Exercice 3 5 points

Les trois parties A, B et C peuvent être ytaitées de façon indépendante.

Partie A

Le chikungunya est une maladie virale transmise d'un être humain à un autre par les piqûres de moustiques femelles infectées.
Un test a été mis au point pour le dépistage de ce virus. Le laboratoire fabriquant ce test fournit les caractéristiques suivantes :

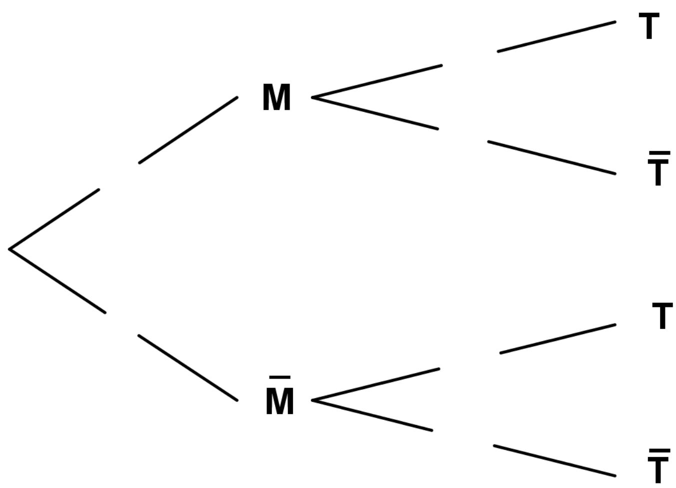
. la probabilité qu'une personne atteinte par le virus ait un test positif est 0,98 ;
. la probabilité qu'une personne non atteinte par le virus ait un test positif est de 0,01.

On procède à un test de dépistage systématique dans une population « cible ». Un individu est choisi au hasard dans cette population. On appelle :

. M l'événement : « L'individu choisi est atteint du chikungunya »
. T l'événement : « Le test de l'individu choisi est positif ».

On notera $\bar{M}$ (respectivement $\bar{T}$) l'événement contraire de l'événement M (respectivement T).
On note p $(0 \leqslant p \leqslant 1)$ la proportion de personnes atteintes par la maladie dans la population cible.

1.a. Recopier et compléter l'arbre de probabilité ci-dessous.



b. Exprimer $P(M \cap T)$ et $P(\bar{M} \cap T)$ puis $P(T)$ en fonction de p.

2.a. Démontrer que la probabilité de M sachant T est donnée par la fonction f définie sur [0;1] par :

$$f(p)=\frac{98p}{97p+1}.$$

b. Etudier les variations de la fonction f.

3. On considère que le test est fiable lorsque la probabilité qu'une personne ayant un test positif soit réellement atteinte du chikungunya est supérieure à 0,95.
En utilisant les résultats de la question **2.** à partir de quelle proportion p de malades dans la popula-tion le test est-il fiable ?

Partie B

En juillet 2014, l'institut de vielle sanitaire d'une île, s'appuyant sur des données remontées par les médecins, publie que 15 % de la population est atteinte par le virus.
Comme certaines personnes ne consulte pas forcément leur médecin, on pense que la proportion est en réalité plus importante.

Pour s'en assurer, on sepropose d'étudier un échantillon de 1 000 personnes choisies au hasard dans cette île. La population est suffisamment importante pour considérer qu'un échantillon résulte d'un tirage avec remise.
On désigne par X la variable aléatoire qui, à tout échantillon de 1 000 personnes choisies au hasard, lui fait corrspondre le nombre de personnes atteintes par le virus et par F la variable aléatoire don-nant la fréquence associée.

1.a. Sous l'hypothèse p = 0,15, déterminer la loi de X.
b. Dans un échantillon de 1 000 personnes choisies au hasard dans l'île, on dénombre 197 person-nes atteintes par le virus.
Quelle conclusion peut-on tirer de cette observation à propos du chiffre de 15 % publié par l'institut de vielle sanitaire ?
Justifier. (On pourra s'aider du calcul d'un intervalle de fluctuation au seuil de 95%).

2. On considère désormais que la valeur de p est inconnue.
En utilisant l'échantillon de la question **1.b.** proposer un de confiance de la valeur de p, au niveau de confiance de 95 %.

Partie C

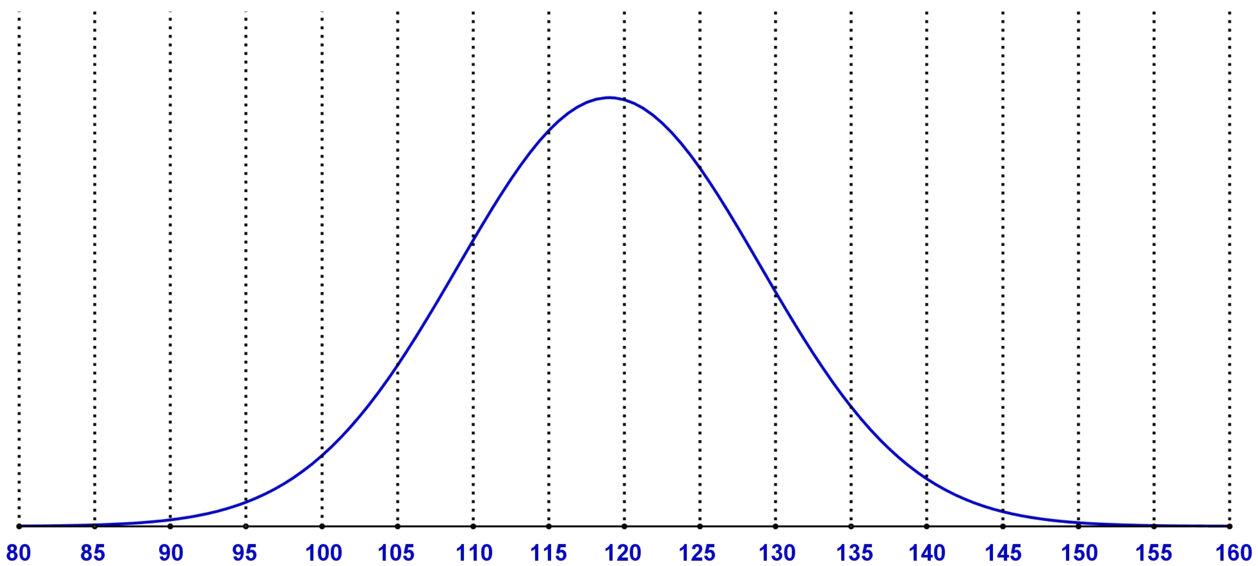
Le temps d'incubation, exprimé en heures, du virus peut être modélisé par une variable aléatoire T suivant une loi normale d'écart type $\sigma=10$.
On souhaite déterminer sa moyenne μ.
La représentation graphique de la fonction densité de probabilité de T est donnée en annexe.

1.a. Conjecturer, à l'aide du graphique, une valeur approchée de μ.
b. On donne $P(T<110)=0{,}18$. Hachurer sur le graphique un domaine dont l'aire correspond à la probabilité donnée.

2. On note T' la variable aléatoire égale à $\dfrac{T-\mu}{10}$

a. Quelle loi la variable aléatoire T' suit-elle ?
b. Déterminer une valeur approchée à l'unité près de la moyenne μ de la variable aléatoire T et vérifier la conjedture de la question 1.

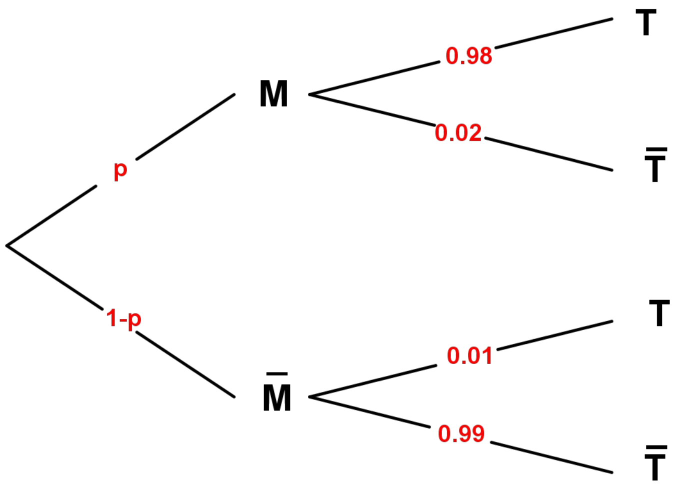
ANNEXE
(à rendre avec la copie)



CORRECTION

Partie A

1.a. L'énoncé précise :
« On notera p la proportion des personnes atteintes par la maladie dans la population cible »
donc $P(M)=p$ et $P(\bar{M})=1-P(M)=1-p$.
« La probabilité qu'une personne atteinte par le virus ait un test positif est de 0,98 »
donc $P_M(T)=0{,}98$ et $P_M(\bar{T})=1-P_M(T)=1-0{,}98=0{,}02$
« La probabilité qu'une personne non atteinte par le virus ait un test positif est 0,01 »
donc $P_{\bar{M}}(T)=0{,}01$ et $P_{\bar{M}}(\bar{T})=1-P_{\bar{M}}(T)=1-0{,}01=0{,}99$.
On obtient l'arbre pondéré suivant :



b. $P(M\cap T)=P(M)\times P_M(T)=p\times 0{,}98$ = **0,98p**
$P(\bar{M}\cap T)=P(\bar{M})\times P_{\bar{M}}(T)=(1-p)\times 0{,}01$ = **0,01- 0,01p**
En utilisant l'arbre pondéré ou la formule des probabilités totales, on obtient :
$P(T)=P(M\cap T)+P(\bar{M}\cap T)=0{,}98\,p+0{,}01-0{,}01\,p$ = **0,01+ 0,97p**

2.a. On nous demande d'exprimer en fonction de p : $P_T(M)$.

$$P_T(M)=\frac{P(M\cap T)}{P(T)}=\frac{0{,}98\,p}{1+0{,}97\,p}=\frac{98\,p}{1+97\,p}$$

On note : $f(p)=\dfrac{98\,p}{1+97\,p}$

b. f est dérivable sur [0;1]

$$f'(p)=\frac{98\times(1+97\,p)-98\,p\times 97}{(1+97\,p)^2}=\frac{98}{(1+97\,p)^2}>0$$ f est strictement croissante sur [0;1]

On donne le tableau de variation de f

p	0		1
f '(p)		+	
f(p)	0	↗	1

3. Le test est fiable si et seulement si $P_T(M)=f(p)>0{,}95$

$\frac{98p}{1+97p}>0{,}95 \Leftrightarrow 98p>0{,}95\times(1+97p)\Leftrightarrow(98-0{,}95\times97)p>0{,}95$

$\Leftrightarrow 5{,}85p>0{,}95\Leftrightarrow p>\frac{0{,}95}{5{,}85}=\frac{95}{585}=\frac{19}{117}$ = **0,1624** à 10^{-4} près.

Le test est fiable siet seulement si plus de 16,24 % de la population cible est atteinte par le virus.

Partie B

1.a. On considère l'épreuve de Bernoulli suivante : on choisit une personne de l'île au hasard
succès S « cette personne est atteinte du virus » P(S) = 0,15
échec $\overline{S}$ « cette personne n'est pas atteinte par le virus » $P(\overline{S})=1{,}0{,}15=0{,}85$.
On choisit au hasard 1000 personnes et on suppose que les tirages sont considérés avec remise,
on obtient un schéma de Bernoulli de paramètres n = 1000 et p = 0,15.
X est la variable aléatoire égale au nombre de succès en 1000 épreuves.
La loi de probabilité de X est la loi binomiale de paramètres : n = 1000 et p = 0,15.

b. p = 0,15
$n=1000\geqslant30$ et $np=150\geqslant5$ et $n(1-p)=850\geqslant5$
On put déterminer, pour la fréquence associée, un intervalle de fluctuation asymptotique au seuil de 95 %.

$$I=\left[0{,}15-1{,}96\times\frac{\sqrt{0,15\times0{,}85}}{1000};0{,}15+1{,}96\times\sqrt{\frac{0{,}15\times085}{1000}}\right]$$

$1{,}96\times\sqrt{\frac{0{,}15\times0{,}85}{1000}}=0{,}022$ à 10^{-3} près

$I=[0{,}128;0{,}172]$.

La proportion obtenue dans l'échantillon de 1000personnes est : $\frac{197}{1000}=0{,}197$.

Or 0,197 n'appartient pas à l'intervalle I.
On rejette au seuil de 95 %, l'hypothèse que p = 0,15.

2. La proportion obtenue dans l'échantillon est : 0,197.
$n=1000\geqslant30$ et $n\times0{,}197=197\geqslant5$ et $n\times(1-0{,}97)=803\geqslant5$
On considère l'intervalle de confiance J au niveau de confiance 95 %.

$$J=\left[0{,}197-\frac{1}{\sqrt{1000}};0{,}197+\frac{1}{\sqrt{1000}}\right]$$

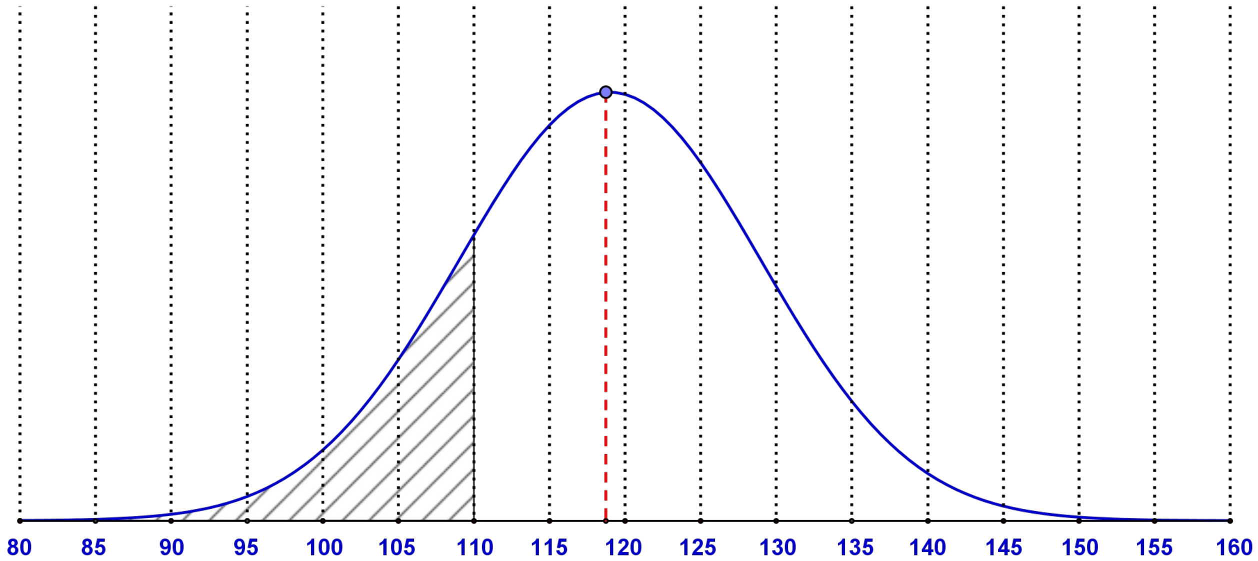
$\frac{1}{\sqrt{1000}}$ = 0,032 à 10^{-3} près

$J=[0{,}165;0{,}229]$

Partie C

1.a. La moyenne μ de la variable aléatoire T, suivant une loi normale, est l'abscisse du maximum de la fonction densité de probabilité de T.
Graphiquement, on obtient **μ = 119.**

b. $P(T<110)$ est l'aire en U.A. de la partie de plan comprise entre la courbe densit é de probabilité de T et l'axe des abscisses sur l'intervalle : $]-\infty;110[$.
On hachure cette partie sur le graphique donné en annexe.
L'énoncé précise que : $P(T<110)=0{,}18$.



2. On note $T' = \dfrac{T-\mu}{10}$.

a. $\sigma = 10$ donc T' suit la loi normale centrée est ré duite.

b. $T < 110 \Leftrightarrow T' = \dfrac{T-\mu}{10} < \dfrac{110-\mu}{10}$

$P(T<110) = 0{,}18 \Leftrightarrow P\left(T' < \dfrac{110-\mu}{10}\right)$

En utilisant la calculatrice, on détermine la valeur de α telle que $P(T'<\alpha) = 0{,}18$

on obtient $\alpha = -0{,}9154$ à 10^{-4} près.

Donc $\dfrac{110-\mu}{10} = -9154$

$110-\mu = 9{,}154$

$\mu = 119{,}154$

On arrondit à l'unité près on obtient **μ = 119**

(c'est la valeur que l'on avait trouvée par lecture graphique).