

Corrigé : Épreuves indépendantes et loi binomiale

Fiche Défi

D1.

1. Une issue contenant exactement k succès est déterminée par les k places occupées par la lettre S parmi les n places : c'est le choix d'une partie à k éléments d'un ensemble à n éléments, il y en a $\binom{n}{k}$. Les épreuves étant indépendantes, la probabilité d'une issue est le produit des probabilités de ses lettres, donc $p^k q^{n-k}$ pour une issue à k succès et $n - k$ échecs. Ces issues sont deux à deux incompatibles, d'où

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

2. La variable X prend ses valeurs dans $\{0; 1; \dots; n\}$: les événements $(X = 0), \dots, (X = n)$ sont deux à deux incompatibles et leur réunion est l'univers. Donc $\sum_{k=0}^n P(X = k) = \sum_{k=0}^n \binom{n}{k} p^k q^{n-k} = 1$, pour toute valeur de n et de p .
3. Pour $1 \leq k \leq n$:

$$k \binom{n}{k} = k \frac{n!}{k! (n-k)!} = \frac{n!}{(k-1)! (n-k)!} = n \frac{(n-1)!}{(k-1)! ((n-1) - (k-1))!} = n \binom{n-1}{k-1}.$$

Le terme $k = 0$ de $E(X)$ est nul. Avec le changement d'indice $j = k - 1$, la somme obtenue est la somme des probabilités de la loi $\mathcal{B}(n-1; p)$, égale à 1 d'après la question 2 :

$$\begin{aligned} E(X) &= \sum_{k=1}^n k \binom{n}{k} p^k q^{n-k} = np \sum_{k=1}^n \binom{n-1}{k-1} p^{k-1} q^{(n-1)-(k-1)} \\ &= np \sum_{j=0}^{n-1} \binom{n-1}{j} p^j q^{n-1-j} = np \times 1 = np. \end{aligned}$$

4. Pour $2 \leq k \leq n$:

$$k(k-1) \binom{n}{k} = \frac{n!}{(k-2)! (n-k)!} = n(n-1) \frac{(n-2)!}{(k-2)! ((n-2) - (k-2))!} = n(n-1) \binom{n-2}{k-2}.$$

Les termes $k = 0$ et $k = 1$ de $E(X(X-1)) = \sum k(k-1)P(X = k)$ sont nuls. Avec $j = k - 2$, la somme obtenue est celle de la loi $\mathcal{B}(n-2; p)$, égale à 1 :

$$E(X(X-1)) = n(n-1)p^2 \sum_{j=0}^{n-2} \binom{n-2}{j} p^j q^{n-2-j} = n(n-1)p^2 \times 1 = n(n-1)p^2.$$

Comme $X^2 = X(X-1) + X$, $E(X^2) = n(n-1)p^2 + np$. Alors

$$V(X) = E(X^2) - E(X)^2 = n(n-1)p^2 + np - n^2p^2 = np - np^2 = np(1-p).$$

Conclusion : pour $X \sim \mathcal{B}(n; p)$, $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$, $E(X) = np$ et $V(X) = np(1-p)$.

D2.

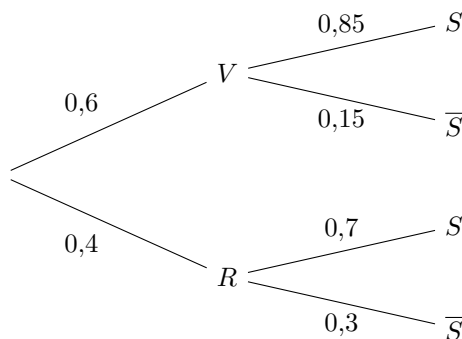
1. L'univers est $\{PP, PF, FP, FF\}$; les lancers sont indépendants et les pièces équilibrées, donc chaque issue a pour probabilité $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$. Alors $A = \{PP, PF\}$, $B = \{PP, FP\}$, $C = \{PP, FF\}$ et $P(A) = P(B) = P(C) = \frac{1}{2}$.
2. $A \cap B = \{PP\}$, $A \cap C = \{PP\}$ et $B \cap C = \{PP\}$ ont chacun pour probabilité $\frac{1}{4} = \frac{1}{2} \times \frac{1}{2}$. Donc A, B sont indépendants, A, C le sont, et B, C le sont.

3. $A \cap B \cap C = \{PP\}$, donc $P(A \cap B \cap C) = \frac{1}{4}$, alors que $P(A)P(B)P(C) = \frac{1}{8}$. Les deux valeurs sont différentes.
4. $P_{A \cap B}(C) = \frac{P(A \cap B \cap C)}{P(A \cap B)} = \frac{1/4}{1/4} = 1 \neq P(C) = \frac{1}{2}$. Savoir que A et B sont réalisés détermine entièrement C : l'indépendance deux à deux ne suffit pas pour que C soit indépendant de l'information « A et B réalisés ».
5. Les trois lancers sont indépendants : $P(P_1 \cap P_2 \cap P_3) = \left(\frac{1}{2}\right)^3 = \frac{1}{8} = P(P_1)P(P_2)P(P_3)$. Ici chaque événement P_i ne dépend que de la i -ème épreuve, et c'est l'indépendance des *épreuves* qui donne la formule du produit. L'événement C de la question précédente dépend, lui, des deux épreuves à la fois.

Conclusion : trois événements peuvent être indépendants deux à deux sans que $P(A \cap B \cap C) = P(A)P(B)P(C)$.

D3. Partie A. Les données : $P(V) = 0,6$, $P(R) = 0,4$, $P_V(S) = 0,85$, $P_R(S) = 0,7$.

1. L'arbre pondéré :



2. (V, R) est une partition de l'univers. D'après la formule des probabilités totales : $P(S) = 0,6 \times 0,85 + 0,4 \times 0,7 = 0,51 + 0,28 = 0,79$.
3. $P_S(V) = \frac{P(V \cap S)}{P(S)} = \frac{0,51}{0,79} = \frac{51}{79} \approx 0,646$.
4. $P(R \cap \bar{S}) = 0,4 \times 0,3 = 0,12$ et $P(\bar{S}) = 1 - 0,79 = 0,21$. Donc $P_{\bar{S}}(R) = \frac{0,12}{0,21} = \frac{4}{7} \approx 0,571$.

Partie B.

5. On répète 25 fois, de façon identique et indépendante (tirages assimilés à des tirages avec remise), une épreuve de Bernoulli de succès « l'élève réussit le quiz », de probabilité $P(S) = 0,79$. Donc X suit $\mathcal{B}(25; 0,79)$.
6. $P(X = 25) = 0,79^{25} \approx 0,00276 \approx 0,003$. Pour $P(X \geq 20) = 1 - P(X \leq 19)$, la calculatrice (fonction de répartition) donne $P(X \leq 19) \approx 0,4325$, d'où $P(X \geq 20) \approx 0,5675 \approx 0,568$.
7. $E(X) = 25 \times 0,79 = 19,75$ et $\sigma(X) = \sqrt{25 \times 0,79 \times 0,21} = \sqrt{4,1475} \approx 2,04$. Interprétation : sur un grand nombre de groupes de 25 élèves, on compte en moyenne 19,75 réussites par groupe.
8. Avec n élèves, le nombre de réussites suit $\mathcal{B}(n; 0,79)$ et « au moins un échec » est le contraire de « n réussites » :

$$1 - 0,79^n \geq 0,99 \iff 0,79^n \leq 0,01 \iff n \ln 0,79 \leq \ln 0,01 \iff n \geq \frac{\ln 0,01}{\ln 0,79} \approx 19,54.$$

Le sens est renversé car $\ln 0,79 < 0$. Il faut $n = 20$ (vérification : $1 - 0,79^{19} \approx 0,9887 < 0,99$ et $1 - 0,79^{20} \approx 0,9910 \geq 0,99$).

9. `def seuil_echec(q):`
`n = 1`
`while 1 - 0.79**n < q:`
`n = n + 1`
`return n`

seuil_echec(0.99) renvoie 20.

Conclusion : un élève ayant réussi a suivi le cours en direct avec la probabilité $\frac{51}{79}$; parmi 25 élèves, il y a au moins 20 réussites avec la probabilité d'environ 0,568 ; il faut interroger 20 élèves pour observer au moins un échec avec une probabilité d'au moins 0,99.

D4.

1. X_n suit $\mathcal{B}(n; 0,1)$ et $u_n = 1 - P(X_n = 0) - P(X_n = 1) = 1 - 0,9^n - n \times 0,1 \times 0,9^{n-1}$.
2. Si $X_n \geq 2$, alors $X_{n+1} \geq 2$. Si $X_n = 1$, alors $X_{n+1} \geq 2$ si et seulement si le $(n+1)$ -ème composant est anormal. Si $X_n = 0$, alors $X_{n+1} \leq 1$. Donc « $X_{n+1} \geq 2$ » est la réunion des deux événements « $X_n \geq 2$ » et « $X_n = 1$ et le $(n+1)$ -ème composant est anormal », qui sont incompatibles. Le $(n+1)$ -ème composant étant indépendant des n premiers :

$$u_{n+1} = u_n + P(X_n = 1) \times 0,1.$$

Comme $P(X_n = 1) \geq 0$, $u_{n+1} \geq u_n$: la suite (u_n) est croissante.

3. n = 1

```
u = 1 - 0.9**n - n * 0.1 * 0.9**(n - 1)
while u < 0.9:
    n = n + 1
    u = 1 - 0.9**n - n * 0.1 * 0.9**(n - 1)
print(n)
```

4. Le script s'arrête au premier rang n pour lequel $u_n \geq 0,9$. Grâce à la croissance de (u_n) , on a alors $u_m \geq 0,9$ pour tout $m \geq n$: tester au moins n composants suffit, et tester moins ne suffit pas. Le script renvoie $n = 38$, avec $u_{37} \approx 0,8964 < 0,9$ et $u_{38} \approx 0,9047 \geq 0,9$ (on admet que u_n tend vers 1, ce qui garantit que la boucle se termine).

Conclusion : il faut tester au moins 38 composants pour déclencher l'analyse avec une probabilité d'au moins 0,9.

D5.

1. Prolongeons fictivement la série jusqu'à $2m-1$ manches. Si A gagne la série, il a atteint m victoires, et ce nombre ne diminue pas : il a au moins m victoires sur les $2m-1$ manches. Réciproquement, si A a au moins m victoires, B en a au plus $m-1$, donc B n'a jamais atteint m victoires, alors que A l'a fait : A est le vainqueur de la série. Les $2m-1$ manches sont identiques et indépendantes, de probabilité de succès p pour A : le nombre Y de victoires de A suit $\mathcal{B}(2m-1; p)$ et $P_m(p) = P(Y \geq m)$.
2. Pour $m = 2$, $Y \sim \mathcal{B}(3; p)$: $P_2(p) = P(Y = 2) + P(Y = 3) = 3p^2(1-p) + p^3 = 3p^2 - 2p^3$. Alors

$$P_2(p) - p = 3p^2 - 2p^3 - p = -p(2p^2 - 3p + 1) = -p(2p-1)(p-1) = p(1-p)(2p-1).$$

Comme $p(1-p) > 0$, $P_2(p) > p$ si et seulement si $p > \frac{1}{2}$: une série avantage le meilleur joueur (et désavantage le plus faible).

3. Pour $m = 3$, $Y \sim \mathcal{B}(5; p)$: $P_3(p) = 10p^3(1-p)^2 + 5p^4(1-p) + p^5$. Pour $p = \frac{1}{2}$: $P_3(\frac{1}{2}) = \frac{10 + 5 + 1}{32} = \frac{1}{2}$. Deux joueurs de même niveau ont autant de chances de gagner la série.
4. $P_1(0,6) = 0,6$, $P_2(0,6) = 3 \times 0,36 - 2 \times 0,216 = 0,648$ et $P_3(0,6) = 10 \times 0,216 \times 0,16 + 5 \times 0,1296 \times 0,4 + 0,07776 = 0,3456 + 0,2592 + 0,07776 = 0,68256$. Plus la série est longue, plus le meilleur joueur est favorisé.

5. from math import comb

```
def P(m, p):
    return sum(comb(2*m - 1, k) * p**k * (1 - p)**(2*m - 1 - k)
               for k in range(m, 2*m))
```

```
m = 1
```

```
while P(m, 0.6) < 0.9:
```

```
m = m + 1
print(m)
```

Le script renvoie $m = 21$ (série de 41 manches au plus), avec $P_{20}(0,6) \approx 0,8979 < 0,9$ et $P_{21}(0,6) \approx 0,9035 \geq 0,9$.

Conclusion : $P_m(p) = P(Y \geq m)$ avec $Y \sim \mathcal{B}(2m - 1; p)$; pour $p = 0,6$, il faut une série au meilleur de 41 manches pour que le meilleur joueur gagne avec la probabilité d'au moins 0,9.

D6.

- On a $X_{n+1} = X_n + Y$, où Y vaut 1 si la $(n + 1)$ -ième épreuve est un succès et 0 sinon, Y étant indépendante des n premières épreuves. Le nombre X_{n+1} est pair dans deux cas incompatibles : la dernière épreuve est un échec (probabilité q) et X_n est pair (probabilité a_n), ou la dernière épreuve est un succès (probabilité p) et X_n est impair (probabilité $1 - a_n$). Par indépendance,

$$a_{n+1} = q a_n + p (1 - a_n) = (q - p) a_n + p = (1 - 2p) a_n + p.$$

- Notons $r = 1 - 2p$ et $\mathcal{P}(n)$: « $a_n = \frac{1 + r^n}{2}$ ».

Initialisation : $a_0 = 1$ (zéro succès est un nombre pair) et $\frac{1 + r^0}{2} = 1$.

Hérédité : on suppose $\mathcal{P}(n)$. Alors

$$a_{n+1} = r \times \frac{1 + r^n}{2} + p = \frac{r + r^{n+1}}{2} + \frac{2p}{2} = \frac{(1 - 2p) + 2p + r^{n+1}}{2} = \frac{1 + r^{n+1}}{2},$$

donc $\mathcal{P}(n + 1)$ est vraie. Par récurrence, $a_n = \frac{1 + (1 - 2p)^n}{2}$ pour tout $n \geq 0$.

- Pour $n = 1$: X_1 est pair si et seulement si $X_1 = 0$, de probabilité $1 - p$, et $\frac{1 + (1 - 2p)}{2} = 1 - p$. Pour $n = 2$: $P(X_2 \text{ pair}) = P(X_2 = 0) + P(X_2 = 2) = (1 - p)^2 + p^2 = 1 - 2p + 2p^2$ et $\frac{1 + (1 - 2p)^2}{2} = \frac{2 - 4p + 4p^2}{2} = 1 - 2p + 2p^2$.

- Avec $n = 10$ et $p = 0,1$: $a_{10} = \frac{1 + 0,8^{10}}{2} \approx \frac{1 + 0,10737}{2} \approx 0,554$.

- Si $p = \frac{1}{2}$, $1 - 2p = 0$ et la probabilité vaut $\frac{1}{2}$ pour tout $n \geq 1$: pair et impair sont équiprobables.

Si $0 < p < 1$, alors $|1 - 2p| < 1$, donc $(1 - 2p)^n \rightarrow 0$ et $a_n \rightarrow \frac{1}{2}$: pour un grand nombre d'épreuves, la parité du nombre de succès est quasiment équiprobable.

Conclusion : $P(X_n \text{ pair}) = \frac{1 + (1 - 2p)^n}{2}$, soit environ 0,554 pour $\mathcal{B}(10; 0,1)$.

D7.

- Les 4 lancers sont indépendants. Le contraire de « au moins un 6 » est « aucun 6 » : $P = 1 - \left(\frac{5}{6}\right)^4 = 1 - \frac{625}{1296} = \frac{671}{1296} \approx 0,518$.

- Un double six a la probabilité $\frac{1}{36}$ par lancer. Le contraire de « au moins un double six » est « aucun double six en 24 lancers » : $P = 1 - \left(\frac{35}{36}\right)^{24} \approx 1 - 0,5086 = 0,491$. Seul le premier pari est favorable ($0,518 > \frac{1}{2} > 0,491$).

- $1 - \left(\frac{35}{36}\right)^n \geq \frac{1}{2} \iff \left(\frac{35}{36}\right)^n \leq \frac{1}{2} \iff n \ln \frac{35}{36} \leq -\ln 2 \iff n \geq \frac{\ln 2}{\ln \frac{36}{35}} \approx 24,61$. Il faut $n = 25$ lancers (on a $1 - \left(\frac{35}{36}\right)^{25} \approx 0,5055$).

- n = 1

```
while 1 - (35/36)**n < 0.5:
```

```

n = n + 1
print(n)

```

Le script affiche 25.

Conclusion : le pari à un dé et 4 lancers est favorable ($\approx 0,518$), celui à deux dés et 24 lancers ne l'est pas ($\approx 0,491$) ; il faut 25 lancers de deux dés pour que le second le devienne.

D8.

1. Chaque passager se présente ou non (épreuve de Bernoulli de paramètre 0,9), indépendamment des autres : $X_n \sim \mathcal{B}(n; 0,9)$. Pour $n = 105$: $E(X_{105}) = 105 \times 0,9 = 94,5$ et $\sigma(X_{105}) = \sqrt{105 \times 0,9 \times 0,1} = \sqrt{9,45} \approx 3,07$.
2. `p_sup(n, p, k)` renvoie $P(X > k) = \sum_{i=k+1}^n P(X = i)$ pour $X \sim \mathcal{B}(n; p)$. Avec $k = 100$, c'est la probabilité qu'au moins un passager soit refusé : environ 1,7 % pour 105 billets vendus et environ 4,0 % pour 106 billets. Le risque augmente avec le nombre de billets vendus.
3. La boucle doit continuer tant que le risque est trop grand : `while p_sup(n, p, k) > alpha:`. Elle renvoie le premier k pour lequel $P(X > k) \leq \alpha$. Pour $n = 105$: $P(X > 98) \approx 0,0899 > 0,05$ et $P(X > 99) \approx 0,0424 \leq 0,05$, donc la fonction renvoie 99. (La boucle se termine toujours car $P(X > n) = 0 \leq \alpha$.)
4. Si $X_n \leq 100$, aucun refus et $G_n = 120n$; sinon $X_n - 100$ passagers sont refusés et $G_n = 120n - 500(X_n - 100)$. Donc $G_n = 120n - 500 \max(X_n - 100; 0)$ et, par définition de l'espérance, $E(G_n) = 120n - 500 \sum_{k=101}^n (k - 100)P(X_n = k)$ (les termes $k \leq 100$ sont nuls).
5. Le script cherche, parmi $n = 100, \dots, 115$, la valeur qui maximise le gain moyen : la compagnie maximise son gain moyen en vendant $n = 109$ billets, pour un gain moyen d'environ 12 837,08 euros. Pour $n = 100$ aucun refus n'est possible et $E(G_{100}) = 120 \times 100 = 12\,000$ euros : la surréservation augmente le gain moyen d'environ 837 euros. Toutefois, avec 109 billets, la probabilité d'au moins un refus est d'environ 22,7 % : la compagnie peut préférer une valeur plus prudente (par exemple 105 ou 106 billets, avec un risque inférieur à 4 %) pour préserver la satisfaction des clients, au prix d'un gain moyen plus faible (12 588,26 euros pour 105 billets, 12 689,20 euros pour 106).