

Plaquette 3 - Formule des probabilités totales

Probabilités conditionnelles — Première générale

Corrections détaillées

Boîte à outils

Formule des probabilités totales et inversion des conditionnelles.

- **Partition.** Des événements $A_1, A_2, \dots, A_n$ forment une partition de l'univers s'ils sont deux à deux **incompatibles** et si leur réunion est l'univers entier. Le cas le plus simple est $\{A; \bar{A}\}$.

- **Formule des probabilités totales**, dans le cas de deux événements :

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

- **Cas général**, avec une partition $A_1, \dots, A_n$:

$$P(B) = \sum_{i=1}^n P(A_i) \times P_{A_i}(B)$$

Sur un arbre, cela revient à additionner **tous** les chemins qui aboutissent à B .

- **Inverser une conditionnelle**, en deux étapes : 1. calculer $P(B)$ par la formule des probabilités totales ; 2. appliquer la définition :

$$P_B(A) = \frac{P(A \cap B)}{P(B)} = \frac{P(A) \times P_A(B)}{P(B)}$$

- **Les deux opérations de l'arbre**, à ne jamais confondre : on **multiplie** le long d'un chemin, on **additionne** entre chemins.
- **Le contrôle systématique.** La somme des conditionnelles inversées sur une partition vaut 1 :

$$P_B(A) + P_B(\bar{A}) = 1$$

- **L'erreur de proportion.** Un taux élevé dans un petit groupe peut peser moins qu'un taux faible dans un grand groupe : la formule des probabilités totales **pondère** chaque conditionnelle par la taille de son groupe. C'est tout son intérêt.

Correction de l'exercice 1 - Première application de la formule

- a) On applique la formule du produit à chacun des deux chemins menant à B :

$$P(A \cap B) = P(A) \times P_A(B) = 0,40 \times 0,70 = 0,28$$

$$P(\bar{A} \cap B) = P(\bar{A}) \times P_{\bar{A}}(B) = 0,60 \times 0,25 = 0,15$$

- b) Les deux chemins sont incompatibles et couvrent tous les cas où B se réalise : la formule des probabilités totales donne

$$P(B) = P(A \cap B) + P(\bar{A} \cap B) = 0,28 + 0,15 = 0,43$$

- c) **Première façon**, par l'événement contraire : $P(\bar{B}) = 1 - 0,43 = 0,57$.

Seconde façon, par la formule des probabilités totales appliquée à $\bar{B}$:

$$P(\bar{B}) = 0,40 \times 0,30 + 0,60 \times 0,75 = 0,12 + 0,45 = 0,57$$

Les deux méthodes concordent ✓

Ce que la formule fait vraiment. Elle calcule une **moyenne pondérée** des deux conditionnelles, les poids étant $P(A) = 0,40$ et $P(\bar{A}) = 0,60$. Le résultat 0,43 est bien compris entre 0,25 et 0,70, et plus proche de 0,25 puisque le groupe $\bar{A}$ est le plus gros.

Le contrôle d'encadrement. $P(B)$ est toujours comprise entre la plus petite et la plus grande des conditionnelles. Si le résultat sort de cet encadrement, il y a une erreur.

Réponse. a) 0,28 et 0,15 — b) $P(B) = 0,43$ — c) $P(\bar{B}) = 0,57$ par les deux méthodes ✓

Correction de l'exercice 2 - Le taux de pièces défectueuses

a) On note A et D comme d'habitude. La formule des probabilités totales donne

$$P(D) = P(A) \times P_A(D) + P(\bar{A}) \times P_{\bar{A}}(D)$$

$$P(D) = 0,60 \times 0,02 + 0,40 \times 0,05 = 0,012 + 0,020 = 0,032$$

Le taux global de défauts est donc de 3,2 %.

b) Non. La moyenne arithmétique de 2 % et 5 % vaut $\frac{2+5}{2} = 3,5$ %, alors que le taux réel est 3,2 %. La différence vient de ce que les deux machines ne produisent **pas** la même quantité de pièces : la formule des probabilités totales calcule une moyenne **pondérée** par les parts de production.

$$0,60 \times 2 \% + 0,40 \times 5 \% = 1,2 \% + 2,0 \% = 3,2 \%$$

Le taux réel est tiré vers le bas parce que la machine la plus fiable (A , 2 %) est aussi la plus productive (60 %).

c) Sur 10 000 pièces :

$$10\,000 \times 0,032 = 320 \text{ pièces défectueuses}$$

Contrôle par les effectifs : 6 000 pièces de A donnent 120 défauts, 4 000 pièces de B en donnent 200, soit 320 au total ✓

La règle à retenir. Une moyenne de taux n'a de sens que si les groupes ont le **même effectif**. Sinon il faut pondérer — et c'est exactement ce que fait la formule des probabilités totales.

Réponse. a) $P(D) = 0,032$ — b) non : 3,2 % et non 3,5 %, car la moyenne doit être pondérée par les parts de production — c) 320 pièces.

Correction de l'exercice 3 - Une partition de trois événements

a) Les trois événements sont **incompatibles** (une pièce provient d'un seul fournisseur) et leur réunion couvre toute la production, puisque

$$P(A) + P(B) + P(C) = 0,50 + 0,30 + 0,20 = 1$$

Ils forment donc bien une partition de l'univers, et la formule des probabilités totales s'applique avec trois termes.

b) On additionne les trois chemins menant à D :

$$P(D) = P(A)P_A(D) + P(B)P_B(D) + P(C)P_C(D)$$

$$P(D) = 0,50 \times 0,01 + 0,30 \times 0,04 + 0,20 \times 0,06$$

$$P(D) = 0,005 + 0,012 + 0,012 = 0,029$$

Le taux global de défauts est donc de 2,9 %.

c) On compare les trois contributions :

Fournisseur	A	B	C
Contribution	0,005	0,012	0,012
Part des défauts	17,2 %	41,4 %	41,4 %

Les fournisseurs *B* et *C* contribuent à **égalité**, chacun pour environ 41,4 % des pièces défectueuses, et *A* pour seulement 17,2 %.

Le paradoxe apparent. *C* fournit moins de pièces que *B* (20 % contre 30 %), mais contribue autant aux défauts, car son taux est une fois et demie plus élevé. Et *A*, qui fournit la moitié des pièces, ne pèse que pour un sixième des défauts.

L'usage en gestion de la qualité. C'est exactement le calcul que fait un responsable qualité pour décider où agir : la contribution aux défauts, et non le taux seul, désigne la priorité.

Réponse. a) somme des probabilités égale à 1 et événements incompatibles — b) $P(D) = 0,029$ — c) *B* et *C* à égalité, avec 41,4 % des défauts chacun.

Correction de l'exercice 4 - Inverser une conditionnelle

a) On applique la définition de la conditionnelle, avec le chemin déjà calculé au numérateur :

$$P_D(A) = \frac{P(A \cap D)}{P(D)} = \frac{0,012}{0,032} = 0,375$$

b) De même pour l'autre machine :

$$P_D(\bar{A}) = \frac{P(\bar{A} \cap D)}{P(D)} = \frac{0,020}{0,032} = 0,625$$

Contrôle : $0,375 + 0,625 = 1$ ✓ C'est attendu : une pièce défectueuse vient forcément de l'une des deux machines.

c) $P_D(A) = 0,375$ signifie : **parmi les pièces défectueuses, 37,5 % viennent de la machine A.**

À comparer à $P_A(D) = 0,02$, qui dit que **parmi les pièces de A, 2 % sont défectueuses**. Les deux nombres n'ont rien à voir : le premier porte sur la population des pièces défectueuses, le second sur celle des pièces de A.

La méthode en deux temps, à retenir. Pour inverser une conditionnelle :

$$1) P(D) \text{ par les probabilités totales} \quad 2) P_D(A) = \frac{P(A) \times P_A(D)}{P(D)}$$

La lecture utile pour l'usine. La machine A produit 60 % des pièces mais seulement 37,5 % des défauts : elle est bien la plus fiable des deux. Le contrôle qualité devrait donc se concentrer sur B.

Réponse. a) $P_D(A) = 0,375$ — b) $P_D(\bar{A}) = 0,625$, de somme 1 ✓ — c) 37,5 % des pièces défectueuses viennent de A, à ne pas confondre avec les 2 % de défauts parmi les pièces de A.

Correction de l'exercice 5 - La vraie fiabilité d'un test

a) La formule des probabilités totales, avec la partition $\{M; \bar{M}\}$:

$$P(T) = P(M)P_M(T) + P(\bar{M})P_{\bar{M}}(T)$$

$$P(T) = 0,02 \times 0,95 + 0,98 \times 0,03 = 0,019 + 0,0294 = 0,0484$$

Soit 4,84 % de tests positifs dans la population.

b) On applique la définition de la conditionnelle :

$$P_T(M) = \frac{P(M \cap T)}{P(T)} = \frac{0,019}{0,0484} \approx 0,393$$

c) **Non, absolument pas.** Le 95 % annoncé est $P_M(T)$: parmi les **malades**, 95 % ont un test positif. Mais parmi les **tests positifs**, seulement 39,3 % correspondent à de vrais malades : un test positif est donc **plus souvent faux que vrai**.

L'explication est arithmétique. Les non-malades représentent 98 % de la population : même avec un taux d'erreur de seulement 3 %, ils produisent

$$0,98 \times 0,03 = 0,0294$$

de faux positifs, soit plus que les 0,019 de vrais positifs. Sur 10 000 personnes testées : 190 vrais positifs contre 294 faux positifs.

La conséquence pratique. C'est pourquoi un dépistage positif est toujours suivi d'un **second test** de confirmation, et pourquoi on ne dépiste massivement que les maladies suffisamment fréquentes ou les populations à risque — quand $P(M)$ est plus grand, la proportion de faux positifs s'effondre.

Le contrôle. $P_T(\overline{M}) = \frac{0,0294}{0,0484} \approx 0,607$, et $0,393 + 0,607 = 1 \checkmark$

Réponse. a) $P(T) = 0,0484$ — b) $P_T(M) \approx 0,393$ — c) non : un test positif n'indique un vrai malade que dans 39,3 % des cas, à cause du poids des faux positifs.

Correction de l'exercice 6 - L'effet de la prévalence

a) Avec $P(M) = 0,30$ et $P(\overline{M}) = 0,70$:

$$P(T) = 0,30 \times 0,95 + 0,70 \times 0,03 = 0,285 + 0,021 = 0,306$$

b) On inverse la conditionnelle :

$$P_T(M) = \frac{0,285}{0,306} \approx 0,931$$

c) La comparaison est spectaculaire :

Prévalence $P(M)$	2 %	30 %
$P_T(M)$	0,393	0,931

Avec **le même test**, un résultat positif est exact dans 39,3 % des cas en population générale, mais dans 93,1 % des cas en population à risque. Le test n'a pas changé : c'est la **prévalence** de la maladie qui change tout.

La raison : quand $P(M)$ augmente, les vrais positifs (0,285) deviennent bien plus nombreux que les faux positifs (0,021), alors qu'ils étaient minoritaires avec une prévalence de 2 %.

La conclusion de santé publique. La valeur d'un dépistage dépend autant de la **population testée** que de la qualité du test. C'est la justification mathématique du ciblage des dépistages : tester une population à risque rend le test utile, tester tout le monde le rend peu informatif.

Le contrôle. $P_T(\overline{M}) = \frac{0,021}{0,306} \approx 0,069$, et $0,931 + 0,069 = 1 \checkmark$

Réponse. a) $P(T) = 0,306$ — b) $P_T(M) \approx 0,931$ — c) avec le même test, la fiabilité d'un positif passe de 39,3 % à 93,1 % : c'est la prévalence qui décide.

Correction de l'exercice 7 - Un sondage par tranches d'âge

On note A_1, A_2, A_3 les trois tranches d'âge et F l'événement « la réponse est favorable ».

a) La formule des probabilités totales, avec trois termes :

$$P(F) = 0,30 \times 0,60 + 0,40 \times 0,45 + 0,30 \times 0,30$$

$$P(F) = 0,18 + 0,18 + 0,09 = 0,45$$

Soit 45 % de réponses favorables dans l'ensemble des sondés.

b) On inverse la conditionnelle pour la première tranche :

$$P_{\bar{F}}(A_1) = \frac{P(A_1 \cap F)}{P(F)} = \frac{0,18}{0,45} = 0,40$$

Soit 40 % des personnes favorables qui ont entre 18 et 30 ans — alors que cette tranche ne représente que 30 % des sondés. Elle est donc **surreprésentée** parmi les favorables.

c) Les deux autres inversions :

$$P_{\bar{F}}(A_2) = \frac{0,18}{0,45} = 0,40 \quad P_{\bar{F}}(A_3) = \frac{0,09}{0,45} = 0,20$$

Somme : $0,40 + 0,40 + 0,20 = 1$ ✓ C'est cohérent, puisque toute personne favorable appartient à exactement une tranche d'âge.

La lecture du résultat. Les 18-30 ans passent de 30 % de l'échantillon à 40 % des favorables ; les plus de 50 ans tombent de 30 % à 20 %. Les tranches où le taux d'adhésion est élevé sont mécaniquement surreprésentées parmi les favorables.

La vérification à faire toujours. Sur une partition, la somme des conditionnelles inversées vaut 1. C'est le contrôle qui détecte immédiatement une erreur de division.

Réponse. a) $P(F) = 0,45$ — b) $P_{\bar{F}}(A_1) = 0,40$ — c) $0,40 + 0,40 + 0,20 = 1$ ✓

Correction de l'exercice 8 - Ponctualité de deux lignes

On note B l'événement « le trajet se fait en bus » et P l'événement « le trajet est ponctuel ».

a) La formule des probabilités totales :

$$P(P) = 0,55 \times 0,90 + 0,45 \times 0,97 = 0,495 + 0,4365 = 0,9315$$

Soit 93,15 % de trajets ponctuels.

b) Par l'événement contraire :

$$P(\bar{P}) = 1 - 0,9315 = 0,0685$$

On peut le retrouver directement : $0,55 \times 0,10 + 0,45 \times 0,03 = 0,055 + 0,0135 = 0,0685$ ✓

c) On inverse la conditionnelle, en prenant le chemin « bus et en retard » au numérateur :

$$P_{\bar{P}}(B) = \frac{P(B \cap \bar{P})}{P(\bar{P})} = \frac{0,055}{0,0685} \approx 0,803$$

Soit environ 80,3 % des trajets en retard qui se font en bus, alors que le bus ne représente que 55 % des trajets.

Le renversement à commenter. Le bus est très majoritaire parmi les retards, pour deux raisons qui s'additionnent : il représente la majorité des trajets **et** son taux de retard est plus de trois fois supérieur (10 % contre 3 %). Les deux effets vont dans le même sens, d'où l'écart considérable.

Le contrôle. $P_{\overline{A}}(\overline{B}) = \frac{0,0135}{0,0685} \approx 0,197$, et $0,803 + 0,197 = 1$ ✓

Réponse. a) $P(P) = 0,9315$ — b) $0,0685$ — c) environ 80,3 % des retards se font en bus.

Correction de l'exercice 9 - Un filtre anti-spam

On note S l'événement « le message est un spam » et K « le message contient le mot-clé ».

a) La formule des probabilités totales :

$$P(K) = P(S)P_S(K) + P(\overline{S})P_{\overline{S}}(K)$$

$$P(K) = 0,40 \times 0,85 + 0,60 \times 0,05 = 0,34 + 0,03 = 0,37$$

Soit 37 % des messages qui contiennent le mot-clé.

b) On inverse la conditionnelle :

$$P_K(S) = \frac{P(S \cap K)}{P(K)} = \frac{0,34}{0,37} \approx 0,919$$

Environ 91,9 % des messages contenant le mot-clé sont des spams.

c) Le filtre est **plutôt bon, mais pas sans risque**. Deux angles :

- il attrape 85 % des spams, ce qui est efficace ;
- mais 8,1 % des messages bloqués sont **légitimes** : sur 1 000 messages reçus, 370 seraient bloqués, dont 30 à tort.

Or bloquer un message légitime (un faux positif) est souvent bien plus grave que laisser passer un spam : un courrier important perdu coûte davantage qu'une publicité à supprimer. Un filtre acceptable devrait donc **mettre en quarantaine** plutôt que supprimer, ou combiner plusieurs indices au lieu d'un seul mot-clé.

La comparaison avec le test médical. Le même calcul donne ici 91,9 % contre 39,3 % pour le dépistage, parce que les spams sont beaucoup plus fréquents (40 %) que la maladie (2 %). La prévalence commande, encore une fois, l'utilité du signal.

Réponse. a) $P(K) = 0,37$ — b) $P_K(S) \approx 0,919$ — c) efficace, mais 8,1 % des messages bloqués sont légitimes : mieux vaut une quarantaine qu'une suppression.

Correction de l'exercice 10 - Sinistres et primes d'assurance

On note J l'événement « l'assuré a moins de 25 ans » et S « il déclare un sinistre ».

a) La formule des probabilités totales :

$$P(S) = 0,25 \times 0,18 + 0,75 \times 0,07 = 0,045 + 0,0525 = 0,0975$$

Soit 9,75 % des assurés qui déclarent un sinistre dans l'année.

b) On inverse la conditionnelle :

$$P_S(J) = \frac{0,045}{0,0975} \approx 0,462$$

Environ 46,2 % des sinistres sont déclarés par des conducteurs de moins de 25 ans, alors qu'ils ne représentent que 25 % des assurés. Ils sont donc très largement surreprésentés.

c) Sur 40 000 assurés :

$$40\,000 \times 0,0975 = 3\,900 \text{ sinistres}$$

dont, pour les jeunes conducteurs :

$$40\,000 \times 0,045 = 1\,800 \text{ sinistres}$$

Contrôle : $\frac{1\,800}{3\,900} \approx 0,462$ ✓ cohérent avec la question b).

La lecture économique. Les jeunes conducteurs coûtent près de la moitié des sinistres pour un quart des cotisants : c'est la justification actuarielle de la surprime qui leur est appliquée. Le calcul de $P_S(J)$ est exactement celui que fait l'assureur.

La nuance à garder. Une corrélation statistique entre âge et sinistralité ne dit rien d'un conducteur **individuel** : elle décrit un groupe, pas une personne.

Réponse. a) $P(S) = 0,0975$ — b) environ 46,2 % — c) 3 900 sinistres, dont 1 800 de jeunes conducteurs.

Correction de l'exercice 11 - Retrouver une conditionnelle manquante

a) Par la formule du produit :

$$P(A \cap B) = P(A) \times P_A(B) = 0,30 \times 0,60 = 0,18$$

b) La formule des probabilités totales s'écrit

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

On isole le terme inconnu :

$$P(\bar{A} \cap B) = P(B) - P(A \cap B) = 0,39 - 0,18 = 0,21$$

Puis, avec $P(\bar{A}) = 0,70$:

$$P_{\bar{A}}(B) = \frac{P(\bar{A} \cap B)}{P(\bar{A})} = \frac{0,21}{0,70} = 0,30$$

c) Vérification par la formule complète :

$$P(B) = 0,30 \times 0,60 + 0,70 \times 0,30 = 0,18 + 0,21 = 0,39$$

On retrouve la valeur donnée dans l'énoncé ✓

La formule lue à l'envers. Utilisée de gauche à droite, elle calcule $P(B)$ à partir des conditionnelles. Utilisée de droite à gauche, comme ici, elle permet de **retrouver** une conditionnelle manquante à partir de $P(B)$. C'est un exercice classique, souvent posé avant une inversion.

Le contrôle d'encadrement. $P(B) = 0,39$ doit être comprise entre les deux conditionnelles 0,30 et 0,60, et plus proche de la première puisque le groupe $\bar{A}$ pèse 70 % ✓

Réponse. a) $P(A \cap B) = 0,18$ — b) $P(\bar{A} \cap B) = 0,21$ et $P_{\bar{A}}(B) = 0,30$ — c) $0,18 + 0,21 = 0,39$ ✓

Correction de l'exercice 12 - Un club et ses compétitions

On note M l'événement « l'adhérent est majeur » et C « il participe aux compétitions ».

a) La formule des probabilités totales :

$$P(C) = 0,70 \times 0,35 + 0,30 \times 0,60 = 0,245 + 0,180 = 0,425$$

Soit 42,5 % des adhérents qui participent aux compétitions.

b) On inverse la conditionnelle :

$$P_C(M) = \frac{P(M \cap C)}{P(C)} = \frac{0,245}{0,425} \approx 0,576$$

Environ 57,6 % des compétiteurs sont majeurs.

c) Les majeurs représentent 70 % du club mais seulement 57,6 % des compétiteurs : ils sont **sous-représentés** parmi les compétiteurs.

C'est logique, puisque leur taux de participation (35 %) est bien plus faible que celui des mineurs (60 %). Le groupe le plus nombreux n'est pas forcément le plus représenté dans une sous-population : tout dépend du taux de participation de chacun.

Le contrôle. $P_C(\overline{M}) = \frac{0,180}{0,425} \approx 0,424$, et $0,576 + 0,424 = 1 \checkmark$

Le point de vigilance. « 70 % des adhérents sont majeurs » et « 57,6 % des compétiteurs sont majeurs » sont deux affirmations compatibles et vraies simultanément : elles portent sur deux populations différentes. Les opposer serait une erreur de lecture.

Réponse. a) $P(C) = 0,425$ — b) environ 57,6 % — c) les majeurs sont sous-représentés (57,6 % contre 70 % du club), car leur taux de participation est plus faible.

Correction de l'exercice 13 - Deux entrepôts et leurs retards

a) La formule des probabilités totales, avec R l'événement « le colis est en retard » :

$$P(R) = 0,65 \times 0,08 + 0,35 \times 0,15 = 0,052 + 0,0525 = 0,1045$$

Soit 10,45 % des colis en retard.

b) On inverse la conditionnelle :

$$P_R(E_1) = \frac{0,052}{0,1045} \approx 0,498$$

Environ 49,8 % des colis en retard viennent de E_1 .

c) De E_2 , mais de très peu :

$$P_R(E_2) = \frac{0,0525}{0,1045} \approx 0,502$$

Les deux valeurs sont quasiment égales (49,8 % contre 50,2 %) : un colis en retard a pratiquement autant de chances de venir de l'un ou de l'autre entrepôt. Contrôle : $0,498 + 0,502 = 1 \checkmark$

L'équilibre trompeur. E_1 traite presque deux fois plus de colis que E_2 , mais son taux de retard est presque deux fois plus faible : les deux effets se **compensent** presque exactement, d'où le partage à égalité des retards.

$$0,65 \times 0,08 = 0,052 \quad 0,35 \times 0,15 = 0,0525$$

Ce que cela implique pour l'action. Comme les deux entrepôts pèsent autant en volume de retards, le choix de la priorité ne peut pas se faire sur ce seul critère : c'est le **taux** (15 % contre 8 %) qui désigne E_2 comme anormal, donc comme la cible la plus prometteuse.

Réponse. a) $P(R) = 0,1045$ — b) environ 49,8 % — c) très légèrement E_2 (50,2 %), les deux contributions étant presque égales.

Correction de l'exercice 14 - Compléter une donnée par les probabilités totales

On note L l'événement « l'article vient de la ligne 1 » et N « l'article est non conforme ». Les données sont $P(L) = 0,40$, $P(N) = 0,06$ et $P_L(N) = 0,04$.

a) Par la formule du produit :

$$P(L \cap N) = P(L) \times P_L(N) = 0,40 \times 0,04 = 0,016$$

b) La formule des probabilités totales donne l'autre chemin par soustraction :

$$P(\bar{L} \cap N) = P(N) - P(L \cap N) = 0,06 - 0,016 = 0,044$$

Puis, avec $P(\bar{L}) = 0,60$:

$$P_{\bar{L}}(N) = \frac{0,044}{0,60} \approx 0,0733$$

La ligne 2 produit donc environ 7,33 % de non-conformes, presque le double de la ligne 1.

c) On inverse la conditionnelle :

$$P_N(\bar{L}) = \frac{0,044}{0,06} \approx 0,733$$

Environ 73,3 % des articles non conformes viennent de la ligne 2.

Le diagnostic industriel. La ligne 2 traite 60 % des articles mais produit 73,3 % des défauts : elle est clairement la source du problème, avec un taux presque double. C'est elle qu'il faut régler en priorité, et cette fois sans ambiguïté — contrairement au cas des deux entrepôts, les deux critères (taux et volume) désignent la même ligne.

Le contrôle final. $0,40 \times 0,04 + 0,60 \times 0,0733 = 0,016 + 0,044 = 0,060 = P(N)$ ✓

Réponse. a) $P(L \cap N) = 0,016$ — b) environ 7,33 % pour la ligne 2 — c) environ 73,3 % des non-conformes viennent de la ligne 2.

Correction de l'exercice 15 - Synthèse : contrôle qualité et décision

On note D l'événement « le composant est défectueux ».

a) Les trois lignes sont incompatibles et

$$P(L_1) + P(L_2) + P(L_3) = 0,45 + 0,35 + 0,20 = 1$$

Elles forment donc une partition. La formule des probabilités totales donne

$$P(D) = 0,45 \times 0,02 + 0,35 \times 0,05 + 0,20 \times 0,08$$

$$P(D) = 0,009 + 0,0175 + 0,016 = 0,0425$$

Soit un taux global de 4,25 % de défauts.

b) Chaque contribution, rapportée au total 0,0425 :

Ligne	L_1	L_2	L_3
Contribution	0,009	0,0175	0,016
Part des défauts	21,2 %	41,2 %	37,6 %

Détail : $\frac{0,009}{0,0425} \approx 0,212$, $\frac{0,0175}{0,0425} \approx 0,412$ et $\frac{0,016}{0,0425} \approx 0,376$. Somme : 1 ✓

c) Sur 50 000 composants :

$$50\,000 \times 0,009 = 450 \quad 50\,000 \times 0,0175 = 875 \quad 50\,000 \times 0,016 = 800$$

soit 2 125 défauts au total, dont 450 de L_1 , 875 de L_2 et 800 de L_3 .

d) Le gain obtenu en ramenant une ligne à 1 % est la baisse de sa contribution, soit $P(L_i)$ multiplié par la baisse de son taux.

$$L_1: 0,45 \times (0,02 - 0,01) = 0,0045$$

$$L_2: 0,35 \times (0,05 - 0,01) = 0,35 \times 0,04 = 0,014$$

$$L_3: 0,20 \times (0,08 - 0,01) = 0,20 \times 0,07 = 0,014$$

Les lignes L_2 et L_3 donnent **exactement le même gain**, 0,014, qui ferait passer le taux global de 4,25 % à 2,85 %. La ligne L_1 n'apporterait que 0,0045, soit trois fois moins.

Le choix se fera donc entre L_2 et L_3 sur d'autres critères : coût de l'intervention, faisabilité technique, ancienneté des machines.

e) On inverse les conditionnelles, ce qui redonne les parts calculées en b) :

$$P_D(L_2) \approx 0,412 \quad P_D(L_3) \approx 0,376 \quad P_D(L_1) \approx 0,212$$

Un composant défectueux vient donc le plus probablement de la ligne L_2 , avec environ 41,2 % de chances.

Le bilan de la plaquette. La formule des probabilités totales additionne les chemins d'un arbre en les **pondérant** par la taille de chaque groupe ; l'inversion des conditionnelles retourne ensuite la question. Les deux ensemble répondent à toutes les situations du chapitre — et la question d) montre qu'un calcul de probabilités peut directement fonder une décision.

Réponse. a) somme des parts égale à 1 et $P(D) = 0,0425$ — b) 21,2 %, 41,2 % et 37,6 % — c) 450, 875 et 800 défauts — d) L_2 ou L_3 , à égalité avec un gain de 0,014 — e) de L_2 , avec 41,2 % de chances.