

DS de Méthodes Statistiques en Informatique

DS-Corrigé

Exercice .1 *Test de grippe*

Soit les événements

- G : *personne atteinte de la grippe*
- P : *personne pour laquelle le test est positif*

L'énoncé nous indique que :

$$\begin{aligned}\Pr(G) &= \frac{1}{10000} \\ \Pr(P|G) &= 0.99 \\ \Pr(P|\bar{G}) &= 0.001\end{aligned}$$

Nous recherchons $\Pr(G|P)$:

$$\begin{aligned}\Pr(G|P) &= \frac{\Pr(G \wedge P)}{\Pr(P)} \\ \Pr(G|P) &= \frac{\Pr(P|G)\Pr(G)}{\Pr(P)}\end{aligned}$$

Or

$$\begin{aligned}\Pr(P) &= \Pr(P \wedge G) + \Pr(P \wedge \bar{G}) \\ \Pr(P) &= \Pr(P|G)\Pr(G) + \Pr(P|\bar{G})\Pr(\bar{G}) \\ \Pr(P) &= 0.99 \times 0.0001 + 0.001 \times (1 - 0.0001) \\ \Pr(P) &= 0.0010989\end{aligned}$$

Donc, nous pouvons conclure :

$$\begin{aligned}\Pr(G|P) &= \frac{\Pr(P|G)\Pr(G)}{\Pr(P)} \\ \Pr(G|P) &= \frac{0.99 \times 0.0001}{0.0010989} \\ \Pr(G|P) &\approx 0.09009 \\ \Pr(\bar{G}|P) &\approx 1 - 0.09009 = 0.90991\end{aligned}$$

Il y a donc 9% de chance qu'une personne dont le test est positif soit effectivement atteinte du virus de la grippe. On peut donc en déduire qu'il y a 91% de chances qu'une personne dont le test est positif ne soit pas en réalité atteinte par le virus.

Exercice .2 *Compilation*

1. Si on suppose que le programmeur a compilé au moins une fois, on a $X \in \mathbb{N}^*$
2. Il s'agit d'une épreuve avec une suite de réalisation conduisant à un échec et qui s'arrête lorsque il y a réussite, c'est donc une loi géométrique de paramètre $(1-p)$. On a donc

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

3. On cherche $P(X = n + k | X \geq n)$. Il s'agit donc d'une probabilité conditionnelle. On a

$$P(X = n + k | X \geq n) = \frac{P(X = n + k \text{ et } X \geq n)}{P(X \geq n)}$$

De plus

$$P(X = n + k \text{ et } X \geq n) = P(X = n + k)$$

et

$$P(X \geq n) = \sum_{i=n}^{\infty} (1 - p)^{i-1} p$$

$$P(X \geq n) = 1 - \sum_{i=1}^{n-1} (1 - p)^{i-1} p$$

$$P(X \geq n) = (1 - p)^{n-1}$$

Donc

$$P(X = n + k | X \geq n) = (1 - p)^k p.$$

4. On cherche $P((X \geq n \text{ et } X \leq n + k) | X \geq n)$. On a

$$(X \geq n \text{ et } X \leq n + k) = (X = n) \text{ ou } (X = n + 1) \text{ ou } \dots (X = n + k - 1) \text{ ou } (X = n + k)$$

Remarquons que pour tout i, j dans $[0..k]$, $i \neq j$, $(X=n+i)$ et $(X=n+j)$ s'excluent. Par suite, en utilisant la distributivité du "et" sur le "ou"

$$P((X \geq n \text{ et } X \leq n + k) | X \geq n) = \frac{\sum_{i=0}^k P(X = n + i \text{ et } X \geq n)}{P(X \geq n)}$$

$$P((X \geq n \text{ et } X \leq n + k) | X \geq n) = \sum_{i=0}^k \frac{P(X = n + i \text{ et } X \geq n)}{P(X \geq n)}$$

$$P((X \geq n \text{ et } X \leq n + k) | X \geq n) = \sum_{i=0}^k P(X = n + i | X \geq n)$$

En utilisant la question précédente on a donc

$$P((X \geq n \text{ et } X \leq n + k) | X \geq n) = \sum_{i=0}^k (1 - p)^i p$$

et

$$P((X \geq n \text{ et } X \leq n + k) | X \geq n) = 1 - (1 - p)^{k+1}.$$

Exercice .3 *Lois binomiales*

$$G_X(z) = \sum_{k=0}^{\infty} \Pr(X = k) \cdot z^k. \quad (1)$$

0- Soit X suivant une loi binomiale $(2, p)$.

$$\begin{aligned} G_X(z) &= C_2^0 p^0 z^0 (1-p)^2 + C_2^1 p^1 z^1 (1-p)^1 + C_2^2 p^2 z^2 (1-p)^0 \\ &= (pz + 1 - p)^2 \end{aligned}$$

1- Soit X suivant une loi binomiale (n, p) .

$$\begin{aligned} G_X(z) &= \sum_{k=0}^n C_n^k p^k (1-p)^{n-k} \cdot z^k \\ &= \sum_{k=0}^n C_n^k (pz)^k (1-p)^{n-k} \\ &= (pz + 1 - p)^n \text{ (d'après la formule du binôme)}. \end{aligned}$$

2- Supposons que $G_X(z) = (pz + 1 - p)^n$. Soit X' une v.a. suivant une loi binomiale (n, p) . Par la question 1, $\forall z \in \mathbb{C}, G_X(z) = G_{X'}(z)$. Les fonctions $G_X, G_{X'}$ sont des fonctions entières (i.e. définies par des séries entières, voir (1)) qui coïncident sur un voisinage de 0 (à savoir \mathbb{C} tout entier). Donc les coefficients de ces séries entières sont les mêmes, i.e.

$$\forall k \in \mathbb{N}, \Pr(X = k) = \Pr'(X' = k).$$

Mais par hypothèse X' suit une loi binomiale (n, p) . Donc X suit une loi binomiale (n, p) .

3- Soient X, Y deux v.a. sur Ω , indépendantes. Remarquons que, pour tout v.a. U à valeurs entières, $G_U(z) = E(z^U)$. Donc

$$G_{X+Y}(z) = E(z^{X+Y}) = E(z^X \cdot z^Y).$$

Comme X, Y sont indépendantes, z^X, z^Y sont des v.a., à valeurs complexes, qui sont aussi indépendantes. Donc

$$E(z^X \cdot z^Y) = E(z^X) \cdot E(z^Y) = G_X(z) \cdot G_Y(z).$$

Finalement on a bien $G_{X+Y}(z) = G_X(z) \cdot G_Y(z)$.

4-

4.1- Par la question 3, $G_{X+Y}(z) = G_X(z) \cdot G_Y(z)$ et en substituant les formules obtenues à la question 1, on obtient :

$$G_{X+Y}(z) = (pz + 1 - p)^n \cdot (pz + 1 - p)^m = (pz + 1 - p)^{n+m}.$$

4.2- Par la question 2 on conclut que $X + Y$ suit une loi binomiale $(n + m, p)$.