Exercices : Probabilités, modélisation et statistique

Polytech Montpellier – DevOps3 – Semestre 6

Année 2025–2026

Raphaël Benerradi

Contenu pédagogique : Gwladys Toulemonde, Chloé Serre-Combe et Raphaël Benerradi

Chapitre 1: Probabilité : Introduction et rappels

Exercice 1.1: Soit A, B et C trois événements. Exprimez en fonction de A, B et C les événements suivants :

a) A et B ont lieu mais pas C
b) Un de ces événements et un seul a lieu.
c) A seul a lieu.
d) Au moins un de ces événements a lieu.
e) Exactement deux de ces événements ont lieu.
f) Aucun de ces événements n'a lieu.
g) Au moins deux de ces événements ont lieu.
h) Pas plus de deux de ces événements n'ont lieu.

Exercice 1.2: On jette trois dés. Calculez :

a) la probabilité d'avoir les trois faces avec le même chiffre.
b) la probabilité d'obtenir au moins un 6.
c) la probabilité d'obtenir au moins deux faces avec le même chiffre.

Exercice 1.3: Dans un restaurant, le menu offre un choix de quatre entrées, trois plats principaux et cinq desserts.

a) Combien y a-t-il de déjeuners complets possibles?
b) Si l'on suppose maintenant la possibilité de ne pas prendre d'entrée, ou de plat principal ou de dessert, combien y a-t-il de déjeuners possibles?

Exercice 1.4: On assimile une partie de poker au tirage sans remise de 5 cartes parmi 32.

a) Calculez la probabilité d'avoir au moins un coeur.
b) Calculez la probabilité de tirer un carré d'as.
c) Calculez la probabilité de tirer une quinte floche (5 cartes consécutives de la même couleur).
d) Calculez la probabilité de tirer une couleur (5 cartes non consécutives de la même couleur).

On rappelle que l'on entend par couleur non pas noir ou rouge mais trèfle, carreau, pique ou coeur.

Exercice 1.5: Dans une loterie, un joueur doit choisir 8 nombres entre 1 et 40. Le tirage sélectionne 8 numéros parmi ces 40 nombres. En admettant que le tirage est équiprobable pour les C_{40}^{8} combinaisons, quelle est la probabilité que le joueur ait

a) les 8 bons numéros?

b) 7 numéros parmi les 8 bons?

c) au moins 6 numéros parmi les 8 bons?

Exercice 1.6: On considère 24 paires de chaussures différentes (soit 48 chaussures!) dont 14 pour homme et 10 pour femme. Les pointures associées à ces 24 paires de chaussures vont du 37 au 44 comme indiqué dans le tableau ci-dessous:

	37	38	39	40	41	42	43	44	Total
Paires de chaussures homme	0	0	1	2	2	4	3	2	14
Paires de chaussures femme	1	3	2	2	2	0	0	0	10

Toutes ces chaussures ne sont pas rangées et se trouvent dans une grande caisse.

a) Combien peut-on construire de paires de chaussures? (il s'agit de donner le nombre total de couples de chaussures possibles comportant une chaussure gauche et une chaussure droite)

Dans la suite, je tire 6 chaussures de la caisse au hasard.

b) Quelle est la probabilité de ne tirer aucune chaussure femme?

c) Quelle est la probabilité de tirer exactement une chaussure femme?

d) Quelle est la probabilité de tirer au moins deux chaussures femme?

Exercice 1.7: Baptiste a u chemises unies et r chemises à rayures. Chaque matin il met une chemise choisie au hasard (équiprobabilité), puis la met à laver à la fin de la journée (tirage sans remise). Quelle est la probabilité pour qu'il ait choisi successivement deux chemises à rayures suivies d'une chemise unie ?

Exercice 1.8: Lors du dernier week-end d'intégration de Polytech Montpellier, les organisateurs ont demandé aux n participants de se munir d'une tente individuelle. A l'issue de la première soirée, bien arrosée, il n'y a plus aucun(e) élève lucide et chacun, en rentrant se coucher choisit une tente au hasard.
Quelle est la probabilité que toutes les tentes soient occupées par leur propriétaire ?

Exercice 1.9: Dans une usine, deux machines A et B fabriquent des micro-processeurs. Ceux issus de A (resp. B) sont défectueux avec une probabilité de 2% (resp. 4%). La chaîne A (resp. B) produit 300 (resp. 200) micro-processeurs par jour. On choisit au hasard un micro-processeur sur la chaîne de fabrication.

a) Avec quelle probabilité est-il défectueux ?

b) S'il est défectueux, quelle est la probabilité qu'il ait été produit par la chaîne B ?

Exercice 1.10:

a) Quelle est la probabilité pour que les anniversaires de r personnes tombent tous à des dates différentes ? (par souci de simplicité, on suppose que toutes les années sont non bissextiles).

b) En déduire la probabilité pour qu'au moins deux personnes de ce groupe aient leur anniversaire le même jour de l'année.

Exercice 1.11: Dans la région Languedoc-Roussillon, 5% des PME font faillite dans une année. Ce pourcentage est de 1% pour les grandes entreprises. Une entreprise du Languedoc-Roussillon fait faillite. Quelle est la probabilité que cela soit une PME sachant qu'il y a 80% de PME dans la région ?

Exercice 1.12: Une grossesse ectopique a deux fois plus de chance de se développer lorsque la femme enceinte fume que lorsqu'elle est non-fumeuse. Si 32% des femmes en âge de maternité fument, quel pourcentage de femmes ayant une grossesse ectopique sont fumeuses ?

Exercice 1.13: Un fumeur a deux boites d'allumettes, chacune dans une poche différente. Lorsque les boites sont pleines, elles contiennent chacune n allumettes. Sachant que chaque fois qu'il allume une cigarette, il prend une allumette dans une des deux boites choisie au hasard, quelle est la probabilité p_r pour que, lorsqu'il s'aperçoit qu'une des deux boites est vide, l'autre contienne exactement r allumettes ?

Exercice 1.14: Au sein d'une population, on s'intéresse à un échantillon de couples avec 2 enfants. On suppose que le genre de chaque enfant est équiprobable (garçon ou fille avec probabilité 1/2) et indépendant.

a) **Le cas de base :** On choisit au hasard une famille parmi celles dont au moins un enfant est un garçon.
Quelle est la probabilité que l'autre enfant soit un garçon ?

b) **Variante du jour de la semaine :** On choisit au hasard une famille parmi celles dont au moins un enfant est un garçon né un mardi.
Quelle est la probabilité que l'autre enfant soit aussi un garçon ?

c) **Variante du prénom :** On choisit au hasard une famille parmi celles dont au moins un enfant est un garçon prénomé Jacques. On suppose que la probabilité qu'un garçon s'appelle Jacques est de 0.3% et que le nom attribué à un enfant est indépendant de celui de ses frères et sœurs.
Quelle est la probabilité que l'autre enfant soit aussi un garçon ?

d) **Cas général :** On choisit au hasard une famille parmi celles dont au moins un enfant est un garçon vérifiant une propriété $\mathcal{P}$. On suppose que cette propriété a une probabilité p d'être vérifiée, indépendament du genre et entre frères et sœurs.
Quelle est la probabilité que l'autre enfant soit un garçon ?

Indications :
Dans l'expérience où on tire un enfant au hasard au sein de la population, notons G l'événement "l'enfant est un garçon", M l'événement "l'enfant est né un mardi", J l'événement "l'enfant s'appelle Jacques", et P l'événement "l'enfant vérifie la propriété $\mathcal{P}$". On va indicer ces événements par 1 pour ceux qui concernent l'aîné(e), par 2 pour ceux qui concernent l'enfant cadet.

Si vous voulez en savoir plus sur ce paradoxe, vous pouvez regarder la page Paradoxe des deux enfants" sur Wikipédia, ou la vidéo de Science4All sur "3 variantes mindfucks des 2 enfants".

Chapitre 2: Variables aléatoires

1 Variables aléatoires discrètes

Exercice 2.1: Un examen est passé sous la forme d'une liste de six questions à trois choix multiples chacune. Un étudiant répond au hasard à cet examen. La variable aléatoire X compte le nombre de réponses correctes.

a) Déterminer la loi de X.
b) Quelle est la probabilité que l'étudiant obtienne au moins 5 bonnes réponses ?
c) Quels sont l'espérance et l'écart-type de sa note si chaque réponse correcte rapporte un point?
d) Si l'examen ne peut comporter que six questions, quel doit être le nombre minimal de choix multiples pour que la probabilité d'obtenir, en répondant au hasard, une note égale ou supérieure à 1 ne dépasse pas 50% ?

Exercice 2.2: Soit X une variable aléatoire discrète ne prenant que les valeurs $\{1, 2, 3, 4, 5\}$ telle que $P(X = 1) = P(X = 5) = x$, $P(X = 2) = P(X = 4) = 2x$ et $P(X = 3) = 4x$.

a) Déterminez x, $P(X > 2)$ et $P(X^2 \leq 7)$.
b) Calculez $P(X = 4|X \geq 3)$, $P(X = 2|X \geq 3.2)$.
c) Calculez $\mathbb{E}(X)$

Exercice 2.3: On dispose de deux urnes. La première contient 4 boules numérotées 2, 2, 4 et 6. La deuxième contient 2 boules numérotées 1 et 3. On tire au hasard une boule dans chaque urne, et l'on note X la somme des deux numéros obtenus. Déterminer la loi de la variable aléatoire X, son espérance et sa variance.

Exercice 2.4: Une machine à sous fonctionne de la manière suivante : on introduit une pièce de 1 euro et 3 roues tournent; ces roues présentent les dix chiffres de 0 à 9 et chaque roue s'arrête en montrant un chiffre au hasard; si les 3 chiffres sont différents, le joueur perd sa mise, s'il y a un double, le joueur touche 2 euros et s'il y a un triplet, le joueur touche x euros.

a) Déterminer le nombre de possibilités d'obtenir 3 chiffres différents ? exactement 2 chiffres identiques ? 3 chiffres identiques ?
b) Soit X la variable aléatoire associée au gain du joueur. Déterminer la loi de X.
c) Jusqu'à quelle valeur de x le jeu est-il profitable au propriétaire de la machine à sous?

Exercice 2.5: On jette deux fois un dé. Quelles sont les valeurs que peuvent prendre les variables aléatoires suivantes et donner leurs fonctions de répartition associées :

a) le plus grand des deux chiffres obtenus,
b) le plus petit des deux chiffres obtenus.

Indication :
On remarquera que si le maximum des 2 valeurs lues sur les dés est plus petit qu'une valeur k fixée alors les 2 dés ont des valeurs plus petites que la valeur k.

Exercice 2.6: On effectue des essais indépendants de probabilité de succès constante et égale à p avec $0 < p < 1$, jusqu'à obtenir un nombre r, fixé à l'avance, de succès. Soit X_r le nombre total d'essais nécessaires. Déterminez la loi de X_r.

2 Variables aléatoires continues

Exercice 2.7: Soit X une v.a. suivant une loi uniforme sur $[a, b]$. Rappelez la densité, la fonction de répartition, l'espérance et la variance de X.

Exercice 2.8: On suppose que la durée de vie d'un livre d'informatique est une variable aléatoire Y_1 de loi exponentielle de paramètre $\theta = 0.5$. Pour rappel la densité associée est égale à $f_\theta(x) = \theta e^{-\theta x}$ pour $x > 0$. On pose $Z = 2Y_1$.

a) Calculer l'espérance de Y_1. En supposant que la durée de vie s'exprime en années, comment interprétez-vous ce résultat ?
b) Calculer la fonction de répartition de Y_1. En déduire $\mathbb{P}(Y_1 > y)$.
c) Quelle est la loi de Z ?
d) En déduire l'espérance de Z.

De la même manière, la durée de vie d'un livre de gestion est une variable aléatoire Y_2 de loi exponentielle de paramètre $\theta = 1/6$ et la durée de vie d'un livre de mathématiques est également une variable aléatoire notée Y_3 de loi exponentielle de paramètre $\theta = 1/6$. Les trois variables aléatoires Y_1, Y_2 et Y_3 sont supposées indépendantes. On s'interesse à la variable aléatoire $M = min(Y_1, Y_2, Y_3)$.

e) Calculer la fonction de répartition de M soit $\mathbb{P}(M \leq y)$. En déduire la loi de M.

Exercice 2.9: On suppose que le temps de survie d'une savonnette est une variable aléatoire T dont la loi admet une densité de la forme $\lambda^2 t \exp(-\lambda t)$.

a) Donnez le nom de cette loi et retrouvez l'espérance associée.
b) Calculez la fonction de répartition de T.
c) Une longue expérience indique que la moyenne $E(T) = 20$ jours. Calculez le paramètre λ.

Exercice 2.10: Trois personnes A, B et C arrivent à un bureau de poste en même temps. Il y a deux guichets qu'occupent immédiatement A et B. C remplace le premier sorti. On note X, Y et Z les temps d'occupation respectifs des guichet par A, B et C. On suppose que ce sont des v.a. exponentielles indépendantes de paramètre λ.

a) Calculez la loi de $U = \min(X, Y)$.
b) Exprimez en fonction de X, Y et Z l'événement : $\{C \text{ sort le dernier}\}$.

Exercice 2.11:

a) Soit $X \sim N(0, 1)$. Calculez $P(X > 1)$ et $P(X^2 < 3,84)$.
b) Trouvez a et b tels que $P(X < a) = P(|X| < b) = 0,95$.
c) Soit $X \sim N(3, 4)$, calculez $P(2 < X < 5)$.
d) Soit X une v.a. telle que $\ln(X - 2) \sim N(1, 4)$. Calculez $P(2,1 < X < 3)$.

Chapitre 3: Vecteurs aléatoires et indépendance

Exercice 3.1: Soit X et Y deux variables aléatoires Gaussiennes centrées et indépendantes. On connait également la variance de X qui vaut 1 et celle de Y qui vaut 4.

a) Donnez la loi du vecteur aléatoire suivant : $\begin{pmatrix} X \\ Y \end{pmatrix}$.

Donnez précisément son espérance et sa matrice de variance-covariance Σ.

b) Soit le vecteur aléatoire $\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} = \begin{pmatrix} 3X + Y + 4 \\ 2X + 2 \end{pmatrix}$.

Donnez sa loi en précisant son espérance et sa matrice de variance-covariance.

c) Les variables Z_1 et Z_2 sont-elles indépendantes ?

Exercice 3.2: On suppose que la taille (en cm) et le poids (en kg) d'un•e étudiant•e de l'UM2 forment un vecteur aléatoire $\mathbf{Z}$ de loi Gaussienne. En notant X la variable aléatoire associée à la taille et Y celle associée au poids, on a donc :

$$\mathbf{Z} = \begin{pmatrix} X \\ Y \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} , \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix} \right)$$

On suppose que tous les paramètres sont connus. On a $\mu_1 = 163$, $\mu_2 = 57$, $\sigma_1^2 = 9$, $\sigma_{12} = 4$ et $\sigma_2^2 = 4$.

a) Quelle est la loi de X ?

b) Quelle est la probabilité qu'un individu, choisi au hasard dans cette population, mesure plus de 169 cm ?

c) Quelle est la probabilité qu'un individu, choisi au hasard dans cette population, mesure entre 160 cm et 166 cm ?

d) Les variables X et Y sont-elles indépendantes ?

e) Calculez le coefficient de corrélation linéaire entre X et Y ? Quelle est l'intérêt de ce coefficient? Comment interprétez-vous le résultat alors obtenu ?

Chapitre 4: Convergence des variables aléatoires

Exercice 4.1: Soit U une v.a. de paramètre $p \in]0,1[$ définie par $\mathbb{P}(U = 1) = p$ et $\mathbb{P}(U = -1) = 1 - p$. On dit alors que U suit une loi de Rademacher de paramètre p. Soit $U_1, ..., U_n$ une suite i.i.d. de variables aléatoires de variable parente U. On définit $V_n = \Pi_{i=1}^n U_i$.

a) Calculer $\mathbb{E}(V_n)$. Peut-on en déduire la loi de V_n ?

b) Montrez que la suite V_n converge en loi.

Exercice 4.2: Soit $X_1, ..., X_n$ une suite de v.a. i.i.d. suivant une loi uniforme sur $[a, b]$, $a < b$. Montrez que $M_n = \max_{1 \leq i \leq n} X_i$ tend en probabilité vers b et que $m_n = \min_{1 \leq i \leq n} X_i$ tend en probabilité vers a.

Exercice 4.3: On rappelle que pour tout x réel $\left(1 + \frac{x}{n}\right)^n$ tend vers e^x quand n tend vers $+\infty$. Soit $X_1, ..., X_n$ une suite de v.a. i.i.d. suivant la loi exponentielle de paramètre λ. On pose $M_n = \max_{1 \leq i \leq n} X_i$. Montrez que

$$M_n - \frac{\ln n}{\lambda} \xrightarrow{\mathcal{L}} L$$

où L est une variable aléatoire dont la fonction de répartition est $F_L(x) = \exp(-\exp(-\lambda x))$.

Indication :

On montrera la convergence de la fonction de répartition de $M_n - \frac{\ln n}{\lambda}$.

Exercice 4.4: Soit $\{X_n\}$ une suite de variables aléatoires indépendantes et de même loi $\mathcal{B}(p)$ avec $p \in]0,1[$. Pour tout k on pose $Y_k = X_k X_{k+1}$ et $S_n = \sum_{k=1}^n Y_k$.

a) Déterminer la loi de Y_k. Peut-on directement appliquer à S_n la LGN vue en cours ?

b) Montrez que $V(S_n) = np^2(1 - p^2) + 2(n - 1)p^3(1 - p)$.

c) Par une inégalité vue dans le cours, montrez enfin que pour tout $\varepsilon > 0$,

$$\mathbb{P}\left(\left|\frac{S_n}{n} - p^2\right| > \varepsilon\right) \xrightarrow[n \to \infty]{} 0.$$

Exercice 4.5:

a) Rappelez la loi de $X_n = \sum_{i=1}^{n} Y_i$ avec $Y_1, \ldots, Y_n$ des variables aléatoires indépendantes suivant une loi de Bernouilli de paramètre p, $p \in]0,1[$.

b) Montrez à l'aide du Théorème Central Limite que $\frac{X_n - np}{\sqrt{np(1-p)}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1)$.

Ce résultat permet alors d'obtenir une approximation de la loi $\mathcal{B}(n, p)$ quand np et $n(1-p)$ sont supérieurs à 10.

Chapitre 5: Statistique inférentielle

Exercice 5.1: Selon l'Insee (2010), 43.3% des travailleurs en France sont des femmes. Cette même année, une enquête a été réalisée sur un échantillon de $n = 10000$ femmes parmi lesquelles 3465 ont déclaré être cadres. Dans la suite, on notera $T_1, \ldots, T_n$ l'échantillon des n variables aléatoires représentant les réponses (Cadre ? : oui - 1 /non - 0) des n femmes intérrogées.

a) Donnez précisément l'expression de l'estimateur $\hat{p}_n$ de la proportion inconnue p de cadres chez les femmes en 2010 (bien définir les quantités impliquées).
b) Détaillez l'obtention de l'espérance et de la variance de $\hat{p}_n$.
c) Citez les propriétés connues de $\hat{p}_n$.
d) Donnez une estimation de la proportion inconnue p de cadres chez les femmes en 2010.
e) Proposez un intervalle de confiance pour p à 95%.
f) Proposez un intervalle de confiance pour p à 80%. Commentez la différence entre l'intervalle obtenu à cette question et celui obtenu à la question précédente.

Exercice 5.2: Les parties 1 et 2 de cet exercice sont indépendantes.
Partie 1 : Une usine utilise une machine pour remplir automatiquement des flacons contenant une épice. Par suite de variations aléatoires dans le mécanisme, le poids de poudre par flacon est une variable aléatoire de loi normale de moyenne m et d'écart-type $1,1$ g. Les flacons sont vendus comme contenant 100 g de produit.

a) La machine est réglée sur $m = 101,2$ g. Quelle est la probabilité que le poids de produit dans un flacon soit
   i. inférieur au poids annoncé de 100 g ?
   ii. supérieur à 105 g ?
b) Sur quelle valeur de m faut-il régler la machine pour que moins de 4% des flacons aient un poids inférieur au poids annoncé de 100 g ?

Partie 2 : Les sulfites ont des propriétés anti-oxydantes et sont donc parfois utilisés comme conservateur. En Europe, la déclaration de la présence de sulfites en tant qu'allergène dans les aliments est obligatoire dès que leur concentration atteint 1 mg pour 100 grammes. Comme il est impossible de tester la contenance en sulfites de tous les flacons, l'idée est d'en ouvrir quelques uns (ici $n = 10$), d'y mesurer le taux de sulfites et d'utiliser vos connaissances en statistiques. Vous obtenez les résultats suivants:

0.83 ; 1.31 ; 1.25 ; 0.42 ; 0.64 ; 1.15 ; 1.04 ; 0.35 ; 0.46 ; 0.68.

On admet que le taux de sulfites dans le flacon i est une variable aléatoire X_i qui suit une loi normale de moyenne inconnue μ et et d'écart-type 0.35 mg et que les observations sont indépendantes.

a) Calculez la médiane associée à cette série et interprétez-la.
b) Quelle est la loi de la moyenne empirique $\overline{X}_n$? (vous indiquerez les paramètres associés à cette loi). Quelles sont les propriétés de cette quantité ?
c) Construire un intervalle de confiance pour μ de la forme $]-\infty; a]$ au niveau 95%, c'est-à-dire déterminer a tel que $P(\mu > a) = 0.05$.
   Indication: On pourra considérer $a = \overline{X}_n + t$ et se contenter de déterminer t.
d) Pensez-vous qu'il faut faire figurer la présence de sulfites sur les flacons d'épices?

e) Si on n'avait pas eu connaissance de la valeur de l'écart-type ($\sigma = 0.35$), expliquez comment vous auriez procédé et faites-le.

Exercice 5.3: Soit $X_1, ..., X_n$ un échantillon de v.a. i.i.d. suivant la loi uniforme sur $[0, \theta]$.

On veut estimer θ par $S = \frac{2}{n} \sum_{i=1}^{n} X_i$. Expliquez pourquoi S a quelques chances d'être un estimateur pertinent de θ. Cet estimateur converge-t-il vers θ? Donnez son biais et sa variance.