Corrigé

BCPST2 **D04S** 11/10/2014

[/3] **Présentation**

Problème :

Le but de ce problème est d'exploiter des données statistiques portant sur le nombre d'oisillons issus d'une couvée de faisans pour tenter de modéliser la loi de probabilité du nombre de naissances à l'issue d'une couvée, et d'éventuellement étendre ce modèle à d'autres espèces d'oiseaux.On dispose d'une étude portant sur 500 couvées numérotées de 0 à 499. Les résultats de cette étude nous sont fournis sous forme d'une liste Python nommée *couv* telle que *couv[k]* est le nombre d'oisillons issus de la couvée numéro k.Nous allons tenter d'évaluer puis de modéliser la probabilité p_k pour qu'une couvée donne naissance à k oisillons, pour tout $k \in \mathbb{N}$.

1. Afin d'évaluer les probabilités p_k, on va tout d'abord créer l'histogramme associé à nos données statistiques, puis en déduire la liste des fréquences d'apparition des différents effectifs des couvées.

[/3] **1.a** L'histogramme associé à nos données est une liste *histo* telle que pour tout k, *histo[k]* est le nombre de couvées qui donnent naissance à k oisillons. Par exemple, si *couv=[3,1,4,3,2,0,6,2]* alors *histo=[1,1,2,2,1,0,1]*.

Ecrire une fonction Python nommé *creation-histo* qui prend en entrée une liste du type de *couv* et qui produit en sortie l'histogramme associé (attention : on ne demande pas ici de représenter graphiquement l'histogramme obtenu !)

Solution:

```python
def creation_histo(c):
    m = max(c)
    h = [0 for k in range(m+1)]
    for x in c:
        h[x] = h[x] + 1
    return(h)
```

[/1] **1.b** A partir de l'histogramme obtenu, on construit la liste des fréquences d'apparition des différents effectifs en divisant les valeurs de l'histogramme par le nombre total de couvées étudiées. Avec l'exemple ci-dessus, qui porte sur 8 couvées, on obtient une liste *frequences=[1/8,1/8,1/4,1/4,1/8,0,1/8]*.

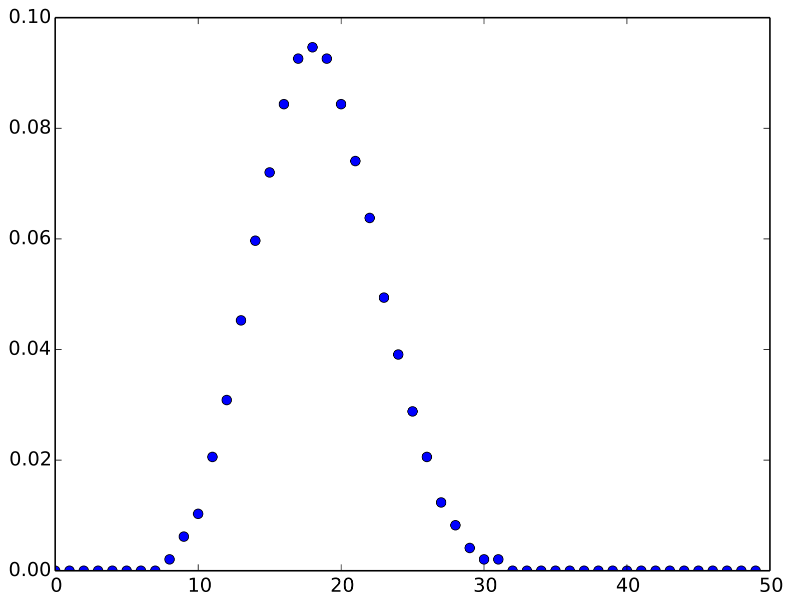
Ecrire une fonction Python qui, à partir d'une liste du type de *couv* et qui retourne la liste des fréquences.

Solution:

```python
def frequences(c):
    histo=creation_histo(c)
    f = [histo[k]/len(c) for k in range(len(histo))]
    return f
```

Les fréquences obtenues après traitement des données expérimentales sont les approximations des probabilités théoriques p_k que l'on va utiliser pour notre modélisation.

2. Afin de modéliser les probabilités p_k, on représente graphiquement les fréquences obtenues (donc *frequence[k]* en fonction de k) :



Pour nos calculs, on note f_k la fréquence associée à la valeur k (plutôt que*frequence[k]*).

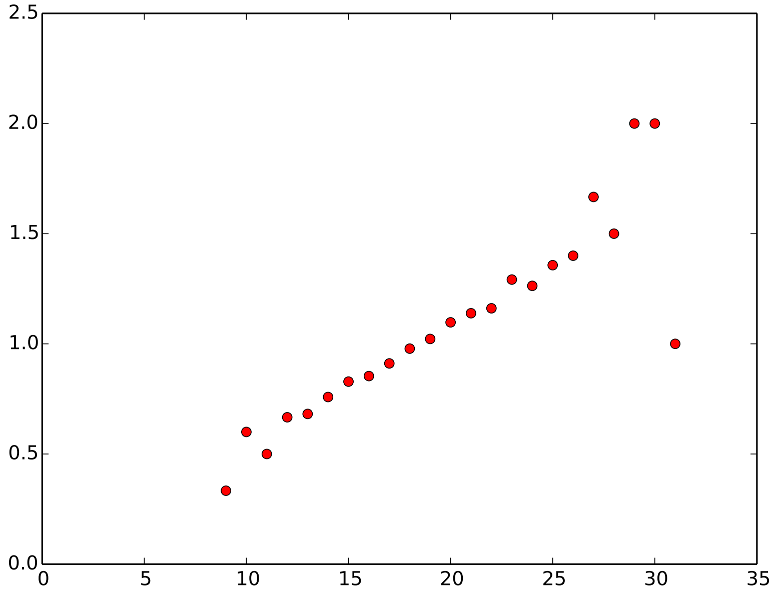
[/3] **2.a** On étudie les variations de f_k en fonction de k. Pour ce faire on introduit la valeur $\alpha_k = \dfrac{f_{k-1}}{f_k}$ pour tout k tel que f_k et f_{k-1} sont non nuls.

Ecrire une fonction Python qui, à partir de la liste des fréquences, retourne deux listes : celle des entiers k tels que α_k est défini (notée Lk), et celle des valeurs de α_k (notée Lalphak).

Solution:

```python
def les_alpha(f):
    Lk = []
    Lalphak = []
    for k in range(1,len(f)):
        if ( f[k]!=0 ) and ( f[k-1]!=0 ) :
            Lk.append(k)
            Lalphak.append(f[k-1]/f[k])
    return Lk , Lalphak
```

[/1] **2.b** On représente alors les α_k en fonction de k et on obtient :



Quelle conjecture formulez-vous au vu de ce graphique quant aux variations de α_k en fonction de k ?

Solution: Hormis quelques valeurs extrêmes, il semble que α_k soit proportionnel à k.

[/2] **2.c** On veut tenter de déterminer la pente de la droite passant par l'origine (donc d'équation de la forme $y = ax$, avec a à déterminer) qui approxime le mieux les points (k, α_k). On procède selon le principe de la régression linéaire par les moindres carrés, on introduit la fonction :

$$\varphi(a) = \sum_{k \in Lk} (\alpha_k - ak)^2$$

Montrer que $\varphi(a)$ est minimale pour la valeur :

$$a_0 = \frac{\sum\limits_{k \in Lk} k\alpha_k}{\sum\limits_{k \in Lk} k^2}$$

Solution: Calculons la dérivée de la fonction φ :

$$\begin{aligned} \varphi'(a) &= \sum_{k \in Lk} -2k(\alpha_k - ak) \\ &= 2\Big(\sum_{k \in Lk} k^2\Big)a - 2\sum_{k \in Lk} k\alpha_k \end{aligned}$$

φ' est croissante et s'annule pour $a_0 = \dfrac{\sum\limits_{k \in Lk} k\alpha_k}{\sum\limits_{k \in Lk} k^2}$.

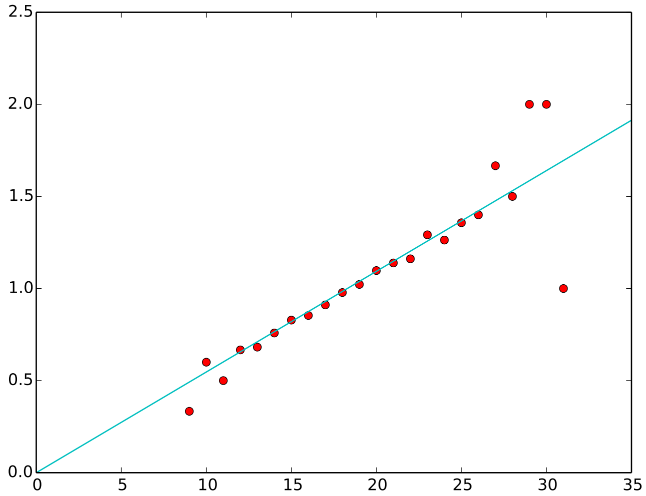
Elle est donc positive avant a_0 et négative après, ce qui montre que φ est minimale en a_0

[/2] **2.d** En déduire une fonction Python qui, à partir des listes Lk et Lalphak, détermine la pente optimale voulue.

Solution:

```
def pente(Lk,Lalphak):
    u = sum([Lk[i]*Lalphak[i] for i in range(len(Lk))])
    v = sum([k**2 for k in Lk])
    return u/v
```

Le calcul de la pente optimale donne le résultat numérique $a_0 = 0.547$.
On trace la droite passant par l'origine et de pente a_0 pour la confronter aux valeurs expérimentales :



[/2] **2.e** Commenter ce résultat.

Solution: L'approximation par une droite est très bonnes pour la plupart des valeurs. Seules les valeurs extrêmes sont éloignées de la droite. Pour les grandes et petites valeurs de k, il est probable qu'il n'y ait que peu de couvées présentant ces valeurs comme nombre d'oisillons, et une petite variation expérimentale produit de grands écarts sur le quotient des fréquences.

3. A l'aide des résultats précédents, on décide donc de modéliser nos probabilités p_k selon le principe :

$$\forall k \in \mathbb{N}^*,\ \frac{p_{k-1}}{p_k} = a_0 k$$

On pose $\lambda = \dfrac{1}{a_0}$.

[/2] **3.a** Exprimer p_1 et p_2 en fonction de p_0 et λ.

Solution:

$$\begin{aligned} p_1 &= \lambda p_0 \\ p_2 &= \frac{\lambda^2}{2} p_0 \end{aligned}$$

[/2] **3.b** Pour tout $k \in \mathbb{N}$, exprimer p_k en fonction de p_0, λ et k.

Solution: Une récurrence facile montre que :

$$\forall k \in \mathbb{N},\ p_k = \frac{\lambda^k}{k!} p_0$$

[/2] **3.c** En utilisant le fait que $\sum_{k=0}^{+\infty} p_k = 1$, déterminer p_0 et donner enfin une expression de p_k en fonction de λ et k.

Solution:

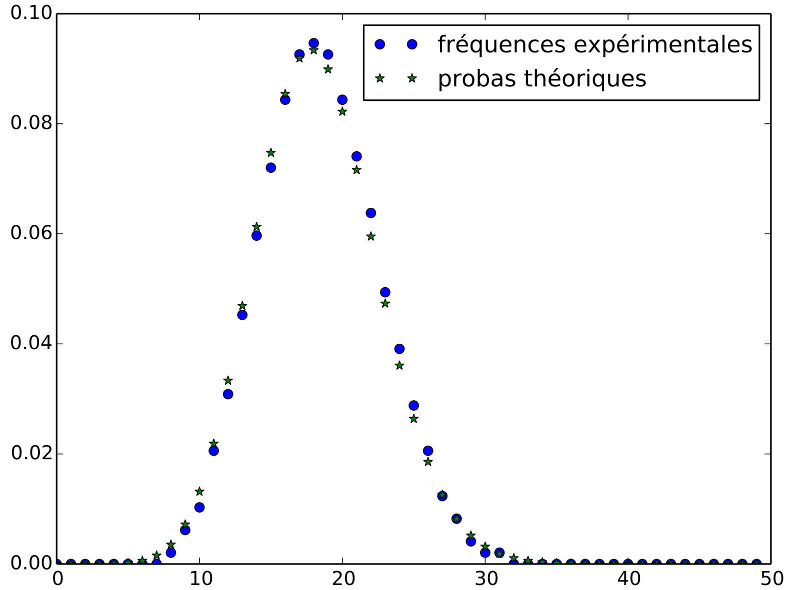
$$\begin{aligned} 1 &= \sum_{k=0}^{+\infty} p_k \\ &= \sum_{k=0}^{+\infty} \frac{\lambda^k}{k!} p_0 \\ &= p_0 \sum_{k=0}^{+\infty} \frac{\lambda^k}{k!} \\ &= p_0 e^{\lambda} \end{aligned}$$

On en déduit donc que $p_0 = e^{-\lambda}$ et donc que :

$$\forall k \in \mathbb{N},\ p_k = e^{-\lambda}\frac{\lambda^k}{k!}$$

(nous retrouverons cette loi, dite loi de Poisson, dans un chapitre de probabilités ultérieur ...)

On représente sur un même graphique les fréquences expérimentales déjà visualisées plus haut et les probabilités théoriques p_k qui viennent d'être déterminées :



[/1] **3.d** Que pensez-vous de la pertinence de notre modèle ?

Solution: Le modèle est tellement proche de la réalité que ça en est suspect ...

4. On suppose que la probabilité qu'une couvée de faisans donne naissance à k oisillons est :

$$p_k = e^{-\lambda}\frac{\lambda^k}{k!}$$

avec $\lambda = 18.3$

[/3] **4.a** Calculer le nombre moyen d'oisillons issus d'une couvée, donné par la formule $\sum_{k=0}^{+\infty} k.p_k$.

Solution:

$$\begin{aligned}
\sum_{k=0}^{+\infty} k.p_k &= \sum_{k=0}^{+\infty} k e^{-\lambda}\frac{\lambda^k}{k!} \\
&= e^{-\lambda}\sum_{k=1}^{+\infty} k\frac{\lambda^k}{k!} \\
&= e^{-\lambda}\sum_{k=1}^{+\infty} \frac{\lambda^k}{(k-1)!} \\
&= e^{-\lambda}\lambda\sum_{k=1}^{+\infty} \frac{\lambda^{k-1}}{(k-1)!} \\
&= e^{-\lambda}\lambda\sum_{k=0}^{+\infty} \frac{\lambda^k}{k!} \\
&= e^{-\lambda}\lambda e^{\lambda} \\
&= \lambda
\end{aligned}$$

Le nombre moyen d'oisillons issus d'une couvée est donc de $\lambda = 18.3$.

[/4] **4.b** On se demande si une couvée a plus de chances de donner naissance à un nombre pair d'oisillons qu'à un nombre impair. On introduit donc la probabilité d'avoir un nombre pair de naissances :

$$p = \sum_{k=0}^{+\infty} p_{2k}$$

et celle d'avoir un nombre impair de naissances :

$$q = \sum_{k=0}^{+\infty} p_{2k+1}$$

Calculer $p+q$ et $p-q$. En déduire les valeurs de p et q, et répondre à la question.

Solution:

$$\begin{aligned}
p+q &= \sum_{k=0}^{+\infty} p_{2k} + \sum_{k=0}^{+\infty} p_{2k+1} \\
&= \sum_{k=0}^{+\infty} p_k \\
&= 1
\end{aligned}$$

$$
\begin{aligned}
p-q &= \sum_{k=0}^{+\infty} p_{2k} - \sum_{k=0}^{+\infty} p_{2k+1} \\
&= \sum_{k=0}^{+\infty} e^{-\lambda}\frac{\lambda^{2k}}{(2k)!} - \sum_{k=0}^{+\infty} e^{-\lambda}\frac{\lambda^{2k+1}}{(2k+1)!} \\
&= \sum_{k=0}^{+\infty} e^{-\lambda}\frac{(-\lambda)^{2k}}{(2k)!} + \sum_{k=0}^{+\infty} e^{-\lambda}\frac{(-\lambda)^{2k+1}}{(2k+1)!} \\
&= e^{-\lambda}\Big(\sum_{k=0}^{+\infty} \frac{(-\lambda)^{k}}{k!} \\
&= e^{-\lambda}e^{-\lambda} \\
&= e^{-2\lambda}
\end{aligned}
$$

De ces deux relations on tire facilement que :

$$
\begin{aligned}
p &= \frac{1}{2}\Big((p+q)+(p-q)\Big) \\
&= \frac{1}{2}(1+e^{-2\lambda})
\end{aligned}
$$

et :

$$
\begin{aligned}
q &= \frac{1}{2}\Big((p+q)-(p-q)\Big) \\
&= \frac{1}{2}(1-e^{-2\lambda})
\end{aligned}
$$

Et on constate alors que $p > q$: il est plus probable d'avoir un nombre pair qu'un nombre impair d'oisillons dans une couvée (étonnant, non ?).

Total pour Pb:	/28

Exercice :

Pour tout n entier naturel, on note $I_n = \int_0^{\frac{\pi}{2}} \cos^n t \, dt$.

[/3] **1.** A l'aide d'une intégration par parties, prouver que pour tout n entier naturel :

$$I_{n+2} = \frac{n+1}{n+2} I_n$$

Solution:

$$\begin{aligned}
I_{n+2} &= \int_0^{\frac{\pi}{2}} \cos^{n+2} t \, \mathrm{d}t \\
&= \int_0^{\frac{\pi}{2}} \cos t \cos^{n+1} t \, \mathrm{d}t \\
&= \Big[\sin t \cos^{n+1} t\Big]_0^{\frac{\pi}{2}} - \int_0^{\frac{\pi}{2}} \sin t (n+1)(-\sin t) \cos^n t \, \mathrm{d}t \\
&= 0 + (n+1) \int_0^{\frac{\pi}{2}} \sin^2 t \cos^n t \, \mathrm{d}t \\
&= (n+1) \int_0^{\frac{\pi}{2}} (1 - \cos^2 t) \cos^n t \, \mathrm{d}t \\
&= (n+1) I_n - (n+1) I_{n+2}
\end{aligned}$$

D'où on tire que $(n+2)I_{n+2} = (n+1)I_n$ et donc que $I_{n+2} = \dfrac{n+1}{n+2} I_n$

[/2] **2.** Prouver que pour tout n entier naturel, $1 \geq \frac{I_{n+1}}{I_n} \geq \frac{n+1}{n+2}$.

Solution: Pour tout $t \in [0, \frac{\pi}{2}]$ on a $0 \leq \cos t \leq 1$.
Donc $\cos^n t \geq \cos^{n+1} t$.
Et donc on a $I_n \geq I_{n+1}$. La suite (I_n) est donc décroissante.
Donc pour tout entier n on a :

$$I_n \geq I_{n+1} \geq I_{n+2}$$

Et, en divisant par I_n et compte-tenu du résultat de la question précédente, il vient :

$$1 \geq \frac{I_{n+1}}{I_n} \geq \frac{n+1}{n+2}$$

[/2] **3.** **3.a** Calculer I_0 et I_1.

Solution:

$$\begin{aligned}
I_0 &= \int_0^{\frac{\pi}{2}} \mathrm{d}t \\
&= \frac{\pi}{2}
\end{aligned}$$

$$\begin{aligned} I_1 &= \int_0^{\frac{\pi}{2}} \cos t \, \mathrm{d}t \\ &= \big[\sin t\big]_0^{\frac{\pi}{2}} \\ &= 1 \end{aligned}$$

[/2] **3.b** Prouver que pour tout n entier naturel, $(n+1)I_{n+1}.I_n = \frac{\pi}{2}$.

Solution: Notons pour tout entier n : $\alpha_n = (n+1)I_{n+1}I_n$.
On a alors :

$$\begin{aligned} \alpha_{n+1} &= (n+2)I_{n+2}I_{n+1} \\ &= (n+2)\frac{n+1}{n+2}I_nI_{n+1} \\ &= (n+1)I_{n+1}I_n \\ &= \alpha_n \end{aligned}$$

La suite (α_n) est donc constante, égale au premier terme $\alpha_0 = I_1I_0 = \frac{\pi}{2}$.
Donc $\forall n \in \mathbb{N}$, $(n+1)I_{n+1}.I_n = \dfrac{\pi}{2}$.

[/3] **4.** En déduire un équivalent de $\sqrt{n}I_n$ quand n tend vers $+\infty$.

Solution: L'encadrement obtenu à la question **2.** montre que $I_{n+1} \sim I_n$.
(signalons au passage que ceci n'est pas forcément vrai pour n'importe quelle suite. Par exemple avec $u_n = e^n$ on a $u_{n+1} = e^{n+1} = eu_n$ et donc u_n et u_{n+1} ne sont pas équivalents)
Donc $(n+1)I_nI_{n+1} \sim (n+1)I_n^2 \sim nI_n^2$.
D'après la question précédente on a $nI_n^2 \to \dfrac{\pi}{2}$.
Donc $\sqrt{n}I_n \to \sqrt{\dfrac{\pi}{2}}$.
Donc au final $\sqrt{n}I_n \sim \dfrac{\sqrt{\pi}}{\sqrt{2}}$.

[/3] **5.** Prouver que pour tout n entier naturel :

$$\forall x \in [0, \sqrt{n}] \quad (1 - \frac{x^2}{n})^n \leq e^{-x^2} \leq \frac{1}{(1+\frac{x^2}{n})^n}$$

Indication : on pourra commencer par montrer que $\forall u \in \mathbb{R}$, $e^u \geq 1 + u$, puis appliquer cette inégalité en des valeurs judicieusement choisies.

Solution: Soit $f(u) = e^u - u - 1$. f est continue et dérivable sur $\mathbb{R}$ et $f'(u) = e^u - 1$, et donc f est décroissante sur $]-\infty, 0]$ et croissante sur $[0, +\infty[$. Elle admet donc un minimum absolu en 0. Or $f(0) = 0$, donc f est positive sur $\mathbb{R}$.
Soit $x \in [0, \sqrt{n}]$. En appliquant l'inégalité précédente à la valeur $-\frac{x^2}{n}$ il vient :

$$e^{-\frac{x^2}{n}} \geq 1 - \frac{x^2}{n}$$

Et en élevant à la puissance n il vient :

$$e^{-x^2} \geq (1 - \frac{x^2}{n})^n$$

En appliquant l'inégalité à la valeur $\frac{x^2}{n}$ il vient :

$$e^{\frac{x^2}{n}} \geq 1 + \frac{x^2}{n}$$

Et en élevant à la puissance n il vient :

$$e^{x^2} \geq (1 + \frac{x^2}{n})^n$$

et par passage à l'inverse pour ces deux quantités positives :

$$e^{-x^2} \leq \frac{1}{(1 + \frac{x^2}{n})^n}$$

[/1] **6.** **6.a** Transformer l'intégrale $\int_0^{\sqrt{n}} (1 - \frac{x^2}{n})^n dx$ en posant $x = \sqrt{n} \sin t$.

Solution: En posant $x = \sqrt{n} \sin t$ et donc $\mathrm{d}x = \sqrt{n} \cos t \, \mathrm{d}t$ on obtient :

$$\begin{aligned}
\int_0^{\sqrt{n}} (1 - \frac{x^2}{n})^n dx &= \sqrt{n} \int_0^{\frac{\pi}{2}} (1 - \sin^2 t)^n \cos t \, \mathrm{d}t \\
&= \sqrt{n} \int_0^{\frac{\pi}{2}} \cos^{2n+1} t \, \mathrm{d}t \\
&= \sqrt{n} I_{2n+1}
\end{aligned}$$

[/1] **6.b** Transformer l'intégrale $\int_0^{\sqrt{n}} \frac{1}{(1 + \frac{x^2}{n})^n} dx$ en posant $x = \sqrt{n} \tan t$.

Solution: En posant $x = \sqrt{n} \tan t$ et donc $\mathrm{d}x = \sqrt{n}(1 + \tan^2 t) \, \mathrm{d}t$ on obtient :

$$\begin{aligned}
\int_0^{\sqrt{n}} \frac{1}{(1 + \frac{x^2}{n})^n} \, \mathrm{d}x &= \int_0^{\frac{\pi}{4}} \sqrt{n} \frac{(1 + \tan^2 t) \, \mathrm{d}t}{(1 + \tan^2 t)^n} \\
&= \sqrt{n} \int_0^{\frac{\pi}{4}} (\cos t)^{2n-2} \, \mathrm{d}t \ \left(\text{car } 1 + \tan^2 t = \frac{1}{\cos^2 t}\right)
\end{aligned}$$

[/3] **7.** Donner un encadrement de $\int_0^{\sqrt{n}} e^{-x^2} \, \mathrm{d}x$ à l'aide de n, I_{2n+1} et I_{2n-2}.

Solution: En passant à l'intégrale l'encadrement obtenu à la question **5.**, il vient :

$$\int_0^{\sqrt{n}} (1 - \frac{x^2}{n})^n \, \mathrm{d}x \leq \int_0^{\sqrt{n}} e^{-x^2} \, \mathrm{d}x \leq \int_0^{\sqrt{n}} \frac{1}{(1 + \frac{x^2}{n})^n} \, \mathrm{d}x$$

C'est à dire :

$$\sqrt{n} I_{2n+1} \leq \int_0^{\sqrt{n}} e^{-x^2} \, \mathrm{d}x \leq \sqrt{n} \int_0^{\frac{\pi}{4}} (\cos t)^{2n-2} \, \mathrm{d}t$$

Et puisque on a de plus :

$$\int_0^{\frac{\pi}{4}} (\cos t)^{2n-2} \, \mathrm{d}t \leq \int_0^{\frac{\pi}{2}} (\cos t)^{2n-2} \, \mathrm{d}t$$

il vient finalement que :

$$\sqrt{n} I_{2n+1} \leq \int_0^{\sqrt{n}} e^{-x^2} \, \mathrm{d}x \leq \sqrt{n} I_{2n-2}$$

[/4] **8.** En déduire la convergence et la valeur des intégrales $\int_0^{+\infty} e^{-x^2}\,dx$ puis $\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}}\,dx$.

Solution: On a vu à la question **4.** que $I_n \sim \dfrac{\sqrt{\pi}}{\sqrt{2n}}$.
Donc $\sqrt{n}I_{2n+1} \sim \sqrt{n}\frac{\sqrt{\pi}}{\sqrt{4n+2}} \to \frac{\sqrt{\pi}}{2}$.
Et $\sqrt{n}I_{2n-2} \sim \sqrt{n}\frac{\sqrt{\pi}}{\sqrt{4n-4}} \to \frac{\sqrt{\pi}}{2}$.
Par le théorème des gendarmes, il vient que :

$$\int_0^{\sqrt{n}} e^{-x^2}\,\mathrm{d}x \longrightarrow \frac{\sqrt{\pi}}{2}$$

L'application $\varphi : y \mapsto \displaystyle\int_0^{y} e^{-x^2}\,\mathrm{d}x$ est croissante et on vient de voir que $\varphi(\sqrt{n})$ tend vers $\dfrac{\sqrt{\pi}}{2}$.
Donc φ tend vers $\dfrac{\sqrt{\pi}}{2}$ en $+\infty$.
En d'autres termes, $\displaystyle\int_0^{+\infty} e^{-x^2}\,\mathrm{d}x$ converge, et :

$$\int_0^{+\infty} e^{-x^2}\,\mathrm{d}x = \frac{\sqrt{\pi}}{2}$$

Par parité de la fonction $x \mapsto e^{-x^2}$ il vient que :

$$\int_{-\infty}^{+\infty} e^{-x^2}\,\mathrm{d}x = \sqrt{\pi}$$

Et en posant $x = \dfrac{t}{\sqrt{2}}$ et donc $\mathrm{d}x = \dfrac{\mathrm{d}t}{\sqrt{2}}$ on trouve la valeur de cette fameuse intégrale de Gauss :

$$\int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}}\,\mathrm{d}t = \sqrt{2\pi}$$

Total pour exo: /24

Total : /55