Exercice sur la règle de Bayes
Quelques subtilités

On a vu en classe que la règle de Bayes n'était pas toujours très utile en intelligence artificielle, et que le facteur crucial est l'incidence dans la population du défaut à identifier *a posteriori*.

Revoici le problème libellé dans un autre contexte que celui de l'identification des sidéens. Tous les tests diagnostiques sont d'ailleurs un peu dans ce cas

On dispose d'un test pour juger de la conformité d'un produit à une norme. Ce test n'est pas infaillible.

On note l'événement 'Être conforme' par C, 'non conforme' par NC ; l'événement 'Être positif à un test de non conformité', i.e. non conforme, par '$+$' et 'Être négatif à ce test', i.e. conforme, par '$-$'. Supposons qu'on ait les probabilités suivantes pour ce test, dites les probabilités *a priori* :

$$\mathrm{P}\big(\,+\mid NC\,\big) = 0{,}999, \quad \mathrm{P}\big(\,-\mid C\,\big) = 0{,}995\,.$$

On sait pas ailleurs que $\mathrm{P}\big(\,NC\,\big) = 0{,}001$.

1. (a) Obtenez les probabilités *a posteriori* $\mathrm{P}\big(\,C\mid +\,\big)$ & $\mathrm{P}\big(\,NC\mid -\,\big)$. En déduire $\mathrm{P}\big(\,NC\mid +\,\big)$ & $\mathrm{P}\big(\,C\mid -\,\big)$.

   (b) Écrire une petite macro Excel où les diverses valeurs sont modifiables avec les *a posteriori* en 'sortie' qui se calculent selon les *a priori* et la valeur de $\mathrm{P}\big(\,C\,\big)$.

   (c) Si on supposait que le test était infaillible, contrairement à ce qu'on donne dans les *a priori* de l'énoncé, en moyenne combien faudrait-il en tester avant de trouver le premier non conforme lors d'essais indépendants successifs. Énoncez au passage la loi qui modélise cette probabilité avec ses paramètres.

   (d) Si on teste 1 000 produits et qu'on examine le nombre de non conformes, encore dans le cas énoncé en (1c), énoncez la loi du nombre de produits non conformes pour ces 1 000 essais. Énoncez la loi qui modélise ce nombre, et les hypothèses nécessaires à son utilisation.

(e) Utilisez une approximation appropriée de cette loi pour déterminer la probabilité d'observer plus de 2 produits dans l'ensemble des 1 000 produits.

(f) Comme on a vu dans la manuel ou à l'item (1a), la règle de Bayes ne permet pas de conclure. À noter qu'on n'a pas d'autre possibilité que d'utiliser cette règle pour conclure...

Supposons maintenant qu'on effectue alors deux fois le test, de façon indépendante l'un de l'autre. Obtenez les diverses probabilités

$$\mathrm{P}\big(NC \mid ++\big)\,, \quad \mathrm{P}\big(NC \mid +-\big)\,, \quad \mathrm{P}\big(NC \mid --\big)\,,$$

de même que $\mathrm{P}\big(C \mid ++\big)$, etc.