TD4 : Exemples de lois de probabilité

Fabio Pietrucci, Alice Sinatra, Sorbonne Université

February 12, 2024

Motivation : nous avons besoin d'une théorie des probabilités pour comprendre l'entropie, étudier la marche aléatoire, introduire l'ensemble canonique (c'est-à-dire un grand système isolé divisé en une petite partie et une plus grande partie, les deux étant presque indépendantes), analyser les expériences et les trajectoires de simulation, etc. Une bonne introduction à la théorie des probabilités peut être trouvée dans le chapitre 2 de la référence [1].

1 Échantillonnage avec et sans remise

Les schemas d'échantillonnage sont dits "sans remise" si aucun élément ne peut être sélectionné plus d'une fois dans le même échantillon, ou "avec remise" si un élément peut apparaître plusieurs fois dans un même échantillon[1]. En anglais "sampling with/without replacement".

(a) Quelle est la probabilité qu'en lançant un dé 6 fois on obtienne 6 faces différentes ?

(b) De même, quelle est la probabilité que dans un groupe de M étudiants tous les anniversaires soient différents ?

(c) Quelle est la probabilité d'obtenir 10 têtes sur 10 lancers de pièces de monnaie ?

2 Loi binomiale

Pour n évènements indépendants, chacun pouvant produire deux résultats (pile ou face) avec probabilité respective p et $q = 1 - p$, la probabilité que l'évènement "pile" se produise k fois sur les n est donnée par la distribution binomiale

$$P(k; n, p) = \binom{n}{k} p^k q^{n-k}$$

Considérons n atomes d'un gaz parfait dans un volume V : quelle est la probabilité d'observer k atomes dans une sous-région du volume v ? Estimez la moyenne $\langle k \rangle$ et les fluctuations. Quelle est la probabilité d'observer $k = 0$ ou $k = n$ pour $v \ll V$ et $n \gg 1$?

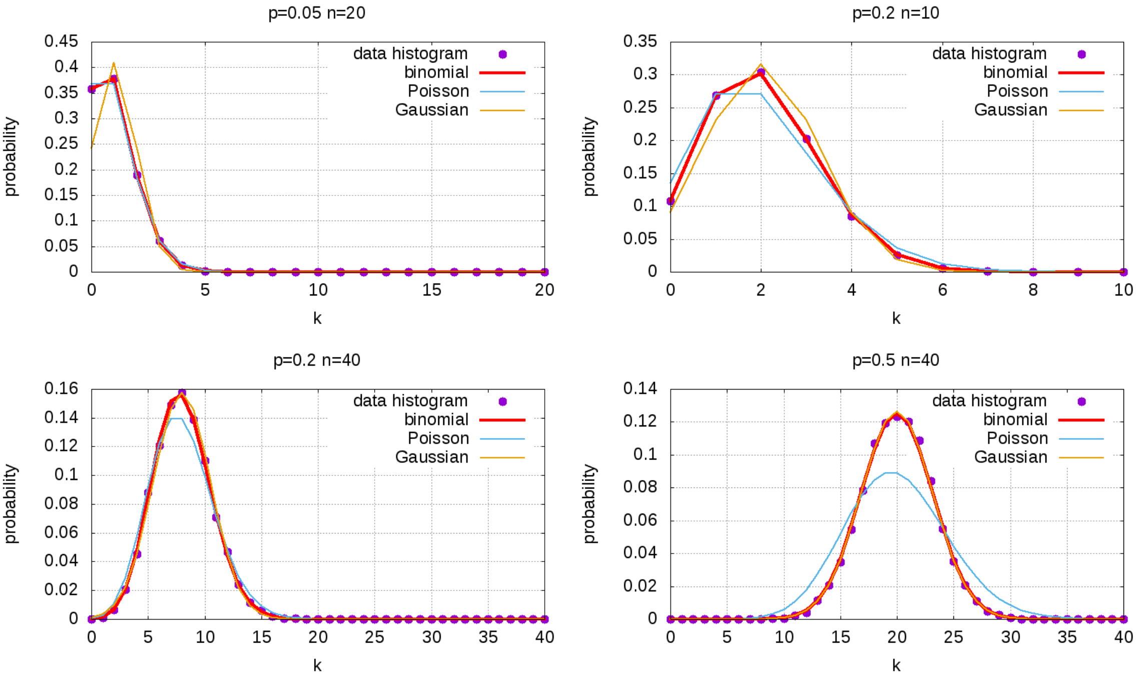
Imaginons que le gaz, initialement confiné dans le sous-volume v, soit relaché dans le volume V. Peut-on utiliser les résultats précédents pour faire un commentaire sur l'irréversibilité macroscopique ?

3 Loi de Poisson et loi gaussienne

La loi de Poisson décrit le comportement du nombre d'événements indépendants se produisant dans un intervalle de temps fixé avec une fréquence connue. Si le nombre moyen d'occurrences dans un intervalle de temps fixé est λ, alors la probabilité qu'il existe exactement k occurrences (k est entier) est

$$P_\lambda(k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

[1] Si c'était du Loto, on peut imaginer que "avec remise" veut dire qu'après avoir tiré un numéro on le remet dans l'urne, si bien qu'on peut le tirer une deuxième fois.




```
awk -v p=$1 -v n=$2 -v nexp=$3 'BEGIN{
  for(k=0;k<=n;k++){
    prob[k]=0             # here we will store the probability histogram
  }
  for(i=0;i<nexp;i++){   # we repeat many times the whole experiment
    k=0                   # number of successes
    for(j=0;j<n;j++){    # we draw n random numbers
      r=rand()            # random number uniformly distributed between 0 and 1
      if (r<p) { k++ }    # true = success, probability of success = p
    }
    prob[k]++             # we increment the histogram
  }
  print "# p n np npq =",p,n,n*p,n*p*(1.0-p)
  for(k=0;k<=n;k++){
    printf "%4d %9.6f\n", k,prob[k]/nexp
  }
}'
```

Figure: Example of code (awk) to generate histograms of Bernoulli trials

```
awk -v p=$1 -v n=$2 'BEGIN{
  fact[0]=1
  for (i=1;i<=n;i++) {
    fact[i]=fact[i-1]*i
  }
  pi=3.14159265359
  print "# k binomial Poisson Gauss"
  for (k=0;k<=n;k++) {
    pb=(fact[n]/(fact[k]*fact[n-k]))*(p**k)*((1.0-p)**(n-k))
    pp=exp(-n*p)*((n*p)**k)/fact[k]
    pg=exp(-((k-n*p)**2)/(2*n*p*(1.0-p)))/sqrt(n*p*(1.0-p)*2*pi)
    printf "%4d %9.6f %9.6f %9.6f\n", k,pb,pp,pg
  }
}'
```

Figure: Example of code (awk) to print the three distributions

Figure 1: Histogrammes de données (10^4 expériences) par rapport aux distributions théoriques.

3.1

Obtenir la distribution de Poisson à partir de la distribution binomiale dans la limite des grands n avec $np = \lambda$ fixé.

3.2

Calculez les premier et deuxième moments des distributions binomiale et de Poisson à l'aide de la fonction génératrice (des moments ou des cumulants).

3.3

Loi normale est une loi de probabilité continue qui dépend de deux paramètres : son espérance μ et son écart type σ.

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Soit $p_B(k; n, p)$ la loi binomiale. Montrez que dans la limite des grands n, avec $n \to \infty$ et $k \to \infty$ à p fixé, $\ln[p_B(k; n, p)]$ a un maximum en $k = \langle k \rangle = np$.

Dans la même limite, obtenir la distribution normale (à une constante de normalisation près) à partir de la distribution binomiale, en développant $\ln[p_B(k; n, p)]$ au second ordre autour de la valeur maximale.

3.4

Calculer la constante de normalisation de la distribution normale.

3.5 À la maison : sur la distribution de Poisson

Dans un moteur de recherche (par exemple, leboncoin) contez le nombre de résultats à une distance fixe d'une adresse choisie. Changez ensuite l'adresse (toujours en lassant la distance fixée) et regardez pour différents adresses si le nombre de réponses suit une distribution de Poisson, c'est-à-dire si les résultats sont répartis géographiquement de manière aléatoire.

3.6 À la maison : sur le théorème de la limite centrale

Tracer la distribution pour la somme de 3 ou 4 variables aléatoires (taille des élèves, pointure, ...).

3.7 À la maison (très important!) : les barres d'erreur

Réfléchissez aux raisons pour lesquelles on peut estimer avec $\sigma_A/\sqrt{M}$ la barre d'erreur sur la moyenne $\langle A \rangle$ pour un jeu $\{A_k\}_{k=1,\ldots,M}$ de M valeurs statistiquement décorrélées d'une observable A, tandis que le seul écart-type σ_A n'est pas une bonne barre d'erreur.

References

[1] Christophe Texier and Guillaume Roux. *Physique statistique: des processus élémentaires aux phénomènes collectifs.* Dunod, 2020.