

# Epreuve d'admissibilité de mathématiques et statistiques

Concours Externe INSEE 2020

## Exercice 1 (3 points)

*Les deux parties sont indépendantes*

### Partie 1

La compagnie Frazul commercialise un vol Marseille-Lille avec une escale à Paris. Durant cette escale, certains passagers descendent mais aucun passager ne monte dans l'avion. Le billet coûte 90€ jusqu'à Paris et 30€ de plus pour aller jusqu'à Lille. Au total, aujourd'hui, 185 personnes sont montées dans l'avion à Marseille et le vol a rapporté 21000€ à la compagnie. Combien de passagers sont descendus à Paris ?

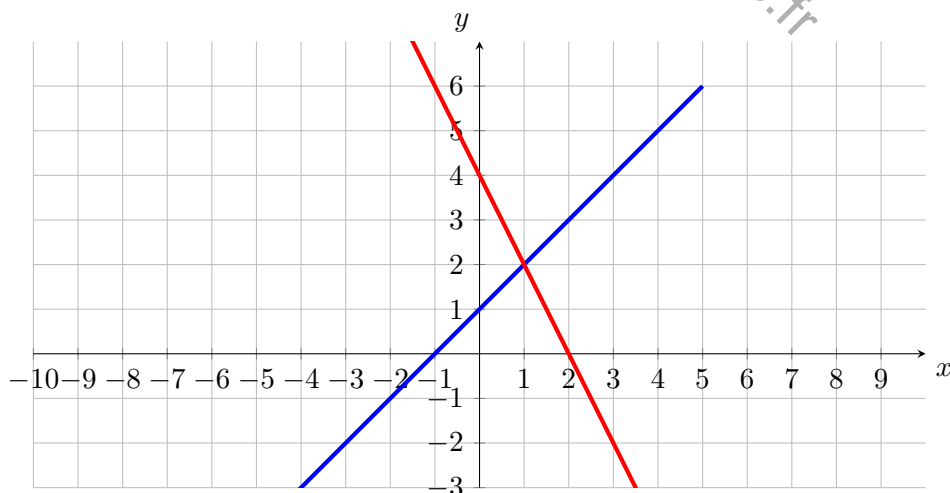
Soit  $x$  le nombre de passagers descendus à Paris. Nous avons donc  $185 - x$  personnes qui sont allées jusqu'à Lille.

Le prix total des billets pour les passagers descendus à Paris est de  $90 \times x$  ; celui pour les passagers descendus à Lille est de  $120 \times (185 - x)$ . Ainsi,

$$\begin{aligned} 90x + 120 \times (185 - x) &= 21000 \\ 90x + 22200 - 120x &= 21000 \\ -30x &= -1200 \\ x &= 40 \end{aligned}$$

40 passagers sont descendus à Paris aujourd'hui.

### Partie 2



1. Écrire les équations des deux droites présentées dans le graphique

Il suffit de lire le graphique pour obtenir ordonnées à l'origine et coefficients directeurs : l'équation de la droite rouge est  $y = -2x + 4$ . L'équation de la droite bleue est  $y = x + 1$ .

2. Déterminer, par le calcul, les coordonnées du point d'intersection I des deux droites.

Les deux équations de droites donnent la même ordonnée à l'intersection. Ainsi,

$$\begin{aligned}x_I + 1 &= -2x_I + 4 \\3x_I &= 3 \\x_I &= 1\end{aligned}$$

On obtient l'ordonnée en remplaçant  $x_I$  dans l'équation de l'une des deux droites et l'on obtient  $y_I = 2$ . Le point d'intersection I des deux droites a pour coordonnées  $(1; 2)$ .

## Exercice 2 (4 points)

*Dans l'ensemble de l'exercice, on ne demande pas de justifier les réponses, et une seule proposition est exacte.*

Un laboratoire souhaite expérimenter un nouveau vaccin contre une maladie contagieuse. Pour cela, elle réalise un protocole de test auprès de la population de Zombiland en vaccinant le quart de la population. Au cours de l'épidémie, elle constate qu'il y a :

- parmi les malades, un vacciné pour 4 non vaccinés
- un malade sur dix parmi les vaccinés

On note  $V$  l'évènement "être vacciné". On note  $\bar{V}$  l'évènement "ne pas être vacciné". On note  $M$  l'évènement "être malade". On note  $\bar{M}$  l'évènement "ne pas être malade". On note  $P(M/V)$  la probabilité que la personne "soit malade sachant qu'elle a été vaccinée".

1. Dans l'énoncé,  $\frac{1}{10}$  correspond à ?

La fraction  $\frac{1}{10}$  correspond (d'après l'énoncé) à la probabilité d'être malade sachant que l'on est vacciné, ce qui se note  $P(M/V)$ , qui est la réponse E dans l'énoncé.

2. En vous appuyant sur l'énoncé, quelle est l'expression correcte ?

D'après l'énoncé, il y a parmi les malades un vacciné pour 4 non vaccinés. Ainsi, la probabilité  $P(\bar{V}/M)$  d'être non vacciné sachant que l'on est malade est quatre fois supérieure à la probabilité  $P(V/M)$  d'être vacciné sachant que l'on est malade. Ainsi,  $P(\bar{V}/M) = 4 \times (V/M)$ . il s'agit de la réponse A dans l'énoncé.

3. Quelle est la probabilité  $P(V/M)$  ?

Sur 5 malades, 1 est vacciné et 4 ne le sont pas. Ainsi,  $P(V/M) = \frac{1}{5}$ , ce qui correspond à la réponse B dans l'énoncé.

4. Comment peut-on écrire  $P(M)$  ?

Il s'agit d'utiliser la formule des probabilités totales :  $P(M) = P(M \cap V) + P(M \cap \bar{V})$ .  
Réponse B.

5. Quelle est l'expression correcte pour  $P(M/\bar{V})$  ?

Par définition,  $P(M/\bar{V}) = \frac{P(M \cap \bar{V})}{P(\bar{V})}$ . Il s'agit de la réponse A.

6. Quelle est l'expression correcte pour  $P(M/\bar{V})$  en fonction de  $P(M)$  ?

L'énoncé nous donne plusieurs informations. On sait que la probabilité de ne pas être vacciné sachant que l'on est malade est de  $\frac{4}{5}$ , soit  $P(\bar{V}/M) = \frac{4}{5}$ .

Or on sait, par définition, que  $P(\bar{V}/M) = \frac{P(\bar{V} \cap M)}{P(M)}$ . De plus,  $P(\bar{V} \cap M) = P(M \cap \bar{V})$ . Donc  $P(M \cap \bar{V}) = P(\bar{V}/M) \times P(M) = \frac{4}{5}P(M)$ .

De plus, on sait que seul un quart de la population a été vaccinée, c'est à dire que  $P(\bar{V}) = \frac{3}{4}$ .

En remplaçant ces deux points dans la définition trouvée à la question 5, on obtient que :

$$\begin{aligned} P(M/\bar{V}) &= \frac{P(M \cap \bar{V})}{P(\bar{V})} \\ &= \frac{\frac{4}{5}P(M)}{\frac{3}{4}} \\ &= \frac{4}{3} \times \frac{4}{5}P(M) \\ &= \frac{16}{15}P(M) \end{aligned}$$

Réponse C.

7. Quelle est la probabilité d'être malade ?

La probabilité d'être malade et d'être vacciné correspond à  $P(M \cap V) = P(V) \times P(M/V) = \frac{1}{4} \times \frac{1}{10} = \frac{1}{40}$ . En reprenant la formule des probabilités totales énoncée à la question 4 on obtient :

$$\begin{aligned} P(M) &= P(M \cap V) + P(M \cap \bar{V}) \\ P(M) &= \frac{1}{40} + \frac{4}{5}P(M) \\ \frac{1}{5}P(M) &= \frac{1}{40} \\ P(M) &= \frac{1}{8} \end{aligned}$$

Réponse C

Remarque : La réponse B ( $\frac{5}{40}$ ) est techniquement correcte, mais il est indiquée qu'une seule réponse est la bonne, donc autant prendre la fraction irréductible...

8. Quelle est la probabilité d'être malade sachant que l'on n'est pas vacciné ?

Il suffit de reprendre les résultats des deux questions précédentes, et l'on trouve  $P(M/\bar{V}) = \frac{2}{15}$ , ce qui correspond à la réponse D.

### Exercice 3 (5 points)

Soit  $f$  la fonction définie par  $f(x) = x + 1 + \frac{2}{e^x + 1}$ .

1. Donner l'ensemble de définition de cette fonction, on appellera cet ensemble  $Df$ .

- La fonction  $x \mapsto x + 1$  est définie sur  $\mathbb{R}$  par propriété des fonctions affines.
- La fonction  $x \mapsto e^x$  est strictement positive et ne s'annule donc jamais sur  $\mathbb{R}$  par propriété de l'exponentielle, donc en particulier la fonction  $x \mapsto e^x + 1$  ne s'annule pas sur  $\mathbb{R}$ .
- La fonction  $x \mapsto \frac{2}{e^x + 1}$  est définie sur  $\mathbb{R}$  comme quotient de fonctions définies sur  $\mathbb{R}$  dont le dénominateur ne s'annule pas sur  $\mathbb{R}$ .
- Conclusion : La fonction  $f$  est définie sur  $\mathbb{R}$  comme somme de fonctions définies sur  $\mathbb{R}$ .  $Df = \mathbb{R}$ .

2. Déterminer les limites aux bornes de  $Df$ .

#### Limite en $+\infty$

On sait que  $\lim_{x \rightarrow +\infty} e^x = +\infty$  donc  $\lim_{x \rightarrow +\infty} \frac{2}{e^x + 1} = 0$ .

De plus,  $\lim_{x \rightarrow +\infty} x + 1 = +\infty$

D'où  $\lim_{x \rightarrow +\infty} f(x) = +\infty$

#### Limite en $-\infty$

On sait que  $\lim_{x \rightarrow -\infty} e^x = 0$  donc  $\lim_{x \rightarrow -\infty} \frac{2}{e^x + 1} = 2$ .

De plus,  $\lim_{x \rightarrow -\infty} x + 1 = -\infty$

D'où  $\lim_{x \rightarrow -\infty} f(x) = -\infty$

3. Étudier le sens de variations de  $f$  et dresser son tableau de variations.

Si l'on souhaite étudier les variations de  $f$ , sa dérivée est un outil puissant. Une étude analogue à celle de la question 1. en remplaçant "définie" par "dérivable" nous assure de la dérivabilité de  $f$  sur  $Df$ .

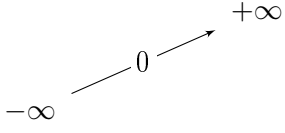
On sait que  $(x + 1)' = 1$ .

De plus,  $x \mapsto \frac{1}{e^x + 1}$  est une fonction de la forme  $\frac{1}{u}$  de dérivée  $-\frac{u'}{u^2}$  avec  $u(x) = e^x + 1$  et  $u'(x) = e^x$ .

Conclusion :  $f'(x) = 1 - \frac{2e^x}{(e^x+1)^2}$ .

Remarquons que  $(e^x+1)^2 = e^{2x} + 2e^x + 1 > 2e^x$  donc pour tout  $x$  réel,  $\frac{2e^x}{(e^x+1)^2} < 1$ .

Conclusion :  $f'$  est strictement positive sur  $Df$ . Dressons maintenant le tableau de variations de  $f$ .

$x$	$-\infty$ $\approx -2,9$ $+\infty$
signe de $f'$	+
variations de $f$	

4. Démontrer que  $A(0;2)$  est centre de symétrie à la courbe représentative de  $f$

L'énoncé rappelle la propriété suivante :

**Propriété 1.** — Soit  $(a, b) \in \mathbb{R}^2$  et  $x \in Df$  tel que  $(a+x) \in Df$  et  $(a-x) \in Df$ .  
Si  $f(a-x) + f(a+x) = 2b$ , alors le point de coordonnées  $(a; b)$  est un centre de symétrie de la courbe représentative de  $f$ .

Vérifions la formule pour  $a = 0$  et  $b = 2$ .

$$\begin{aligned}
 f(a-x) + f(a+x) &= f(-x) + f(x) \\
 &= -x + 1 + \frac{2}{e^{-x}+1} + x + 1 + \frac{2}{e^x+1} \\
 &= 2 \left( 1 + \frac{e^x+1 + e^{-x}+1}{(e^x+1)(e^{-x}+1)} \right) \\
 &= 2 \left( 1 + \frac{(e^x+1)(e^{-x}+1)}{(e^x+1)(e^{-x}+1)} \right) \\
 &= 4 \\
 f(a-x) + f(a+x) &= 2 \times b
 \end{aligned}$$

Conclusion : Le point  $A(0;2)$  est centre de symétrie à la courbe représentative de  $f$ .

5. Démontrer que l'équation  $y = x + 1$  et la droite  $y = x + 3$  sont des asymptotes à la courbe représentative de la fonction  $f$

$$\lim_{x \rightarrow +\infty} f(x) - (x+1) = \lim_{x \rightarrow +\infty} \frac{2}{e^x+1} = 0$$

Conclusion : la droite d'équation  $y = x + 1$  est une asymptote oblique à la courbe représentative de  $f$  en  $+\infty$

De même  $\lim_{x \rightarrow -\infty} f(x) - (x+3) = \lim_{x \rightarrow -\infty} x+1 + \frac{2}{e^x+1} - (x+3) = \lim_{x \rightarrow -\infty} \frac{2}{e^x+1} - 2 = 0$

Conclusion : la droite d'équation  $y = x + 3$  est une asymptote oblique à la courbe représentative de  $f$  en  $-\infty$ .

6. Étudier les positions relatives de la courbe représentative de  $f$  et de ses asymptotes.

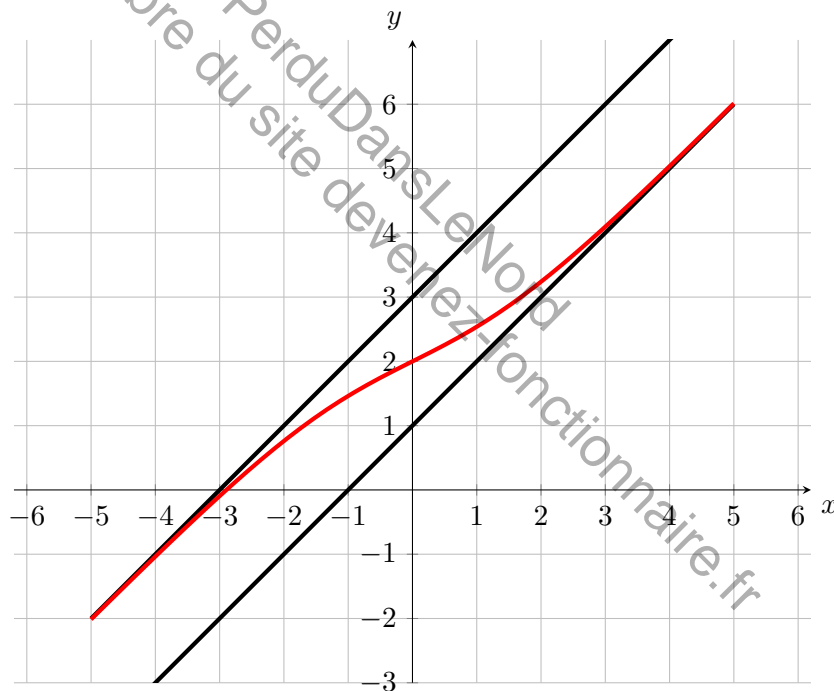
$$\lim_{x \rightarrow +\infty} \frac{2}{e^x+1} = 0^+$$

La courbe s'approche de son asymptote par valeurs positives, elle est donc située au-dessus de son asymptote en  $+\infty$ .

de même,  $\lim_{x \rightarrow -\infty} \frac{2}{e^x+1} = 2^-$ , donc  $\lim_{x \rightarrow -\infty} \frac{2}{e^x+1} - 2 = 0^-$

La courbe s'approche de son asymptote par valeurs négatives, elle est donc située en dessous de son asymptote en  $-\infty$ .

7. Tracer  $f$  et ses asymptotes



8. Que représente  $\int_0^3 |x+3 - f(x)| dx$  ?

Cette intégrale représente l'aire de la figure délimitée par les droites d'équations  $y = x + 3$ ,  $x = 0$ ,  $x = 3$  et la courbe représentative de la fonction  $f$ .

## Exercice 4 (3,5 points)

Sophie souhaiterait réaliser son arbre généalogique mais avant de commencer, elle veut se faire une idée du nombre d'ancêtres qu'elle pourrait retrouver.

Elle part du principe qu'elle a 2 parents, 4 grands-parents, 8 arrière grands-parents, etc.

En supposant que nous appartenons à la génération 1, que nos parents appartiennent à la génération 2, nos grands-parents à la génération 3, etc, on note  $(U_n)$  la suite représentant le nombre d'individus à la génération  $n$  (avec  $n$  un entier strictement positif).

1. Exprimer  $U_{n+1}$  en fonction de  $U_n$

Remarque : L'énoncé fait la confusion entre la suite  $(U_n)$  et le terme de rang  $n$  noté  $U_n$ , ce qui est corrigé ici.

Si chaque individu au rang  $n$  possède deux parents, on a directement que  $U_{n+1} = 2 \times U_n$ .

2. La suite est-elle géométrique ou arithmétique ? Vous justifierez la réponse en donnant les paramètres de cette suite.

Il s'agit par définition d'une suite géométrique de raison  $q = 2$  et de premier terme  $U_1 = 1$ .

3. Exprimer  $U_n$  en fonction de  $n$ .

Par propriété des suites géométriques (ou par récurrence immédiate) on a que  $U_n = 2^{n-1}$ .

4. Combien d'ancêtres de Sophie figurent à la 5ème génération ?

En utilisant la formule que l'on vient de trouver,  $U_5 = 2^{5-1} = 2^4 = 16$ . Sophie a 16 ascendants figurant à la 5ème génération.

5. Combien de personnes l'arbre généalogique de Sophie comporterait si l'on remontait jusqu'à la 20ème génération ? Selon qu'il s'agit d'une suite géométrique ou arithmétique, vous utiliserez la formule des sommes des termes d'une suite arithmétique ou géométrique.

On utilise la formule conseillée : la somme  $S_n$  des  $n$  premiers termes d'une suite géométrique de raison  $q$  et de premier terme  $U_1$  peut se calculer à l'aide de la formule

$$S_n = U_1 \frac{1 - q^n}{1 - q}$$

D'où ici

$$\begin{aligned}
S_{20} &= U_1 \frac{1 - q^{20}}{1 - q} \\
&= 1 \times \frac{1 - 2^{20}}{1 - 2} \\
S_{20} &= 1048575
\end{aligned}$$

En remontant à 20 générations, l'arbre de Sophie comporterait un peu plus d'un millions d'ancêtres.

**6.** Quelle(s) hypothèse(s) implicite(s) dans l'énoncé et en matière de consanguinité a rendu possible vos calculs ?

La principale hypothèse est que tous les ancêtres sont distincts les uns des autres et que donc il n'y a aucune consanguinité à aucun degré entre les ascendants de Sophie.

### Exercice 5 (4,5 points)

Le tableau suivant indique la population de l'île Tagada d'après les recensements depuis 1946.

Année	Rang $x_i$ de l'année $n$	Population $y_i$ (nombre d'habitants)
1946	0	80878
1954	8	85412
1962	16	92850
1968	22	99424
1975	29	105184
1982	36	108670
1990	44	113230
1999	53	116832

Les nombres à déterminer seront arrondis à trois décimales.

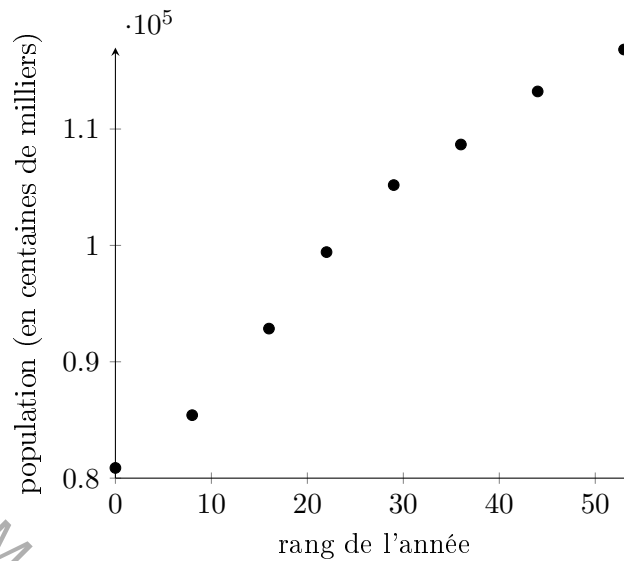
**1.** Quel est le taux d'évolution annuel moyen entre 1946 et 1954 et entre 1990 et 1999 ?

$$\begin{aligned}
\tau_{46 \rightarrow 54} &= \frac{85412 - 80878}{8 - 0} = 566,750. \\
\tau_{90 \rightarrow 99} &= \frac{116832 - 113230}{53 - 44} = 400,222
\end{aligned}$$

Le taux d'évolution annuel moyen entre 1946 et 1954 est de 566,750 hab/an tandis qu'entre 1990 et 1999 il est de 400,222 hab/an.

**2.** Le plan est rapporté à un repère orthogonal, les unités graphiques étant : 0,25cm sur l'axe des abscisses ; 1cm sur l'axe des ordonnées pour 5000 habitants, la graduation des ordonnées débutant à 80000. Construire le nuage de points  $M(x_i; y_i)$





3. Quel le coefficient de corrélation linéaire entre  $x$  et  $y$  ?

La calculatrice donne un coefficient de corrélation linéaire de  $r = 0,989$ . Il est pertinent de réaliser un ajustement affine car le coefficient de corrélation est très proche de 1, montrant que les deux variables  $x$  et  $y$  sont corrélées linéairement.

4. Indiquer les coordonnées du point moyen  $G$  associé à la série  $(x_i, y_i)$  et placer ce point sur le graphique précédent.

Les coordonnées du point moyen sont obtenues en réalisant la moyenne  $\bar{x}$  des  $x_i$  et la moyenne  $\bar{y}$  des  $y_i$ . Ainsi,

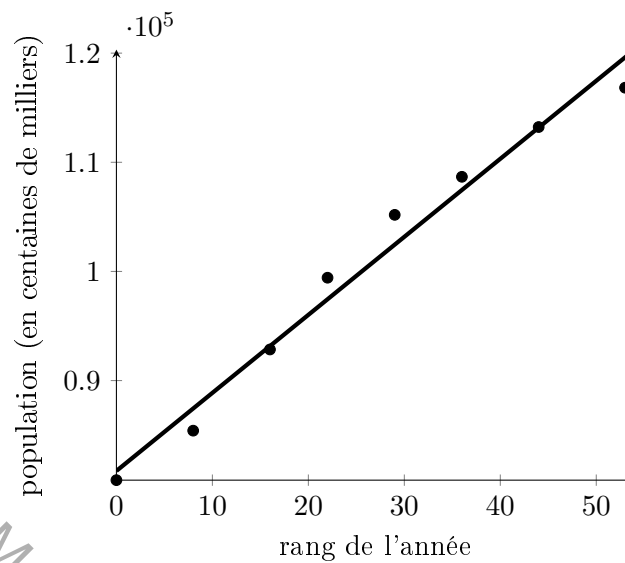
$$\begin{aligned}\bar{x} &= \frac{0 + 8 + \dots + 53}{8} = 26 \\ \bar{y} &= \frac{80878 + \dots + 1168332}{8} = 100310\end{aligned}$$

Les coordonnées du point  $G$  sont donc  $(26 ; 100310)$ .

5. Déterminer une équation de la droite d'ajustement affine de  $y$  en  $x$  par la méthode des moindres carrés.

La calculatrice nous donne  $y = 714,845x + 81724$ .

6. Tracer cette droite (D) sur le graphique précédent



Remarque : par la méthode des moindres carrés, la droite passe toujours par le point moyen  $G$ .

7. En supposant que cette évolution de la population se poursuive, donner une estimation de la population en 2020.

Si l'année 1999 représente le rang 53, alors l'an 2020 a pour rang  $x = 74$ . En remplaçant dans l'équation de la droite d'ajustement, on obtient une population estimée de 134623 personnes en 2020 sous l'hypothèse que l'évolution constatée de la population se poursuive entre 1999 et 2020.