Statistique

06 November 2019

Dans ce TP, nous utiliserons plusieurs fonction de `numpy` et `scipy`. En particulier :

— Distribution aléatoire : package `numpy.random`, fonctions `normal`, `poisson`, `binomial`. En général ces fonctions ont un paramètres `size` qui permet de faire un grand nombres de tirages. Il est aussi possible de leur envoyer un tableau de paramètre.

— Histogramme : on pourra utiliser la fonction `numpy.histogram` ou la fonction `matplotlib.pyplot.hist`. Il est important d'utiliser l'argument optionel `bins` en mettant une séquence. Par exemple `numpy.histogram(data, bins=(a,b))` calcule le nombre de point entre a et b.

— Pour un v.a. discrete, on peut utiliser la fonction `numpy.unique` pour tracer l'histogramme.

— Dans le package scipy.stats, il existe des fonctions qui peuvent être utiles. Par exemple `scipy.stats.poisson.pmf` permet de calculer la fonction de distribution d'une loi de Poisson.

Distribution uniforme

On considère une variable aléatoire X qui suit une distribution uniforme entre 0 et 1. A partir d'un échantillon de N mesures, l'objectif est de déterminer le centre de cette distribution. On utilisera deux estimateurs : le premier est simplement la moyenne et le second est la moyenne entre le minimum et le maximum de l'échantillon. Comparez la distribution de ces deux estimateurs et leur erreur quadratique. On pourra prendre $N = 10$.

Loi de Poisson et lame semi-réfléchissante

La loi de Poisson est très présente en physique. Une raison est que cette loi est stable par certaine transformation. Par exemple, la somme de deux variables aléatoires qui suivent une loi de Poisson suit aussi une loi de Poisson. Réciproquement, si on réduit l'intensité par un processus aléatoire, alors on retrouve un loi de Poisson. C'est ce que l'on propose de vérifier.

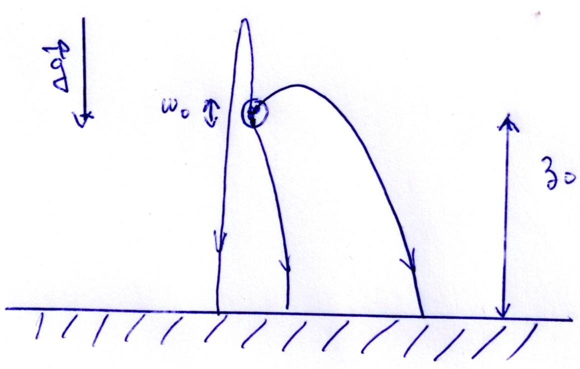
On considère une source de lumière. Les photons sortent selon un processus poissonien. La mesure du nombre de photon pendant un durée Δt suit donc une loi poissonienne de paramètre $\lambda = \Phi\Delta t$ où Φ est le flux de photon. On prendra $\lambda = 10$). Une lame semi-réfléchissante réfléchi aléatoirement un photon sur deux.

1. Si on suppose que pendant la durée Δt, le nombre de photons sortant de la source est k, quelle est la distribution de probabilité au niveau du détecteur, après la lame semi-réfléchissante ?
2. Vérifiez graphiquement que la distribution est une distribution poissonienne de paramètre $\lambda/2$. Calculez son espérance et sa valeur moyenne

3. On modifie maintenant le fonctionnement de la lame semi refléchissante et on suppose qu'elle réfléchi alternativement un photon sur deux. Dans le cas d'un nombre pair de photons on a donc $k/2$ photons et dans le cas impair, on $(k-1)/2$ ou $(k+1)/2$ photons avec une probabilité $1/2$. Vérifiez que la distribution n'est alors plus poissonienne. Calculez son espérance et sa valeur moyenne

Mesure de la gravité sur de l'antimatière

La matière attire la matière. A priori, c'est aussi le cas pour de l'antimatière. Par contre la question de savoir si de la matière attire ou pas l'antimatière et comment reste ouverte (et pourrait expliquer la dissymétrie matière/antimatière dans l'univers). Pour cela une expérience est en cours de réalisation au CERN. Il s'agit du projet GBAR qui consiste à laisser tomber un anti-atome d'hydrogène lent dans le champ de gravité de la Terre et à mesurer le temps de chute. Dans le meilleur des cas, on espère pouvoir réaliser 1000 fois l'expérience. Le but de ce problème est d'estimer l'incertitude sur la mesure de g à laquelle on s'attend.



A l'instant $t = 0$, la distribution en position des atomes est une gaussienne centrée autour de $z_0 = 30\mathrm{cm}$ avec un écart type w_0 de l'ordre de $1\mu m$ et une distribution en vitesse gaussienne d'écart type σ_v de l'ordre de $1\ m/s$ centrée en 0. On prendra $g = 10$ orienté vers le bas.

1. Ecrire l'équation du second degrés dérivant la trajectoire de l'atome selon l'axe vertical. Résoudre analytiquement pour déterminer le temps de chute.
2. Simuler un jeu de $N = 1000$ données et tracer l'histogramme.
3. A partir d'un jeu de données t_i, on peut calculer la moyenne du temps de chute $\langle t_i \rangle$ et ainsi estimer simplement g par la formule :

$$\hat{g}(t_1, ..., t_N) = \frac{2z_0}{\langle t_i \rangle^2}$$

   Implémenter et caractériser cet estimateur. Pour ce, on simulera l'expérience M fois pour calculer l'espérance et l'écart type de l'estimateur.
4. En supposant que l'incertitude sur la position initiale est négligeable, calculer analytiquement la distribution de probabilité de la v.a. T représentant le temps de chute. Vérifier graphiquement la fonction.
5. Implémenter et caractériser l'estimateur de maximum de vraisemblance.