Devoir Surveillé n°7 Correction

TES

Probabilités et Échantillonnage

Durée 1 heure - Coeff. 5

Noté sur 20 points

Exercice 1. Probabilités conditionnelles, loi normale , IFA, IC et loi uniforme 16 points

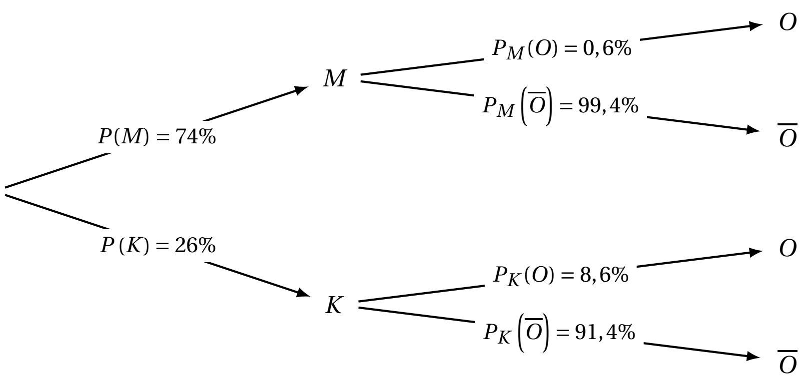
D'après une étude récente il y a 216 762 médecins en France métropolitaine parmi lesquels 0,6 % pratiquent l'ostéopathie et on compte 75 164 kinésithérapeutes parmi lesquels 8,6 % pratiquent l'ostéopathie,

Partie A : Probabilités conditionnelles (4 points)

On note les évènements suivants :

- M : « la personne choisie est médecin » ;
- K : « la personne choisie est kinésithérapeute » ;
- O : « la personne choisie pratique l'ostéopathie ».

1. Reproduire l'arbre de probabilité puis le compléter.



- « *il y a 216 762 médecins en France métropolitaine parmi lesquels 0,6 % pratiquent l'ostéopathie* » donc $P_M(O) = 0,6\% = 0,006$;
- « *on compte 75 164 kinésithérapeutes parmi lesquels 8,6 % pratiquent l'ostéopathie,* » donc $P_K(O) = 8,6\% = 0,086$.

2. Montrer que la probabilité $P(O)$ de l'évènement O est égale à 0,026 8.

Les évènements K et M formant une partition de l'univers on a, d'après la formule des probabilités totales :

$$P(O) = P(O \cap M) + P(O \cap K)$$
$$P(O) = P_M(O) \times P(M) + P_K(O) \times P(K)$$
$$P(O) = 0,006 \times 0,74 + 0,086 \times 0,26$$

$$\boxed{P(O) = 0,0268}$$

3. Un patient vient de suivre une séance d'ostéopathie chez un praticien d'une des deux catégories. Déterminer la probabilité que le praticien soit un kinésithérapeute. Donner le résultat arrondi au centième.

La probabilités cherchée est $P_O(K)$ or :

$$\begin{aligned} P_O(K) &= \frac{P(O \cap K)}{P(O)} \\ &= \frac{P_K(O) \times P(K)}{0,0268} \\ &= \frac{0,086 \times 0,26}{0,0268} \end{aligned}$$

On a donc, arrondi au centième :

$$\boxed{P_O(K) \approx 0{,}83}$$

La probabilité que le praticien soit un kinésithérapeute est donc, arrondie au centième, de $0,83$.

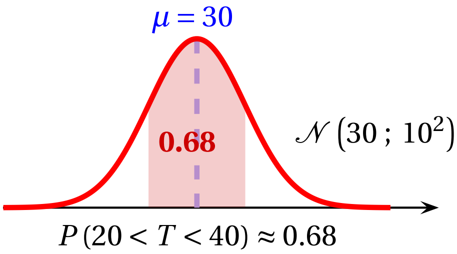
Partie B : loi normale (4 points)

On note T la variable aléatoire associant à chaque patient la durée de visite, en minutes, chez un médecin-ostéopathe. On admet que T suit la loi normale d'espérance 30 et d'écart-type 10.
Dans cette partie, les résultats seront arrondis au centième.

1. Déterminer la probabilité $P(20 \leqslant T \leqslant 40)$.

La variable aléatoire T suit une loi normale d'espérance $\mu = 30$ et d'écart-type $\sigma = 10$. La calculatrice nous donne à 10^{-2} près :

$$T \sim \mathcal{N}\left(30\,;\,10^2\right) \Longrightarrow P\left(20 < T < 40\right) \approx \underline{0{,}68}$$



> *Calculatrices*
>
> - *Sur la TI Voyage 200 : TIStat.normFDR*$(20\,,\,40\,,\,30\,,\,10) \approx \underline{0{,}6826894809}$
> - *Sur TI82/83+ : normalcdf*$(20\,,\,40\,,\,30\,,\,10)$ *ou (fr.) normalfrép*$(20\,,\,40\,,\,30\,,\,10)$
> - *Sur Casio 35+ ou 75 : Menu STAT/DIST/NORM/Ncd* $\Rightarrow$ *NormCD*$(20\,,\,40,\,10\,,\,30\,)$

2. Déterminer la probabilité qu'une visite dure plus de trois quart d'heure.

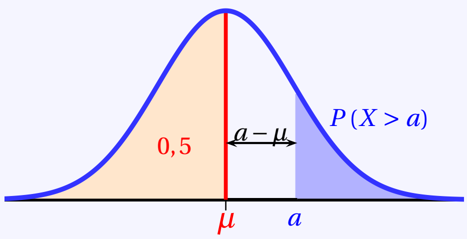
Propriété 1 $\left(P\left(X > a\right)\,;\,a > \mu\right)$

Si la variable X suit une loi normale $\mathcal{N}\left(\mu\,;\,\sigma^2\right)$ alors :

$$P\left(X < \mu\right) = 0{,}5 = P\left(X > \mu\right)$$

De plus pour tout réel a avec $a > \mu$:

$$\boxed{P\left(X > a\right) = 0{,}5 - P\left(\mu < X < a\right)}$$



Donc ici avec $\mu = 30$ et $a = 45 > \mu$ on a :

$$P(T \geq 45) = 0{,}5 - P(30 \leq T \leq 45) \approx 0{,}07$$

La probabilité qu'une visite dure plus de trois quart d'heure est donc, arrondie au centième, de $\underline{0{,}07}$.

> *Calculatrices*
>
> - *Sur la TI Voyage 200 :* $\left(0{,}5 - \text{ TIStat.normFDR}(30\,,\,45\,,\,30\,,\,10)\right) \approx \underline{0{,}0668072292}$
> - *Sur TI82/83+ : normalcdf*$(30\,,\,45\,,\,30\,,\,10)$ *ou (fr.) normalfrép*$(30\,,\,45\,,\,30\,,\,10)$
> - *Sur Casio 35+ ou 75 : Menu STAT/DIST/NORM/Ncd* $\Rightarrow$ *NormCD*$(30\,,\,45,\,10\,,\,30\,)$

3. Déterminer k tel que $P(T \geq k) = 0{,}8$ (arrondir à l'unité) et interpréter le résultat dans le cadre de l'exercice.

On cherche k tel que $P\left(T \leq k\right) = 0{,}2$ où T qui suit une loi normale $\mathcal{N}\left(30\,;\,10^2\right)$. La calculatrice nous donne alors avec la répartition normale réciproque, arrondi à 10^{-0} près :

$$\boxed{P\left(T \leq k\right) = 0{,}2 \Longleftrightarrow k \approx \underline{22}}$$

> *Calculatrices*
>
> - *Sur la TI Voyage 200 : TIStat.invNorm*$(0.2\,,\,30\,,\,10) \approx \underline{21{,}583788}$
> - *Sur TI82/83+ : invNorm*$(0.2\,,\,30\,,\,10)$ *ou (fr.) FracNormale*$(0.2\,,\,30\,,\,10)$
> - *Sur Casio 35+ ou 75 : Menu STAT/DIST/NORM/InvN* $\Rightarrow$ *InvNormCD*$(0.2\,,\,10\,,\,30)$

Cela signifie qu'environ 80% des patients on une durée de visite supérieure ou égale à 22 minutes.

Partie C : IFA (4 points)

On rappelle qu'en France métropolitaine 0,6 % des médecins pratiquent l'ostéopathie. Une région compte 47 000 médecins dont 164 *médecins-ostéopathes. On note I l'intervalle de fluctuation asymptotique au seuil de 95 % de la fréquence de médecins ostéopathes de la région.*

1.

1. a. Vérifier que les conditions d'utilisation de cet intervalle sont remplies.

Théorème 1 (Intervalle de fluctuation asymptotique)

Si les conditions suivantes sont remplies : $\begin{cases} \checkmark & n \geq 30 \\ \checkmark & np \geq 5 \\ \checkmark & n(1-p) \geq 5 \end{cases}$

Alors un intervalle de fluctuation asymptotique au seuil de confiance de 95% de la fréquence F_n d'un caractère dans un échantillon de taille n est, si p désigne la proportion de ce caractère dans la population :

$$I_n = \left[p - 1{,}96\frac{\sqrt{p(1-p)}}{\sqrt{n}} \ ; \ p + 1{,}96\frac{\sqrt{p(1-p)}}{\sqrt{n}}\right]$$

- **1. Analyse des données** :
  - « *Une région compte 47 000 médecins dont* 164 *médecins-ostéopathes.* ». Donc la fréquence observée de médecins-ostéopathes est
  $$f = \frac{164}{47\,000} \approx 0{,}003\,489 \approx \mathbf{0.3489\%}$$
  - « *En France métropolitaine* 0,6 % *des médecins pratiquent l'ostéopathie*, donc la proportion de médecins pratiquant l'ostéopathie est $p = 0{,}6\,\%$.
- **2. Intervalle de fluctuation** : On a $n = 47000$, $p = 0{,}6\%$ alors on sait que puisque :
$$\begin{cases} \checkmark & n = 47\,000 \geq 30 \\ \checkmark & np = 47\,000 \times 0{,}6\% = 282 \geq 5 \\ \checkmark & n(1-p) = 47\,000 \times 99{,}4\% = 46\,718 \geq 5 \end{cases}$$
Les conditions de validité sont réunies pour déterminer un intervalle de fluctuation asymptotique au seuil de confiance de 95%.

1. b. Justifier que $I = [0{,}0053 \ ; \ 0{,}0067]$, les bornes ayant été arrondies à 10^{-4} près. Peut-on considérer que pour la pratique de l'ostéopathie par les médecins, cette région est représentative, privilégiée ou défavorisée par rapport à la situation en France métropolitaine ? Justifier la réponse.

- **Intervalle de fluctuation**
Les conditions de validité sont réunies donc l'intervalle de fluctuation au seuil 95% pour la fréquence $F_{47\,000}$.
On a pour le cas étudié , $n = 47000$, $p = 0{,}6$ %. Un intervalle fluctuation asymptotique au seuil de confiance de 95% de la fréquence f dans un échantillon de taille $n = 47000$: est alors :
$$I_n = \left[p - 1{,}96\frac{\sqrt{p(1-p)}}{\sqrt{n}} \ ; \ p + 1{,}96\frac{\sqrt{p(1-p)}}{\sqrt{n}}\right] = \left[0{,}006 - 1{,}96\ \frac{\sqrt{0{,}006 \times 0{,}994}}{\sqrt{47000}} \ ; 0{,}006 + 1{,}96\ \frac{\sqrt{0{,}006 \times 0{,}994}}{\sqrt{47000}}\right]$$
Soit puisque les borne sont :
  - $p - 1{,}96\ \frac{\sqrt{p(1-p)}}{\sqrt{n}} \approx 0{,}0053$. On arrondit la borne inférieure par défaut 10^{-4} près soit $\underline{0{,}0053}$.
  - $p + 1{,}96\ \frac{\sqrt{p(1-p)}}{\sqrt{n}} \approx 0{,}0067$. On arrondit la borne supérieure par exc 10^{-4} près soit $\underline{0{,}0067}$.
$$\boxed{I_{47000} \approx [0{,}0053 \ ; \ 0{,}0067]}$$
- **Conclusion** : La fréquence observée n'appartient pas à l'intervalle de fluctuation
$$\boxed{f = \frac{164}{47\,000} \approx 0{,}003\,489 \notin I_{47\,000}}$$
donc **la région est donc défavorisée par rapport à la situation en France métropolitaine.** .

Partie D : intervalle de confiance (3 points)

Le ministre de la santé souhaite savoir si les français sont satisfaits de leur médecin généraliste. Une enquête de satisfaction est réalisée sur un échantillon de 500 français et 304 se déclarent satisfaits de leur médecin.

1. **Donner un intervalle de confiance au niveau de 95 % de la proportion de patients satisfaits.**

- **Analyse des données** :
  - « *Sur un échantillon de* $n = 500$ *patients. Il est constaté que* 304 *sont satisfaits.* ». Donc la fréquence observée de patients satisfaits est

$$f = 304 \div 500 = 0{,}608 \text{ soit } f = \underline{0{,}608}$$

- **Intervalle de confiance** :
  On a pour le cas étudié, $n = 500$, $f = 0{,}608$. Vérifions les conditions d'application du théore :

$$\begin{cases} \checkmark & n = 500 \geq 30 \\ \checkmark & nf = 500 \times \dfrac{304}{500} = 304 \geq 5 \\ \checkmark & n(1-f) = 500 \times \dfrac{196}{500} = 196 \geq 5 \end{cases}$$

  Un intervalle de confiance au seuil de confiance de 95% de la proportion p est alors :

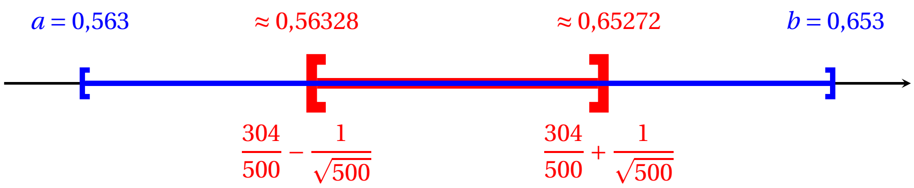
$$I_n = \left[f - \frac{1}{\sqrt{n}} \; ; \; f + \frac{1}{\sqrt{n}}\right] = \left[\frac{304}{500} - \frac{1}{\sqrt{500}} \; ; \; \frac{304}{500} + \frac{1}{\sqrt{500}}\right]$$

  Soit puisque les borne sont :

  - $f - \dfrac{1}{\sqrt{n}} = \dfrac{304}{500} - \dfrac{1}{\sqrt{500}} \approx 0{,}56328$. On arrondit la borne inférieure par défaut 10^{-3} prms soit $\underline{0{,}563}$.
  - $f + \dfrac{1}{\sqrt{n}} = \dfrac{304}{500} + \dfrac{1}{\sqrt{500}} \approx 0{,}65272$. On arrondit la borne supérieure par excms 10^{-3} prms soit $\underline{0{,}653}$.

$$\boxed{I_{500} \approx [0{,}563 \; ; \; 0{,}653]}$$

- **Conclusion**
  Cet intervalle contient la proportion p des patients satisfaits , au niveau de confiance 95% (ou au risque d'erreur de 5%). La proportion des patients satisfaits se situe donc, au niveau de confiance 95%, <u>entre 56.3% et 65.3%</u>.



2. **Le directeur souhaite cependant avoir une estimation plus précise et donc veut un intervalle de confiance au niveau de 95 % d'amplitude 0,05. Déterminer le nombre de personnes à interroger pour obtenir un tel intervalle.**
   L'amplitude de l'intervalle de confiance est de : $\dfrac{2}{\sqrt{n}}$ donc on cherche l'entier n tel que :

$$\frac{2}{\sqrt{n}} = 0{,}05 \iff \frac{4}{n} = 0{,}05^2 \iff n = \frac{4}{0{,}05^2} = \underline{1\,600}$$

Partie E : loi uniforme (1 point)

Un client est choisi au hasard chez un médecin généraliste et on note A la durée (en minutes) qui s'est écoulée entre l'arrivée du client dans la salle d'attente du cabinet, et sa prise ne charge par le médecin. On admet que A est une variable aléatoire qui suit la loi uniforme sur l'intervalle $[0 \; ; \; 90]$. *Déterminer la probabilité que le client choisi attende moins de 30 minutes avant d'être pris en charge par le médecin.*

$$P(A < 30) = P(0 < A < 30) = \frac{30-0}{90-0} = \frac{1}{3} \approx \underline{0{,}33}$$

Exercice 2. Loi binomiale 4 points

Un opérateur de téléphonie mobile organise une campagne de démarchage par téléphone pour proposer la souscription d'un nouveau forfait à sa clientèle. Les relevés réalisés au cours de ces premières journées permettent de constater que 12 % des personnes interrogées souscrivent à ce nouveau forfait. Chaque employé de l'opérateur effectue 60 appels par jour. On suppose le fichier suffisamment important pour que les choix soient considérés réalisés de façon indépendante et dans des conditions identiques. On note X la variable aléatoire qui comptabilise le nombre de souscriptions réalisées par un employé donné un jour donné.

1. Justifier que la variable aléatoire X suit une loi binomiale dont on donnera les paramètres.

Modélisation

Il y a répétition de $n = 60$ événements indépendants et identiques (on tire client).
Chaque tirage a deux issues possibles (épreuve de Bernoulli) :

- succès de probabilité $p = 0,12$ quand client est souscripteur d'un nouveau forfait;
- et échec de probabilité $1 - p = 0,88$ sinon.

Donc la variable aléatoire X qui est égale au nombre de succès au cours de ces $n = 60$ épreuves indépendantes de Bernoulli de paramètre $p = 0,12$ suit une loi binomiale de paramètres $n = 60$ et $p = 0,12$.
On peut écrire :

$$X \text{ suit } \mathscr{B}\left(60\ ;\ 0,12\right) \text{ ou } X \sim \mathscr{B}\left(60\ ;\ 0,12\right).$$

2. Déterminer la probabilité que l'employé obtienne 5 souscriptions un jour donné. (*Arrondir au centième*).

Puisque X suit une loi Binomiale de paramètre $n = 60$ et $p = 0,12$ on a :

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} = \binom{60}{k} \times 0,12^k \times (0,88)^{60-k}$$

Et donc

$$p(X = 5) = \binom{60}{5} \times 0,12^5 \times 0,88^{55}$$

Soit :

$$\boxed{p(X = 5) \approx 0.12}$$

Calculatrices

- *Sur la TI Voyage 200 : TIStat.binomDdP* (60 , 0,12 , 5) ≈ 0,12
- *Sur TI82/83+ : Menu Distrib ⇒ binomFdp* (60 , 0,12 , 5) ≈ 0,12
- *Sur Casio 35+ ou 75 : Menu Opt/STAT/DIST/DINM ⇒ binomialPD* (5 , 60 , 0,12) ≈ 0,12

3. Déterminer la probabilité que l'employé obtienne au moins trois souscriptions un jour donné. *Arrondir à* 10^{-3}.

- Méthode 1 :
  La probabilité d'avoir au moins trois souscriptions un jour donné est $P(X \geq 3)$ soit en passant à l'évènement contraire :

$$P(X \geq 3) = 1 - P(X = 0) - P(X = 1) - P(X = 2) \approx \underline{0.94}$$

- Méthode 2 :
  On a : $P(X \geq 3) = 1 - P(X \leq 2) \approx \underline{0.94}$
  On utilise alors la calculatrice directement.

  Calculatrices

  - *Sur la TI Voyage 200 :* 1 − *TIStat.binomFdR* (60 , 0,12 , 2) ≈ 0,94
  - *Sur TI82/83+ : Menu Distrib ⇒* 1 − *binomFrép* (60 , 0,12 , 2) ≈ 0,94
  - *Sur Casio 35+ ou 75 :*
    Menu Opt/STAT/DIST/DINM ⇒ 1 − *binomialCD* (2 , 60 , 0,12) ≈ 0,94

4. Déterminer $E(X)$, l'espérance mathématique de X, et interpréter le résultat dans le cadre de l'exercice.

La variable X suit une loi binomiale de paramètre $n = 60$ et $p = 0,12$ donc son espérance mathématiques est :

$$E(X) = n \times p = 60 \times 0,12 = \underline{7,2}$$

Ceci signifie que le nombre moyen de souscriptions réalisées par un employé donné un jour donné est d'environ 7.

Fin du devoir