aYzL **Exercice** - *Étude d'une méthode de Monte Carlo*

On souhaite calculer à l'aide d'une méthode de Monte Carlo une approximation de l'intégrale

$$I = \int_0^1 \sin(\sqrt{x})dx$$

Soit $(X_i)_{i\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes identiquement distribuées selon une loi uniforme $\mathcal{U}([0;1])$.

1. Démontrer que si on définit la suite de variables aléatoires (I_n) par

$$I_n = \frac{1}{n}\sum_{k=1}^{n} \sin\left(\sqrt{X_k}\right)$$

alors la suite (I_n) converge presque sûrement vers la constante I.

> Les variables $\left(\sin\left(\sqrt{X_k}\right)\right)_i$ sont indépendantes et identiquement distribuées, elles admettent en outre une espérance qui se calcule à l'aide du théorème de transfert. Soit f la densité d'une loi uniforme sur $[0;1]$. Alors $\mathbb{E}(X_1) = \int \sin\left(\sqrt{x}\right) f(x)dx = \int_0^1 \sin\left(\sqrt{x}\right)dx = I$.
> D'après la loi forte de grands nombres, la suite de variables aléatoires (I_n) converge donc presque sûrement vers $\mathbb{E}(X_1) = I$.

2. Compléter le code Python suivant, comportant deux champs manquants, afin qu'il affiche une approximation de I.

```
n=1000
S=0
for i in range(n):
  u = ...
  S= S+sin(sqrt(u))
print("Valeur approchée de I = ")
print(...)
```

> Il manque `u=rand()` et `print(S/n)`

3. Pour tout $k \in \mathbb{N}^*$, on pose $Y_k = \sin\left(\sqrt{X_k}\right)$. Les variables aléatoires (X_k) étant i.i.d., les variables aléatoires (Y_k) le sont aussi et on note μ leur espérance et σ^2 leur variance.

(a) Calculer l'espérance et la variance de la variable aléatoire I_n en fonction de μ, σ et n.

> Par linéarité de l'espérance, $\mathbb{E}(I_n) = \frac{1}{n} \times n \times I = I$. Par propriétés de la variance et indépendance des variables dans la somme, $V(I_n) = \frac{1}{n^2} \times n \times V(Y_1) = \frac{V(Y_1)}{n} = \frac{\sigma^2}{n}$.

(b) Déterminer, en justifiant, une approximation de la loi de la variable aléatoire

$$\frac{\sqrt{n}}{\sigma}(I_n - I)$$

lorsque n est suffisamment grand.

> Les variables Y_k sont indépendantes, identiquement distribuées, admettent une espérance et une variance finies, donc d'après le Théorème Central Limite, la variable
>
> $$\frac{\frac{1}{n}\sum_{k=1}^n Y_k - I}{\frac{\sigma}{\sqrt{n}}} = \frac{\sum_{k=1}^n \frac{Y_k}{n} - \mathbb{E}\left(\sum_{k=1}^n \frac{Y_k}{n}\right)}{\sigma\left(\sum_{k=1}^n \frac{Y_k}{n}\right)}$$
>
> suit approximativement une loi normale centrée réduite.

(c) En déduire le nombre d'itérations N_0 (dépendant de σ) à partir duquel la suite (I_n) réalise une approximation de I à 10^{-3} près avec une confiance de 95%.

> On cherche le rang à partir duquel $\mathrm{P}(|I_n - I| < \varepsilon) \geqslant 0.95$ où $\varepsilon = 10^{-3}$. Or
>
> $$\begin{aligned}\mathrm{P}(|I_n - I| < \varepsilon) &= \mathrm{P}(-\varepsilon < I_n - I < \varepsilon)\\ &= \mathrm{P}\left(-\frac{\sqrt{n}}{\sigma}\varepsilon < \frac{\sqrt{n}}{\sigma}(I_n - I) < \frac{\sqrt{n}}{\sigma}\varepsilon\right)\\ &\approx \mathrm{P}\left(-\frac{\sqrt{n}}{\sigma}\varepsilon < Z < \frac{\sqrt{n}}{\sigma}\varepsilon\right) \quad \text{où } Z \sim \mathcal{N}(0,1)\end{aligned}$$
>
> Or par lecture de tables, $\mathrm{P}(-1.96 < Z < 1.96) \approx 0.95$ donc il suffit de prendre n tel que $\frac{\sqrt{n}}{\sigma}\varepsilon \geqslant 1.96$ i.e.
>
> $$n \geqslant 10^6\,(1.96\sigma)^2$$

(d) Soit la variable

$$V_n = \frac{1}{2n}\sum_{k=1}^{n}(Y_{2k} - Y_{2k-1})^2$$

Vérifier que la suite (V_n) permet d'approcher la valeur de σ^2.

> On calcule
>
> $$\begin{aligned}\mathbb{E}((Y_{2k} - Y_{2k-1})^2) &= \mathbb{E}(Y_{2k}^2 - 2Y_{2k}Y_{2k-1} + Y_{2k-1}^2)\\ &= \mathbb{E}(Y_{2k}^2) - 2\mathbb{E}(Y_{2k}Y_{2k-1}) + \mathbb{E}(Y_{2k-1}^2) \text{ par linéarité}\\ &= 2\mathbb{E}(Y_{2k}^2) - 2\mathbb{E}(Y_{2k}Y_{2k-1}) \text{ par identique distribution}\\ &= 2\mathbb{E}(Y_{2k}^2) - 2\mathbb{E}(Y_{2k})\mathbb{E}(Y_{2k-1}) \text{ par indépendance}\\ &= 2\mathbb{E}(Y_{2k}^2) - 2\mathbb{E}(Y_{2k})^2\\ &= 2\mathbb{E}(Y_1^2) - 2\mathbb{E}(Y_1)^2\\ &= 2V(Y_1)\end{aligned}$$
>
> donc d'après la loi forte des grands nombres, $\frac{1}{n}\sum_{k=1}^n (Y_{2k} - Y_{2k-1})^2$ converge simplement vers $2V(Y_1)$, ce qui répond à la question posée.