ꕥ INTERRO ꕥ
PROBABILITÉS ET ESTIMATION

Exercice 1 :

PARTIE A : **Etude du nombre de sinistres dans une équipe de 15 conducteurs**

On répète 15 fois de manière identique et indépendante l'expérience « choisir un conducteur au hasard ». Cette expérience est :

- ⇝ soit un succès (le conducteur n'a pas eu de sinistre) de probabilité 0,6.
- ⇝ soit un échec (le conducteur a eu un sinistre)

X compte le nombre de succès dans ce schéma de Bernoulli. Donc X suit la loi binomiale de paramètres $n = 15$ et $p = 0.6$. Ainsi $P(X = 10) \simeq 0.186$.

PARTIE B : **Estimation de la fréquence des sinistres**

1. $p_1 = \dfrac{136}{150} \simeq 0.907$
2. $u_{0.05} \simeq 1.96$ Donc $\quad IC_1 \simeq \left[0.907 - 1.96\sqrt{\dfrac{0.907 \times 0.093}{150}} ; 0.907 + 1.96\sqrt{\dfrac{0.907 \times 0.093}{150}}\right]$
   Donc $IC_1 \simeq [0.860; 0.954]$
3. ⇝ Si on augmente la taille de l'échantillon, on obtient un intervalle de confiance de plus faible amplitude. Donc IC_2 est plus précis.
   ⇝ Par contre, cela n'augmente pas le coefficient de confiance : sur un grand nombre d'échantillon de type IC_2, on n'en aura toujours qu'environ 95% qui contiendront la vraie valeur de p.
   Les échantillons IC_1 et IC_2 ont donc la même fiabilité.
   ⇝ Bien sûr !
   ⇝ Cette phrase est fausse, d'après ce que j'ai dit dans le deuxième point de cette question.

PARTIE C : **Estimation du coût des sinistres**

1. $\mu \simeq 1200$ et $s_n = 200$ donc $\sigma \simeq 200\sqrt{\dfrac{100}{99}} \simeq 201.008$
2. ⇝ $u_{0.1} \simeq 1.64$ donc $\quad I_{10\%} = \left[1200 - 1.64\dfrac{201,008}{\sqrt{100}} ; 1200 + 1.64\dfrac{201,008}{\sqrt{100}}\right]$
   Donc $I_{10\%} \simeq [1167.034; 1232.966]$
   ⇝ $u_{0.01} \simeq 2.58$ donc $\quad I_{1\%} = \left[1200 - 2.58\dfrac{201,008}{\sqrt{100}} ; 1200 + 2.58\dfrac{201,008}{\sqrt{100}}\right]$
   Donc $I_{1\%} \simeq [1148.139; 1251.861]$
3. Lequel de ces deux intervalles est :
   ⇝ L'intervalle le plus précis est celui qui a l'amplitude la moins grande, donc c'est le premier.
   ⇝ Cependant, il s'agit aussi de l'intervalle le moins fiable, puisque sur un grand nombre d'échantillons de ce type, seulement environ 90% contiendront réellement μ alors que pour l'autre intervalle, il y en aura environ 99%.

PARTIE D : **Etude du coût des sinistres**

On considère la variable aléatoire C qui, à chaque sinistre tiré au hasard parmi les sinistres de cette catégorie, associe son coût en euros. On suppose que C suit la loi normale de paramètre $\mu = 1200$ et $\sigma = 201$.

1. $P(1000 \leq Z \leq 1500) \simeq 0.772$
2. $\alpha \simeq 1457.592$
3. Environ 90% des voitures sinistrées ont un sinistre de moins de 1457.592€