs réelles dépendent du hasard, on est en présence d'une variable aléatoire réelle.

4.2 Définition

Soit E un référentiel, et soit f une fonction ($f : E \longrightarrow \mathbb{R}$). On dit que f définit une **variable aléatoire réelle**. En d'autres termes, une variable aléatoire réelle est une fonction dont la valeur dépend du résultat d'une épreuve. A partir d'une même épreuve, on peut définir plusieurs variables aléatoires réelles.

4.3 Exemple

On lance une pièce de monnaie. On nomme A l'événement obtention de pile et B l'événement obtention de face.

On peut définir une première variable aléatoire par $f(A) = 1$, $f(B) = 1$, et dans ce cas $P(f = 1) = 1$.

Mais, on peut définir une seconde variable aléatoire par $f(A) = 1$, $f(B) = 0$, et dans ce cas $P(f = 1) = P(A) = p$ et $P(f = 0) = P(B) = q = 1 - p$.

4.4 Remarques

Usuellement, on confond la fonction définissant la variable aléatoire réelle et la variable aléatoire réelle, elle-même.

De la même façon que l'ensemble des événements constitue une algèbre, on admettra que les variables aléatoires réelles forment, elles aussi, une algèbre.

4.5 Convergence en probabilité d'une variable aléatoire

On dit que la suite de variables aléatoires réelles $(X_i)_i$ converge en probabilité vers le réel x si

$$\forall \varepsilon > 0, \lim_{n\to\infty} P(|X_n - x| > \varepsilon) = 0.$$

Exercice 45 [9, exercice 8.I, p. 240] Sur un espace probabilisé $(\Omega, \mathcal{A}, P)$, on considère deux suites $(X_n)_n$ et $(Y_n)_n$ de variables aléatoires.

On suppose que $X_n \xrightarrow{P} x$ et $Y_n \xrightarrow{P} y$

1. Montrer que $X_n + Y_n \xrightarrow{P} x + y$.
2. Montrer que si $(\lambda_n)_n$ est une suite de réels convergeant vers λ, alors $\lambda_n X_n \xrightarrow{P} \lambda x$.
3. Montrer que $X_n^2 \xrightarrow{P} x^2$.

5 Variables aléatoires réelles discrètes

5.1 Définition

Une variable aléatoire réelle est dite discrète si son domaine de variations est fini ou dénombrable.

5.1.1 Remarque

Si le système fondamental est lui-même fini ou dénombrable, la variable aléatoire sera automatiquement discrète.

5.1.2 Cas particulier d'un vecteur de variables aléatoires réelles discrètes

Un vecteur de variables aléatoires réelles discrètes est un vecteur dont toutes les composantes sont des variables aléatoires réelles discrètes.

5.2 Lois de probabilités

Soit une variable aléatoire réelle discrète X, $X : E \longrightarrow \mathbb{R}$ pouvant prendre les valeurs $x_1, x_2, \ldots, x_n, \ldots$

Par définition, la probabilité que X soit égale à x_i est la probabilité des éléments de E ayant pour image x_i par X. Elle est notée : $P(X = x_i)$.

La loi de probabilité d'une variable aléatoire réelle discrète X est définie par les valeurs de ces probabilités $P(X = x_i)$ qui doivent vérifier $\sum_i P(X = x_i) = 1$.

5.2.1 Exemple

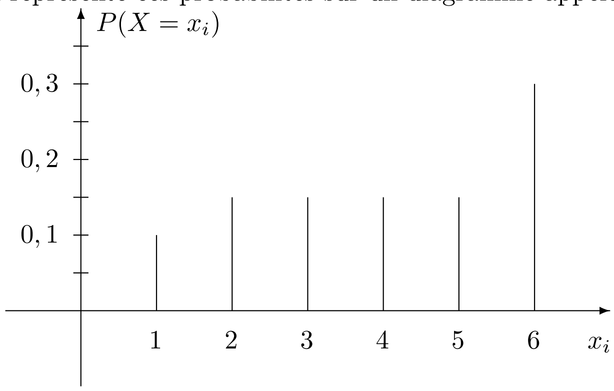
On considère le lancer d'un dé pipé dont les probabilités des 6 événements possibles sont données dans le tableau suivant :

Face supérieure	1	2	3	4	5	6
Probabilité	0,1	0,15	0,15	0,15	0,15	0,3

On est bien en présence d'une loi de probabilité discrète, car

$$\sum_i P(X = x_i) = 1.$$

On représente ces probabilités sur un diagramme appelé diagramme en bâtons.



5.2.2 Loi d'un vecteur de variables aléatoires réelles discrètes et loi marginale

Soit X un vecteur de variables aléatoires réelles discrètes, de composantes $X_i : E \longrightarrow \mathbb{R}$, $X \in E^d$ prend ses valeurs dans $\mathbb{R}^d$.

Par définition, la probabilité que $X = (X_1, X_2, \ldots, X_d)$ soit égale à $(x_{i_1}, x_{i_2}, \ldots, x_{i_d})$ est la probabilité des éléments de E^d ayant pour image $(x_{i_1}, x_{i_2}, \ldots, x_{i_d})$ par X. Elle est notée : $P(X = (x_{i_1}, x_{i_2}, \ldots, x_{i_d}))$.

La loi de probabilité d'une variable aléatoire réelle discrète X est définie par les valeurs de ces probabilités $P(X = (x_{i_1}, x_{i_2}, \ldots, x_{i_d}))$ qui doivent vérifier $\sum_{i_1, i_2, \ldots, i_d} P(X = (x_{i_1}, x_{i_2}, \ldots, x_{i_d})) = 1$.

Si on suppose donnée la loi de X, on appelle loi marginale de I la loi de X_I
(où $I = \{a_1, a_2, \ldots, a_r\} \subset K = [[1, d]]$, $J = \{b_1, b_2, \ldots, b_s\} \subset K$ tels que $I \bigcup J = K$, $I \bigcap J = \emptyset$, avec $r + s = d$) dont la probabilité $P(X_I = (x_{i_{a_1}}, x_{i_{a_2}}, \ldots, x_{i_{a_r}}))$ est donnée par la formule

$$P(X_I = (x_{i_{a_1}}, x_{i_{a_2}}, \ldots, x_{i_{a_r}})) = \sum_{i_{b_1}, i_{b_2}, \ldots, i_{b_s}} P(X = (x_{i_1}, x_{i_2}, \ldots, x_{i_d})).$$

Exercice 46 Soient X et Y deux variables alétoires discrètes prenant leurs valeurs dans $\mathbb{N}$. On suppose que pour

$p \in]0,1[$ et pour $\lambda \in \mathbb{R}_+$, la loi du couple (X,Y) est donnée par

$$P((X,Y)=(i,j)) = e^{-\lambda}\frac{(\lambda p)^i}{i!}\frac{(\lambda(1-p))^j}{j!}.$$

1. Déterminer les lois marginales de X et Y.
2. Montrer que X et Y sont indépendantes.

Exercice 47 [9, exercice 7, p. 134] On jette trois dés simultanément.
On désigne par X le minimum et par Y le maximum des trois scores obtenus.
Quelle est la loi de (X,Y) ? X et Y sont-elles indépendantes ?

5.2.3 Loi conditionnelle, indépendance de deux variables aléatoires réelles discrètes et indépendance de n variables aléatoires réelles discrètes

Soient X et Y deux variables aléatoires réelles discrètes, on peut donner la définition de la loi conditionnelle de la variable aléatoire réelle discrète $Y/X=k$ (on lit Y sachant que $X=k$) par la probabilité :

$$P((Y/X=k)=l) = \frac{P(Y=l\bigcap X=k)}{P(X=k)}.$$

Soient X et Y deux variables aléatoires réelles discrètes, on dit que X et Y sont deux variables aléatoires réelles discrètes indépendantes si

$$P(X=k\bigcap Y=l) = P(X=k)P(Y=l).$$

On peut définir l'indépendance deux à deux, mutuelle ou totale de n variables aléatoires réelles discrètes en calquant celles sur n événements.

5.3 Fonction de répartition ou cumulative

On appelle fonction de répartition ou fonction cumulative d'une variable aléatoire discrète X, la fonction réelle F partout définie :

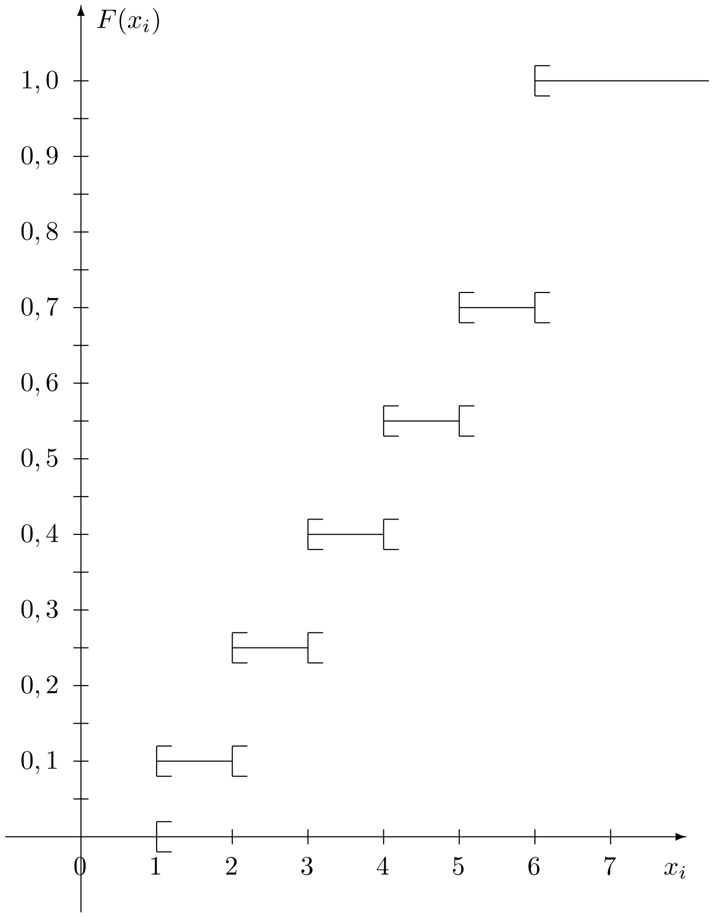
$$\begin{aligned} F:\ & \mathbb{R} \longrightarrow \mathbb{R}^+ \\ & x \longrightarrow F(x) = P(X \leq x) \end{aligned}$$

En particulier, si $x_1 < x_2 < \ldots < x_i < \ldots$, on a

$$F(x_i) = P(X=x_1\bigcup X=x_2\bigcup\ldots\bigcup X=x_i).$$

5.3.1 Exemple -suite du 5.2.1.-

La fonction de répartition est :



5.4 Les moments d'ordre n

Soit X une variable aléatoire réelle discrète. On appelle moment d'ordre n d'une variable aléatoire réelle discrète X, et on note $E(X^n)$ la quantité :

$$E(X^n) = \sum_i x_i^n P(X = x_i).$$

Ce moment d'ordre n n'existe, par définition, que si $\sum_i |x_i|^n P(X = x_i)$ est défini.

5.5 L'espérance mathématique

Soit X une variable aléatoire réelle discrète. Lorsque la quantité $\sum_i |x_i| P(X = x_i)$ est définie, on appelle espérance mathématique d'une variable aléatoire réelle discrète X, et on note $E(X)$ la quantité :

$$E(X) = \sum_i x_i P(X = x_i).$$

L'espérance mathématique est donc le moment d'ordre 1.

5.5.1 Exemple -suite du 5.2.1.-

$E(X) = 4$.

Exercice 48 [9, exercice 12, p. 136–137] Soit X une variable équirépartie sur $[[1, n]]$ (i.e. $P(X = i) = \frac{1}{n}$, $\forall i \in [[1, n]]$).

Soient $X_1, \dots, X_k$ des variables aléatoires de même loi que X et indépendantes.

On pose $Y = \max(X_1, \dots, X_k)$.

1. Déterminer la loi de Y.
2. Montrer que
$$E(Y) = n - \frac{1}{n^k}\sum_{i=1}^{n-1} i^k.$$

5.5.2 Propriétés de l'espérance mathématique

– Pour toute constante a, $E(a) = a$ (trivial)

– Pour toute constante a, $E(aX) = aE(X)$ (trivial)

– $E(X+Y) = E(X) + E(Y)$

Démonstration

Montrons d'abord que si $E(X)$ et si $E(Y)$ existent, alors, $E(X+Y)$ aussi.

On a $P(X+Y=z_k) = \sum_{i,j,x_i+y_j=z_k} P(X=x_i, Y=y_j)$.

Cependant,

$$\begin{aligned}\sum_k |z_k| P(X+Y=z_k) &= \sum_{i,j,k,x_i+y_j=z_k} |x_i+y_j| P(X=x_i,Y=y_j)\\ &\leq \sum_{i,j,k,x_i+y_j=z_k} |x_i| P(X=x_i,Y=y_j) + \sum_{i,j,k,x_i+y_j=z_k} |y_j| P(X=x_i,Y=y_j)\\ &\leq \sum_{i,j} |x_i| P(X=x_i,Y=y_j) + \sum_{i,j} |y_j| P(X=x_i,Y=y_j)\end{aligned}$$

car $\sum_{i,j,k,x_i+y_j=z_k} |x_i| P(X=x_i,Y=y_j) = \sum_{i,j} |x_i| P(X=x_i,Y=y_j)$

et car $\sum_{i,j,k,x_i+y_j=z_k} |y_j| P(X=x_i,Y=y_j) = \sum_{i,j} |y_j| P(X=x_i,Y=y_j)$.

Or $P(X=x_i) = \sum_j P(X=x_i,Y=y_j)$ et $P(Y=y_j) = \sum_i P(X=x_i,Y=y_j)$, d'après la définition d'une loi marginale. Donc
$$\sum_k |z_k| P(X+Y=z_k) \leq \sum_i |x_i| P(X=x_i) + \sum_j |y_j| P(Y=y_j).$$

Il s'ensuit que la série $E(X+Y)$ est absolument convergente et donc convergente.

Pour montrer que $E(X+Y) = E(X)+E(Y)$, il suffit maintenant de réécrire les lignes précédentes en enlevant les valeurs absolues et en remplaçant les inégalités par des égalités.

– $E(aX+bY) = aE(X) + bE(Y)$ (trivial)

– Lorsque X et Y sont indépendantes, on a $E(XY) = E(X)E(Y)$

Démonstration

Montrons d'abord que si $E(X)$ et si $E(Y)$ existent, alors, $E(XY)$ aussi.

On a :

$$\begin{aligned}\sum_k |z_k| P(XY=z_k) &= \sum_{i,j,k,x_iy_j=z_k} |x_iy_j| P(X=x_i,Y=y_j)\\ &= \sum_{i,j} |x_i||y_j| P(X=x_i)P(Y=y_j)\\ &= \sum_i |x_i| P(X=x_i) \times \sum_j |y_j| P(Y=y_j)\end{aligned}$$

On a pu écrire $P(X=x_i,Y=y_j) = P(X=x_i)P(Y=y_j)$ car les variables sont indépendantes.

On a montré que la série $E(XY)$ est absolument convergente et donc convergente.

Pour montrer que $E(XY) = E(X)E(Y)$, il suffit maintenant de réécrire les lignes précédentes en enlevant les valeurs absolues.

Exercice 49 [9, exercice 14, p. 138] N boules numérotées de 1 à N sont réparties dans deux urnes :
- M dans U_1 ;
- $N - M$ dans U_2.

On a d'autre part N jetons (équilibrés et indiscernables) numérotés dans une boîte.

Un mouvement consiste à tirer au hasard un jeton dans la boîte et de changer d'urne la boule qui porte le numéro du jeton tiré, puis à remettre le jeton dans la boîte.

X_n désigne le nombre de boules dans U_1 après n mouvements.

1. Déterminer la loi de X_1 et prouver que $E(X_1) = M + 1 - \frac{2M}{N}$.
2. Prouver que :
   (a) $P(X_{n+1} = 0) = \frac{1}{N} P(X_n = 1)$;
   (b) $P(X_{n+1} = k) = \frac{N-k+1}{N} P(X_n = k-1) + \frac{k+1}{N} P(X_n = k+1)$, $\forall k \in \{1, \ldots, N-1\}$;
   (c) $P(X_{n+1} = N) = \frac{1}{N} P(X_n = N-1)$.
3. Prouver que
$$E(X_{n+1}) = (1 - \frac{2}{N})E(X_n) + 1.$$
   En déduire que
$$E(X_n) = \frac{N}{2} + (M - \frac{N}{2})(1 - \frac{2}{N})^n.$$

5.5.3 Variable centrée

On dit qu'une variable aléatoire réelle discrète X est centrée si elle est d'espérance mathématique nulle.

Si on considère une variable aléatoire réelle discrète X d'espérance mathématique $E(X)$ (non nécessairement nulle), alors la variable aléatoire réelle discrète $X - E(X)$ est centrée.

En effet, $E(X - E(X)) = E(X) - E(E(X)) = E(X) - E(X) = 0$.

5.5.4 Espérance conditionnelle

Soient X et Y deux variables aléatoires réelles discrètes. Lorsque la quantité $\sum_k |x_k| P(X = x_k / Y = y_l)$ est définie, on appelle espérance conditionnelle de $X/Y = y_l$ (on lit X sachant que $Y = y_l$) la quantité :
$$E(X/Y = y_l) = \sum_k x_k P(X = x_k / Y = y_l).$$

Il s'ensuit que nous obtenons le résultat suivant :
$$E(X) = \sum_l E(X/Y = y_l) P(Y = y_l).$$

Démonstration

$$\begin{aligned}
\sum_l E(X/Y=y_l)P(Y=y_l) &= \sum_l \sum_k x_k P(X=x_k/Y=y_l)P(Y=y_l) \\
&= \sum_l \sum_k x_k P(X=x_k \bigcap Y=y_l) \\
&= \sum_k x_k \sum_l P(X=x_k \bigcap Y=y_l) \\
&= \sum_k x_k P(X=x_k) \\
& \quad \text{(par définition d'une loi marginale)} \\
&= E(X)
\end{aligned}$$

5.6 Variance et écart-type

Lorsque les quantités $\sum_i |x_i|P(X=x_i)$ et $\sum_i |x_i|^2 P(X=x_i)$ sont définies, on appelle variance d'une variable aléatoire réelle discrète X, et on note $V(X)$ la quantité :

$$V(X) = \sum_i (x_i - E(X))^2 P(X=x_i).$$

L'écart-type est alors défini par $\sigma(X) = \sqrt{V(X)}$.

5.6.1 Exemple -suite du 5.2.1.-

$V(X) = 3$ et $\sigma(X) = \sqrt{3}$

5.6.2 Propriétés de la variance

- $V(X) \geq 0$. (trivial)
- $V(X) = E((X-E(X))^2)$. (par définition)
- $V(X) = E(X^2) - (E(X))^2$.
  La variance est donc la différence entre le moment d'ordre 2 et le carré du moment d'ordre 1.
  Démonstration

$$\begin{aligned}
V(X) &= E((X-E(X))^2) \\
&= E(X^2) - E(2XE(X)) + E(E(X)^2) \\
&= E(X^2) - 2(E(X))^2 + (E(X))^2 \\
&= E(X^2) - (E(X))^2.
\end{aligned}$$

- Pour toute constante a, $V(a) = 0$. (en effet, $V(a) = E(a^2) - (E(a))^2 = a^2 - a^2 = 0$)
- Pour toute constante a, $V(aX) = a^2 V(X)$.
  (en effet, $V(aX) = E((aX)^2) - (E(aX))^2 = a^2[E(X^2) - (E(X))^2] = a^2 V(X)$)
- Lorsque X et Y sont indépendantes, on a $V(X+Y) = V(X) + V(Y)$.
  Démonstration

$$\begin{aligned}
V(X+Y) &= E((X+Y)^2) - (E(X+Y))^2 \\
&= E(X^2 + 2XY + Y^2) - (E(X) + E(Y))^2 \\
&= E(X^2) - (E(X))^2 + E(Y^2) - (E(Y))^2 - 2[E(XY) - E(X)E(Y)] \\
&= V(X) + V(Y)
\end{aligned}$$

car puisque X et Y sont indépendantes, $E(XY) = E(X)E(Y)$.

– Lorsque X et Y sont indépendantes, on a $V(aX + bY) = a^2V(X) + b^2V(Y)$. (trivial)

Exercice 50 [9, exercice 10, p. 135] Soit X une variable équirépartie sur trois valeurs : $\{0, 1, 2\}$.

On pose $Y = X_1 + \ldots + X_n$ où les X_i sont indépendantes et suivent la même loi que X.

Calculer l'espérance et la variance de Y.

Exercice 51 [9, exercice 5, p. 157] Soient X et Y deux variables alétoires discrètes prenant leurs valeurs dans $\mathbb{N}$ dont les lois sont définies par $\forall k \in \mathbb{N}$, $P(X = k) = p_k$ et $P(Y = k) = q_k$.

Soit $\theta \in [0; 1]$ et $\forall k \in \mathbb{N}$, $r_k = \theta p_k + (1 - \theta)q_k$.

1. Vérifier que $\displaystyle\sum_{k\geq 0} r_k = 1$.

Soit Z une variable alétoire discrète prenant ses valeurs dans $\mathbb{N}$ dont la loi est définie par $\forall k \in \mathbb{N}$, $P(Z = k) = r_k$.

On suppose que X et Y sont d'ordre 2 (i.e. $E(X)$, $E(X^2)$, $E(Y)$ et $E(Y^2)$ existent).

2. Calculer $E(Z)$ et $V(Z)$ en fonction de $E(X)$, $V(X)$, $E(Y)$ et $V(Y)$.

Remarque. La loi de Z est appelée mélange des lois de X et de Y avec les pondérations respectives θ et $1 - \theta$.

Exercice 52 [9, exercice 7, p. 159] Soit $p \in [0; 1]$. Soit X une variable alétoire discrète prenant ses valeurs dans $\mathbb{N}^*$ dont la loi est définie par $\forall k \in \mathbb{N}^*$, $P(X = k) = ck(1 - p)^{k-1}$.

1. Déterminer c.
2. Calculer $E(X)$ et $V(X)$.

5.6.3 Variable réduite

On dit qu'une variable aléatoire réelle discrète X est réduite si elle est de variance 1.

Si on considère une variable aléatoire réelle discrète X de variance $V(X) = (\sigma(X))^2 \neq 0$ (non nécessairement égale à 1), alors la variable aléatoire réelle discrète $X/\sigma(X)$ est réduite.

En effet, $V(X/\sigma(X)) = 1/(\sigma(X))^2V(X) = V(X)/V(X) = 1$.

Si on considère une variable aléatoire réelle discrète X d'espérance mathématique $E(X)$ et de variance $V(X) = (\sigma(X))^2$, alors la variable aléatoire réelle discrète $(X - E(X))/\sigma(X)$ est centrée et réduite.

5.6.4 Variance conditionnelle

Soient X et Y deux variables aléatoires réelles discrètes. Lorsque les quantités $\displaystyle\sum_k |x_k|P(X = x_k/Y = y_l)$ et $\displaystyle\sum_k |x_k|^2P(X = x_k/Y = y_l)$ sont définies, on appelle variance conditionnelle de $X/Y = y_l$ (on lit X sachant que $Y = y_l$) la quantité :

$$V(X/Y = y_l) = \sum_k (x_k - E(X/Y = y_l))^2 P(X = x_k/Y = y_l),$$

ou

$$V(X/Y = y_l) = E(X^2/Y = y_l) - [E(X/Y = y_l)]^2,$$

d'après la troisième propriété relative à la variance.

Il s'ensuit que nous obtenons le résultat suivant :

$$V(X) = \sum_l \left(V(X/Y = y_l) + [E(X/Y = y_l)]^2\right) P(Y = y_l) \ - [E(X)]^2.$$

Démonstration

On a :

$$\begin{aligned}
& \sum_l V(X/Y=y_l)P(Y=y_l) \\
= & \sum_l \sum_k (x_k - E(X/Y=y_l))^2 P(X=x_k/Y=y_l)P(Y=y_l) \\
= & \sum_l \sum_k x_k^2 P(X=x_k/Y=y_l)P(Y=y_l) \\
& -2\sum_l \sum_k E(X/Y=y_l) x_k P(X=x_k/Y=y_l)P(Y=y_l) \\
& +\sum_l \sum_k [E(X/Y=y_l)]^2 P(X=x_k/Y=y_l)P(Y=y_l) \\
= & \sum_l \sum_k x_k^2 P(X=x_k \bigcap Y=y_l) \\
& -2\sum_l E(X/Y=y_l)^2 P(Y=y_l) \\
& +\sum_l \sum_k [E(X/Y=y_l)]^2 P(X=x_k \bigcap Y=y_l) \\
= & \sum_k x_k^2 P(X=x_k) - 2\sum_l E(X/Y=y_l)^2 P(Y=y_l) + \sum_l [E(X/Y=y_l)]^2 P(Y=y_l) \\
& \text{(par définition d'une loi marginale)} \\
= & E(X^2) - \sum_l [E(X/Y=y_l)]^2 P(Y=y_l)
\end{aligned}$$

Donc,

$$\begin{aligned}
V(X) &= E(X^2) - [E(X)]^2 \\
&= \sum_l \left(V(X/Y=y_l) + [E(X/Y=y_l)]^2\right) P(Y=y_l) \ - [E(X)]^2
\end{aligned}$$

5.7 Covariance, régression et corrélation

5.7.1 Covariance

Soient X et Y deux variables aléatoires réelles discrètes. Lorsque les quantités $\sum_{i,j} |x_i y_j| P(X=x_i \bigcap Y=y_j)$, $\sum_i |x_i| P(X=x_i)$ et $\sum_j |y_j| P(Y=y_j)$ sont définies, on appelle covariance de deux variables aléatoires réelles discrètes X et Y, et on note $Cov(X,Y)$ la quantité :

$$Cov(X,Y) = \sum_{i,j} (x_i - E(X))(y_j - E(Y)) P(X=x_i \bigcap Y=y_j).$$

La covariance de deux variables aléatoires réelles discrètes indépendantes est nulle, et on a :

$$Cov(X,Y) = E(XY) - E(X)E(Y).$$

En retournant à la formule donnant la variance de la somme de deux variables aléatoires réelles discrètes, on trouve :

$$V(X+Y) = V(X) + V(Y) + 2Cov(X,Y).$$

En généralisant à n variables,

$$V(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} V(X_i) + 2\sum_{i<j} Cov(X_i, X_j).$$

Et ensuite, si les variables sont indépendantes deux à deux :

$$V(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} V(X_i).$$

Exercice 53 [9, exercice 8, p. 161] Soit (X, Y) un vecteur aléatoire discret à valeurs dans $\mathbb{R}^2$ dont la loi de probabilité est définie par $P(X = x_i, Y = y_j),\ \forall (i, j) \in \mathbb{N}^2$.

On suppose que $\sum_i |x_i| P(X = x_i)$, $\sum_j |y_j| P(Y = y_j)$, $\sum_i x_i^2 P(X = x_i)$ et $\sum_j y_j^2 P(Y = y_j)$ sont définies.

Montrer que $Cov(X, Y)$ est aussi définie.

5.7.2 Droite de régression et corrélation

On nomme coefficient de corrélation et l'on note $r(X, Y)$ la quantité $r(X, Y) = \frac{Cov(X,Y)}{\sqrt{V(X)V(Y)}}$.

Soient X et Y deux variables aléatoires réelles discrètes.

La droite D d'équation $y = ax + b$ où $a = \frac{Cov(X,Y)}{V(X)}$ et $b = E(Y) - aE(X)$, est appelée la droite de régression linéaire de Y par rapport à X ; cette droite D minimise la somme des carrés des écarts en ordonnée entre les représentations graphiques des valeurs et leurs projections sur D selon (0y), pondérés de leur probabilité.

Démonstration

Il nous faut minimiser en fonction de a et b,

$$\begin{aligned}
& \sum_{i,j} (y_j - ax_i - b)^2 P(X = x_i \bigcap Y = y_j) \\
= \ & [a^2 V(X) - 2aCov(X, Y) + V(Y)] + [(E(Y) - aE(X) - b)^2] \\
= \ & V(X)[a - \frac{Cov(X, Y)}{V(X)}]^2 + [(E(Y) - aE(X) - b)^2] + V(Y)(1 - (r(X, Y))^2)
\end{aligned}$$

Cette quantité est minimale lorsque b est tel que $E(Y) - aE(X) - b = 0$ (i.e. le point moyen appartient à cette droite) et lorsque a est tel que $a - \frac{Cov(X,Y)}{V(X)} = 0$. Dans ce cas, le minimum est $V(Y)(1 - (r(X, Y))^2)$.

Conséquences

1. Les points sont alignés si et seulement si $|r(X, Y)| = 1$.
2. La droite D' d'équation $x = a'y + b'$ où $a' = \frac{Cov(X,Y)}{V(Y)}$ et $b' = E(X) - a'E(Y)$, est appelée la droite de régression linéaire de X par rapport à Y ; cette droite D' minimise la somme des carrés des écarts en abscisse entre les représentations graphiques des valeurs et leurs projections sur D' selon (0x), pondérés de leur probabilité.
3. Le point de coordonnées $(E(X), E(Y))$ est commun aux deux droites D et D'.
4. Les deux droites sont confondues si et seulement si $|r(X, Y)| = 1$.

   Lorsque le coefficient de corrélation est proche de 1 en valeur absolue, on parle de bon ajustement et, dans ce cas, les droites D et D' sont *presque* confondues.

   Inversement, lorsque ce coefficient est proche de 0, on parle de mauvais ajustement, et dans ce cas, les variables X et Y sont presque non covariées (i.e. elles n'ont rien à voir entre-elles).

   Ainsi, ce coefficient rend compte de la validité de la régression linéaire.

Propriété

Si $V(X) \neq 0$ et si $V(Y) \neq 0$, alors $-1 \leq r(X,Y) \leq 1$.

En effet,

$$\begin{aligned} 0 \leq V(\frac{X}{\sigma(X)} + \frac{Y}{\sigma(Y)}) &= \frac{V(X)}{(\sigma(X))^2} + \frac{V(Y)}{(\sigma(Y))^2} + 2\frac{Cov(X,Y)}{\sigma(X)\sigma(Y)} \\ &= 2[1 + r(X,Y)] \end{aligned}$$

et donc $-1 \leq r(X,Y)$. Puis,

$$\begin{aligned} 0 \leq V(\frac{X}{\sigma(X)} - \frac{Y}{\sigma(Y)}) &= \frac{V(X)}{(\sigma(X))^2} + \frac{V(Y)}{(\sigma(Y))^2} - 2\frac{Cov(X,Y)}{\sigma(X)\sigma(Y)} \\ &= 2[1 - r(X,Y)] \end{aligned}$$

et donc $1 \geq r(X,Y)$.

Exercice 54 Soient X et Y deux variables alétoires discrètes prenant leurs valeurs dans $\mathbb{N}$. On suppose que pour $p \in]0,1[$ et pour $\lambda \in \mathbb{R}_+$, la loi du couple (X,Y) est donnée par

$$P((X,Y) = (i,j)) = e^{-\lambda} \frac{(\lambda p)^i}{i!} \frac{(\lambda(1-p))^j}{j!}.$$

1. Déterminer les lois marginales de X et Y.
2. Montrer que X et Y sont indépendantes.

Exercice 55 Soit (X,Y) une série double représentée dans le plan rapporté à un repère $(O, \vec{i}, \vec{j})$ par les points $M_k(x_k, y_k)$, $k \in \{1, 2, \ldots, n\}$.

On considère le point moyen $G(m(X), m(Y))$ et le changement de variables $x' = x - m(X)$, $y' = y - m(Y)$ où $m(X) = \frac{1}{n}\sum_{k=0}^{n} x_k$ et $m(Y) = \frac{1}{n}\sum_{k=0}^{n} y_k$.

Étant donnée une droite Δ dont une équation dans le repère $(G, \vec{i}, \vec{j})$ est

$$x' \cos\alpha + y' \sin\alpha - \rho = 0.$$

On note H_k la projection orthogonale de M_k sur la droite Δ.

On se propose de déterminer Δ (i.e. α et ρ) de manière à minimiser

$$S = \sum_{k=1}^{n} \overline{H_k M_k}^2.$$

1. Montrer que $S = \sum_{k=1}^{n} (x'_k \cos\alpha + y'_k \sin\alpha - \rho)^2$.
2. Montrer que s'il existe une droite Δ minimisant S, cette droite passe par G.
3. Déterminer le coefficient directeur de Δ qui rend S minimum (préciser dans quel cas, la solution est unique).

   La droite Δ ajustée est appelée droite de régression orthogonale.
4. On suppose que la droite de régression orthogonale est unique. Soient D et D' les droites de régression de Y par rapport à X et de X par rapport à Y. Étudier la position relative des droites D, D' et Δ.
5. Soit $T = \sum_{k=1}^{n} \overline{GM_k}^2$ et $S' = \sum_{k=1}^{n} \overline{GH_k}^2$.

   Montrer que minimiser S revient à maximiser S'.

   On pose $q = \frac{S'}{T}$.

   Montrer que q est un indice de qualité de l'ajustement compris entre $\frac{1}{2}$ et 1.

   Quelle est la significaion des valeurs $q = \frac{1}{2}$ et $q = 1$?

6. On donne les notes d'un groupe de 8 étudiants dans deux disciplines : mathématiques (note x sur 20) et informatique (note y sur 20).

x	12	15	8	6	10	12	11	9
y	10	16	11	4	9	14	17	8

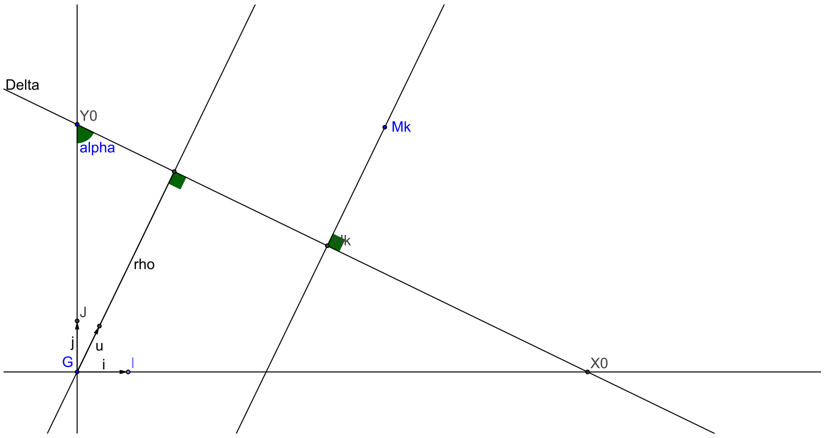
Ajuster sur ces données les deux droites de régression D et D' et calculer le coefficient de corrélation r.

Ajuster sur ces données la droite de régression orthogonale Δ et calculer le coefficient de qualité q de cet ajustement.

Dans le repère $(G, \vec{i}, \vec{j})$, la droite Δ a pour équation cartésienne

$$x' \cos\alpha + y' \sin\alpha - \rho = 0,$$

coupe l'axe des abscisses en le point X_0 qui a pour coordonnées $(\frac{\rho}{\cos(\alpha)}; 0)$ et l'axe des ordonnées en le point Y_0 qui a pour coordonnées $(0; \frac{\rho}{\sin(\alpha)})$.



1. Dans le repère $(G, \vec{i}, \vec{j})$, le vecteur $\vec{u}$ de coordonnées $(\cos(\alpha); \sin(\alpha))$ est unitaire et normal à la droite Δ. Donc, $\overline{H_kM_k}^2 = \left\langle \overrightarrow{H_kM_k}, \vec{u} \right\rangle^2 = \left(\left\langle \overrightarrow{GM_k}, \vec{u} \right\rangle - \left\langle \overrightarrow{GH_k}, \vec{u} \right\rangle \right)^2$.

   Or $\left\langle \overrightarrow{GM_k}, \vec{u} \right\rangle = x'_k \cos(\alpha) + y' k \sin(\alpha)$ et $\left\langle \overrightarrow{GH_k}, \vec{u} \right\rangle = \rho$. Donc, $S = \sum_{k=1}^n (x'_k \cos\alpha + y'_k \sin\alpha - \rho)^2$.
2. $S = \sum_{k=1}^n (x'_k \cos\alpha + y'_k \sin\alpha)^2 + n\rho^2$, car $\sum_{k=1}^n x'_k = \sum_{k=1}^n y'_k = 0$. Il est alors évident que S sera minimale lorsque $\rho = 0$ et donc lorsque cette droite passe par G.
3. Comme $\rho = 0$, il reste à minimiser $\frac{1}{n} \sum_{k=1}^n (x'_k \cos\alpha + y'_k \sin\alpha)^2$.

$$\begin{aligned}
\frac{1}{n} \sum_{k=1}^n (x'_k \cos\alpha + y'_k \sin\alpha)^2 &= V(X)(\cos(\alpha))^2 + V(Y)(\sin(\alpha))^2 + 2Cov(X,Y)\sin(\alpha)\cos(\alpha) \\
&= V(X)\frac{1+\cos(2\alpha)}{2} + V(Y)\frac{1-\cos(2\alpha)}{2} + Cov(X,Y)\sin(2\alpha) \\
&= \frac{V(X)-V(Y)}{2}\cos(2\alpha) + Cov(X,Y)\sin(2\alpha) + \frac{V(X)+V(Y)}{2}
\end{aligned}$$

Cas 1 $V(X) = V(Y)$.

- Si $Cov(X,Y) = 0$, $\frac{1}{n} \sum_{k=1}^n (x'_k \cos\alpha + y'_k \sin\alpha)^2 = \frac{V(X)+V(Y)}{2}$ quel que soit α.
- Si $Cov(X,Y) > 0$, $\frac{1}{n} \sum_{k=1}^n (x'_k \cos\alpha + y'_k \sin\alpha)^2$ est minimum lorsque $2\alpha = -\frac{\pi}{2} + 2k\pi$ ou $\alpha = -\frac{\pi}{4} + k\pi$. Le coefficient directeur de Δ est donc $t = \tan(\alpha - \frac{\pi}{2}) = 1$.

– Si $Cov(X,Y) < 0$, $\frac{1}{n}\sum_{k=1}^{n}(x'_k\cos\alpha + y'_k\sin\alpha)^2$ est minimum lorsque $2\alpha = \frac{\pi}{2} + 2k\pi$ ou $\alpha = \frac{\pi}{4} + k\pi$. Le coefficient directeur de Δ est donc $t = \tan(\alpha - \frac{\pi}{2}) = -1$.

Cas 2 $V(X) \neq V(Y)$.

On pose $\tan(\theta) = \frac{2Cov(X,Y}{V(X)-V(Y)}$ avec $-\frac{\pi}{2} < \theta < \frac{\pi}{2}$. Ainsi, $\frac{1}{n}\sum_{k=1}^{n}(x'_k\cos\alpha + y'_k\sin\alpha)^2 = \frac{V(X)-V(Y)}{2\cos(\theta)}\cos(2\alpha - \theta) + \frac{V(X)+V(Y)}{2}$.

Cas 2.a $V(X) - V(Y) > 0$.

Le minimum de $\frac{1}{n}\sum_{k=1}^{n}(x'_k\cos\alpha + y'_k\sin\alpha)^2$ est obtenu pour $2\alpha - \theta = \pi + 2k\pi$ ou $\alpha = \frac{\theta}{2} + \frac{\pi}{2} + k\pi$. Le coefficient directeur de Δ est donc $t = \tan(\alpha - \frac{\pi}{2}) = \tan(\frac{\theta}{2})$.

Comme $\tan(\theta) = \frac{2t}{1-t^2}$, il découle $\frac{2Cov(X,Y}{V(X)-V(Y)} = \frac{2t}{1-t^2}$ ou $t^2Cov(X,Y) + t(V(X) - V(Y)) - Cov(X,Y) = 0$. Et, comme $-\frac{\pi}{4} < \frac{\theta}{2} < \frac{\pi}{4}$, t est l'unique racine de $t^2Cov(X,Y) + t(V(X) - V(Y)) - Cov(X,Y) = 0$ qui soit comprise entre -1 et 1 (il n'y a qu'une racine qui soit comprise entre -1 et 1 car le produit des racines est égal à 1).

Si $f(t) = t^2Cov(X,Y) + t(V(X) - V(Y)) - Cov(X,Y)$, on a $f(-1) = V(Y) - V(X) < 0$ et $f(1) = V(X) - V(Y) > 0$. Donc

– si $Cov(X,Y) = 0$, on a $t = 0$;

– si $Cov(X,Y) > 0$, t est la plus grande des racines ou $t = \frac{V(Y)-V(X)+\sqrt{(V(X)-V(Y))^2+4(Cov(X,Y))^2}}{2Cov(X,Y)}$;

– si $Cov(X,Y) < 0$, t est la plus petite des racines ou $t = \frac{V(Y)-V(X)+\sqrt{(V(X)-V(Y))^2+4(Cov(X,Y))^2}}{2Cov(X,Y)}$ (même expression que précédemment).

Cas 2.b $V(X) - V(Y) < 0$.

Le minimum de $\frac{1}{n}\sum_{k=1}^{n}(x'_k\cos\alpha + y'_k\sin\alpha)^2$ est obtenu pour $2\alpha - \theta = 2k\pi$ ou $\alpha = \frac{\theta}{2} + k\pi$. Le coefficient directeur de Δ est donc $t = \tan(\alpha - \frac{\pi}{2}) = -\frac{1}{\tan(\frac{\theta}{2})}$.

Comme $\tan(\theta) = \frac{2(\frac{-1}{t})}{1-(\frac{-1}{t})^2}$, il découle $\frac{2Cov(X,Y}{V(X)-V(Y)} = \frac{2t}{1-t^2}$ ou $t^2Cov(X,Y) + t(V(X) - V(Y)) - Cov(X,Y) = 0$. Et, comme $-\frac{\pi}{4} < \frac{\theta}{2} < \frac{\pi}{4}$, t est l'unique racine de $t^2Cov(X,Y) + t(V(X) - V(Y)) - Cov(X,Y) = 0$ qui ne soit pas comprise entre -1 et 1 (il n'y a qu'une racine qui ne soit pas comprise entre -1 et 1 car le produit des racines est égal à 1).

Si $f(t) = t^2Cov(X,Y) + t(V(X) - V(Y)) - Cov(X,Y)$, on a $f(-1) = V(Y) - V(X) > 0$ et $f(1) = V(X) - V(Y) < 0$. Donc

– si $Cov(X,Y) = 0$, on a $t = +\infty$;

– si $Cov(X,Y) > 0$, t est la plus grande des racines ou $t = \frac{V(Y)-V(X)+\sqrt{(V(X)-V(Y))^2+4(Cov(X,Y))^2}}{2Cov(X,Y)}$ (même expression que dans le cas 2.a) ;

– si $Cov(X,Y) < 0$, t est la plus petite des racines ou $t = \frac{V(Y)-V(X)+\sqrt{(V(X)-V(Y))^2+4(Cov(X,Y))^2}}{2Cov(X,Y)}$ (même expression que précédemment).

La droite Δ ajustée est appelée droite de régression orthogonale. Le seul cas où elle n'est pas déterminée correspond à $V(X) = V(Y)$ et $Cov(X,Y) = 0$.

4. Pour rappel, dans le repère $(G, \vec{i}, \vec{j})$, la droite D a pour équation cartésienne $y' = ax'$ avec $a = \frac{Cov(X,Y)}{V(X)}$ et la droite D' a pour équation cartésienne $y' = bx'$ avec $b = \frac{V(Y)}{Cov(X,Y)}$.

Cas 1 $V(X) = V(Y)$.

– Si $Cov(X,Y) = 0$, la droite Δ n'est pas déterminée.

– Si $Cov(X,Y) > 0$, la droite Δ est la première bissectrice, D a un coefficient directeur dans $]0,1]$ et D' a un coefficient directeur dans $[1,\infty[$.

– Si $Cov(X,Y) < 0$, la droite Δ est la deuxième bissectrice, D a un coefficient directeur dans $[-1,0[$ et D' a un coefficient directeur dans $]-\infty,1]$.

Cas 2 $V(X) \neq V(Y)$.

Cas 2.a $V(X) - V(Y) > 0$.

- Si $Cov(X,Y) = 0$, on a $t = 0$, et la droite Δ est l'axe des abscisses, D est l'axe des abscisses et D' est l'axe des ordonnées ;
- On a $(V(X)-V(Y))^2+4(Cov(X,Y))^2 = (V(X))^2+(V(X))^2-2V(X)V(Y)+4(Cov(X,Y))^2 \leq (V(X))^2+(V(X))^2+2V(X)V(Y) = (V(X)+V(Y))^2$. Ainsi, si $Cov(X,Y) > 0$, alors $t = \frac{V(Y)-V(X)+\sqrt{(V(X)-V(Y))^2+4(Cov(X,Y))^2}}{2Cov(X,Y)} \geq \frac{V(Y)}{Cov(X,Y)} = b$ et $t = \frac{2Cov(X,Y)}{V(Y)-V(X)-\sqrt{(V(X)-V(Y))^2+4(Cov(X,Y))^2}} \geq \frac{Cov(X,Y)}{V(X)} = a$ (cette seconde expression de t provient du fait que le produit des racines de $f(t)$ vaut 1) ;
- On a toujours $(V(X)-V(Y))^2+4(Cov(X,Y))^2 = (V(X))^2+(V(X))^2-2V(X)V(Y)+4(Cov(X,Y))^2 \leq (V(X))^2+(V(X))^2+2V(X)V(Y) = (V(X)+V(Y))^2$. Ainsi, si $Cov(X,Y) < 0$, alors $t = \frac{V(Y)-V(X)+\sqrt{(V(X)-V(Y))^2+4(C\cdots}}{2Cov(X,Y)}$ $\frac{V(Y)}{Cov(X,Y)} = b$ et $t = \frac{2Cov(X,Y)}{V(Y)-V(X)-\sqrt{(V(X)-V(Y))^2+4(Cov(X,Y))^2}} \leq \frac{Cov(X,Y)}{V(X)} = a$ (cette seconde expression de t provient du fait que le produit des racines de $f(t)$ vaut 1).

Cas 2.b $V(X) - V(Y) < 0$.

- si $Cov(X,Y) = 0$, on a $t = +\infty$, et la droite Δ est l'axe des ordonnées, D est l'axe des abscisses et D' est l'axe des ordonnées ;
- On a $(V(X)-V(Y))^2+4(Cov(X,Y))^2 = (V(X))^2+(V(X))^2-2V(X)V(Y)+4(Cov(X,Y))^2 \leq (V(X))^2+(V(X))^2+2V(X)V(Y) = (V(X)+V(Y))^2$. Ainsi, si $Cov(X,Y) > 0$, alors $t = \frac{V(Y)-V(X)+\sqrt{(V(X)-V(Y))^2+4(Cov(X,Y))^2}}{2Cov(X,Y)}$ $\frac{V(Y)}{Cov(X,Y)} = b$ et $t = \frac{2Cov(X,Y)}{V(Y)-V(X)-\sqrt{(V(X)-V(Y))^2+4(Cov(X,Y))^2}} \geq \frac{Cov(X,Y)}{V(X)} = a$ (cette seconde expression de t provient du fait que le produit des racines de $f(t)$ vaut 1) ;
- On a toujours $(V(X)-V(Y))^2+4(Cov(X,Y))^2 = (V(X))^2+(V(X))^2-2V(X)V(Y)+4(Cov(X,Y))^2 \leq (V(X))^2+(V(X))^2+2V(X)V(Y) = (V(X)+V(Y))^2$. Ainsi, si $Cov(X,Y) < 0$, alors $t = \frac{V(Y)-V(X)+\sqrt{(V(X)-V(Y))^2+4(C\cdots}}{2Cov(X,Y)}$ $\frac{V(Y)}{Cov(X,Y)} = b$ et $t = \frac{2Cov(X,Y)}{V(Y)-V(X)-\sqrt{(V(X)-V(Y))^2+4(Cov(X,Y))^2}} \leq \frac{Cov(X,Y)}{V(X)} = a$ (cette seconde expression de t provient du fait que le produit des racines de $f(t)$ vaut 1).

5. Soit $T = \sum_{k=1}^n \overline{GM_k}^2$ et $S' = \sum_{k=1}^n \overline{GH_k}^2$.

   On a $T = S' + S$, d'après le théorème de Pythagore, et comme T est constant, minimiser S revient à maximiser S'.

   On a $q = \frac{S'}{T} = 1 - \frac{S}{T}$, et par suite, $q \leq 1$. Le cas $q = 1$ correspond à $H_kM_k = 0$, $\forall k$ et donc au cas où les points sont alignés.

   Soit Δ' la perpendiculaire à Δ passant par G et soit H'_k la projection orthogonale de M_k sur la droite Δ'. On a $S = \sum_{k=1}^n \overline{GH'_k}^2$. Par suite, on a $S \leq S'$ (car S est calculée avec la droite réalisant le minimum de la somme) puis $q \geq \frac{1}{2}$. Le cas $q = \frac{1}{2}$ correspond à $S = S'$ et donc au cas où l'on ne peut déterminer une unique droite Δ.

6. On donne les notes d'un groupe de 8 étudiants dans deux disciplines : mathématiques (note x sur 20) et informatique (note y sur 20).

x	12	15	8	6	10	12	11	9	83
y	10	16	11	4	9	14	17	8	89
x^2	144	225	64	36	100	144	121	81	915
y^2	100	256	121	16	81	196	289	64	1123
xy	120	240	88	24	90	168	187	72	989

$m(X) = \frac{83}{8} \approx 10,38$; $m(Y) = \frac{89}{8} \approx 11,13$; $V(X) = \frac{915}{8} - (\frac{83}{8})^2 \approx 6,73$; $V(Y) = \frac{1123}{8} - (\frac{89}{8})^2 \approx 16,61$; $Cov(X,Y) = \frac{989}{8} - \frac{83}{8} \times \frac{89}{8} \approx 8,20$.

Pour D et D', $a = \frac{Cov(X,Y)}{V(X)} \approx 1,22$; $b = \frac{V(Y)}{Cov(X,Y)} \approx 2,02$; $r = \frac{Cov(X,Y)}{\sqrt{V(X)V(Y)}} \approx 0,78$.

Pour Δ, avec $\alpha = \frac{\theta}{2} + k\pi$, $t = \underbrace{\frac{V(Y) - V(X) + \sqrt{(V(X) - V(Y))^2 + 4(Cov(X,Y))^2}}{2Cov(X,Y)}}_{\approx 1,77} = -\frac{1}{\tan(\frac{\theta}{2})}$ ou $t' = \tan(\frac{\theta}{2}) \approx -0,56$, puis $\sin(2\alpha) = \sin(\theta) = \frac{2t'}{1+t'^2} \approx -0,85$ et $\cos(2\alpha) = \cos(\theta) = \frac{1-t'^2}{1+t'^2} \approx 0,52$.
Ainsi $\frac{S}{8} = \frac{V(X)-V(Y)}{2}\cos(2\alpha) + Cov(X,Y)\sin(2\alpha) + \frac{V(X)+V(Y)}{2} \approx 2,10$.
Or, $\frac{T}{8} = V(X) + V(Y) \approx 23,34$.
Donc, $q = 1 - \frac{S}{T} \approx 0,91$.

5.7.3 Covariance conditionnelle

Soient X, Y et Z trois variables aléatoires réelles discrètes. Lorsque les quantités $\sum_k |x_k| P((X = x_k \bigcap Y = y_l)/Z = z_m)$ et $\sum_l |y_l| P((X = x_k \bigcap Y = y_l)/Z = z_m)$ et $\sum_{k,l} |x_k y_l| P((X = x_k \bigcap Y = y_l)/Z = z_m)$ sont définies, on appelle covariance conditionnelle de $X, Y/Z = z_m$ (on lit X et Y sachant que $Z = z_m$) la quantité :

$$Cov(X,Y/Z = z_m) = \sum_{k,l}(x_k - E(X/Z = z_m))(y_l - E(Y/Z = z_m))P((X = x_k \bigcap Y = y_l)/Z = z_m),$$

ou

$$Cov(X,Y/Z = z_m) = E(XY/Z = z_m) - E(X/Z = z_m)E(Y/Z = z_m),$$

d'après une propriété citée plus haut, relative à la covariance.

Il s'ensuit que nous obtenons le résultat suivant :

$$Cov(X,Y) = \sum_m \left(Cov(X,Y/Z = z_m) + E(X/Z = z_m)E(Y/Z = z_m)\right) P(Z = z_m) - E(X)E(Y).$$

Démonstration

On a :

$$\begin{aligned}
& \sum_m Cov(X,Y/Z=z_m)P(Z=z_m) \\
= & \sum_m \sum_{k,l} (x_k - E(X/Z=z_m))(y_l - E(Y/Z=z_m))P((X=x_k \bigcap Y=y_l)/Z=z_m)P(Z=z_m) \\
= & \sum_m \sum_{k,l} x_k y_l P((X=x_k \bigcap Y=y_l)/Z=z_m)P(Z=z_m) \\
& -\sum_m \sum_{k,l} E(Y/Z=z_m) x_k P((X=x_k \bigcap Y=y_l)/Z=z_m)P(Z=z_m) \\
& -\sum_m \sum_{k,l} E(X/Z=z_m) y_l P((X=x_k \bigcap Y=y_l)/Z=z_m)P(Z=z_m) \\
& +\sum_m \sum_{k,l} E(X/Z=z_m)E(Y/Z=z_m)P((X=x_k \bigcap Y=y_l)/Z=z_m)P(Z=z_m) \\
= & \sum_m \sum_{k,l} x_k y_l P(X=x_k \bigcap Y=y_l \bigcap Z=z_m) \\
& -\sum_m E(Y/Z=z_m)E(X/Z=z_m)P(Z=z_m) \\
& -\sum_m E(X/Z=z_m)E(Y/Z=z_m)P(Z=z_m) \\
& +\sum_m \sum_{k,l} E(X/Z=z_m)E(Y/Z=z_m)P(X=x_k \bigcap Y=y_l \bigcap Z=z_m) \\
= & \sum_{k,l} x_k y_l P(X=x_k \bigcap Y=y_l) - \sum_m E(Y/Z=z_m)E(X/Z=z_m)P(Z=z_m) \\
& \text{(par définition d'une loi marginale)} \\
= & E(XY) - \sum_m E(Y/Z=z_m)E(X/Z=z_m)P(Z=z_m)
\end{aligned}$$

Donc,

$$\begin{aligned}
Cov(X,Y) &= E(XY) - E(X)E(Y) \\
&= \sum_m \left(Cov(X,Y/Z=z_m) + E(X/Z=z_m)E(Y/Z=z_m)\right) P(Z=z_m) \\
&\quad -E(X)E(Y)
\end{aligned}$$

Exercice 56 Soit X la variable aléatoire dont la valeur est le nombre de points indiqués par la face supérieure d'un dé à 6 faces non pipé, lors de son lancer.

On considère une urne contenant 6 boules indiscernables au toucher dont 3 noires, 2 blanches et 1 verte.

On extrait au hasard une boule de l'urne. Si celle-ci est noire N, on lance 3 dés, si elle est blanche B, 2 dés et si elle est verte V, un dé.

Le jeu coûte 8€et rapporte 1€par nombre Y de point(s) indiqué(s) par le(s) dé(s). Combien peut-on espérer gagner G ou perdre au cours d'une partie de ce jeu ?

Solution : $E(X) = \frac{7}{2}$; $E(Y/V) = E(X) = \frac{7}{2}$; $E(Y/B) = 2E(X) = 7$; $E(Y/N) = 3E(X) = \frac{21}{2}$; $P(N) = \frac{1}{2}$; $P(B) = \frac{1}{3}$; $P(V) = \frac{1}{6}$.

Ainsi, $E(Y) = \frac{7}{2}\frac{1}{6} + 7\frac{1}{3} + \frac{21}{2}\frac{1}{2} = \frac{49}{6}$.

Et, $E(G) = E(Y) - 8 = \frac{1}{6}$.

Exercice 57 Soit X la variable aléatoire dont la valeur est le nombre de points indiqués par la face supérieure d'un dé à 6 faces non pipé, lors de son lancer.

On appelle k le résultat obtenu à ce lancer.

Soit Y la variable aléatoire totalisant le nombre de fois que l'on obtient pile en jetant k pièces de monnaie équilibrées.

Calculer $E(X)$, $V(X)$, $E(Y)$, $V(Y)$, et $Cov(X,Y)$.

Solution : Pour calculer $E(X)$ et $V(X)$, on utilise $P(X=k)=\frac{1}{6}, \forall k\in\{1,2,\dots,6\}$, et il s'ensuit que $E(X)=\frac{1}{6}\sum_{k=1}^{6}k=\frac{7}{2}$; $V(X)=\frac{1}{6}\sum_{k=1}^{6}(k-\frac{7}{2})^2=\frac{35}{12}$.

Afin d'obtenir $E(Y)$ et $V(Y)$, on utilise $P(Y=i/X=k)=\frac{C_k^i}{2^k}, \forall i\in\{1,2,\dots,k\}$, et il s'ensuit que $E(Y/X=k)=\sum_{i=1}^{k}i\frac{C_k^i}{2^k}=\frac{k}{2}$ et que $E(Y^2/X=k)=\sum_{i=1}^{k}i^2\frac{C_k^i}{2^k}=\frac{k(k+1)}{4}$ puis que $V(Y/X=k)=E(Y^2/X=k)-(E(Y/X=k))^2=\frac{k}{4}$.

En poursuivant, $E(Y)=\frac{1}{6}\sum_{k=1}^{6}\frac{k}{2}=\frac{7}{4}$; $E(Y^2)=\frac{1}{6}\sum_{k=1}^{6}\frac{k(k+1)}{4}=\frac{14}{3}$ et donc $V(Y)=E(Y^2)-(E(Y))^2=\frac{77}{48}$ ou encore, $V(Y)=\frac{1}{6}\sum_{k=1}^{6}(\frac{k}{4}+\frac{k^2}{2^2})-(\frac{7}{4})^2=\frac{77}{48}$, d'après la formule de la variance conditionnelle.

Par ailleurs, il est évident que $E(XY/X=k)=kE(Y/X=k)=\frac{k^2}{2}$ et que $Cov(X,Y/X=k)=E(XY/X=k)-E(X/X=k)E(Y/X=k)=0$.

Puis $E(XY)=\frac{1}{6}\sum_{k=1}^{6}\frac{k^2}{2}=\frac{91}{12}$, et donc $Cov(X,Y)=E(XY)-E(X)E(Y)=\frac{35}{24}$ ou encore, $Cov(X,Y)=\frac{1}{6}\sum_{k=1}^{6}(0+\frac{k^2}{2})-\frac{7}{2}\cdot\frac{7}{4}=\frac{35}{24}$, d'après la formule de la covariance conditionnelle.

Donner les deux droites de régression. Expliquer.

Solution : $D_{Y/X} : y=\frac{x}{2}$; $D_{X/Y} : x=\frac{10y+21}{11}$; $r(X,Y)=\sqrt{\frac{5}{11}}$. (mauvais ajustement linéaire)

Tracer et explications ; donner un nuage statistique de points ayant les mêmes caractéristiques.

Exercice 58 On considère une urne contenant N (on suppose $N\geq 1$) boules noires et B (on suppose $B\geq 2$) boules blanches, toutes indiscernables.

Le joueur A extrait des boules de l'urne, sans les remettre, une à une, jusqu'à tirer une boule blanche. Ensuite, le joueur B tire des boules de l'urne, en les remettant, une à une, jusqu'à tirer une boule blanche.

Le vainqueur du jeu est celui qui a tiré le plus de boules noires. Qui a le plus de chance de gagner ? Si le joueur B parie B mises alors que le joueur A ne parie que $B-1$ mises, le pari est-il équilibré ?

Solution :

Soit X la variable aléatoire qui retourne le nombre de boules noires tirées par A avant de tirer une boule blanche.
$P(X=k)=\left(\frac{N}{N+B}\frac{N-1}{N+B-1}\cdots\frac{N+1-k}{N+B+1-k}\right)\frac{B}{N+B-k}$.
Et donc, $P(X=k)=\frac{N!B}{(N+B)!}\frac{(N+B-1-k)!}{(N-k)!}$.
On remarque que $\sum_{k=0}^{N}C_{N+B-1-k}^{B-1}=C_{N+B}^{B}$.
Ainsi,

$$\begin{aligned}
E(N-X) &= \frac{N!B}{(N+B)!}\sum_{k=0}^{N}(N-k)\frac{(N+B-1-k)!}{(N-k)!}\\
&= \frac{N!B}{(N+B)!}\sum_{k=0}^{N-1}(N-k)\frac{(N+B-1-k)!}{(N-k)!}\\
&= \frac{N!B}{(N+B)!}\sum_{k=0}^{N-1}\frac{(N+B-1-k)!}{(N-1-k)!}\\
&= \frac{N!B!B}{(N+B)!}\sum_{k=0}^{N-1}C_{(N-1)+(B+1)-1-k}^{B}\\
&= \frac{N!B!B}{(N+B)!}C_{N+B}^{B+1}\\
&= \frac{BN}{B+1},
\end{aligned}$$

et donc $E(X) = \frac{N}{B+1}$.

Soit Y la variable aléatoire qui retourne le nombre de boules noires tirées par B avant de tirer une boule blanche.

D'autre part, $P(Y = l/X = k) = \frac{B-1}{B+N-1-k}\left(\frac{N-k}{B+N-1-k}\right)^l$.

Donc

$$\begin{aligned} E(Y/X = k) &= \frac{B-1}{B+N-1-k}\frac{N-k}{B+N-1-k}\sum_{l\geq 1} l\left(\frac{N-k}{B+N-1-k}\right)^{l-1} \\ &= \frac{B-1}{B+N-1-k}\frac{N-k}{B+N-1-k}\left(\frac{B+N-1-k}{B-1}\right)^2 \\ &= \frac{N-k}{B-1}, \end{aligned}$$

puis,

$$\begin{aligned} E(Y) &= \frac{N!B}{(N+B)!(B-1)}\sum_{k=0}^{N}(N-k)\frac{(N+B-1-k)!}{(N-k)!} \\ &= \frac{N!B}{(N+B)!(B-1)}\sum_{k=0}^{N-1}(N-k)\frac{(N+B-1-k)!}{(N-k)!} \\ &= \frac{N!B}{(N+B)!(B-1)}\sum_{k=0}^{N-1}\frac{(N+B-1-k)!}{(N-1-k)!} \\ &= \frac{N!B!B}{(N+B)!(B-1)}\sum_{k=0}^{N-1}C^B_{(N-1)+(B+1)-1-k} \\ &= \frac{N!B!B}{(N+B)!(B-1)}C^{B+1}_{N+B} \\ &= \frac{BN}{(B-1)(B+1)}, \end{aligned}$$

Il s'ensuit que le joueur B a plus de chance de gagner que le joueur A car $\frac{B}{B-1} \geq 1$. Par contre, le pari dont la cote est fixée plus haut est équitable.

Exercice 59 [9, exercice 13, p. 163–164] Soit (X, Y) un vecteur aléatoire discret à valeurs dans $\mathbb{N}^2$ de loi définie par

$$\forall (j,k) \in \mathbb{N}^2, P(X = j, Y = k) = \frac{(j+k)\lambda^{j+k}}{e(j!)(k!)}.$$

1. Montrer que $\lambda = \frac{1}{2}$.
2. Déterminer les lois marginales.

   Les variables aléatoires X et Y sont-elles indépendantes ?
3. Déterminer la loi conditionnelle de Y liée par $X = 0$.

   Calculer sa moyenne et sa variance.
4. Calculer $E(X)$, $E(Y)$, $V(X)$, $V(Y)$, $Cov(X, Y)$, $r(X, Y)$.
5. Soit $Z = 2^{X+Y}$.

   Calculer $E(Z)$ et $V(Z)$.

5.8 Les fonctions génératrices de moments

Soit X une variable aléatoire réelle ne prenant que des valeurs dans $\mathbb{N}$.

On appelle $G_X(t) = \sum_{n\geq 0} t^n P(X = n)$ la fonction génératrice des moments de la variable X.
Elle est au moins définie sur $]-1, 1[$.
On a :
- $G_X^{(k)}(0) = k!P(X = k)$; (application de la formule de Taylor)
- $G_X^{(k)}(1) = E(X(X-1)\dots(X-k+1))$.

Exercice 60 [9, exercice 10, p. 162] Montrer que la fonction génératrice des moments de la somme de deux variables aléatoires à valeurs dans $\mathbb{N}$ indépendantes est le produit des fonctions génératrices de ces deux varables aléatoires.

5.9 Inégalité de Markov

Soit la variable aléatoire réelle discrète X à valeurs non négatives, d'espérance mathématique $E(X) = m$. Alors,

$$P(X \geq a) \leq \frac{m}{a}, \forall a > 0.$$

Démonstration

$E(X) = m = \sum_k x_k P(X = x_k)$.
Donc,

$$\begin{aligned}
\forall a > 0, m &= \sum_{k, x_k \geq a} x_k P(X = x_k) + \sum_{k, x_k < a} x_k P(X = x_k) \\
&\geq \sum_{k, x_k \geq a} x_k P(X = x_k) \\
&\geq a \sum_{k, x_k \geq a} P(X = x_k) \\
&\geq aP(X \geq a).
\end{aligned}$$

Il s'ensuit que

$$P(X \geq a) \leq \frac{m}{a}, \forall a > 0.$$

5.10 Inégalité de Bienaymé-Tchebychev

Soit la variable aléatoire réelle discrète X, d'espérance mathématique $E(X) = m$ et de variance $V(X) = \sigma^2$. Alors,

$$P(|X - m| > \lambda\sigma) \leq \frac{1}{\lambda^2}.$$

Démonstration

Cette inégalité de Bienaymé-Tchebychev provient directement de celle de Markov.
Par ailleurs, si on suppose l'inégalité de Markov non connue, l'inégalité de Bienaymé-Tchebychev peut se montrer comme suit :
$V(X) = \sigma^2 = E((X-m)^2) = \sum_k (x_k - m)^2 P(X = x_k)$.
Donc,

$$\begin{aligned}
\forall \varepsilon > 0, \sigma^2 &= \sum_{k, |x_k - m| > \varepsilon} (x_k - m)^2 P(X = x_k) + \sum_{k, |x_k - m| \leq \varepsilon} (x_k - m)^2 P(X = x_k) \\
&\geq \sum_{k, |x_k - m| > \varepsilon} (x_k - m)^2 P(X = x_k) \\
&\geq \varepsilon^2 \sum_{k, |x_k - m| > \varepsilon} P(X = x_k) \\
&\geq \varepsilon^2 P(|X - m| > \varepsilon).
\end{aligned}$$

On pose alors $\varepsilon = \lambda\sigma$, et il s'ensuit que

$$P(|X-m| > \lambda\sigma) \leq \frac{\sigma^2}{\lambda^2\sigma^2} = \frac{1}{\lambda^2}.$$

5.11 Enonçé de la loi faible des grands nombres

Soient $X_1, X_2, \ldots, X_n$, n variables aléatoires réelles discrètes indépendantes de même loi. Si on note $E(X_i) = m$, alors :

$$\lim_{n\to\infty} P(\left|\frac{S_n}{n} - m\right| > \varepsilon) = 0,$$

pour tout $\varepsilon > 0$, avec $S_n = X_1 + X_2 + \ldots + X_n$.

Démonstration

On pose $V(X_i) = \sigma^2,\ \forall i \in \{1, 2, \ldots, n\}$. Dans ce cas, $E(\frac{S_n}{n}) = m$ et $V(\frac{S_n}{n}) = \frac{\sigma^2}{n}$.
Ainsi, d'après l'inégalité de Bienaymé-Tchebychev, il résulte :

$$P(\left|\frac{S_n}{n} - m\right| > \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2}.$$

Puis,

$$\lim_{n\to\infty} P(\left|\frac{S_n}{n} - m\right| > \varepsilon) = 0,$$

pour tout $\varepsilon > 0$, avec $S_n = X_1 + X_2 + \ldots + X_n$

6 Lois discrètes usuelles

6.1 La loi constante

6.1.1 Définition

La variable aléatoire X constante ne peut prendre qu'une seule valeur, a, et on a $P(X = a) = 1$.

6.1.2 Caractéristiques

$E(X) = a$.
$V(X) = 0$.

6.2 La variable de Bernoulli (ou loi de Bernoulli)

C'est un modèle simple et introductif qui premet d'introduire des lois plus complexes.

6.2.1 Définition

La variable aléatoire X de Bernoulli peut prendre deux valeurs, 0 et 1, et P (qui dépend du paramètre p –probabilité de succès–) est représentée dans le tableau suivant :

x_i	0	1
$P(X = x_i)$	$q = 1 - p$	p

6.2.2 Caractéristiques

$E(X) = p$ et $V(X) = p(1-p)$.

Démonstration

$E(X) = \sum_{i=0}^{1} x_i P(X = x_i) = 0 + p = p$.

$V(X) = \sum_{i=0}^{1} (x_i - p)^2 P(X = x_i) = (-p)^2(1-p) + (1-p)^2 p = p(1-p)$.

Remarque

$G_X(t) = (1-p) + pt$.

6.2.3 Exemples typiques

Le lancer d'une pièce de monnaie ; l'extraction d'une boule dans une urne ne contenant que deux types de boules ; la réponse à une question d'un vrai ou faux ; ...

6.3 La loi binomiale

C'est un modèle composé de n lois de Bernoulli indépendantes. Nous avons déjà rencontré cette loi dans les épreuves itératives.

6.3.1 Définition

La variable aléatoire X binomiale peut prendre $n+1$ valeurs, $0, 1, \ldots, n$, et P (qui dépend des paramètres n –nombre d'essais– et p –probabilité de succès ($q = 1-p$)–) est donnée par :

$$P(X = k) = C_n^k p^k q^{n-k}.$$

On a bien

$$\sum_{k=0}^{n} P(X = k) = \sum_{k=0}^{n} C_n^k p^k q^{n-k} = (p+q)^n = 1.$$

X suit une loi binomiale de paramètres n et p est noté $X \hookrightarrow B(n, p)$. Ainsi, X suit une loi de Bernoulli de paramètre p est noté $X \hookrightarrow B(1, p)$.

6.3.2 Caractéristiques

$E(X) = np$ et $V(X) = np(1-p) = npq$.

Démonstration

$X = X_1 + X_2 + \ldots + X_n$ où X_i est une variable de Bernoulli dont la probabilité s'écrit dans le tableau suivant :

x	0	1
$P(X_i = x)$	$q = 1-p$	p

$E(X) = \sum_{i=1}^{n} E(X_i) = np$ car $E(X_i) = p$

et

$$\begin{aligned} V(X) &= \sum_{i=1}^{n} V(X_i) \text{ car les variables sont indépendantes} \\ &= npq = np(1-p) \text{ car } V(X_i) = pq = p(1-p). \end{aligned}$$

Remarque

$G_X(t) = ((1-p) + pt)^n$.

Démonstration

$$\begin{aligned} G_X(t) &= \sum_{k=0}^{n} t^k C_n^k p^k (1-p)^{n-k} \\ &= \sum_{k=0}^{n} C_n^k (pt)^k (1-p)^{n-k} \\ &= ((1-p)+pt)^n. \end{aligned}$$

6.3.3 Exemples typiques

Le lancer de n pièces de monnaie (on s'intéresse au nombre d'obtentions de pile, ...) ; le tirage avec remise de n boules dans une urne ne contenant que deux types de boules (on s'intéresse au nombre de boules extraites d'un même type, ...) ; ...

6.3.4 La loi binomiale des fréquences relatives

$X \hookrightarrow B(n,p)$. On note alors F_n la variable aléatoire $F_n = \frac{X}{n}$. On a alors : $E(F_n) = p$ et $V(F_n) = pq/n = p(1-p)/n$.
Ainsi, $\lim\limits_{n\to\infty} E(F_n) = p$ et $\lim\limits_{n\to\infty} V(F_n) = 0$.

En augmentant le nombre de tirages, la fréquence F_n va tendre vers la proportion p dans un schéma de type d'urne. Ce résultat est un cas particulier de la loi des grands nombres.

6.3.5 Additivité de deux variables aléatoires binomiales indépendantes

On suppose que $X \hookrightarrow B(n,p)$ et que $Y \hookrightarrow B(n',p)$.
Lorsque X et Y sont indépendantes, $X+Y \hookrightarrow B(n+n',p)$.

Démonstration

$$\begin{aligned} P(X+Y=k+l) &= \sum_{j=0}^{k+l} P(X=j \bigcap Y=k+l-j) \\ &= \sum_{j=0}^{k+l} P(X=j)P(Y=k+l-j) \text{ (\textit{car les variables sont indépendantes})} \\ &= \sum_{j=0}^{k+l} C_n^j p^j q^{n-j} C_{n'}^{k+l-j} p^{k+l-j} q^{n'-k-l+j} \text{ (\textit{car les variables sont binomiales})} \\ &= p^{k+l} q^{n+n'-(k+l)} \sum_{j=0}^{k+l} C_n^j C_{n'}^{k+l-j} \\ &= C_{n+n'}^{k+l} p^{k+l} q^{n+n'-(k+l)} \text{ (\textit{car} } \sum_{j=0}^{k+l} C_n^j C_{n'}^{k+l-j} = C_{n+n'}^{k+l}, \\ &\quad \text{\textit{par récurrence sur} } n+n' \text{ \textit{– voir} 6.4.1–)}, \end{aligned}$$

et donc, $X+Y$ suit une loi binomiale de paramètres $n+n', p$.

$\boxed{\textbf{Exercice 61}}$ [9, exercice 5, p. 133] Déterminer la (ou les) valeur (ou valeurs) la (ou les) plus probable (ou probables) de X lorsque X suit une loi binomiale de paramètres n et p.

$\boxed{\textbf{Exercice 62}}$ On considère un lot de 40 pièces. La probabilité pour qu'une pièce soit défectueuse est $p = 0,09$. Calculez les probabilités

– de n'avoir aucune pièce défectueuse dans le lot
(Réponse : $0,91^{40} = 0,023...$)
– d'avoir au plus deux pièces défectueuses dans le lot
(Réponse : $0,91^{40} + C_{40}^{1}0,09.0,91^{39} + C_{40}^{2}0,09^{2}.0,91^{38} = 0,289...$)

Sur ces 40 pièces, à votre avis, combien vont être défectueuses ?
(Réponse : $E(X) = 40 \times 0,09 = 3,6$, et donc on doit s'attendre à ce que 3 ou 4 d'entre elles soient défectueuses)

6.4 La loi hypergéométrique

C'est un modèle composé de n lois de Bernoulli dépendantes. Nous l'avons déjà rencontrée dans les épreuves itératives.

6.4.1 Définition

La variable aléatoire X hypergéométrique peut prendre $n+1$ valeurs, $0, 1, \dots, n$, et P (qui dépend des paramètres N –population totale–, n –nombre d'échantillons dans la population totale– et p –probabilité de succès ($q = 1-p$)– tel que $Np \in \mathbb{N}$) est donnée par :

$$P(X=k) = \frac{C_{Np}^{k} C_{Nq}^{n-k}}{C_N^n}.$$

Montrons que,

$$\sum_{k=0}^{n} P(X=k) = \sum_{k=0}^{n} \frac{C_{Np}^{k} C_{Nq}^{n-k}}{C_N^n} = 1.$$

Ceci revient à montrer que

$$\sum_{p=0}^{m} C_n^p C_{N-n}^{m-p} = C_N^m.$$

Par récurrence sur N,

– Pour $N = n$, $\sum_{p=0}^{m} C_n^p C_0^{m-p} = C_n^m = C_N^m$.

–

$$\begin{aligned}
\sum_{p=0}^{m} C_n^p C_{N+1-n}^{m-p} &= \sum_{p=0}^{m} C_n^p (C_{N-n}^{m-p-1} + C_{N-n}^{m-p}) \\
&= \sum_{p=0}^{m} C_n^p C_{N-n}^{m-p-1} + \sum_{p=0}^{m} C_n^p C_{N-n}^{m-p} \\
&= \sum_{p=0}^{m-1} C_n^p C_{N-n}^{m-p-1} + \sum_{p=0}^{m} C_n^p C_{N-n}^{m-p} \\
&= C_N^{m-1} + C_N^m \\
&= C_{N+1}^m
\end{aligned}$$

X suit une loi hypergéométrique de paramètres N, n et$p$ est noté $X \hookrightarrow H(N, n, p)$.

6.4.2 Caractéristiques

$E(X) = np$ et $V(X) = np(1-p)\frac{N-n}{N-1} = npq\frac{N-n}{N-1}$.
$\frac{N-n}{N-1}$ est appelé coefficient d'exhaustivité.

Démonstration

$$
\begin{aligned}
E(X) &= \sum_{k=0}^{n} kP(X=k) \\
&= \sum_{k=0}^{n} k\frac{C_{Np}^{k}C_{Nq}^{n-k}}{C_N^n} \\
&= \frac{Np}{C_N^n}\sum_{k=1}^{n} \frac{k}{Np}\frac{(Np)!}{k!(Np-k)!}C_{Nq}^{n-k} \\
&= \frac{Np}{C_N^n}\sum_{k=1}^{n} \frac{(Np-1)!}{(k-1)!(Np-k)!}C_{Nq}^{n-k} \\
&= \frac{Np}{C_N^n}\sum_{m=0}^{n-1} C_{Np-1}^{m}C_{Nq}^{n-m-1} \ (avec\ m=k-1) \\
&= \frac{Np}{C_N^n}C_{N-1}^{n-1} \ (car\ \sum_{m=0}^{n-1}\frac{C_{Np-1}^{m}C_{Nq}^{n-m-1}}{C_{N-1}^{n-1}}=1) \\
&= Np\frac{n}{N} \\
&= np.
\end{aligned}
$$

$$
\begin{aligned}
E(X(X-1)) &= \sum_{k=0}^{n} k(k-1)P(X=k) \\
&= \sum_{k=0}^{n} k(k-1)\frac{C_{Np}^{k}C_{Nq}^{n-k}}{C_N^n} \\
&= \frac{Np(Np-1)}{C_N^n}\sum_{k=2}^{n} \frac{k(k-1)}{Np(Np-1)}\frac{(Np)!}{k!(Np-k)!}C_{Nq}^{n-k} \\
&= \frac{Np(Np-1)}{C_N^n}\sum_{k=2}^{n} \frac{(Np-2)!}{(k-2)!(Np-k)!}C_{Nq}^{n-k} \\
&= \frac{Np(Np-1)}{C_N^n}\sum_{k=2}^{n} C_{Np-2}^{k-2}C_{Nq}^{n-k} \\
&= \frac{Np(Np-1)}{C_N^n}\sum_{m=0}^{n-2} C_{Np-2}^{m}C_{Nq}^{n-m-2} \quad (avec\ m=k-2) \\
&= \frac{Np(Np-1)}{C_N^n}C_{N-2}^{n-2} \ (car\ \sum_{m=0}^{n-2}\frac{C_{Np-2}^{m}C_{Nq}^{n-m-2}}{C_{N-2}^{n-2}}=1) \\
&= Np(Np-1)\frac{n(n-1)}{N(N-1)} \\
&= np(Np-1)\frac{n-1}{N-1}
\end{aligned}
$$

Or

$$\begin{aligned}
V(X) &= E(X^2) - (E(X))^2 \\
&= E(X(X-1)) + E(X) - (E(X))^2 \\
&= np(Np-1)\frac{n-1}{N-1} + np - n^2p^2 \\
&= \frac{np}{N-1}((Np-1)(N-1) + N - 1 - Nnp + np) \\
&= \frac{np}{N-1}(N - n - p(N-n)) \\
&= \frac{np(N-n)(1-p)}{N-1}
\end{aligned}$$

Donc $V(X) = np(1-p)\frac{N-n}{N-1} = npq\frac{N-n}{N-1}$.

6.4.3 Exemples typiques

Le tirage sans remise de n boules dans une urne ne contenant que deux types de boules (on s'intéresse au nombre de boules extraites d'un même type, ...) ; ...

6.4.4 Approximation de la loi hypergéométrique par la loi binomiale

La loi binomiale approche la loi hypergéométrique lorsque N tend vers l'infini, pour $p \in [\rho, 1-\rho]$ où $\rho > 0$.

On suppose donc que $X_N \hookrightarrow H(N, n, p)$ et donc $P(X_N = k) = \frac{C_{Np}^k C_{Nq}^{n-k}}{C_N^n}$. Nous nous proposons de montrer que sous les conditions précédentes,

$$P(X_N = k) = \frac{C_{Np}^k C_{Nq}^{n-k}}{C_N^n} \sim_{N\to\infty} C_n^k p^k q^{n-k} = P(X = k).$$

Calcul de $\frac{P(X_N=k)}{P(X=k)}$:

$$\begin{aligned}
\frac{P(X_N = k)}{P(X = k)} &= \frac{C_{Np}^k C_{Nq}^{n-k}}{C_N^n C_n^k p^k q^{n-k}} \\
&= \frac{(Np)!(Nq)!n!(N-n)!(n-k)!k!}{(Np-k)!(Nq-n+k)!k!(n-k)!N!n!p^k q^{n-k}} \\
&= \frac{(\frac{Np}{p})(\frac{Np-1}{p})\dots(\frac{Np-k+1}{p})(\frac{Nq}{q})(\frac{Nq-1}{q})\dots(\frac{Nq-n+k+1}{q})}{N(N-1)\dots(N-n+1)}
\end{aligned}$$

Donc

$$P(X_N = k) \sim_{N\to\infty} P(X = k)$$

car le précédent quotient tend vers 1 lorsque N tend vers l'infini.

$\boxed{\textbf{Exercice 63}}$ Un électricien achète des composants par paquets de 10. Sa technique de contrôle est de n'examiner que trois des composants, tirés au hasard dans le paquet et de n'acheter le lot des 10 paquets que si les trois composants examinés sont sans défaut. Si 5 pourcents des paquets contiennent deux composants à malfaçon, si 25 pourcents n'en contiennent qu'un et si 70 pourcents n'en contiennent aucun, quelle est la probabilité que l'électricien achète un paquet.

Solution : On note A l'événement "l'électricien achète un paquet" et B_i l'événement "le paquet contient i composants à malfaçon". On a

$$\begin{aligned} P(A) &= P(A/B_2)P(B_2) + P(A/B_1)P(B_1) + P(A/B_0)P(B_0) \\ &= \frac{C_8^3 C_2^0}{C_{10}^3} 0,05 + \frac{C_9^3 C_1^0}{C_{10}^3} 0,25 + \frac{C_{10}^3 C_0^0}{C_{10}^3} 0,70 \\ &= \frac{539}{600} \end{aligned}$$

6.5 La loi géométrique

6.5.1 Définition

La variable aléatoire X géométrique peut prendre une infinité de valeurs, $1, 2, \ldots$, et P (qui dépend du paramètre p –probabilité de succès $(q = 1 - p)$–) est donnée par :

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

On a bien

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

X suit une loi géométrique de paramètre p est noté $X \hookrightarrow G(p)$.

6.5.2 Caractéristiques

$E(X) = \frac{1}{p}$ et $V(X) = \frac{q}{p^2} = \frac{1-p}{p^2}$.

Démonstration

$E(X) = \sum_{k \geq 1} kq^{k-1}p = p \sum_{k \geq 0} kq^{k-1} = \frac{p}{(1-q)^2} = \frac{1}{p}$.

Pour calculer $V(X)$, on calcule tout d'abord $E(X(X-1))$.

$$\begin{aligned} E(X(X-1)) &= \sum_{k \geq 1} k(k-1)P(X = k) \\ &= \sum_{k \geq 0} k(k-1)q^{k-1}p \\ &= 2\frac{pq}{(1-q)^3} \\ &= 2\frac{q}{p^2} \end{aligned}$$

Or

$$\begin{aligned} V(X) &= E(X^2) - (E(X))^2 \\ &= E(X(X-1)) + E(X) - (E(X))^2 \\ &= 2\frac{q}{p^2} + \frac{1}{p} - \frac{1}{p^2} \end{aligned}$$

Donc $V(X) = \frac{q}{p^2} = \frac{1-p}{p^2}$.

$G_X(t) = \frac{p}{1-p} \frac{1}{1-(1-p)t}$.

Démonstration

$$\begin{aligned} G_X(t) &= \sum_{k \geq 0} (1-p)^{k-1} p t^k \\ &= \frac{p}{1-p} \sum_{k \geq 0} ((1-p)t)^k \end{aligned}$$

6.5.3 Exemples typiques

Le tirage avec remise de n boules dans une urne ne contenant que deux types de boules (on s'intéresse à l'indice de la première obtention d'une boule d'un certain type) ; ...

Exercice 64 On tire avec remise une boule dans une urne contenant 113 boules blanches et 7 boules noires. A priori, combien devra-t-on effectuer de tirages pour obtenir une boule noire pour la première fois ?

Solution : Si p est la proportion de boules noires, $p = \frac{7}{120}$. D'après la formule précédente sur l'espérance mathématique, $E(X) = \frac{1}{p} = \frac{120}{7}$ (où X est implicitement la variable géométrique propre à l'expérience en question), et donc, il faudra s'attendre à exécuter entre 17 et 18 tirages pour voir apparaître pour la première fois une boule noire.

6.6 La loi de Poisson

6.6.1 Définition

La variable aléatoire X de Poisson peut prendre une infinité de valeurs, $0, 1, \ldots$, et P (qui dépend du paramètre λ) est donnée par :

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

On a bien

$$\sum_{k \geq 0} P(X = k) = \sum_{k \geq 0} \frac{\lambda^k e^{-\lambda}}{k!} = 1.$$

X suit une loi de Poisson de paramètre λ est noté $X \hookrightarrow P(\lambda)$.

6.6.2 Caractéristiques

$E(X) = \lambda$ et $V(X) = \lambda$.

Démonstration

$$\begin{aligned} E(X) &= \sum_{k \geq 0} k \frac{\lambda^k e^{-\lambda}}{k!} \\ &= e^{-\lambda} \sum_{k \geq 1} \frac{\lambda^k}{(k-1)!} \\ &= e^{-\lambda} \lambda e^{\lambda} \\ &= \lambda. \end{aligned}$$

Pour calculer $V(X)$, on calcule tout d'abord $E(X(X-1))$.

$$\begin{aligned} E(X(X-1)) &= \sum_{k \geq 0} k(k-1) P(X = k) \\ &= \sum_{k \geq 0} k(k-1) \frac{\lambda^k e^{-\lambda}}{k!} \\ &= e^{-\lambda} \sum_{k \geq 0} k(k-1) \frac{\lambda^k}{k!} \\ &= e^{-\lambda} \lambda^2 e^{\lambda} \\ &= \lambda^2. \end{aligned}$$

Or

$$\begin{aligned} V(X) &= E(X^2) - (E(X))^2 \\ &= E(X(X-1)) + E(X) - (E(X))^2 \\ &= \lambda^2 + \lambda - \lambda^2 \end{aligned}$$

Donc $V(X) = \lambda$.

$G_X(t) = e^{\lambda(t-1)}$.

Démonstration

$$\begin{aligned} G_X(t) &= \sum_{k\geq 0} t^k \frac{\lambda^k e^{-\lambda}}{k!} \\ &= e^{-\lambda} \sum_{k\geq 0} \frac{(\lambda t)^k}{k!} \\ &= e^{-\lambda} e^{\lambda t} \\ &= e^{\lambda(t-1)}. \end{aligned}$$

6.6.3 Approximation de la loi de Poisson par la loi binomiale

La loi binomiale approche la loi de Poisson lorsque n tend vers l'infini, pour $\lambda = np_n < \infty$.

On suppose donc que $X_n \hookrightarrow B(n, p_n)$ et donc $P(X_n = k) = C_n^k p_n^k q_n^{n-k}$. Nous nous proposons de montrer que sous les conditions précédentes,

$$P(X_n = k) = C_n^k p_n^k q_n^{n-k} \sim_{n\to\infty} \frac{\lambda^k e^{-\lambda}}{k!} = P(X = k).$$

On a : $C_n^k = \frac{n!}{k!(n-k)!} = \frac{n(n-1)\ldots(n-k+1)}{k!}$

Par ailleurs, $p_n^k = \frac{(np_n)^k}{n^k} = \frac{\lambda^k}{n^k}$

et $(1-p_n)^{n-k} = (1-\frac{\lambda}{n})^{n-k} = \frac{(1-\frac{\lambda}{n})^n}{(1-\frac{\lambda}{n})^k}$

Donc $P(X_n = k) = C_n^k p_n^k q_n^{n-k} = \frac{n(n-1)\ldots(n-k+1)}{k!} \frac{\lambda^k}{n^k} \frac{(1-\frac{\lambda}{n})^n}{(1-\frac{\lambda}{n})^k}$.

Or $\lim_{n\to\infty} \frac{n(n-1)\ldots(n-k+1)}{n^k} = 1$,

$\lim_{n\to\infty} (1-\frac{\lambda}{n})^k = 1$

et $(1-\frac{\lambda}{n})^n \sim_{n\to\infty} e^{-\lambda}$ (par passage au logarithme).

Et donc $P(X_n = k) \sim_{n\to\infty} P(X = k)$ lorsque $\lambda = np_n < \infty$.

6.6.4 Additivité de deux variables aléatoires de Poisson indépendantes

On suppose que $X \hookrightarrow P(\lambda)$ et que $Y \hookrightarrow P(\lambda')$.

Lorsque X et Y sont indépendantes, $X + Y \hookrightarrow P(\lambda + \lambda')$.

Démonstration

$$\begin{aligned}
P(X+Y=k+l) &= \sum_{j=0}^{k+l} P(X=j \bigcap Y=k+l-j) \\
&= \sum_{j=0}^{k+l} P(X=j)P(Y=k+l-j) \ (\textit{car les variables sont indépendantes}) \\
&= \sum_{j=0}^{k+l} \frac{\lambda^j e^{-\lambda}}{j!} \frac{\lambda'^{k+l-j} e^{-\lambda'}}{(k+l-j)!} \ (\textit{car les variables sont de Poisson}) \\
&= e^{-(\lambda+\lambda')} \sum_{j=0}^{k+l} \frac{\lambda^j}{j!} \frac{\lambda'^{k+l-j}}{(k+l-j)!} \\
&= e^{-(\lambda+\lambda')} \frac{(\lambda+\lambda')^{k+l}}{(k+l)!} \ (\textit{car } (\lambda+\lambda')^{k+l} = \sum_{j=0}^{k+l} C_{k+l}^j \lambda^j \lambda'^{k+l-j}),
\end{aligned}$$

et donc, $X+Y$ suit une loi de Poisson de paramètre $\lambda+\lambda'$.

Exercice 65 [9, exercice 3, p. 156] Soit X une variable aléatoire suivant une loi de Poisson de paramètre λ.

Déterminer le mode de X, c'est-à-dire la valeur la plus probable de X (on distinguera les cas $\lambda \in \mathbb{N}$ et $\lambda \notin \mathbb{N}$).

Exercice 66 [13] Dans un hôtel, il arrive en moyenne 1,25 personne par 10 minutes, entre 15h et 21h. On prend pour variable aléatoire X le nombre de personnes arrivant dans cet hôtel en 10 minutes, dans cet horaire particulier. On admet que X suit une loi de Poisson.

- Déterminez la probabilité pour qu'en 10 minutes, il arrive k personnes.
  Réponse : $P(X=k) = \frac{1{,}25^k e^{-1{,}25}}{k!}$ car la moyenne donne la valeur du paramètre λ.
- Déterminez la probabilité pour qu'en 10 minutes, il arrive 2 personnes.
  Réponse : $P(X=2) = \frac{1{,}25^2 e^{-1{,}25}}{2!} = 0,22...$
- Déterminez la probabilité pour qu'en 10 minutes, il arrive 4 personnes au plus.
  Réponse : $P(X=0)+P(X=1)+P(X=2)+P(X=3)+P(X=4)$
  $= \frac{e^{-1{,}25}}{0!} + \frac{1{,}25e^{-1{,}25}}{1!} + \frac{1{,}25^2 e^{-1{,}25}}{2!} + \frac{1{,}25^3 e^{-1{,}25}}{3!} + \frac{1{,}25^4 e^{-1{,}25}}{4!} = 0,99...$
- Déterminez la probabilité pour qu'en 10 minutes, il arrive 3 personnes au moins.
  Réponse : $1 - P(X=0) - P(X=1) - P(X=2) = 1 - \frac{e^{-1{,}25}}{0!} - \frac{1{,}25e^{-1{,}25}}{1!} - \frac{1{,}25^2 e^{-1{,}25}}{2!} = 0,13...$

Exercice 67 [13] La probabilité pour qu'une personne soit allergique au médicament M est $p = 10^{-3}$. On considère un échantillon de 1000 personnes.

Soit X la variable aléatoire dont la valeur est le nombre de personnes allergiques dans l'échantillon.

- Déterminez la loi de probabilité de X.
  Réponse : $P(X=k) = C_n^k p^k (1-p)^{n-k}$.
- En utilisant l'approximation justifiée, calculez la probabilité que 2 personnes soient allergiques dans l'échantillon, puis celle qu'au moins deux personnes soient allergiques dans l'échantillon.
  Réponse :
  $P(X=k) = \frac{1^k e^{-1}}{k!} = \frac{1}{e.k!}$ car on pose $\lambda = np = 1$.
  $P(X=2) = \frac{1}{2e} = 0,18...$
  $P(X \geq 2) = 1 - P(X=0) - P(X=1) = 1 - \frac{1}{e} - \frac{1}{e} = 1 - \frac{2}{e} = 0,26...$

6.7 La loi binomiale négative

6.7.1 Définition

La variable aléatoire X binomiale négative peut prendre une infinité de valeurs, $r, r+1, \ldots$, et P (qui dépend des paramètres p –probabilité de succès $(p = 1-q)$– et r) est donnée par :

$$P(X=k) = C_{k-1}^{r-1} q^{k-r} p^r.$$

On a bien

$$\sum_{k \geq r} P(X=k) = \sum_{k \geq r} C_{k-1}^{r-1} q^{k-r} p^r = 1.$$

En effet, si on pose $g(q) = \frac{1}{1-q} = \sum_{l \geq 1} q^{l-1}$, on obtient par dérivations successives que

$$\begin{aligned} g^{(r-1)}(q) = \frac{(r-1)!}{(1-q)^r} &= \sum_{l \geq 1} (l-1)(l-2)\ldots(l-r+1) q^{l-r} \\ &= \sum_{l \geq r} (l-1)(l-2)\ldots(l-r+1) q^{l-r} \\ &= \sum_{l \geq r} \frac{(l-1)!}{(l-r)!} q^{l-r}, \end{aligned}$$

et donc $\frac{1}{p^r} = \frac{1}{(1-q)^r} = \sum_{l \geq r} C_{l-1}^{r-1} q^{l-r}$.

X suit une loi binomiale négative de paramètres p, r est noté $X \hookrightarrow BN(p, r)$.

En particulier, $X \hookrightarrow BN(p, 1) \iff X \hookrightarrow G(p)$.

6.7.2 Caractéristiques

$E(X) = \frac{r}{p}$ et $V(X) = \frac{rq}{p^2}$.

Démonstration

$X = X_1 + X_2 + \ldots + X_r$ où $X_i \hookrightarrow G(p)$ et où les X_i sont indépendantes (on rappelle $E(X_i) = \frac{1}{p}$ et $V(X_i) = \frac{q}{p^2}$).
Donc, $E(X) = E(X_1) + E(X_2) + \ldots + E(X_r) = \frac{r}{p}$
et $V(X) = V(X_1) + V(X_2) + \ldots + V(X_r) = \frac{rq}{p^2}$.

6.7.3 Exemples typiques

Le tirage avec remise de n boules dans une urne ne contenant que deux types de boules (on s'intéresse à l'indice de la $r^{ème}$ obtention d'une boule d'un certain type) ; ...

Exercice 68 *Le problème des allumettes de Banach.* [10] Un mathématicien se trouve être également fumeur de pipe et il porte à tout moment deux boîtes d'allumettes, une dans chacune de ses poches. Chaque fois qu'il a besoin d'une allumette, il a une chance sur deux d'aller la chercher dans sa poche gauche et autant pour l'autre. Il découvre subittement que la boîte tirée est vide. Les deux boîtes contenaient au départ N allumettes chacune. Quelle est la probabilité p qu'il lui reste k allumettes dans l'autre boîte ?

On désigne par A l'événement "le mathématicien découvre que sa poche droite est vide alors qu'il lui reste k allumettes dans l'autre". Cet événement n'arrive que lorsqu'il choisit la boîte droite pour la $(N+1)^{ème}$ fois lors du $(N+1+N-k)^{ème}$ tirage.
Il s'ensuit que $P(A) = C_{2N-k}^{N} (\frac{1}{2})^{N+1} (\frac{1}{2})^{N-k}$, puis $P(A) = C_{2N-k}^{N} (\frac{1}{2})^{2N+1-k}$.

Comme la probabilité est la même que ce soit sa poche gauche qui soit vide alors qu'il lui reste k allumettes dans la droite,

$$p = 2P(A) = C_{2N-k}^{N}(\frac{1}{2})^{2N-k}.$$

7 Variables aléatoires réelles continues

7.1 Définition

Une variable aléatoire réelle est dite continue si son domaine de variations est $\mathbb{R}$ ou un intervalle de $\mathbb{R}$.

7.1.1 Remarque

Si le système fondamental est fini ou dénombrable, il est assurément impossible que la variable aléatoire réelle soit continue sur $\mathbb{R}$.

7.2 Lois de probabilités

Soit une variable aléatoire réelle continue X, $X : E \longrightarrow \mathbb{R}$ prenant ses valeurs dans $\mathbb{R}$.

On définit la probabilité d'un événement de type $[a, b[$ comme la probabilité que la variable X prenne sa valeur dans $[a, b[$.

La loi de probabilité d'une variable aléatoire réelle continue X est définie par la densité de probabilité f qui est une fonction positive, continue sauf peut-être en un nombre fini ou dénombrable de points et qui est telle que $\int_{-\infty}^{\infty} f(x)dx = 1$.

La probabilité que l'événement soit de type $[a, b[$ est alors donnée par la valeur $\int_a^b f(x)dx = P(X \in [a, b[)$ (si $a \leq b$).

On pourra remarquer que dans le cas d'une variable aléatoire réelle continue, $P(X = x) = 0$.

7.3 Fonction de répartition

La fonction de répartition F d'une variable aléatoire continue X ayant une densité de probabilité f est partout définie et vaut :

$$\begin{array}{rl} F: & \mathbb{R} \longrightarrow \mathbb{R}^+ \\ & x \longrightarrow F(x) = P(X < x) = \int_{-\infty}^{x} f(t)dt \end{array}$$

Propriétés

1. $P(a \leq X < b) = F(b) - F(a) = \int_a^b f(x)dx$
2. F est une fonction croissante car si $x > y$, alors, $F(x) > F(y)$
3. $\lim_{x \to -\infty} F(x) = 0$
4. $\lim_{x \to \infty} F(x) = 1$

Exercice 69 [13] Un système de monitoring est constitué de modules qui peuvent être montés en série ou en parallèle. Le montage est dit *série* si la défaillance d'un seul module entraîne la défaillance de tout le système ; il est dit *parallèle* si la défaillance de tous les modules est nécessaire pour qu'il y ait défaillance de tout le système. La fiabilité (ou probabilité de survie) d'un module est f ; sa probabilité de défaillance est $d = 1 - f$.

Système *série*

- Quelle est la fiabilité F_1 d'un système composé de 10 modules, de fiabilités $f_1 = 0,98$, montés en série ? (Réponse : $F_1 = f_1^{10} = 0,81...$)
- Quelle serait la fiabilité F_2 d'un système composé de 9 modules de fiabilités f_1 et d'1 module de fiabilité $f_2 = 0,60$, montés en série ? (Réponse : $F_2 = f_1^9 f_2 = 0,50...$)
- Quelle valeur devrait avoir la fiabilité f_3 de 10 modules montés en série pour que la fiabilité du système composé de ces modules, soit supérieure ou égale à $0,98$? (Réponse : $f_3^{10} \geq 0,98$, d'où $f_3 \geq 0,99...$)
- 10 modules sont choisis au hasard dans un lot de fabrication dont on sait que 96 pourcents des modules ont une fiabilité f_1 et les autres ont une fiabilité $f_4 = 0,90$.
  - Quelle est la probabilité p_1 d'obtenir un système *série* de fiabilité maximale ? (Réponse : $p_1 = 0,96^{10} = 0,66...$)
  - Quelle est la probabilité p_2 d'obtenir un système *série* de fiabilité minimale ? (Réponse : $p_2 = 0,04^{10} = 0,00...$)
  - Un contrôle de fabrication permettrait de diminuer le pourcentage des modules de fiabilités f_4. A quel niveau de contrôle doit-il amener ce pourcentage t pour que la probabilité p_3 d'obtenir un système *série* de fiabilité maximale soit supérieure à $0,95$? (Réponse : $t^{10} \geq 0,95$ d'où $t \geq 0,99...$)

Système *parallèle*

- Quelle est la fiabilité F_3 d'un système composé de 10 modules, de fiabilités f_1, montés en parallèle ? (Réponse : $F_3 = 1 - (1 - f_1)^{10} = 0,99...$)
- Quelle serait la fiabilité F_4 d'un système composé de 9 modules de fiabilités f_1 et d'1 module de fiabilité f_2, montés en parallèle ? (Réponse : $F_4 = 1 - (1 - f_1)^9(1 - f_2) = 0,99...$)
- Quelle valeur devrait avoir la fiabilité f_5 de 10 modules montés en parallèle pour que la fiabilité du système composé de ces modules, soit supérieure ou égale à $0,98$? (Réponse : $1 - (1 - f_5)^{10} \geq 0,98$, d'où $f_5 \geq 0,32...$)

Protection par redondance Afin d'assurer le bon fonctionnement du monitoring, on envisage de monter en parallèle des ensembles (E) de modules montés en série.

- Quelle serait la fiabilité F_5 d'un montage en parallèle de deux ensembles E_1 constitués de dix modules de même fiabilité f_1 montés en série ? (Réponse : $F_5 = 1 - (1 - f_1^{10})^2 = 0,96...$)
- Quelle serait la fiabilité F_6 d'un montage en parallèle de deux ensembles E_2 constitués de dix modules de même fiabilité f_4 montés en série ? (Réponse : $F_6 = 1 - (1 - f_4^{10})^2 = 0,57...$)
- Combien d'ensembles E_2 devrait-on monter en parallèle pour que la fiabilité de ce montage soit au moins égale à F_5 ? (Réponse : $1 - (1 - f_4^{10})^n \geq F_5$ d'où $n \geq 8$)

Evolution temporelle d'une fiabilité : modèle exponentiel Dans ce modèle, on admet que la probabilité de défaillance d'un module dans l'intervalle de temps $[t, t + \Delta t[$, sachant qu'elle ne survient pas dans l'intervalle de temps $[0, t]$, est proportionnelle à la durée Δt de cet intervalle, c'est-à-dire de la forme $g\Delta t$, expression dans laquelle g est une constante de proportionalité.

- Soit $D(t)$ la fonction de répartition temporelle des probabilités de défaillance. Montrer que $D(t + \Delta t) - D(t) = (1 - D(t))g\Delta t$. (Réponse : $P(t \leq T < t + \Delta t) = D(t + \Delta t) - D(t)$ où T est la variable aléatoire donnant l'instant où la défaillance survient)
- En faisant tendre Δt vers 0, trouver une équation différentielle à variables séparables dont on déduira $D(t)$ et la fonction de densité $d(t)$ correspondante. (Réponse : $\frac{dD}{dt} = (1 - D)g$; $D(t) = 1 - e^{-gt}$ et $d(t) = ge^{-gt}$)
- Déduire de $D(t)$, l'expression de la fonction de répartition temporelle $F(t)$ des fiabilités. Calculer la probabilité d'une fiabilité de 1 an en supposant $g = 0,02/mois$. (Réponse : $F(t) = e^{-gt}$; $F(1an) = 0,78...$)
- Calculer le temps moyen $\overline{T}$ des défaillance et en déduire le temps moyen des fiabilités. (Réponse : $\overline{T} = \frac{1}{g} = 50mois$; $\overline{T} = E(F)$)
- Calculer la probabilité d'observer une défaillance entre 1 et 2 ans. (Réponse : $D(2ans) - D(1an) = 0,16...$)
- Exprimer, en fonction du temps, la fiabilité d'un système *série*, puis d'un système *parallèle*, constitués de 10 modules. Calculer la probabilité de survie à 1 an de chacun de ces systèmes. (Réponse : $F(t)^{10} = 0,09...$ *série* ;

$1-(1-F(t))^{10}=0,99...$ *parallèle)*

7.4 Les moments d'ordre n

On appelle moment d'ordre n d'une variable aléatoire continue X, et on note $E(X^n)$ la quantité :

$$E(X^n)=\int_{-\infty}^{\infty} t^n f(t)dt.$$

Ce moment d'ordre n n'existe, par définition, que si $\int_{-\infty}^{\infty}|t|^n f(t)dt$ existe.

$\boxed{\textbf{Exercice 70}}$ **Démonstration**.

Soit X une variable aléatoire réelle continue de densité f.

Exprimer la densité de $Y=X^p$, $p\in\mathbb{N}^*$.

Montrer que si $\int_{-\infty}^{\infty}|t|^p f(t)dt$ existe, alors :

$$E(X^p)=\int_{-\infty}^{\infty} t^p f(t)dt.$$

Premier cas : p **impair.**

$P(Y\le y)=P(X\le y^{1/p})=\int_{-\infty}^{y^{1/p}} f(x)dx=\frac{1}{p}\int_{-\infty}^{y} f(u^{1/p})u^{1/p-1}du$ et la densité de Y est $g(y)=\frac{1}{p}f(y^{1/p})y^{1/p-1}$, $\forall y\in\mathbb{R}$.

Il s'ensuit :

$$\begin{aligned}E(Y) &= \frac{1}{p}\int_{-\infty}^{\infty} uf(u^{1/p})u^{1/p-1}du\\ &= \frac{1}{p}\int_{-\infty}^{\infty} f(u^{1/p})u^{1/p}du\\ &= \int_{-\infty}^{\infty} t^p f(t)dt \qquad u=t^p.\end{aligned}$$

Second cas : p **pair.**

$P(Y\le y)=P(-y^{1/p}\le X\le y^{1/p})=\int_0^{y^{1/p}}(f(x)+f(-x))dx=\frac{1}{p}\int_0^y[f(u^{1/p})+f(-u^{1/p})]u^{1/p-1}du$ et la densité de Y est $g(y)=\frac{1}{p}[f(y^{1/p})+f(-y^{1/p})]y^{1/p-1},\ \forall y\in\mathbb{R}^+$ et $g(y)=0,\ \forall y\in\mathbb{R}^-$.

Il s'ensuit :

$$\begin{aligned}E(Y) &= \frac{1}{p}\int_0^{\infty} u[f(u^{1/p})+f(-u^{1/p})]u^{1/p-1}du\\ &= \frac{1}{p}\int_0^{\infty}[f(u^{1/p})+f(-u^{1/p})]u^{1/p}du\\ &= \int_0^{\infty} t^p[f(t)+f(-t)]dt \qquad u=t^p\\ &= \int_{-\infty}^{\infty} t^p f(t)dt.\end{aligned}$$

Application directe : Si X a pour densité $f(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$, quelle est la densité de X^2 ?

Si $y>0, g(y)=\frac{1}{\sqrt{2\pi y}}e^{-y/2}$.

Si $y\le 0, g(y)=0$.

7.5 L'espérance mathématique

Lorsque la quantité $\int_{-\infty}^{\infty} |x|f(x)dx$ est définie, on appelle espérance mathématique d'une variable aléatoire continue X, notée $E(X)$, la quantité :

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx.$$

L'espérance mathématique est donc le moment d'ordre 1.

7.5.1 Propriétés de l'espérance mathématique

- Pour toute constante a, $E(a) = a$ (trivial)
- Pour toute constante a, $E(aX) = aE(X)$ (trivial)
- $E(X + Y) = E(X) + E(Y)$ (démonstration hors programme)
  De plus, on a le théorème suivant : **Théorème** (admis)
  Si X et Y sont deux variables aléatoires réelles continues indépendantes de densités respectives $f(x)$ et $g(x)$, alors, $X + Y$ est une variable aléatoire réelle continue de densité $h(x) = \int_{-\infty}^{\infty} f(x - t)g(t)dt$.
- $E(aX + bY) = aE(X) + bE(Y)$ (trivial)
- Lorsque X et Y sont indépendantes, on a $E(XY) = E(X)E(Y)$ (démonstration hors programme)

7.5.2 Variable centrée

La définition est exactement la même que pour les variables aléatoires réelles discrètes.

7.6 Variance et écart-type

Lorsque les quantités $\int_{-\infty}^{\infty} |x|f(x)dx$ et $\int_{-\infty}^{\infty} |x|^2 f(x)dx$ sont définies, on appelle variance d'une variable aléatoire réelle continue X, et on note $V(X)$ la quantité :

$$V(X) = \int_{-\infty}^{\infty} (x - E(X))^2 f(x)dx.$$

L'écart-type est alors défini par $\sigma(X) = \sqrt{V(X)}$.

7.6.1 Propriétés de la variance

Elles sont identique à celles du cas discret et les démonstrations demeurent valables. Pour rappel,

- $V(X) \geq 0$.
- $V(X) = E((X - E(X))^2)$. (par définition)
- $V(X) = E(X^2) - (E(X))^2$.
- Pour toute constante a, $V(a) = 0$.
- Pour toute constante a, $V(aX) = a^2V(X)$.
- Lorsque X et Y sont indépendantes, on a $V(X + Y) = V(X) + V(Y)$.
- Lorsque X et Y sont indépendantes, on a $V(aX + bY) = a^2V(X) + b^2V(Y)$.

7.6.2 Variable réduite

La définition est strictement la même que pour les variables aléatoires réelles discrètes.

Exercice 71 [13] On suppose que la durée de vie d'un individu est une variable aléatoire continue dont la densité de probabilité f est donnée par

$$\begin{cases} f(t) = kt^2(100-t)^2 & si\ 0 \leq t \leq 100 \\ f(t) = 0 & sinon \end{cases}$$

1. Déterminez k pour que f soit effectivement une densité de probabilité

   Réponse :
   $\int_{-\infty}^{\infty} f(t)dt = 1$
   $\iff \int_0^{100} kt^2(100-t)^2 dt = 1$
   $\iff \int_0^{100} (t^4 - 2.10^2 t^3 + 10^4 t^2)dt = \frac{1}{k}$
   $\iff \left[\frac{t^5}{5} - 10^2\frac{t^4}{2} + 10^4\frac{t^3}{3}\right]_0^{100} = \frac{1}{k}$
   $\iff \frac{10^{10}}{5} - \frac{10^{10}}{2} + \frac{10^{10}}{3} = \frac{1}{k}$
   $\iff \frac{10^{10}}{30} = \frac{1}{k}$
   $\iff k = \frac{3}{10^9}$.

2. Calculez l'espérance mathématique de la durée de vie d'un individu, puis l'écart-type.

   Réponses :
   $E(X) = \int_0^{100} kt^3(100-t)^2 dt$
   $\iff E(X) = k\left[\frac{t^6}{6} - 2.10^2\frac{t^5}{5} + 10^4\frac{t^4}{4}\right]_0^{100}$
   $\iff E(X) = k(\frac{10^{12}}{6} - 2\frac{10^{12}}{5} + \frac{10^{12}}{4})$
   $\iff E(X) = \frac{3}{10^9}\frac{10^{12}}{60} = 50.$
   D'autre part, $E(X^2) = \int_0^{100} kt^4(100-t)^2 dt$
   $\iff E(X^2) = k\left[\frac{t^7}{7} - 2.10^2\frac{t^6}{6} + 10^4\frac{t^5}{5}\right]_0^{100}$
   $\iff E(X^2) = k(\frac{10^{14}}{7} - 2\frac{10^{14}}{6} + \frac{10^{14}}{5})$
   $\iff E(X^2) = \frac{3}{10^9}\frac{10^{14}}{105} = 3\frac{10^5}{105}$.
   Il s'ensuit que $\sigma(X) = \sqrt{V(X)} = \sqrt{E(X^2) - (E(X))^2} = 10\sqrt{\frac{25}{7}}$

3. Calculez la probabilité pour qu'un individu meure entre 30 et 60 ans

   Réponse :
   $P(30 \leq X \leq 60)$
   $= \int_{30}^{60} \frac{3}{10^9} t^2(100-t)^2 dt$
   $= \frac{3}{10^9}\left[\frac{t^5}{5} - 10^2\frac{t^4}{2} + 10^4\frac{t^3}{3}\right]_{30}^{60}$
   $= \frac{3}{10^5}(15066 - 60750 + 63000) = \frac{51948}{10^5}$.

Exercice 72 [9, exercice 16, p. 190] Etant donnés $x_0 > 0$ et $\alpha > 0$.

Soit X une variable aléatoire réelle de fonction de répartition : $F : \mathbb{R} \to \mathbb{R}; x \mapsto F(x)$ tel que $F(x) = 0$ si $x \leq x_0$ et $F(x) = 1 - \frac{x_0^\alpha}{x^\alpha}$ si $x > x_0$.

1. Montrer que X est continue et déterminer la densité f de X.
2. Lorsque $E(X)$ existe (pour quelles valeurs de α ?), calculer $E(X)$.
3. Lorsque $V(X)$ existe (pour quelles valeurs de α ?), calculer $V(X)$.

Exercice 73 [9, exercice 15, p. 190] Soit X une variable aléatoire de densité $f : \mathbb{R} \to \mathbb{R}\ ;\ x \mapsto \dfrac{a}{\pi(a^2+x^2)}$ avec $a \in \mathbb{R}_+^*$. La loi de X est appelée loi de Cauchy de paramètre a.

1. Vérifier que f est bien une densité.
2. Pour quelles valeurs de α la variable $|X|^\alpha$ est-elle intégrable ?
3. Montrer que si X suit une loi de Cauchy de paramètre a, alors $\dfrac{X}{a}$ suit une loi de Cauchy de paramètre 1.

7.7 Inégalité de Markov

Soit la variable aléatoire réelle continue X à valeurs non négatives, d'espérance mathématique $E(X) = m$. Alors,

$$P(X \geq a) \leq \frac{m}{a}, \forall a > 0.$$

Démonstration

$E(X) = m = \int_0^\infty xf(x)dx$.

Donc,

$$\begin{aligned}
\forall a > 0, m &= \int_0^a xf(x)dx + \int_a^\infty xf(x)dx \\
&\geq \int_a^\infty xf(x)dx \\
&\geq a\int_a^\infty f(x)dx \\
&\geq aP(X \geq a).
\end{aligned}$$

Il s'ensuit que

$$P(X \geq a) \leq \frac{m}{a}, \forall a > 0.$$

7.8 Inégalité de Bienaymé-Tchebychev

Soit la variable aléatoire réelle continue X, d'espérance mathématique $E(X) = m$ et de variance $V(X) = \sigma^2$. Alors,

$$P(|X - m| > \lambda\sigma) \leq \frac{1}{\lambda^2}.$$

Démonstration

Cette inégalité de Bienaymé-Tchebychev provient directement de celle de Markov.

Par ailleurs, si on suppose l'inégalité de Markov non connue, l'inégalité de Bienaymé-Tchebychev peut se montrer comme suit :

$V(X) = \sigma^2 = E((X - m)^2) = \int_{-\infty}^\infty (t - m)^2 f(t)dt$.

Donc,

$$\begin{aligned}
\forall \varepsilon > 0, \sigma^2 &= \int_{-\infty}^{m-\varepsilon} (t-m)^2 f(t)dt + \int_{m-\varepsilon}^{m+\varepsilon} (t-m)^2 f(t)dt + \int_{m+\varepsilon}^{\infty} (t-m)^2 f(t)dt \\
&\geq \int_{-\infty}^{m-\varepsilon} (t-m)^2 f(t)dt + \int_{m+\varepsilon}^{\infty} (t-m)^2 f(t)dt \\
&\geq \varepsilon^2 (\int_{-\infty}^{m-\varepsilon} f(t)dt + \int_{m+\varepsilon}^{\infty} f(t)dt) \\
&\geq \varepsilon^2 P(|X - m| > \varepsilon).
\end{aligned}$$

On pose alors $\varepsilon = \lambda\sigma$, et il s'ensuit que

$$P(|X - m| > \lambda\sigma) \leq \frac{\sigma^2}{\lambda^2\sigma^2} = \frac{1}{\lambda^2}.$$

7.9 Enonçé de la loi faible des grands nombres

Soient $X_1, X_2, \ldots, X_n$, n variables aléatoires réelles continues indépendantes de même loi. Si on note $E(x_i) = m$, alors :

$$\lim_{n\to\infty} P(\left|\frac{S_n}{n} - m\right| > \varepsilon) = 0,$$

pour tout $\varepsilon > 0$, avec $S_n = X_1 + X_2 + \ldots + X_n$.

Démonstration

On pose $V(X) = \sigma^2$. Dans ce cas, $E(\frac{S_n}{n}) = m$ et $V(\frac{S_n}{n}) = \frac{\sigma^2}{n}$.
Ainsi, d'après l'inégalité de Bienaymé-Tchebychev, il résulte :

$$P(\left|\frac{S_n}{n} - m\right| > \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2}.$$

Puis,

$$\lim_{n\to\infty} P(\left|\frac{S_n}{n} - m\right| > \varepsilon) = 0,$$

pour tout $\varepsilon > 0$, avec $S_n = X_1 + X_2 + \ldots + X_n$.

8 Lois continues usuelles

8.1 La loi uniforme

8.1.1 Définition

La densité f de la variable aléatoire X uniforme peut prendre deux valeurs, 0 et $\frac{1}{b-a}$ où a et b sont deux constantes données ($0 \leq a \leq b$). Elle est définie par :

$$f(x) = \begin{cases} \frac{1}{b-a} & si\ x \in [a,b] \\ 0 & sinon \end{cases}.$$

On a bien $\int_a^b \frac{1}{b-a} dt = 1$.

La fonction de répartition F de la précédente variable aléatoire X est donnée par

$$F(x) = \begin{cases} \frac{x-a}{b-a} & si\ x \in [a,b] \\ 0 & si\ x < a \\ 1 & si\ x > b \end{cases}.$$

8.1.2 Caractéristiques

$E(X) = \int_a^b \frac{1}{b-a} t dt = \frac{a+b}{2}$ et $V(X) = \int_a^b \frac{1}{b-a}(t - E(X))^2 dt = \frac{a^2+ab+b^2}{3} - \frac{(a+b)^2}{4} = \frac{(b-a)^2}{12}$.

Exercice 74 *Problème de l'aiguille de Buffon.* [10] Sur une table, on trace des lignes parallèles espacées d'un écart D les unes des autres. On y jette une aiguille de longueur L, avec $L \leq D$. Quelle est la probabilité que l'aiguille coupe une ligne (l'alternative étant que l'aiguille soit complètement située dans une des bandes délimitées par les lignes).

Solution.

On repère la position de l'aiguille grâce à la distance entre le centre de celle-ci et la parallèle la plus proche, et grâce à l'angle Θ entre l'aiguille et une perpendiculaire aux lignes. L'aiguille coupe une ligne si $\frac{X}{cos\Theta} < \frac{L}{2}$. La variable X varie entre 0 et $D/2$ et Θ entre 0 et $\pi/2$. On admet que les variables X et Θ sont uniformément réparties et indépendantes. Ainsi,

$$\begin{aligned} P(\frac{X}{cos\Theta} < \frac{L}{2}) &= \frac{4}{\pi D}\int_0^{\frac{\pi}{2}}\int_0^{\frac{L}{2}cos\theta} dxd\theta \\ &= \frac{4}{\pi D}\int_0^{\frac{\pi}{2}} \frac{L}{2}cos\theta d\theta \\ &= \frac{2L}{\pi D}. \end{aligned}$$

Exercice 75 [9, exercice 23, p. 198] Etant données X_1, X_2, X_3 des variables aléatoires indépendantes suivant la loi uniforme "[0,1]".

1. Déterminer la loi de $X_1 + X_2$. Calculer $E(X_1 + X_2)$ et $V(X_1 + X_2)$.
2. Déterminer la loi de $X_1 + X_2 + X_3$. Calculer $E(X_1 + X_2 + X_3)$ et $V(X_1 + X_2 + X_3)$.

Exercice 76 [9, exercice 24, p. 199] Etant données X, Y des variables aléatoires indépendantes suivant la loi uniforme "[0,1]".

Calculer la moyenne de $|X - Y|$.

8.2 La loi de Gamma

8.2.1 Définition

La variable aléatoire X de Gamma est donnée par sa densité f telle que :

$$f(x) = \begin{cases} \frac{1}{\Gamma(a)b^a}x^{a-1}e^{\frac{-x}{b}} & si\ x \geq 0 \\ 0 & sinon \end{cases},$$

où $\Gamma(a) = \displaystyle\int_0^\infty x^{a-1}e^{-x}dx$ et $b > 0$ (on a en particulier $\Gamma(a) = (a-1)\Gamma(a-1)$ et $\Gamma(n+1) = (n)!$, $\forall n \in \mathbb{N}$).

On a bien $\displaystyle\int_0^\infty \frac{1}{\Gamma(a)b^a}x^{a-1}e^{\frac{-x}{b}}dx = 1$.

X suit une loi de Gamma de paramètres a, b est noté $X \hookrightarrow G(a, b)$.

Un cas particulier de cette loi est la loi exponentielle où l'on impose $a = 1$ et $b = \frac{1}{\lambda}$

8.2.2 Caractéristiques

Si $X \hookrightarrow G(a, b)$, $E(X) = ab$ et $V(X) = ab^2$.

On a :

$$\begin{aligned} \int_0^\infty x^a e^{\frac{-x}{b}}dx &= \underbrace{[x^a(-b)e^{\frac{-x}{b}}]_0^\infty}_{=0} - \int_0^\infty ax^{a-1}(-b)e^{\frac{-x}{b}}dx \\ &= ab\int_0^\infty x^{a-1}e^{\frac{-x}{b}}dx \\ & et \\ \int_0^\infty x^{a+1}e^{\frac{-x}{b}}dx &= \underbrace{[x^{a+1}(-b)e^{\frac{-x}{b}}]_0^\infty}_{=0} - \int_0^\infty (a+1)x^a(-b)e^{\frac{-x}{b}}dx \\ &= (a+1)b\int_0^\infty x^a e^{\frac{-x}{b}}dx \\ &= a(a+1)b^2\int_0^\infty x^{a-1}e^{\frac{-x}{b}}dx \end{aligned}$$

Donc,

$$E(X) = \int_0^{\infty} \frac{1}{\Gamma(a)b^a} x^a e^{\frac{-x}{b}}\, dx = ab.$$

$$V(X) = \int_0^{\infty} \frac{1}{\Gamma(a)b^a} x^{a+1} e^{\frac{-x}{b}}\, dx - (E(X))^2 = a(a+1)b^2 - a^2b^2 = ab^2.$$

Il s'ensuit que si $X \hookrightarrow G(1, \frac{1}{\lambda})$, $E(X) = \frac{1}{\lambda}$ et $V(X) = \frac{1}{\lambda^2}$.

Exercice 77 [10] On suppose que la durée d'une conversation téléphonique, mesurée en minutes, est une variable aléatoire exponentielle de paramètre $\lambda = \frac{1}{10}$. J'arrive à une cabine et quelqu'un passe juste avant moi. Avec quelle probabilité dois-je attendre

1. plus de 10 minutes ?
   Réponse : X désigne la durée de la conversation de la personne qui m'a devancé.
   $P(X > 10) = \int_{10}^{\infty} \frac{1}{10} e^{\frac{-x}{10}}\, dx = [-e^{\frac{-x}{10}}]_{10}^{\infty} = \frac{1}{e}$.
2. entre 10 et 20 minutes ?
   Réponse : $P(10 < X < 20) = \int_{10}^{20} \frac{1}{10} e^{\frac{-x}{10}}\, dx = [-e^{\frac{-x}{10}}]_{10}^{20} = \frac{1}{e} - \frac{1}{e^2}$.

Exercice 78 Durée de vie d'un composant électronique (Bac série S, France métropolitaine, 2004).

On s'intéresse à la durée de vie, exprimée en semaines, d'un composant électronique. On modélise cette situation par une loi de probabilité P de durée de vie sans vieillissement définie sur l'intervalle $[0; +\infty[$ par

$$P([0; t[) = \int_0^t \lambda \exp(-\lambda x) dx,$$

représentant la probabilité que le composant ne soit plus en état de marche au bout de t semaines.

Une étude statistique montrant qu'environ 50% d'un lot important de ces composants sont encore en état de marche au bout de 200 semaines, permet de poser $P([0; 200[) = 0,5$.

1. Montrer que
$$\lambda = \frac{\ln 2}{200}.$$
2. Quelle est la probabilité qu'un de ces composants pris au hasard ait une durée de vie supérieure à 300 semaines ?
3. On admet que la durée de vie moyenne d_m de ces composants est la limite quand A tend vers $+\infty$ de $\displaystyle\int_0^A \lambda x \exp(-\lambda x) dx$.
   Montrer que $\displaystyle\int_0^A \lambda x \exp(-\lambda x) dx = \frac{-\lambda A \exp(-\lambda A) - \exp(-\lambda A) + 1}{\lambda}$ et en déduire d_m à la semaine près.

Exercice 79 Soient $X_1, X_2, \ldots, X_n$ des variables aléatoires indépendantes de loi exponentielle de paramètre $\lambda > 0$.

1. Calculer la loi (i.e. déterminer sa densité) de
$$\max_{i \in [[1,n]]} X_i.$$
   (Indication : calculer la fonction de répartition.)
2. Calculer la loi (i.e. déterminer sa densité) de
$$\min_{i \in [[1,n]]} X_i.$$

.

8.3 La loi de Laplace-Gauss (ou loi normale)

8.3.1 Définition

La variable aléatoire X normale est donnée par sa densité f telle que :

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

où $m \in \mathbb{R}$ et $\sigma \in \mathbb{R}_+^*$.

On a bien

$$\begin{aligned}
\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-m)^2}{2\sigma^2}} dx &= \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-t^2} \sigma\sqrt{2} dt \\
&\quad (\textit{en posant } t = \frac{x-m}{\sigma\sqrt{2}}), \\
&= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-t^2} dt \\
&= 1.
\end{aligned}$$

En effet,

$$\begin{aligned}
\left(\int_{-\infty}^{\infty} e^{-t^2} dt\right)^2 &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-x^2-y^2} dxdy \\
&= \int_{0}^{2\pi}\int_{0}^{\infty} e^{-r^2} r dr d\theta \\
&\quad (\textit{par changement de variables en polaires}), \\
&= 2\pi[\frac{-1}{2} e^{-r^2}]_0^{\infty} \\
&= \pi
\end{aligned}$$

De plus, on remarquera que f est symétrique par rapport à m puisque $\forall x \in \mathbb{R}, f(m+x) = f(m-x)$.

X suit une loi normale de paramètres m, σ est noté $X \hookrightarrow N(m, \sigma)$.

8.3.2 Caractéristiques

$E(X) = m$ et $V(X) = \sigma^2$.

On a :

$$\begin{aligned}
E(X) &= \int_{-\infty}^{\infty} \frac{x}{\sigma\sqrt{2\pi}} e^{\frac{-(x-m)^2}{2\sigma^2}} dx \\
&= \int_{-\infty}^{\infty} \frac{t\sigma\sqrt{2}+m}{\sigma\sqrt{2\pi}} e^{-t^2} \sigma\sqrt{2} dt \\
&\quad (\textit{en posant } t = \frac{x-m}{\sigma\sqrt{2}}), \\
&= m\frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-t^2} dt + \frac{\sigma\sqrt{2}}{\sqrt{\pi}} \int_{-\infty}^{\infty} t e^{-t^2} dt \\
&= m + \frac{\sigma\sqrt{2}}{\sqrt{\pi}} \underbrace{[\frac{-1}{2} e^{-t^2}]_{-\infty}^{\infty}}_{=0} \\
&= m.
\end{aligned}$$

Et,

$$\begin{aligned}
V(X) &= \int_{-\infty}^{\infty} \frac{(x-m)^2}{\sigma\sqrt{2\pi}} e^{\frac{-(x-m)^2}{2\sigma^2}} dx \\
&= \int_{-\infty}^{\infty} \frac{2\sigma^2 t^2}{\sigma\sqrt{2\pi}} e^{-t^2} \sigma\sqrt{2} dt \\
&\quad (en\ posant\ t = \frac{x-m}{\sigma\sqrt{2}}), \\
&= \frac{2\sigma^2}{\sqrt{\pi}} \int_{-\infty}^{\infty} t^2 e^{-t^2} dt \\
&= \frac{2\sigma^2}{\sqrt{\pi}} \left(\underbrace{[\frac{-1}{2} t e^{-t^2}]_{-\infty}^{\infty}}_{=0} - \int_{-\infty}^{\infty} \frac{-1}{2} e^{-t^2} dt \right) \\
&= \sigma^2.
\end{aligned}$$

8.3.3 La loi normale centrée réduite

On appelle loi normale centrée réduite la loi normale de paramètres $m = 0$ et $\sigma = 1$.

Ainsi, la densité de probabilité devient

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

La fonction de répartition F s'écrit

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{-t^2}{2}} dt.$$

La courbe représentative de F passe par le point de coordonnées $(0, \frac{1}{2})$, possède une asymptote au voisinage de $-\infty$ d'équation $y = 0$ (et une asymptote au voisinage de ∞ d'équation $y = 1$) et est symétrique par rapport au point de coordonnées $(0, \frac{1}{2})$.

On remarque que si $X \hookrightarrow N(m, \sigma)$ (loi normale), alors, $\frac{X-m}{\sigma} \hookrightarrow N(0, 1)$ (loi normale centrée réduite).

Propriétés

Ces propriétés servent à utiliser la table de la loi normale centrée réduite ne donnant qu'une partie des valeurs.

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

1. $\int_{-\infty}^{-x} f(t)dt = \int_{x}^{\infty} f(t)dt$
   En effet, $f(t) = f(-t), \forall t \in \mathbb{R}$.
2. $\int_{-x}^{x} f(t)dt = 2\int_{-\infty}^{x} f(t)dt - 1$
   En effet, $\int_{-x}^{x} f(t)dt$
   $= \int_{-\infty}^{x} f(t)dt - \int_{-\infty}^{-x} f(t)dt$
   $= \int_{-\infty}^{x} f(t)dt - (1 - \int_{-x}^{\infty} f(t)dt)$
   $= \int_{-\infty}^{x} f(t)dt - (1 - \int_{-\infty}^{x} f(t)dt)$
   $= 2\int_{-\infty}^{x} f(t)dt - 1.$

Exercice 80 [9, exercice 24, p. 199] On suppose que X suit une loi normale de paramètres m et σ.

On a $P(X > 15) \approx 0,0668$ et $P(X < 10) \approx 0,1587$

Evaluer m et σ.

8.3.4 Valeurs remarquables pour $N(m,\sigma)$

Si $X \hookrightarrow N(m,\sigma)$, alors :
$P(m-1,96\sigma < X < m+1,96\sigma) = 0,95...$
$P(m-2,58\sigma < X < m+2,58\sigma) = 0,99...$

8.3.5 Additivité de deux lois normales indépendantes

Si $X_1 \hookrightarrow N(m_1,\sigma_1)$ et si $X_2 \hookrightarrow N(m_2,\sigma_2)$ sont indépendantes, alors $X_1+X_2 \hookrightarrow N(m_1+m_2, \sqrt{\sigma_1^2+\sigma_2^2})$.
Démonstration
La densité de probabilité de X_1+X_2 est

$$\begin{aligned} h(x) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-t-m_1)^2}{2\sigma_1^2}} \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{(t-m_2)^2}{2\sigma_2^2}} dt \\ &= \frac{1}{2\pi\sigma_1\sigma_2} \int_{-\infty}^{\infty} e^{-\frac{\sigma_2^2(x-t-m_1)^2+\sigma_1^2(t-m_2)^2}{2\sigma_1^2\sigma_2^2}} dt \end{aligned}$$

Ensuite, le polynôme $\sigma_2^2(x-t-m_1)^2+\sigma_1^2(t-m_2)^2$ de degré deux en t est mis sous forme canonique :

$$\sigma_2^2(x-t-m_1)^2+\sigma_1^2(t-m_2)^2 = (\sigma_1^2+\sigma_2^2)[t-\frac{\sigma_2^2(x-m_1)+\sigma_1^2 m_2}{\sigma_1^2+\sigma_2^2}]^2 + \frac{\sigma_1^2\sigma_2^2(x-m_1-m_2)^2}{\sigma_1^2+\sigma_2^2}.$$

Ainsi,

$$\begin{aligned} h(x) &= \frac{1}{\sqrt{2\pi(\sigma_1^2+\sigma_2^2)}} e^{-\frac{(x-m_1-m_2)^2}{2(\sigma_1^2+\sigma_2^2)}} \int_{-\infty}^{\infty} \frac{\sqrt{\sigma_1^2+\sigma_2^2}}{\sqrt{2\pi}\sigma_1\sigma_2} e^{-\frac{\sigma_1^2+\sigma_2^2}{2\sigma_1^2\sigma_2^2}[t-\frac{\sigma_2^2(x-m_1)+\sigma_1^2 m_2}{\sigma_1^2+\sigma_2^2}]^2} dt \\ &= \frac{1}{\sqrt{2\pi(\sigma_1^2+\sigma_2^2)}} e^{-\frac{(x-m_1-m_2)^2}{2(\sigma_1^2+\sigma_2^2)}} \end{aligned}$$

Et, le résultat demandé est ainsi prouvé.

8.3.6 Approximation de la loi binomiale par la loi normale ou théorème de De Moivre-Laplace

Soit $p \in]0,1[$.
Soit X_n une suite de variables aléatoires suivant la loi binomiale de paramètres p $(q=1-p)$ et n. Ainsi,

$$P(X_n = k_n) = C_n^{k_n} p^{k_n} q^{n-k_n},\ \forall k_n \in \{0,1,\dots,n\}.$$

Ainsi, on a :

$$P(c < \frac{X_n - np}{\sqrt{npq}} < d) \sim_{n\sim\infty} \frac{1}{\sqrt{2\pi}} \int_c^d e^{\frac{-t^2}{2}dt},\ \forall c < d.$$

En d'autres termes, la loi binomiale approche la loi normale.
Idée de la démonstration
Rappel de la formule de Stirling :

$$n! \sim_{n\sim\infty} \sqrt{2\pi n} n^n e^{-n}.$$

On pose $t(k_n) = \frac{k_n - np}{\sqrt{npq}}$.
On a

$$\begin{aligned} k_n &= np + t(k_n)\sqrt{npq} \\ &= np(1+\frac{t(k_n)}{\sqrt{n}}\sqrt{\frac{q}{p}}) \\ &\sim_{n\sim\infty} np\ (car\ \lim_{n\to\infty} \frac{t(k_n)}{\sqrt{n}} = 0\ puisque\ t(k_n) \in [c,d]). \end{aligned}$$

De même,

$$\begin{aligned} n-k_n &= nq - t(k_n)\sqrt{npq} \\ &= nq(1-\frac{t(k_n)}{\sqrt{n}}\sqrt{\frac{p}{q}}) \\ &\sim_{n\sim\infty} nq. \end{aligned}$$

En appliquant la formule de Stirling,

$$P(X_n = k_n) \sim_{n\sim\infty} \frac{1}{\sqrt{2\pi}}\sqrt{\frac{n}{k_n(n-k_n)}}\left(\frac{np}{k_n}\right)^{k_n}\left(\frac{nq}{n-k_n}\right)^{n-k_n}$$

On pose alors

$$Q = \sqrt{\frac{n}{k_n(n-k_n)}},$$

$$R = \left(\frac{np}{k_n}\right)^{k_n}\left(\frac{nq}{n-k_n}\right)^{n-k_n}$$

et

$$u = \frac{t(k_n)}{\sqrt{n}}.$$

On obtient

$$\begin{aligned} Q &= \frac{1}{\sqrt{npq\left(1+u\sqrt{\frac{q}{p}}\right)\left(1-u\sqrt{\frac{p}{q}}\right)}} \\ &\sim_{n\sim\infty} \frac{1}{\sqrt{npq}}. \end{aligned}$$

et

$$\begin{aligned} \ln(R) &= k_n \ln\left(\frac{np}{k_n}\right) + (n-k_n)\ln\left(\frac{nq}{n-k_n}\right) \\ &= n(p+u\sqrt{pq})\ln\left(\frac{1}{1+u\sqrt{\frac{q}{p}}}\right) + n(q-u\sqrt{pq})\ln\left(\frac{1}{1-u\sqrt{\frac{p}{q}}}\right) \\ &= n(-\frac{u^2}{2} + O(u^3)). \end{aligned}$$

Ensuite,

$$R \sim_{n\sim\infty} e^{\frac{-(t(k_n))^2}{2}}.$$

Il s'ensuit :

$$P(X_n = k_n) \sim_{n\sim\infty} \frac{1}{\sqrt{2\pi npq}} e^{-\frac{(k_n-np)^2}{2npq}}.$$

Ensuite, en admettant qu'on peut faire la somme infinie de ces équivalents (le démontrer n'est pas chose aisée), on obtient :

$$\begin{aligned} \sum_{np+c\sqrt{npq}\le k_n\le np+d\sqrt{npq}} P(X_n = k_n) &\sim_{n\sim\infty} \frac{1}{\sqrt{2\pi npq}} \sum_{np+c\sqrt{npq}\le k_n\le np+d\sqrt{npq}} e^{-\frac{(k_n-np)^2}{2npq}} \\ &\sim_{n\sim\infty} \frac{1}{\sqrt{2\pi}}\int_c^d e^{-\frac{t^2}{2}}\,dt \ (\textit{somme de Riemann}). \end{aligned}$$

8.3.7 Le théorème *central limit*

Le théorème de De Moivre-Laplace est un cas particulier du théorème suivant.

Enoncé

Soient $X_1, X_2, \ldots, X_n$, n variables aléatoires réelles indépendantes de même loi. Si on note $E(X_i) = m$ et $V(X_i) = \sigma^2$, alors :

$$\lim_{n\to\infty} P(a \leq \frac{S_n - nm}{\sqrt{n}\sigma} \leq b) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{\frac{-t^2}{2}} dt$$

lorsque $a \leq b$ et $S_n = X_1 + X_2 + \ldots + X_n$

ou, avec une conclusion formulée autrement, $\frac{S_n - nm}{\sqrt{n}\sigma} \hookrightarrow N(0,1)$.

Ce théorème n'est pas démontré.

$\boxed{\textbf{Exercice 81}}$ [10] Lors d'un procès en attribution de paternité, un expert témoigne que la durée de la grossesse, en jours, c'est-à-dire le laps de temps entre la conception et la naissance de l'enfant, est de distribution approximativement normale avec paramètres $m = 270$ et $\sigma^2 = 100$. L'un des pères putatifs est en mesure de prouver son absence du pays pendant une période s'étendant entre le $290^{ème}$ et le $240^{ème}$ jour précédent l'accouchement. Quelle est la probabilité que la conception ait eu lieu à ce moment ?

Solution : $P(X > 290 \bigcup X < 240)$

$= P(X > 290) + P(X < 240)$

$= P(\frac{X-270}{10} > 2) + P(\frac{X-270}{10} < -3)$

$= 0,0241...$, après consultation de la table fournissant certaines valeurs de la loi normale centrée réduite.

8.4 La loi de Laplace

8.4.1 Définition

La variable aléatoire X de Laplace est donnée par sa densité f telle que :

$$f(x) = \frac{1}{2a} e^{\frac{-|x|}{a}}$$

où $a \in \mathbb{R}^+ \backslash \{0\}$.

On a bien

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{1}{2a} e^{\frac{-|x|}{a}} dx &= 2 \int_0^{\infty} \frac{1}{2a} e^{\frac{-x}{a}} dx \\ &= \frac{1}{a} [(-a) e^{\frac{-x}{a}}]_0^{\infty} \\ &= 1. \end{aligned}$$

De plus, on a que f est symétrique par rapport à 0 puisque $\forall x \in \mathbb{R}, f(x) = f(-x)$.

X suit une loi de Laplace de paramètre a est noté $X \hookrightarrow L(a)$.

8.4.2 Caractéristiques

$E(X) = 0$ et $V(X) = 2a^2$.

On a :

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} \frac{x}{2a} e^{\frac{-|x|}{a}} dx \\ &= 0, \end{aligned}$$

car la fonction qui à x fait correspondre $\frac{x}{2a}e^{\frac{-|x|}{a}}$ est impaire. Et,

$$\begin{aligned}
V(X) &= \int_{-\infty}^{\infty} \frac{(x-0)^2}{2a} e^{\frac{-|x|}{a}} dx \\
&= 2\int_0^{\infty} \frac{x^2}{2a} e^{\frac{-x}{a}} dx \\
&= \frac{1}{a}\int_0^{\infty} x^2 e^{\frac{-x}{a}} dx \\
&= \frac{1}{a}\left(-\int_0^{\infty} 2x(-a)e^{\frac{-x}{a}} dx + \underbrace{[x^2(-a)e^{\frac{-x}{a}}]_0^{\infty}}_{=0}\right) \\
&= \int_0^{\infty} 2xe^{\frac{-x}{a}} dx \\
&= -\int_0^{\infty} 2(-a)e^{\frac{-x}{a}} dx + \underbrace{[2x(-a)e^{\frac{-x}{a}}]_0^{\infty}}_{=0} \\
&= 2a[(-a)e^{\frac{-x}{a}}]_0^{\infty} \\
&= 2a^2.
\end{aligned}$$

9 Statistiques inférentielles

On note $I(a,b) = [a-b, a+b]$.

9.1 Echantillonnage et estimation

L'**échantillonnage** correspond au problème suivant : connaissant une population, en déduire les caractéristiques d'un échantillon d'effectif donné.

L'**estimation** répond au problème réciproque : connaissant un échantillon donné, en induire les caractéristiques de la population dont il est issu. Ce deuxième problème comporte des incertitudes. Il ne pourra être résolu que moyennant un certain "risque d'erreur".

Dans cette section, les éléments des échantillons sont supposées être des **variables indépendantes**.

9.1.1 Echantillonnage

A. Etude de la moyenne d'échantillon

Soit une population (P) d'effectif N de moyenne μ et d'écart-type σ donnée.

Soit X un élément quelconque de (P), inconnu. X suit une variable aléatoire, dotée d'une espérance mathématique $E(X)$ et d'un écart-type σ_X.

Soit, par exemple, le tableau de distribution de (P) d'effectif N de moyenne μ et d'écart-type σ.

Variable X	x_1	x_2	$\dots$	x_i	$\dots$	x_r
Effectif K	k_1	k_2	$\dots$	k_i	$\dots$	k_r

Remarque : $\sum_{i=1}^{r} k_i = N$.

En prélevant dans cette population un élément X inconnu, la probabilité qu'il ait la valeur x_i est $p_i = \frac{k_i}{N}$.

Son espérance mathématique est : $E(X) = \sum_{i=1}^{r} p_i x_i = \sum_{i=1}^{r} \frac{k_i}{N} x_i = \mu$.

De même, sa variance est : $V(X) = E(X^2) - (E(X))^2 = \sum_{i=1}^{r} p_i x_i^2 - \mu^2 = \sum_{i=1}^{r} \frac{k_i}{N} x_i^2 - \mu^2 = \sigma^2$.

Conclusion : un élément quelconque X inconnu, d'une population de moyenne μ et d'écart-type σ est une variable aléatoire d'espérance mathématique $E(X) = \mu$ et d'écart-type $\sigma_X = \sigma$.

On en déduit que si on prélève dans une population (P) d'effectif N de moyenne μ et d'écart-type σ un échantillon de taille n, sa moyenne $\overline{X} = \sum_{i=1}^{n} \frac{X_i}{n}$ est une variable aléatoire d'espérance $E(\overline{X}) = \mu$ et d'écart-type $\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$.

Si la variable X suit une loi normale de paramètres μ et σ, $\overline{X}$ suit une loi normale de paramètres μ et $\frac{\sigma}{\sqrt{n}}$.

Si la variable X suit une autre loi que la loi normale de paramètres μ et σ, on peut faire comme si c'était le cas (lorsque n est suffisamment grand), d'après le théorème "central-limit".

Exemple.

Une machine effectue l'ensachage d'un produit.

On sait que les sacs ont un poids moyen de $250g$ avec un écart-type de $25g$.

Quelles sont les caractéristiques de la moyenne des poids d'un échantillon de 100 sacs ?

Solution.

$\overline{X}$ suit la loi normale de paramètres 250 et $2,5$.

B. Etude des fréquences

Il s'agit de la fréquence d'apparition, dans une population (P), d'une propriété caractérisée par une alternative (oui/non).

Fréquence, pourcentage et proportion sont des termes équivalents.

Soit une population (P) d'effectif N de moyenne μ et d'écart-type σ, dont les éléments X possèdent ou non une propriété donnée. Soit K le nombre d'éléments qui possèdent cette propriété, la fréquence est donc $\frac{K}{N}$, et chaque élément de (P) considéré comme une variable aléatoire a donc la probabilité $p = \frac{K}{N}$ de posséder la propriété.

Soit (E) un échantillon de taille n : $(E) = \{X_1, X_2, \ldots, X_n\}$. Soit k le nombre d'éléments de (E) possédant la propriété, la variable k suit donc la loi binomiale de paramètres n et p. Soit $f = \frac{k}{n}$ la fréquence de succès dans l'échantillon, on déduit $E(f) = p$ et $\sigma(f) = \sqrt{\frac{p(1-p)}{n}}$.

On peut estimer que f suit une loi normale de paramètres p et $\sqrt{\frac{p(1-p)}{n}}$ (lorsque n est suffisamment grand), d'après le théorème "central-limit".

9.1.2 Estimation

A. Estimation ponctuelle

Elle consiste à estimer une caractéristique c de la population par la valeur ponctuelle γ d'une caractéristique de l'échantillon observé. γ s'appelle l'estimateur de c et il doit pour cela remplir une condition : γ doit tendre vers c quand l'effectif de l'échantillon augmente, avec un écart-type nul au voisinage de l'infini.

Soit une population (P) d'effectif N de moyenne μ et d'écart-type σ et un échantillon observé (E) d'effectif n de moyenne m et d'écart-type s.

Nous savons qu'un échantillon quelconque d'effectif n a une moyenne $\overline{X}$ telle que $E(\overline{X}) = \mu$ et $\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$.

- Estimation de la moyenne μ :

m est un des $\overline{X}$ possibles.

Cependant, quand n tend vers l'infini, $\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$ tend vers 0. Donc, au voisinage de l'infini : $m = E(\overline{X}) = \mu$.

Il est donc naturel que m soit l'estimateur ponctuel de μ.

- Estimation de l'écart-type σ :

Le meilleur estimateur ponctuel de σ n'est pas s mais $s\sqrt{\frac{n}{n-1}}$.

En effet, soit un échantillon quelconque de taille n, d'éléments X_i (dans la suite, l'ensemble de ces éléments X_i est considéré comme une variable aléatoire notée abusivement X_i).

Sa moyenne $E(X_i) = \overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}$.

Sa variance est $V(X_i) = s^2 = \frac{\sum_{i=1}^{n}(X_i^2 - \overline{X}^2)}{n}$, et, par conséquent, $E(s^2) = \frac{\sum_{i=1}^{n}(E(X_i^2) - E(\overline{X}^2)}{n}$.

Or tous les X_i ont même espérance mathématique, puisque éléments de la population (P) : $E(X_i) = E(X)$ et $E(X_i^2) = E(X^2)$. Par suite, $E(s^2) = \frac{nE(X^2) - nE(\overline{X}^2)}{n} = E(X^2) - E(\overline{X}^2)$.

Dans la population (P), $V(X) = \sigma^2 = E(X^2) - (E(X))^2 = E(X^2) - \mu^2$.

D'où,

$$E(X^2) = \sigma^2 + \mu^2.$$

Dans l'échantillon (E), $V(\overline{X}) = \frac{\sigma^2}{n} = E(\overline{X}^2) - (E(\overline{X}))^2 = E(\overline{X}^2) - \mu^2$ (car $m = E(\overline{X}) = \mu$).

D'où,

$$E(\overline{X}^2) = \frac{\sigma^2}{n} + \mu^2.$$

Par suite,

$$E(s^2) = E(X^2) - E(\overline{X}^2) = (\sigma^2 + \mu^2) - (\frac{\sigma^2}{n} + \mu^2) = \sigma^2 \frac{n-1}{n},$$

puis,

$$\sigma^2 = E(s^2)\frac{n}{n-1}.$$

Il est donc naturel d'estimer σ par $s\sqrt{\frac{n}{n-1}}$ ...

- **Estimation d'une fréquence** f :

Soit une population (P) dont les éléments possèdent (succès) ou ne possèdent pas (échec) une propriété donnée. Il s'agit d'estimer la probabilité p de succès d'un élément quelconque de (P).

Soit la fréquence de succès dans l'échantillon $f = \frac{k}{n}$, nous avons vu que $E(f) = p$ et $\sigma(f) = \sqrt{\frac{p(1-p)}{n}}$. Cependant, $\sqrt{\frac{p(1-p)}{n}}$ tend vers 0 lorsque n tend vers l'infini, donc f est l'estimateur de p.

Estimation par un intervalle de confiance

L'estimation ponctuelle n'indique pas le degré de confiance que l'on peut lui accorder. D'où l'estimation par un intervalle de confiance, au seuil de probabilité donnée c, appelé seuil de confiance.

- **Intervalle de confiance d'une moyenne** :

Il est supposé que l'échantillon a une taille suffisamment importante pour que l'on puisse admettre que sa moyenne $\overline{X}$ suit une loi normale.

Soit une population $P(N, \mu, \sigma)$ et un échantillon observé $E(n, m, s)$.

L'échantillon observé (E) est l'un des échantillons possibles de même taille. La moyenne m est donc l'un des $\overline{X}$, variable aléatoire, moyenne d'un échantillon quelconque de taille n. Et, comme $\overline{X}$, m suit une loi normale de paramètres μ et $\frac{\sigma}{\sqrt{n}}$.

$\longrightarrow$ Si on connaît σ :

Soit c une probabilité donnée.

Cherchons l'intervalle $I(\mu, \alpha)$ tel que la probabilité que m soit dans $I(\mu, \alpha)$ soit c.

Si Π est la fonction de répartition de la loi normale centrée réduite, on déduit que $\Pi(\frac{\alpha}{\frac{\sigma}{\sqrt{n}}}) - \Pi(\frac{-\alpha}{\frac{\sigma}{\sqrt{n}}}) = c$ ou que $\Pi(\frac{\alpha}{\frac{\sigma}{\sqrt{n}}}) = \frac{1+c}{2}$, ce qui donne α par lecture de la table de la loi normale centrée réduite.

Cependant, $\mu - \alpha \leq m \leq \mu + \alpha$ équivaut à $m - \alpha \leq \mu \leq m + \alpha$ ou encore m est dans $I(\mu, \alpha)$ équivaut à μ est dans $I(m, \alpha)$ ou encore la probabilité que m soit dans $I(\mu, \alpha)$ est aussi la probabilité que μ soit dans $I(m, \alpha)$. On conclut que μ suit la loi normale de paramètres m et $\frac{\sigma}{\sqrt{n}}$.

L'intervalle $I(m, \alpha)$ est appelé intervalle de confiance de μ au seuil c.

$\longrightarrow$ Si on ne connaît pas σ :

On l'estime selon la loi d'estimation ponctuelle : $\sigma_{est} = s\sqrt{\frac{n}{n-1}}$.

Exemple.

Une machine effectue l'ensachage d'un produit.

Le poids théorique des sachets est de $250g$.

Pour contrôler la fabrication, on cherche à déterminer la moyenne μ du poids des sachets, au risque d'erreur de 5%.

Pour ce faire, on prélève un échantillon de 100 sachets. La moyenne observée est $m = 251g$, avec un écart-type $s = 25g$.

Solution.

Compte tenu de la taille de l'échantillon, on admettra que l'écart-type σ de la variable X, le poids d'un sachet, est égal à celui de l'échantillon (on néglige le facteur $\sqrt{\frac{n}{n-1}}$).

On peut donc considérer que la moyenne μ suit la loi normale de paramètres 251 et 2, 5. Ceci fournit $\Pi(\frac{\alpha}{2,5}) = 0,975$ (le risque r est donné par $r = 1 - c$), puis $\alpha \approx 4,9$.

D'où le poids moyen des sachets est compris entre $\approx 246,1g$ et $\approx 255,9g$, au risque d'erreur de 5%.

- Intervalle de confiance d'une fréquence :

Soit une population (P) dont les éléments possèdent (succès) ou ne possèdent pas (échec) une propriété donnée.

Il s'agit d'estimer la probabilité p de succès d'un élément quelconque de (P).

Soit la fréquence de succès dans l'échantillon $f = \frac{k}{n}$.

Dans le cas où on peut estimer que f suit la loi normale de paramètres p et $\sqrt{\frac{p(1-p)}{n}}$, l'objectif est d'estimer p dans un intervalle de confiance de probabilité donnée c.

Par un processus semblable à celui pour chercher l'intervalle de confiance d'une moyenne, si f suit la loi normale de paramètres p et $\sqrt{\frac{p(1-p)}{n}}$, p suivra la loi normale de paramètres f et $\sqrt{\frac{f(1-f)}{n}}$.

On cherchera donc dans cette loi le nombre α tel que la probabilité que p soit dans $I(f, \alpha)$ soit c.

Exemple.

Dans l'analyse de la qualité de la fabrication d'une machine, on cherche la probabilité pour qu'une pièce soit défectueuse : p.

Pour ce faire, on prélève un échantillon de 125 pièces dans lequel on a relevé 10 pièces défectueuses.

On cherche p dans un intervalle de confiance à 95%.

Solution.

Dans l'échantillon, la fréquence de pièces défectueuses est $f = \frac{10}{125} = 0,08$.

On peut donc considérer que la probabilité p suit la loi normale de paramètres 0, 08 et $\approx 0,024$. Ceci fournit $\Pi(\frac{\alpha}{0,024}) = 0,975$, puis $\alpha \approx 0,047$.

D'où la probabilité pour qu'une pièce soit défectueuse est comprise entre $\approx 0,033$ et $\approx 0,127$, au risque d'erreur de 5%.

9.2 Tests de validité d'hypothèse

A partir d'un échantillon, on est souvent amené à prendre une décision au sujet de la population-mère. Deux questions se présentent :

– l'échantillon est-il représentatif d'une population théorique donnée ?
– deux échantillons sont-ils homogènes, c'est-à-dire issus d'une même population ou de deux populations de mêmes caractéristiques ?

9.2.1 Comparaison à une norme

A. Comparaison d'une moyenne à une norme

Soit un échantillon (E) d'effectif n de moyenne m et d'écart-type s prélevé sur une population (P) dont on estime la moyenne M à la valeur μ et l'écart-type à la valeur σ.

Le problème est le suivant : de l'observation m, peut-on conclure raisonnablement que la moyenne de la population est effectivement μ ?

On se place exclusivement dans le cas où la population est réputée suivre une loi normale (ou l'échantillon suffisamment grand pour pouvoir appliquer le théorème "central-limit").

On adopte une seuil c, par exemple un risque d'erreur de $c = 5\%$.

La moyenne d'un échantillon de taille n suit une loi normale de paramètres μ et $\frac{\sigma}{\sqrt{n}}$, si μ est effectivement la moyenne de la population-mère. Dans ces conditions, cette moyenne $\overline{X}$ est telle que la probabilité que $\overline{X}$ soit dans $I(\mu, \alpha)$ soit de $\underbrace{0,95}_{dans\ l'exemple}$, avec $\alpha = \underbrace{1,96}_{dans\ l'exemple} \frac{\sigma}{\sqrt{n}}$.

Conclusion : si m est extérieur à l'intervalle $I(\mu, \alpha)$ ainsi défini, il n'y a que $\underbrace{5}_{dans\ l'exemple}$ chances sur 100 que ce soit dû au seul hasard de l'échantillonnage, et $\underbrace{95}_{dans\ l'exemple}$ chances sur 100 que ce soit dû à une différence entre la moyenne théorique μ et la moyenne réelle M de la population-mère.

Dans ces conditions, on est amené à rejeter l'hypothèse "$M = \mu$" au seuil de $\underbrace{5}_{dans\ l'exemple}$ %, c'est-à-dire avec un risque de $\underbrace{5}_{dans\ l'exemple}$ % de rejeter l'hypothèse à tort.

Si m est intérieur à l'intervalle $I(\mu, \alpha)$, on adoptera l'hypothèse. Cela ne veut pas dire que "$M = \mu$", mais que l'on n'a aucune raison véritable d'en douter.

Si on ne connaît pas l'écart-type, on l'estime selon la loi d'estimation ponctuelle : $\sigma_{est} = s\sqrt{\frac{n}{n-1}}$.

Exemple.

Soit un stock important dont on estime qu'un caractère X suit une loi normale, de moyenne $\mu = 240$ et d'écart-type $\sigma = 50$.

On prélève un échantillon de 40 unités, dont la moyenne observée est $m = 260$.

Au risque de 5% peut-on considérer que μ est effectivement égal à 240 ?

Solution.

Si $\mu = 240$, $\overline{X}$ suit une loi normale de paramètres 240 et $\frac{50}{\sqrt{40}}$.

On déduit $\alpha = 1,96\frac{50}{\sqrt{40}} \approx 15,5$.

D'où $m = 260$ est extérieure à l'intervalle $I(240, 15, 5)$. On refuse donc l'hypothèse au seuil de 5%.

Remarque : au seuil de 1%, on trouve $\alpha \approx 20,3$ et on pourra accepter l'hypothèse.

B. Comparaison d'une fréquence à une norme

Soit une population (P) où la fréquence d'apparition ρ d'une certaine propriété est supposée égale à p. On prélève un échantillon (E) de taille n, et on note f la fréquence observée dans l'échantillon.

Soit f' la variable aléatoire, fréquence des échantillons de taille n.

On adopte une seuil c, par exemple un risque d'erreur de $c = 5\%$.

La fréquence d'un échantillon de taille n suit une loi normale de paramètres p et $\sqrt{\frac{p(1-p)}{n}}$, si p est effectivement la fréquence d'apparition de la population-mère. Dans ces conditions, cette fréquence f' est telle que la probabilité que f' soit dans $I(p, \alpha)$ soit de $\underbrace{0,95}_{dans\ l'exemple}$, avec $\alpha = \underbrace{1,96}_{dans\ l'exemple} \sqrt{\frac{p(1-p)}{n}}$.

Conclusion : si f est extérieur à l'intervalle $I(p, \alpha)$ ainsi défini, il n'y a que $\underbrace{5}_{dans\ l'exemple}$ chances sur 100 que ce soit dû au seul hasard de l'échantillonnage, et $\underbrace{95}_{dans\ l'exemple}$ chances sur 100 que ce soit dû à une différence entre la fréquence d'apparition théorique p et la fréquence d'apparition réelle ρ de la population-mère.

Dans ces conditions, on est amené à rejeter l'hypothèse "$\rho = p$" au seuil de $\underbrace{5}_{dans\ l'exemple}$ %, c'est-à-dire avec un risque de $\underbrace{5}_{dans\ l'exemple}$ % de rejeter l'hypothèse à tort.

Si f est intérieur à l'intervalle $I(p, \alpha)$, on adoptera l'hypothèse. Cela ne veut pas dire que "$\rho = p$", mais que l'on n'a aucune raison véritable d'en douter.

Exemple.

Soit la population constituée par la production d'une machine. On estime sa fréquence de pièces hors tolérance à $p = 0,03$.

On prélève au hasard un échantillon de 500 pièces, dans lequel on relève 18 pièces hors tolérance.

Peut-on, au seuil de 5%, adopter l'hypothèse $p = 0,03$?

Solution.

Au vu de l'effectif de l'échantillon, on admet que f', la fréquence des pièces hors tolérance dans les échantillons de taille 500, suit la loi normale de paramètres $0,03$ et $\sqrt{\frac{0,03\times 0,97}{500}} \approx 0,0076$.

On déduit $\alpha = 1,96\sqrt{\frac{0,03\times 0,97}{500}} \approx 0,015$.

D'où $f = \frac{18}{500} = 0,036$ est intérieure à l'intervalle $I(0,03, 0,015)$. On accepte donc l'hypothèse au seuil de 5%.

9.2.2 Comparaison de deux échantillons

A. Comparaison de deux moyennes observées

Soient deux échantillons (E_1) et (E_2) prélevés sur une population théoriquement inchangée entre les deux prélèvements et assez importante pour que l'on puisse considérer que les deux échantillons sont indépendants.

On suppose en outre que la population suit une loi normale ou que les échantillons ont des effectifs suffisamment grand pour pouvoir appliquer le théorème "central-limit".

On analyse un caractère mesurable : soit (E_1) de taille n_1, de moyenne m_1 et d'écart-type s_1, et soit (E_2) de taille n_2, de moyenne m_2 et d'écart-type s_2.

Soit $\overline{X_1}$ la variable aléatoire, moyenne d'un échantillon de taille n_1, et $\overline{X_2}$ la variable aléatoire, moyenne d'un échantillon de taille n_2.

Les éléments de (E_1) et (E_2), X_1 et X_2 appartiennent à la même population.

Donc : $E(\overline{X_1}) = E(X_1) = E(X_2) = E(\overline{X_2})$.

Par ailleurs, $V(\overline{X_1}) = \frac{V(X_1)}{n_1} = \frac{s_1^2}{n_1}$ et $V(\overline{X_2}) = \frac{V(X_2)}{n_2} = \frac{s_2^2}{n_2}$.

Soit la variable $\overline{Y} = \overline{X_1} - \overline{X_2}$.

On a $E(\overline{Y}) = E(\overline{X_1}) - E(\overline{X_2}) = 0$ et $V(\overline{Y}) = V(\overline{X_1}) + V(\overline{X_2}) = \frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}$.

Comme $\overline{X_1}$ et $\overline{X_2}$ suivent des lois normales, $\overline{Y}$ suit également une loi normale. Et $\overline{Y}$ suit une loi normale de paramètres 0 et $\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$.

On pourra estimer, à un seuil donné, que les deux échantillons appartiennent à la même population si $d = m_1 - m_2$ appartient à l'intervalle déterminé pour $\overline{Y}$, au même seuil.

Exemple.

Une machine fabrique des pièces de cote X. On cherche à détecter une dérive de réglage dans le temps. Pour cela, on prélève des échantillons à des temps différents.

soit (E_1) de taille $n_1 = 100$, de moyenne $m_1 = 248$ et d'écart-type $s_1 = 28$, et soit (E_2) de taille $n_2 = 200$, de moyenne $m_2 = 256$ et d'écart-type $s_2 = 22$.

Peut-on estimer, au seuil de 5% que le réglage de la machine n'a pas changé entre ces deux événements ?

Solution.

Soit $d = m_2 - m_1 = 8$.

$\overline{Y} = \overline{X_1} - \overline{X_2}$ suit une loi normale de paramètres 0 et $\sqrt{\frac{28^2}{100} + \frac{22^2}{200}} \approx 3,20$.

On déduit $\alpha \approx 1,96 \times 3,20 \approx 6,27$.

D'où $d = 8$ est extérieure à l'intervalle $I(0, 6, 27)$.

On doit donc estimer au seuil de 5% que le réglage de la machine a changé.

B. Comparaison de deux fréquences observées

Soit f_1 et f_2 les fréquences d'apparition d'une propriété donnée, dans des échantillons (E_1), d'effctifs n_1, et (E_2), d'effctifs n_2, tirés d'une même population.

Soit p la fréquence théorique de la population, relativement à cette propriété.

La variable aléatoire f', fréquence d'un échantillon de taille n, suit la loi normale de paramètres p et $\sqrt{\frac{p(1-p)}{n}}$.

Ainsi, f'_1, fréquence d'un échantillon de taille n_1, suit approximativement une loi normale de paramètres p et $\sqrt{\frac{f_1(1-f_1)}{n_1}}$. De même, f'_2, fréquence d'un échantillon de taille n_2, suit approximativement une loi normale de paramètres p et $\sqrt{\frac{f_2(1-f_2)}{n_2}}$.

Soit $d = f_1 - f_2$ la différence entre les deux fréquences observées et $d' = f'_1 - f'_2$ la variable aléatoire correspondante.

d' suit approximativement la loi normale de paramètres 0 et $\sqrt{\frac{f_1(1-f_1)}{n_1} + \frac{f_2(1-f_2)}{n_2}}$.

Au niveau du seuil retenu, on obtiendra encore un intervalle dans lequel doit se trouver d pour que l'on puisse admettre que les deux échantillons proviennent de la même population.

Exemple.

On cherche à étudier l'évolution dans le temps de la proportion de pièces hors tolérance produites par une machine.

Un premier échantillon est prélevé, d'effectif $n_1 = 125$ et pour lequel, on décèle 10 pièces défectueuses. Un deuxième échantillon est prélevé, d'effectif $n_2 = 200$ et pour lequel, on décèle 12 pièces défectueuses.

Peut-on, au seuil de 5%, estimer que la production a gardé la même qualité entre les deux prélèvements ?

Solution.

$f_1 = \frac{10}{125} = 0,08$ et $f_2 = \frac{12}{200} = 0,06$.

On peut estimer que les fréquences suivent une loi normale de paramètres p et $\sqrt{\frac{p(1-p)}{n}}$. La variable aléatoire f'_1 suivra approximativement une loi normale de paramètres p et $\sqrt{\frac{0,08\times0,92}{125}}$ et la variable aléatoire f'_2 suivra approximativement une loi normale de paramètres p et $\sqrt{\frac{0,06\times0,94}{200}}$.

La différence $d' = f'_1 - f'_2$ suivra par conséquent approximativement une loi normale de paramètres 0 et $\sqrt{\sqrt{\frac{0,08\times0,92}{125}} + \sqrt{\frac{0,06\times0,94}{200}}} \approx 0,03$.

Au seuil de 5%, on obtient l'intervalle $I(0, \underbrace{1,96 \times 0,03}_{\approx 0,058})$.

Or $d = f_1 - f_2 = 0,02$ est dans cet intervalle $I(0, \underbrace{1,96 \times 0,03}_{\approx 0,058})$.

Ainsi, au seuil de 5%, la qualité de la production n'a pas changé de façon significative.

Exercice 82 [1, exercice 40, p. 116] La production de sacs de grains du poids nominal de $50kgs$ suit une loi normale d'espérance mathématique $\mu = 50kgs$ et d'écart-type σ.

1. 95% des sacs ont un poids compris entre $48kgs$ et $52kgs$. Calculer l'écart-type σ.
2. On constitue des lots de 40 sacs pris au hasard dans la production.
   (a) Quelle est la loi suivie par la variable aléatoire $\overline{X}$, poids moyen des sacs dans un lot quelconque ?
   (b) Quelle est la probabilité p pour que ce poids moyen soit inférieur à $49,5kgs$?
3. Le contrat liant le fournisseur et un client stipule que le poids total d'un lot doit être d'au moins $2000kgs$.

   Le fournisseur cherche à remplir ce contrat avec la quasi-certitude aportée par la probabilité $q = 1 - p$.

Quelle devra être l'espérance mathématique μ qui permettra d'obtenir ce résultat, sans toucher à l'écart-type ?

Exercice 83 [1, exercice 48, p. 125] Une pièce est produite par deux machines, de même type et réglées pour la même production. La cote minimale des pièces est $84,20mm$, avec une tolérance de $\pm 0,20mm$.

On estime que la cote X d'une pièce suit, par machine, une loi normale.

Un échantillon de 200 pièces prélevées indistinctement sur l'ensemble de la production donne une moyenne $\overline{X} = 84,27mm$.

L'histogramme correspondant présente toutefois une dissymétrie par rapport à cette valeur, ce qui donne des doutes dur les réglages respectifs des deux machines.

On cherche donc, en séparant les deux productions, à comparer les résultats des deux machines dans le même temps, pour déceler un éventuel défaut d'homogénéité.

Le résultat est le suivant :

Machine	Effectif	Moyenne	Ecart-type	Nombre de pièces satisfaisantes
M_1	$n_1 = 106$	$m_1 = 84,21$	$s_1 = 0,127$	$k_1 = 94$
M_2	$n_2 = 94$	$m_2 = 84,34$	$s_2 = 0,129$	$k_2 = 71$

Peut-on accepter, au seuil de 5% l'hypothèse selon laquelle la production des deux machines est homogène :

1. en utilisant les moyennes ?
2. en utilisant les fréquences ?

Quelle décision prendre ?

Références

[1] H. EGON, *Statistiques et Probabilités, BTS Secteur industriel, DUT*, Hachette, 1992.

[2] E. FAVRO, *Statistiques et Probabilités*, Dunod, 1991.

[3] J. FOURASTIE, JF. LASLIER, *Probabilités et Statistiques*, Dunod, 1987.

[4] P. JAFFARD, *Initiation aux Méthodes de la Statistiques et du Calcul des Probabilités*, Masson, 1986.

[5] C. LEBOEUF, JL. ROQUE, J. GUEGAND, *Cours de Probabilités et Statistiques*, Ellipses, Edition Marketing, 1987.

[6] S. LIPSCHUTZ, *Probabilités, Cours et Problèmes*, Mc Graw-Hill, 1993.

[7] J. MARCEIL, *Aide Mémoire de Probabilités et Statistiques*, Ellipses, 1992.

[8] JC. RADIX, *Pratique Moderne des Probabilités*, Lavoisier Tec et Doc, 1991.

[9] C. RAFFIN, *Statistiques et Probabilités, DEUG Scientifiques 2e année*, Flash U, Armand Colin, 1987.

[10] M. ROSS, *Initiation aux Probabilités*, Presses Polytechniques et Universitaires romandes, 1988.

[11] A. RUEGG, *Probabilités et Statistiques*, Presses Polytechniques et Universitaires romandes, 1985.

[12] G. SAPORTA, *Probabilités, Analyse des Données et Statistique*, Editions Technip, 1990.

[13] C. TROUSSET, JF MORIN, *Mathématiques pour les Sciences de la Vie, Probabilités et Statistiques*, Mc Graw-Hill, 1991.

[14] C. VIGNERON, E. LOGAK *Probabilités et Statistiques 1*, Diderot Editeur, Art et Sciences, 1995.

[15] Exercices corrigés - Dénombrement et probabilités :
http ://www.bibmath.net/exercices/index.php3 ?action=affiche&quoi=proba

Table de la loi normale centrée réduite :

	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	0,7703	0,7734	0,7764	0,7793	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	0,8906	0,8925	0,8943	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	0,9986	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