# Introduction à la théorie des probabilités - Exercices

**Préambule**

Deux propositions $A$ et $B$ sont dites *indépendantes* (étant donnée une information contextuelle $C$) si $\mathrm{P}(AB|C) = \mathrm{P}(A|C)\ \mathrm{P}(B|C)$. Cette information contextuelle est souvent sous-entendue: on exprime alors l'indépendance par $\mathrm{P}(AB) = \mathrm{P}(A)\ \mathrm{P}(B)$. Comme, de maniére générale,

$$\mathrm{P}(AB) = \mathrm{P}(A|B)\ \mathrm{P}(B) = \mathrm{P}(B|A)\ \mathrm{P}(A),$$

cela revient à dire que $\mathrm{P}(A|B) = \mathrm{P}(A)$ et $\mathrm{P}(B|A) = \mathrm{P}(A)$: le fait de savoir que $B$ ($A$) est vraie ne modifie pas la plausibilité de $A$ ($B$).

**Exercices élémentaires de probabilité**

**(1)** La distribution des groupes sanguins aux USA est

| Groupe sanguin | 0 | A | B | AB |
|---|---|---|---|---|
| Probabilité | 0.45 | 0.40 | 0.11 | ? |

**-a-** Que vaut $\mathrm{P}(AB)$? (Rép: 0.04).
**-b-** Maria est du groupe $B$. Elle peut recevoir une transfusion des personnes des groupes $O$ et $B$. Quelle est la probabilité qu'un Américain choisi au hasard puisse lui donner du sang sans risque? (Rép: 0.56).

**(2)** La distribution des groupes sanguins en Chine est donnée par

| Groupe sanguin | 0 | A | B | AB |
|---|---|---|---|---|
| Probabilité | 0.35 | 0.27 | 0.26 | 0.12 |

Si un Américain et un Chinois sont choisis indépendamment et au hasard,
**-a-** quelle est la probabilité qu'ils soient tous les deux du groupe 0? (Rép: 0.1575).
**-b-** quelle est la probabilité qu'ils appartiennent au même groupe sanguin? (Rép: 0.2989).

**(3)** Les couleurs les plus populaires pour les voitures aux USA sont

| Couleur | Grise | Blanche | Noire | Vert foncé | Bleu foncé | Vert moyen |
|---|---|---|---|---|---|---|
| Probabilité | 0.176 | 0.172 | 0.113 | 0.089 | 0.088 | 0.067 |

**-a-** Quelle est la plausibilité que le véhicule acheté par un consommateur américain (dont on ne sait rien a priori) soit d'une autre couleur que les six citées? (R: 0.295).
**-b-** Quelle est la plausibilité que le véhicule de cet Américain soit gris ou blanc? (R: 0.348).
**-c-** Quelle est la plausibilité que deux Américains (dont on ne sait rien d'autre a priori que ce qui suit) possédant chacun un véhicule aient tous les deux une voiture de couleur grise ou blanche? (R: 0.121104)

**Manipulation de la table $\mathcal{N}(0,1)$**

**(4)** A l'aide d'une table appropriée et en supposant que $Z \sim \mathcal{N}(0,1)$, vérifier:
**-a-** $\mathrm{P}(Z < 1.5) = 0.933$ ; $\mathrm{P}(Z < -2.85) = 0.0022$ ; $\mathrm{P}(Z > 1.8) = 0.036$ ; $\mathrm{P}(Z > -2.1) = 0.982$.
**-b-** Les quantiles 2.5%, 25%, 50%, 75% et 97.5% de la normale $\mathcal{N}(0,1)$ sont respectivement -1.96, -0.67, 0, 0.67 et 1.96.

**(5)** La durée d'une grossesse chez l'humain suit approximativement une distribution normale de moyenne égale à 266 jours et d'écart-type égal à 16 jours. Vérifier, à l'aide d'une table appropriée, que
**-a-** le pourcentage de grossesses de moins de 240 jours (càd 8 mois) est 5.2%.
**-b-** le pourcentage de grossesses d'une durée comprise entre 240 et 270 jours (càd entre 8 et 9 mois) est 54.7%.
**-c-** les 20% de grossesses les plus longues ont une durée au moins égale à 279.5 jours.

**Compréhension de la probabilité conditionnelle**

**(6)** Un recensement de la population américaine révèle la répartition suivante des femmes adultes (en milliers):

| | Age (en années) | | | |
|---|---|---|---|---|
| | 18-29 | 30-64 | $\geq 65$ | Total |
| Mariée | 7842 | 43808 | 8270 | 59920 |
| Jamais mariée | 13930 | 7184 | 751 | 21865 |
| Veuve | 36 | 2523 | 8385 | 10944 |
| Divorcée | 704 | 9174 | 1263 | 11141 |
| Total | 22512 | 62689 | 18669 | 103870 |

Supposons qu'une américaine adulte soit choisie au hasard.
**-a-** Quelle est la probabilité qu'elle ait moins de 30 ans? (R: 0.217).

**-b-** Quelle est la probabilité qu'elle ait moins de 30 ans et soit mariée? (R: 0.075).
**-c-** Sachant qu'elle a moins de 30 ans, quelle est la probabilité qu'elle soit mariée? (R: 0.348).
**-d-** Sachant qu'elle est mariée, quelle est la probabilité qu'elle ait moins de 30 ans? (R: 0.131).

**Si ...,** $\mathrm{P}(A) = \sum_k \mathrm{P}(A|B_k)\ \mathrm{P}(B_k)$

Pour rappel, si $B_1$, $B_2$,...,$B_n$ sont $n$ propositions telles que
$\rightarrow$ si $B_i$ est vraie, alors $B_j$ ($j \neq i$) est nécessairement fausse (on dit que $B_1$, $B_2$,...,$B_n$ sont *mutuellement exclusives*).
$\rightarrow$ nécessairement, une des $B_i$ ($i = 1, \dots, n$) est vraie,
alors

$$\mathrm{P}(A) = \mathrm{P}(A|B_1)\mathrm{P}(B_1) + \ldots + \mathrm{P}(A|B_n)\mathrm{P}(B_n)$$

En particulier,

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

**(7)** Une compagnie démarche des acheteurs par téléphone. Elle choisit ses correspondants au hasard dans le botin. L'expérience montre que 70% des appels n'aboutissent pas (pas de réponse ou refus de parler), 20% conduisent à une interlocutrice et 10% à un interlocuteur. Sachant que 30% des femmes et 20% des hommes engageant la discussion finissent par acheter quelque chose, quel est le pourcentage d'appels conduisant à une vente? (R: 0.08).

**(8)** Les électeurs d'une grande ville américaine sont constitués de 40% de blancs, 40% de noirs et 20% d'hispaniques. Un candidat noir à la fonction de Maire espère attirer 30% des votes chez les blancs, 90% chez les noirs et 50% chez les hispaniques. Quel pourcentage des votes globaux espère-t-il donc réunir? (R: 58%).

**Distribution, moyenne et variance d'une variable aléatoire discrète**

**(10)** Des tickets de loterie sont vendus un euro. Un ticket sur 1000 est gagnant et rapporte une somme de 500 euros ; tous les autres sont perdants.
**-a-** Quelle est la distribution de la variable aléatoire $X$ (=*gain brut* ne tenant pas compte de la mise de départ)? (R: $\mathcal{E} = \{0, 500\}$ avec probs. associées 0.999 et 0.001).
**-b-** Que valent la moyenne et la variance de $X$? (R: $\mu_X = 0.5$ ; $\sigma_X^2 = 249.75$).

**-c-** Que valent la moyenne et la variance du gain net $W$ (obtenu en soustrayant le coût du ticket du gain brut)? (R: $\mu_W = -0.5$ ; $\sigma_W^2 = 249.75$).
**-d-** Que valent la moyenne et l'écart-type du gain brut $Z = X_1 + X_2$ résultant de l'achat de 2 tickets? (R: $\mu_Z = 1.0$ ; $\sigma_Z = 22.35$).

**(11)** Vous disposez de 2 dés équilibrés. Le premier présentent sur ses faces les valeurs 1, 3, 4, 5, 6 et 8. Le deuxième présente les valeurs 1, 2 , 2, 3, 3 et 4.
**-a-** Quelle est la distribution de probabilité pour le résultat du 1er dé? Même question pour le 2ème dé. En déduire la moyenne du résultat obtenu pour un lancé pour chacun des dés. (R: $\mu_1 = 4.5$ ; $\mu_2 = 2.5$).
**-b-** Quelle est la distribution de probabilité de la somme des résultats obtenus par les lancés simultanés des 2 dés?
(R:

| 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1/36 | 2/36 | 3/36 | 4/36 | 5/36 | 6/36 | 5/36 | 4/36 | 3/36 | 2/36 | 1/36 |

)
**-c-** Que vaut la moyenne de la somme des résultats obtenus par les lancés simultanés des 2 dés? Touvez la réponse par 2 méthodes: 1: en exploitant le fait que le total est la somme de 2 variables aléatoires (indépendantes) ; 2: en utilisant la distribution de probabilité obtenue en -b-. (R: 7).

**La distribution binomiale et son approximation normale**

Pour rappel, lorsque $np > 5$ et $n(1-p) > 5$, on peut approximer la binomiale $\text{Bin}(n, p)$ par la normale $\mathcal{N}(np, np(1-p))$. En particulier, si $Y \sim \text{Bin}(n, p)$, alors, sous les conditions précédentes, $\hat{p} = \frac{Y}{n} \sim N(p, p(1-p)/n)$.

**(12)** Un moyen de vérifier si les individus sélectionnés pour participer à un sondage sont "représentatifs" de la population cible est de comparer certaines caractéristiques démographiques de l'échantillon avec les caractéristiques correspondantes de la population. Supposons qu'un sondage soit organisé aux USA auprès d'adultes (18 ans et plus). On sait, d'après un recensement, que 23772494 des 209128094 adultes se qualifient comme étant "noir" (ou "afro-américain").
**-a-** Quelle est la proportion $p$ de noirs aux USA? (R: 0.1136743)
**-b-** Un sondage est organisé auprès de 1500 adultes choisis au hasard. Quelle est la distribution du nombre de noirs dans cet échantillon? Que valent la moyenne et la variance de cette variable aléatoire? (R: $\text{Bin}(1500, p)$ ; 170.5 ; 151.13).
**-c-** Si une approximation le permet (vérifier), calculer la probabilité que

l'échantillon contienne au moins 170 noirs. (R: 0.52).

**(13)** Une étude a été menée pour évaluer les capacités humaines de perception extra-sensorielle. A cette fin, le maître de cérémonie dispose de 30 cartes portant un des quatre symboles habituels (pic, coeur, trèfle ou carreau). Il est demandé à chaque personne testée de deviner, un par un, le symbole sur chacune des 30 cartes. (La proportion de carte de chaque type n'est pas annoncée.)
**-a-** Si la personne testée ne dipose d'aucun don particulier en ce domaine, quelle est la probabilité qu'il annonce le symbole correct pour une carte donnée? Quelle est la distribution du nombre de cartes correctement classées? (R: 1/4 ; $\text{Bin}(30, 1/4)$).
**-b-** En utilisant l'approximation normale (vérifier les conditions), calculer la probabilité qu'un sujet sans don particulier donne un résultat correct pour au moins 15 des 30 cartes. (R: $1 - \text{P}(Z \geq 3.16) = 0.0008$.)
**-c-** Calculer les quantiles 2.5% et 97.5% de cette approximation normale. En déduire dans quel intervalle on retrouve 95% du nombre de symboles correctement devinés par des individus sans don particulier. (R: (3,12) après arrondi de 2.85 et 12.15)