

---

## TD9 – Révisions

---

### Exercice 1.

Deux évènements  $A$  et  $B$  sont indépendants. De plus, ils vérifient  $P(B|A \cup B) = 2/3$ , et  $P(A|B) = 1/2$ . Que vaut  $P(B)$  ?

**Solution:** Traduisons, sans probabilités conditionnelles, les 3 assertions :

$$\begin{aligned} A \text{ et } B \text{ indépendants} &\Leftrightarrow P(A \cap B) = P(A)P(B) \\ P(B|A \cup B) = 2/3 &\Leftrightarrow P(A \cup B) = \frac{3}{2}P(B) \\ P(A|B) = 1/2 &\Leftrightarrow P(A \cap B) = \frac{1}{2}P(B) \end{aligned}$$

Des affirmations (1) et (3), on déduit que  $P(A) = P(A|B) = 1/2$  (indépendance). Enfin, la loi générarale pour l'union de deux ensembles nous dit que :

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ \frac{3}{2}P(B) &= \frac{1}{2} + P(B) - \frac{1}{2}P(B) \\ P(B) &= \frac{1}{2}. \end{aligned}$$

### Exercice 2.

Un fabricant produit des vis “premier prix”. Chaque vis produite possède une probabilité  $p = 5\%$  d'être défectueuse. Les vis sont conditionnées par paquets de 500. On note  $X$  le nombre de vis défectueuses dans un paquet.

1. Quelle est la loi de probabilité de  $X$  ?

**Solution:**  $X$  suit une loi binomiale de paramètres  $N = 500$ ,  $p = 0.05$ .

2. Justifiez qu'on peut approcher  $X$  par une loi normale, dont vous préciserez l'espérance  $\mu$  et l'écart-type  $\sigma$ .

**Solution:** Comme toute loi binomiale, on peut décrire  $X$  comme une somme de  $N$  tirages de Bernoulli indépendants :

$$X = \sum_{i=1}^{500} Y_i$$

Puisque 500 est un grand nombre,  $X$  est une somme d'un grand nombre de variables indépendantes. En vertu du **théorème central limite**,  $X$  a donc une distribution approximativement normale :

$$X \sim \mathcal{N}(\mu, \sigma)$$

Les nombres  $\mu$  et  $\sigma$ , l'espérance et l'écart-type de  $X$ , sont ceux de la vraie loi de  $X$  (loi binomiale), donc

$$\mu = E(X) = Np = 25$$

et

$$\sigma = \sqrt{V(X)} = \sqrt{Np(1-p)} = \sqrt{23.75} \simeq 4.87$$

3. En utilisant cette approximation, donnez la probabilité que le nombre de vis défectueuses soit compris entre 20 et 30.

**Solution:** On a

$$P(20 \leq X \leq 30) = P\left(\frac{20 - \mu}{\sigma} \leq Y \leq \frac{30 - \mu}{\sigma}\right) = P(-1.026 \leq Y \leq 1.026)$$

pour  $Y$  une loi normale centrée réduite. On peut donc aller lire dans le tableau :

$$\begin{aligned} P(20 \leq X \leq 30) &= \Phi(1.026) - \Phi(-1.026) \\ &= \Phi(1.026) - (1 - \Phi(1.026)) && \text{(par symétrie)} \\ &= 2\Phi(1.026) - 1 \\ &\simeq 2 * 0.847 - 1 \\ &\simeq 0.694 \end{aligned}$$

4. Lorsqu'on achète un paquet de vis, quel est le nombre minimal de vis valides qu'on est garanti d'avoir avec une probabilité de 99% ?

**Solution:** On cherche  $a$  tel que

$$P(X \leq a) = 0.99$$

Autrement dit :

$$P(Y \leq \frac{a - \mu}{\sigma}) = 0.99$$

et donc

$$\frac{a - \mu}{\sigma} = z$$

pour le nombre  $z$  tel que

$$\begin{aligned} \Phi(z) &= 0.99 \\ z &\simeq 2.33 \end{aligned}$$

d'où

$$a = \mu + z.\sigma \simeq 25 + 2.33 \times 4.87 \simeq 36.35$$

Il y a donc 99% de chances d'avoir moins de 37 vis défectueuses. Et donc, un nombre de vis valides garanti dans un paquet de 500 :

$$500 - 37 = 463 \text{ vis valides}$$

### Exercice 3.

Un sismographe (appareil de mesure de secousses sismiques) surveille l'activité d'un volcan dangereux. Le volcan a deux phases :

- phase active (notée  $A$ ) où on observe en moyenne 0.5 micro-secousses par heure.
- phase inactive (notée  $\bar{A}$ ) où on observe en moyenne 0.1 micro-secousses par heure.

Le volcan passe la majorité de son temps en phase inactive : on sait que  $P(A) = \frac{1}{1000}$ .

Il y a un danger d'éruption uniquement pendant la phase active.

1. Quel est le nombre moyen de micro-secousses attendues sur une journée de 24 heures, lorsque le volcan est en phase active ? Et lorsque le volcan est en phase inactive ?

**Solution:** En phase active, il y a en moyenne 12 secousses par 24 heures. En phase inactive, il y a en moyenne 2.4 secousses par 24 heures.

2. Chaque jour, l'appareil fournit une variable aléatoire  $X$  qui représente le *nombre de micro-secousses enregistrées sur les dernières 24 heures*.

(a) Un statisticien nous suggère de modéliser  $X$  par une loi de Poisson. Pourquoi ?

(b) En phase active, on modélise  $X$  par une loi de Poisson de paramètre  $\lambda$  :

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

Quelle valeur doit-on donner à  $\lambda$  ?

(c) En phase inactive, on modélise  $X$  par une loi de Poisson de paramètre  $\mu$  :

$$P(X = k | \bar{A}) = e^{-\mu} \frac{\mu^k}{k!}$$

Quelle valeur doit-on donner à  $\mu$  ?

**Solution:**

(a) Le fait d'utiliser une loi de Poisson implique que les micro-secousses surviennent de manière indépendante les unes des autres — autrement dit, sans aucun 'pattern' particulier entre les temps des différentes secousses.

Le seul paramètre existant dans une loi de Poisson est l'espérance de la variable, c'est-à-dire, le *nombre moyen* de secousses attendues sur 24 heures.

(b) On sait que  $\lambda$  représente l'espérance de  $X$  (en phase active), soit donc 12.

(c) De même, en phase inactive, on utilise le paramètre  $\mu = 2.4$ .

3. (a) Calculez la probabilité d'être en phase active *et* d'observer 9 secousses, soit

$$P(A \cap \{X = 9\}).$$

(b) De même, calculez  $P(\bar{A} \cap \{X = 9\})$ .

(c) Déduisez-en  $P(X = 9)$ , et enfin  $P(A | X = 9)$ .

(d) Un jour donné, l'appareil a enregistré 9 secousses. Faut-il s'alarmer ?

**Solution:**

$$(a) P(A \cap \{X = 9\}) = P(A).P(X = 9|A) = \frac{1}{1000}e^{-12} \frac{12^9}{9!} \simeq 8.736 \times 10^{-5}$$

$$(b) P(\bar{A} \cap \{X = 9\}) = P(\bar{A}).P(X = 9|\bar{A}) = \frac{999}{1000}e^{-2.4} \frac{2.4^9}{9!} \simeq 6.598 \times 10^{-4}$$

(c) Par la « loi de la probabilité totale »,

$$P(X = 9) = P(A \cap \{X = 9\}) + P(\bar{A} \cap \{X = 9\}) \simeq 7.471 \times 10^{-4}.$$

Puis, avec la formule de Bayes ‘dans l’autre sens’

$$P(A | X = 9) = \frac{P(A \cap \{X = 9\})}{P(X = 9)} \simeq \frac{2.736 \times 10^{-5}}{7.471 \times 10^{-4}} \simeq 12\%$$

(d) Si un jour donné, l’appareil a enregistré 9 secousses, il y a environ 12% de chances que le volcan soit en phase active. Vu le danger potentiel, c’est un niveau assez important pour être signalé et surveiller avec attention le volcan.

4. Le jour suivant, l’appareil enregistre 8 secousses. Faut-il s’alarmer ?

*Indice :* considérez la variable  $Y = X_1 + X_2$  donnant le compte de micro-secousses sur les dernières 48 heures.

**Solution:** La variable  $Y$  (compte de secousses sur 48 heures) peut elle aussi être modélisée par une loi de Poisson. Les calculs sont donc identiques aux questions précédentes, mais avec des valeurs différentes :

- En phase active, on attend un nombre moyen  $\lambda = 24$  micro-secousses.
- En phase inactive, on attend un nombre moyen  $\mu = 4.8$  micro-secousses.
- Sur les deux derniers jours, on a observé  $k = 9 + 8 = 17$  secousses.

On a donc

$$P(A | Y = 17) = \frac{\frac{1}{1000} e^{-24} \frac{24^{17}}{17!}}{\frac{1}{1000} e^{-24} \frac{24^{17}}{17!} + \frac{999}{1000} e^{-4.8} \frac{4.8^{17}}{17!}} = \frac{e^{-24} 24^{17}}{e^{-24} 24^{17} + 999 e^{-4.8} 4.8^{17}} \simeq 78\%$$

Avec des mots, le score élevé de 8 observé le deuxième jour laisse à penser que le score de 9 observé la veille n’était pas un « accident ». Un tel niveau de micro-secousses observé sur *deux jours consécutifs* rend très plausible le fait que le volcan soit rentré en phase active.

#### Exercice 4.

Un système électronique fait intervenir deux composants de durées de vies respectives  $X$  et  $Y$ , montés en série. Ces deux variables aléatoires suivent des lois exponentielles indépendantes de paramètres respectifs  $\lambda_1$  et  $\lambda_2$ .

1. Expliquer pourquoi la durée de vie  $D$  du système s’exprime par :  $D = \min(X, Y)$ .
2. Soit un nombre  $a > 0$ . Exprimer  $P(D > a)$  à l’aide des variables  $X$  et  $Y$ .
3. En déduire  $F_D(a)$ , la fonction de répartition de  $D$ . Quelle loi suit la variable  $D$  ?
4. Donner l’espérance  $E(D)$ . Comment ce nombre se compare-t-il à  $E(X)$  et  $E(Y)$  ?

#### Solution:

1. Composants montés en série  $\rightarrow$  le système est en panne dès que l’un de deux composants est en panne. Par suite,  $D = \min(X, Y)$ .

2. Pour tout  $a \geq 0$  on a :

$$\begin{aligned}P(D > a) &= P(\min(X, Y) > a) \\&= P(X > a \cap Y > a) \\&= P(X > a) \times P(Y > a)\end{aligned}$$

car  $X$  et  $Y$  sont indépendantes. D'autre part, on a :

$$\begin{aligned}P(X > a) &= 1 - P(X \leq a) \\&= 1 - F_X(a) \\&= 1 - (1 - e^{-\lambda_1 a}) \\&= e^{-\lambda_1 a}\end{aligned}$$

De même,  $P(Y > a) = e^{-\lambda_2 a}$ . Par conséquent,

$$\begin{aligned}P(D > a) &= e^{-\lambda_1 a} \times e^{-\lambda_2 a} \\&= e^{-(\lambda_1 + \lambda_2)a}\end{aligned}$$

3. La fonction de répartition de  $D$  est :

$$\begin{aligned}F_D(a) &= P(D \leq a) \\&= 1 - P(D > a) \\&= 1 - e^{-(\lambda_1 + \lambda_2)a}\end{aligned}$$

Par conséquent,  $D$  suit la loi exponentielle de paramètre  $\lambda_1 + \lambda_2$  :  $D \sim \mathcal{E}(\lambda_1 + \lambda_2)$ .

4. L'espérance de  $D$  est

$$E(D) = \frac{1}{\lambda_1 + \lambda_2}$$

Par comparaison,  $E(X) = \frac{1}{\lambda_1}$ , on a donc  $E(D) < E(X)$ . Ce qui est normal puisque, par définition  $D$  est toujours inférieur (ou égal) à  $X$ . Le même raisonnement tient également pour  $Y$ .

### Exercice 5.

Dans une colonie de marmottes de Sibérie orientale, on note la durée de vie  $X$  d'une marmotte au hasard. On suppose que  $X$  suit une loi exponentielle, de paramètre  $\lambda = 0.5 \text{ an}^{-1}$ .

1. On définit  $Y$  la variable aléatoire donnant la moyenne des durées de vie de 3 marmottes prises au hasard. C'est-à-dire  $Y = (X_1 + X_2 + X_3)/3$ , avec  $X_1, X_2, X_3$  trois variables aléatoires indépendantes suivant une loi exponentielle de paramètre  $\lambda$ . Donnez les valeurs de  $E(Y)$  et  $V(Y)$ .

**Solution:** Déjà, d'après les propriétés de la distribution exponentielle, on a  $E(X) = 1/\lambda = 2 \text{ ans}$ , et  $V(X) = 1/\lambda^2 = 4 \text{ ans}^2$ .

Ensuite, par linéarité de l'espérance,  $E(Y) = (E(X_1) + E(X_2) + E(X_3))/3 = (2 + 2 + 2)/3 = 2 \text{ ans}$ .

Ensuite, comme les variables  $X_1, X_2, X_3$  sont **indépendantes**, un résultat de cours garantit que leurs variances s'additionnent. On a donc

$$V(X_1 + X_2 + X_3) = V(X_1) + V(X_2) + V(X_3) = 4 + 4 + 4 = 12$$

Enfin, une variance représente une valeur moyenne du *carré* de la variable. Donc lorsqu'on multiplie la variable par une constante  $c$ , sa variance est multipliée par  $c^2$ . (Preuve :  $V(cX) = E(c^2X^2) - E(cX)^2 = c^2(E(X^2) - E(X)^2) = c^2V(X)$ .)

On en déduit que

$$V(Y) = V\left(\frac{X_1 + X_2 + X_3}{3}\right) = \frac{V(X_1 + X_2 + X_3)}{9} = \frac{12}{9} = \frac{4}{3}$$

2. On définit  $Z$  la variable aléatoire donnant la moyenne des durées de vie de 100 marmottes prises au hasard. C'est-à-dire  $Z = (X_1 + \dots + X_{100})/100$ , avec  $X_1, \dots, X_{100}$  cent variables aléatoires indépendantes suivant une loi exponentielle de paramètre  $\lambda$ . Donnez les valeurs de  $E(Z)$  et  $V(Z)$ .

**Solution:** Par le même argument mais en remplaçant 3 par 100, on a

$$E(Z) = 2$$

et

$$V(Z) = \frac{V(X)}{100} = \frac{4}{100} = 0.04$$

3. Donnez, en le justifiant, une bonne approximation pour la loi de la variable  $Z$ . Utilisez cette approximation pour calculer  $P(1.8 \leq Z \leq 2.2)$ .

**Solution:**  $Z$  est une somme d'un grand nombre de variables indépendantes. D'après le théorème central limite, la distribution de  $Z$  est donc approximativement une loi normale.

De plus, on a calculé au-dessus l'espérance (2) et l'écart-type ( $\sqrt{0.04} = 0.2$ ) de  $Z$ . Donc

$$Z \sim \mathcal{N}(2, 0.2)$$

En particulier,

$$P(1.8 \leq Z \leq 2.2) = P\left(\frac{1.8 - 2}{0.2} \leq \frac{Z - 2}{0.2} \leq \frac{2.2 - 2}{0.2}\right) = P(-1 \leq T \leq 1)$$

pour  $T$  une loi normale centrée réduite. Il n'y a plus qu'à aller lire le tableau.

4. Soit un nombre entier  $N$ . On note  $W_N$  la variable aléatoire donnant la moyenne des durées de vie de  $N$  marmottes prises au hasard. Quelle valeur minimum doit avoir  $N$  pour qu'on ait  $P(1.99 \leq W_N \leq 2.01) \geq 0.95$  ?

**Solution:** En répétant le même raisonnement, la variable  $W_N$  a toujours une espérance de 2, et une variance de

$$\text{Var}(W_N) = \frac{\text{Var}(X)}{N} = \frac{4}{N}$$

et de plus, dès que  $N$  est suffisamment grand, la distribution de  $W_N$  devient une loi normale. D'après le tableau pour la loi normale, le  $z$  correspondant à un intervalle symétrique à 95% est  $z \simeq 1.96$ . On doit donc résoudre

$$1.96 \times \sqrt{\frac{4}{N}} \leq 0.01$$

ce qui donne  $N \geq 4 \times 196^2 \simeq 160000$ .