STA240 : Probabilités

Rappel de cours :

1 Axiomes des probabilités

- La *probabilité d'un événement* dans une population est la proportion des individus pour lesquels l'événement est réalisé.
- $P(A \text{ ou } B) = P(A) + P(B) - P(A \text{ et } B)$
- $P(A) = 1 - P(\overline{A})$
- La *probabilité conditionnelle de A sachant B* est la proportion d'individus pour lesquels A est réalisé *parmi ceux pour lesquels B l'est aussi*. C'est le rapport de la probabilité de "A et B" à la probabilité de B :

$$\mathbb{P}[A \mid B] = \frac{\mathbb{P}[A \text{ et } B]}{\mathbb{P}[B]} \; .$$

- Si A et B sont *indépendants*, $P(A/B) = P(A)$ et $P(A \text{ et } B) = P(A) * P(B)$
- La *formule des probabilités totales* donne la probabilité d'un événement A en fonction des probabilités conditionnelles sachant un autre événement B et son contraire $\overline{B}$:

$$\mathbb{P}[A] = \mathbb{P}[A \mid B]\,\mathbb{P}[B] + \mathbb{P}[A \mid \overline{B}]\,\mathbb{P}[\overline{B}] \; .$$

- La *formule de Bayes* permet d'échanger l'ordre des probabilités conditionnelles :

$$\mathbb{P}[B \mid A] = \frac{\mathbb{P}[A \mid B]\,\mathbb{P}[B]}{\mathbb{P}[A \mid B]\,\mathbb{P}[B] + \mathbb{P}[A \mid \overline{B}]\,\mathbb{P}[\overline{B}]} \; .$$

2 Loi binomiale

- Au cours de n *expériences répétées indépendamment*, la variable aléatoire X égale au nombre de réalisations d'un même *événement de probabilité p*, suit la loi *binomiale* de paramètres n et p.
- La variable X peut prendre toutes les valeurs entières entre 0 et n.
- Pour tout entier k entre 0 et n, la variable X prend la valeur k avec la probabilité :

$$\mathbb{P}[X = k] = \binom{n}{k} p^k \, (1-p)^{n-k} \; ,$$

où

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n \times (n-1) \cdots \times (n-k+1)}{k \times (k-1) \cdots \times 3 \times 2 \times 1}$$

est le nombre de manières de choisir k objets parmi n.
- L'espérance de X est np, sa variance est $np(1-p)$.

3 Loi hypergéométrique

- Dans un ensemble de N éléments, parmi lesquels m sont marqués, on en choisit au hasard n distincts. La variable aléatoire X égale au nombre d'éléments marqués parmi l'échantillon de n suit la la loi *hypergéométrique de paramètres* N, m, n.
- Dans le cas où $n \leq m$ et $n \leq N - m$, X peut prendre toutes les valeurs entières entre 0 et n.
- Pour tout entier k entre 0 et n, X prend la valeur k avec probabilité :

$$\mathbb{P}[X = k] = \frac{\binom{m}{k}\binom{N-m}{n-k}}{\binom{N}{n}} .$$

- L'espérance de X est nm/N.

4 Loi normale

- Si on n'a pas de logiciel à disposition, on lit dans les tables pour la loi $\mathcal{N}(0,1)$:
  - ⋆ les valeurs de la fonction de répartition F : pour une valeur de x la table retourne la probabilité $p = P[X \leqslant x] = F(x)$.
  - ⋆ les valeurs de la fonction quantile F^{-1} : pour une probabilité p la table retourne la valeur de $x = F^{-1}(p)$ telle que $p = \mathbb{P}[X \leqslant x]$.
- La densité de la loi $\mathcal{N}(0,1)$ est symétrique :

$$\mathbb{P}[X \leqslant -x] = \mathbb{P}[X \geqslant x] .$$

- Si une variable aléatoire X suit la loi $\mathcal{N}(\mu, \sigma^2)$, alors $(X - \mu)/\sqrt{\sigma^2}$ suit la loi $\mathcal{N}(0,1)$. Ainsi :

$$\begin{aligned} \mathbb{P}[a \leqslant X \leqslant b] &= P\left[\frac{a-\mu}{\sqrt{\sigma^2}} \leqslant \frac{X-\mu}{\sqrt{\sigma^2}} \leqslant \frac{b-\mu}{\sqrt{\sigma^2}}\right] \\ &= F\left(\frac{b-\mu}{\sqrt{\sigma^2}}\right) - F\left(\frac{a-\mu}{\sqrt{\sigma^2}}\right) , \end{aligned}$$

  où F est la fonction de répartition de la loi $\mathcal{N}(0,1)$.
- Si X et Y sont deux variables aléatoires indépendantes, de lois respectives $\mathcal{N}(\mu_x, \sigma_x^2)$ et $\mathcal{N}(\mu_y, \sigma_y^2)$, alors $X + Y$ suit la loi $\mathcal{N}(\mu_x + \mu_y, \sigma_x^2 + \sigma_y^2)$ et $X - Y$ suit la loi $\mathcal{N}(\mu_x - \mu_y, \sigma_x^2 + \sigma_y^2)$.

Exercices :

Exercice 1. Dans un élevage de moutons, on estime que 30% sont atteints par une certaine maladie. On dispose d'un test pour cette maladie. Si un mouton n'est pas atteint, il a 9 chances sur 10 d'avoir une réaction négative au test ; s'il est atteint, il a 8 chances sur 10 d'avoir une réaction positive. On soumet tous les moutons de l'élevage au test.

Pour tout l'exercice, on note M l'événement "le mouton est malade" et T l'événement "le mouton a une réaction positive au test. L'énoncé donne :

$$\mathbb{P}[M] = 0.3 \;, \quad \mathbb{P}[\overline{T} \,|\, \overline{M}] = 0.9 \;, \quad \mathbb{P}[T \,|\, M] = 0.8 \;.$$

1. Quelle est la probabilité qu'un mouton de cet élevage ne soit pas malade ?

$$\mathbb{P}[\overline{M}] = 1 - \mathbb{P}[M] = 1 - 0.3 = 0.7 \;.$$

2. Quelle est la probabilité conditionnelle qu'un mouton ait une réaction positive au test sachant qu'il n'est pas malade ?

$$\mathbb{P}[T \,|\, \overline{M}] = 1 - \mathbb{P}[\overline{T} \,|\, \overline{M}] = 1 - 0.9 = 0.1 \;.$$

3. Quelle est la probabilité qu'un mouton ne soit pas malade et ait une réaction positive au test ?

$$\mathbb{P}[T \text{ et } \overline{M}] = \mathbb{P}[T \,|\, \overline{M}]\, \mathbb{P}[\overline{M}] = 0.1 \times 0.7 = 0.07 \;.$$

4. Quelle proportion des moutons de l'élevage réagit positivement au test ?

   On peut utiliser la formule des probabilités totales ou raisonner directement, en distinguant, parmi les moutons ayant réagi positivement, ceux qui sont malades de ceux qui ne le sont pas.

$$\begin{aligned} \mathbb{P}[T] &= \mathbb{P}[T \text{ et } M] + \mathbb{P}[T \text{ et } \overline{M}] \\ &= \mathbb{P}[T \,|\, M]\, \mathbb{P}[M] + \mathbb{P}[T \,|\, \overline{M}]\, \mathbb{P}[\overline{M}] \\ &= 0.8 \times 0.3 + 0.1 \times 0.7 = 0.24 + 0.07 = 0.31 \;. \end{aligned}$$

5. Quelle est la probabilité qu'un mouton soit malade, sachant qu'il a réagi positivement ?

   On peut utiliser directement la formule de Bayes ou bien la retrouver comme suit.

$$\begin{aligned} \mathbb{P}[M \,|\, T] &= \frac{\mathbb{P}[T \text{ et } M]}{\mathbb{P}[T]} \\ &= \frac{\mathbb{P}[T \,|\, M]\, \mathbb{P}[M]}{\mathbb{P}[T \,|\, M]\, \mathbb{P}[M] + \mathbb{P}[T \,|\, \overline{M}]\, \mathbb{P}[\overline{M}]} \\ &= \frac{0.8 \times 0.3}{0.8 \times 0.3 + 0.1 \times 0.7} \simeq 0.774 \;. \end{aligned}$$

6. Quelle est la probabilité qu'un mouton ne soit pas malade, sachant qu'il a réagi négativement ?

   On peut utiliser directement la formule de Bayes ou bien la retrouver comme suit.

$$\begin{aligned}
\mathbb{P}[\overline{M} \,|\, \overline{T}] &= \frac{\mathbb{P}[\overline{T} \text{ et } \overline{M}]}{\mathbb{P}[\overline{T}]} \\
&= \frac{\mathbb{P}[\overline{T} \,|\, \overline{M}]\, \mathbb{P}[\overline{M}]}{\mathbb{P}[\overline{T} \,|\, \overline{M}]\, \mathbb{P}[\overline{M}] + \mathbb{P}[\overline{T} \,|\, M]\, \mathbb{P}[M]} \\
&= \frac{0.9 \times 0.7}{0.9 \times 0.7 + 0.2 \times 0.3} \simeq 0.913 \;.
\end{aligned}$$

Exercice 2. On sait par expérience qu'une certaine opération chirurgicale a 90% de chances de réussir. On s'apprête à réaliser l'opération sur 5 patients. Soit X la variable aléatoire égale au nombre de réussites de l'opération sur les 5 tentatives.

1. Quel modèle proposez-vous pour X ?

   En supposant que les résultats (succès ou échec) des 5 opérations soient indépendants entre eux, le nombre de succès suit la loi binomiale de paramètres 5 *et* 0.9. *La variable aléatoire X prend ses valeurs dans l'ensemble* $\{0, 1, 2, 3, 4, 5\}$*, et pour tout entier k dans cet ensemble :*

$$\mathbb{P}[X = k] = \binom{5}{k} 0.9^k \, 0.1^{5-k} \;.$$

2. Quelle est la probabilité que l'opération rate les 5 fois ?

$$\mathbb{P}[X = 0] = 0.1^5 = 0.00001 \;.$$

3. Quelle est la probabilité que l'opération rate exactement 3 fois ?

$$\mathbb{P}[X = 2] = \binom{5}{2} 0.9^2 \, 0.1^3 = 0.0081 \;.$$

4. Quelle est la probabilité que l'opération réussisse au moins 3 fois ?

$$\begin{aligned}
\mathbb{P}[X \geqslant 3] &= \mathbb{P}[X = 3] + \mathbb{P}[X = 4] + \mathbb{P}[X = 5] \\
&= \binom{5}{3} 0.9^3 \, 0.1^2 + \binom{5}{4} 0.9^4 \, 0.1^1 + \binom{5}{5} 0.9^5 \, 0.1^0 \\
&= 0.0729 + 0.32805 + 0.59049 = 0.99144 \;.
\end{aligned}$$

Exercice 3. Un groupe d'étudiants est composé de 18 filles et de 11 garçons. On choisit au hasard dans ce groupe un échantillon de 5 personnes. Soit X la variable aléatoire égale au nombre de filles dans cet échantillon.

1. Quel modèle proposez-vous pour X ?

   La loi de X est la loi hypergéométrique de paramètres $N = 29$ (nombre total d'individus), $m = 18$ (les individus "marqués" sont les filles) et $n = 5$ (la taille de l'échantillon extrait). Les valeurs prises sont les entiers de 0 à 5. Pour tout entier $k = 0, 1 \ldots, 5$, on a :

$$\mathbb{P}[X = k] = \frac{\binom{18}{k}\binom{11}{5-k}}{\binom{29}{5}} .$$

2. Donner l'espérance de X.

   L'espérance de X est $5 \times 18/29 \simeq 3.1$. C'est la taille de l'échantillon, multipliée par la proportion de filles dans le groupe.

3. Calculer la probabilité que l'échantillon ne contienne que des filles.

$$\mathbb{P}[X = 5] = \frac{\binom{18}{5}}{\binom{29}{5}} \simeq 0.072 .$$

4. Calculer la probabilité que l'échantillon contienne au moins une fille.

   On doit calculer $\mathbb{P}[X \geqslant 1]$. On pourrait calculer $\mathbb{P}[X = 1] + \mathbb{P}[X = 2] + \mathbb{P}[X = 3] + \mathbb{P}[X = 4] + \mathbb{P}[X = 5]$, mais il est plus rapide de calculer $1 - \mathbb{P}[X = 0]$, ce qui revient au même :

$$\mathbb{P}[X \geqslant 1] = 1 - \mathbb{P}[X = 0] = 1 - \frac{\binom{11}{5}}{\binom{29}{5}} \simeq 0.996 .$$

5. Calculer la probabilité que l'échantillon contienne exactement 3 filles.

$$\mathbb{P}[X = 3] = \frac{\binom{18}{3}\binom{11}{2}}{\binom{29}{5}} \simeq 0.378 .$$

Exercice 4. La taille X des hommes en France est modélisée par une loi normale $\mathcal{N}(172, 196)$ (unité : le cm).

1. Quelle proportion de français a une taille inférieure à 160 cm ?

$$\mathbb{P}[X < 160] = \mathbb{P}\left[\frac{X - 172}{\sqrt{196}} < \frac{160 - 172}{\sqrt{196}}\right] = F(-0.857) = 1 - F(0.857) = 0.1957 ,$$

   où F désigne (comme dans tout l'exercice) la fonction de répartition de la loi $\mathcal{N}(0, 1)$.

2. Quelle proportion de français mesure plus de deux mètres ?

$$\mathbb{P}[X > 200] = \mathbb{P}\left[\frac{X-172}{\sqrt{196}} > \frac{200-172}{\sqrt{196}}\right] = 1 - F(2) = 0.02275\ .$$

3. Quelle proportion des français mesure entre 165 et 185 centimètres ?

$$\begin{aligned}\mathbb{P}[165 < X < 185] &= \mathbb{P}\left[\frac{165-172}{\sqrt{196}} < \frac{X-172}{\sqrt{196}} < \frac{185-172}{\sqrt{196}}\right] \\ &= F(0.928) - F(-0.5) = 0.8234 - 0.3085 = 0.5149\ .\end{aligned}$$

4. Si on classait dix mille français choisis au hasard par ordre de taille croissante, quelle serait la taille du 9000-ième ?

   La question revient à trouver la taille telle que 90% *des français aient une taille inférieure, à savoir le quantile d'ordre* 0.9*, ou encore le neuvième décile. Soit x cette taille.*

   $$\mathbb{P}[X < x] = \mathbb{P}\left[\frac{X-172}{\sqrt{196}} < \frac{x-172}{\sqrt{196}}\right] = 0.9$$

   Donc $\frac{x-172}{\sqrt{196}}$ est la valeur de la fonction quantile de la loi $\mathcal{N}(0,1)$ au point 0.9*, à savoir* 1.2816*. On en déduit :*

   $$x = 172 + 1.2816 \times \sqrt{196} \simeq 190\ cm.$$

5. La taille Y des françaises est modélisée par une loi normale $\mathcal{N}(162, 144)$ (en centimètres). Quelle est la probabilité pour qu'un homme choisi au hasard soit plus grand qu'une femme choisie au hasard ?

   Si X désigne la taille de l'homme et Y la taille de la femme, supposées indépendantes, alors $X - Y$ suit la loi normale $\mathcal{N}(10, 340)$. La probabilité que X soit supérieure à Y est la probabilité que la différence soit positive :

   $$\mathbb{P}[X - Y > 0] = \mathbb{P}\left[\frac{(X-Y)-10}{\sqrt{340}} > \frac{0-10}{\sqrt{340}}\right] = 1 - F(-0.5423) = 0.7062\ .$$

Exercice 5. Les laboratoires pharmaceutiques indiquent pour chaque test sa sensibilité α, qui est la probabilité que le test soit positif si le sujet est malade et sa sensibilité β qui est la probabilité que le test soit négatif si le sujet est sain. Sachant qu'en moyenne il y a un malade sur 1000 personnes, quelle est la probabilité que vous soyiez un sujet sain alors que votre test est positif, avec $\alpha = 98\%$ et $\beta = 97\%$. Calculez la probabilité d'être malade alors que le test est négatif. Commentaires.

Exercice 6. Dans un lot de pièces fabriquées, il y a 5% de pièces défectueuses. On contrôle les pièces, mais le mécanisme de contrôle est aléatoire. Si la pièce est bonne, elle est acceptée avec une probabilité égale à 0.96 ; si la pièce est mauvaise, elle est refusée avec probabilité 0.98. On choisit au hasard une pièce que l'on contrôle.

1. Quelle est la probabilité que cette pièce soit refusée ?
2. Quelle est la probabilité que cette pièce soit bonne, sachant qu'elle est refusée ?
3. Quelle est la probabilité que cette pièce soit mauvaise sachant qu'elle est acceptée ?
4. Quelle est la probabilité qu'il y ait une erreur dans le contrôle (une bonne pièce est refusée ou une mauvaise est acceptée) ?

Exercice 7. On suppose que l'on a autant de chances d'avoir une fille ou un garcon à la naissance. Votre voisin de palier vous dit qu'il a 2 enfants.

1. Quelle est la probabilité qu'il aît au moins un garcon ?
2. Quelle est la probabilité qu'il aît un garcon sachant que l'ainé est une fille ?
3. Quelle est la probabilité qu'il aît un garcon sachant qu'il a au moins une fille ?
4. Vous téléphonez à votre voisin. Une fille decroche au téléphone. Vous savez que dans les familles avec un garcon et une fille, la fille décroche le téléphone avec une probabilité p. Quelle est la probabilité que votre voisin aît un garcon ?
5. Vous sonnez à la porte de votre voisin. Une fille ouvre la porte. Sachant que l'ainé(e) ouvre la porte avec une probabilité p, quelle est la probabilité que votre voisin aît un garcon ?

Exercice 8. Le gène qui détermine les yeux bleus est récessif. Pour avoir les yeux bleus il faut donc avoir le phénotype bb. Les génotypes bm et mm donnent les yeux marron. On suppose que les parents transmettent indifféremment un de leus gènes à leurs enfants La femme et la sœur d 'Adrien ont les yeux bleus mais se parents ont les yeux marron.

1. Quelle est la probabilit/'e pour qu'Adrien aît les yeux bleus ?
2. Quelle est la probabilit/'e pour que le premier enfant d'Adrien aît les yeu bleus sachant qu'Adrien a les yeux marron ?
3. Quelle est la probabilit/'e pour que le deuxième enfant d'Adrien aît les yeu bleus sachant que le premier a les yeux marron ?
4. Comment expliquez vous la différence de résultats entre les 2 dernières questions ?

Exercice 9. La France a eu 38 médailles dont 13 d'or aux jeux olympiques de Sydney en 2000. sur 928 médailles dont 301 d'or. On estime la population à 6 milliards dont 60 millions en France. Peut on dire que les sportifs de haut niveau sont uniformément répartis dans la population mondiale ?

Exercice 10. On jette 3 dés.

1. Quelle est la probabilité d'avoir exactement un 6 ?
2. Quelle est la probabilité d'avoir au moins un 6 ?
3. Quelle est la probabilité d'avoir un 6 au premier tirage puis 2 autres tirages autres que 6 ?

4. Quelle est la probabilité d'avoir un 6 au premier tirage ?
5. Quelle est la probabilité d'avoir un 6 , un 2 et un 1 ?

Exercice 11. On pose 20 questions à un candidat ; Pour chaque question, k réponses sont proposées dont une seule est la bonne. Le candidat choisit au hasard une des réponses proosées.

1. On lui attribue un point par bonne réponse. Soit X le nombre de points obtenus. Quelle est al loi de X ?
2. Lorsque le candidat donne une mauvaise réponse, il peut choisir à nouveau une des réponses proposées. On lui attribue alors 1/2 point par bonne réponse. Soit Y le nombre de points obtenus lors de ces seconds choix. Quelle est la loi de Y ?
3. Soit S le nombre total de points obtenus. Déterminez k pour que le candidat obtienne en moyenne 5 sur 20.

Exercice 12. Lors d'une séance d'identification, on propose à 6 témoins de désigner un coupable parmi 4 suspects, dont vous faites partie.

1. Si chacun des 6 témoins choisissait au hasard, quelles seraient vos chances :
   (a) de n'être jamais désigné ?
   (b) d'être désigné exactement une fois ?
   (c) d'être désigné deux fois ou plus ?
2. Il se trouve que 2 des 6 témoins vous ont désigné comme le coupable. Par référence au résultat de la question 1 (c), pensez-vous que le juge pourra attribuer cela au hasard ?
3. Et si 4 des 6 témoins vous avaient désigné ?

Exercice 13. Dans chacune des situations suivantes, on donnera la loi de probabilité de la variable aléatoire X et son espérance. On calculera la probabilité que X soit égal à 0, puis la probabilité que X soit supérieur ou égal à 2.

1. À la belote, huit cartes sont distribuées à chacun des quatre joueurs. Soit X le nombre d'as que reçoit un joueur donné.
2. À la belote, les quatre joueurs jouent par équipes de deux. Soit X le nombre de piques d'une équipe donnée.
3. Au bridge, treize cartes sont distribuées à chacun des quatre joueurs. Soit X le nombre de figures (valet, dame ou roi) d'un joueur donné.
4. Au loto, vous avez coché 6 numéros sur une grille qui en comporte 49. Soit X le nombre de bons numéros sur votre grille.

Exercice 14. Soit X une variable aléatoire de loi $\mathcal{N}(0,1)$.

1. Exprimer à l'aide de la fonction de répartition de X, puis calculer à l'aide de la table les probabilités suivantes.

(a) $\mathbb{P}[X > 1.45]$
(b) $\mathbb{P}[-1.65 \leqslant X \leqslant 1.34]$
(c) $\mathbb{P}[|X| < 2.05]$

2. Déterminer la valeur de u dans les cas suivants.
   (a) $\mathbb{P}[X < u] = 0.63$
   (b) $\mathbb{P}[X \geqslant u] = 0.63$
   (c) $\mathbb{P}[|X| < u] = 0.63$

Exercice 15. Soit X une variable aléatoire suivant la loi $\mathcal{N}(0,1)$. On pose $Y = 2X - 3$.

1. Quelle est la loi de Y ?
2. Calculer $\mathbb{P}[Y < -4]$.
3. Calculer $\mathbb{P}[-2 < Y < 3]$.

Exercice 16. Soit X une variable aléatoire de loi $\mathcal{N}(3, 25)$.

1. Exprimer à l'aide de la fonction de répartition de la loi $\mathcal{N}(0,1)$, puis calculer à l'aide de la table les probabilités suivantes.
   (a) $\mathbb{P}[X < 6]$
   (b) $\mathbb{P}[X > -2]$
   (c) $\mathbb{P}[-1 \leqslant X \leqslant 1.5]$
2. Déterminer la valeur de u dans les cas suivants.
   (a) $\mathbb{P}[X < u] = 0.63$
   (b) $\mathbb{P}[X > u] = 0.63$
   (c) $\mathbb{P}[|X - 3| \leqslant u] = 0.63$

Exercice 17. Dans un pays donné, le taux de cholestérol sérique d'un individu pris au hasard est modélisé par une loi normale avec une moyenne de 200 mg/100 mL et un écart-type de 20 mg/100 mL.

1. Quelle est la probabilité qu'un individu pris au hasard dans ce pays ait un taux de cholestérol inférieur à 160 mg/100 mL ?
2. Quelle proportion de la population a un taux de cholestérol compris entre 170 et 230 mg/100 mL ?
3. Dans un autre pays, le taux moyen de cholestérol sérique est de 190 mg/100 mL, pour le même écart-type. Reprendre les questions précédentes.
4. On choisit un individu au hasard dans le premier pays, puis dans le second. Quelle est la probabilité que le premier individu ait un taux supérieur au second ?

Exercice 18. La taille d'un épi de blé dans un champ est modélisée par une variable aléatoire X de loi normale $\mathcal{N}(15, 36)$ (unité : le cm).

1. Quelle est la probabilité pour qu'un épi ait une taille inférieure à 16 cm ?
2. On admet qu'il y a environ 15 millions d'épis dans le champ, donner une estimation du nombre d'épis de plus de 20 cm.
3. Quelle est la probabilité pour que 10 épis prélevés dans le champ aient tous leur taille dans l'intervalle $[16\ ;\ 20]$?
4. On suppose que la taille d'un épi de blé d'un autre champ est modélisée par une variable aléatoire Y de loi normale $\mathcal{N}(10, 16)$ et que X et Y sont des variables indépendantes. Quelle est la probabilité pour qu'un épi pris dans le premier champ soit plus grand qu'un épi pris dans le second ?