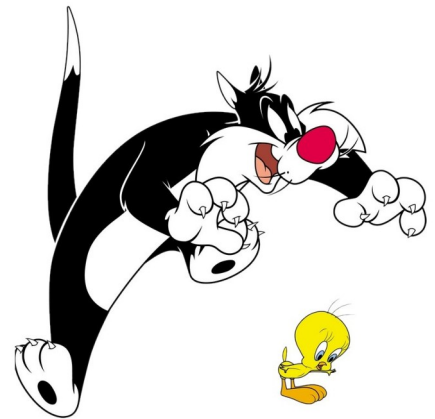


**Introduction aux logiciels pour les statistiques (IS2A3) - 2024/2025****Examen**

*Le but de l'examen est de vous permettre de montrer vos acquis. Le sujet est volontairement long. Il est donc conseillé de traiter en premier lieu les questions sur lesquelles vous vous sentez le plus à l'aise. Les langages R et Python sont les seuls autorisés pour répondre au sujet. Vous répondrez à chaque question en utilisant au moins une ligne de code et en formulant au moins une phrase de réponse. Recopier une valeur lue visuellement dans la table de données ou l'environnement pour répondre à une question ne rapporte aucun point. Une valeur de la table ne pourra être utilisée qu'à condition de l'avoir trouvée par du code. Pour répondre à une question, vous pouvez admettre le résultat de la question précédente à condition de le mentionner. La clarté des commentaires et la qualité du code constitueront un élément important dans l'appréciation de votre travail. Pour chaque question, vous êtes invité(e) à faire figurer toute trace de recherche même infructueuse.*

**Partie 1 : Le musée des oiseaux (R)**

En pleine nuit, la banque de Polytech's town a été cambriolée. Heureusement l'équipe de détectives – composée de Mémé, Sylvestre et Titi – a suivi les traces du voleur ! Ce dernier s'est réfugié dans le musée municipal. On suppose qu'il comporte 3 immenses salles A, B et C contenant chacune des oiseaux empaillés. On considère que le contenu de la salle A est représentatif de l'ensemble des oiseaux du musée. Sylvestre explore la salle A tandis que Mémé et Titi explorent les salles B et C. Tout en cherchant le voleur, Sylvestre admire les oiseaux dans les vitrines car il en raffole...



1. Combien d'oiseaux sont décrits dans la table ?
2. Quel est le nom latin de l'oiseau ayant pour numéro d'inventaire "MDG-O-02030" ?
- 3 Sylvestre sait que le véritable nom latin de la Guifette noire est en réalité *Titi Polytechus IS2Aius*. Effectuez la modification dans la table.
4. Sylvestre souhaite partir en vacances dans la ville où a été découvert le plus d'oiseaux. Dans quelle ville doit-il partir ?
5. Sylvestre souhaite faire une belle photo souvenir avec des oiseaux en arrière-plan. Il hésite entre poser devant la vitrine avec la famille *Linaria cannabina* et celle avec la famille *Haematopus ostralegus*. Existe-il plus d'oiseaux de la famille *Linaria cannabina* en mauvais état que ceux de la famille *Haematopus ostralegus* ? En déduire le conseil à donner à Sylvestre.
6. Représentez graphiquement la variable *etat\_de\_conservation* le plus pertinemment possible. Expliquez en une ou deux phrases simples à Sylvestre l'enseignement à tirer du graphique.
7. Sylvestre souhaite manger les oiseaux – s'ils étaient non empaillés – avec des champignons, de la sauce ou du fromage.
  - a) Créez un vecteur *ingredients* avec 700 valeurs aléatoires parmi *champignons*, *sauce* et *fromage*. On suppose le tirage des modalités équiprobable.
  - b) Ajoutez le vecteur *ingredients* à la table data.

8. Selon Sylvestre, la qualité gustative d'un oiseau dans la nature est proportionnelle au prix de ses pairs empaillés acquis par le musée.
- Combien d'oiseaux le musée a-t-il acheté plus cher que la valeur estimée ?
  - Quelles hypothèses formuleriez-vous pour expliquer ce résultat ?
  - Quel est le prix d'achat médian d'un oiseau par famille d'oiseau ?
  - En déduire la famille pour laquelle le prix médian est le moins élevé.
  - Expliquez en 1 ou 2 phrases le résultat obtenu à Sylvestre.
  - En déduire des noms d'oiseaux que Sylvestre gagnerait à ne pas manger.
9. Calculez de 2 façons différentes la variance non corrigée de la variable *prix\_achat*.
10. On rappelle que le contenu de la salle A est représentatif de l'ensemble des oiseaux du musée.
- Donnez une estimation de la variance du prix d'achat pour l'ensemble des oiseaux du musée.
  - Donnez une estimation de la moyenne du prix d'achat pour l'ensemble des oiseaux du musée.
  - Conseillerez-vous à Sylvestre d'explorer les autres salles du musée pour identifier les oiseaux qui pourraient lui servir de casse-croûte?
- Bonus :** Combien d'oiseaux ont été identifiés entre 1869 et 1883 ?

## Partie 2 : Course poursuite (R)

Soudain un cri retentit dans le musée. Titi a repéré le voleur ! Malheureusement ce dernier saute par la fenêtre du rez-de-chaussée et s'enfuit en scooter. Qu'à cela ne tienne ! Sylvestre et Titi rejoignent Mémé dans sa voiture. Ils se lancent ensemble à la poursuite du fuyard.



- La vitesse de la voiture de Mémé est donnée par la formule en annexe.
  - Écrivez une fonction  $v(d,t,k)$  qui renvoie la vitesse de la voiture de Mémé.
  - Testez la fonction pour  $d=2$ ,  $t=10$  et  $k=5$ .
- Exprimez  $d$  en fonction de  $t$ ,  $v$  et  $k$ .
  - Écrivez une fonction  $d(t,v,k)$  qui renvoie la distance parcourue par la voiture de Mémé.
  - Testez la fonction pour  $t=10$ ,  $v=0$  et  $k=5$ .
  - On fixe  $t=2$  et  $v=2$ . On fait donc varier uniquement  $k$ . Représentez  $d$  en fonction de  $k$ .
  - En déduire si Mémé doit prendre le chemin passant par le marais ( $k=6$ ) où celui vers la campagne ( $k=4$ ) pour rejoindre rapidement le voleur.

Le voleur s'est réfugié au rez-de-chaussée d'une maison abandonnée. Mémé, Sylvestre et Titi sont au centre du couloir qui comprend 4 portes. Chacun ouvre simultanément une porte. Si l'un d'entre eux voit le voleur derrière la porte alors il crie. Ensuite Mémé, Sylvestre et Titi se remettent au centre de la pièce et l'expérience recommence pour un nouvel essai. Le temps étant long dans le monde des dessins animés, l'équipe de détectives effectue 100 essais !

- On note  $X$  la variable aléatoire donnant le résultat d'un essai.
  - Justifiez que  $X$  suit une loi binomiale de paramètres  $p=0.75$  et  $n=100$ .
  - Quelle fonction déjà implémentée permet calculer la valeur de loi de densité en un point  $x$  ?
  - Écrivez une fonction *maFonction* équivalente à cette fonction et la tester.
  - Représentez la loi de densité de  $X$  en veillant à la qualité du graphique obtenu.
  - Quelle est la probabilité que moins de 10 essais soient nécessaires pour attraper le voleur ?
  - Quelle est la probabilité qu'entre 5 et 22 essais soient nécessaires pour attraper le voleur ?
  - Déterminez le réel  $x$  vérifiant  $P(X \leq x)=0.98$ .

3. Alors que Mémé vient d'ouvrir la porte derrière laquelle se trouve le voleur, ce dernier s'enfuit encore une fois par la fenêtre en courant. Il porte un sac de la banque sur son dos mais en a abandonné

un près de la fenêtre. Sylvestre veut vérifier que les pièces dans ce sac sont bien en or. Si elles ne respectent pas la règle du Bureau de l'or, alors les pièces sont fausses. Pour la quantité d'or présent dans une pièce, la moyenne de référence est de 6.8 microgrammes par gramme. Sylvestre prélève 9 pièces dans le sac et mesure la teneur en or dans chacune d'entre elle. Les résultats sont : 6.7 – 7.2 – 7.5 – 9.3 – 6.8 – 7.0 – 30.1 – 7.7 – 7.6. On suppose que la teneur d'or microgramme (mg) présent dans une pièce peut être modélisée par une variable aléatoire  $X$  suivant une loi normale. On fixe  $\alpha=5\%$ . La problématique est la suivante : Peut-on affirmer, avec un faible risque de se tromper, que les pièces du sac ne sont pas conformes à la règle concernant la teneur en or ?

### Partie 3 : Identité du voleur (R ou Python)

Le voleur court avec son sac aussi vite qu'il peut avec Mémé, Sylvestre et Titi sur les talons. Or le sac du voleur est craqué. Il perd par voie de conséquence des pièces sur la route.

1. Le nombre de pièce perdues au bout de  $n$  enjambées est donné par la suite en annexe.
  - a) Combien de pièces le voleur aura perdu au bout de 12 enjambées ?
  - b) Au bout de combien d'enjambées le voleur aura-t-il perdu la totalité des 2024 pièces du trésor ?
2. Dans une clairière, le voleur trébuche *in fine* sur la branche d'un arbre et s'étale sur le sol. Mémé, Sylvestre, Titi et un policier décide d'encercler le voleur. On représente la clairière par une matrice carrée de côté  $n$  (où  $n$  est un entier impair).
  - a) Écrivez une fonction *scenario*( $n$ ) qui renvoie la matrice avec :
    - *au centre le voleur (V)*
    - *en haut à gauche : titi (T)*
    - *en haut à droite : sylvestre (S)*
    - *en bas à gauche : mémé (M)*
    - *en bas à droite : un policier (P)*
 On représentera les personnages par les lettres entre parenthèses.
    - b) Que pensez-vous de la modélisation avec la matrice ?
3. Le nom du voleur n'est autre que celui ayant le plus long (espaces compris) dans la liste des personnages connus. Écrivez une fonction *voleur* qui prend en paramètre la liste des personnages et renvoie le nom du voleur.



#### Annexe

- Formule de la vitesse  $v$  de la voiture de Mémé :  $v(d, t, k) = \sqrt{\ln\left(\frac{d \cdot k}{t}\right)}$ 
  - $d$  : la distance parcourue par la voiture en km.
  - $t$  : le temps en heure
  - $k$  : le coefficient *toon* qui varie selon le terrain sur lequel roule la voiture
- Suite donnant le nombre de pièces perdues au bout de  $n$  enjambées :

$$\begin{cases} p_0 = 8 \\ p_{n+1} = p_n + 2n - 10 \end{cases}$$