

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.

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].

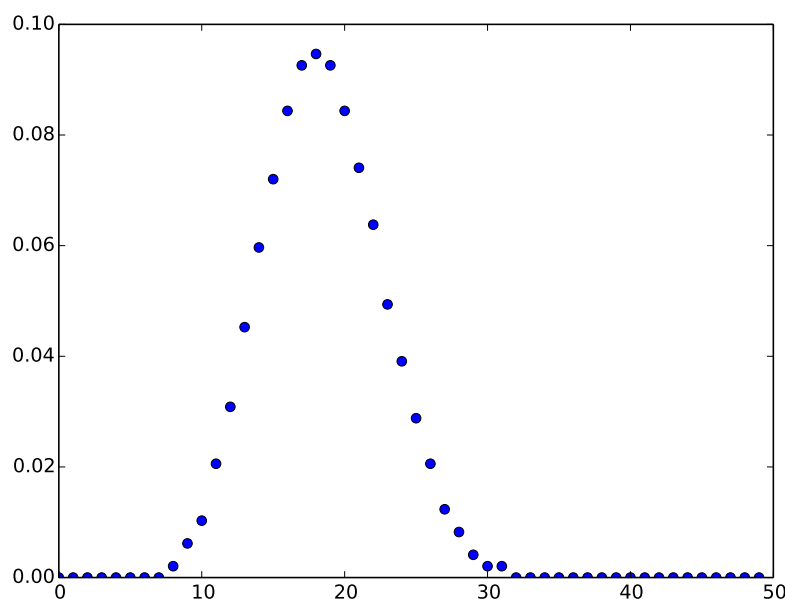
Ecrire une fonction Python nommé *creation-histo* qui prend en entrée une liste du type de *couv* et qui produit en sortie l'histogramme associé (attention : on ne demande pas ici de représenter graphiquement l'histogramme obtenu !)

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 *fréquences*=[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.

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 *fréquence[k]* en fonction de *k*) :

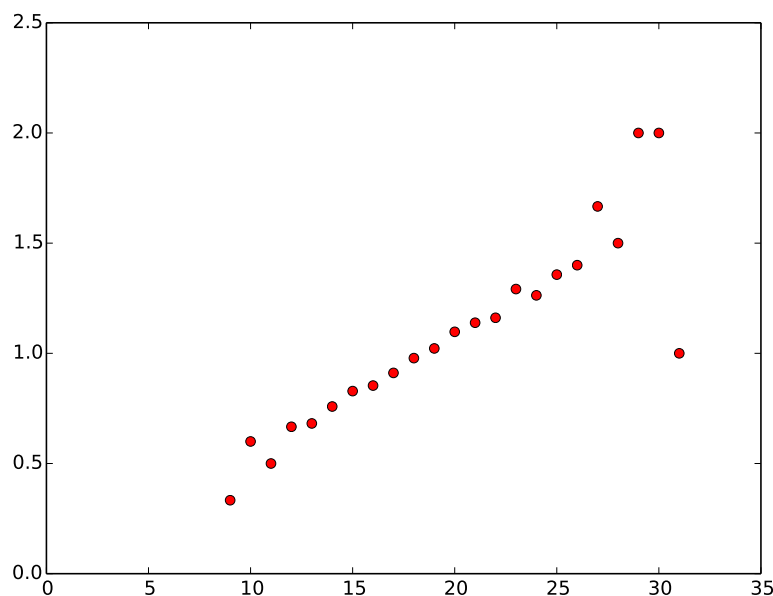


Pour nos calculs, on note f_k la fréquence associée à la valeur k (plutôt que $frequency[k]$).

2.a On étudie les variations de f_k en fonction de k . Pour ce faire on introduit la valeur $\alpha_k = \frac{f_{k-1}}{f_k}$ 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 $Lalphak$).

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 ?

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 :

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

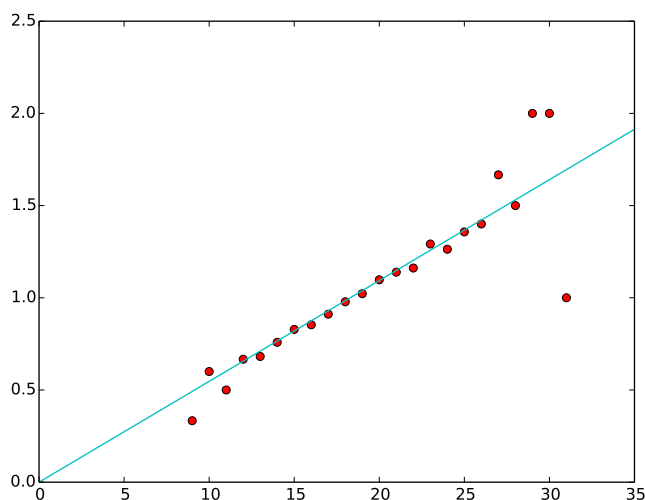
Montrer que $\varphi(a)$ est minimale pour la valeur :

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

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

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.e Commenter ce résultat.

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

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

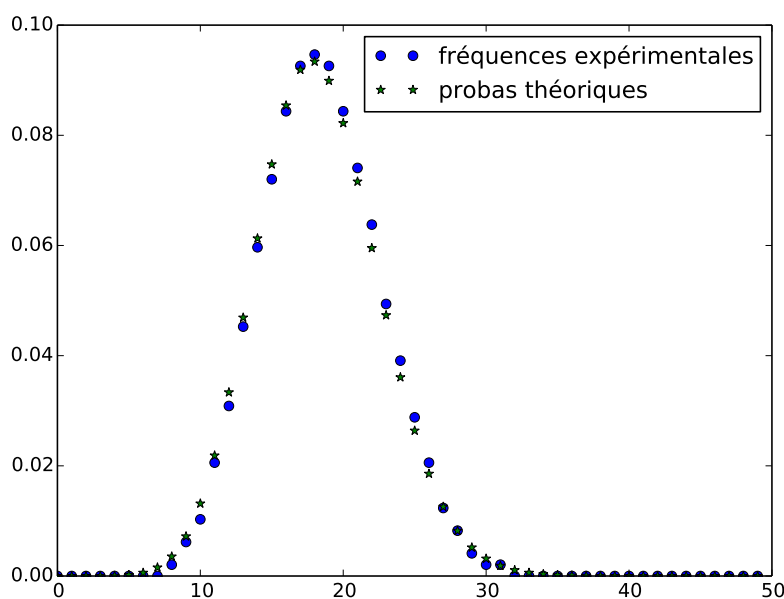
On pose $\lambda = \frac{1}{a_0}$.

3.a Exprimer p_1 et p_2 en fonction de p_0 et λ .

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

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 .

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



3.d Que pensez-vous de la pertinence de notre modèle ?

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 $\lambda = 18.3$

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

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.

Exercice :

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

1°) A l'aide d'une intégration par parties, prouver que pour tout n entier naturel :

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

2°) Prouver que pour tout n entier naturel, $1 \geq \frac{I_{n+1}}{I_n} \geq \frac{n+1}{n+2}$.

3°) **3.a** Calculer I_0 et I_1 .

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

4°) En déduire un équivalent de $\sqrt{n}I_n$ quand n tend vers $+\infty$.

5°) Prouver que pour tout n entier naturel :

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

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.

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

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

7°) Donner un encadrement de $\int_0^{\sqrt{n}} e^{-x^2} dx$ à l'aide de n, I_{2n+1} et I_{2n-2} .

8°) En déduire la convergence et la valeur des intégrales $\int_0^{+\infty} e^{-x^2} dx$ puis $\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx$.

