ISIMA (1ère année)

PROBABILITES
(Exercices)

2021-2022

TABLE DES MATIERES

CHAPITRE 1 : ÉLEMENTS DE BASE DES PROBABILITES

Exercice 1 : Opérations ensemblistes

Soient A,B,C, trois événements. Exprimer en fonction de A,B,C et des opérations ensemblistes (intersection, union, complémentaire), les événements suivants:

1) A seul se produit
2) A et C se produisent, mais non B
3) les trois évènements se produisent
4) l'un au moins des évènements se produit
5) deux évènements au moins se produisent
6) un évènement au plus se produit
7) aucun des trois évènements ne se produit
8) deux évènements exactement se produisent
9) pas plus de deux évènements ne se produisent

1) $A \cap \overline{B} \cap \overline{C}$
2) $A \cap C \cap \overline{B}$
3) $A \cap B \cap C$
4) $A \cup B \cup C$
5) $(A \cap B) \cup (B \cap C) \cup (A \cap C)$ ou $(A \cap B \cap \overline{C}) \cup (\overline{A} \cap B \cap C) \cup (A \cap \overline{B} \cap C) \cup (A \cap B \cap C)$
6) $(A \cap \overline{B} \cap \overline{C}) \cup (\overline{A} \cap B \cap \overline{C}) \cup (\overline{A} \cap \overline{B} \cap C) \cup (\overline{A} \cap \overline{B} \cap \overline{C})$
7) $\overline{A} \cap \overline{B} \cap \overline{C}$
8) $(A \cap B \cap \overline{C}) \cup (B \cap C \cap \overline{A}) \cup (A \cap C \cap \overline{B})$
9) $\overline{A} \cup \overline{B} \cup \overline{C}$

Exercice 2 : Probabilités

Soit A , B et C trois évènements de probabilités $P(A) = 0.3$, $P(B) = 0.4$ et $P(C) = 0.7$. Supposons que l'on a de plus $P(A \cap B) = 0.2$ et $P_C(A) = 0.4$.

1) Calculer $P(\bar{A})$.

Solution : $P(\bar{A}) = 1 - P(A) = 0.7$.

2) Calculer $P(A \cup B)$.

Solution :

$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.3 + 0.4 - 0.2 = 0.5$.

3) Calculer $P_A(B)$ et $P_B(A)$.

Solution : $P_A(B) = \frac{P(A \cap B)}{P(A)} = \frac{0.2}{0.3} \simeq 0.67$. De même $P_B(A) = \frac{P(A \cap B)}{P(B)} = \frac{0.2}{0.4} = 0.5$.

4) Calculer $P_A(C)$.

Solution :

$P_A(C) = \frac{P(A \cap C)}{P(A)}$. Calculons $P(A \cap C)$. On a $P(A \cap C) = P(C \cap A) = P(C) \cdot P_C(A) = 0.7 \cdot 0.4 = 0.28$. Ceci implique $P_A(C) = \frac{0.28}{0.3} \simeq 0.9$

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

Solution :

$P(A \cup C) = P(A) + P(C) - P(A \cap C) = 0.3 + 0.7 - 0.28 = 0.72$.

Exercice 3 : Lancer un dé

On considère l'expérience aléatoire consistant à lancer un dé à 6 faces et à observer le résultat. On suppose dans un premier temps que le dé n'est pas pipé.

1) Quels sont les événements élémentaires ? Quel est l'univers des événements?

Solution : L'univers des événements, en général noté Ω , est l'ensemble des résultats possibles.

$$\Omega = \{1,2,3,4,5,6\}$$

Tout élément $\omega \in \Omega$ est un événement élémentaire :

$$\{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{6\}$$

2) Donner la probabilité de chaque événement élémentaire.

Solution : Le dé n'étant pas pipé, les événements élémentaires sont équiprobables.

$$P(\{\omega\}) = \frac{card(\{\omega\})}{card(\Omega)} = \frac{1}{6}, \forall \omega \in \Omega$$

Par la suite, nous noterons $P(\omega)$ pour $P(\{\omega\})$.

3) On considère les événements (non élémentaires) suivants :

- Evénement E_1 : Le résultat est pair
- Evénement E_2 : Le résultat est différent de 2
- Evénement E_3 : Le résultat est strictement inférieur à 10

Donner la probabilité de chacun de ces événéments.

Solution : Ω est un ensemble discret et fini. De plus, les événements élémentaires sont équiprobables. Nous pouvons donc écrire que la probabilité d'un événement A est :

$$P(A) = \frac{card(A)}{card(\Omega)}$$

Ici $card(E_1) = 3$, $card(E_2) = 5$, $card(E_3) = 6$ et $card(\Omega) = 6$. Il vient alors :

$$P(E_1) = \frac{1}{2}, P(E_2) = \frac{5}{6}, P(E_3) = 1$$

4) Calculer les probabilités des événements $E_1 \cap E_2$, $E_1 \cup E_2$ et $\bar{E_1}$

Solution : Ici $E_1 \cap E_2 = \{4,6\}$, $E_1 \cup E_2 = \{1,2,3,4,5,6\}$ et $\bar{E_1} = \{1,3,5\}$. Il vient, en utilisant à nouveau $P(A) = card(A) \,/\, card(\Omega)$:

$$P(E_1 \cap E_2) = 1 \,/\, 3, P(E_1 \cup E_2) = 1, P(\bar{E_1}) = 1 \,/\, 2$$

On aurait pu aussi calculer $P(E_1 \cup E_2)$ et $P(\bar{E_1})$ comme suit :

$$P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2) = 1 \,/\, 2 + 5 \,/\, 6 - 1 \,/\, 3 = 1$$
$$P(\bar{E_1}) = 1 - P(E_1) = 1 \,/\, 2$$

5) Calculer la probabilité que le résultat soit différent de 2, sachant que le résultat est pair.

Solution :

$$P_{E_1}(E_2) = \frac{P(E_1 \cap E_2)}{P(E_1)} = \frac{1 \,/\, 3}{1 \,/\, 2} = 2 \,/\, 3$$

6) On suppose maintenant que le dé est pipé. Les événements élémentaires ont désormais les probabilités suivantes :

$$P(1) = P(2) = P(3) = P(4) = \frac{1}{12}, P(5) = P(6) = \frac{1}{3}$$

Reprendre les questions précédentes.

Solution : Les événements élémentaires n'étant plus équiprobables, on ne peut plus utiliser la relation $P(A) = card(A) \,/\, card(\Omega)$. On peut en revanche toujours utiliser la relation $P(A_1 \cup A_2) = P(A_1) + P(A_2)$ lorsque les événements A_1 et A_2 sont incompatibles ($A_1 \cap A_2 = \emptyset$).

$$P(E_1) = \frac{1}{12} + \frac{1}{12} + \frac{1}{3} = \frac{1}{2}$$
$$P(E_2) = 1 - \frac{1}{12} = \frac{11}{12}$$
$$P(E_3) = 1$$
$$P(E_1 \cap E_2) = \frac{1}{12} + \frac{1}{3} = \frac{5}{12}$$
$$P(E_1 \cup E_2) = 1$$
$$P(\bar{E_1}) = \frac{1}{2}$$
$$P_{E_1}(E_2) = \frac{P(E_1 \cap E_2)}{P(E_1)} = \frac{5/12}{1/2} = \frac{5}{6}$$

Exercice 4 : Lancer deux dés

1) On lance deux dés non pipés. Quelle est la probabilité d'obtenir au moins un 6 sachant que les deux résultats sont différents ? Vous détaillerez votre modélisation en précisant bien quelle est l'expérience aléatoire et l'univers des événements élémentaires.

Solution :

Modélisation 1 Soit l'expérience aléatoire suivante. On considère deux dés distinguables (bleu et rouge). On lance ces deux dés et on observe le résultat obtenu sous la forme d'un couple (x_1, x_2) où x_1 est le résultat du dé bleu et x_2 est le résultat du dé rouge. Notons que le couple (x_1, x_2) est différent du couple (x_2, x_1) si $x_1 \neq x_2$.

L'univers des événements Ω est alors l'ensemble des couples possibles :

$$\Omega = \{1,2,3,4,5,6\} \times \{1,2,3,4,5,6\} .$$

Ainsi nous avons $card(\Omega) = 36$.

Soit A l'événement "il y a au moins un 6" et B l'événement "les deux dés donnent un résultat différent". Nous cherchons à calculer $P(A|B)$ dont la définition est

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

Un calcul rapide donne :

$card(A) = 11, card(B) = 30, card(A \cap B) = 10$

Les événements élémentaires étant équiprobables, on obtient alors

$$P(A) = \frac{card(A)}{card(\Omega)} = \frac{11}{36}, P(B) = \frac{30}{36} = \frac{15}{18}, P(A \cap B) = \frac{10}{36} = \frac{5}{18}$$

Nous pouvons conclure :

$$P(A|B) = 1 / 3$$

Modélisation 2 : Soit l'expérience aléatoire suivante. On considère deux dés identiques (par exemple deux dés bleus). On lance ces deux dés et on observe le résultat obtenu sous la forme d'une paire $\{x_1, x_2\}$ non ordonnée. Notons que la paire $\{x_1, x_2\}$ est alors équivalente à la paire $\{x_2, x_1\}$.

L'univers des événements élémentaires Ω est alors l'ensemble des paires posssibles et $card(\Omega) = 21$. Les événements élémentaires ne sont plus équiprobables. En effet, la probabilité d'obtenir la paire (1,1) est de 1/36 tandis que la probabilité d'obtenir la paire (1,2) est de 2/36=1/18.

Nous avons $A = \{\{1,6\}, \{2,6\}, \{3,6\}, \{4,6\}, \{5,6\}, \{6,6\}\}$. Les 5 premiers événements élémentaires de A sont de probabilité 1/18 tandis que la paire (6,6) a une probabilité de 1/36. Ainsi :

$$P(A) = 5 \times \frac{1}{18} + \frac{1}{36} = \frac{11}{36}$$

L'événement B correspond à 15 paires, chacune de probabilité 1/18, et on obtient bien $P(B) = 15/18$. Enfin, nous avons $A \cap B = \{\{1,6\}, \{2,6\}, \{3,6\}, \{4,6\}, \{5,6\}\}$ et $P(A \cap B) = 5/18$. On conclue de la même manière que dans la modélisation précédente.

2) Reprendre la question précédente avec des dés pipés. Les probabilités pour un dé sont les suivantes :

$$P(1) = P(2) = P(3) = P(4) = 1/12\,, P(5) = P(6) = 1\,/\,3$$

Solution : En supposant que les résultats des dés sont indépendants, on peut calculer facilement les probabilités $P(i,j)$ d'obtenir un couple (i,j) :

$$P(i,j) = P(i) \times P(j) = \begin{cases} \frac{1}{144} & si\ i \leq 4\ et\ j \leq 4 \\ \frac{1}{9} & si\ i \geq 5\ et\ j \geq 5 \\ \frac{1}{36} & si\ (i \leq 4\ \ et\ j \geq 5)\ ou\ (i \geq 5\ et\ j \leq 4) \end{cases}$$

Ainsi :

$$P(A \cap B) = \frac{8}{36} + \frac{2}{9} = \frac{4}{9}$$
$$P(\bar{B}) = \frac{4}{144} + \frac{2}{9} = \frac{1}{4}$$
$$P(B) = \frac{3}{4}$$
$$P_B(A) = \frac{P(A \cap B)}{P(B)} = \frac{16}{27}$$

Exercice 5 : Jumeaux et plus !

Une femme est enceinte et se rend chez l'échographe avec son mari. Surprise, elle attend des faux jumeaux (deux ovules fécondés). Cependant, elle ne veut pas connaître les sexe des enfants avant la naissance.

1) Quelle est la probabilité qu'elle attende un garçon et une fille ? Vous préciserez bien la modélisation.

Solution :

Modélisation 1 : On observe le résultat sous la forme d'un couple (x_1, x_2) où x_1 représente le sexe du premier enfant né et x_2 celui du deuxième enfant né. L'univers des événements élémentaires est

$$\Omega = \{(G,F), (G,G), (F,G), (F,F)\}$$

avec F pour fille et G pour garçon.

Soit l'événément $A = \{(G,F),(F,G)\}$ correspondant à avoir un garçon et une fille. Si l'on suppose que les événements élémentaires sont équiprobables, on obtient $P(A) = card(A)/card(\Omega) = 1\,/\,2$.

Modélisation 2: Soit A_i l'événement le i-ème enfant né est un garçon. Nous supposerons que les événements A_1 et A_2 sont indépendants. Soit B l'événement "avoir une fille et un garçon (peu importe l'ordre)". Nous avons :

$$B = (A_1 \cap \bar{A}_2) \cup (\bar{A}_1 \cap A_2)$$

Comme $(A_1 \cap \bar{A}_2) \cap (\bar{A}_1 \cap A_2) = \emptyset$, nous obtenons :

$$P(B) = P(A_1 \cap \bar{A}_2) + P(\bar{A}_1 \cap A_2) = P(A_1)P(\bar{A}_2) + P(\bar{A}_1)P(A_2) = \tfrac{1}{2} \times \tfrac{1}{2} + \tfrac{1}{2} \times \tfrac{1}{2} = \tfrac{1}{2}$$

Nous avons utilisé le fait que les événements A_1 et A_2 étaient indépendants dans la deuxième égalité.

2) Quelle est la probabilité que les deux enfants soient des garçons sachant qu'au moins l'un des deux est un garçon ?

Solution :

Modélisation 1 :

Soit B l'événement "les deux enfants sont des garçons" et C l'événement "au moins l'un des deux est un garçon". Nous avons :

$$B = \{(G,G)\} \qquad C = \{(G,F),(G,G),(F,G)\} \qquad B \cap C = \{(G,G)\}$$

puis

$$P(B|C) = \frac{P(B \cap C)}{P(C)} = \frac{1\,/\,4}{3\,/\,4} = \frac{1}{3}$$

Modélisation 2 : Avec la 2ème modélisation, nous avons :

$$B = A_1 \cap A_2, C = A_1 \cup A_2$$

$$\begin{aligned} P(B|C) &= \frac{P((A_1 \cap A_2) \cap (A_1 \cup A_2))}{P(A_1 \cup A_2)} \\ &= \frac{P(A_1 \cap A_2)}{P(A_1) + P(A_2) - P(A_1)P(A_2)} \\ &= \frac{P(A_1)P(A_2)}{P(A_1) + P(A_2) - P(A_1)P(A_2)} \\ &= \frac{1}{3} \end{aligned}$$

3) Quelle est la probabilité que les deux enfants soient des garçons sachant que le premier enfant né est un garçon ?

Solution :

Modélisation 1 : Soit D l'événement "le premier enfant né est un garçon".

$$D = \{(G,F),(G,G)\}, B \cap D = \{(G,G)\}$$

Puis :

$$P(B|D) = \frac{P(B \cap D)}{P(D)} = \frac{1\,/\,4}{1\,/\,2} = 1\,/\,2$$

Modélisation 2 : Avec la 2ème modélisation, nous avons :

$$P(A_1 \cap A_2|A_1) = \frac{P((A_1 \cap A_2) \cap A_1)}{P(A_1)} = \frac{P(A_1 \cap A_2)}{P(A_1)} = \frac{P(A_1)P(A_2)}{P(A_1)} = P(A_2) = \frac{1}{2}$$

4) Reprendre les questions précédentes en supposant qu'il naît 106 garçons pour 100 filles.

Solution : Avec la 2ème modélisation, nous avons désormais :

$$P(A_i) = \frac{106}{100 + 106} = \frac{53}{103} \simeq 0.515$$

Les calculs sont alors identiques :

$$P(A) \simeq 0.5$$

$$P(B|C) \simeq 0.346$$

$$P(B|D) \simeq 0.515$$

Exercice 6 : Fiabilité de composants en série ou en parallèle

On considère un système constitué de 3 composants identiques. Le système fonctionne tant que le composant 1 fonctionne ET au moins l'un des deux autres composants fonctionnent. On sait que la probabilité qu'un composant ne tombe pas en panne au cours de l'année est de 0.8.

On suppose que les composants fonctionnent indépendamment. Quelle est la probabilité que le système tombe en panne au cours de l'année ? Vous justifierez soigneusement votre raisonnement en définissant bien les événements que vous considérez.

Solution : Soit A_i l'événement "le composant i fonctionne". Nous avons $P(A_i) = 0.8$. Soit B l'événement "le composant 1 fonctionne ET au moins l'un des deux autres composants fonctionnent". Nous avons :

$$B = A_1 \cap (A_2 \cup A_3)$$

Remarquons que les événements A_1 et $(A_2 \cup A_3)$ sont indépendants :

$$\begin{aligned} P(B) &= P((A_1 \cap A_2) \cup (A_1 \cap A_3)) \\ &= P(A_1 \cap A_2) + P(A_1 \cap A_3) - P(A_1 \cap A_2 \cap A_3) \\ &= P(A_1)P(A_2) + P(A_1)P(A_3) - P(A_1)P(A_2)P(A_3) \\ &= 0.8 \times [0.8 + 0.8 - 0.8^2] = 0.768 \end{aligned}$$

On conclut alors que :

$$P(\bar{B}) = 1 - P(B) = 0.232$$

On aurait pu aussi partir dans l'autre sens :

$$\begin{aligned} P(\bar{B}) &= P(\bar{A}_1 \cup (\bar{A}_2 \cap \bar{A}_3)) \\ &= P(\bar{A}_1) + P(\bar{A}_2 \cap \bar{A}_3) - P(\bar{A}_1 \cap \bar{A}_2 \cap \bar{A}_3) \\ &= P(\bar{A}_1) + P(\bar{A}_2)P(\bar{A}_3) - P(\bar{A}_1)P(\bar{A}_2)P(\bar{A}_3) \\ &= 0.2 + 0.2^2 - 0.2^3 = 0.232 \end{aligned}$$

Exercice 7 : Système de contrôle de pièces

Dans un lot de pièces fabriquées, il y a 5% de pièces défectueuses. On contrôle les pièces mais le mécanisme du contrôle est aléatoire : si la pièce est bonne, elle est acceptée avec une probabilité égale à 0.96; si la pièce est mauvaise, elle est refusée avec une probabilité égale à 0.98. On choisit au hasard une pièce que l'on contrôle. Déterminer les probabilités pour que :

1) il y ait une erreur dans le contrôle,

Solution : Soit B l'évènement "la pièce contrôlée est bonne", $\bar{B}$ l'évènement "la pièce contrôlée est mauvaise", A l'évènement "la pièce contrôlée est acceptée", et $\bar{A}$ l'évènement "la pièce contrôlée est refusée". L'énoncé nous dit que $P(B) = 0{,}95$, $P(\bar{B}) = 1 - P(B) = 0{,}05$, $P_B(A) = 0{,}96$, et $P_{\bar{B}}(\bar{A}) = 0{,}98$.

La probabilité qu'il y ait une erreur dans le contrôle est égale à

$$P((B \cap \bar{A}) \cup (\bar{B} \cap A)) = P(B \cap \bar{A}) + P(\bar{B} \cap A),$$

car les évènements $B \cap \bar{A}$ et $\bar{B} \cap A$ sont incompatibles. Or on a

$$P(B \cap \bar{A}) = P(B) \times P_B(\bar{A}) = P(B) \times (1 - P_B(A))$$

et

$$P(\bar{B} \cap A) = P(\bar{B}) \times P_{\bar{B}}(A) = P(\bar{B}) \times (1 - P_{\bar{B}}(\bar{A})),$$

d'où

$P(B \cap \bar{A}) + P(\bar{B} \cap A) = 0{,}95 \times 0{,}04 + 0{,}05 \times 0{,}02 = 0{,}039.$

2) la pièce soit bonne sachant qu'elle est refusée,

Solution : On a

$$P_{\bar{A}}(B) = P(\bar{A} \cap B) \,/\, P(\bar{A})$$

et

$$P(\bar{A}) = P(\bar{A} \cap B) + P(\bar{A} \cap \bar{B}) = P(B) \times P_B(\bar{A}) + P(\bar{B}) \times P_{\bar{B}}(\bar{A})$$

d'où

$$P_{\bar{A}}(B) = \frac{1}{1 + \frac{0{,}05 \times 0{,}98}{0{,}95 \times 0{,}04}} \simeq 0{,}437.$$

3) la pièce soit mauvaise sachant qu'elle est acceptée. Que pensez-vous de ce système de contrôle ?

Solution : On a

$$P_A(\bar{B}) = P(\bar{B} \cap A) \;/\; P(A)$$

et

$$P(\bar{B} \cap A) = P(\bar{B}) \times P_{\bar{B}}(A),$$

d'où

$$P_A(\bar{B}) = \frac{1}{1 + \frac{0,95 \times 0,96}{0,05 \times 0,02}} \simeq 0,001$$

grâce à la question précédente.

En conclusion, on peut dire que ce système de contrôle est très fiable, dans le sens ou il ne laisse passer que très peu de mauvaises pièces : plus de 99% des pièces acceptées sont bonnes d'après la question précédente. Cependant, ce système de contrôle fait beaucoup de gâchis : plus de 40 % des pièces refusées sont en fait bonnes d'après la question 1 !

Exercice 8 : Pièces défectueuses

Trois machines M_1 , M_2 et M_3 produisent respectivement 50%, 30% et 20% du nombre total de pièces fabriquées dans cet atelier. Les pourcentages de pièces défectueuses de ces machines sont 3%, 4% et 5%.

1) Si on prend une pièce au hasard, quelle est la probabilité pour que cette pièce soit défectueuse ?

Solution : Soit HS l'évènement "la pièce prises est défectueuse", et M_i l'évènement "la pièce a été produite par la machine M_i ", $i = 1,2,3$. L'énoncé nous donne les valeurs de $P \quad (HS)$ ainsi que $P(M_i)$, pour $i = 1,2,3$.

La probabilité que la pièce prise soit défectueuse est

$$\begin{aligned} P(HS) \quad &= \sum_{i=1}^{3} P\,(HS \cap M_i) \\ &= \sum_{i=1}^{3} P_{M_i}\,(HS) \times P(M_i) \\ &= 0,03 \times 0,5 + 0,04 \times 0,3 + 0,05 \times 0,2 \\ &= 0,037. \end{aligned}$$

Il y a donc 3,7% de pièces défectueuses.

2) Quelle est la probabilité qu'une pièce défectueuse ait été produite par la machine M_1 ?

Solution : La probabilité que la pièce ait été fabriquée par M_1 sachant qu'elle est défectueuse est égale à

$$\begin{aligned} P_{HS}(M_1) \quad &= P(M_1 \cap HS) \;/\; P(HS) \\ &= P_{M_1}(HS) \times P(M_1) \;/\; P(HS) \\ &= 0,03 \times 0,5/0,037 \\ &= 0,405. \end{aligned}$$

Si une pièce est défectueuse, alors il y a 40,5 % de chances qu'elle ait été fabriquée par la machine M_1 .

Exercice 9 : Le meilleur fournisseur

Pour une même pièce, un industriel s'adresse à 2 fournisseurs A et B. Le fournisseur A fournit 75% des besoins de l'industriel et le pourcentage de pièces défectueuses qu'il produit est de 1%. Le pourcentage de pièces défectueuses parmi les pièces produites par B est de 2%.

L'industriel prend une pièce au hasard dans son stock.

1) Quelle est la probabilité que cette pièce soit défectueuse ?

Solution : Soit A l'évènement "la pièce prise provient du fournisseur A ", B l'évènement "la pièce prise provient du fournisseur B ", OK l'évènement "la pièce prise est bonne", et HS l'évènement "la pièce prise est défectueuse". On a bien sûr $A = \bar{B}$ et $OK = \bar{HS}$. L'énoncé nous donne les valeurs de $P(A)$, $P(B)$, $P_A(HS)$, et $P_B(HS)$.

La probabilité qu'une pièce soit défectueuse est

$$\begin{aligned} P(HS) &= P(HS \cap A \cup HS \cap B) \\ &= P(HS \cap A) + P(HS \cap B) \;\; \textit{car les événements sont disjoints} \\ &= P(A) \times P_A(HS) + P(B) \times P_B(HS) \\ &= 0{,}75 \times 0{,}01 + 0{,}25 \times 0{,}02 \\ &= 0{,}0125. \end{aligned}$$

2) Dans le cas où elle est bonne, quelle est la probabilité qu'elle provienne du fournisseur B ?

Solution : La probabilité qu'une pièce bonne ait été produite par le fournisseur B est

$$\begin{aligned} P_{OK}(B) &= P(B \cap OK) \,/\, P(OK) \\ &= P_B(OK) \times P(B) \,/\, P(OK) \\ &= 0{,}98 \times 0{,}25 \,/\, (1 - 0{,}0125) \\ &\simeq 0{,}248. \end{aligned}$$

3) Dans le cas où elle est défectueuse, quelle est la probabilité qu'elle provienne de B ?

Solution : La probabilité qu'une pièce défectueuse provienne de B est

$$\begin{aligned} P_{HS}(B) &= P(B \cap HS) \,/\, P(HS) \\ &= P_B(HS) \times P(B) \,/\, P(HS) \\ &= 0{,}02 \times 0{,}25 \,/\, 0{,}0125 \\ &\simeq 0{,}4. \end{aligned}$$

4) L'industriel constate qu'elle est défectueuse et s'exclame "elle doit provenir de chez B". Ce jugement vous parait-il valable ?

Solution : Le raisonnement ne tient pas debout : B a certes un taux de défaillance supérieur à celui de A, mais il ne faut tout de même pas oublier que c'est A qui fournit la majorité des pièces à l'industriel. Ces deux effets se compensent, et, d'après la question précédente, il n'y a que 2 chances sur 5 qu'une pièce défectueuse provienne de B.

Exercice 10 : Formule de Sylvester-Poincaré

On considère un ensemble de pièces géométriques qui peuvent avoir trois formes différentes (carrée, triangulaire ou circulaire), quatre couleurs (rouge, vert, bleu ou jaune), deux tailles (petite ou grande) et deux épaisseurs (fine ou épaisse). Chaque pièce étant unique, le nombre de pièces est donc de $3 \times 4 \times 2 \times 2 = 48$.

On choisit une pièce au hasard. Quelle est la probabilité pour qu'elle soit petite ou fine ou ni verte ni triangulaire ?

Solution :

Notations des évènements :

"Petite" : P ; "Fine" : F ; "Non verte" : $\bar{V}$; "Non triangulaire" : $\bar{T}$;

La probabilité demandée est $\mathrm{P}(P \cup F \cup (\bar{V} \cap \bar{T}))$. D'après la formule de Sylvester-Poincaré, on a :

$\mathrm{P}(P \cup F \cup (\bar{V} \cap \bar{T})) = \mathrm{P}(P) + \mathrm{P}(F) + \mathrm{P}(\bar{V} \cap \bar{T}) - \mathrm{P}(P \cap F) - \mathrm{P}(P \cap \bar{V} \cap \bar{T}) - \mathrm{P}(F \cap \bar{V} \cap \bar{T})$
$+ \mathrm{P}(P \cap F \cap \bar{V} \cap \bar{T}) = \frac{24}{48} + \frac{24}{48} + \frac{24}{48} - \frac{12}{48} - \frac{12}{48} - \frac{12}{48} + \frac{6}{48} = \frac{42}{48} = \frac{7}{8}$.

Exercice 11 : Combinatoire

On lance trois dés non truqués. Soient les événements suivants :

A : "on obtient au moins un 1" ;

B : "on obtient au moins deux faces identiques" ;

C : "la somme des nombres obtenus est paire".

1°) Calculer $P(A), P(B), P(C)$.

2°) Calculer $P(A \cap B), P(A \cap C), P(B \cap C), P(A \cap B \cap C)$.

3°) Calculer $P(A \cup B), P(A \cup C), P(B \cup C), P(A \cup B \cup C)$.

Solution :

1°) L'univers des possibles est $\Omega = [\![1, 6]\!]^3$. Comme tous les lancers sont équiprobables, on a

$P(A) = \frac{Card(A)}{Card(\Omega)}, P(B) = \frac{Card(B)}{Card(\Omega)}$ et $P(C) = \frac{Card(C)}{Card(\Omega)}$.

$\mathrm{Card}(\Omega) = 6^3 = 216$.

$\mathrm{Card}(\bar{A}) = 5^3 = 125$, d'où $\mathrm{Card}(A) = \mathrm{Card}(\Omega) - \mathrm{Card}(\bar{A}) = 91$ et $\mathbf{P}(A) = \frac{91}{216}$.

On peut déterminer $\mathrm{Card}(B)$ de deux manières :

$\mathrm{Card}(\bar{B}) = 6 \times 5 \times 4 = 120$, d'où $\mathrm{Card}(B) = 216 - 120 = 96$;

$B = B_2 \cup B_3$, où B_2 désigne l'événement " on obtient deux faces identiques" et B_3 l'événement " on obtient trois faces identiques".

Alors, Card(B_2) = $C_3^2 \times 6 \times 5 = 90$ (C_3^2 : nombre de choix des deux faces identiques, 6 : nombre de choix pour la valeur commune de ces deux faces, 5 : nombre de choix pour la valeur de l'autre face) et Card(B_3) = 6, d'où le résultat.

On a donc $\mathbf{P}(B) = \frac{96}{216} = \frac{12}{27} = \frac{4}{9}$.

Enfin $C = C_3 \cup C_1$, où C_3 est l'événement "les trois nombres obtenus sont pairs" et C_1 l'événement "un seul des nombres obtenus est pair".

On a Card(C_3) = $3^3 = 27$ et Card(C_1) = $C_3^1 \times 3^2 \times 3 = 81$ (C_3^1 : nombre de choix du dé qui donne un nombre pair, 3^2: nombre de choix pour les deux lancers impairs, 3 : nombre de choix pour le lancer pair).

On a donc $\mathbf{P}(C) = \frac{108}{216} = \frac{1}{2}$.

2°) Comme $(A \cap B) \cup (\bar{A} \cap B) = (A \cup \bar{A}) \cap B = \Omega \cap B = B$, on a

Card($A \cap B$) = Card(B) – Card($\bar{A} \cap B$).

Or Card($\bar{A} \cap B$) = Card($\bar{A} \cap B_2$) + Card($\bar{A} \cap B_3$) = $C_3^2 \times 5 \times 4 + 5 = 65$.

Finalement Card($A \cap B$) = 96 – 65 = 31, d'où $\mathbf{P}(A \cap B) = \frac{31}{216}$.

Remarques :

a) une autre possibilité pour calculer Card($\bar{A} \cap B$) est la suivante :

Card($\bar{A} \cap B$) = Card($\bar{A}$) – Card($\bar{A} \cap \bar{B}$) = 125 – (5×4×3) = 65.

b) On peut calculer directement Card($A \cap B$) sans passer par $\bar{A}$ en posant $A = A_1 \cup A_2 \cup A_3$, où A_i désigne l'évènement "on obtient exactement i as".

Alors $A \cap B = (A_1 \cup A_2 \cup A_3) \cap (B_2 \cup B_3)$, d'où $Card(A \cap B) = \sum_{\substack{i \in \{1,2,3\} \\ j \in \{2,3\}}} Card(A_i \cap A_j)$, les valeurs de $Card(A_i \cap A_j)$étant consignées dans le tableau ci-dessous :

	A_1	A_2	A_3
B_2	$C_3^1 \times 5$	$C_3^2 \times 5$	0
B_3	0	0	1

<u>Première possibilité :</u>

Card ($\bar{A} \cap C$) = Card ($\bar{A} \cap C_1$) + Card ($\bar{A} \cap C_3$) = $C_3^1 \times 3 \times 2^2 + 3^3 = 36 + 27 = 63$

$= C_3^1 \times 3 \times 2^2 + 3^3 = 36 + 27 = 63$.

Dans le premier terme de la somme C_3^1 est le nombre de manière de choisir le dé donnant le nombre pair, 3 est le nombre de choix du nombre pair et 2^2 est le nombre de choix des deux nombres impairs restant (qui ne sont pas des as). Dans le second terme, il y a trois nombres pairs possibles pour chaque dé.

Donc Card ($A \cap C$) = Card (C) – Card ($\bar{A} \cap C$) = 108 – 63 = 45, et $\mathbf{P}(A \cap C) = \frac{45}{216} = \frac{5}{24}$.

Deuxième possibilité :

$A \cap C = (A_1 \cup A_2 \cup A_3) \cap (C_1 \cup C_3)$, d'où Card($A \cap C$) = $\sum_{j \in \{1,3\}}$ $i \in \{1,2,3\}$, les valeurs de Card($A_i \cap C_j$) étant consignées dans le tableau ci-dessous :

	A_1	A_2	A_3
C_1	$C_3^1 \times 3 \times C_2^1 \times 2$	$C_3^2 \times 3$	0
C_3	0	0	0

Dans la case donnant Card($A_1 \cap C_1$), C_3^1 est le nombre de manière de choisir le dé donnant le nombre pair, 3 est le nombre de choix du nombre pair, C_2^1 est le nombre de choix restant pour le dé donnant l'as et 2 et le nombre de choix du deuxième nombre impair, qui n'est pas un as.

Card ($B \cap C$) = Card $\big((B_2 \cup B_3) \cap (C_1 \cup C_3)\big)$

= Card($B_2 \cap C_1$) + Card($B_2 \cap C_3$) + Card($B_3 \cap C_1$) + Card($B_3 \cap C_3$)

= $C_3^2 \times 3 \times 3 + C_3^2 \times 3 \times 2 + 0 + 3$

= 48.

$\mathbf{P}(B \cap C) = \frac{48}{216} = \frac{6}{27} = \frac{2}{9}$.

Card ($A \cap B \cap C$) + Card ($\bar{A} \cap B \cap C$) = Card ($B \cap C$)

$\Rightarrow$ Card ($A \cap B \cap C$) = Card ($B \cap C$) – Card ($\bar{A} \cap B \cap C$).

Or Card ($\bar{A} \cap B \cap C$) = Card $\big(\bar{A} \cap (B_2 \cup B_3) \cap (C_1 \cup C_3)\big)$

= Card($\bar{A} \cap B_2 \cap C_1$) + Card($\bar{A} \cap B_2 \cap C_3$) + Card($\bar{A} \cap B_3 \cap C_1$) + Card($\bar{A} \cap B_3 \cap C_3$)

= $C_3^2 \times 2 \times 3 + C_3^2 \times 3 \times 2 + 0 + 3 = 39$.

Donc Card ($A \cap B \cap C$) = 48 – 39 = 9 et $\mathbf{P}(A \cap B \cap C) = \frac{9}{216} = \frac{1}{24}$.

3°) $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B) = \frac{91+96-31}{216} = \frac{156}{216} = \frac{13}{18}$.

$\mathbf{P}(A \cup C) = \mathbf{P}(A) + \mathbf{P}(C) - \mathbf{P}(A \cap C) = \frac{91+108-45}{216} = \frac{154}{216} = \frac{77}{108}$.

$\mathbf{P}(B \cup C) = \mathbf{P}(B) + \mathbf{P}(C) - \mathbf{P}(B \cap C) = \frac{96+108-48}{216} = \frac{156}{216} = \frac{13}{18}$.

$\mathbf{P}(A \cup B \cup C) = \mathbf{P}(A) + \mathbf{P}(B) + \mathbf{P}(C) - \mathbf{P}(A \cap B) - \mathbf{P}(A \cap C) - \mathbf{P}(B \cap C) + \mathbf{P}(A \cap B \cap C)$

$$= \frac{91+96+108-31-45-48+9}{216} = \frac{180}{216} = \frac{5}{6}.$$

Remarque :

On peut aussi calculer Card($\overline{A \cup B \cup C}$) = Card($\bar{A} \cap \bar{B} \cap \bar{C}$) en posant $\bar{C} = C_0 \cup C_2$ où C_i représente l'évènement "on obtient *i* nombres pairs".

On a Card($\bar{A} \cap \bar{B} \cap C_0$) = 0 et Card($\bar{A} \cap \bar{B} \cap C_2$) = $C_3^2 \times 3 \times 2 \times 2$, où C_3^2 représente le nombre de manières de choisir les deux dés donnant des nombres pairs, 3×2 représente le nombre de choix possibles pour ces deux nombres pairs forcément distincts ($\bar{B}$) et le dernier 2 représente le nombre de choix possible pour le nombre impair, différent de l'as ($\bar{A}$).

Exercice 12 : Magiciens

Douze magiciens s'assoient autour d'une table ronde. Cinq sont des magiciens blancs et les sept autres des magiciens noirs. Ils s'assoient au hasard autour de la table.

1°) Quelle est la probabilité pour qu'aucun magicien blanc ne soit encadré par deux magiciens noirs ?

2°) Ils décident de se séparer en trois groupes de quatre. Quelle est la probabilité pour qu'un groupe ne comporte que des magiciens noirs ?

Solution :

1°) Le nombre total de répartitions possibles autour de la table est le nombre de permutations de 12 éléments, soit 12!.

Pour qu'aucun magicien blanc ne soit encadré par deux magiciens noirs, il faut que, soit les 5 magiciens blancs soient groupés, côte à côte, soit qu'ils se divisent en 2 groupes de 2 et 3 magiciens placés côte à côte (dans toutes les autres configurations, un magicien blanc se retrouve isolé). Pour la première configuration, compte tenu de la circularité de la table, il y a 12 positions possibles du groupe. Pour chaque position, il y a 5! répartitions possibles des magiciens blancs, et pour chaque répartition des magiciens blancs il y a 7! répartitions possibles des magiciens noirs. Il y a donc $12 \times 5! \times 7!$ configurations du premier type.

Pour la seconde configuration, il faut d'abord compter le nombre de positions possibles sur la table : il y a 12 manières de placer le groupe de 3, puis une fois celui-ci placé, il n'y a plus que 6 manières de placer le groupe de 2 (il ne doit pas être contigu au groupe de 3 : faire un dessin) ; Une fois ces groupes placés, il faut répartir les magiciens blancs dans ces deux groupes : il y a C_5^3 manières de choisir les magiciens qui seront dans le groupe de 3 (les groupes sont donc formés) ; il y a ensuite $3! \times 2!$ manières de positionner chaque magicien à chaque place. Il y a donc $12 \times 6 \times C_5^3 \times 3! \times 2\,! \times 7!$ configurations du deuxième type.

Le nombre total de configurations favorables est donc :

$12 \times 5! \times 7! + 12 \times 6 \times C_5^3 \times 3! \times 2\,! \times 7! = 12 \times 7! \times (5! + 6 \times C_5^3 \times 3! \times 2\,!)$

$12 \times 7! \times (120 + 6 \times 10 \times 6 \times 2) = 840 \times 12 \times 7!$.

La probabilité est finalement $\frac{840 \times 12 \times 7!}{12!} = \frac{840}{11 \times 10 \times 9 \times 8} = \frac{21}{11 \times 9 \times 2} = \frac{7}{66}$.

<u>Remarque</u> : Des simplifications peuvent être apportées à cette solution :

On peut remarquer que le nombre de manières de faire un groupe de 2 et un groupe de 3, puis de répartir les 5 magiciens blancs dans ces deux groupes, $C_5^3 \times 3! \times 2\,!$, revient tout simplement au nombre de permutations des 5 magiciens blancs, c'est-à-dire 5!. Le nombre total de configurations favorables devient donc :

$12 \times 7! \times (5! + 6 \times C_5^3 \times 3! \times 2\,!) = 12 \times 7! \times 5! \times (1 + 6) = 12 \times 7 \times 7\,! \times 5\,!$

Et comme le nombre de cas possibles est 12 ! la probabilité est $\frac{12 \times 7 \times 7! \times 5!}{12!}$.

On peut aussi ne pas considérer le nombre de manières de répartir les magiciens noirs. En effet, leur répartition n'intervient pas dans le problème dont l'énoncé aurait aussi pu être : « Cinq magiciens s'installent sur une table ronde qui contient 12 places. Calculer la probabilité pour qu'aucun magicien blanc ne soit isolé. »

Dans ce cas, le nombre de cas favorables devient $12 \times 7 \times 5\,!$ et le nombre de cas possibles est le nombre de manières de répartir les 5 magiciens blancs autour de la table en tenant compte de l'ordre, c'est-à-dire l'arrangement $A_{12}^5 = \frac{12!}{7\,!}$. On retrouve bien sûr la même probabilité.

Enfin, on peut remarquer que l'ordre des magiciens n'intervient pas dans le problème. A ce moment, il est inutile de dénombrer le nombre de manières de répartir les 5 magiciens blancs dans leur(s) groupe(s), et le nombre de cas favorables devient très simplement 12×7, c'est-à-dire le nombre de manières de placer les groupes. En revanche, le nombre de cas possibles est le nombre de manières de répartir les 5 magiciens blancs autour de la table sans tenir compte de l'ordre, c'est-à-dire la combinaison $\binom{12}{5} = \frac{12!}{7\,! \times 5\,!}$. On retrouve encore la même probabilité.

2°) Le nombre total de manières de former trois groupes peut s'obtenir en prenant quatre magiciens parmi 12 (ni l'ordre des magiciens, ni l'ordre des groupes, n'intervient), puis pour chacun de ces choix, quatre autres magiciens parmi les huit restants.

Le nombre manières de former trois groupes avec un groupe ne comportant que des magiciens noirs peut s'obtenir en prenant quatre magiciens parmi les 7 magiciens noirs, puis pour chacun de ces choix, quatre autres magiciens parmi les huit restants.

La probabilité demandée est donc $\frac{C_7^4 \times C_8^4}{C_{12}^4 \times C_8^4} = \frac{C_7^4}{C_{12}^4} = \frac{A_7^4}{A_{12}^4} = \frac{7\times6\times5\times4}{12\times11\times10\times9} = \frac{7}{99}$.

Exercice 13 : Trois dés

1°) On lance trois dés, respectivement rouge, vert et bleu. On forme un nombre entre 111 et 666 en prenant le tirage du dé rouge (respectivement, vert et bleu) comme chiffre des centaines (respectivement, dizaines et unités). Quelle est la probabilité pour que le nombre obtenu soit supérieur ou égal à 421.

2°) Même question en lançant trois dés de couleur identique et en formant le plus grand nombre possible avec leurs trois tirages.

Solution :

1°) Il y a $6^3 = 216$ tirages possibles. Le nombre de tirages favorables se compte ainsi :

le dé rouge fait 5 ou 6 : $2 \times 6^2 = 72$ possibilités ;

le dé rouge fait 4 et le dé vert fait 2, 3, 4, 5 ou 6 : $5 \times 6 = 30$ possibilités.

La probabilité demandée est donc $\frac{102}{216} = \frac{17}{36}$.

2°) Le problème est ici que les nombres obtenus ne sont pas équiprobables, et il est préférable de se replacer dans le cas de l'équiprobabilité, où il y a 216 triplets possibles :

les 6 triplets où les trois chiffres sont identiques fournissent 6 nombres ;

les $3 \times 6 \times 5 = 90$ triplets où seulement deux chiffres sont identiques et le troisième diffère ne fournissent que $6 \times 5 = 30$ nombres car chaque nombre correspond à trois triplets (par exemple (2, 4, 4), (4, 2, 4) et (4, 4, 2) correspondent tous les trois au nombre 442) ;

les $6 \times 5 \times 4 = 120$ triplets où tous les chiffres sont distincts ne fournissent que $5 \times 4 = 20$ nombres car chaque nombre correspond à six triplets (par exemple (2, 4, 1), (2, 1, 4), (4, 1, 2), (4, 2, 1), (1, 4, 2) et (1, 2, 4) correspondent tous les six au nombre 421).

Comptons donc le nombre de triplets correspondant aux nombres supérieurs ou égaux à 421 :

parmi les nombres avec 3 chiffres identiques, 3 d'entre eux (444, 555 et 666) sont supérieurs ou égaux à 421 ;

parmi les nombres dont deux chiffres sont identiques et le troisième différent, 511, 611, 422, 522, 622, 433, 533, 633, et les 3×5 nombres commençant par 44, 55 ou 66 sont supérieurs ou égaux à 421, soit un total de 23 nombres et donc de $23 \times 3 = 69$ triplets ;

comme on ordonne les chiffres pour obtenir le nombre maximal, tous les nombres avec trois chiffres distincts sont supérieurs ou égaux à 421, sauf 321, soit 19 nombres, donc $19 \times 6 = 114$ triplets.

La probabilité demandée est donc $\frac{114+69+3}{216} = \frac{186}{216} = \frac{93}{108}$.

Exercice 14 : Théorème des chapeaux

Soit $n \in \mathbb{N}^*$. On considère n invités qui laissent leur chapeau au vestiaire et repartent les uns après les autres en prenant un chapeau au hasard.

Pour $i \in [\![1, n]\!]$, notons A_i l'événement "l'invité i repart avec son chapeau" et A l'événement "aucun invité ne repart avec son chapeau". On souhaite calculer $P(A)$.

1°) Exprimer l'événement $\bar{A}$ en fonction des événements $A_1, \ldots, A_n$.

2°) Calculer $P(A_i), P(A\_i \cap A_j), \ldots, P(A_{i_1} \cap A_{i_2} \ldots \cap A_{i_k})$.

3°) En déduire $P(\bar{A})$ en appliquant la formule de Sylvester-Poincaré.

4°) En déduire $P(A)$ ainsi que sa limite quand n tend vers l'infini.

Solution :

1°) $\bar{A} = \bigcup_{i=1}^{n} A_i$.

2°) $\forall\ i \in [\![1, n]\!]$, $\mathbf{P}(A_i) = \frac{Card(A_i)}{Card(\Omega)}$, $\mathrm{Card}(\Omega) = n!$ et $\mathrm{Card}(A_i) = (n-1)!$, donc $\mathbf{P}(A_i) = \frac{1}{n}$.

De même, $\mathbf{P}(A_i \cap A_j) = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}$,

et $P(A_{i_1} \cap A_{i_2} \ldots \cap A_{i_k}) = \frac{(n-k)!}{n!} = \frac{1}{n(n-1)\ldots(n-k+1)}$.

3°) $\mathbf{P}(\bar{A})$ s'obtient en appliquant la formule de Sylvester-Poincaré :

$\mathbf{P}(\bar{A}) = \mathbf{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{k=1}^{n}(-1)^{k-1} \sum_{J \in P_k(\{1,\ldots,n\})} P\left(\bigcap_{j\in J} A_j\right)$, où $\mathcal{P}_k(E) = \{ P \in \mathcal{P}(E) \ / \ \mathrm{Card}(P) = k \}$,

c'est-à-dire $\mathbf{P}(\bar{A}) = \mathbf{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k) - \ldots$

$+ (-1)^{k-1} \sum_{i_1<i_2<\ldots<i_k} P(A_{i_1} \cap A_{i_2} \ldots \cap A_{i_k})$

$+ \ldots + (-1)^{n-1} P(\cap A_i)$.

Donc, pour tout $J \in \mathcal{P}_k([\![1, n]\!])$, $P\left(\bigcap_{j\in J} A_j\right) = \frac{(n-k)!}{n!}$, et comme $\mathrm{Card}\left(\mathcal{P}_k([\![1, n]\!])\right) = C_n^k$,

on a $\sum_{J \in P_k(\{1,\ldots,n\})} P\left(\bigcap_{j\in J} A_j\right) = C_n^k \frac{(n-k)!}{n!} = \frac{1}{k!}$.

Finalement $\mathbf{P}(\bar{A}) = \sum_{k=1}^{n}(-1)^{k-1} \frac{1}{k!}$.

4°) On a $\mathbf{P}(A) = 1 - \mathbf{P}(\bar{A}) = 1 + \sum_{k=1}^{n}(-1)^{k} \frac{1}{k!} = \sum_{k=0}^{n}(-1)^{k} \frac{1}{k!}$ et sa limite quand n tend vers l'infini est donc $\sum_{k=0}^{\infty}(-1)^{k} \frac{1}{k!} = \frac{1}{e}$.

CHAPITRE 2 : VARIABLES ALEATOIRES

Variables aléatoires discrètes

Exercice 15 : Échauffement

Pour une loi géométrique de paramètre $p = 0.5$, calculer $P(X^3 = 8)$, $P(X^3 = 7)$ et $P((X - 4)^2 = 4)$ avec $p = 0.5$.

Solution :

$P(X^3 = 8) = P(X = 2) = 0.5 \times 0.5 = 0.25$

$P(X^3 = 7) = P(X = \sqrt[3]{7}) = 0$

$$\begin{aligned} P((X-4)^2 = 4) &= P(X = 2 \; ou \; X = 6) = P(X = 2) + P(X = 6) \\ &= 0.5^2 + 0.5 \times 0.5^5 \approx 0.266 \end{aligned}$$

Exercice 16 : Loi discrète uniforme

Soit X une variable aléatoire réelle discrète uniforme sur $[\![1000, 2000]\!]$.

Calculer $E[X]$ et $\sigma[X]$.

Remarque : $\sum_{k=0}^{n} k^2$est donné par $\frac{n(n+1)(2n+1)}{6}$.

Solution : On a $X(\Omega) = [\![1000, 2000]\!]$, et donc $card\big(X(\Omega)\big) = 2000 - 1000 + 1 = 1001.$

Comme X est uniforme, $\forall k \in \Omega, P(X = k) = \frac{1}{1001}$

On a alors :

$$E(X) = \frac{1}{1001}\sum_{k=1000}^{2000} k = \frac{1}{1001}\sum_{k=0}^{1000}(k + 1000) = 1000 + \frac{1}{1001}\sum_{k=0}^{1000} k = 1500$$

$$\begin{aligned} Var(X) &= \frac{1}{1001}\sum_{k=1000}^{2000}(k - 1500)^2 = \frac{1}{1001}\sum_{k=-500}^{500} k^2 = \frac{2}{1001}\frac{500 \times 501 \times 1001}{6} \\ &= 83500 \end{aligned}$$

Puis $\sigma(X) = 10\sqrt{835} \approx 289{,}0$.

Exercice 17 : Quelques lois classiques

On considère les lois discrètes suivantes :

- Loi de Bernoulli de paramètre p (notée $\mathcal{B}(p)$)
- Loi uniforme sur $\{2, \cdots, 5\}$
- Loi binômiale de paramètres (n, p) (notée $\mathcal{B}(n, p)$)
- Loi géométrique de paramètres p (notée $G(p)$)
- Loi de Poisson de paramètre λ (notée $P(\lambda)$)

On notera à chaque fois X la variable aléatoire.

Pour toutes ces lois :

- Donner la distribution des probabilités.
- Donner la fonction de répartition $F(x)$.
- Calculer $E(X)$, l'espérance de X.
- Calculer $E(X^2)$, l'espérance de X^2.
- Calculer $Var(X)$, la variance de X.
- Calculer $\sigma(X)$, l'écart-type de X.

Remarque : On rappelle quelques résultats qui peuvent être utiles dans les calculs. Pour $|x| < 1$, nous avons :

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x} \qquad \sum_{k=0}^{\infty} k\, x^k = \frac{x}{(1-x)^2} \qquad \sum_{k=0}^{\infty} k^2\, x^k = \frac{x(1+x)}{(1-x)^3}$$

Solution : Nous utiliserons dans la suite les définitions suivantes

$$\begin{aligned} &F(x) = P(X \leq x) \\ &E(X) = \sum_x x\, P(X = x) \\ &E(g(X)) = \sum_x g\,(x) P(X = x) \\ &E(X^2) = \sum_x x^2\, P(X = x) \\ &Var(X) = E[(X - E(X))^2] \\ &\sigma(X) = \sqrt{Var(X)} \end{aligned}$$

ainsi que la propriété $Var(X) = E(X^2) - E(X)^2$.

Loi de Bernoulli : $X \sim \mathcal{B}(p)$

Une variable aléatoire (v.a.) X suit une loi de Bernoulli de paramètre p , notée $\mathcal{B}(p)$, si et seulement si sa distribution de probabilités est la suivante :

$$\begin{aligned} &P(X = 1) = p \\ &P(X = 0) = 1 - p \end{aligned}$$

La loi de Bernoulli correspond à une expérience aléatoire dont l'issue est soit un succès ($X = 1$) avec une probabilité p , soit un échec ($X = 0$) avec une probabilité $q = 1 - p$.

$$F(x) = \begin{cases} 0 & si\ x < 0 \\ 1 - p & si\ 0 \leq x < 1 \\ 1 & si\ x \geq 1 \end{cases}$$

$$E(X) = 1 \times p + 0 \times q = p$$
$$E(X^2) = 1^2 \times p + 0^2 \times q = p$$
$$Var(X) = E(X^2) - E(X)^2 = p - p^2 = p(1-p) = pq$$
$$\sigma(X) = \sqrt{pq}$$

Loi uniforme sur $\{2, \cdots, 5\}$

$$P(X = k) = \frac{1}{4} \; pour \; k \in \{2, \cdots, 5\}$$
$$E(X) = \frac{2+3+4+5}{4} = 7\,/\,2$$
$$E(X^2) = \frac{4+9+16+25}{4} = 27\,/\,2$$
$$Var(X) = \frac{5}{4}$$
$$\sigma(X) = \frac{\sqrt{5}}{2}$$
$$F(x) = \begin{cases} 0 & si\ x < 2 \\ 1/4 & si\ 2 \leq x < 3 \\ 1/2 & si\ 3 \leq x < 4 \\ 3/4 & 4 \leq x < 5 \\ 1 & si\ x \geq 5 \end{cases}$$

Loi binômiale : $Y \sim \mathcal{B}(n, p)$

Une v.a. Y suit une loi binômiale de paramètre n et p , notée $\mathcal{B}(n,p)$, si et seulement si sa distribution de probabilités est la suivante :

$$P(Y = k) = \begin{cases} C_n^k p^k (1-p)^{n-k} & pour\ k \in [\![0, n]\!] \\ 0 & sinon \end{cases}$$

Il n'y a pas d'expression simple de la fonction de répartition :

$$F(x) = \begin{cases} 0 & si\ x < 0 \\ \sum_{k=0}^{\lfloor x \rfloor} C_n^k p^k (1-p)^{n-k} & si\ 0 \leq x < n \\ 1 & si\ x \geq n \end{cases}$$

<u>Calcul de l'espérance</u>

E $[X]$ = $\sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k}$ = $\sum_{k=1}^{n} k \frac{n!}{k!\,(n-k)!} p^k (1-p)^{n-k}$

= $\sum_{k=1}^{n} \frac{n!}{(k-1)!\,(n-k)!} p^k (1-p)^{n-k}$ = $np \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!\,[(n-1)-(k-1)]!} p^{k-1} (1-p)^{(n-1)-(k-1)}$

= $np \sum_{k=1}^{n} \binom{n-1}{k-1} p^{k-1} (1-p)^{(n-1)-(k-1)}$

= $np \sum_{k=0}^{n-1} \binom{n-1}{k} p^k (1-p)^{(n-1)-k}$ = $np\ [\ p + (1-p)\ ]^{n-1} = np$.

<u>Calcul de la variance</u>

Utilisons la formule de Koenig : V [X] = E [X^2] – (E [X^2])^2.

E [X^2] = $\sum_{k=0}^{n} k^2 \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=0}^{n} [\, k(k-1) + k \,] \binom{n}{k} p^k (1-p)^{n-k}$

$= \sum_{k=2}^{n} k(k-1) \binom{n}{k} p^k (1-p)^{n-k}$ + E [X].

Calculons séparément $\sum_{k=2}^{n} k(k-1) \binom{n}{k} p^k (1-p)^{n-k}$:

$\sum_{k=2}^{n} k(k-1) \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=2}^{n} k\,(k-1) \frac{n!}{k!\,(n-k)!} p^k (1-p)^{n-k}$

$= \sum_{k=2}^{n} \frac{n!}{(k-2)!\,(n-k)!} p^k (1-p)^{n-k}$

$= n(n-1)\, p^2 \sum_{k=2}^{n} \frac{(n-2)!}{(k-2)!\,[(n-2)-(k-2)]!} p^{k-2} (1-p)^{(n-2)-(k-2)}$

$= n(n-1)\, p^2 \sum_{k=2}^{n} \binom{n-2}{k-2} p^{k-2} (1-p)^{(n-2)-(k-2)}$

$= n(n-1)\, p^2 \sum_{k=2}^{n-1} \binom{n-2}{k} p^k (1-p)^{(n-2)-k} = n(n-1)\, p^2\, [\, p + (1-p)\,]^{n-2} = n(n-1)\, p^2.$

On a donc E [X^2] = $n(n-1)\, p^2$ + E [X] = $n(n-1)\, p^2 + np$.

Finalement :

V [X] = E [X^2] – E [X]^2 = $n(n-1)\, p^2 + np - n^2p^2 = np\,\big((n-1)p + 1 - np\big) = np\,(1-p) = npq$.

Correction alternative :

X représente le nombre de succès au cours de n expériences de Bernoulli indépendantes, chacune ayant une probabilité de succès p . Ainsi, si $X_1, \cdots, X_n$ sont n v.a. indépendantes de Bernoulli $\mathcal{B}(p)$, alors $X = \sum_{i=1}^{n} X_i$ suit une loi binômiale $\mathcal{B}(n,p)$. En utilisant la linéarité de l'espérance (cf fin du chapitre 2), nous obtenons

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

En utilisant la propriété que la variance de la somme est égale à la somme des variances pour des variables aléatoires indépendantes (cf fin chapitre 2), il vient :

$$Var(X) = Var[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} V\,ar(X_i) = npq$$

$$\sigma(X) = \sqrt{npq}$$

Loi géométrique : $X \sim G(p)$

Une v.a. X est distribuée suivant une loi géométrique de paramètre p ($q = 1 - p$), notée $G(p)$, si et seulement si, pour $k \in \mathbb{N}^*$:

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

X représente le nombre d'épreuves indépendantes de Bernoulli de paramètre p nécessaires avant d'obtenir un succès.

Pour une variable aléatoire entière, on se contente souvent de donner la fonction de répartition pour les valeurs où la probabilité est non nulle. Pour $x \in \mathbb{N}$:

$$F(x) = \sum_{k=1}^{x} q^{k-1}\, p = \frac{1-q^x}{1-q} p = 1 - q^x$$

$$\begin{aligned} E(X) \quad &= \sum_{k=1}^{\infty} k\, P(X=k) = \sum_{k=1}^{\infty} k\, q^{k-1} p \\ &= \frac{p}{q} \sum_{k=1}^{\infty} k\, q^k = \frac{p}{q}\frac{q}{p^2} = \frac{1}{p} \end{aligned}$$

$$\begin{aligned} E(X^2) \quad &= \sum_{k=1}^{\infty} k^2\, P(X=k) = \sum_{k=1}^{\infty} k^2\, q^{k-1} p \\ &= \frac{p}{q} \sum_{k=1}^{\infty} k^2\, q^k = \frac{\frac{p}{q} \times q(1+q)}{p^3} = \frac{1+q}{p^2} \end{aligned}$$

$$Var(X) = E(X^2) - E(X)^2 = \frac{1+q}{p^2} - \frac{1}{p^2} = \frac{q}{p^2}$$

$$\sigma(X) = \sqrt{Var(X)} = \frac{\sqrt{q}}{p}$$

Loi de Poisson : $\boldsymbol{X \sim \mathcal{P}(\lambda)}$

Une v.a. X est distribuée suivant une loi de Poisson de paramètre λ, notée $\mathcal{P}(\lambda)$, si et seulement si pour $k \in \mathbb{N}$:

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

Il n'y a pas d'expression simple de la fonction de répartition. Pour $x \in IN$:

$$F(x) = \frac{\sum_{k=0}^{x} \lambda^k e^{-\lambda}}{k!}$$

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

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

Puis :

$$
Var(X) = E(X^2) - E(X)^2 = \lambda
$$
$$
\sigma(X) = \sqrt{\lambda}
$$

Exercice 18 : Urne

Une urne contient 9 boules blanches et une boule rouge.

1°) On répète dix fois de suite la procédure suivante : prélever une boule au hasard, noter sa couleur, remettre la boule dans l'urne et mélanger.

a) Donner la loi de la variable aléatoire X qui retourne le nombre de tirages où la boule rouge a été tirée.

Solution : X est distribuée suivant une loi binomiale de paramètre $n = 10$ et $p = 1/10$. On a donc :

1. $X(\Omega) = [\![0, 10]\!]$;
2. $\forall\, k \in X(\Omega)$, $P(X = k) = C_{10}^{k}\, (0{,}1)^k\, (0{,}9)^{10-k}$.

b) Calculer $P(X = 2)$.

Solution : $P(X = 2) = C_{10}^{2}\, (0{,}1)^2\, (0{,}9)^8 \approx 45 \times 10^{-2} \times 0{,}430 \approx 19{,}37\ \%$

2°) On répète la procédure décrite ci-dessus jusqu'à ce que l'on obtienne la boule rouge.

a) Donner la loi de la variable aléatoire Y qui retourne le nombre de tirages nécessaires.

Solution : Y est distribuée suivant une loi géométrique de paramètre $p = 1/10$. On a donc :

$Y(\Omega) = \mathbb{N}^*$ et $\forall\, k \in \mathbb{N}^*$, $\mathbf{P}\,(Y = k) = \left(\frac{9}{10}\right)^{k-1} \times \frac{1}{10} = \frac{9^{k-1}}{10^k}$.

b) Calculer $P(Y > E[Y])$.

Solution : D'après le cours, on a $E[Y] = \frac{1}{p} = 10$.

Alors $P(Y > 10) = \sum_{k=11}^{\infty} P(Y = k) = \frac{1}{10} \sum_{k=11}^{\infty} \left(\frac{9}{10}\right)^{k-1} = \frac{1}{10} \times \left(\frac{9}{10}\right)^{10} \sum_{k=0}^{\infty} \left(\frac{9}{10}\right)^{k}$

En faisant le changement de variable $i = k - 11$, il vient

$$P(Y > 10) = \frac{1}{10}\sum_{i=0}^{\infty}\left(\frac{9}{10}\right)^{i+10} = \frac{1}{10}\left(\frac{9}{10}\right)^{10}\sum_{i=0}^{\infty}\left(\frac{9}{10}\right)^{i} = \frac{1}{10}\left(\frac{9}{10}\right)^{10}\frac{1}{1-\frac{9}{10}} = \left(\frac{9}{10}\right)^{10}$$

$\approx$ 34,87 %.

Exercice 19 : Approximations successives

On prélève au hasard (et sans remise) 15 pièces dans un lot de 500 pièces contenant 10 pièces défectueuses. Soit X la variable aléatoire comptant le nombre de pièces défectueuses trouvées.

1°) a) Préciser la loi de la variable aléatoire X ?

Solution : X suit une loi hypergéométrique de paramètres $N = 500, n = 15$ et $p = \frac{10}{500} = 0{,}02$.

b) Calculer alors une valeur approchée de $P(X = 0), P(X = 1), P(X = 2), P(X \geq 3)$ à 10^{-3} près.

Solution : Posons $N_1 = Np = 10$ (le nombre de pièces défectueuses) et $N_2 = N - N_1 = 490$ (le nombre de pièces correctes).

On a $X(\Omega) = [\![\max(0, n - N_2), \min(n, N_1)]\!] = [\![0, 10]\!]$ = et

$\forall\, k \in [\![0, 10]\!]$, $\mathbf{P}(X = k) = \frac{C_{N_1}^{k} C_{N_2}^{n-k}}{C_N^n} = \frac{C_{10}^{k} C_{490}^{15-k}}{C_{500}^{15}}$

Donc $\mathbf{P}(X = 0) = \frac{C_{490}^{15}}{C_{500}^{15}} = \frac{\frac{490\times\cdots\times476}{15!}}{\frac{500\times\cdots\times486}{15!}} = \frac{490\times\cdots\times476}{500\times\cdots\times486} \approx 0{,}735$,

$\mathbf{P}(X = 1) = \frac{10\times C_{490}^{14}}{C_{500}^{15}} = \frac{10\times\frac{490\times\cdots\times477}{14!}}{\frac{500\times\cdots\times486}{15!}} = \frac{490\times\cdots\times477\times15\times10}{500\times\cdots\times486} \approx 0{,}232$,

$\mathbf{P}(X = 2) = \frac{C_{10}^{2}\times C_{490}^{13}}{C_{500}^{15}} = \frac{45\times\frac{490\times\cdots\times478}{13!}}{\frac{500\times\cdots\times486}{15!}} = \frac{490\times\cdots\times478\times15\times14\times45}{500\times\cdots\times486} \approx 0{,}031$ et

$\mathbf{P}(X \geq 3) = 1 - \mathbf{P}(X = 0) - \mathbf{P}(X = 1) - \mathbf{P}(X = 2) \approx 0{,}002$.

2°) a) Par quelle loi Y la loi de X peut-elle être approchée, et pourquoi ?

Solution : X peut être approchée par une loi binomiale de paramètres $n = 15$ et $p = 0{,}02$, car le rapport $\frac{n}{N}$ est suffisamment petit pour que l'approximation asymptotique ($N \to \infty$) soit acceptable. Plus concrètement, cela revient à considérer que comme $n = 15$ est petit devant $N = 500$, c'est approximativement comme si l'on faisait des tirages avec remise, d'où la loi binomiale.

b) Reprendre alors, à l'aide de cette loi, les calculs de la question 1°) b).

Solution : On a donc $Y(\Omega) = [\![0, 15]\!]$ = et

$\forall\, k \in [\![0, 15]\!]$, $\mathbf{P}(X = k) = C_n^k\, p^k (1-p)^{n-k} = C_{15}^{k}\, (0{,}02)^k (0{,}98)^{n-k}$.

Donc $\mathbf{P}(X = 0) = (0{,}98)^{15} \approx 0{,}739$, $\mathbf{P}(X = 1) = 15 \times 0{,}02 \times (0{,}98)^{14} \approx 0{,}226$,

$\mathbf{P}(X = 2) = 105 \times (0{,}02)^2 \times (0{,}98)^{13} \approx 0{,}032$ et

$\mathbf{P}(X \geq 3) = 1 - \mathbf{P}(X = 0) - \mathbf{P}(X = 1) - \mathbf{P}(X = 2) \approx 0{,}003$.

3°) a) Par quelle loi Z la loi de Y peut-elle être approchée, et pourquoi ?

Solution :

Y peut être approchée par une loi de Poisson de paramètre $\lambda = n \times p = 0{,}3$, car n peut être considéré comme grand et p comme petit.

b) Reprendre alors, à l'aide de cette loi, les calculs de la question 1°) b).

Solution :

On a donc $Z(\Omega) = \mathbb{N}$ et $\forall\, k \in \mathbb{N}$, $\mathbf{P}(X = k) = e^{-\lambda}\,\frac{\lambda^k}{k!}$.

Donc $\mathbf{P}(X = 0) = e^{-0{,}3} \approx 0{,}741$, $\mathbf{P}(X = 1) = 0{,}3 \times e^{-0{,}3} \approx 0{,}222$,

$\mathbf{P}(X = 2) = \frac{0{,}3^2}{2} \times e^{-0{,}3} \approx 0{,}033$ et $\mathbf{P}(X \geq 3) = 1 - \mathbf{P}(X = 0) - \mathbf{P}(X = 1) - \mathbf{P}(X = 2) \approx 0{,}004$.

Exercice 20 : Déduire la loi à partir de la fonction de répartition

Une urne contient N boules numérotées de 1 à N. On en tire n au hasard ($n < N$) sans remise. Soit X le plus grand numéro tiré.

a) Donner la fonction de répartition de X.

Solution :

On a $X(\Omega) = [\![n , N]\!]$ et $\forall\, k \in [\![n , N]\!]$,

$\mathbf{P}(X \leq k) = \frac{C_k^n}{C_N^n} = \frac{k!}{n!(k-n)!} \times \frac{n!(N-n)!}{N!} = \frac{k!}{(k-n)!} \times \frac{(N-n)!}{N!} = \frac{k(k-1)\ldots(k-n+1)}{N(N-1)\ldots(N-n+1)}$.

b) En déduire la loi de X.

Solution :

On a $X(\Omega) = [\![n , N]\!]$ et donc $\mathbf{P}(X = n) = \mathbf{P}(X \leq n) = \frac{C_n^n}{C_N^n} = \frac{1}{C_N^n}$ et

$\forall\, k \in [\![n+1 , N]\!]$, $\mathbf{P}(X = k) = \mathbf{P}(X \leq k) - \mathbf{P}(X \leq k-1) = \frac{C_k^n - C_{k-1}^n}{C_N^n}$.

c) Application numérique à $N = 5$ et $n = 2$.

Solution :

On a $X(\Omega) = [\![2 , 5]\!]$ et $\mathbf{P}(X = 2) = \frac{1}{C_5^2} = \frac{1}{10} = 0{,}1$, $\mathbf{P}(X \leq 3) = \frac{C_3^2}{C_5^2} = \frac{3}{10} = 0{,}3$,

$\mathbf{P}(X \leq 4) = \frac{C_4^2}{C_5^2} = \frac{6}{10} = 0{,}6$ et $\mathbf{P}(X \leq 5) = 1$, d'où $\mathbf{P}(X = 3) = 0{,}3 - 0{,}1 = 0{,}2$,

$\mathbf{P}(X = 4) = 0{,}6 - 0{,}3 = 0{,}3$ et $\mathbf{P}(X = 5) = 1 - 0{,}6 = 0{,}4$.

Exercice 21 : Les contrôles de parité

On note τ_b le *taux d'erreur* par bit d'un canal de transmission, c'est-à-dire la probabilité pour que, si l'on transmet un bit sur celui-ci, le bit reçu soit inversé. On considère que τ_b est très petit devant 1.

Le contrôle de parité par bit de parité est un des systèmes de contrôle les plus simples. Il consiste à ajouter un bit supplémentaire à un mot de code de 7 bits (pour former un octet avec le bit de parité) dont la valeur (0 ou 1) est telle que le nombre total de bits à 1 soit pair, par exemple. On parle, dans ce cas, de bit de parité paire. Pour être plus explicite le contrôle de parité avec un bit de parité paire consiste à ajouter un 1 si le nombre de bits du mot de code est impair, et un 0 dans le cas contraire.

A la réception de l'octet transmis, le nombre de 1 dans celui-ci est contrôlé, et si ce nombre est impair, une demande de retransmission est retournée à l'émetteur.

On fait l'hypothèse (très forte) que les erreurs de transmission sur chaque bit se produisent de manière indépendante entre elles.

1°) Quelle est, dans ce cas, la loi de la variable aléatoire N qui donne le nombre de bits mal transmis dans un octet ?

Solution : Il s'agit d'une loi binomiale de paramètres $n = 8$ et $p = \tau_b$.

2°) Quelle est, en fonction de τ_b, la probabilité pour qu'un octet soit transmis correctement (donner la formule exacte, puis une approximation du premier ordre) ?

Solution : $\mathbf{P}\ (N = 0) = (1-\tau_b)^8 \approx 1 - 8\tau_b$.

3°) Quelle est, sachant qu'un octet est mal transmis, la probabilité τ_0 pour que l'erreur ne soit pas détectée par le contrôle de parité (donner la formule exacte, puis une approximation du premier ordre) ?

Solution :

$\tau_o = \mathbf{P}\ (N \in \{2, 4, 6, 8\} \mid N > 0) = \frac{P\ (N \in \{2,4,6,8\})}{P\ (N>0)}$, et

$\mathbf{P}\ (N \in \{2, 4, 6, 8\}) = \sum_{k=1}^{4} P(N = 2k) = \sum_{k=1}^{4} C_8^{2k} \tau_{b^{2k}} (1-\tau_b)^{8-2k}$

$= 28\ \tau_b^2\ (1-\tau_b)^6 + 70\ \tau_b^4 (1-\tau_b)^4 + 28\ \tau_b^6 (1-\tau_b)^2 + \tau_b^8 \approx 28\ \tau_b^2\ (1-\tau_b)^6 \approx 28\ \tau_b^2$.

$\mathbf{P}\ (N > 0) = 1 - \mathbf{P}\ (N = 0) = 1 - (1-\tau_b)^8 \approx 1 - (1 - 8\tau_b) = 8\tau_b$.

Finalement $\tau_o \approx \frac{28\ \tau_{b^2}}{8\ \tau_b} = 3{,}5\ \tau_b$.

4°) Quelle est la loi de la variable aléatoire T qui donne le nombre de transmissions nécessaires pour qu'un octet soit accepté (donner la loi exacte, puis des lois approximatives au premier et au deuxième ordre) ?

Solution : T est une loi géométrique de paramètre p, avec :

$p = \mathbf{P}\,(N \in \{0, 2, 4, 6, 8\}) = \sum_{k=0}^{4} P(N = 2k) = \sum_{k=0}^{4} C_8^{2k} {\tau_b}^{2k} (1-\tau_b)^{8-2k}$

$= (1-\tau_b)^8 + 28\,\tau_b^2\,(1-\tau_b)^6 + 70\,\tau_b^4(1-\tau_b)^4 + 28\,\tau_b^6(1-\tau_b)^2 + \tau_b^8 \approx 1 - 8\tau_b$.

Donc $T(\Omega) = I\!N^*$ et $\mathbf{P}\,(T = k) = p\,(1-p)^{k-1}$.

<u>Au premier ordre</u> : on a $p \approx 1 - 8\tau_b$, d'où $\mathbf{P}\,(T = k) = p\,(1-p)^{k-1} \approx (1 - 8\tau_b)\,(8\tau_b)^{k-1}$.

Alors, on a approximativement $T(\Omega) = \{1, 2\}$ car $\mathbf{P}\,(T = 1) \approx 1 - 8\tau_b$ et $\mathbf{P}\,(T = 2) \approx (1 - 8\tau_b)\,(8\tau_b) \approx 8\tau_b$.

<u>Au deuxième ordre</u> : on a $p \approx 1 - 8\tau_b + 28\,\tau_b^2 + 28\,\tau_b^2 = 1 - 8\tau_b + 56\,\tau_b^2$, d'où :

$\mathbf{P}\,(T = k) = p\,(1-p)^{k-1} \approx (1 - 8\tau_b + 56\,\tau_b^2)\,(8\tau_b - 56\,\tau_b^2)^{k-1}$.

Alors, on a approximativement $T(\Omega) = \{1, 2, 3\}$ car $\mathbf{P}\,(T = 1) \approx 1 - 8\tau_b + 56\,\tau_b^2$ et $\mathbf{P}\,(T = 2) \approx (1 - 8\tau_b + 56\,\tau_b^2)\,(8\tau_b - 56\,\tau_b^2) \approx 8\tau_b - 56\,\tau_b^2 - 64\,\tau_b^2 = 8\tau_b - 120\,\tau_b^2$,

et $\mathbf{P}\,(T = 3) \approx 64\,\tau_b^2$.

5°) Quelle est la probabilité pour que la transmission finalement acceptée par le récepteur soit correcte (donner la formule exacte, puis une approximation au premier ordre et au deuxième ordre) ?

Solution :

Notons C l'évènement dont on recherche la probabilité.

D'après le théorème des probabilités totales :

$\mathbf{P}\,(C) = \mathbf{P}\,(C \mid N = 0) \times \mathbf{P}\,(N = 0) + \mathbf{P}\,(C \mid N \in \{1, 3, 5, 7\}) \times \mathbf{P}\,(N \in \{1, 3, 5, 7\})$

$\quad + \mathbf{P}\,(C \mid N \in \{2, 4, 6, 8\}) \times \mathbf{P}\,(N \in \{2, 4, 6, 8\})$

$\quad = \mathbf{P}\,(N = 0) + \mathbf{P}\,(C) \times \mathbf{P}\,(N \in \{1, 3, 5, 7\}) + 0 \times \mathbf{P}\,(N \in \{2, 4, 6, 8\})$.

Explication : $\mathbf{P}\,(C \mid N = 0) = 1$ est évident.

$\mathbf{P}\,(C \mid N \in \{1, 3, 5, 7\}) = \mathbf{P}\,(C)$ parce que les retransmissions sont indépendantes entre elles (on repart à zéro).

$\mathbf{P}\,(C \mid N \in \{2, 4, 6, 8\}) = 0$ parce que, dans ce cas, l'octet mal transmis n'est pas détecté par le contrôle de parité.

On obtient donc $\mathbf{P}\,(C) = \frac{P\,(N=0)}{1 - P\,(N \in \{1,3,5,7\})} = \frac{P\,(N=0)}{P\,(N \in \{0,2,4,6,8\})} = \frac{P\,(N=0)}{p}$.

avec $\mathbf{P}\,(N = 0) = (1-\tau_b)^8 \approx 1 - 8\tau_b$,

et $p = (1-\tau_b)^8 + 28\,\tau_b^2\,(1-\tau_b)^6 + 70\,\tau_b^4(1-\tau_b)^4 + 28\,\tau_b^6(1-\tau_b)^2 + \tau_b^8$.

Au premier ordre : $\mathbf{P}\,(N = 0) = (1-\tau_b)^8 \approx 1 - 8\tau_b$ et $p \approx 1 - 8\tau_b$, d'où $\mathbf{P}\,(C) \approx 1$, ce qui est une approximation guère satisfaisante.

Au deuxième ordre : $\mathbf{P}\,(N = 0) = (1-\tau_b)^8 \approx 1 - 8\tau_b + 28\,\tau_b^2$

et $p \approx 1 - 8\tau_b + 28\,\tau_b^2 + 28\,\tau_b^2 = 1 - 8\tau_b + 56\,\tau_b^2$, d'où $\mathbf{P}(C) \approx \frac{1-8\tau_b+28\,\tau_{b2}}{1-8\tau_b+56\,\tau_{b2}} \approx (1 - 8\tau_b + 28\,\tau_b^2) \times (1 + 8\tau_b - 56\,\tau_b^2 + 32\,\tau_b^2)$

$= (1 - 8\tau_b + 28\,\tau_b^2) \times (1 + 8\tau_b - 24\,\tau_b^2) \approx 1 + 8\tau_b - 24\,\tau_b^2 - 8\tau_b - 64\tau_b^2 + 28\tau_b^2 = 1 - 60\,\tau_b^2$.

Exercice 22 : Allumettes de Banach

Tout en faisant avancer l'analyse fonctionnelle, Stefan Banach fumait la pipe. Comme il devait souvent la ré-allumer et qu'il ne souhaitait pas perdre de temps, il mettait dans chaque poche une boîte d'allumette comportant au départ N allumettes chacune, et prélevait de temps en temps une allumette dans l'une ou l'autre poche de manière équiprobable. Il arrivait forcément un moment où il retirait la dernière allumette de l'une des deux boîtes …

L'objectif de ce problème est de déterminer, pour $k \in [\![1, N]\!]$, la probabilité pour que quand une des deux boîtes est vide, il reste k allumettes dans la seconde. Pour cela, il faut déterminer la loi X qui donne le nombre d'allumettes dans la boîte restante. Ce nombre est naturellement compris entre 1 et N et on a donc $X(\Omega) = [\![1, N]\!]$.

1°) Montrer que, du fait de la symétrie du problème, on peut considérer sans perte de généralité que c'est la boîte dans la poche droite, par exemple, qui est vidée la première et que X compte donc le nombre d'allumettes restant dans la boîte de la poche gauche.

Solution :

On peut utiliser la symétrie du problème : si X est la variable qui compte le nombre d'allumettes restantes, quelle que soit la boîte qui a été vidée, on peut écrire, en utilisant le théorème des probabilités totales :

$\mathbf{P}(X = k) = \mathbf{P}(X = k \mid DV) \times \mathbf{P}(DV) + \mathbf{P}(X = k \mid GV) \times \mathbf{P}(GV)$

où DV et GV désignent les événements « la boîte de droite (respectivement, de gauche) est vide. » La symétrie du problème permet de considérer que $\mathbf{P}(X = k \mid DV) = \mathbf{P}(X = k \mid GV)$ et que $\mathbf{P}(DV) = \mathbf{P}(GV) = \frac{1}{2}$. On a donc, par exemple, $\mathbf{P}(X = k) = \mathbf{P}(X = k \mid DV)$, et on peut alors se placer dans le cas où c'est la boîte de la poche de droite que l'on vide.

2°) Montrer que $P(X = k) = C_{2N-k-1}^{N-1} \left(\frac{1}{2}\right)^{2N-k-1}$

Indication : Pour qu'il reste k allumettes dans la boîte de gauche quand on retire la dernière allumette de la boîte de droite, il faut avoir auparavant retiré $N - k$ allumettes de la boîte de gauche et $N - 1$ allumettes de la boîte de droite.

Solution :

$\left(\frac{1}{2}\right)^{2N-k-1}$ est la probabilité d'une répartition particulière des $2N - k - 1$ tirages (probabilité de 1/2 de tirer dans la poche de gauche et probabilité de 1/2 de tirer dans la poche de droite).

L'exposant est bien $2N - k - 1$ et non $2N - k$, car la dernière allumette est forcément tirée dans la poche de droite car on calcule une probabilité conditionnelle sachant DV.

C_{2N-k-1}^{N-1} compte le nombre de manières de répartir les $N - 1$ prises d'allumettes dans la poche de droite parmi les $(N - 1) + (N - k) = 2N - k - 1$ essais qui précèdent le dernier.

Remarque :

En considérant l'expérience de Bernoulli où prendre une allumette dans la poche de droite est un succès et prendre une allumette dans la poche de gauche un échec, Alors, pour qu'il reste k allumettes dans la boîte de gauche quand la boîte de droite se vide, il faut subir $N - k$ échecs avant d'atteindre la $N^{\text{ième}}$ réussite.

La formule obtenue peut donc faire penser à celle d'une loi binomiale négative (nombre d'échecs pour atteindre la $r^{\text{ième}}$ réussite), $\mathbf{P}\,(N_{\text{échecs}} = k\,) = C_{k+r-1}^{r-1}\, p^r q^k$, appliquée à $N - k$, avec $p = q = 1/2$ et $r = N$, mais on obtiendrait $\left(\frac{1}{2}\right)^{2N-k}$ au lieu de $\left(\frac{1}{2}\right)^{2N-k-1}$, ce qui donnerait, pour $N = 1$, une probabilité $\mathbf{P}\,(\,X = 1\,) = ½$, ce qui est faux. Le problème vient du fait que l'on ne sait pas quelle est la poche qui sera vidée en premier, donc on ne sait pas définir au préalable ce qui est un échec et ce qui est une réussite. On est donc obligé d'appliquer le théorème des probabilités totales pour envisager ces deux possibilités, et le calcul des probabilités conditionnelles ne considère qu'il n'y a que $2N - k - 1$ tirages qui sont aléatoires, le dernier étant déterminé par la condition (sa probabilité conditionnelle étant donc égale à 1).

Variables aléatoires continues

Exercice 23 : Loi uniforme

Soit X une variable aléatoire suivant une loi uniforme sur $[a, b]$. On notera f sa densité et X sa fonction de répartition.

1°) Donner $X(\Omega)$ et la densité f de X.

2°) Vérifier que $P(X \in \mathbb{R}) = 1$.

3°) Rappeler ce que signifie la fonction de répartition en terme de probabilité. Calculer la fonction de répartition de X.

4°) Représenter graphiquement la densité et la fonction de répartition.

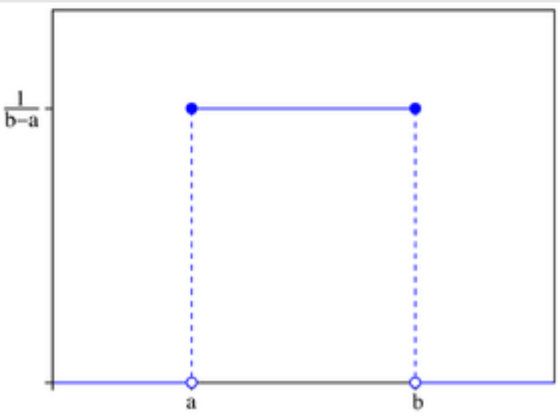
5°) Calculer l'espérance et la variance de X.

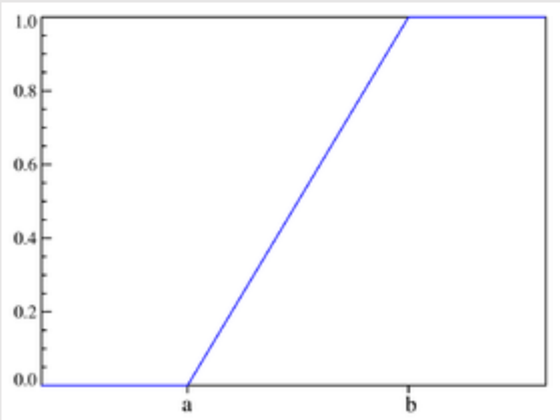
1°) $X(\Omega) = [a, b]$. La densité f vaut $\frac{1}{b-a}$ pour $x \in [a, b]$ et 0 ailleurs.

2°) $P(X \in \mathbb{R}) = \int_{-\infty}^{+\infty} f(x)dx = \int_a^b \frac{1}{b-a} dx = 1$

3°) $$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt = \begin{cases} \int_{-\infty}^{x} 0dt = 0 & si\ x < a \\ \int_{-\infty}^{a} 0dt + \int_a^x \frac{1}{b-a} dt = \frac{x-a}{b-a} & si\ a \leq x < b \\ \int_{-\infty}^{a} 0dt + \int_a^b \frac{1}{b-a} dt + \int_b^x 0dt = 1 & si\ x \geq b \end{cases}$$

4°) A gauche la densité, à droite la fonction de répartition (source des images : Article *Loi uniforme continue* de Wikipédia en français) :





5°) $E(X) = \int_{-\infty}^{+\infty} xf(x)dt = \int_a^b \frac{x}{b-a} dt = \left[\frac{x^2}{2(b-a)}\right]_a^b = \frac{b^2-a^2}{2(b-a)} = \frac{(b-a)(b+a)}{2(b-a)} = \frac{a+b}{2}$

$$E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x)dt = \int_a^b \frac{x^2}{b-a} dt = \left[\frac{x^3}{3(b-a)}\right]_a^b = \frac{b^3-a^3}{3(b-a)}$$

$$= \frac{(b-a)(b^2+ab+a^2)}{3(b-a)} = \frac{b^2+ab+a^2}{3}$$

Puis

$$Var(X) = E(X^2) - E(X)^2 = \frac{b^2 + ab + a^2}{3} - \frac{(a+b)^2}{4} = \frac{4(b^2 + ab + a^2) - 3(a+b)^2}{12}$$

$$= \frac{b^2 + a^2 - 2ab}{12} = \frac{(b-a)^2}{12}$$

Exercice 24 : Densité de probabilité et fonction de répartition

Soit X une variable aléatoire réelle continue définie sur [0,10], de densité de probabilité f définie par $f(x) = C(x+3)$ où C une constante réelle.

1°) Calculer la constante C.

2°) Calculer la fonction de répartition de X.

3°) Calculer $P(5 < X < 6)$.

4°) Calculer l'espérance et la variance de X.

Solution :

1°) On a $\int_0^{10} C(x+3)dx = C\left[\frac{x^2}{2} + 3x\right]_0^{10} = 80C = 1$, d'où $C = \frac{1}{80}$.

2°) Pour $x \leq 0$, $F_X(x) = 0$ et pour $x \geq 10$, $F_X(x) = 1$.

Pour $x \in [0,10]$, $F_{X(x)} = \int_0^x C(t+3)dt \frac{1}{80}\left[\frac{t^2}{2} + 3t\right]_0^x = \frac{\frac{x^2}{2}+3x}{80} = \frac{x^2+6x}{160}$.

3°) Comme F_X est continue, on peut écrire $P(5 < X < 6) = P(5 < X \leq 6) = F_X(6) - F_X(5)$ $= \frac{72-55}{160} = \frac{17}{160}$.

4°) $E(X) = \frac{1}{80}\int_0^{10}(x^2 + 3x)dx = \frac{1}{80}\left[\frac{x^3}{3} + 3\frac{x^2}{2}\right]_0^{10} = \frac{1}{80} \times \frac{1450}{3} = \frac{145}{24} \approx 6{,}04$.

$E(X^2) = \frac{1}{80}\int_0^{10}(x^3 + 3x^2)dx = \frac{1}{80}\left[\frac{x^4}{4} + 3\frac{x^3}{3}\right]_0^{10} = \frac{3500}{80} = \frac{175}{4}$.

$V(X) = E(X^2) - (E(X))^2 = \frac{175}{4} - \frac{21025}{576} = \frac{25200-21025}{576} = \frac{4175}{576} \approx 7{,}25$.

Exercice 25 : Loi normale 1

Le kilométrage que des pneus d'une marque et d'un type donné peuvent parcourir avant d'atteindre leur limite d'usure est distribué suivant une loi normale d'espérance 40 000 km et d'écart-type 6 400 km.

Quelle est la probabilité pour que :

1°) un pneu parcourt plus de 40 000 km ?

2°) un pneu parcourt entre 30 000 et 35 000 km ?

3°) un pneu parcourt encore au moins 10 000 km supplémentaires, étant donné qu'il a déjà parcouru 30 000 km ?

Solution :

Soit K la variable aléatoire représentant le kilométrage que les pneus peuvent parcourir avant d'atteindre leur limite d'usure. On note U la loi normale centrée-réduite.

Les probabilités demandées sont :

1°) $\mathbf{P}\,(K > 40\,000) = P\left(\frac{K-40000}{6400} > 0\right) = 0{,}5$;

2°) $\mathbf{P}\,(30\,000 \leq K \leq 35\,000) = P\left(\frac{30000-40000}{6400} \leq \frac{K-40000}{6400} \leq \frac{35000-40000}{6400}\right)$

$= \mathbf{P}\,(-1{,}5625 \leq U \leq -0{,}78125) = \mathbf{P}\,(0{,}78125 \leq U \leq 1{,}5625) \approx 0{,}9409 - 0{,}7827 = 15{,}82\ \%$

3°) $\mathbf{P}\,(K > 40\,000 \mid K \geq 30\,000) = P\left(\frac{K-40000}{6400} > 0 \mid \frac{K-40000}{6400} > \frac{30000-40000}{6400}\right)$

$= \mathbf{P}\,(U > 0 \mid U \geq -1{,}5625) = \frac{P(U>0\ \cap\ U\geq -1{,}5625)}{P(U\geq -1{,}5625)} = \frac{P(U>0)}{P(U\geq -1{,}5625)}.$

Or $\mathbf{P}\,(U \geq -1{,}5625) = \mathbf{P}\,(U \leq 1{,}5625) \approx 0{,}9409$,

d'où $\mathbf{P}\,(K > 40\,000 \mid K \geq 30\,000) \approx \frac{0{,}5}{0{,}9409} \approx 53{,}14\ \%$

Exercice 26 : Loi normale 2

Si votre temps de déplacement de votre domicile à l'école suit une loi normale d'espérance 40 minutes et d'écart-type 7 minutes.

1°) A quelle heure devez-vous quitter votre domicile pour avoir une probabilité d'au moins 90 % de ne pas être en retard en cours à 8 heures ?

2°) Si vous partez tous les matins à 7h 05, quelle est la probabilité pour que, sur une période de 22 semaines, vous soyez en retard plus d'une fois ? plus de deux fois ?

Solution :

1°) Soit T la variable aléatoire représentant le temps de déplacement de votre domicile à l'école.

Cherchons le temps de déplacement t_d tel que $\mathbf{P}\,(T \leq t_d) \geq 0{,}9$.

On a $\mathbf{P}\,(T \leq t_d) \geq 0{,}9 \Leftrightarrow P\left(\frac{T-40}{7} \leq \frac{t_d-40}{7}\right) \geq 0{,}9 \Leftrightarrow \Phi\left(\frac{t_d-40}{7}\right) \geq 0{,}9$

$\Leftrightarrow \frac{t_d-40}{7} \geq \Phi^{-1}(0{,}9) = u_{0,9}$, fractile d'ordre 0,9 de la loi gaussienne centrée-réduite, cette dernière inégalité étant justifiée par le fait que la fonction Φ est croissante.

La table des fractiles donne $\frac{t_d-40}{7} \geq 1{,}2816 \Rightarrow t_d \geq 49$ minutes.

Il faut donc partir avant 7h 11.

2°) On note U la loi normale centrée-réduite.

Calculons d'abord la probabilité d'arriver en retard en partant à 7h 05 :

$\mathbf{P}\,(T > 55) = P\left(\frac{T-40}{7} > \frac{55-40}{7}\right) = \mathbf{P}\,(U > 2{,}1429) = 1 - \mathbf{P}\,(U \leq 2{,}1429) \approx 1 - 0{,}9839$

$= 1{,}61$ %.

Le nombre N d'arrivées en retard au cours de la période est donc distribué suivant une loi binomiale de paramètres $n = 22 \times 5 = 110$ et $p = 0{,}0161$. On a donc :

$\mathbf{P}\,(N > 1) = 1 - \mathbf{P}\,(N \leq 1)$

$= 1 - 0{,}9839^{110} - (110 \times 0{,}0161 \times 0{,}9839^{109}) \approx 53$ %.

$\mathbf{P}\,(N > 2) = 1 - \mathbf{P}\,(N \leq 2)$

$= 1 - 0{,}9839^{110} - (110 \times 0{,}0161 \times 0{,}9839^{109}) - \left(\frac{110\times109}{2} \times 0{,}0161^2 \times 0{,}9839^{108}\right) \approx 26$ %.

Exercice 27 : Durée de fonctionnement d'un appareil électrique

Un appareil électrique fonctionne avec trois piles P_j, $j \in \{1, 2, 3\}$. La durée de vie de la pile P_j est une variable aléatoire notée T_j. On suppose que les variables aléatoires T_1, T_2 et T_3 sont indépendantes entre elles et qu'elles sont toutes distribuées suivant une même loi exponentielle de paramètre $\lambda \in \mathbb{R}_+^*$

1°) Soit D la variable aléatoire égale à la durée de fonctionnement de l'appareil électrique à plein régime, c'est-à-dire avec les trois piles en état de marche. Déterminer la loi de D.

Solution :

On a $D(\Omega) = \mathbb{R}_+$.

$D = \mathrm{Min}\,(T_1, T_2, T_3)$, donc $P\,(D > t) = P\,((T_1 > t) \cap (T_2 > t) \cap (T_3 > t))$

$= P\,(T_1 > t) \times P\,(T_2 > t) \times P\,(T_3 > t)$ (car T_1, T_2 et T_3 sont indépendantes)

Or $P\,(T_1 > t) = 1 - F_{T_1}(t) = e^{-\lambda t}$

D'où $P(D > t) = \left(e^{-\lambda t}\right)^3 = e^{-3\lambda t}$

Puis $P\,(D \leq t) = 1 - e^{-3\lambda t}$.

D est donc distribuée suivant une loi exponentielle de paramètre 3λ.

2°) Soit $t \in \mathbb{R}_+^*$. On note N_t la variable aléatoire donnant le nombre de piles en état de marche à l'instant t. Déterminer la loi de N_t.

Solution :

Chaque pile j fonctionne à l'instant t avec une probabilité $\mathbf{P}\,(T_j > t) = e^{-\lambda t}$. Comme les variables aléatoires T_1, T_2 et T_3 sont indépendantes entre elles, la variable aléatoire N_t donnant le nombre de piles en état de marche à l'instant t est donc distribuée suivant une loi binomiale de paramètres $n = 3$ et $p = e^{-\lambda t}$.

3°) L'appareil s'arrête de fonctionner dès que deux piles sont usées. On note T la variable aléatoire égale au temps de fonctionnement de l'appareil électrique. Déterminer la loi de T.

Solution :

On a $T(\Omega) = I\!R_+$.

$F_T(t) = \mathbf{P}(T \leq t) = \mathbf{P}(T < t) = \mathbf{P}(N_t < 2) = 3\,e^{-\lambda t}(1 - e^{-\lambda t})^2 + (1 - e^{-\lambda t})^3$

$= 3\,e^{-\lambda t} - 6\,e^{-2\lambda t} + 3\,e^{-3\lambda t} + 1 - 3\,e^{-\lambda t} + 3\,e^{-2\lambda t} - e^{-3\lambda t} = 1 - 3\,e^{-2\lambda t} + 2e^{-3\lambda t}$

et donc $f_T(t) = F_T'(t) = 6\lambda\,(e^{-2\lambda t} - e^{-3\lambda t})$.

3°) Calculer l'espérance et l'écart-type de T.

Solution :

Pour toute variable aléatoire X distribuée suivant une loi exponentielle de paramètre λ, on a $\mathrm{E}[X] = \frac{1}{\lambda} = \int_0^{+\infty} t(\lambda e^{-\lambda t})dt$, d'où :

$\mathrm{E}[T] = 6\lambda \int_0^{+\infty} t\left(e^{-2\lambda t} - e^{-3\lambda t}\right)dt = \int_0^{+\infty} t(2\lambda e^{-2\lambda t})dt - 2\int_0^{+\infty} t(3\lambda e^{-3\lambda t})dt = \frac{3}{2\lambda} - \frac{2}{3\lambda} = \frac{5}{6\lambda}$.

De même, pour toute variable aléatoire X distribuée suivant une loi exponentielle de paramètre λ, on a $\mathrm{E}[X^2] = \mathrm{V}[X] + (\mathrm{E}[X])^2 = \frac{2}{\lambda^2} = \int_0^{+\infty} t^2(\lambda e^{-\lambda t})dt$, d'où :

$\mathrm{E}[T^2] = 6\lambda \int t^2\left(e^{-2\lambda t} - e^{-3\lambda t}\right)dt = 3\int_0^{+\infty} t^2(2\lambda e^{-2\lambda t})dt - 2\int_0^{+\infty} t^2(3\lambda e^{-3\lambda t})dt$

$= \frac{6}{(2\lambda)^2} - \frac{4}{(3\lambda)^2} = \frac{3}{2\lambda^2} - \frac{4}{9\lambda^2} = \frac{19}{18\,\lambda^2}$.

On a donc $\mathrm{V}[T] = \frac{19}{18\lambda^2} - \frac{25}{36\,\lambda^2} = \frac{13}{36\,\lambda^2}$.

Finalement, l'écart-type de T vaut $\frac{\sqrt{13}}{6\,\lambda}$.

Exercice 28 : Lois exponentielles et fiabilité

On étudie la fiabilité d'un système en s'intéressant à sa durée de fonctionnement sans panne, donnée par une variable aléatoire notée T, de densité f et de fonction de répartition F.

1°) On définit la fonction de fiabilité R du système comme une fonction du temps t telle que $R(t)$ représente la probabilité pour que le système soit en bon état de fonctionnement à la date t. Montrer que $R = 1 - F$.

Indication : Si, à la date t, le système est en bon état de fonctionnement, sa durée de fonctionnement sans panne T est forcément supérieure à t. Et réciproquement.

Solution : $\forall\, t \in I\!R_+$ on a : $R(t) = \mathbf{P}(T > t) = 1 - \mathbf{P}(T \leq t) = 1 - F(t)$, donc $R = 1 - F$.

2°) On suppose dans cette question que la durée de fonctionnement sans panne est distribuée suivant une loi exponentielle de paramètre λ.

a) Exprimer la fonction de fiabilité R.

Solution : $R(t) = 1 – F(t) = e^{–\lambda t}$.

b) Vérifier que $E[T] = \int_0^{+\infty} R(t)dt$.

Solution : $\mathbf{E}[\ T\ ] = \int_0^{+\infty} e^{-\lambda t}dt = \frac{1}{\lambda}$.

c) Si le système est en état de fonctionnement depuis une durée t, quelle est la probabilité pour qu'il continue encore à fonctionner sans panne pendant une durée supplémentaire Δt ?

Solution : $\mathbf{P}\ (T > t + \Delta t \mid T > t\ ) = \frac{P(T>(t+\Delta t)\cap T>t)}{P(T>t)} = \frac{P(T>(t+\Delta t))}{P(T>t)} = \frac{R(t+\Delta t)}{R(t)}$

$= \frac{e^{-\lambda(t+\Delta t)}}{e^{-\lambda t}} = e^{-\lambda\Delta t}$.

d) Pourquoi dit-on que la loi exponentielle est sans mémoire ?

Solution : Lorsque le système a fonctionné sans panne pendant une durée t, la probabilité pour qu'il continue encore à fonctionner sans panne pendant une durée supplémentaire Δt est exactement la même qu'à l'instant initial. Le fonctionnement passé du système n'a donc aucune influence sur sa fiabilité future.

3°) On définit le taux de défaillance τ du système comme une fonction du temps t telle que

$\tau(t) = \lim\limits_{dt\to 0} \frac{P(t<T<t+dt|T>t)}{dt}$.

Montrer que $\tau = -\frac{R'}{R}$.

Solution :

$\forall\ t \in \mathbb{R}_+$ on a :

$\tau(t) = \lim\limits_{dt\to 0} \frac{P(t<T<t+dt|T>t)}{dt} = \lim\limits_{dt\to 0} \frac{P(t<T<t+dt)}{dt} \times \frac{1}{P(T>t)} = \frac{f(t)}{1-F(t)}$,

donc $\tau = \frac{f}{1-F} = \frac{f}{R}$. Mais $f = F\,' = – R'$, d'où finalement : $\tau = -\frac{R'}{R}$.

4°) Montrer que lorsque T est distribuée suivant une loi exponentielle le taux de défaillance du système est constant.

Solution :

Si T est distribuée suivant une loi exponentielle de paramètre λ :

$\forall\ t \in \mathbb{R}_+$ on a : $R(t) = \mathbf{P}\ (\ T > t\ ) = 1 – \mathbf{P}\ (\ T \le t\ ) = 1 – (\ 1 – e^{-\lambda t}\ ) = e^{-\lambda t}$,

donc $R'(t) = –\lambda\ e^{-\lambda t} = –\lambda R(t)$, donc $R' = –\lambda R$.

Il en découle que τ est une fonction constante du temps, égale à λ.

5°) On suppose que T est distribuée suivant une loi de Weibull de paramètre α et λ de fonction de densité $f : t \mapsto \alpha\lambda\ (\lambda t)^{(\alpha-1)}\ e^{-(\lambda t)^{\alpha}}$, sur le support $\mathbb{R}_+$, avec $\alpha > 0$ et $\lambda > 0$.

a) En remarquant que f s'exprime comme l'opposé une dérivée, montrer que $R(t)\ =\ e^{-(\lambda t)^{\alpha}}$.

Solution :

En remarquant que f est la dérivée de l'application $t \mapsto -e^{-(\lambda t)^{\alpha}}$, on a donc

$F(t) = \int_0^t f(t)dt = \left[-e^{-(\lambda u)^{\alpha}}\right]_0^t = 1 - e^{-(\lambda t)^{\alpha}}$, d'où $R(t) = e^{-(\lambda t)^{\alpha}}$.

b) En déduire que $\tau\ (t) = \alpha\lambda\ (\lambda t)^{(\alpha-1)}$.

Solution :

$\tau(t) = -\frac{R'(t)}{R(t)} = -\frac{-\alpha\lambda^{\alpha}t^{\alpha-1}e^{-(\lambda t)^{\alpha}}}{e^{-(\lambda t)^{\alpha}}} = \alpha\lambda\ (\lambda t)^{\alpha-1}$.

c) Interpréter, en fonction des valeurs du paramètre α les variations de τ en fonction du temps. Donner des exemples concrets de systèmes avec $\alpha > 1$, $\alpha = 1$, $\alpha < 1$.

Solution :

Il est clair que le taux de défaillance est une fonction croissante du temps pour $\alpha > 1$, décroissante pour $\alpha < 1$, et constante (et égale à λ) pour $\alpha = 1$.

On peut donc considérer qu'une loi de Weibull modélise :

- pour $\alpha > 1$, la durée de bon fonctionnement avant défaillance de tout système soumis à une usure, ce qui est le cas de la plus grande part des systèmes matériels. Exemples : cartes électroniques dans un ordinateur, pièces mécaniques d'usures dans une automobile, etc. ;
- pour $\alpha = 1$, on remarquera que l'on retombe dans le cas de la loi exponentielle qui modélise la durée de bon fonctionnement avant défaillance de tout système qui ne s'use pas. Ceci ne signifie pas que la durée de bon fonctionnement est infinie, mais simplement que la probabilité de défaillance est la même pour un matériel neuf ou un matériel vieux, c'est-à-dire qu'elle est de caractère accidentel. Exemples : carosserie de l'ordinateur, meubles, vaisselle, outils simples, comme les tournevis, les clés (mécaniques ou autres) que l'on perd mais qui ne s'usent pas vraiment …
- pour $\alpha < 1$, la durée de bon fonctionnement avant défaillance de tout système qui se bonifie avec le temps. Ceci est particulièrement le cas durant la phase transitoire d'adaptation d'un système à un nouvel environnement, où un certain nombre de pannes, dites "de jeunesse" ont tendance à devenir de moins en moins fréquentes avec le temps. Exemple : réglages initiaux inadaptés d'un système matériel, bugs informatiques d'un logiciel récemment développé, …

Fonction d'une variable aléatoire

Exercice 29 :

Soit X une variable aléatoire discrète ayant la distribution suivante :

x	-1	0	1	2
P(X=x)	1/6	1/3	1/3	1/6

Soit $Y = X^2$. Donner la loi de Y.

$Y(\Omega) = \{0,1,4\}$

$P(Y = 0) = P(X^2 = 0) = P(X = 0) = \frac{1}{3}$

$P(Y = 1) = P(X^2 = 1) = P(X = 1 \text{ ou } X = -1) = P(X = 1\,) + \mathrm{P}(X = -1) = \frac{1}{6} + \frac{1}{3} = 1/2$

…

y	0	1	4
$P(Y = y)$	1/3	1/2	1/6

Exercice 30 : Espérance et variance d'une loi et de son inverse

Soit X une variable aléatoire continue de densité $f(x) = x$ sur le support $[\,0\,,\,c\,]$.

1°) Déterminer la valeur de la constante c.

2°) Calculer l'espérance et la variance de X.

3°) Soit $Y = \frac{1}{X}$.

a) Déterminer la loi de Y.

b) Déterminer **E [** Y **]** de deux manières différentes (dont une qui est immédiate).

c) Que peut-on dire de la variance de Y ?

<u>Solution :</u>

1°) Pour que f soit une densité on doit avoir $\int_0^C x dx = \left[\frac{x^2}{2}\right]_0^C = \frac{c^2}{2} = 1$, d'où $c = \sqrt{2}$.

2°) $\mathbf{E}\,[\,X\,] = \int_0^{\sqrt{2}} x^2 dx = \left[\frac{x^3}{3}\right]_0^{\sqrt{2}} = \frac{\sqrt{8}}{3}$.

$\mathbf{E}\,[\,X^2\,] = \int_0^{\sqrt{2}} x^3 dx = \left[\frac{x^4}{4}\right]_0^{\sqrt{2}} = 1$, d'où $\mathbf{V}\,[\,X\,] = \mathbf{E}\,[\,X^2\,] - \mathbf{E}\,[\,X\,]^2 = 1 - \frac{8}{9} = \frac{1}{9}$.

3°) a) On a $Y(\Omega) = [\frac{1}{\sqrt{2}}\,; +\infty[$.

Pour $x \in Y(\Omega)$, $F_Y(x) = \mathbf{P}(Y \leq x) = P\left(\frac{1}{X} \leq x\right) = P\left(X \geq \frac{1}{x}\right) = 1 - F_X\left(\frac{1}{x}\right)$,

d'où $f_Y(x) = \frac{1}{x^2} f_X\left(\frac{1}{x}\right) = \frac{1}{x^3}$.

b) $\mathbf{E}[Y]$ peut être déterminé simplement en appliquant la formule $\mathbf{E}[\varphi(X)] = \int \phi(x) f(x) dx$.

On a donc $\mathbf{E}[Y] = \int_0^{\sqrt{2}} \frac{x}{x} dx = \int_0^{\sqrt{2}} dx = \sqrt{2}$.

On peut également déterminer $\mathbf{E}[Y]$ à partir de la loi de Y :

$\mathbf{E}[Y] = \int_{1/\sqrt{2}}^{+\infty} \frac{x}{x^3} dx = \int_{1/\sqrt{2}}^{+\infty} \frac{dx}{x} = \left[-\frac{1}{x}\right]_{1/\sqrt{2}}^{+\infty} = \sqrt{2}$.

c) En revanche, $\mathbf{E}[Y^2]$ n'existe pas car la fonction $x \mapsto \frac{x^2}{x^3} = \frac{1}{x}$ n'est pas intégrable au voisinage de l'infini.

Conclusion : La variance de Y n'existe pas.

Exercice 31 : Génération d'une loi continue

Soit X une variable aléatoire continue dont l'inverse F_X^{-1} de la fonction de répartition F_X est connue. Soient U une loi uniforme sur [0 ; 1] et $Y = F_X^{-1}(U)$.

Quelle est la distribution de Y ?

Application : En déduire un algorithme pour générer des nombres pseudo-aléatoires distribués suivant une loi exponentielle en utilisant un générateur pseudo-aléatoire uniforme sur [0 ; 1].

Solution :

Soit F_Y la fonction de répartition de répartition de Y. Pour tout x de $X(\Omega)$, on a :

$F_Y(x) = \mathbf{P}(Y \leq x) = \mathbf{P}\left(F_X^{-1}(U) \leq x\right) = \mathbf{P}\left(U \leq F_X(x)\right) = F_X(x)$.

On en déduit que Y a la même distribution que X.

Application :

Pour $X \sim \mathcal{E}(\lambda)$, on a $F_X : x \mapsto 1 - e^{-\lambda x}$, d'où $F_X^{-1} : y \mapsto -\frac{1}{\lambda}\ln(1 - y)$.

Donc, en remarquant que si $U \sim \mathcal{U}_{[0\,;\,1]}$, alors $1 - U \sim \mathcal{U}_{[0\,;\,1]}$, la loi $Y = -\frac{1}{\lambda}\ln(U)$ est distribuée suivant une loi exponentielle de paramètre λ, d'où l'algorithme de génération ci-dessous :

$u \leftarrow$ tirage uniforme sur [0 ; 1] ;

$y \leftarrow -\frac{1}{\lambda}\ln(u)$;

Retourner y.

Exercice 32 : Fonctions d'une variable aléatoire

1°) *Cas discret*

Soit X une variable aléatoire uniforme sur $[\![1, n]\!]$.

a) Déterminer la loi de $Y = X^2$.

b) Déterminer la loi de $Z = X^2 - \frac{n}{2}X$.

Indication : Les valeurs prises par Z sont les $k^2 - \frac{n}{2}k$, pour $k \in [\![1, n]\!]$.

Application numérique : $n = 8, 9, 10$.

Solution :

1°) On a $X(\Omega) = [\![1, n]\!]$ et $\forall\, k \in X(\Omega)$, $\mathbf{P}\,(X = k) = \frac{1}{n}$.

a) Comme l'application $x \mapsto x^2$ est bijective, on a simplement :

$Y(\Omega) = \{\, k^2\, ;\, k \in [\![1, n]\!]\, \}$, et comme $\forall\, K \in Y(\Omega)$, $\exists!\, k \in X(\Omega)$ tel que $K = k^2$,

$\mathbf{P}\,(Y = K) = \mathbf{P}\,(X = k) = \frac{1}{n}$.

b) On a $Z(\Omega) = \left\{k^2 - \frac{n}{2}k; \mathrm{k} \in [\![0, n]\!]\right\}$.

Donc, $\forall\, K \in Z(\Omega)$, $\exists\, k \in X(\Omega)$ tel que $K = k^2 - \frac{n}{2}k$, mais k n'est pas forcément unique.

Dans ce cas,

$$\begin{aligned} P(Z = k) &= P\left(X^2 - \frac{n}{2}X = k^2 - \frac{n}{2}k\right) = P\left((X-k)(X+k) - \frac{n}{2}(X-k) = 0\right) \\ &= P\left((X-k)\left(X + k - \frac{n}{2}\right) = 0\right) = P\left(X = k \text{ ou } X = \frac{n}{2} - k\right) \\ &= P(X = k) + P\left(X = \frac{n}{2} - k\right) - P\left(X = k \text{ et } X = \frac{n}{2} - k\right). \end{aligned}$$

Si n est impair, l'évènement $X = \frac{n}{2} - k$ est impossible, et l'on retrouve la relation

$\mathbf{P}\,(Z = K) = \mathbf{P}\,(X = k) = \frac{1}{n}$.

Si n est pair, l'évènement "$X = k$ et $X = \frac{n}{2} - k$" n'est pas l'évènement impossible si et seulement si $X = \frac{n}{2} - k \Leftrightarrow k = \frac{n}{4}$.

Donc, si n est pair, mais non multiple de 4, l'évènement "$X = k$ et $X = \frac{n}{2} - k$" est impossible

et il en découle que $\mathbf{P}\,(Z = K) = \mathbf{P}\,(X = k) + \mathbf{P}\left[X = \frac{n}{2} - k\right]$:

- pour $k \in [\![1, n/2 - 1]\!]$, $\mathbf{P}\,(Z = K) = \frac{2}{n}$;

- pour $k \in [\![n/2, n]\!]$, $\mathbf{P}\,(Z = K) = \frac{1}{n}$, car comme $\frac{n}{2} - k < 1$ on a $\mathbf{P}\left[X = \frac{n}{2} - k\right] = 0$.

Enfin, si n est multiple de 4, on a :

- pour $k \in [\![0, n/2[\![\setminus \{n/4\}$, $\mathbf{P}(Z = K) = \frac{2}{n}$;

- pour $k \in \{{}^{n}/_{4}\} \cup [\![n/2, n]\!]$, $\mathbf{P}(Z = K) = \frac{1}{n}$.

Application :

a) $n = 8$: $X(\Omega) = [\![1, 8]\!]$ et $Z = X^2 - 4X$, donc $Z(\Omega) = \{-3, 0, 5, 12, 21, 32\}$ comme le montre le tableau ci-dessous :

X	1	2	3	4	5	6	7	8
$X^2 - 4X$	-3	-4	-3	0	5	12	21	32

On a $\mathbf{P}(Z = -3) = \mathbf{P}(X = 1) + \mathbf{P}(X = 3) = \frac{1}{4}$ et

$\forall\, k \in \{{}^{8}/_{4} = 2\} \cup [\![{}^{8}/_{2} = 4, 8]\!]$, $\mathbf{P}(Z = k^2 - 4k) = \mathbf{P}(X = k) = \frac{1}{8}$.

b) $n = 9$: $X(\Omega) = [\![1, 9]\!]$ et $Z = X^2 - {}^{9}/_{2}\,X$, $Z(\Omega) = \{-{}^{7}/_{2}, -5, -{}^{9}/_{2}, -2, {}^{5}/_{2}, 9, {}^{35}/_{2}, 28, {}^{153}/_{2}\}$ comme le montre le tableau ci-dessous :

X	1	2	3	4	5	6	7	8	9
$X^2 - \frac{9}{2}X$	$-\frac{7}{2}$	-5	$-\frac{9}{2}$	-2	$\frac{5}{2}$	9	$\frac{35}{2}$	28	$\frac{81}{2}$

$\forall\, k \in [\![1, 9]\!]$, $\mathbf{P}(Z = k^2 - 4k) = \mathbf{P}(X = k) = \frac{1}{9}$.

c) $n = 10$: $X(\Omega) = [\![1, 10]\!]$ et $Z = X^2 - 5X$, donc $Z(\Omega) = \{-4, -6, 0, 6, 14, 24, 36, 50\}$ comme le montre le tableau ci-dessous :

X	1	2	3	4	5	6	7	8	9	10
$X^2 - 5X$	-4	-6	-6	-4	0	6	14	24	36	50

On a $\mathbf{P}(Z = -4) = \mathbf{P}(X = 1) + \mathbf{P}(X = 4) = \frac{1}{5}$ et $\mathbf{P}(Z = -6) = \mathbf{P}(X = 2) + \mathbf{P}(X = 3) = \frac{1}{5}$.

$\forall\, k \in [\![{}^{10}/_{2} = 5, 10]\!]$, $\mathbf{P}(Z = k^2 - 4k) = \mathbf{P}(X = k) = \frac{1}{10}$.

2°) *Cas continu*

Soit X une variable aléatoire uniforme sur $[0, n]$.

a) Déterminer la loi de $Y = X^2$.

b) Déterminer la loi de $Z = X^2 - \frac{n}{2}X$.

Indication : Etudier la fonction $\varphi : x \mapsto x^2 - \frac{n}{2}x$ sur $[0, n]$, résoudre l'inéquation

$$\varphi(X) \leq x \Leftrightarrow X^2 - \frac{n}{2}X - x \leq 0$$

et remarquer que, X étant uniforme, sa fonction de répartition s'exprime très simplement.

Solution :

2°) On a $X(\Omega) = [\ 0\ ,\ n\ ]$ et $f_X(x) = \frac{1}{n}$ et donc, $\forall\ x \in X(\Omega)$, $F_X(x) = \frac{x}{n}$.

a) On a $Y(\Omega) = [\ 0\ ,\ n^2\ ]$.

$F_Y(x) = \mathbf{P}\ (\ Y \leq x\ ) = \mathbf{P}\ (\ X^2 \leq x\ ) = P\left(X \leq \sqrt{x}\right) = F_X\left(\sqrt{x}\right) = \frac{\sqrt{x}}{n}$, d'où $f_Y(x) = \frac{1}{2n\sqrt{x}}$.

b) En étudiant la fonction $\varphi : x \mapsto x^2 - \frac{n}{2}x$ sur $[\ 0\ ,\ n\ ]$, on obtient $Z(\Omega) = \left[-\frac{n^2}{16} ; \frac{n^2}{2}\right]$.

Pour $x \leq 0$, on a $F_Z(x) = \mathbf{P}\ (\ Z \leq x\ ) = \mathbf{P}\ [\ X^2 - \frac{n}{2}X \leq x\ ] = \mathbf{P}\ [\ X^2 - \frac{n}{2}X - x \leq 0\ ]$

$= \mathbf{P}\ (\ X \in [\ x_1\ ;\ x_2\ ]\ ) = \frac{x_2 - x_1}{n}$ avec $x_1 = \frac{n}{4} - \frac{1}{2}\sqrt{\frac{n^2}{4} + 4x}$ et $x_2 = \frac{n}{4} + \frac{1}{2}\sqrt{\frac{n^2}{4} + 4x}$.

Donc $F_Z(x) = \frac{1}{n}\sqrt{\frac{n^2}{4} + 4x}$, d'où $f_Z(x) = \frac{2}{n\sqrt{\frac{n^2}{4}+4x}}$.

<u>Remarque</u> : $F_Z\left(-\frac{n^2}{16}\right) = 0$ et $F_Z(0) = \frac{1}{2}$.

Pour $x \geq 0$, on a $F_Z(x) = \mathbf{P}\ (\ Z \leq x\ ) = \mathbf{P}\ [\ X^2 - \frac{n}{2}X \leq x\ ] = \mathbf{P}\ [\ X^2 - \frac{n}{2}X - x \leq 0\ ]$

$= \mathbf{P}\ (\ X \leq x_2\ ) = \frac{x_2}{n} = \frac{1}{4} + \frac{1}{2n}\sqrt{\frac{n^2}{4} + 4x}$, d'où $f_Z(x) = \frac{1}{n\sqrt{\frac{n^2}{4}+4x}}$.

<u>Remarque</u> : $F_Z(0) = \frac{1}{2}$ et $F_Z\left(\frac{n^2}{2}\right) = \frac{1}{4} + \frac{1}{2n}\sqrt{\frac{n^2}{4} + 2n^2} = \frac{1}{4} + \frac{1}{2n}\sqrt{\frac{9n^2}{4}} = 1$.

CHAPITRE 3 : VECTEURS ALEATOIRES

Vecteurs discrets

Exercice 33 :

On considère le domaine D du plan correspondant au triangle dont les sommets sont les points (0,0), (0,5) et (3,0). Le domaine D est donc délimité par les trois droites d'équations respectives $x = 0$, $y = 0$ et $5x + 3y = 15$.

On considère un vecteur discret $V = (X, Y)$ de support $D_d = \{ (x, y) \in D$, t.q. $x \in \mathbb{N}$ et $y \in \mathbb{N} \}$ (c'est-à-dire l'ensemble des points de D de coordonnée entière).

1°) Dans le cas où V est uniforme :

a) Déterminer les lois marginales de X et Y.

b) Déterminer la loi de la somme $S = X + Y$.

2°) Mêmes questions pour V défini sur le même support par $\mathbf{P}\left(V = (x, y)\right) = k\,(x + 1)$, où k est une constante à déterminer.

Solution :

1°) On a Card $\left(\text{V}(\Omega)\right) = 13$,

donc $\forall$ (x, y) $\in$ V(Ω), $\mathbf{P}\left(\text{V} = (\text{x}, \text{y})\right) = \frac{1}{13}$.

On a X(Ω) = $[\![0, 3]\!]$, et $\mathbf{P}(\text{X} = 0) = \frac{6}{13}$, $\mathbf{P}(\text{X} = 1) = \frac{4}{13}$,

$\mathbf{P}(\text{X} = 2) = \frac{2}{13}$ et $\mathbf{P}(\text{X} = 3) = \frac{1}{13}$.

On a Y(Ω) = $[\![0, 5]\!]$, et $\mathbf{P}(\text{Y} = 0) = \frac{4}{13}$, $\mathbf{P}(\text{Y} = 1) = \frac{3}{13}$,

$\mathbf{P}(\text{Y} = 2) = \mathbf{P}(\text{Y} = 3) = \frac{2}{13}$ et $\mathbf{P}(\text{Y} = 4) = \mathbf{P}(\text{Y} = 5) = \frac{1}{13}$.

$\mathbf{P}(\text{X} = 3) \times \mathbf{P}(\text{Y} = 5) = \frac{1}{169} \neq \mathbf{P}\left(\text{V} = (3, 5)\right) = 0$, donc les variables X et Y ne sont pas indépendantes.

On a S(Ω) = $[\![1, 5]\!]$, et $\mathbf{P}(\text{S} = 0) = \frac{1}{13}$, $\mathbf{P}(\text{S} = 1) = \frac{2}{13}$, $\mathbf{P}(\text{S} = 2) = \frac{3}{13}$, $\mathbf{P}(\text{S} = 3) = \frac{4}{13}$,

$\mathbf{P}(\text{S} = 4) = \frac{2}{13}$ et $\mathbf{P}(\text{S} = 5) = \frac{1}{13}$.

2°) On a k (6×1 + 4×2 + 2×3 + 1×4) = 24k = 1 $\Rightarrow$ k = $\frac{1}{24}$.

On a X(Ω) = $[\![1, 3]\!]$, et $\mathbf{P}(\text{X} = 0) = \frac{6}{24} = \frac{1}{4}$, $\mathbf{P}(\text{X} = 1) = \frac{2\times4}{24} = \frac{8}{24} = \frac{1}{3}$, $\mathbf{P}(\text{X} = 2) = \frac{3\times2}{24} = \frac{6}{24} = \frac{1}{4}$ et $\mathbf{P}(\text{X} = 3) = \frac{4}{24} = \frac{1}{6}$.

On a $Y(\Omega) = [\![1, 5]\!]$, et $\mathbf{P}(Y = 0) = \frac{1+2+3+4}{21} = \frac{10}{24}$, $\mathbf{P}(Y = 1) = \frac{1+2+3}{24} = \frac{6}{24}$,

$\mathbf{P}(Y = 2) = \mathbf{P}(Y = 3) = \frac{1+2}{24} = \frac{3}{24}$ et $\mathbf{P}(Y = 4) = \mathbf{P}(Y = 5) = \frac{1}{24}$.

On a $S(\Omega) = [\![1, 5]\!]$, et $\mathbf{P}(S = 0) = \frac{1}{24}$, $\mathbf{P}(S = 1) = \frac{3}{24}$, $\mathbf{P}(S = 2) = \frac{6}{24}$, $\mathbf{P}(S = 3) = \frac{10}{24}$,

$\mathbf{P}(S = 4) = \frac{3}{24}$ et $\mathbf{P}(S = 5) = \frac{1}{24}$.

Exercice 34 : Jetons

Une urne contient un jeton numéroté 1, deux jetons numérotés 2, trois jetons numérotés 3 et quatre jetons numérotés 4. On tire, sans remise, deux jetons. Soit X_1 la variable aléatoire donnant le numéro du premier jeton tiré et X_2 la variable aléatoire donnant le numéro du second jeton tiré.

1°) Déterminer la loi de X_1.

Solution :

On a 10 jetons.

$X_1(\Omega) = [\![1, 4]\!]$ et $\mathbf{P}(X_1 = 1) = \frac{1}{10}$, $\mathbf{P}(X_1 = 2) = \frac{2}{10}$, $\mathbf{P}(X_1 = 3) = \frac{3}{10}$, $\mathbf{P}(X_1 = 4) = \frac{4}{10}$.

2°) Déterminer la loi du vecteur $V = (X_1, X_2)$.

Solution :

On a $V(\Omega) = [\![1, 4]\!]^2 - \{(1,1)\}$.

Pour $i \neq j$, $\mathbf{P}\,(X_1 = i \cap X_2 = j) = \frac{i\times j}{90}$ et $\mathbf{P}\,(X_1 = i \cap X_2 = i) = \frac{i\times(i-1)}{90}$.

On obtient ainsi le tableau de probabilités ci-dessous (données 1/90$^{\text{èmes}}$) :

$X_2 \setminus X_1$	1	2	3	4	$\mathbf{P}(X_2 = i)$
1	0	2	3	4	9
2	2	2	6	8	18
3	3	6	6	12	27
4	4	8	12	12	36
$\mathbf{P}(X_1 = j)$	9	18	27	36	

3°) En déduire la loi de X_2.

Solution : Il s'agit de la loi marginale donnée sur la dernière colonne du tableau. On retrouve la même loi que celle de X_1.

4°) Déterminer $\mathbf{P}(X_2 \geq X_1)$.

Solution : Il suffit de sommer les probabilités correspondant à cet évènement dans le tableau. On obtient $\mathbf{P}(X_2 \geq X_1) = \frac{55}{90}$.

5°) Déterminer la loi de $R = \frac{X_2}{X_1}$ et calculer $\mathbf{E}[R]$.

Solution :

Rajoutons les valeurs de $\frac{X_2}{X_1}$ dans le tableau des probabilités de V :

$X_2 \setminus X_1$	1	2	3	4
1	0 ; 1	2 ; $^1/_2$	3 ; $^1/_3$	4 ; $^1/_4$
2	2 ; 2	2 ; 1	6 ; $^2/_3$	8 ; $^1/_2$
3	3 ; 3	6 ; $^3/_2$	6 ; 1	12 ; $^3/_4$
4	4 ; 4	8 ; 2	12 ; $^4/_3$	12 ; 1

La loi de R est donc donnée par le tableau ci-dessous :

r	$^1/_4$	$^1/_3$	$^1/_2$	$^2/_3$	$^3/_4$	1	$^4/_3$	$^3/_2$	2	3	4
$\mathbf{P}(R = r)$	$^4/_{90}$	$^3/_{90}$	$^{10}/_{90}$	$^6/_{90}$	$^{12}/_{90}$	$^{20}/_{90}$	$^{12}/_{90}$	$^6/_{90}$	$^{10}/_{90}$	$^3/_{90}$	$^4/_{90}$

On a donc $\mathbf{E}[R] = \frac{1}{90}(1 + 1 + 5 + 4 + 9 + 20 + 16 + 9 + 20 + 9 + 16) = \frac{110}{90}$.

Exercice 35 :

On considère un vecteur discret $V = (X, Y)$ défini sur $V(\Omega) = \{ (x, y) \in \mathbb{Z}^2 \text{ t.q. } |x| + |y| \leq 2 \}$.

1°) Dessiner le support $V(\Omega)$.

2°) On a $P(V = (x, y)) = \frac{|2x-y-1|}{C}$.

a) Sur la figure, écrire les valeurs de $|2x - y - 1|$ à côté de chaque point (x, y).

b) Déterminer la valeur de la constante C.

3°) Déterminer les lois marginales de V.

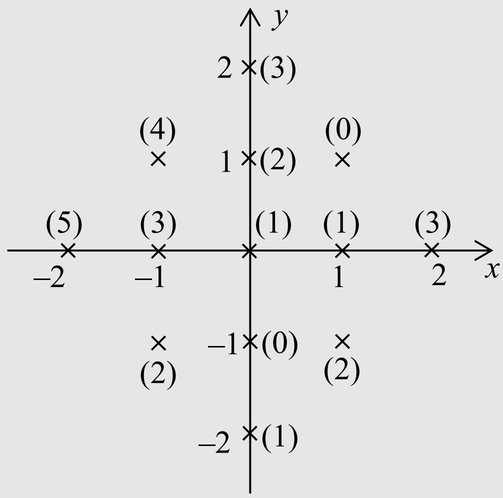
4°) Déterminer la loi de $S = X + Y$.

5°) Calculer E [XY].

Solution :

1°) $V(\Omega)$ est contient les points représentés par les croix :

2°) a) Les valeurs sont écrites entre parenthèses.



b) La somme des probabilités est égale à $\frac{27}{C}$, donc $C = 27$.

3°) Pour trouver la loi de X, on somme les probabilités des points qui se trouvent sur une même droite verticale (x constant).

$X(\Omega) = [\![-2, 2]\!]$, $P(X = -2) = \frac{5}{27}$, $P(X = -1) = \frac{9}{27}$,

$P(X = 0) = \frac{7}{27}$, $P(X = 1) = \frac{3}{27}$ et $P(X = 2) = \frac{3}{27}$.

Pour trouver la loi de Y, on somme les probabilités des points qui se trouvent sur une même droite horizontale (y constant).

$Y(\Omega) = [\![-2, 2]\!]$, $P(Y = -2) = \frac{1}{27}$, $P(Y = -1) = \frac{4}{27}$,

$P(Y = 0) = \frac{13}{27}$, $P(Y = 1) = \frac{6}{27}$ et $P(Y = 2) = \frac{3}{27}$.

4°) Pour trouver la loi de S, on somme les probabilités des points qui se trouvent sur une même droite de pente égale à –1 ($x + y$ constant)

$S(\Omega) = [\![-2, 2]\!]$, $P(S = -2) = \frac{8}{27}$, $P(S = -1) = \frac{3}{27}$,

$P(S = 0) = \frac{7}{27}$, $P(S = 1) = \frac{3}{27}$ et $P(S = 2) = \frac{6}{27}$.

5°) On applique la formule de transfert : $E[XY] = \sum_{(x,y)\in V(\Omega)} xyP(V = (x,y))$.

Le produit xy s'annule pour tous les points qui sont sur l'axe. On a donc :

$E[XY] = \frac{4}{27} \times -1 \times 1 + \frac{0}{27} \times 1 \times 1 + \frac{2}{27} \times -1 \times -1 + \frac{2}{27} \times 1 \times -1 = -\frac{4}{27}$.

Exercice 36 :

Soit p et q deux réels vérifiant $p \in]0 ; 1[$ et $q = 1 - p$. Soient X, Y deux variables aléatoires indépendantes et telles que :

$$X(\Omega) = Y(\Omega) = \mathbb{N} \text{ et } \forall k \in \mathbb{N}, \mathbf{P}(X = k) = \mathbf{P}(Y = k) = pq^k.$$

1°) Soient $S = X + Y$ et Z une loi ayant même distribution que X et Y et indépendante de S.

a) Déterminer la loi de S.

b) Pour tout entier $k \in \mathbb{N}$, calculer $\mathbf{P}(Z \geq k)$.

c) Calculer la probabilité de l'événement $S \leq Z$.

Indication : utiliser le résultat de la question précédente.

2°) On pose $V = \inf(X, Y)$ et $W = X - Y$.

a) Déterminer la loi du vecteur (V, W).

Indication : en fonction du signe de W, V sera égale à X ou à Y.

b) Déterminer les lois de V et de W.

Solution :

1°) a) On a $S(\Omega) = \mathbb{N}$, et comme X et Y sont indépendantes :

$\mathbf{P}(S = k) = \sum_{i=0}^{k} P(X = i) \cdot P(Y = k - i) = \sum_{i=0}^{k} (pq^i) \cdot (pq^{k-i}) = \sum_{i=0}^{k} p^2 q^k = (k+1)\, p^2 q^k$.

b) $\mathbf{P}(Z \geq k) = \sum_{i=k}^{+\infty} P(Z = i) = \sum_{i=k}^{+\infty} pq^i = pq^k \sum_{i=0}^{+\infty} q^i = \frac{pq^k}{1-q} = q^k$.

c) En appliquant le théorème des probabilités totales :

$\mathbf{P}(S \leq Z) = \sum_{k=0}^{+\infty} P\big((S \leq Z) \cap (S = k)\big) = \sum_{k=0}^{+\infty} P\big((Z \geq k) \cap (S = k)\big)$.

Comme Z et S sont indépendantes :

$\mathbf{P}(S \leq Z) = \sum_{k=0}^{+\infty} P(Z \geq k) \cdot P(S = k) = \sum_{k=0}^{+\infty} q^k (k+1) p^2 q^k = p^2 \sum_{k=0}^{+\infty} (k+1) q^{2k} = p^2 \sum_{k=1}^{+\infty} k (q^2)^{k-1}$.

Or d'après un résultat du cours, $\sum_{k=1}^{+\infty} k x^{k-1} = \frac{1}{(1-x)^2}$.

Donc, on obtient $\mathbf{P}(S \leq Z) = \frac{p^2}{(1-q^2)^2} = \frac{1}{(1+q)^2}$.

2°) a) Notons $U = (V, W)$. On a $V(\Omega) = \mathbb{N}$ et $W(\Omega) = \mathbb{Z}$, donc $U(\Omega) = \mathbb{N} \times \mathbb{Z}$.

$\forall (k, l) \in \mathbb{N} \times \mathbb{Z}$, $\mathbf{P}[U = (k, l)] = \mathbf{P}[(\inf(X, Y) = k) \cap (X - Y = l)]$.

1er cas : $l \geq 0$, donc $X \geq Y$, donc $\inf(X, Y) = Y$, donc :

$\mathbf{P}[U = (k, l)] = \mathbf{P}(Y = k \cap X = k + l) = \mathbf{P}(X = k + l) \,.\, \mathbf{P}(Y = k) = pq^{k+l} pq^k = p^2 q^{2k+l}$.

2ème cas : $l < 0$, donc $X < Y$, donc $\inf(X, Y) = X$, donc :

$\mathbf{P}[U = (k, l)] = \mathbf{P}(X = k \cap Y = k - l) = \mathbf{P}(X = k) \,.\, \mathbf{P}(Y = k - l) = pq^k pq^{k-l} = p^2 q^{2k-l}$.

Finalement, $\mathbf{P}[U = (k, l)] = p^2 q^{2k+|l|}$.

b) Loi de V : $V(\Omega) = \mathbb{N}$ et $\mathbf{P}(V = k) = \sum_{l=-\infty}^{+\infty} P\big((V = k) \cap (W = l)\big) = \sum_{l=-\infty}^{+\infty} p^2 q^{2k} q^{|l|}$

$= p^2 q^{2k} (1 + 2 \sum_{l=1}^{+\infty} q^l) = p^2 q^{2k} \left(1 + \frac{2q}{1-q}\right) = pq^{2k}(1+q)$.

Loi de W : $W(\Omega) = \mathbb{Z}$ et $\mathbf{P}(W = k) = \sum_{k=0}^{+\infty} P\big((V = k) \cap (W = l)\big) = \sum_{k=0}^{+\infty} p^2 q^{2k} q^{|l|}$

$= p^2 q^{|l|} \sum_{k=0}^{+\infty} (q^2)^k = \frac{p^2 q^{|l|}}{1-q^2} = \frac{pq^{|l|}}{1+q}$.

Exercice 37 :

Soit $V = (X, Y)$ un vecteur défini sur $V(\Omega) = \{(x, y) \in \mathbb{N}^2 \text{ t. q. } y > 0, x > y \; et \; 2x + y \leq 12\}$.

1°) Dessiner le support $V(\Omega)$.

2°) On a $P(V = (x, y)) = k\,\frac{x}{y}$

Sur la figure, écrire les valeurs de $\frac{x}{y}$ à côté de chaque point (x, y).

Déterminer la valeur la constante k.

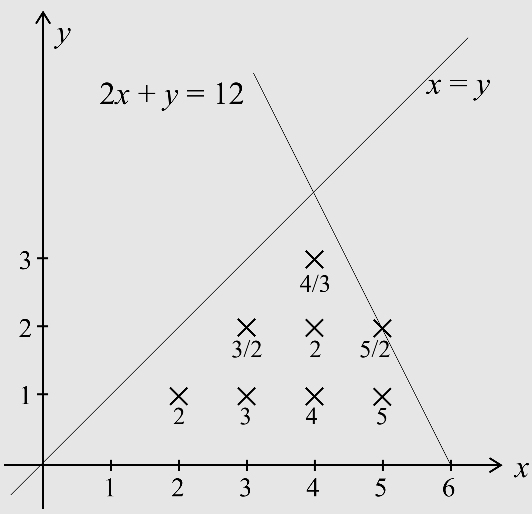
3°) Déterminer les lois marginales de V.

4°) Déterminer la loi de $S = X + Y$.

5°) Calculer $E\left[\frac{Y}{X}\right]$.

<u>Solution :</u>

1°) $V(\Omega)$ est contient les points représentés par les croix :



2°) b) On a $k\,(\,2 + 3 + 4 + 5 + {}^{3}/_{2} + 2 + {}^{5}/_{2} + {}^{4}/_{3}\,) = 1 \Rightarrow \frac{64}{3}k = 1 \Rightarrow k = \frac{3}{64}$.

3°) Pour trouver la loi de X, on somme les probabilités des points qui se trouvent sur une même droite verticale (x constant).

$X(\Omega) = [\![2, 5]\!]$, $P\,(X = 2) = \frac{6}{64} = \frac{3}{32}$, $P\,(X = 3) = \frac{27}{128}$, $P\,(X = 4) = \frac{22}{64} = \frac{11}{32}$

et $P\,(X = 5) = \frac{45}{128}$.

Pour trouver la loi de Y, on somme les probabilités des points qui se trouvent sur une même droite horizontale (y constant).

$Y(\Omega) = [\![1, 3\,]\!]$, $P\,(Y = 1) = \frac{42}{64} = \frac{21}{32}$, $P\,(Y = 2) = \frac{18}{64} = \frac{9}{32}$ et $P\,(Y = 3) = \frac{4}{64} = \frac{1}{16}$.

4°) Pour trouver la loi de S, on somme les probabilités des points qui se trouvent sur une même droite de pente égale à –1 ($x + y$ constant)

$S(\Omega) = [\![3, 7]\!]$, $P\ (S = 3) = \frac{6}{64} = \frac{3}{32}$, $P\ (S = 4) = \frac{9}{64}$, $P\ (S = 5) = \frac{33}{128}$, $P\ (S = 6) = \frac{21}{64}$

et $P\ (S = 7) = \frac{23}{128}$.

5°) On applique la formule de transfert :

$\mathrm{E}\left[\frac{Y}{X}\right] = \sum_{(x,y)\in V(\Omega)} \frac{y}{x} P(V = (x,y)) = \sum_{(x,y)\in V(\Omega)} \frac{y}{x} \times k\,\frac{x}{y} = \mathrm{Card}\,(V(\Omega)) \times k = 8 \times \frac{3}{64} = \frac{3}{8}$.

Vecteurs continus

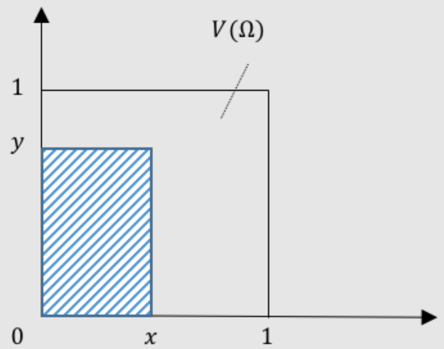
Exercice 38 : Fonction de répartition d'un couple

Soit V un vecteur uniforme sur $V(\Omega) = [0,1]^2$ et soit F sa fonction de répartition.

Que vaut $F(x,y)$ pour $(x,y) \in V(\Omega)$?

Pour la loi uniforme d'un couple, la probabilité d'appartenir à un domaine (inclus dans $V(\Omega)$) est proportionnel à l'aire du domaine.

$$F(x,y) = P(X \leq x, Y \leq y) = \frac{\text{Aire du rectangle hachuré } [0,x] \times [0,y]}{\text{Aire du carré } [0,1]^2} = xy$$



Exercice 39 :

Soit $V = (X, Y)$ un vecteur aléatoire continu de densité $f_V(x, y) = k\,x^3 y^2$ sur le support $V(\Omega) = [\,0\,;3\,] \times [\,0\,;2\,]$.

1°) Déterminer la constante k.

2°) Déterminer la fonction de répartition de V.

3°) Calculer **P** ($V \in [\,0\,;1\,]^2$).

4°) Déterminer la loi de la composante X.

5°) Calculer **P** ($1 \leq X \leq 2$).

Solution :

1°) $\int_0^3 \int_0^2 kx^3y^2\,dx\,dy = \int_0^3 kx^3 \left[\frac{y^3}{3}\right]_0^2 dx = \frac{8}{3}k\left[\frac{x^4}{4}\right]_0^3 = 54k = 1 \Rightarrow k = \frac{1}{54}$.

2°) Si $x \leq 0$ ou $y \leq 0$, $F_V(x, y) = 0$ et si $x \geq 3$ et $y \geq 2$, $F_V(x, y) = 1$.

Si $(x, y) \in V(\Omega)$, $F_V(x, y) = k\int_0^x \int_0^y u^3v^2\,dvdu = \frac{1}{54}\int_0^x u^3\left[\frac{v^3}{3}\right]_0^y dx = \frac{y^3}{162}\left[\frac{u^4}{4}\right]_0^x = \frac{x^4y^3}{648}$.

Si $x \geq 3$ et $y \leq 2$, $F_V(x, y) = F_V(3, y) = \frac{y^3}{8}$.

Si $x \leq 3$ et $y \geq 2$, $F_V(x, y) = F_V(x, 2) = \frac{x^4}{81}$.

3°) $\mathbf{P}$ ($V \in [\,0\,;\,1\,]^2$) = $F_V(1, 1) = \frac{1}{648}$.

4°) On a $X(\Omega) = [\,0\,;\,3\,]$ et $F_X(x) = F_V(x, 2) = \frac{x^4}{81}$, d'où $f_X(x) = \frac{4x^3}{81}$.

5°) $\mathbf{P}$ ($1 \leq X \leq 2$) = $F_X(2) - F_X(1) = \frac{15}{81}$.

Exercice 40 :

Considérons le couple de variables aléatoires $V = (X, Y)$ de densité $f(x, y) = kxy$ pour $(x, y) \in \mathrm{V}(\Omega)$ avec $\mathrm{V}(\Omega) = \{(x, y) \in \mathbb{R}_+^2$ tels que $2x + y \leq 2\}$.

1°) Dessiner $\mathrm{V}(\Omega)$

2°) Déterminer k .

3°) Déterminer la loi marginale de X .

1°)

2°) Intégrons la densité sur $\mathrm{V}(\Omega)$:.

$$\int_{V(\Omega)} f(x,y)dxdy \int_0^1 \int_0^{2(1-x)} kxy\, dydx \quad = k\int_0^1 x\left[\frac{y^2}{2}\right]_0^{2(1-x)} dx = 2k\int_0^1 x(1-x)^2 dx$$

$$= 2k\int_0^1 (x - 2x^2 + x^3)dx = 2k\left[\frac{x^2}{2} - \frac{2x^3}{3} + \frac{x^4}{4}\right]_0^1$$

$$= 2k\left(\frac{1}{2} - \frac{2}{3} + \frac{1}{4}\right) = \frac{1}{6}k$$

Cette intégrale vaut par ailleurs 1, ce qui donne $k = 6$.

3°) Nous avons pour $x \in \mathrm{X}(\Omega) = [0,1]$:

$$f_X(x) = \int_0^{2(1-x)} kxy\, dy = 2kx(1-x)^2$$

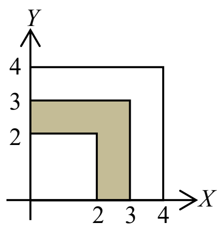
Exercice 41 :

Soit $V = (X, Y)$ un vecteur aléatoire (X, Y) tel que $V(\Omega) = [0\,;\,4]^2$ et de fonction de répartition $F_V(x, y) = k\,(x^2y + 2xy^2)$ pour $(x, y) \in [0; 4]^2$.

1°) Déterminer la valeur de la constante k.

2°) Montrer que la fonction de densité de V est $f_V(x, y) = 2k\,(x + 2y)$.

3°) Calculer la probabilité pour que V appartienne à la zone grisée ci-dessous.



Solution :

1°) On a F_V (4, 4) = 1 ⇒ 192k = 1 ⇒ $k = \frac{1}{192}$.

2°) $f_V(x, y) = \frac{\partial}{\partial x}\left(\frac{\partial}{\partial y} k(x^2y + 2xy^2)\right) = k\frac{\partial}{\partial x}(x^2 + 4xy) = 2k\,(x + 2y)$.

3°) Cette probabilité est égale à $F_V(3, 3) - F_V(2, 2) = \frac{81-24}{192} = \frac{57}{192}$.

Exercice 42 :

Soit $V = (X, Y)$ un vecteur aléatoire dont la densité est définie par :

$$f_{(X,Y)}(x, y) = \begin{cases} \frac{k}{\pi} & \text{si } (x-1)^2 + y^2 \leq 1 \\ 0 & \text{sinon} \end{cases}.$$

1°) A quel type de loi de probabilité appartient la loi du vecteur V ?

Solution : V est distribué suivant une loi uniforme (sur le disque de rayon 1 et de centre (1, 0)).

2°) Quel est le support de la fonction de densité de V ?

Solution :

Le support de la fonction de densité de V est le disque D de rayon 1 et centré en (1,0).

3°) Déterminer la valeur de k.

Solution :

On doit avoir $\int_D f_{(X,Y)}(x, y)dxdy = \frac{k}{\pi} \times$ Aire(D) = k = 1.

4°) Calculer **P** ($X + Y > 0$).

Solution :

Comme la distribution est uniforme, **P** ($X + Y > 0$) est égale au rapport de la surface “au dessus de” la courbe $y = -x$ divisé par la surface du disque.

L’aire du disque est égale à π, et celle de la surface “au dessus de” la courbe $y = -x$ est égale au trois quart du disque, $\frac{3\pi}{4}$, plus l’aire du triangle, $\frac{1}{2}$.

Finalement, **P** ($X + Y > 0$) = $\frac{\frac{3\pi}{4}+\frac{1}{2}}{\pi} = \frac{3}{4} + \frac{1}{2\pi} \approx 0{,}909$.

Exercice 43 :

Soit $V = (X, Y)$ un vecteur de densité $\sqrt{\frac{x}{y}}$ sur $[\ 0\ ;\ 1]^2$. Calculer $\mathbf{E}\left[\sqrt{Y}\right]$ de deux manières différentes :

1°) En déterminant la loi de Y à partir de celle de V, puis en calculant de manière classique.

2°) En considérant la fonction $\Phi : (x, y) \mapsto \sqrt{y}$ comme une fonction de V.

Solution :

1°) On a $Y(\Omega) = [\ 0\ ;\ 1\ ]$ et $f_Y(y) = \int_0^1 \sqrt{\frac{x}{y}}\,dx = \frac{2}{3\sqrt{y}}$.

En utilisant la formule de transfert, on obtient $\mathbf{E}\left[\sqrt{Y}\right] = \int_0^1 \sqrt{y}\,\frac{2}{3\sqrt{y}}\,dy = \frac{2}{3}$.

2°) En considérant la fonction $\Phi : (x, y) \mapsto \sqrt{y}$ comme une fonction de V, on peut directement utiliser la formule de transfert pour un vecteur :

$$E\left[\sqrt{Y}\right] = E[\Phi(V)] = \int_0^1\int_0^1 \Phi(x,y) f_V(x,y)\,dydx = \int_0^1\int_0^1 \sqrt{y}\sqrt{\frac{x}{y}}\,dydx = \int_0^1\int_0^1 \sqrt{x}\,dydx$$

$$= \int_0^1 \sqrt{x}\,dx = \frac{2}{3}$$

Exercice 44 :

Soit (X, Y) un vecteur défini sur le support $[\ -1\ ;\ 1\ ]^2$ par la fonction de densité

$$f(x, y) = k\,(x^2 + y^2).$$

1°) Calculer la valeur de k.

Solution :

On a : $\int_{-1}^1\int_{-1}^1 k(x^2+y^2)\,dxdy = 4k\int_0^1\int_0^1 k(x^2+y^2)\,dxdy = 8k\int_0^1\int_0^1 x^2\,dxdy$

(car $\int_0^1\int_0^1 y^2\,dxdy = \int_0^1\int_0^1 y^2\,dydx$ (interversion des intégrales)

$= \int_0^1\int_0^1 x^2\,dxdy$ (interversion des variables d'intégration))

$= 8k\int_0^1 \left[\frac{x^3}{3}\right]_0^1 dy = \frac{8k}{3} = 1 \Rightarrow k = \frac{3}{8}$.

2°) Calculer les probabilités suivantes :

a) $\mathbf{P}\,(V \in [-½\ ;\ ½\,]^2\,)$;

Solution :

$\mathbf{P}\,(V \in [-½\ ;\ ½\,]^2\,) = \frac{3}{8}\int_{-\frac{1}{2}}^{\frac{1}{2}}\int_{-\frac{1}{2}}^{\frac{1}{2}}(x^2+y^2)\,dxdy = 3\int_0^{\frac{1}{2}}\int_0^{\frac{1}{2}} x^2\,dxdy = \frac{1}{16}$

b) $\mathbf{P}\,(X > Y)$;

Solution :

On a Ω = “$X > Y$” ∪ “$X < Y$” ∪ “$X = Y$”, donc **P** ($X > Y$) + **P** ($X < Y$) + **P** ($X = Y$) = 1.

Or **P** ($X = Y$) = $\frac{3}{8}\int_{-\frac{1}{2}}^{\frac{1}{2}}\int_{x}^{x}(x^2+y^2)dxdy = 0$ et, par symétrie, **P** ($X > Y$) = **P** ($X < Y$), d'où

P ($X > Y$) = ½.

c) **P** ($Y \geq 0{,}5 \cap X = 0$).

Solution :

P ($Y \geq 0{,}5 \cap X = 0$) = $\frac{3}{8}\int_{0}^{0}\int_{0.5}^{1}(x^2+y^2)dydx = 0$.

3°) Déterminer la loi marginale de X et de Y.

Solution :

On a $X(\Omega) = [-1 ; 1]$ et $f_X(x) = \frac{3}{8}\int_{-1}^{1}(x^2+y^2)dy = \frac{3}{4}\left[x^2y+\frac{y^3}{3}\right]_0^1 = \frac{3}{4}\left(x^2+\frac{1}{3}\right)$

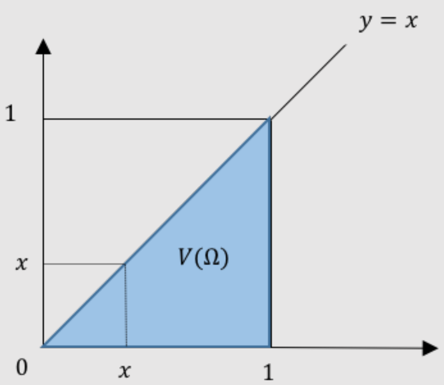
Par symétrie on a $Y(\Omega) = [-1 ; 1]$ et $f_Y(y) = \frac{3}{4}\left(y^2+\frac{1}{3}\right)$.

Exercice 45 : Formule de transfert

Soit $V = (X, Y)$ un vecteur continu, de densité f défini sur $V(\Omega) = \{(x,y) \in [0,1]^2 : x \geq y\}$ par :

$$\forall (x,y) \in V(\Omega), f(x,y) = \frac{9}{2}\sqrt{xy}.$$

Que vaut $E\left[\sqrt{\frac{X}{2Y}}\right]$?



$$E\left[\sqrt{\frac{X}{2Y}}\right] = \int_{V(\Omega)}\sqrt{\frac{x}{2y}}f(x,y)dxdy = \int_{V(\Omega)}\frac{9x}{2\sqrt{2}}dxdy = \int_0^1\int_0^x\frac{9x}{2\sqrt{2}}dydx = \int_0^1\frac{9x^2}{2\sqrt{2}}dx$$
$$= \frac{3}{2\sqrt{2}}$$

Exercice 46 : Transformation en coordonnées polaires

Soit $V = (X, Y)$ et $W = (R, \Theta)$, avec $R(\Omega) = \mathbb{R}^+, \Theta(\Omega) = [0,2\pi[$ et

$$\begin{cases} X = R\cos\Theta \\ Y = R\sin\Theta \end{cases}$$

On note f_W et f_V les densités respectives de W et V.

1°) Donner la fonction ψ telle que $V = \psi(\Omega)$.

2°) Donner le Jacobien de ψ.

3°) En déduire f_W si $f_V(x,y) = 4xy$ et $V(\Omega) = \{(x,y) \in \mathbb{R}^2$ tels que $x \geq 0, y \geq 0$ et $x^2 + y^2 \leq 1\}$.

1°) $\psi : (r,\theta) \to (r\cos\theta\,, r\sin\theta)$

2°) Le Jacobien de Ψ vaut $J_\psi(r,\theta) = \begin{vmatrix} \cos\theta & -\mathrm{r}\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r$

3°) Nous savons que $f_W(r,\theta) = |J_\psi(r,\theta)| \times f_W[\psi(y)]$. D'où

$$f_W(r,\theta) = \mathrm{r}f_V(r\cos\theta\,, r\sin\theta) = r^3\cos\theta\sin\theta = \frac{\mathrm{r}^3}{2}\sin 2\theta$$

pour $(r,\theta) \in W(\Omega) = [0,1] \times [0,\frac{\pi}{2}]$.

Exercice 47 :

Soit $V = (X, Y)$ un vecteur aléatoire défini par :

$$V(\Omega) = [0\,;\,1]^2 \text{ et } \forall\ (x, y) \in V(\Omega), f_V(x, y) = x + y.$$

1°) Déterminer la loi de $S = X + Y$.

2°) Soit $W = (X, S)$.

a) Déterminer et représenter graphiquement $W(\Omega)$.

b) Déterminer la fonction de densité de W.

c) Déduire de la question précédente la loi de S.

<u>Solution :</u>

1°) On a $S(\Omega) = [0\,;\,2]$.

Si $s \leq 1$: $F_S(s) = \int_0^s \int_0^{s-x} (x+y)dydx = \int_0^s \left[xy + \frac{y^2}{2}\right]_0^{s-x} dx$

$= \int_0^s \left(x(s-x) + \frac{(s-x)^2}{2}\right) dx = \frac{1}{2}\int_0^s (s^2 - x^2)dx = \frac{1}{2}\left[s^2x - \frac{x^3}{3}\right]_0^s = \frac{s^3}{3}$

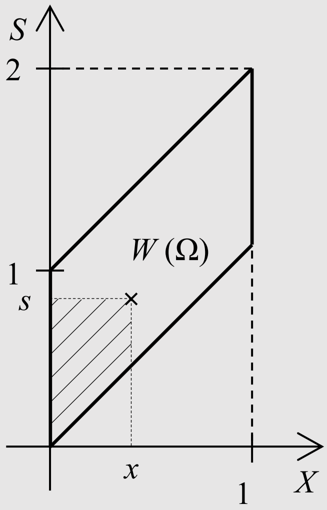
Donc, pour $s \leq 1$, $f_S(s) = s^2$.

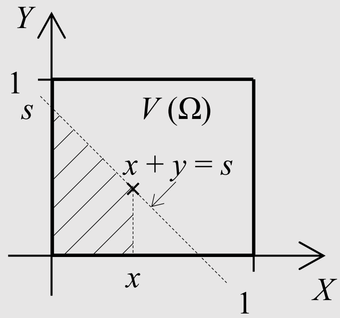
Si $s \geq 1$: $F_S(s) = \int_0^{s-1}\int_0^1 (x+y)dydx + \int_{s-1}^1\int_0^{s-x}(x+y)dydx = -\frac{s^3}{3} + s^2 - \frac{1}{3}$

Donc $f_S(s) = 2s - s^2$

2°) a) $W = (X, S) = (X, X + Y)$, donc $W(\Omega) = \{ (x, x+y), x, y \in [0 ; 1] \}$.

Représentation graphique de $W(\Omega)$:





b) La relation $W = \varphi(V) = \varphi(X, Y)$ s'écrit matriciellement $\begin{pmatrix} X \\ S \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} X \\ Y \end{pmatrix}$.

La relation réciproque $V = \varphi^{-1}(V) = \psi(W)$ s'écrit donc $\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} X \\ S \end{pmatrix}$.

D'après le cours, on a $f_W(x, s) = f_V(\psi(x, s)) \times |J_\psi(x, s)| = f_V(x, y) \times |J_\psi(x, s)|$, où $|J_\psi(x, s)|$ désigne le jacobien de ψ, c'est-à-dire le déterminant de la matrice jacobienne de ψ.

Ici, ψ est une application linéaire $(x, s) \mapsto \begin{pmatrix} \psi_1(x,s) \\ \psi_2(x,s) \end{pmatrix} = \begin{pmatrix} x \\ -x+s \end{pmatrix}$, donc la matrice jacobienne de ψ est $J_\psi(x, s) = \begin{pmatrix} \frac{\partial \psi_1}{\partial x} & \frac{\partial \psi_1}{\partial s} \\ \frac{\partial \psi_2}{\partial x} & \frac{\partial \psi_2}{\partial s} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$, d'où $|J_\psi(x, s)| = 1$.

D'autre part, $f_V(x, y) = x + y = s$, donc, finalement :

$\forall (x, s) \in W(\Omega), f_W(x, s) = s$.

c) Il suffit de déterminer la loi marginale de S sur $W(\Omega)$:

Si $s \leq 1$, $f_S(s) = \int_0^s sdx\ [sx]_0^s = s^2$,

et si $s \geq 1$, $f_S(s) = \int_{s-1}^1 sdx = [sx]_{s-1}^1 = s - s(s-1) = 2s - s^2$.

CHAPITRE 4 : INDEPENDANCE ET CORRELATION

Exercice 48 :

Soit $V = (X, Y)$ un vecteur aléatoire discret et uniforme sur

$$V(\Omega) = \{(1,2), (2,3), (3,4)\}$$

1°) Donner la covariance de X et Y.

2°) Donner le coefficient de corrélation linéaire de X et Y.

3°) X et Y sont-elles indépendantes ? Sont-elles corrélées linéairement ?

Solution :

1°) Nous avons : $E(X) = 2, E(Y) = 3, E(XY) = \frac{20}{3}$.

Puis

$$cov(X, Y) = E(XY) - E(X)E(Y) = \frac{2}{3}$$

2°) Les variables X et Y ont une relation affine positive ($Y = X + 1$) donc $\rho(X, Y) = +1$.

On peut retrouver ce résultat par le calcul. Nous avons

$$Var(X) = E(X^2) - E(X)^2 = \frac{14}{3} - 4 = \frac{2}{3}$$

$$Var(Y) = E(Y\text{^}2) - E(Y)^2 = \frac{29}{3} - 9 = \frac{2}{3}.$$

Puis :

$$\rho(X, Y) = \frac{cov(X, Y)}{\sigma(X)\sigma(Y)} = \frac{cov(X, Y)}{\sqrt{Var(X)Var(Y)}} = 1$$

3°) $cov(X, Y) \neq 0$ donc X et Y sont corrélées. Il suit que X et Y ne sont pas indépendantes.

Exercice 49 :

Soit X uniforme sur $\{-1,0,1\}$ et $Y = X^2$. Montrer que X et Y ne sont pas indépendantes mais sont en revanche corrélées linéairement.

Exercice 50 :

Soit $V = (X, Y)$ un vecteur aléatoire de densité $f_V(x,y) = k(x + y)$ sur le domaine

D = $\{ (x,y) \in I\!R^2 : x \geq 0, y \geq 0, x + y \leq 1 \}$.

1°) Déterminer la valeur de k.

2°) Déterminer les densités marginales des composantes X et Y.

3°) Calculer la covariance de X et Y.

4°) Calculer le coefficient de corrélation linéaire de X et Y.

5°) X et Y sont-elles indépendantes ? Sont-elles corrélées linéairement ?

Solution :

1°) $\int f_V(x,y)dydx = k\int\int(x+y)dydx$

$= k\int[xy + y^2/2]_0^{1-x}dx$

$= k\int\left[x(1-x) + \frac{(1-x)^2}{2}\right]dx$

$= \frac{k}{2}\int(1-x^2)dx$

$= \frac{k}{2}\left[x - \frac{x^3}{3}\right]_0^1 = \frac{k}{3} = 1 \Rightarrow k = 3.$

2°) $f_X(x) = 3\int(x+y)dy = \frac{3}{2}(1-x^2)$, donc par symétrie $f_Y(y) = \frac{3}{2}(1-y^2)$.

3°) $\mathbf{cov}(X,Y) = \mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y]$.

$\mathrm{E}[X] = \frac{3}{2}\int x(1-x^2)dx = \frac{3}{2}\left[\frac{x^2}{2} - \frac{x^4}{4}\right]_0^1 = \frac{3}{8} = \mathrm{E}[Y].$

$\mathrm{E}[XY] = 3\int\int xy(x+y)dydx$

$= 3\int\left[\frac{x^2y^2}{2} + \frac{xy^3}{3}\right]_0^{1-x}dx$

$= 3\int\left[\frac{x^2(1-x)^2}{2} + \frac{x(1-x)^3}{3}\right]dx$

$= \frac{1}{2}\int(2x - 3x^2 + x^4)dx$

$= \frac{1}{2}\left[x^2 - x^3 + \frac{x^5}{5}\right]_0^1 = \frac{1}{10}.$

D'où finalement $\mathbf{cov}(X,Y) = \frac{1}{10} - \frac{9}{64} = -\frac{13}{320}$.

4°) X et Y ne sont pas indépendantes car elles sont corrélées ($\mathbf{cov}(X,Y) \neq 0$).

Exercice 51 :

Soit $V = (X, Y, Z)$ un vecteur aléatoire de support $[\,0\,;\,1\,]^3$ et de densité conjointe f_V définie par :

$\forall\ (x, y, z) \in [\,0\,;\,1\,]^3, f_V(x, y, z) = k\,(xy + yz + xz)$.

1°) Déterminer la valeur de la constante k.

Solution :

On a $k\int\int\int(xy + yz + xz)dz\,dy\,dx = k\int\int\left(xy + \frac{y}{2} + \frac{x}{2}\right)dy\,dx = k\int\left(\frac{x}{2} + \frac{1}{4} + \frac{x}{2}\right)dx$

$= k\int \left(x+\frac{1}{4}\right)dx = k\left(\frac{1}{2}+\frac{1}{4}\right) = \frac{3}{4}k = 1 \Rightarrow k = \frac{4}{3}$.

2°) a) Montrer que $F_V(x, y, z) = \frac{x^2y^2z+xy^2z^2+x^2yz^2}{3}$ sur [0,1]^3.

Solution :

$F_V(x, y, z) = \frac{4}{3}\int\int\int(uv+vt+ut)dt\,dv\,du = \frac{4}{3}\int\int\left(uvz+\frac{vz^2}{2}+\frac{uz^2}{2}\right)dv\,du$

$= \frac{4}{3}\int\left(\frac{uy^2z}{2}+\frac{y^2z^2}{4}+\frac{uyz^2}{2}\right)du = \frac{4}{3}\left(\frac{x^2y^2z}{4}+\frac{xy^2z^2}{4}+\frac{x^2yz^2}{4}\right) = \dfrac{x^2y^2z+xy^2z^2+x^2yz^2}{3}$.

b) Retrouver la densité f_V à partir de la fonction de répartition F_V.

Solution :

$\frac{\partial^3}{\partial x\partial y\partial z}F_V(x,y,z) = \frac{1}{3}\frac{\partial}{\partial x}\left(\frac{\partial}{\partial y}\left(\frac{\partial}{\partial z}(x^2y^2z+xy^2z^2+x^2yz^2)\right)\right)$

$= \frac{1}{3}\frac{\partial}{\partial x}\left(\frac{\partial}{\partial y}(x^2y^2+2xy^2z+2x^2yz)\right) = \frac{1}{3}\frac{\partial}{\partial x}(2x^2y+4xyz+2x^2z)$

$= \frac{1}{3}(4xy+4yz+4xz) = f_V(x, y, z)$

3°) a) Montrer que $F_X(x) = F_V(x,1, 1)$.

Solution : $F_X(x) = \mathbf{P}(X \leq x) = \mathbf{P}(X \leq x \cap Y \leq 1 \cap Z \leq 1) = F_V(x,1, 1)$.

b) En déduire F_X et f_X, puis f_Y et f_Z.

Solution :

$F_X(x) = F_V(x,1, 1) = \frac{2x^2+x}{3} \Rightarrow f_X(x) = \dfrac{1}{3}(4x+1)$.

Par symétrie du problème, on déduit que $f_Y(y) = \frac{1}{3}(4y+1)$ et $f_Z(z) = \frac{1}{3}(4z+1)$.

c) Retrouver directement les densités marginales f_X, f_Y et f_Z à partir de la densité conjointe f_V.

Solution :

$f_X(x) = \frac{4}{3}\int\int(xy+yz+xz)dz\,dy = \frac{4}{3}\int\left(xy+\frac{y}{2}+\frac{x}{2}\right)dy = \frac{4}{3}\left(\frac{x}{2}+\frac{1}{4}+\frac{x}{2}\right) = \dfrac{1}{3}(4x+1)$,

et on obtient donc les mêmes formules pour f_Y et f_Z.

d) Les composantes X, Y et Z du vecteur V sont-elles indépendantes entre elles ?

Solution :

On a clairement $f_V(x, y, z) \neq f_X(x) \times f_Y(y) \times f_Z(z)$, donc les composantes X, Y et Z du vecteur V ne sont pas indépendantes entre elles.

4°) a) Montrer que $F_{(X, Y)}(x, y) = F_V(x,y, 1)$.

Solution :

$F_{(X, Y)}(x, y) = \mathbf{P}(X \leq x \cap Y \leq y) = \mathbf{P}(X \leq x \cap Y \leq y \cap Z \leq 1) = F_V(x,y, 1)$.

b) En déduire $F_{(X,\,Y)}$, puis $f_{(X,\,Y)}$.

Solution :

$F_{(X,\,Y)}(x, y) = F_V(x,y, 1) = \frac{x^2y^2+xy^2+x^2y}{3}$, d'où $f_{(X,\,Y)}(x, y) = \frac{1}{3}\frac{\partial}{\partial x}\left(\frac{\partial}{\partial y}(x^2y^2 + xy^2 + x^2y)\right)$

$= \frac{1}{3}\frac{\partial}{\partial x}(2x^2y + 2xy + x^2) = \frac{1}{3}(4xy + 2y + 2x)$.

c) Retrouver directement la densité $f_{(X,\,Y)}$ à partir de la densité conjointe f_V.

Solution :

$f_{(X,\,Y)}(x, y) = \frac{4}{3}\int(xy + yz + xz)dz = \frac{4}{3}\left(xy + \frac{y}{2} + \frac{x}{2}\right) = \frac{1}{3}(4xy + 2y + 2x)$.

d) Les composantes X, Y et Z du vecteur V sont-elles indépendantes deux à deux ?

Solution :

On a clairement $f_{(X,\,Y)}(x, y) \neq f_X(x) \times f_Y(y)$, donc les composantes X et Y du vecteur V ne sont pas indépendantes deux à deux. Par symétrie, il en est de même de X et Z et de Y et Z.

5°) a) Calculer $\mathbf{E}[XY]$ en utilisant la densité $f_{(X,\,Y)}$ du vecteur (X, Y).

Solution :

$\mathbf{E}[XY] = \frac{1}{3}\int\int xy(4xy + 2y + 2x)dy\,dx = \frac{1}{3}\int\int(4x^2y^2 + 2xy^2 + 2x^2y)dy\,dx$

$= \frac{1}{3}\int\left(\frac{4x^2}{3} + \frac{2x}{3} + x^2\right)dx = \frac{1}{3}\left(\frac{4}{9} + \frac{1}{3} + \frac{1}{3}\right) = \frac{10}{27}$.

b) Calculer $\mathbf{E}[XY]$ en utilisant la densité f_V du vecteur V.

Solution :

$\mathbf{E}[XY] = \frac{4}{3}\int\int\int xy(xy + yz + xz)dz\,dy\,dx = \frac{4}{3}\int\int\int(x^2y^2 + xy^2z + x^2yz)dz\,dy\,dx$

$= \frac{4}{3}\int\int\left(x^2y^2 + \frac{xy^2}{2} + \frac{x^2y}{2}\right)dy\,dx = \frac{4}{3}\int\left(\frac{x^2}{3} + \frac{x}{6} + \frac{x^2}{4}\right)dx = \frac{4}{3}\left(\frac{1}{9} + \frac{1}{12} + \frac{1}{12}\right)$

$= \frac{4}{3}\left(\frac{1}{9} + \frac{1}{6}\right) = \frac{4}{3} \times \frac{5}{18} = \frac{10}{27}$.

c) Calculer le coefficient de corrélation entre X et Y.

Solution :

On a $\rho(X, Y) = \frac{cov(X,Y)}{\sigma_X\sigma_Y} = \frac{E[XY]-E[X]\,E[Y]}{\sigma_X\sigma_Y}$.

On a : $\mathbf{E}[X] = \frac{1}{3}\int x(4x + 1)dx = \frac{1}{3}\int(4x^2 + x)dx = \frac{1}{3}\left(\frac{4}{3} + \frac{1}{2}\right) = \frac{11}{18} = \mathbf{E}[Y]$.

$\mathbf{E}[X^2] = \frac{1}{3}\int x^2(4x + 1)dx = \frac{1}{3}\int(4x^3 + x^2)dx = \frac{1}{3}\left(1 + \frac{1}{3}\right) = \frac{4}{9} = \mathbf{E}[Y^2]$.

Donc, $\mathbf{V}[X] = \mathbf{E}[X^2] - (\mathbf{E}[X])^2 = \frac{4}{9} - \frac{121}{324} = \frac{144-121}{324} = \frac{23}{324}$.

Comme $\sigma_X = \sigma_Y$, $\sigma_X \times \sigma_Y = \sigma_X^2 = \frac{23}{324}$.

Finalement $\rho(X, Y) = \frac{\frac{10}{27}-\frac{121}{324}}{\frac{23}{324}} = \frac{\frac{120}{324}-\frac{121}{324}}{\frac{23}{324}} = -\frac{1}{23}$.

Exercice 52 :

Lors d'un jeu télévisé, chaque candidat tire au hasard 3 questions parmi 9 questions comprenant 3 thèmes différents : histoire, sport et connaissance des arts (3 questions par thème). X (respectivement Y) représente le nombre de questions tirées par le candidat qui portent sur l'histoire (respectivement sur le sport).

1°) Déterminer les lois conjointe et marginales du couple (X, Y).

2°) Déterminer cov (X, Y).

3°) X et Y sont-elles indépendantes ? Justifiez votre réponse de deux manières différentes.

<u>Solution :</u>

1°) Il y a C_9^3 possibilités de tirer 3 questions parmi 9, l'ordre ne comptant pas. Il y a C_3^x possibilités de tirer x questions parmi les 3 questions d'histoire, avec $x \in \{0,1,2,3\}$. De même, il y a C_3^y possibilités de tirer y questions parmi les 3 questions de sport, avec $y \in \{0,1,2,3\}$. Sachant que $(x+y)$ questions ont été sélectionnées en histoire et sport, il reste C_3^{3-x-y} possibilités de tirer $(3-x-y)$ questions en connaissance des arts parmi les 3 questions en connaissance des arts. Au final, nous avons pour $0 \leq x+y \leq 3$:

$$P[(X,Y) = (x,y)] = \frac{C_3^x \times C_3^y \times C_3^{3-x-y}}{C_9^3},$$

D'où la loi conjointe donnée dans le tableau ci-dessous :

$y \setminus x$	0	1	2	3	Loi de Y
0	$\frac{1}{84}$	$\frac{3\times1\times3}{84}$	$\frac{3\times1\times3}{84}$	$\frac{1}{84}$	$\frac{20}{84}$
1	$\frac{1\times3\times3}{84}$	$\frac{3\times3\times3}{84}$	$\frac{3\times3\times1}{84}$	0	$\frac{45}{84}$
2	$\frac{1\times3\times3}{84}$	$\frac{3\times3\times1}{84}$	0	0	$\frac{18}{84}$
3	$\frac{1}{84}$	0	0	0	$\frac{1}{84}$
Loi de X	$\frac{20}{84}$	$\frac{45}{84}$	$\frac{18}{84}$	$\frac{1}{84}$	

2°) $E[X] = E[Y] = \frac{45+36+3}{84} = 1$ et $E[XY] = \frac{27+36}{84} = \frac{63}{84}$,

d'où cov $(X, Y) = \frac{63}{84} - 1 = -\frac{21}{84}$.

3°) X et Y ne sont pas indépendantes parce que :

- $P(X = 3 \cap Y = 3) = 0 \neq P(X = 3) \times P(Y = 3)$;
- $\mathrm{cov}(X, Y) \neq 0$;

Exercice 53 :

Première partie :

On considère un vecteur aléatoire $V = (X, Y)$, uniforme sur $V(\Omega) = [-L ; L] \times [-H ; H]$, avec $0 < H < L$.

1°) Déterminer la densité du vecteur V.

Solution :

Puisque V est uniforme, il existe une constante C telle que $f_V(x, y) = C$ en tout point (x, y) de $V(\Omega)$. On doit, de plus avoir $\int_{V(\Omega)} f_V(x, y)dydx = 1$. On a donc :

$$\int_{V(\Omega)} Cdydx = C \times \left(Aire\ de\ V(\Omega)\right) = 4CHL = 1 \Rightarrow C = \frac{1}{4HL}$$

2°) Déterminer le coefficient de corrélation entre X et Y.

Solution :

Vérifions que X et Y sont indépendantes. Déterminons donc tout d'abord les distributions marginales de X et Y :

On a $X(\Omega) = [-L ; L]$, et pour tout x de $X(\Omega)$, $f_X(x) = \frac{1}{4HL}\int_{-H}^{+H} dy = \frac{1}{2L}$.

De même, $Y(\Omega) = [-H ; H]$, et pour tout y de $Y(\Omega)$, $f_Y(y) = \frac{1}{4HL}\int_{-L}^{+L} dx = \frac{1}{2H}$.

Comme on a $f_V(x, y) = f_X(x) \times f_Y(y)$, les variables aléatoires X et Y sont indépendantes. Leur covariance est donc nulle, ainsi que leur coefficient de corrélation.

Deuxième partie :

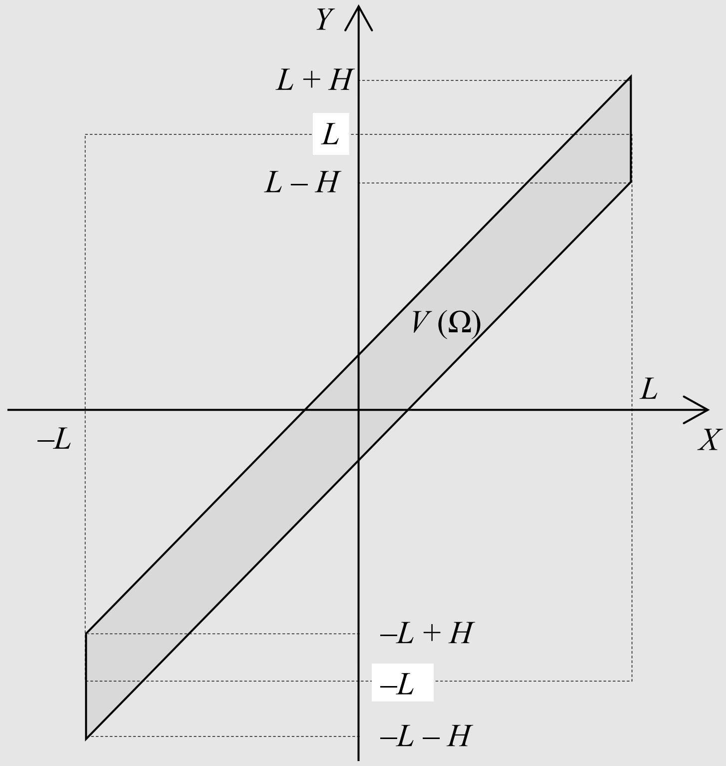
On considère un vecteur aléatoire $V = (X, Y)$, uniforme sur

$V(\Omega) = \{ (x, y) \in \mathbb{R}^2$ tel que $x \in [-L ; L]$ et $y \in [x - H ; x + H] \}$, avec $0 < H < L$.

1°) Déterminer la densité du vecteur V.

Solution :

Représentons tout d'abord graphiquement $V(\Omega)$:



Puisque V est uniforme, il existe une constante C telle que $f_V(x, y) = C$ en tout point (x, y) de $V(\Omega)$. On doit, de plus avoir $\int_{V(\Omega)} f_V(x,y)dydx = 1$. On a donc :

$$C\int_{V(\Omega)} dydx = C \times \big(Aire\ de\ V(\Omega)\big) = 4CHL \Rightarrow C = \frac{1}{4HL}$$

2°) Montrer que le coefficient de corrélation entre X et Y est égal à $\frac{\sigma_X}{\sigma_Y}$ où σ_X et σ_Y désignent les écart-type respectifs de X et Y.

<u>Solution :</u>

Le coefficient de corrélation entre X et Y est :

$\rho(X, Y) = \frac{cov\,(X,Y)}{\sigma_X\, \sigma_Y}$ avec $\mathbf{cov}(X, Y) = \mathbf{E}[XY] - \mathbf{E}[X]\,\mathbf{E}[Y]$.

Déterminons donc les distributions marginales de X et Y :

On a $X(\Omega) = [-L\,;L]$, et pour tout x de $X(\Omega)$, $f_X(x) = \frac{1}{4HL}\int_{x-H}^{x+H} dy = \frac{1}{2L}$

On a donc $X \sim \mathcal{U}_{[-L\,;\,L]}$.

De plus, $Y(\Omega) = [-L-H\,;L+H]$, et pour tout y de $Y(\Omega)$:

si $y \in [-L-H\,;-L+H]$, $f_Y(y) = \frac{1}{4HL}\int_{-L}^{y+H} dx = \frac{y+H+L}{4HL}$

si $y \in [-L+H\,;L-H]$, $f_Y(y) = \frac{1}{4HL}\int_{y-H}^{y+H} dx = \frac{1}{2L}$

si $y \in [L-H\,;L+H]$, $f_Y(y) = \frac{1}{4HL}\int_{y-H}^{L} dx = \frac{L+H-y}{4HL}$

Comme $X \sim \mathcal{U}_{[-L\,;\,L]}$, on a **E** [X] = 0, donc **cov** (X, Y) = **E** [XY] et :

$$E(XY) = \frac{1}{4HL}\int_{-L}^{+L}\int_{x-H}^{x+H} xydy\,dx = \frac{1}{4HL}\int_{-L}^{L} x\frac{(x+H)^2-(x-H)^2}{2}dx = \frac{1}{2L}\int_{-L}^{L} x^2dx$$
$$= \frac{L^2}{3}$$

De plus, **V** [X] = **E** [X^2] – **E** [X]^2 = **E** [X^2] = $\frac{1}{2L}\int_{-L}^{L} x^2dx = \frac{L^2}{3}$, d'où $\sigma_X = \frac{L}{\sqrt{3}}$.

On observe que **cov** (X, Y) = **E** [XY] = σ_X^2, et on a donc finalement :

$$\rho(X,Y) = \frac{cov\,(X,Y)}{\sigma_X\,\sigma_Y} = \frac{\sigma_X^2}{\sigma_X\,\sigma_Y} = \frac{\sigma_X}{\sigma_Y}$$

Somme de variables aléatoires

Exercice 54 :

Soient X une loi de Pascal de paramètres r et p et Y une loi géométrique de paramètre p. X et Y sont indépendantes. Déterminer la loi de $S = X + Y$.

On pourra admettre la formule $\sum_{n=p}^{N}\binom{n}{p} = \binom{N+1}{p+1}$

Solution :

On a $X(\Omega) = [\![r, +\infty [\![$ et $Y(\Omega) = \mathbb{N}^*$, donc $S(\Omega) = [\![r+1, +\infty [\![$.

Notons $q = 1 - p$.

$\forall\, k \geq r+1$, $P(S = k) = \sum_{i=r}^{k-1} P(X = i)\cdot P(Y = k-i) = \sum_{i=r}^{k-1}\binom{i-1}{r-1} p^r q^{i-r} p q^{k-i-1}$

$= p^{r+1} q^{k-(r+1)} \sum_{i=r}^{k-1}\binom{i-1}{r-1} = p^{r+1} q^{k-(r+1)} \sum_{i=r-1}^{k-2}\binom{i}{r-1} = \binom{k-1}{r} p^{r+1} q^{k-(r+1)}$.

S est donc une loi de Pascal de paramètres $r+1$ et p.

Exercice 55 :

Soient X_1 et X_2 deux variables aléatoires discrètes indépendantes, identiquement distribuées sur $\{-1, 1\}$ et définies par $P(X_1 = -1) = P(X_2 = -1) = ½$ et $P(X_1 = 1) = P(X_2 = 1) = ½$.

1°) Déterminer la loi de $Y = X_1 + X_2$.

Solution : $Y(\Omega) = \{-2, 0, 2\}$ et $\mathbf{P}(Y = -2) = \mathbf{P}(Y = 2) = ¼$ et $\mathbf{P}(Y = 0) = ½$.

2°) Soit Z une loi géométrique de paramètre p, indépendante de Y.

Déterminer la loi de $S = Y + Z$.

Solution :

On a $Z(\Omega) = \mathbb{N}^*$, donc $S(\Omega) = \{-1\} \cup \mathbb{N}$.

On a $P(S = k) = P(Z = k-2) \times P(Y = 2) + P(Z = k) \times P(Y = 0) + P(Z = k+2) \times P(Y = -2)$.

Pour $k \geq 3$, $P(S = k) = \frac{1}{4}p(1-p)^{k-3} + \frac{1}{2}p(1-p)^{k-1} + \frac{1}{4}p(1-p)^{k+1}$

$= \frac{1}{4}p(1-p)^{k-3}\left(1 + 2(1-p)^2 + (1-p)^4\right)$.

$P(S = 2) = P(Z = 2) \times P(Y = 0) + P(Z = 4) \times P(Y = -2)$

$= \frac{1}{2}p(1-p) + \frac{1}{4}p(1-p)^3 = \frac{1}{2}p(1-p)\left(1 + \frac{1}{2}(1-p)^2\right)$.

$P(S = 1) = P(Z = 1) \times P(Y = 0) + P(Z = 3) \times P(Y = -2)$

$= \frac{1}{2}p + \frac{1}{4}p(1-p)^2 = \frac{1}{2}p\left(1 + \frac{1}{2}(1-p)^2\right)$.

$P(S = 0) = P(Z = 2) \times P(Y = -2) = \frac{1}{4}p(1-p)$.

$P(S = -1) = P(Z = 1) \times P(Y = -2) = \frac{p}{4}$.

Exercice 56 :

Soit X une loi exponentielle de paramètre $\lambda > 0$.

1°) Déterminer la loi de $Y = X + 1$.

Solution :

$Y(\Omega) = [\ 1\ ;+\infty\ [$.

$F_Y(t) = P\ (X+1 \le t) = P\ (X \le t-1) = F_X(t-1) \Rightarrow f_Y(t) = f_X(t-1) = \lambda e^{-\lambda(t-1)}$.

2°) Soit Z une variable aléatoire continue uniforme sur [0 ; 1], indépendante de Y.

Déterminer la loi de $D = Y - Z$.

Solution :

On a $-Z(\Omega) = [-1\ ;\ 0]$ et $F_{-Z}(t) = P\ (-Z \le t) = P\ (Z \ge -t) = 1 - F_Z(-t) \Rightarrow f_{-Z}(t) = f_Z(t) = 1$, donc $-Z$ est uniforme sur $[-1\ ;\ 0]$.

On a donc $D(\Omega) = I\!R_+$ et, en appliquant la formule de convolution, on obtient

$$f_D(d) = \int_{-\infty}^{+\infty} f_Y(t) f_{-Z}(d-t)dt = \int_{1}^{+\infty} \lambda e^{-\lambda(t-1)} f_{-Z}(d-t)dt$$

Effectuons le changement de variable $u = d - t$. Il vient

$$f_D(d) = -\int_{d-1}^{-\infty} \lambda e^{-\lambda(d-u-1)} f_{-Z}(u)du = \int_{-\infty}^{d-1} \lambda e^{-\lambda(d-u-1)} f_{-Z}(u)du$$

$$= e^{-\lambda(d-1)} \int_{-\infty}^{d-1} \lambda e^{\lambda u} f_{-Z}(u)du$$

Pour $-1 \le d - 1 \le 0 \Leftrightarrow 0 \le d \le 1$,

$$f_D(d) = e^{-\lambda(d-1)} \int_{-1}^{d-1} \lambda e^{\lambda u}\, du = e^{-\lambda(d-1)} \left[e^{\lambda u}\right]_{-1}^{d-1} = 1 - e^{-\lambda d}$$

Pour $d - 1 \ge 0 \Leftrightarrow d \ge 1$,

$$f_D(d) = e^{-\lambda(d-1)} \int_{-1}^{0} \lambda e^{\lambda u}\, du = e^{-\lambda(d-1)} \left[e^{\lambda u}\right]_{-1}^{0} = e^{-\lambda(d-1)}\ (1 - e^{-\lambda})$$

Exercice 57 :

Soient X et Y deux variables aléatoires indépendantes, distribuées suivant une loi exponentielle de paramètre respectif λ et μ. Le but de cet exercice est de déterminer, en fonction de z, la probabilité $\mathbf{P}(X \le Y + z)$.

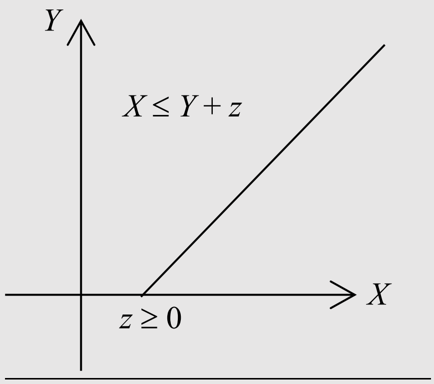
1°) Première méthode : On construit le vecteur $V = (X, Y)$.

a) Déterminer $V(\Omega)$.

Solution : $V(\Omega) = (I\!R_+)^2$.

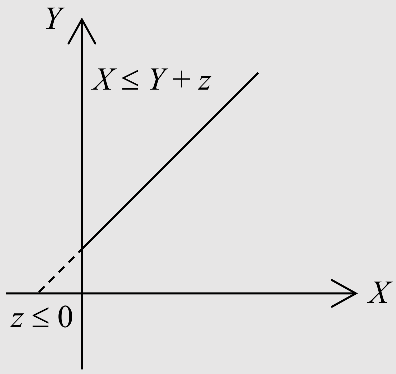
b) Dans le cas où $z \ge 0$, dessiner sur $V(\Omega)$ l'évènement $X \le Y + z$.

Solution :



c) Même question pour $z \leq 0$.

Solution :



d) Calculer $\mathbf{P}(X \leq Y + z)$.

Solution :

Pour $z \leq 0$:

$$P(X \leq Y + z) = \int_0^{+\infty} \int_{x-z}^{+\infty} \lambda\mu e^{-\lambda x} e^{-\mu y} dy\, dx = \int_0^{+\infty} \lambda e^{-\lambda x} [-e^{-\mu y}]_{x-z}^{+\infty} dx$$

$$= \int_0^{+\infty} \lambda e^{-\lambda x} e^{-\mu(x-z)} dx = e^{\mu z} \int_0^{+\infty} \lambda e^{-(\lambda+\mu)x} dx = \frac{\lambda}{\lambda+\mu} e^{\mu z}$$

Pour $z \geq 0$:

$$P(X \leq Y + z) = \int_0^{+\infty} \int_0^{y+z} \lambda\mu e^{-\lambda x} e^{-\mu y} dy\, dx = 1 - \frac{\mu}{\lambda+\mu} e^{-\lambda z}$$

2°) Deuxième méthode : On détermine, par convolution, la loi de $Z = X - Y$.

a) Déterminer la loi de $-Y$.

Solution : $(-Y)(\Omega) = I\!R_-$.

$F_{-Y}(y) = \mathbf{P}(-Y \leq y) = \mathbf{P}(Y \geq -y) = 1 - F_Y(-y) \Rightarrow f_{-Y}(y) = f_Y(-y) = \mu\, e^{\mu y}$.

b) En utilisant la formule de convolution entre X et $-Y$, déterminer la loi de $Z = X - Y$.

Solution : $Z(\Omega) = I\!R$.

$$f_Z(z) = \int_0^{+\infty} \lambda e^{-\lambda t} f_{-Y}(z-t)dt = \int_{-\infty}^{z} \lambda e^{-\lambda(z-u)} f_{-Y}(u)du$$

Pour $z \leq 0$, $f_Z(z) = \int_{-\infty}^{z} \lambda e^{-\lambda(z-u)} \mu e^{\mu u} du = \lambda\mu e^{-\lambda z} \int_{-\infty}^{z} e^{(\lambda+\mu)u} du = \frac{\lambda\mu}{\lambda+\mu} e^{\mu z}$

Pour $z \geq 0$, $f_Z(z) = \int_{-\infty}^{0} \lambda e^{-\lambda(z-u)} \mu e^{\mu u} du = \lambda\mu e^{-\lambda z} \int_{-\infty}^{0} e^{(\lambda+\mu)u} du = \frac{\lambda\mu}{\lambda+\mu} e^{-\lambda z}$

c) En déduire $\mathbf{P}(X \leq Y + z)$.

Solution :

Pour $z \leq 0$,

$$P(X \leq Y + z) = F_Z(z) = \int_{-\infty}^{z} f_Z(t)dt = \int_{-\infty}^{z} \frac{\lambda\mu}{\lambda+\mu} e^{\mu t} dt = \frac{\lambda}{\lambda+\mu} [e^{\mu t}]_{-\infty}^{z} = \frac{\lambda}{\lambda+\mu} e^{\mu z}$$

Pour $z \geq 0$,

$$P(X \leq Y + z) = F_Z(z) = \int_{-\infty}^{z} f_Z(t)dt = \int_{-\infty}^{0} \frac{\lambda\mu}{\lambda+\mu} e^{\mu t} dt + \int_{0}^{z} \frac{\lambda\mu}{\lambda+\mu} e^{-\lambda t} dt$$

$$= \frac{\lambda}{\lambda+\mu} [e^{\mu t}]_{-\infty}^{0} + \frac{\mu}{\lambda+\mu} \left[-e^{-\lambda t}\right]_{0}^{z} = \frac{\lambda}{\lambda+\mu} + \frac{\mu}{\lambda+\mu}\left(1 - e^{-\lambda z}\right)$$

$$= 1 - \frac{\mu}{\lambda+\mu} e^{-\lambda z}$$

Exercice 58 : Loi de Rayleigh

Étant données X_1 et X_2 deux variables aléatoires indépendantes et identiquement distribuées suivant une loi gaussienne centrée de variance σ^2, la variable aléatoire R, définie par $R = \sqrt{X_1^2 + X_2^2}$, est distribuée suivant une loi de Rayleigh de paramètre σ.

On notera $R \sim \mathcal{R}(\sigma)$.

1°) Déterminer la loi de $R \sim \mathcal{R}(\sigma)$.

Solution :

Notons respectivement F_R et f_R les fonctions de répartition et de densité de R.

On a clairement $R(\Omega) = \mathbb{R}_+$.

Pour déterminer F_R puis f_R, on peut considérer le vecteur aléatoire $V = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$.

On a $V(\Omega) = \mathbb{R}^2$, et, comme X_1 et X_2 sont indépendantes :

$f_V(x_1, x_2) = f_{X_1}(x_1) \times f_{X_2}(x_2) = \frac{1}{2\pi\,\sigma^2} e^{-\frac{x_1{}^2 + x_2{}^2}{2\sigma^2}}$.

En notant $D(0, r)$ le disque de $\mathbb{R}^2$ centré en $(0, 0)$ et de rayon r, on a :

$F_R(r) = \mathbf{P}(R \leq r) = \mathbf{P}\big(V \in D(0, r)\big) = \iint \frac{1}{2\pi\,\sigma^2} e^{-\frac{x_1{}^2 + x_2{}^2}{2\sigma^2}} dx_1 dx_2$

$= \frac{1}{2\pi\,\sigma^2} \int\int e^{-\frac{\rho^2}{2\sigma^2}}\,\rho\, d\rho\, d\theta = \frac{1}{\sigma^2}\int e^{-\frac{\rho^2}{2\sigma^2}}\,\rho\, d\rho,$

et donc, par définition de F_R, $f_R\,(r) = \frac{r}{\sigma^2} e^{-\frac{r^2}{2\sigma^2}}$pour $r \in \mathbb{R}_+$ et $f_R\,(r) = 0$ sinon.

2°) Calculer **E** [R].

On rappelle que $\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{x^2}{2\sigma^2}} dx = 1$

<u>Solution :</u>

E [R] $= \int r \times \frac{r}{\sigma^2} e^{-\frac{r^2}{2\sigma^2}}\, dr \overset{IPP}{=} \lim_{R\to+\infty} \left[-r e^{-\frac{r^2}{2\sigma^2}} \right]_0^R + \int e^{-\frac{r^2}{2\sigma^2}}\, dr = \frac{1}{2}\sqrt{2\pi}\sigma = \sqrt{\frac{\pi}{2}}\sigma.$

CHAPITRE 5 : DISTRIBUTIONS CONDITIONNELLES

Exercice 59 :

Soit un couple de variables aléatoires discrètes (X, Y) de distribution conjointe

$$P(X = 1, Y = 1) = 0.5, P(X = 1, Y = 2) = 0.1$$

$$P(X = 2, Y = 1) = 0.1, P(X = 2, Y = 2) = 0.3$$

1°) Donner la loi conditionnelle de X sachant que $Y = 1$ (noté $X_{|Y=1}$).

Nous avons $X_{|Y=1}(\Omega) = \{1,2\}$ et

$$P\left(X_{|Y=1} = 1\right) = P(X = 1|Y = 1) = \frac{P(X = 1, Y = 1)}{P(Y = 1)} = \frac{0.5}{0.6} = \frac{5}{6}$$

$$P\left(X_{|Y=1} = 2\right) = 1 - P\left(X_{|Y=1} = 1\right) = \frac{1}{6}$$

2°) Donner l'espérance conditionnelle de X sachant que $Y = 1$.

$$E(X|Y = 1) = \frac{5}{6} + 2 \times \frac{1}{6} = \frac{7}{6}$$

Exercice 60 :

Soit X une variable aléatoire discrète uniforme sur $[\![1; 6]\!]$. Soit Y une variable discrète telle que la loi conditionnelle $Y_{|X=n}$ est distribuée suivant une loi binomiale de paramètre n et ½. Soit $V = (X, Y)$.

1°) Déterminer $V(\Omega)$ et $Y(\Omega)$.

Solution : $V(\Omega) = \{ (x, y) : 1 \le x \le 6 \text{ et } 0 \le y \le x \}$ et $Y(\Omega) = [\![0, 6]\!]$.

2°) Démontrer que $P\big(V = (n, m)\big) = \frac{\binom{n}{m}}{6\times 2^n}$ pour $(n, m) \in V(\Omega)$.

Solution :

$P\left(V = (n, m)\right) = P\left(X = n \cap Y = m\right) = P\left(X = n\right) \times P_{|X=n}\left(Y = m\right) = P\left(X = n\right) \times P\left(Y_{|X=n} = m\right)$

$= \frac{1}{6}\times\binom{n}{m}\left(\frac{1}{2}\right)^m\left(1 - \frac{1}{2}\right)^{n-m} = \frac{\binom{n}{m}}{6\times 2^n}.$

3°) Calculer $P(Y = 5)$.

Solution : $P\left(Y = 5\right) = P\left(V = (5, 5)\right) + P\left(V = (6, 5)\right) = \frac{1}{6}\times\left(\frac{1}{32} + \frac{6}{64}\right) = \frac{1}{48}$

4°) Déterminer la loi de $X_{|Y=5}$.

Solution : $X_{|Y=5}(\Omega) = \{5, 6\}$ et

$P\left(X_{|Y=5} = 5\right) = P_{|Y=5}\left(X = 5\right) = \frac{P(X=5\cap Y=5)}{P(Y=5)} = \frac{P(V=(5,5))}{P(Y=5)} = \frac{48}{6\times 32} = \frac{1}{4}$;

$P\ (X_{|Y=5} = 6) = P_{|Y=5}\ (X= 6) = \frac{P(X=6 \cap Y=5)}{P(Y=5)} = \frac{P(V=(6,5))}{P(Y=5)} = \frac{48\times6}{6\times64} = \frac{3}{4}$.

Exercice 61 :

Soit X et Y des variables aléatoires indépendantes suivant des lois de Poisson de paramètres λ_1 et λ_2.

1°) Montrer que la loi conditionnelle de X sachant que $X+Y=n$, avec $n \in \mathbb{N}$, suit une loi binomiale de paramètres n et $p = \frac{\lambda_1}{\lambda_1+\lambda_2}$.

Notons en premier lieu que $X+Y$ suit une loi de Poisson de paramètre $\lambda_1+\lambda_2$.

Nous avons pour $k \in X_{|X+Y=n}(\Omega) = \{0, n\}$

$$P(X=k|X+Y=n) = \frac{P(X=k, X+Y=n)}{P(X+Y=n)} = \frac{P(X=k, Y=n-k)}{P(X+Y=n)}$$

$$= \frac{P(X=k)PY=n-k)}{P(X+Y=n)} = \frac{\frac{e^{-\lambda_1}\lambda_1^k}{k!}\frac{e^{-\lambda_2}\lambda_2^{n-k}}{(n-k)!}}{\frac{e^{-\lambda_1-\lambda_2}(\lambda_1+\lambda_2)^n}{n!}}$$

$$= \frac{n!}{k!\,(n-k)!}\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^k\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{n-k} = C_n^k p^k(1-p)^{n-k}$$

2°) Donner l'espérance conditionnelle de X sachant que $X+Y=n$.

L'espérance conditionnelle de X sachant que $X+Y=n$ est l'espérance d'une loi binomiale de paramètres n et p :

$$E(X|X+Y=n) = np = n\frac{\lambda_1}{\lambda_1+\lambda_2}$$

Exercice 62 :

Soit un couple de variables aléatoires continues (X, Y) de distribution conjointe

$$f(x,y) = \begin{cases} 6xy(2-x-y) & \text{si } 0 < x < 1 \text{ et } 0 < y < 1 \\ 0 & \text{sinon} \end{cases}$$

1°) Déterminer la densité conditionnelle de X sachant que $Y=y$.

Pour $y \in X_{|Y=y}(\Omega) =]0,1[$, nous avons :

$$f_{X|Y=y}(x) = \frac{f(x,y)}{f_Y(y)} = \frac{6xy(2-x-y)}{\int_0^1 6xy(2-x-y)dx} = \frac{6xy(2-x-y)}{(4-3y)y} = \frac{6x(2-x-y)}{4-3y}$$

2°) Donner l'espérance conditionnelle de X sachant que $Y=y$, pour $y \in]0,1[$.

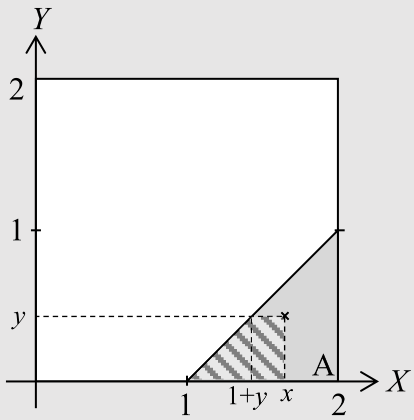
$$E(X|Y=y) = \int_0^1 x f_{X|Y=y}(x)dx = \int_0^1 \frac{6x^2(2-x-y)}{4-3y}dx = \frac{5-4y}{8-6y}$$

Exercice 63 :

Soit $V = (X,Y)$ un vecteur aléatoire uniforme sur $[0 ; 2] \times [0 ; 2]$.

1°) Soit A l'évènement $X - Y \geq 1$. Représenter <u>graphiquement</u> l'espace d'arrivée de V et l'évènement A en hachurant ce dernier.

<u>Solution :</u>



2°) Déterminer **P**(A).

<u>Solution</u> : **P** (A) = $\frac{1}{8}$.

3°) Déterminer F_V, fonction de répartition de V.

<u>Solution :</u>

$F_V(x,y) = \frac{xy}{4}$ pour $(x, y) \in [0 ; 2]^2$;

Si $x \leq 0$ ou $y \leq 0$, alors $F_V(x,y) = 0$;

Si $x \in [0 ; 2]$ et $y \geq 2$, alors $F_V(x,y) = F_V(x,2) = \frac{2x}{4} = \frac{x}{2}$;

Si $y \in [0 ; 2]$ et $x \geq 2$, alors $F_V(x,y) = F_V(2,y) = \frac{2y}{4} = \frac{y}{2}$;

Si $x \geq 2$ et $y \geq 2$, alors $F_V(x,y) = F_V(2,2) = 1$.

4°) Déterminer $F_{V|A}$, la fonction de répartition de V sachant que $V \in$ A.

<u>Solution :</u>

Nous avons par définition

$$F_{V|A}(x,y) = P(X \leq x, Y \leq y | V \in A) = \frac{P(X \leq x \cap Y \leq y \cap V \in A)}{P(V \in A)}$$

Le dénominateur a déjà été calculé : $P(V \in A) = P(A) = \frac{1}{8}$.

Reste à calculer le numérateur pour lequel nous allons devoir distinguer plusieurs cas de figure.

<u>Si $x \leq 1$ ou $y \leq 0$</u>, alors $F_{V|A}(x,y) = 0$.

<u>Si $(x,y) \in A$</u>, alors

$$F_{V|A}(x,y) = P(X \le x, Y \le y | (X,Y) \in A) = \frac{P(X \le x \cap Y \le y \cap V \in A)}{P(V \in A)}$$

Le dénominateur a déjà été calculé : $P(V \in A) = P(A) = \frac{1}{8}$. Reste à calculer le numérateur $P(X \le x \cap Y \le y \cap V \in A))$ qui est la probabilité d'appartenir au quadrilatère hachuré, c'est-à-dire son aire divisée par l'aire du carré [0 ; 2] × [0 ; 2], à savoir 4. L'aire de ce quadrilatère est égale à la somme de l'aire d'un rectangle et de celle d'un triangle. Le rectangle a une largeur égale à $x - (1 + y) = x - y - 1$ et une hauteur égale à y d'où une aire égale à $y(x - y - 1)$. Le triangle est rectangle isocèle, donc son aire est égale à la moitié de celle d'un carré de côté $(1 + y) - 1 = y$, soit $\frac{y^2}{2}$.

Ainsi, $P(X \le x \cap Y \le y \cap (X,Y) \in A)) = \frac{y(x-y-1)}{4} + \frac{y^2}{8}$

Puis

$$F_{V|A}(x,y) = 2y\,(x - y - 1) + y^2 = 2xy - y^2 - 2y \qquad \textbf{(équation 1)}$$

Si $x \in [1,2]$ et $y \ge x - 1$, alors $F_{V|A}(x,y) = F_{V|A}(x, x-1) = (x-1)^2$ d'après l'équation 1.

Si $x \ge 2$ et $y \in [0,1]$, alors $F_{V|A}(x,y) = F_{V|A}(2,y) = 2y - y^2$ d'après l'équation 1.

Si $x \ge 2$ et $y \ge 1$ alors $F_{V|A}(x,y) = F_{V|A}(2,1) = 1$ d'après l'équation 1.

5°) Déterminer $f_{V|A}$, densité de probabilité de V sachant que $V \in$ A.

Solution :

On a $f_{V|A} = \frac{\partial^2 F_{V|A}}{\partial x \partial y}$. On observe que si $(x , y) \notin$ A, alors $f_{V|A}(x, y) = 0$.

Si $(x, y) \in$ A, $f_{V|A}(x, y) = \frac{\partial}{\partial x}\left(\frac{\partial}{\partial y}(2xy - y^2 - 2y)\right) = \frac{\partial}{\partial x}(2x - 2y - 2) = 2$.

6°) La loi conditionnelle de V sachant que $V \in$ A est-elle uniforme ?

Solution : Oui, parce que sa densité est constante.

7°) Déterminer $\mathbf{E}_A[X]$, espérance conditionnelle de X sachant que $V \in$ A.

Solution :

On peut déterminer la loi marginale de la composante X_A de V |A,

$f_{X_A}(x) = \int_0^{x-1} 2\,dy = 2(x-1)$, puis en déduire $\mathbf{E}_A[X] = \int_1^2 2x(x-1)dx = 2\left[\frac{x^3}{3} - \frac{x^2}{2}\right]_1^2 = \frac{5}{3}$.

On peut aussi considérer que X est une fonction du vecteur $V = (X, Y)$ et calculer directement $\mathbf{E}_A[X]$, par la formule $\mathbf{E}_A[X] = \int_1^2 \int_0^{x-1} 2x\, dy\, dx = \int_1^2 2x(x-1)dx = \frac{5}{3}$.

8°) Déterminer $\mathbf{E}_A[Y]$, espérance conditionnelle de Y sachant que $V \in$ A.

Solution :

On peut déterminer la loi marginale de la composante Y_A de $V|A$,

$f_{Y_A}(x) = \int 2dx = 2(1-y)$, puis en déduire $\mathbf{E}_A[Y] = \int 2y(1-y)dy = 2\left[\frac{y^2}{2} - \frac{y^3}{3}\right]_0^1 = \frac{1}{3}$.

On peut aussi considérer que Y est une fonction du vecteur $V = (X, Y)$ et calculer directement $\mathbf{E}_A[Y]$, par la formule $\mathbf{E}_A[Y] = \int \int 2ydy\, dx = \int (x-1)^2 dx = \int x^2 dx = \frac{1}{3}$.

Exercice 64 :

Soit $V = (X, Y)$ un vecteur aléatoire discret représentant le résultat de deux lancers de dés (indépendants).

1°) Quel est l'univers des possibles du vecteur V ?

Solution :

$V(\Omega) = [\![1, 6]\!]^2$.

2°) Donner la loi de V.

Solution :

Si les deux dés sont non truqués et indépendants, V est distribué suivant une loi uniforme.

3°) Calculer l'espérance de X.

Solution :

$\mathbf{E}[X] = \sum_{x=1}^{6} x \times \frac{1}{6} = \frac{1}{6}\frac{6\times 7}{2} = \frac{7}{2}$.

4°) Soit $S = X + Y$ la somme des deux dés. Que vaut l'espérance de S ?

Solution :

$\mathbf{E}[S] = \mathbf{E}[X+Y] = \mathbf{E}[X] + \mathbf{E}[Y] = \frac{7}{2} + \frac{7}{2} = 7$.

5°) Quelle est la loi de S sachant que $X = 1$? sachant que $X = x$.

Solution :

$S_{X=1}(\Omega) = [\![2, 7]\!]$, et $\mathbf{P}(S_{X=1} = s) = \mathbf{P}(S = s \mid X = 1) = \frac{P(S=s\cap X=1)}{P(X=1)}$

$= \frac{P(Y=s-1\cap X=1)}{P(X=1)} = \frac{1/6\times 1/6}{1/6} = \frac{1}{6}$.

$S_{X=1}$ est donc distribué suivant une loi uniforme sur $[\![2, 7]\!]$.

$S_{X=x}(\Omega) = [\![x+1, x+6]\!]$, et $\mathbf{P}(S_{X=x} = s) = \mathbf{P}(S = s \mid X = x) = \frac{P(S=s\cap X=x)}{P(X=x)}$

$= \frac{P(Y=s-x \cap X=x)}{P(X=x)} = \frac{1/6 \times 1/6}{1/6} = \frac{1}{6}$.

$S_{X=x}$ est donc distribué suivant une loi uniforme sur $[\![x+1 , x+6]\!]$.

6°) Quelle est l'espérance de S sachant que $X = 1$? sachant que $X = x$?

En déduire $\mathbf{E}\,[S \mid X]$.

Solution :

$\mathbf{E}\,[S_{|X=1}] = \sum_{s=2}^{7} s \times \frac{1}{6} = 1 + \sum_{s=1}^{6} s \times \frac{1}{6} = 1 + \frac{7}{2} = \frac{9}{2}$.

$\mathbf{E}\,[S_{|X=x}] = \sum_{s=x+1}^{x+6} s \times \frac{1}{6} = x + \sum_{s=1}^{6} s \times \frac{1}{6} = x + \frac{7}{2} = \frac{7+2x}{2}$.

On a alors $\mathbf{E}\,[S \mid X] = X + \frac{7}{2}$.

7°) Quelle est la loi de S^2 sachant que $X = 1$? sachant que $X = x$.

Solution :

$S^2_{|X=1}\,(\Omega) = \{4, 9, 16, 25, 36, 49\}$, et $\forall\, s \in [\![2 , 7]\!]$, $\mathbf{P}\,(S^2_{|X=1} = s^2) = \mathbf{P}\,(\,S = s \mid X = 1\,) = \frac{P(S=s \cap X=1)}{P(X=1)} = \frac{P(Y=s-1 \cap X=1)}{P(X=1)} = \frac{1/6 \times 1/6}{1/6} = \frac{1}{6}$.

$S^2_{|X=1}$ est donc distribué suivant une loi uniforme sur $\{4, 9, 16, 25, 36, 49\}$.

$S^2_{|X=x}\,(\Omega) = \{(x+1)^2, (x+2)^2, (x+3)^2, (x+4)^2, (x+5)^2, (x+6)^2\}$ et $\forall\, s \in S^2_{|X=x}\,(\Omega)$,

$\mathbf{P}\,(S^2_{|X=x} = s^2) = \mathbf{P}\,(\,S = s \mid X = x\,) = \frac{P(S=s \cap X=x)}{P(X=x)} = \frac{P(Y=s-x \cap X=x)}{P(X=x)}$

$= \frac{1/6 \times 1/6}{1/6} = \frac{1}{6}$.

$S^2_{|X=x}$ est donc distribué suivant une loi uniforme sur $\{(x+1)^2, (x+2)^2, (x+3)^2, (x+4)^2, (x+5)^2, (x+6)^2\}$.

8°) Quelle est l'espérance de S^2 sachant que $X = x$. En déduire $\mathbf{E}\,[S^2 \mid X]$.

Remarque : On pourra utiliser la formule $\sum_{k=1}^{n} k^2 = \frac{n(2n+1)(n+1)}{6}$.

Solution :

$\mathbf{E}\,[S^2_{|X=x}] = \sum_{s=x+1}^{x+6} s^2 \times \frac{1}{6} = \frac{1}{6}\sum_{k=1}^{6}(x+k)^2 = \frac{1}{6}(6x^2 + 2x\sum_{k=1}^{6} k + \sum_{k=1}^{6} k^2)$.

En utilisant la formule $\sum_{k=1}^{n} k^2 = \frac{n(2n+1)(n+1)}{6}$ on obtient :

$\mathbf{E}\,[S^2_{|X=x}] = x^2 + \frac{x}{3}\left(\frac{6\times 7}{2}\right) + \frac{1}{6}\left(\frac{6\times 13\times 7}{6}\right) = x^2 + 7x + \frac{91}{6}$.

On a alors $\mathbf{E}\,[S^2 \mid X] = X^2 + 7X + \frac{91}{6}$.

9°) En déduire $\mathbf{V}\,[S \mid X] = \mathbf{E}\,[\,S^2_{|X}\,] - (\mathbf{E}\,[S \mid X])^2$, qui représente l'erreur effectuée en approximant S par $\mathbf{E}\,[S \mid X]$.

Solution :

On a $\mathbf{V}\,[S \mid X] = \left(X^2 + 7X + \frac{91}{6}\right) - \left(X + \frac{7}{2}\right)^2 = \frac{91}{6} - \frac{49}{4} = \frac{182-147}{12} = \frac{35}{12}$.

10°) Montrer que $\mathbf{V}\,[S] = \mathbf{E}\,[\mathbf{V}\,[S \mid X]] + \mathbf{V}\,[\mathbf{E}\,[S \mid X]]$.

<u>Solution :</u>

$\mathbf{V}\,[S \mid X]$ étant une constante, elle est égale à son espérance. Donc, $\mathbf{E}\,[\mathbf{V}\,[S \mid X]] = \frac{35}{12}$.

D'autre part, on a $\mathbf{E}\,[S \mid X] = X + \frac{7}{2}$, donc $\mathbf{V}\,[\mathbf{E}\,[S \mid X]] = V\,\left[X + \frac{7}{2}\right] = \mathbf{V}\,[X]$.

Calculons donc $\mathbf{V}\,[X]$: $\mathbf{V}\,[X] = \mathbf{E}\,[X^2] - (\mathbf{E}\,[X])^2$.

$\mathbf{E}\,[X] = \sum_{x=1}^{6} x \times \frac{1}{6} = \frac{1}{6} \times \frac{6\times 7}{2} = \frac{7}{2}$.

Pour calculer $\mathbf{E}\,[X^2]$, on peut utiliser la formule $\sum_{k=1}^{n} k^2 = \frac{n(2n+1)(n+1)}{6}$:

$\mathbf{E}\,[X^2] = \sum_{x=1}^{6} x^2 \times \frac{1}{6} = \frac{1}{6} \times \frac{6\times 13\times 7}{6} = \frac{91}{6}$.

Alors, $\mathbf{V}\,[X] = \frac{91}{6} - \frac{49}{4} = \frac{182-147}{12} = \frac{35}{12}$, et par conséquent $\mathbf{V}\,[\mathbf{E}\,[S \mid X]] = \frac{35}{12}$.

On a donc $\mathbf{E}\,[\mathbf{V}\,[S \mid X]] + \mathbf{V}\,[\mathbf{E}\,[S \mid X]] = \frac{35}{12} + \frac{35}{12} = \frac{70}{12}$.

Enfin, il suffit de remarquer que, comme X et Y sont indépendantes et identiquement distribuées $\mathbf{V}\,[S] = \mathbf{V}\,[X + Y] = \mathbf{V}\,[X] + \mathbf{V}\,[Y] = 2\,\mathbf{V}\,[X] = \frac{70}{12}$, ce qui prouve l'égalité demandée.

Exercice 65 :

Quand X et Y sont discrets, montrer que $\mathrm{E}(X) = \mathrm{E}[\mathrm{E}(X|Y)]$.

Rappelons que $\mathrm{E}(X|Y) = \phi(Y)$ avec $\phi(y) = \mathrm{E}(X|Y = y) = \sum_x x\mathrm{P}(X = x|Y = y)$.

Nous avons

$$\begin{aligned}
\mathrm{E}[\mathrm{E}(X|Y)] &= \mathrm{E}[\phi(\mathrm{Y})] = \sum_y \phi(\mathrm{y})P(Y = y) = \sum_y \mathrm{E}(X|Y = y)P(Y = y) \\
&= \sum_y \sum_x x\mathrm{P}(X = x|Y = y)P(Y = y) \\
&= \sum_y \sum_x P(X = x, Y = y) = \sum_x x \sum_y P(X = x, Y = y) \\
&= \sum_x xP(X = x) = E(X)
\end{aligned}$$

Exercice 66 :

Soit $X_1, X_2, \ldots$ des variables aléatoires identiquement distribuées, d'espérance finie μ et d'écart-type σ. Soit N une autre variable aléatoire, indépendante des X_i. Soit $S = \sum_{i=1}^{N} \mathrm{X_i}$. Montrer que :

$$\mathrm{E}(S) = \mathrm{E(N)}\mu$$

En supposant de plus que les X_i sont indépendants, montrer que :

$$\mathrm{Var}(S) = \sigma^2 \mathrm{E(N)} + \mu^2 \mathrm{Var}(N)$$

<u>Application :</u> Dans une petite ville, le nombre d'accidents de la route par an est en moyenne de 10. Le nombre moyen de blessés par accident est de 2. Combien y-a-t-il de blessés par an, en moyenne ?

<u>Calcul de l'espérance</u>

Nous allons utiliser la propriété $E(S) = E\big(E(S|\, N)\big)$.

Nous avons

$$E(S|N=n) = E(\textstyle\sum_{i=1}^{N} X_i \mid N=n) = E(\sum_{i=1}^{n} X_i \mid N=n) = E(\sum_{i=1}^{n} X_i\,) = n\mu$$

L'avant dernière égalité est vraie car les X_i sont indépendants de N.

Il suit que

$$\mathrm{E}(S|N) = N\mu \qquad (1)$$

Puis

$$\mathrm{E}(S) = E\big(E(S|\, N)\big) = E(N\mu) = E(N)\mu$$

<u>Calcul de la variance</u>

Nous allons utiliser la propriété $Var(S) = E(Var(S|N)) + Var(E(S|N))$.

Nous avons tout d'abord, en utilisant l'indépendance des X_i et de N

$$\begin{aligned}\mathrm{Var}(S|N=n) &= \mathrm{Var}\left(\sum_{i=1}^{N} \mathrm{X_i} \mid N=n\right) = \mathrm{Var}\left(\sum_{i=1}^{n} \mathrm{X_i} \mid N=n\right)\\ &= \mathrm{Var}\left(\sum_{i=1}^{n} \mathrm{X_i}\right) = n\sigma^2\end{aligned}$$

Puis

$$\mathrm{Var}(S|N) = N\sigma^2$$

et

$$E(\mathrm{Var}(S|N)) = \sigma^2\, E(N)$$

En utilisant l'équation (1), nous obtenons

$$\mathrm{Var(E}(S|N)) = Var(N\mu) = \mu^2\, Var(N)$$

Finalement :

$$Var(S) = E(Var(S|N)) + Var(E(S|N)) = \sigma^2 \mathrm{E}(N) + \mu^2 Var(N)$$

<u>Application :</u> Notons N la variable aléatoire représentant le nombre d'accidents de la route par an, et X_i le nombre de blessés au i-ème accident de l'année. Le nombre d'accidents par an est donc $\sum_{i=1}^{N} X_i$. Si l'on suppose que le nombre de blessés X_i est indépendant du nombre d'accidents N, nous avons alors

$$\mathrm{E}\left(\sum_{i=1}^{N} \mathrm{X_i}\right) = E(N)E(X) = 10 \times 2 = 20$$

CHAPITRE 6 : STATISTIQUES INFERENTIELLES

Exercice 67 :

1°) Soit $(X_1, \ldots, X_n)$ un échantillon aléatoire où les variables aléatoires X_i sont distribuées suivant une loi exponentielle de paramètre λ.

Que vaut la moyenne arithmétique aléatoire de cet échantillon ?

<u>Solution</u> : $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

a) De quelle caractéristique des variables X_i cette moyenne arithmétique aléatoire est-elle un estimateur ?

<u>Solution</u> : Les lois X_i sont toutes d'espérance 1 / λ : $\forall\ i \in [\![1, n]\!]$, $\mathbf{E}\,[\,X_i\,] = 1/\lambda$.

Or, la moyenne arithmétique aléatoire est un estimateur de l'espérance des X_i, donc de l'inverse du paramètre λ.

b) Est-ce un estimateur sans biais ? Pourquoi ?

<u>Solution</u> : $\bar{X}$ est un estimateur sans biais de 1 / λ puisque $E[\bar{X}] = \frac{1}{n}\sum_{i=1}^{n} E\,[X_i] = 1/\lambda$.

2°) Que vaut la variance estimée aléatoire $S_X{}^2$ de l'échantillon ? (on supposera que l'espérance commune des X_i est <u>*inconnue*</u>).

<u>Solution</u> : $S_X^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \left(\frac{1}{n}\sum_{i=1}^{n} X_i^2\right) - (\bar{X})^2$.

3°) a) De quelle caractéristique des variables aléatoires X_i, $S_X{}^2$ est-il un estimateur ?

<u>Solution</u> : S_X^2 est un estimateur de la variance des X_i, donc de $1/\lambda^2$.

b) Est-ce un estimateur sans biais ? Pourquoi ?

<u>Solution</u> : Posons $\sigma^2 = 1/\lambda^2$.

$S_X^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$ est un estimateur biaisé de σ^2 puisque $\mathbf{E}\,[\,S_X^2\,] = \sigma^2 - \frac{\sigma^2}{n} = \frac{n-1}{n}\sigma^2$.

c) $S_X{}^2$ est-il un estimateur asymptotiquement sans biais ? Pourquoi ?

<u>Solution :</u>

Oui, $S_X{}^2$ est un estimateur asymptotiquement sans biais parce que $\lim\limits_{n\to\infty} \mathbf{E}\,[\,S_X^2\,] = \sigma^2$.

4°) Vers quelle valeur tend la moyenne arithmétique de $(X_1, \ldots, X_n)$ quand n tend vers l'infini ?

<u>Solution</u> : D'après la loi des grand nombres, elle tend vers l'espérance des variables aléatoires X_i, c'est-à-dire vers 1 / λ.

Exercice 68 :

Soit ($X_1,\dots, X_n$) un échantillon aléatoire. Les variables aléatoires X_i sont distribuées suivant une loi uniforme *continue* sur [0 ; 1].

On notera $\bar{X}_n$ la moyenne arithmétique aléatoire de l'échantillon d'effectif n.

1°) On suppose que $n = 4$. Que vaut **E** [$\bar{X}_4$] ?

Solution :

$\forall\, n \in \mathbb{N}, \forall\, i \in [\![1, n]\!]$, **E** [$\bar{X}_n$] = **E** [X_i] = ½.

2°) Soit $S_X{}^2$ la variance estimée aléatoire de l'échantillon d'effectif 4. Que vaut **E** [$S_X{}^2$] ?

Solution :

$\forall\, i \in [\![1, 4]\!]$, **V** [X_i] = $\sigma^2 = 1/12$, et **E** [S_X^2] = $\sigma^2 - \frac{\sigma^2}{4} = \frac{3}{4}\sigma^2 = 1/16$.

3°) Vers quelle valeur tend $\bar{X}_n$ quand n tend vers l'infini ?

Solution : D'après la loi des grand nombres, elle tend vers l'espérance des variables aléatoires X_i, c'est-à-dire vers ½.

4°) On suppose maintenant que $n = 1000$.

Par quelle loi peut-on approximer $\bar{X}_n$?

Que vaut **P**($\bar{X}_n > 0{,}51$) ?

Solution :

a) $\bar{X}_n$ peut être approximée par une loi gaussienne d'espérance ½ et de variance $\frac{\sigma^2}{n} = \frac{1}{12000}$.

b) **P**($\bar{X}_n > 0{,}51$) = $P\left(\frac{\bar{X}_n - 1/2}{\sqrt{1/12000}} > \frac{0{,}51-0{,}5}{\sqrt{1/12000}}\right) = 1 - P\left(\frac{\bar{X}_n - 1/2}{\sqrt{1/12000}} \leq 1{,}095\right) = 1 - \Phi(1{,}095)$

$\approx 1 - 0{,}863 = 13{,}7\ \%$

Exercice 69 : Estimateur de l'écart-type d'une loi normale

On va maintenant proposer un estimateur de l'écart-type σ dans le cas particulier où X suit une loi normale $\mathcal{N}(\mu,\sigma)$ et μ est supposé connu. Cet estimateur est défini par :

$$W = \frac{1}{n}\sum_{i=1}^{n} |X_i - \mu|$$

1) Soit Z une variable aléatoire de loi normale centrée réduite. Calculer l'espérance de $|Z|$ (valeur absolue de Z) en fonction de σ .

Réponse : Calculons $E[|Z|]$ avec Z une loi normale centrée réduite. On rappelle que la densité f de Z vaut : $f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$.

Il vient alors:

$$\begin{aligned} E[|Z|] &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} |z| e^{-\frac{z^2}{2}} dz = 2\int_{0}^{+\infty} \frac{1}{\sqrt{2\pi}} |z| e^{-\frac{z^2}{2}} dz \\ &= \sqrt{\frac{2}{\pi}} \left[e^{-\frac{z^2}{2}} \right]_0^{\infty} = \sqrt{\frac{2}{\pi}} \end{aligned}$$

2) En déduire que W est un estimateur biaisé de σ et que $W' = W\sqrt{\frac{2}{\pi}}$ est un estimateur sans biais de σ .

Réponse : On peut écrire $X_i - \mu = Z_i\sigma$ avec Z_i une loi normale centrée réduite. L'espérance de W s'écrit alors :

$$\begin{aligned} E[W] &= \frac{1}{n}\sum_{i=1}^{n} E\,[|X_i - \mu|] = \frac{1}{n}\sum_{i=1}^{n} E\,[|Z_i\sigma|] \\ &= \frac{1}{n}\sigma\sum_{i=1}^{n} E\,[|Z_i|] = \sigma E[|Z|] = \sigma\sqrt{\frac{2}{\pi}} \end{aligned}$$

W est donc un estimateur biaisé de σ tandis que $W' = \sqrt{\frac{\pi}{2}}W$ est un estimateur non biaisé. En effet $E(W') = \sqrt{\frac{\pi}{2}}E(W) = \sigma$

Remarque : Cet estimateur est plus facile à utiliser par des "non-mathématiciens" car il utilise uniquement des additions et des soustractions, contrairement à l'estimateur classique qui demande de calculer aussi des carrés et une racine carrée.

3) Montrer que W' est un estimateur convergent vers σ.

Nous avons $W = \frac{1}{n}\sum_{i=1}^{n} A_i$ où $A_i = |X_i - \mu|$. D'après la loi des grands nombres, W converge vers $E(A_i) = \sigma\sqrt{\frac{2}{\pi}}$. On conclut que W' converge vers σ.

Exercice 70 : Loi du chi-deux

1) Rappeler quelle est la densité f_U d'une variable aléatoire normale centrée réduite U.

Réponse :

$$f_U(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$$

2) Rappelons que le moment d'ordre i de U est l'espérance de U^i , c'est-à-dire $E(U^i)$. Calculer les moments d'ordre 1, 2, 3 et 4 de U.

Réponse : D'après le cours de probabilité, $E(U) = 0$ et $Var(U) = 1$. On en déduit immédiatement : $E(U^2) = Var(U) + E(U)^2 = 1$.

La fonction $t^3 f_U(t)$ étant impaire, il vient immédiatement :

$$E(U^3) = \int_{-\infty}^{+\infty} t^3 f_U(t)dt = 0$$

Procédons maintenant à une intégration par partie pour calculer $E(U^4)$ en posant $f(t) = t^3$ et $g'(t) = te^{-t^2/2}$.

$$\begin{aligned} E(U^4) &= \int_{-\infty}^{+\infty} t^4 f_U(t)dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t^4 e^{-t^2/2} dt \\ &= \frac{1}{\sqrt{2\pi}} \left[-t^3 e^{-t^2/2}\right]_{-\infty}^{+\infty} + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} 3\mathrm{t}^2 e^{-t^2/2} dt \\ &= 0 + 3E(U^2) = 3 \end{aligned}$$

3) Soit n variables aléatoires centrées réduites U_i indépendantes. On définit la variable aléatoire suivante :

$$\chi_n^2 = \sum_{i=1}^{n} U_i^2$$

On dira que χ_n^2 suit une loi du chi-deux à n degrés de liberté. Cette loi nous sera utile dans la suite du cours pour calculer des intervalles de confiance.

Calculer l'espérance et la variance de χ_n^2. Dessiner schématiquement l'allure de la densité de χ_n^2 pour plusieurs valeurs de n .

Réponse :

Sachant que $E(U_i) = 1$, on a immédiatement $E[\chi^2(n)] = n$. Maintenant :

$$\begin{aligned} E[\chi^4(n)] &= E\left[\left(\sum U_i^2\right)^2\right] = E\left[\left(\sum U_i^4\right)\right] + E\left(\sum_{i \neq j} U_i^2\, U_j^2\right) \\ &= \sum E\,(U_i^4) + \sum_{i \neq j} E\,(U_i^2) E(U_j^2) \\ &= 3n + n(n-1) = n(n+2) \end{aligned}$$

Puis :

$$\begin{aligned} Var[\chi^2(n)] &= E[\chi^4(n)] - [E(\chi^2(n))]^2 \\ &= n(n+2) - n^2 = 2n \end{aligned}$$

4) Par quelle loi peut-on approcher la loi du chi-deux quand n est grand ?

Réponse : D'après le théorème central limite, nous avons

$$\frac{\chi^2(n) - n}{\sqrt{2n}}$$

qui converge en loi vers la loi normale centrée réduite quand n tend vers l'infini. On peut donc approcher $\chi^2(n)$ par une loi normale $\mathcal{N}(n, \sqrt{2n})$.

Exercice 71 :

On fabrique des paquets de sucre de 1 kilo. On a prélevé un échantillon de 25 paquets et la moyenne empirique obtenue est 0.992. On modélise le poids d'un paquet par une loi normale d'espérance notée μ et de variance notée σ^2 .

On suppose dans un premier temps que l'écart-type du process est connu et vaut $\sigma = 0.02$.

1) Donner un intervalle de confiance de l'espérance μ de niveau de confiance 0.95.

Réponse :

L'intervalle de confiance à $1-\alpha$ pour μ est $[\bar{X} - u\frac{\sigma}{\sqrt{n}}, \bar{X} + u\frac{\sigma}{\sqrt{n}}]$ où u est tel que $P(-u \leq U \leq u) = 1-\alpha$ avec U suivant une loi $\mathcal{N}(0,1)$.

Nous avons $\bar{x} = 0.992$, $1-\alpha = 95\%$ et $u = 1.96$. L'intervalle de confiance est [0.984,1].

2) On ne suppose plus que l'écart-type du process est connu mais à partir des données de l'échantillon, on a calculé une variance empirique (non biaisé) $s_X^{2'} = 0.029^2$. Donner un intervalle de confiance de l'espérance μ de niveau de confiance 0.95 .

Réponse :

L'intervalle de confiance à $1-\alpha$ pour μ est $[\overline{X} - \frac{S_X'}{\sqrt{n}}t, \overline{X} + \frac{S_X'}{\sqrt{n}}t]$. La valeur de t_α est donnée par $P(-t \leq T \leq t) = 1-\alpha$ avec T suivant une loi de Student de paramètre $(n-1)$.

Nous avons $1-\alpha = 0.95$ et $t = 2.064$. L'intervalle de confiance est $[0.980, 1.004]$. On remarque que l'amplitude de l'intervalle est plus importante que lorsque l'écart-type est connu.

3) Proposer maintenant un intervalle de confiance symétrique de la variance σ^2 de niveau de confiance 0.95.

Réponse : L'intervalle de confiance à $1-\alpha$ pour σ^2 est $\left[\frac{(n-1)S_X^{2'}}{v}, \frac{(n-1)S_X^{2'}}{\mathrm{u}}\right]$ où u et v sont choisis tels que $P(\chi^2_{n-1} \leq u) = \frac{\alpha}{2}$ et $P(\chi^2_{n-1} \leq v) = 1 - \frac{\alpha}{2}$

Ici, $1-\alpha = 0.95$ et nous obtenons $u = 12.4$ et $v = 39.4$. L'intervalle de confiance pour σ^2 est $[0.0005, 0.00163]$.

4) Peut-on utiliser les résultats précédents pour une loi qui n'est pas normale ?

Réponse : Tous les résultats précédents sont valables pour une loi quelconque si la taille de l'échantillon est grande (grâce au théorème central limite). Usuellement, on dira que l'échantillon est suffisamment grand si $n \geq 30$. L'approximation est d'autant meilleure que la taille de l'échantillon est grande.

Exercice 72 : Intervalle de confiance pour un sondage

Un sondage sur la popularité du Premier Ministre indique que 51 % des personnes interrogées sont favorables à sa politique. Soit p la proportion de français favorables à sa politique

1) Donner l'estimateur usuel de p.

Fréquence empirique : $F = \frac{1}{n}\sum_{i=1}^{n} X_i$ où $X_i = 1$ si le i-ème individu est favorable à la politique, 0 sinon.

2) Construire un intervalle de confiance pour p de niveau de confiance 95 %, sachant que le sondage a été réalisé auprès de 100 personnes. Même question si le sondage a été réalisé auprès de 1000 personnes.

Réponse : L'intervalle de confiance utilisé pour une proportion p est

$$[F - u\sqrt{F(1-F)\,/\,n}, F + u\sqrt{F(1-F)\,/\,n}]$$

où z_α est tel que $P(-u \leq U \leq u) = 1 - \alpha$ avec U suivant une loi $\mathcal{N}(0,1)$.

On obtient :

$$\begin{aligned} n = 100 : &\quad IC = [41.2\%, 60.8\%] \\ n = 1000 : &\quad IC = [47.9\%, 54.1\%] \end{aligned}$$

3) Quelle aurait dû être la taille de l'échantillon pour que l'intervalle soit de longueur inférieure à 4 %, en supposant que le sondage donne 51% de personnes favorables ? Reprendre le calcul si les sondages avaient donné une côte de popularité de 90 % .

Réponse : Soit A l'amplitude de l'intervalle. Supposons que $f = 0.51$. On cherche n tel que

$2u\sqrt{f(1-f)\,/\,n} \leq A \Rightarrow n \geq 4f(1-f)\left(\frac{u}{A}\right)^2 = 2399.95$

Il faudrait donc d'interroger 2400 personnes.

En reprenant les calculs avec $f = 0.9$, on obtient $n \geq 865$.

4) Supposons que l'on ne dispose pas d'information sur la valeur estimée de la proportion p. Comment déterminer la taille de l'échantillon pour garantir un niveau de confiance $1 - \alpha$ de longueur inférieure à 4% ?

Réponse : Dans le calcul que nous venons de faire, nous avons supposé que f était connu. Si f n'est pas connu, on peut utiliser le fait que la fonction $f(1-f)$ atteint son maximum $\frac{1}{4}$ en $f = 0.5$ (sur l'intervalle [0,1]). L'intervalle $[F - u\,/\,2\sqrt{n}, F + u\,/\,2\sqrt{n}]$ est alors un intervalle de confiance de niveau de confiance supérieur à $1 - \alpha$.

Cherchons n tel que

$$2u\frac{1}{\sqrt{4n}} \leq 4\% \Rightarrow n \geq \left(\frac{u}{A}\right)^2 = 2400.9$$

La faible différence entre les 2 résultats vient du fait que la fréquence observée est proche de 50% .

Exercice 73 :

Dans une grande ville, une étude vieille de 5 ans a évalué la proportion de fumeurs à 46 %. La municipalité a lancé une campagne contre le tabagisme et désire évaluer son efficacité en interrogeant un échantillon aléatoire de n individus. On note p la proportion de fumeurs dans la population et W la variable aléatoire représentant la proportion de fumeurs dans l'échantillon. Est considéré fumeur dans cette étude, toute personne fumant au moins une fois par semaine.

1) Quelle est la distribution exacte de W (sans utiliser d'approximation) ?

Solution : nW suit une loi binômiale $B(n,p)$ donc pour k entier non nul,

$$P\left(W=\frac{k}{n}\right)=C_n^k p^k(1-p)^{n-k}$$

2) Démontrer que W est un estimateur sans biais de p.

Solution : nW suit une loi binômiale donc $E(nW)=np$ puis $E(W)=p$ et W est un estimateur sans biais de p.

3) Exprimer la variance de W en fonction de p et n.

Solution : nW suit une loi binômiale donc $V[nW]=np(1-p)$ qui implique que $V[W]=\frac{p(1-p)}{n}$ qui tend vers 0 quand n tend vers l'infini. Ainsi W est un estimateur sans biais de p.

4) Par quelle loi peut-on approcher W pour $n \geq 30$? Justifier cette approximation.

Solution : En utilisant le théorème central limite, nous pouvons approcher la loi de W par une loi normale d'espérance p et de variance $\frac{p(1-p)}{n}$.

5) Dans un échantillon de 100 individus, on a dénombré 50 fumeurs. Donner un intervalle de confiance pour p de niveau de confiance 95 %.

Solution : L'intervalle de confiance pour p est $\left[W-u\sqrt{\frac{W(1-W)}{n}}, W+u\sqrt{\frac{W(1-W)}{n}}\right]$ avec u tel que $P(U \leq u)=97.5\%$. En utilisant les tables de loi normales, $u=1.96$. L'estimation de p vaut ici $w=0.5$. Nous obtenons donc comme intervalle de confiance

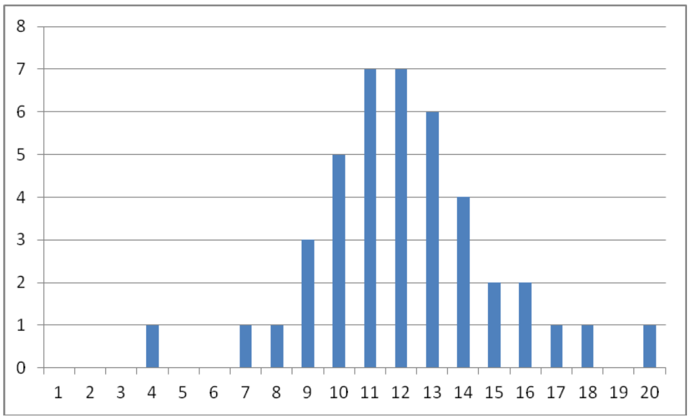
$$\left[0.5-1.96\sqrt{\frac{0.25}{100}}, W+1.96\sqrt{\frac{0.25}{100}}\right]=[0.402,0.598]$$

6) Quelle serait la taille minimale de l'échantillon à prélever pour diviser par 2 l'amplitude de l'intervalle de confiance ?

Solution : L'amplitude de l'intervalle de confiance vaut $2u\sqrt{\frac{W(1-W)}{n}}$. Pour diviser par 2 cette amplitude, il faut multiplier n par 4. Il faudrait donc un échantillon de taille 400.

Exercice 74 :

Les notes de 42 étudiants se répartissent selon l'histogramme suivant :



Les étudiants étant classés par ordre alphabétique, on note (n_1, ..., n_{42}) l'échantillon représenté par l'histogramme ci-dessus. Au vu du profil de cet histogramme, on fait l'hypothèse que l'échantillon est gaussien : ceci revient à considérer que $\forall\ i \in [\![1, 42]\!]$, n_i est la réalisation d'une variable aléatoire N_i, et que les variables aléatoires N_i sont indépendantes entre elles et identiquement distribuées suivant une loi normale de paramètres μ et σ^2.

1°) Calculer la moyenne arithmétique $\bar{n}$ et l'écart-type non biaisé s'_N de cet échantillon.

Solution :

$$\bar{n} = \frac{4+7+8+3\times 9+5\times 10+7\times 11+7\times 12+6\times 13+4\times 14+2\times 15+2\times 16+17+18+20}{42} = \frac{508}{42} \approx 12{,}09.$$

$$s'^2_N = \frac{42}{41} s_N^2 = \frac{42}{41}\left[\frac{1}{42}\left(\sum_{i=1}^{42} n_i^2\right) - \bar{n}^2\right] = \frac{1}{41}\left(\sum_{i=1}^{42} n_i^2\right) - \frac{42}{41}\bar{n}^2.$$

$\sum_{i=1}^{42} n_i^2 = 16 + 49 + 64 + 3 \times 81 + 5 \times 100 + 7 \times 121 + 7 \times 144 + 6 \times 169 + 4 \times 196 + 2 \times 225$
$+ 2 \times 256 + 289 + 324 + 400 = 6500$, d'où

$$s'^2_N = \frac{42}{41}\left[\frac{6500}{42} - \left(\frac{508}{42}\right)^2\right] \approx 8{,}67 \Rightarrow s'_N \approx 2{,}945.$$

2°) Déterminer, au niveau de confiance de 99%, un intervalle de confiance pour μ.

Solution :

L'échantillon est gaussien, mais le paramètre σ étant inconnu, l'intervalle de confiance est

$$\left[\bar{n} + t_{41\,;\,\alpha/2}\frac{s'_N}{\sqrt{n}}\ ;\bar{n} + t_{41\,;\,1-\alpha/2}\frac{s'_N}{\sqrt{n}}\right]$$

$$\approx \left[12{,}09 + t_{41\,;\,0{,}005}\frac{2{,}945}{\sqrt{42}}\ ;12{,}09 + t_{41\,;\,0{,}995}\frac{2{,}945}{\sqrt{42}}\right]$$

$$\approx \left[12{,}09 + 2{,}7\ \times\frac{2{,}945}{\sqrt{42}}\ ;12{,}09 + 2{,}7\ \times\frac{2{,}945}{\sqrt{42}}\right]$$

≈ [12,09 – 1,23 ; 12,09 + 1,23] = [10,86 ; 13,32].

3°) Déterminer, au niveau de confiance de 99%, un intervalle de confiance pour σ.

Solution :

L'échantillon est gaussien, mais le paramètre μ étant inconnu, l'intervalle de confiance sur $s'_N{}^2$ est $\left[\frac{41\ s'_N{}^2}{\chi^2_{41\,;\,1-\alpha/2}}\ ;\ \frac{41\ s'_N{}^2}{\chi^2_{41\,;\,\alpha/2}}\right] = \left[\frac{41\ s'_N{}^2}{\chi^2_{41\,;\,0{,}995}}\ ;\ \frac{41\ s'_N{}^2}{\chi^2_{41\,;\,0{,}005}}\right] \approx \left[\frac{41 \times 8{,}67}{\chi^2_{41\,;\,0{,}995}}\ ;\ \frac{41 \times 8{,}67}{\chi^2_{41\,;\,0{,}005}}\right]$.

On a : $\chi^2_{40\,;\,0{,}99} \approx 63{,}69$ et $\chi^2_{40\,;\,0{,}999} \approx 73{,}40$, donc par interpolation linéaire (par défaut, en l'absence d'informations plus précises), $\chi^2_{40\,;\,0{,}995} \approx 63{,}69 + \frac{5}{9} \times (73{,}40 - 63{,}69) \approx 69{,}1$.

De la même manière, on obtient : $\chi^2_{45\,;\,0{,}995} \approx 69{,}96 + \frac{5}{9} \times (80{,}08 - 69{,}96) \approx 75{,}6$.

Finalement, une dernière interpolation linéaire nous donne

$\chi^2_{41\,;\,0{,}995} \approx 69{,}1 + \frac{1}{5} \times (75{,}6 - 69{,}1) = 70{,}4$.

Les mêmes calculs pour $\chi^2_{41\,;\,0{,}005}$ donnent :

$\chi^2_{40\,;\,0{,}005} \approx 17{,}92 + \frac{4}{9} \times (22{,}16 - 17{,}92) \approx 19{,}8$;

$\chi^2_{45\,;\,0{,}005} \approx 21{,}25 + \frac{4}{9} \times (25{,}90 - 21{,}25) \approx 23{,}3$;

$\chi^2_{41\,;\,0{,}005} \approx 19{,}8 + \frac{1}{5} \times (23{,}3 - 19{,}8) = 20{,}5$.

Finalement, l'intervalle de confiance sur $s'_N{}^2$ est :

$\left[\frac{41 \times 8{,}67}{\chi^2_{41\,;\,0{,}995}}\ ;\ \frac{41 \times 8{,}67}{\chi^2_{41\,;\,0{,}005}}\right] \approx \left[\frac{41 \times 8{,}67}{70{,}4}\ ;\ \frac{41 \times 8{,}67}{20{,}5}\right] \approx [\ 5{,}05\ ;\ 17{,}34\ ]$,

et l'intervalle de confiance sur s'_N est donc $[\sqrt{5{,}05}\ ;\ \sqrt{17{,}34}\ ] \approx [\ 2{,}25\ ;\ 4{,}16\ ]$.

4°) Déterminer, au niveau de confiance de 99%, un intervalle de confiance sur la proportion d'étudiants ayant une note supérieure ou égale à 10.

Solution :

Le nombre d'étudiants ayant une note supérieure ou égale à 10 est égal à 42 – 6 = 36, donc la proportion d'étudiants ayant une note supérieure ou égale à 10 est $f = \frac{36}{42} = \frac{6}{7} \approx 0{,}857$.

Un intervalle de confiance sur la proportion d'étudiants ayant une note supérieure ou égale à 10 est $\left[f + u_{\alpha/2}\sqrt{\frac{f(1-f)}{42}}\ ; f + u_{1-\alpha/2}\sqrt{\frac{f(1-f)}{42}}\right]$

$\approx \left[0{,}857\ - 2{,}5758\ \sqrt{\frac{0{,}857 \times 0{,}143}{42}}\ ; 0{,}857\ + 2{,}5758\ \sqrt{\frac{0{,}857 \times 0{,}143}{42}}\right]$

≈ [0,857 – 0,139 ; 0,857 + 0,139] = [0,718 ; 0,996].

5°) Pour $N \sim \mathcal{N}(\bar{n}, s'_N)$, la probabilité $\mathbf{P}(N \geq 10)$ appartient-elle à l'intervalle de confiance déterminé à la question précédente ?

<u>Solution :</u>

$\mathbf{P}(N \geq 10) = P\left(\frac{N-\bar{n}}{s'_N} \geq \frac{10-\bar{n}}{s'_N} = \frac{10-12{,}09}{2{,}945} \approx -0{,}71\right) = \Phi(0{,}71) \approx 0{,}7611.$

Donc, si $N \sim \mathcal{N}(\bar{n}, s'_N)$, $\mathbf{P}(N \geq 10)$ appartient bien à l'intervalle de confiance sur la proportion d'étudiants ayant une note supérieure ou égale à 10.

Exercice 75 :

Soit $(X_1, ..., X_n)$ un échantillon aléatoire tel que : $\forall\ i \in [\![1, n]\!]$, $E(X_i) = \mu$ et $V(X_i) = \sigma^2$.

1°) On considère $n = 3$ et une réalisation de cet échantillon (x_1, x_2, x_3) telle que $\bar{x} = 1$.

a) Peut-on, pour une loi quelconque, donner une estimation par intervalle de confiance de μ ? Justifiez votre réponse.

<u>Solution :</u>

On ne peut pas faire d'estimation de μ par intervalle de confiance parce que la loi des X_i étant inconnue, celle de $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ est également inconnue.

b) On se place dans le cas d'un échantillon gaussien et on suppose que l'on a $\sigma = 1$.

a. Donner un intervalle de confiance I_1 pour μ au niveau de confiance de 95%.

<u>Solution :</u>

$I_1 = \left[\bar{x} - u_{0,975}\frac{\sigma}{\sqrt{n}}\ ;\ \bar{x} + u_{0,975}\frac{\sigma}{\sqrt{n}}\right] = \left[1 - 1{,}96\frac{1}{\sqrt{3}}\ ;\ 1 + 1{,}96\frac{1}{\sqrt{3}}\right] = [-0{,}132\ ;\ 2{,}132].$

b. Donner un intervalle de confiance I_2 pour μ au niveau de confiance de 99%. Comparez I_1 et I_2 et justifiez vos observations.

<u>Solution :</u>

$I_2 = \left[\bar{x} - u_{0,995}\frac{\sigma}{\sqrt{n}}\ ;\ \bar{x} + u_{0,995}\frac{\sigma}{\sqrt{n}}\right] = \left[1 - 2{,}5758\frac{1}{\sqrt{3}}\ ;\ 1 + 2{,}5758\frac{1}{\sqrt{3}}\right] = [-0{,}487\ ;\ 2{,}487].$

I_2 est plus large que I_1 car, plus le niveau de confiance est élevé, plus l'intervalle est large (car quand l'ordre du fractile tend vers 1, le fractile tend vers l'infini).

c) On suppose maintenant que l'on ne connaît pas σ et que son estimation biaisée vaut $s_X \approx 0{,}82$. Donner un intervalle de confiance I_3 pour μ au niveau de confiance de 95%. Comparer I_1 et I_3 et justifiez vos observations.

<u>Solution :</u>

$I_3 = \left[\bar{x} - t_{2;0,975}\frac{s_X}{\sqrt{n-1}}\ ;\ \bar{x} + t_{2;0,975}\frac{s_X}{\sqrt{n-1}}\right] \approx \left[1 - 4{,}3\frac{0{,}82}{\sqrt{2}}\ ;\ 1 + 4{,}3\frac{0{,}82}{\sqrt{2}}\right] = [-1{,}493\ ;\ 3{,}493].$

I_3 est plus large que I_1 car on ne connaît pas l'écart-type et on est obligé de le remplacer par son estimation. L'élargissement de l'intervalle n'est pas lié ici à la valeur de l'estimation, car on peut remarquer que comme $s_X \approx 0{,}82$, alors $s'_X \approx 1$, donc $\frac{s_X}{\sqrt{2}} = \frac{0{,}82}{\sqrt{2}} \approx \frac{s'_X}{\sqrt{3}} = \frac{1}{\sqrt{3}}$. L'élargissement de l'intervalle est lié à la perte d'information qui fait que l'on remplace une valeur théorique certaine par une estimation, dont on connait la distribution théorique, certes, mais dont la distribution a une forte variance (comme le montre la grande valeur du fractile, 4,3) parce que l'estimation est effectué sur un très petit nombre d'individus.

2°) On considère $n = 60$ une réalisation de cet échantillon $(x_1, ..., x_{60})$ telle que $\bar{x} = 1$.

a) Peut-on, pour une loi quelconque, donner une estimation par intervalle de confiance de μ ? Justifiez votre réponse.

Solution :

Oui, on peut, car comme $n > 30$, le théorème central limite s'applique et permet de considérer que la moyenne arithmétique $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ est approximativement distribuée suivant une loi gaussienne, ce qui permet le calcul d'un intervalle de confiance.

b) On se place dans le cas d'un échantillon gaussien.

a. On suppose que l'on a σ = 1. Donner un intervalle de confiance I_4 pour μ au niveau de confiance de 95%. Comparer I_1 et I_4 et justifiez vos observations.

Solution :

$I_4 = \left[\bar{x} - u_{0,975}\frac{\sigma}{\sqrt{n}} \; ; \; \bar{x} + u_{0,975}\frac{\sigma}{\sqrt{n}}\right] = \left[1 - 1{,}96\frac{1}{\sqrt{60}} \; ; \; 1 + 1{,}96\frac{1}{\sqrt{60}}\right] = [\, 0{,}747 \; ; \; 1{,}253 \,]$.

L'intervalle I_4 est beaucoup plus étroit que I_1. Toutes les autres valeurs étant égales, ceci se justifie uniquement par le passage de n de 3 à 60.

b. On suppose maintenant que l'on ne connaît pas σ et que son estimation non biaisée vaut $s'_X \approx 1{,}00$. Donner un intervalle de confiance I_5 pour μ au niveau de confiance de 95%. Comparer I_4 et I_5 et justifiez vos observations.

Solution :

$I_5 = \left[\bar{x} - t_{59;0,975}\frac{s'_X}{\sqrt{n}} \; ; \; \bar{x} + t_{59;0,975}\frac{s'_X}{\sqrt{n}}\right] \approx \left[1 - 2{,}006\frac{1}{\sqrt{60}} \; ; \; 1 + 2{,}006\frac{1}{\sqrt{60}}\right] = [\, 0{,}741 \; ; \; 1{,}259 \,]$.

L'intervalle I_5 est légèrement plus large que I_4. La raison en est encore que l'on ne connaît pas l'écart-type et que l'on est donc obligé de le remplacer par son estimation. L'élargissement de l'intervalle n'est pas lié ici à la valeur de l'estimation, car la valeur de l'estimation non biaisée est approximativement égale à la valeur théorique, 1, et on continue à diviser par $\sqrt{n} = \sqrt{60}$. L'élargissement de l'intervalle est donc ici aussi lié à la perte d'information qui fait que l'on remplace une valeur théorique certaine par une estimation, mais l'élargissement reste modéré car comme l'échantillon est relativement grand, le fractile de la loi de la Student (2,006) est relativement prroche du fractile de la loi gaussienne (1,96). La valeur du fractile pour la loi de Student à 59 degrés de liberté est trouvé dans la table en interpolant les valeurs des fractiles correspondants des lois de Student à 40 et à 80 degrés de liberté.

c) On se place dans le cas d'un échantillon non gaussien.

a. On suppose que l'on a σ = 1. Donner un intervalle de confiance I_6 pour μ au niveau de confiance de 95%. Comparer I_4 et I_6 et justifiez vos observations.

Solution :

$I_6 \approx \left[\bar{x} - u_{0,975}\frac{\sigma}{\sqrt{n}}\ ;\ \bar{x} + u_{0,975}\frac{\sigma}{\sqrt{n}}\right] = \left[1 - 1{,}96\frac{1}{\sqrt{60}}\ ;\ 1 + 1{,}96\frac{1}{\sqrt{60}}\right]$ = [0,747 ; 1,253].

On peut dans ce cas appliquer le théorème central limite, et on obtient donc, de manière approximative le même intervalle que I_4.

b. On suppose maintenant que l'on ne connaît pas σ et que son estimation non biaisée vaut $s'_X \approx 1{,}00$. Donner un intervalle de confiance I_7 pour μ au niveau de confiance de 95%. Comparer I_5 et I_7 et justifiez vos observations.

Solution :

$I_7 \approx \left[\bar{x} - t_{59;0,975}\frac{s'_X}{\sqrt{n}}\ ;\ \bar{x} + t_{59;0,975}\frac{s'_X}{\sqrt{n}}\right] \approx \left[1 - 2{,}006\frac{1}{\sqrt{60}}\ ;\ 1 + 2{,}006\frac{1}{\sqrt{60}}\right]$ = [0,741 ; 1,259].

On a ici encore égalité approximative entre les intervalles du cas gaussien (I_5) et du cas non gaussien (I_7) car, d'une part, le théorème central limite s'applique, et d'autre part, les correctifs qui sont appliqués dans le cas gaussien lorsqu'on ne connaît pas la valeur théorique de σ (en remplaçant une loi normale par une loi de Student) peuvent également être appliqués sur les distributions approximatives.

3°) On considère n = 600 une réalisation de cet échantillon (x_1, ..., x_{600}) telle que $\bar{x}$ = 1.

a) Peut-on, pour une loi quelconque, donner une estimation par intervalle de confiance de μ ? Justifiez votre réponse.

Solution :

Comme n > 30, la loi des grands nombre s'applique et l'on peut faire une estimation ponctuelle. Par ailleurs, comme l'échantillon est gaussien, on peut toujours faire une estimation par intervalle de confiance.

b) On se place dans le cas d'un échantillon non gaussien.

a. On suppose que l'on a σ = 1. Donner un intervalle de confiance I_8 pour μ au niveau de confiance de 95%.

Solution :

On applique encore le théorème central limite :

$I_8 \approx \left[\bar{x} - u_{0,975}\frac{\sigma}{\sqrt{n}}\ ;\ \bar{x} + u_{0,975}\frac{\sigma}{\sqrt{n}}\right] = \left[1 - 1{,}96\frac{1}{\sqrt{600}}\ ;\ 1 + 1{,}96\frac{1}{\sqrt{600}}\right]$ = [0,920 ; 1,080].

b. On suppose maintenant que l'on ne connaît pas σ et que son estimation non biaisée vaut $s'_X \approx 1{,}00$. Donner un intervalle de confiance I_9 pour μ au niveau de confiance de 95%. Comparer I_8 et I_9 et justifiez vos observations.

Solution :

$I_9 \approx \left[\bar{x} - t_{599;0,975} \frac{s'_X}{\sqrt{n}} \ ; \ \bar{x} + t_{599;0,975} \frac{s'_X}{\sqrt{n}}\right] \approx \left[\bar{x} - u_{0,975} \frac{\sigma}{\sqrt{n}} \ ; \ \bar{x} + u_{0,975} \frac{\sigma}{\sqrt{n}}\right]$

$= \left[1 - 1{,}96 \frac{1}{\sqrt{600}} \ ; \ 1 + 1{,}96 \frac{1}{\sqrt{600}}\right]$ = [0,920 ; 1,080].

On a $I_8 \approx I_9$, car pour n = 600, le fait de remplacer σ par une estimation n'a quasiment aucun effet puisqu'une loi de Student à 599 degrés de liberté est indiscernable d'une loi gaussienne.