Probabilités et simulations.
Feuille de TD 2.

Dans ce qui suit on suppose que l'on travaille dans un espace probabilisé et que tous les ensembles A, B, A_n, B_n,... appartiennent à la tribu.

Exercice 1 Montrer que si $A \cap B = \emptyset$, alors A et B *ne sont pas* indépendants, sauf si $\mathbb{P}(A) = 0$ ou $\mathbb{P}(B) = 0$.

Exercice 2 Si $\mathbb{P}(C) > 0$, montrer que

$$\mathbb{P}(A \cup B|C) = \mathbb{P}(A|C) + \mathbb{P}(B|C) - \mathbb{P}(A \cap B|C).$$

Exercice 3 Si $\mathbb{P}(C) > 0$ et si les A_i sont deux à deux disjoints, montrer que

$$\mathbb{P}(\bigcup_{i=1}^{n} A_i|C) = \sum_{i=1}^{n} \mathbb{P}(A_i|C)$$

Exercice 4 *Principe de la méthode du rejet*
On considère un espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$ et un événement A.

(1) On suppose que la procédure `psimul` simule ω de loi $\mathbb{P}$. On suppose d'autre part que la procédure `testA(psimul)` renvoie le booléen `T` si $\omega \in A$ et `F` sinon.
Expliquer alors ce que réalise l'algorithme suivant pour $n \in \mathbb{N}^*$ (on commencera par préciser ce que renvoie l'algorithme lorsque $n = 1$):
```
function M=rejet(n)
counter = 1;
M = [];
while counter < n+1
x= psimul;
if testA(x) == T
then M(counter)=x; counter=counter+1;
end
end
disp(M);
endfunction
```
(2) Simuler simplement une variable Unif$[0, 9]$. Ecrire alors un programme basé sur la méthode du rejet qui simule n réalisations d'une variable uniforme sur la réunion d'intervalles $[0, 1] \cup [3, 5] \cup [8, 9]$.

Exercice 5 Un entreprise pharmaceutique souhaite commercialiser un "pré-test" pour le VIH. On suppose que ce pré-test a une efficacité $\mathbf{e} := 99\%$ (= probabilité pur que ce test soit positif pour une personne atteinte du SIDA) et une probabilité de fausse alarme de $\mathbf{f} := 5\%$ (=probabilité que le test soit positif pour une personne non atteinte). Enfin, on suppose que la fréquence de séropositivité est $\mathbf{s} := 1/10000$.

On fait les hypothèses habituelles d'indépendance pour les différents dons testés, ainsi que sur les tests.

(1) Quelle est, en fonction de $\mathbf{s}, \mathbf{e}, \mathbf{f}$, la probabilité p pour qu'une personne obtenant un test positif soit effectivement malade ? Qu'obtient-on pour les valeurs de l'énoncé?

Pour $\mathbf{e} = 99\%, \mathbf{s} = 1/10000$ comme ci-dessus, quel $\mathbf{f}$ faudrait-il viser pour que $p \geq 1/4$?

(2) Ecrire un programme qui, en fonction de $\mathbf{e}, \mathbf{f}$ simule 1000000 dons de sang, et qui renvoie le nombre de tests positifs, le nombre de fausses alarmes, et le nombre de malades non détectés. Déduire de la simulation une estimation de la probabilité conditionnelle calculée en 1.

Que constatez-vous pour les valeurs de $\mathbf{e}, \mathbf{f}$ ci-dessus? Au vu des résultats exacts et ceux de votre simulation, pensez-vous qu'un tel test puisse raisonnablement être mis en place?

(3) Ecrire un programme pour estimer la probabilité conditionnelle calculée en 1, en considérant cette fois un échantillon de 50000 tests positifs. On basera la simulation sur une méthode de rejet.

Exercice 6 On modélise le jet d'une pièce avec deux résultats possibles pile (P) et face (F), chacun avec la probabilité 1/2. On jette deux fois cette pièce de sorte que l'ensemble des résultats possibles Ω contient les quatre points PP, FF, PF, FP. On suppose que les deux jets sont indépendants.

(1) Quelle est la probabilité d'obtenir deux fois face sachant que le premier jet donne face ?
(2) Quelle est la probabilité d'obtenir deux fois face sachant que l'un des deux jets au moins donne face?
(3) Simuler les deux expériences précédentes grâce à une méthode de rejet.

Exercice 7 Si A, B, C sont indépendants et $\mathbb{P}(A \cap B) \neq 0$, montrer que

$$\mathbb{P}(C|A \cap B) = \mathbb{P}(C).$$

Exercice 8 *Urnes de Polya.*
Une urne contient r boules rouges et n boules noires. Une boule est choisie au hasard, on note sa couleur, et on la remet avec d boules supplémentaires de la même couleur. Puis on recommence la même procédure aussi souvent que nécessaire.

(1) Trouver la probabilité que la seconde boule tirée est noire.
(2) Trouver la probabilité que la première boule tirée est noire sachant que la seconde est noire.
(3) On note N_m l'événement selon lequel la m-ième boule tirée est noire. Montrer que $\mathbb{P}(N_m) = \mathbb{P}(N_1)$.
(4) Trouver la probabilité pour que la première boule soit noire, sachant que les m suivantes sont noires. Trouver ensuite la limite de cette probabilité lorsque $m \to \infty$.
(5) Ecrire un programme `untirage(r,n,d)` qui modélise le tirage de la première boule. Plus précisément, pour (r, n) donné, `untirage(r,n,d)` simulera le nombre de boules rouges et noires (r', n') présentes dans l'urne après que le premier tirage—remise a été effectué.
Ecrire alors des programmes qui permettent de simuler les fréquences correspondantes aux événements des questions 1,2,3,4.
(6) Simuler 10000 tirages successifs, et faire un graphe des proportions successives de boules rouges. Qu'observez-vous?

(7) Pour $r = n = d = 1$, simuler 1000 réalisations indépendantes de 1000 tirages successifs, et faire un histogramme des proportions de boules rouges à l'issue de ces 1000 tirages (on utilisera la commande `histplot`). Qu'observez-vous? Répéter cette procédure pour différentes valeurs de r, n, d.

Exercice 9 Une compagnie d'assurance assure un nombre égal de conducteurs masculins et féminins. Tous les conducteurs masculins ont chaque année la probabilité α d'avoir un accident, et ceci indépendamment des autres années et des autres conducteurs. Il en va de même des conductrices avec une probabilité d'accident β. La compagnie choisit un numéro d'assuré au hasard.

(1) Quelle est la probabilité pour que la personne sélectionnée ait un accident cette année ?
(2) Quelle est la probabilité pour que la personne sélectionnée ait un accident deux années consécutives?
(3) Soit A_i l'événement "la personne sélectionnée a un accident l'année numéro i". Montrer que $\mathbb{P}(A_2|A_1) \geq \mathbb{P}(A_1)$.
(4) Trouver la probabilité pour que, une année donnée, une personne sélectionnée au hasard parmi celles ayant eu un accident soit une conductrice.

Exercice 10 *Règle de succession de Laplace.* Une boite contient $k+1$ pièces numérotées de 0 à k. Lorsque la i-ième pièce est lancée, $0 \leq i \leq k$, on obtient Pile avec $\frac{i}{k}$. On tire au hasard une pièce dans l'urne pour la jeter ensuite un grand nombre de fois et on suppose alors que le choix de la pièce à lancer étant fait, les jets sont indépendants.

(1) Quelle est la probabilité conditionnelle $p_{k,n}$ que le $(n+1)$-ième lancer donne Pile sachant que les n premiers l'ont fait? Quelle est la limite de $p_{k,n}$ quand k tend vers l'infini?
(2) Sachant que les n premiers lancers ont donné Pile, montrer que la probabilité conditionnelle que les m suivants fassent de même tend, pour k tendant vers l'infini, vers $\frac{n+1}{n+m+1}$.
(3) On suppose que les n premiers jets ont donné r Piles et $n-r$ Faces. Montrer que la probabilité conditionnelle que le $(n+1)$-ième donne Pile tend, pour k tendant vers l'infini, vers $\frac{r+1}{n+2}$.

   Indication: on pourra utiliser l'identité:
$$\int_0^1 y^n(1-y)^m dy = \frac{n!\ m!}{(n+m+1)!}.$$
(4) Quelle est la probabilité que la i-ième pièce ait été tirée si les n premiers jets ont donné Pile?
(5) Ecrire un programme qui modélise le choix initial de la pièce et n jets indépendants de celle-ci.
(6) Utiliser un algorithme de rejet pour estimer la fréquence conditionnelle (sur 1000 réalisations) des événements en 1 et 2. Comparer avec les résultats obtenus en 1,2.
(7) Estimer les fréquences conditionnelles (sur 1000 réalisations) en 3,4 pour $k = 1000$, et
   - $n = 10, m = 5$.
   - $n = 10, r = 3$.

   Comparer avec les limites exactes données en 3,4.