

# Travaux dirigés de statistiques et analyses de données

## Probabilités

### Promotion 115

---

#### Exercice 1

Soit un groupe de 3 animaux. On note :

- $G_1$  : l'évènement «L'individu 1 du groupe est vacciné»
- $G_2$  : l'évènement «L'individu 2 du groupe est vacciné»
- $G_3$  : l'évènement «L'individu 3 du groupe est vacciné»

De plus soient les deux évènements suivants :

- $E_1$  : «Au moins un individu est vacciné»
- $E_2$  : «Au moins 2 individus sont vaccinés»

1. Quels sont les évènements complémentaires de  $E_1$  et  $E_2$  ?
2. Exprimez les évènements  $E_1$  et  $E_2$  à l'aide de  $G_i$ , avec  $i \in \llbracket 1, 3 \rrbracket$ .

[stat-0019]

---

#### Exercice 2

Un nouveau vaccin a été testé sur un troupeau de 12500 vaches : 75 d'entre elles, dont 35 en gestation, ont eu des réactions secondaires nécessitant l'intervention d'un vétérinaire.

1. Sachant que ce vaccin a été administré à 680 vaches en gestation, quelle est la probabilité qu'une vache en gestation ait une réaction secondaire si elle reçoit le vaccin ?
2. Quelle est la probabilité qu'une vache qui n'est pas en gestation ait une réaction secondaire ?

[stat-0020]

---

#### Exercice 3

Dans une population, 45% des personnes interrogées déclarent jouer d'un instrument de musique, 30% avoir un ou plusieurs proches musiciens. Par ailleurs, 10% déclarent être à la fois musicien et avoir au moins un ami musicien.

Si on prend un individu au hasard dans cette population, quelle est la probabilité qu'il ne soit pas ni musicien, ni ami avec un musicien ?

[stat-0021]

---

#### Exercice 4

Un groupe de 70 canards est constitué de 40 malades et 30 non malades. Respectivement 45% des malades et 30% des non malades ne sont pas sortis du bâtiment d'élevage.

On notera

- $M$  : « Etre malade »

- B : « Être resté à l'intérieur du bâtiment »
1. Que représentent les valeurs 45% et 30% ?
  2. Quelle est la probabilité qu'un canard soit malade et qu'il ne soit pas sorti du bâtiment d'élevage ?
  3. Quelle est la probabilité qu'un canard ne soit pas sorti du bâtiment d'élevage ?

[stat-0022]

## Exercice 5

Un vétérinaire examine des chevaux pour déterminer leur aptitude à la compétition. Il constate que

- 1/3 des chevaux ont déjà été blessé,
  - parmi les chevaux qui ont déjà été blessé, la moitié présente une bonne aptitude à la compétition
  - parmi les chevaux qui n'ont jamais été blessé, 60% présentent une bonne aptitude à la compétition
1. Si on prend un cheval au hasard, quelle est la probabilité qu'il ait une bonne aptitude à la compétition ?
  2. Si un cheval a une bonne aptitude à la compétition, quelle est la probabilité pour qu'il ait déjà été blessé ?

[stat-0023]

## Exercice 6

Sur un grand nombre de naissances, on a pu estimer :

- la probabilité d'avoir un bélier :  $P(M) = 0.52$
- la probabilité d'avoir une brebis :  $P(F) = 0.48$

Pour un brebis qui a eu 3 naissances :

1. quelle est la probabilité d'avoir 3 bélier ?
2. quelle est la probabilité d'avoir au moins 1 brebis ?
3. Quelle est la probabilité d'avoir au moins 2 brebis ?

[stat-0024]

## Exercice 7

L'ensemble des animaux soignés par un cabinet vétérinaire peuvent être classés en 3 catégories :

- médecine,
- chirurgie
- et vaccination.

Le cabinet vétérinaire s'interroge sur la nécessité de recruter un vétérinaire supplémentaire spécialisé en chirurgie.

1. Les animaux domestiques soignés au cabinet se répartissent de la façon suivante : médecine 50%, chirurgie 40% et vaccination 10%. De plus, 3/4 de ces "patients" qui ont besoin d'une intervention chirurgicale sont des "patients" ayant une complication digestive et 1/4 des "patients" ont des besoins en orthopédie. Quelle est donc la proportion d'animaux domestiques faisant appel à un chirurgien orthopédiste ?
2. Chez les animaux d'élevage, la répartition des consultations est la suivante : médecine 55%, chirurgie 40% (dont la moitié en orthopédie) et vaccination 5%. Sachant qu'un "patient" (soigné par un vétérinaire du cabinet) sur 5 est un animal domestique, dire la ou les propositions vraies :

- (a) parmi les "patients", la proportion d'animaux d'élevage nécessitant une consultation orthopédique est de 4%
  - (b) parmi les "patients", la proportion d'animaux d'élevage nécessitant une consultation orthopédique est de 20%
  - (c) parmi les d'animaux d'élevage qui se font soigner par le cabinet, la proportion nécessitant une consultation en orthopédie est des 4%
  - (d) parmi les animaux d'élevage qui se font soigner par le cabinet, la proportion nécessitant une consultation en orthopédie est de 20%
  - (e) Aucune des réponses proposées ci-dessus n'est correcte
3. Quelle est la proportion d'adultes parmi les patients aux urgences faisant appel à un chirurgien orthopédiste ?

[stat-0025]

### Exercice 8

Dans une population, deux maladies M1 et M2 sont présentes respectivement chez 10% et 20% des individus (le nombre de ceux qui souffrent des deux maladies est négligeable). On entreprend un dépistage systématique des maladies M1 et M2. Pour cela, on applique un test qui réagit à la maladie sur 90% des malades de M1, sur 70% des malades de M2 et sur 10% des individus qui n'ont aucune de ces deux maladies.

1. Quand on choisit au hasard un individu de la population, quelle est la probabilité pour que le test réagisse ?
2. Sachant que pour cet individu, le test a réagi, donner les probabilités pour que ce soit à cause de la maladie M1, à cause de la maladie M2, sans qu'il n'ait l'une des deux maladies.

[stat-0026]

### Exercice 9

En période de pandémie de grippe aviaire, on considère que 10% de la population des élevages est atteinte de la grippe aviaire.

Quelle est la probabilité que strictement plus de 20 volailles soient malades dans un élevage de 100 individus ?

[stat-0028]

### Exercice 10

On suppose que la production laitière des vaches d'une exploitation agricole donnée est distribuée normalement, avec une moyenne de 7240 L/an et un écart type de 3 L/an. On mesure la production laitière sur une année d'une vache choisie au hasard. Quelle est la probabilité que sa production soit :

1. inférieure à 7236 L/an
2. supérieure à 7243 L/an
3. comprise entre 7233 et 7241 g/L
4. À partir de quelle valeur, une vache a-t-elle une production dans les 20% les plus élevée ?

[stat-0029]

### Exercice 11

Le cheptel bovin français comptait 16,8 millions de têtes en 2023. Dans ce cheptel, il y a environ 20% de vaches allaitantes. On tire au sort un échantillon de 100 individus et on compte le nombre de vaches allaitantes dans l'échantillon. Soit  $X$  la variable aléatoire représentant le nombre de vaches allaitantes dans l'échantillon. Quelle(s) proposition(s) est(sont) vraie(s) ?

1.  $X$  suit une loi binomiale
2.  $X$  suit une loi de Bernoulli
3.  $X$  suit approximativement une loi normale
4. La probabilité d'observer strictement plus de 26 vaches allaitantes dans l'échantillon vaut environ 2.3%
5. La probabilité d'observer moins de 26 vaches allaitantes dans l'échantillon vaut environ 2.3%

[stat-0030]

## Exercice 12

Dans une région sahélienne, le rendement du mil suit une distribution normale, de moyenne 780 kg/ha et d'écart-type 120 kg/ha. On prélève un échantillon aléatoire simple de  $n = 50$  exploitations.

1. *Probabilité pour la moyenne de l'échantillon*
  - (a) Calculer la probabilité que la moyenne de cet échantillon dépasse 820 kg/ha.
  - (b) Calculer la probabilité que la moyenne de cet échantillon soit comprise entre 770 kg/ha et 805 kg/ha.
2. *Détermination de la taille d'échantillon*

On souhaite déterminer la taille d'un échantillon permettant, si c'est possible, d'obtenir une probabilité de 80% pour un rendement moyen compris entre 775 kg/ha et 785 kg/ha.

Calculer cette taille d'échantillon.
3. *Proportion d'exploitations avec engrais subventionnés*

Selon un rapport antérieur, 40% des exploitations produisant du mil ont accès à des engrais subventionnés dans la région concernée. On conserve l'échantillon aléatoire simple de  $n = 50$  exploitations.

Calculer la probabilité que la proportion observée dans l'échantillon soit inférieure à 30%.
4. *Types d'échantillonnage*
  - (a) Citer les deux grands types d'échantillonnage.
  - (b) Pour chacun de ces types, donner deux méthodes d'échantillonnage.

[stat-0033]

## Exercice 13

Au Bénin, la production de coton est organisée en coopératives. Une étude souhaite étudier les rendements des exploitants. On a relevé les rendements (en kg/ha) de 240 exploitants appartenant à 11 coopératives différentes. De plus, on a noté si chaque exploitant utilise de la matière organique (MO) ou non. On suppose que le rendement moyen suit une loi normale.

Rendement (kg/ ha)	Nombre d'exploitants	Nombre d'exploitants utilisant MO
500-599	20	12
600-699	35	20
700-799	50	30
800-899	55	35
900-999	40	25
1000-1099	25	15
1100-1199	15	10
Total	240	147

On considère que cet échantillon est représentatif de la population d'exploitants de coton.

1. Donner la nature des variables *Rendement* et *Utilisation MO*.
2. **(BONUS)** Tracer le boxplot du rendement, également appelé "boîte à moustaches", en distinguant les exploitants ayant utilisé de la matière organique (MO) de ceux ne l'ayant pas utilisée.
3. Proposer un graphique permettant de visualiser l'effet de l'utilisation de MO sur le rendement moyen. Justifiez votre choix.
4. Calculer la probabilité qu'un exploitant de l'échantillon ait un rendement compris entre 800 et 899 kg/ha et qu'il ait utilisé de la MO.
5. Calculer la moyenne et l'écart type des rendements des 240 exploitants.
6. Calculer la proportion d'exploitants utilisant de la matière organique.
7. *Estimation ponctuelle*  
Estimer la moyenne et l'écart type des rendements de production de coton au Bénin. On donnera également une estimation de la proportion des exploitants qui utilisent de la matière organique.
8. *Intervalle de confiance*  
En supposant que les rendements sont normalement distribués, construire un intervalle de confiance à 90% pour :
  - (a) La moyenne des rendements.
  - (b) La proportion d'exploitants utilisant la matière organique.
9. *Intervalle de confiance en rapport avec les données nationales*  
On sait, d'après les données nationales, que l'écart-type du rendement moyen de production du coton au Bénin est 180 kg/ha. Construire un intervalle de confiance à 80% pour le rendement moyen de production du coton, en utilisant les données nationales.

[stat-0034]

## Exercice 14 Production de riz dans la vallée du fleuve Sénégal

Dans la vallée du fleuve Sénégal, le riz irrigué constitue une culture stratégique pour la sécurité alimentaire. Une étude menée par un service agricole régional s'intéresse aux rendements et aux pratiques culturales des exploitations rizicoles.

On admet que le rendement annuel en riz paddy par exploitation suit une loi normale de moyenne 4800 kg/ha et d'écart-type 600 kg/ha.

Un échantillon aléatoire simple de 76 exploitations est sélectionné.

1. *Analyse du rendement moyen*  
On note  $\bar{X}$  la moyenne des rendements observés dans l'échantillon.
  - (a) Calculer l'écart-type de la moyenne  $\bar{X}$ .
  - (b) Calculer la probabilité que la moyenne soit inférieure à 4600 kg/ha.
  - (c) Calculer la probabilité que la moyenne soit comprise entre 4700 kg / ha et 4900 kg/ha.
2. *Fiabilité de l'estimation du rendement*  
Les autorités agricoles souhaitent disposer d'une estimation fiable du rendement moyen. Déterminer la taille minimale de l'échantillon nécessaire pour que la probabilité que la moyenne soit comprise entre 4750 kg/ha et 4850 kg/ha soit égale à 95%.
3. *Adoption des engrais minéraux* Une enquête antérieure indique que 65% des exploitations rizicoles de la vallée utilisent des engrais minéraux. Calculer alors la probabilité que la proportion observée dans l'échantillon soit supérieure à 75%.

#### 4. Choix du plan d'échantillonnage

Pour améliorer la qualité des futures enquêtes agricoles, identifier les deux grandes familles de méthodes d'échantillonnage, et proposer une méthode d'échantillonnage adaptée aux périmètres irrigués, et une méthode adaptée aux villages agricoles dispersés de la vallée.

[stat-0031]

### Exercice 15 Culture de l'arachide au Sénégal

Dans le bassin arachidier du Sénégal, la culture de l'arachide joue un rôle majeur dans les revenus des exploitations familiales. Une étude vise à analyser les rendements des exploitants et l'effet de l'utilisation de semences certifiées (SC) sur la production.

Les rendements (en kg/ha) de 300 exploitants appartenant à 15 villages agricoles ont été relevés. Pour chaque exploitant, on a noté s'il utilise des semences certifiées (SC) ou non. On suppose que le rendement moyen suit une loi normale.

Rendement (kg/ha)	Nombre d'exploitants	Nombre utilisant SC
400–499	30	10
500–599	45	18
600–699	70	35
700–799	65	38
800–899	50	32
900–999	25	18
1000–1099	15	11
<b>Total</b>	<b>300</b>	<b>162</b>

On considère que cet échantillon est représentatif de l'ensemble des exploitants d'arachide du bassin arachidier.

#### 1. Description des variables

- Donner la nature de la variable *Rendement*.
- Donner la nature de la variable *Utilisation des semences certifiées*.

#### 2. Analyse conjointe rendement – pratiques culturales

Calculer la probabilité qu'un exploitant choisi au hasard ait un rendement compris entre 700 et 799 kg/ha et qu'il utilise des semences certifiées.

#### 3. Caractéristiques statistiques des rendements

- Calculer la moyenne des rendements des 300 exploitants.
- Calculer l'écart-type des rendements.

#### 4. Adoption des semences certifiées

- Calculer la proportion d'exploitants utilisant des semences certifiées.
- Interpréter ce résultat dans le contexte de la diffusion des innovations agricoles. (à vous de jouer !)

#### 5. Estimation ponctuelle

À partir de l'échantillon, donner :

- une estimation ponctuelle du rendement moyen et de l'écart type de l'arachide dans le bassin arachidier
- une estimation ponctuelle de la proportion d'exploitants utilisant des semences certifiées.

#### 6. Intervalles de confiance

En supposant que les rendements suivent une loi normale, construire un intervalle de confiance à 95% pour :

- (a) le rendement moyen de l'arachide ;
- (b) la proportion d'exploitants utilisant des semences certifiées.

7. *Lien avec les données nationales*

D'après les statistiques nationales, l'écart-type du rendement de l'arachide au Sénégal est estimé à 150 kg/ha. Construire un intervalle de confiance à 90% pour le rendement moyen de l'arachide dans le bassin arachidier, en utilisant cet écart-type national.

---

[stat-0032]