Probabilités conditionnelles
Problème de recherche : le sondage confidentiel

Le service de contrôle de gestion d'une grande entreprise veut estimer la proportion p de salariés qui ont déjà gonflé une note de frais. Personne n'osera répondre « oui » franchement… On utilise donc un protocole de **réponse aléatoire** :

— chaque salarié lance une pièce équilibrée, en cachette ;
— s'il obtient **pile**, il répond honnêtement à la question « Avez-vous déjà gonflé une note de frais ? » ;
— s'il obtient **face**, il répond « oui », quoi qu'il en soit.

L'enquêteur ne voit que la réponse, jamais le résultat de la pièce. Sur 1 000 salariés interrogés, on obtient 600 réponses « oui ».

La mission

1. Exprimer, en fonction de p, la probabilité qu'un salarié choisi au hasard réponde « oui ».
2. En déduire une estimation de la proportion p.
3. Un directeur veut sanctionner tous les salariés qui ont répondu « oui ». Quelle est la probabilité qu'un salarié ayant répondu « oui » ait réellement gonflé une note de frais ? Le protocole protège-t-il les salariés ?

Pour aller plus loin

Un autre service utilise un dé : sur 1, 2, 3 ou 4, on répond honnêtement ; sur 5, on répond « oui » ; sur 6, on répond « non ». Il obtient 30 % de « oui ». Estimer la proportion de tricheurs dans ce service.

Coups de pouce

Bloqué ? Des indices progressifs sont disponibles en ligne. N'en ouvrir qu'un à la fois, et seulement après avoir cherché !

Éléments de correction

On note P : « la pièce tombe sur pile », T : « le salarié a déjà gonflé une note de frais », O : « le salarié répond oui ».

1. Probabilité de répondre oui

Arbre : P (0,5) puis T (p) → oui, ou $\overline{T}$ ($1-p$) → non ; $\overline{P}$ (0,5) → oui dans tous les cas.

$P(O) = P(P \cap O) + P(\overline{P} \cap O) = 0{,}5 \times p + 0{,}5 \times 1 = 0{,}5p + 0{,}5$.

2. Estimation de p

La fréquence observée des « oui » est $\frac{600}{1\,000} = 0{,}6$. On résout $0{,}5p + 0{,}5 = 0{,}6$: $p = 0{,}2$.

Environ **20 %** des salariés auraient déjà gonflé une note de frais.

3. Que signifie un « oui » ?

On cherche $P_O(T)$. Un salarié tricheur répond toujours « oui » (avec pile comme avec face), donc $P(T \cap O) = P(T) = 0{,}2$. Un honnête répond « oui » s'il obtient face : $P(\overline{T} \cap O) = 0{,}8 \times 0{,}5 = 0{,}4$. On retrouve $P(O) = 0{,}6$, et

$$P_O(T) = \frac{P(T \cap O)}{P(O)} = \frac{0{,}2}{0{,}6} = \frac{1}{3} \approx 0{,}33.$$

Deux salariés sur trois qui ont répondu « oui » sont honnêtes : **on ne peut accuser personne**. Le protocole protège chaque salarié, tout en donnant une estimation fiable pour l'ensemble de l'entreprise.

Pour aller plus loin

Avec le dé : $P(O) = \frac{4}{6}p + \frac{1}{6}$. On résout $\frac{2}{3}p + \frac{1}{6} = 0{,}3 : \frac{2}{3}p = \frac{2}{15}$, donc $p = 0{,}2$. Ici aussi, environ 20 % de tricheurs. Avantage de ce protocole : répondre « non » ne prouve pas non plus qu'on est honnête.

Ce qu'on a utilisé du chapitre

- Arbre pondéré ; formule des probabilités totales.
- Probabilité conditionnelle $P_O(T) = \dfrac{P(T \cap O)}{P(O)}$.