CHAPITRE 1

Probabilité et dénombrement

Objectifs :

- utiliser une distribution de probabilité discrète
- utiliser une distribution de probabilité continue
- dénombrer des configurations

La notion de probabilité intervient de trois façons différentes dans notre description de la nature.

- Depuis la découverte de la mécanique quantique, de nombreuses expériences montrent que le comportement d'un objet à l'échelle microscopique est intrinsèquement probabiliste. C'est-à-dire, que le résultat d'une mesure d'un système physique préparé quantiquement ne peut être prévu qu'avec une certaine probabilité.
- En dehors de cet aspect probabiliste de la physique quantique, les incertitudes sur les conditions initiales d'un système physique ou la complexité de la chaine d'événements que subit le système au cours de son évolution peut rendre le résultat d'une expérience imprédictible. Par exemple, un lancer de dé obéit aux lois de la mécanique classique mais la complexité des rebonds du dé sur la table rend la prédiction de la mesure impossible.
- Les quantités physiques que nous pouvons mesurer sur un système physique sont bien souvent des valeurs moyennes. Prenons l'exemple d'une mesure de la température d'un gaz. Cette mesure de la température est une mesure de l'énergie cinétique moyenne des particules. Étant donné que les particules constitutives du gaz n'ont pas toutes la même énergie cinétique, le calcul de la valeur moyenne de l'énergie cinétique fait intervenir la probabilité de trouver une particule donnée dans un certain intervalle d'énergie cinétique.

C'est cette dernière situation que nous allons rencontrer en permanence en physique statistique. En effet, nous allons principalement chercher dans ce cours à expliquer les propriétés macroscopiques de la matière à partir des lois microscopiques.

A ce tire, nous allons commencer le cours de physique statistique par un chapitre sur les probabilités.

1.1 Distribution de probabilités discrètes

1.1.1 Définition

De manière générale, nous notons p la probabilité d'apparition d'un événement. Il est absolument certain qu'un événement advienne si sa probabilité est de 1. Au contraire, un événement qui n'a aucune chance de se produire à une probabilité de 0.

Nous allons nous intéresser pour l'instant uniquement à une variable aléatoire qui peut prendre des valeurs discrètes x_i avec $i = 0, ..., N$. Nous notons x une telle variable aléatoire et p_i la probabilité d'apparition de la valeur x_i. **L'ensemble des p_i constitue une distribution de probabilités discrètes**.

Notons que tous les événements "une valeur quelconque de x apparait" sont des événements liés dans le sens où une valeur apparait à la place des autres. **les probabilités d'apparition d'événements qui ne sont pas indépendants s'additionnent**.

Exemple

Considérons un lancer de dé à 6 faces. La probabilité que le dé s'arrête sur l'une de ses faces est de $p_i = \frac{1}{6}$. Ainsi, la probabilité d'obtenir un 6 est de $p_6 = \frac{1}{6}$. La probabilité d'obtenir un nombre strictement supérieur à 4 est $p_5 + p_6 = \frac{1}{3}$.

Considérons maintenant l'événement suivant : la valeur que prend x est supérieure ou égale à la plus petite des valeurs possibles x_i. La probabilité correspondant à cet événement est **la somme des probabilités correspondant à chaque occurrence possible x_i de x**. Étant donné qu'un événement certain a une probabilité de 1, nous avons donc par définition :

$$\boxed{\sum_{i=0}^{N} p_i = 1 \quad \text{et} \quad 0 \leqslant P_i \leqslant 1, \forall i} \tag{1.1}$$

où N est le nombre de valeurs possibles de la variable x. Autrement dit, une probabilité est une quantité normalisée à 1.

Supposons maintenant que nous lancions deux dés simultanément, la probabilité d'obtenir un double 6 est de $p_6.p_6 = \frac{1}{36}$. Autrement dit, **les probabilités se multiplient si les deux événements sont indépendants. Les événements sont indépendants au sens où il n'y a aucun lien physique entre les deux lancer de dés**.

Exemple

Nous pouvons calculer la probabilité d'obtenir au moins un 6 lors du lancer de deux dés. Nous avons deux cas de figures à considérer, soit nous obtenons un 6 avec le premier dé soit nous obtenons un 6 avec le deuxième dé. Les probabilités de ces deux événements s'additionnent puisque ces événements ne sont pas indépendants. Nous obtenons donc :

$$\begin{aligned} p_{\text{au moins un 6}} &= \frac{1}{6}.(p_1 + p_2 + p_3 + p_4 + p_5 + p_6) + \frac{5}{6}.\frac{1}{6} \\ &= \frac{1}{6}.1 + \frac{5}{6}.\frac{1}{6} \\ &= \frac{11}{36} \end{aligned}$$

1.1.2 Valeur moyenne

Définition

La valeur moyenne $< x >$ de la variable x est définie par :

$$\boxed{< x >= \sum_{i=0}^{N} x_i p_i} \tag{1.2}$$

Nous pouvons ainsi calculer la valeur de n'importe quelle fonction f de x par :

$$\boxed{< f(x) >= \sum_{i=0}^{N} f(x) p_i} \tag{1.3}$$

En particulier, nous pouvons calculer $< x^2 >= \sum\limits_{i=0}^{N} x_i^2 p_i$.

Exemple

Considérons un lancer de dé à 6 faces. La probabilité que le dé s'arrête sur l'une de ses faces est de $p_i = \frac{1}{6}$ et $x = \{1, 2, 3, 4, 5, 6\}$. Nous avons donc $< x >= \sum\limits_{i=0}^{N} x_i p_i = \frac{1}{6} + \frac{2}{6} + \frac{3}{6} + \frac{4}{6} + \frac{5}{6} + \frac{6}{6} = \frac{21}{6} = 3.5$.

Translation

Nous allons établir ici une propriété intéressante de la valeur moyenne. Posons $y = ax + b$ où a et b sont des constantes. La valeur moyenne de y est donnée par :

$$\begin{aligned}
< y >= &< ax + b > \\
= &\sum_{i=0}^{N} (ax + b) p_i \\
= &a \sum_{i=0}^{N} x p_i + b \sum_{i=0}^{N} p_i \\
= &a < x > + b
\end{aligned}$$

De même, si nous avons $y = ax^2 + bx$ où a et b sont des constantes, nous avons :

$$\begin{aligned}
< y >= &< ax^2 + bx > \\
= &\sum_{i=0}^{N} (ax^2 + bx) p_i \\
= &a \sum_{i=0}^{N} x^2 p_i + b \sum_{i=0}^{N} x p_i \\
= &a < x^2 > + b < x >
\end{aligned}$$

1.1.3 Moments

Variance

Nous allons maintenant chercher à caractériser la dispersion des valeurs de x autour de la valeur moyenne $< x >$. La dispersion d'une valeur particulière de x est donnée par $x_i - < x >$ mais cette notion n'est pas intéressante puisqu'elle ne caractérise pas la dispersion de l'ensemble des valeurs. Nous pouvons par contre calculer $< x - < x >>$. Cependant, d'après la propriété de translation de la valeur moyenne, la quantité $< x - < x >>$ est nulle et ne peut donc pas représenter la dispersion des valeurs autour de la moyenne.

Nous allons ainsi introduire **la variance pour caractériser la dispersion autour de la moyenne** qui est définie par :

$$\boxed{\sigma_x^2 =< (x- < x >)^2 >} \tag{1.4}$$

La racine carré de la variance s'appelle l'écart-type. L'expression précédente de la variance peut se réécrire :

$$\begin{aligned}
\sigma_x^2 &= < (x- < x >)^2 > \\
&= < x^2 - 2x < x > + < x >^2> \\
&= < x^2 > -2 < x >^2 + < x >^2 \\
&= < x^2 > - < x >^2
\end{aligned}$$

Nous pouvons donc retenir le résultat suivant :

$$\boxed{\sigma_x^2 =< x^2 > - < x >^2} \tag{1.5}$$

Exemple

Considérons un lancer de dé à 6 faces. La probabilité que le dé s'arrête sur l'une de ses faces est de $p_i = \frac{1}{6}$ et $x = \{1, 2, 3, 4, 5, 6\}$. Ainsi $< x^2 >= \sum_{i=0}^{N} x_i^2 p_i = \frac{1}{6} + \frac{4}{6} + \frac{9}{6} + \frac{16}{6} + \frac{25}{6} + \frac{36}{6} = \frac{91}{6}$ et $\sigma_x^2 = \frac{546}{36} - \frac{441}{36} = \frac{105}{36}$ d'où $\sigma_x = \sqrt{\sigma_x^2} = 1.7$.

Translation

Posons $y = ax + b$, nous pouvons montrer que $\sigma_y^2 = a^2\sigma_x^2$. En effet :

$$\begin{aligned}
\sigma_y^2 &= < (ax + b)^2 > - < ax + b >^2 \\
&= < a^2x^2 + 2abx + b^2 > -(a < x > +b)^2 \\
&= a^2 < x^2 > +2ab < x > -a^2 < x >^2 -2ab < x > \\
&= a^2(< x^2 > - < x >^2) \\
&= a^2\sigma_x^2
\end{aligned}$$

définition des moments

De manière générale, le k-ième moment d'un ensemble de valeurs est définie par :

$$\boxed{\sigma_x^k =< (x- < x >)^k >} \tag{1.6}$$

Ainsi, σ^2 est la variance, σ^3 est le coefficient d'asymétrie (σ^3 est positif si le nombre de valeurs élevées est important et négatif dans le cas contraire) et σ^4 est le kurtosis.

1.2 Distribution de probabilité continue

1.2.1 Définition

Nous avons vu dans la section précédente une variable aléatoire qui peut prendre uniquement des valeurs discrètes. La distribution de probabilité associée est alors également discrète.

Nous considérons maintenant une variable aléatoire x qui peut prendre n'importe quelle valeur. Nous parlons dans ce cas de variable aléatoire continue et la distribution de probabilité associée est une distribution de probabilité continue. Puisque la variable x peut prendre une infinité de valeurs, la probabilité d'obtenir une valeur donnée est donc nulle.

Exemple

Si nous reprenons l'exemple d'un dé à N face, la probabilité d'obtenir une valeur donnée lors d'un lancer de dé est de $p = \frac{1}{N}$. Supposons maintenant que le nombre de faces tendent vers l'infini, la probabilité est donc $p = \lim\limits_{N\to+\infty} \frac{1}{N} = 0$.

Nous allons donc introduire la probabilité d'obtenir une valeur comprise dans un intervalle infiniment petit dx. Ainsi, une distribution de probabilité continue est définie par la relation :

$$\boxed{P(x)dx \text{ représente la probabilité que la mesure de } x \text{ soit entre } x \text{ et } x + dx} \tag{1.7}$$

La probabilité correspondant à la mesure d'une valeur quelconque de x est de 1, nous avons donc :

$$\boxed{\int P(x)dx = 1} \tag{1.8}$$

1.2.2 Valeur moyenne et moments

La valeur moyenne de x est définie par :

$$\boxed{< x >= \int xP(x)dx} \tag{1.9}$$

La valeur moyenne d'une fonction $f(x)$ est définie par :

$$\boxed{< f(x) >= \int f(x)P(x)dx} \tag{1.10}$$

Exemple

Considérons une particule qui rebondit entre deux murs situés entre les points $x = 0$ et $x = 4l$, la distribution de probabilité est donc égale à $P(x) = \frac{1}{4l}$. La valeur moyenne de la position de la particule est donnée par $< x >= \int_0^{4l} \frac{x}{4l} dx = \frac{1}{4l} 8l^2 = 2l$.

Cette dernière définition permet de retrouver la notion de moyenne temporelle définie par $< f(t) >= \frac{1}{T} \int_t^{t+T} f(t)dt$. Cette définition implique que la distribution de probabilité est uniforme entre t et $t+T$ et est donc égale à $P(t) = \frac{1}{T}$.

La définition des différents moments de la variable aléatoire x est identique au cas discret et nous avons donc :

$$\boxed{\sigma_x^k =< (x- < x >)^k >} \tag{1.11}$$

Nous avons en particulier, la variance qui est définie par :

$$\boxed{\sigma_x^2 =< x^2 > - < x >^2} \tag{1.12}$$

Exemple

Considérons à nouveau une particule qui rebondit entre deux murs situés entre les points $x = 0$ et $x = 4l$, la distribution de probabilité est égale à $P(x) = \frac{1}{4l}$ et $< x >= 2l$. Nous avons $< x^2 >= \int_0^{4l} \frac{x^2}{4l} dx = \frac{1}{4l} \frac{64l^3}{3} = \frac{64}{12} l^2$. Ainsi, la variance a pour expression $\sigma_x^2 = \frac{64}{12} l^2 - \frac{48}{12} l^2 = \frac{4}{3} l^2$ d'où $\sigma_x = \frac{4l}{\sqrt{12}}$.

Complément

La fonction $M(t) = \langle e^{tx} \rangle$ est appelée la fonction génératrice des moments. En effet, nous avons par définition :

$$\begin{aligned} M(t) =& \int_{-\infty}^{+\infty} e^{tx} P(x) dx \\ =& \int_{-\infty}^{+\infty} (1 + tx + \frac{1}{2} t^2 x^2 + ...) P(x) dx \\ =& 1 + t \langle x \rangle + \frac{1}{2} t^2 \langle x^2 \rangle + ... \end{aligned}$$

Ainsi :

$$\begin{aligned} M(t) &= \langle e^{tx} \rangle & M(0) &= 1 \\ M'(t) &= \langle x e^{tx} \rangle & M'(0) &= \langle x \rangle \\ M''(t) &= \langle x^2 e^{tx} \rangle & M''(0) &= \langle x^2 \rangle \end{aligned}$$

1.3 Indépendance et corrélation

1.3.1 Variables indépendantes

Soient x et y deux variables aléatoires continues. La relation $< x+y >=< x > + < y >$ est vrai sans hypothèses particulières sur x et y, par contre la relation $< xy >=< x >< y >$ est vrai uniquement dans le cas où les variables x et y sont indépendantes. Dans ce cas, nous avons :

$$\begin{aligned} < xy > &= \iint xP(x)yP(y)dxdy \\ &= \int xP(x)dx \int yP(y)dy \quad \text{car } x \text{ et } y \text{ sont indépendantes} \\ &= < x >< y > \end{aligned}$$

et

$$\begin{aligned} \sigma^2_{x+y} &= < (x+y)^2 > - < x+y >^2 \\ &= < x^2 > +2 < xy > + < y^2 > - < x^2 > - < y >^2 -2 < x >< y > \\ &= \sigma^2_x + \sigma^2_y \end{aligned}$$

1.3.2 N variables indépendantes

Considérons le cas de n mesures successives de la quantité x d'un même système physique entaché d'une incertitude σ_x. Nous nommons x_i la quantité x mesurée lors de la i-ème mesure. Toutes les quantités x_i ont dans ce cas la même valeur moyenne $< x >$ et la même variance σ_x. Soit la variable $y = x_1 + x_2 + ... + x_n$. La valeur moyenne de y est alors donnée par $< y >=< x_1 > + < x_2 > +...+ < x_n >= n < x >$. Ainsi, $< y >$ a la même valeur que $< x >$ en divisant $< y >$ par le nombre de mesures réalisées. Nous allons maintenant montrer que l'écart-type sur la quantité $\frac{y}{n}$ est moindre que l'écart-type sur la quantité x. Pour ce faire, nous allons calculer la variance σ^2_y :

$$\begin{aligned} \sigma^2_y &= < y^2 > - < y >^2 \\ &= < x_1^2 + ... + x_n^2 + x_1x_2 + x_1x_3 + ...x_2 > -n^2 < x >^2 \\ &= n < x^2 > +n(n-1) < x >^2 -n^2 < x >^2 \\ &= n < x^2 > -n < x >^2 \\ &= n\sigma^2_x \end{aligned}$$

Ainsi, **l'écart-type sur la variable $\frac{y}{n}$ a pour expression $\frac{\sigma_x}{\sqrt{n}}$**. Réaliser plusieurs mesures successives et travailler avec la somme des valeurs mesurées divisée par le nombre de mesures permet donc de diminuer les incertitudes sur la valeur moyenne.

1.3.3 Coefficient de corrélation

Nous voulons maintenant évaluer quantitativement la corrélation entre deux quantités x et y. Nous allons calculer le coefficient de corrélation r entre ces deux quantités. Le calcul de r donne une valeur numérique positive, négative ou nulle. Une valeur nulle de r signifie que les deux quantités x et y ne sont pas corrélées, une valeur positive indique que y augmente si x augmente et une valeur négative de r indique que y diminue si x augmente. Par exemple, le temps d'attente au téléphone d'un service après-vente est corrélé à la satisfaction du service mais le coefficient de corrélation entre les deux quantités est négatif.

Nous allons calculer le coefficient de corrélation entre deux variables aléatoires. Considérons par exemple trois variables aléatoires indépendantes u, v et w. Puisque ces variables sont indépendantes, la corrélation entre deux de ces variables est nulle. Construisons maintenant les variables aléatoires $x = u + w$ et $y = v + w$, les variables aléatoires x et y sont maintenant corrélées puisque la variable w est une variable "cachée" commune aux deux variables aléatoires.

La corrélation entre les deux variables aléatoires x et y peut être mesurée grâce au coefficient de corrélation r définie par :

$$\begin{aligned} r &= \frac{< xy > - < x >< y >}{\sigma_x \sigma_y} \\ &= \frac{\mathrm{Cov}(x, y)}{\sigma_x \sigma_y} \end{aligned}$$

où $\mathrm{Cov}(x, y) =< xy > - < x >< y >$ est la covariance. Cette définition de r implique bien que $r = 0$ pour des variables indépendantes, le dénominateur étant choisi pour avoir un coefficient de corrélation compris entre -1 et 1. Nous pouvons également noter que $r = 1$ si $x = y$. La notion de corrélation intervient notamment en physique statistique pour caractériser une transition de phase.

Exemple

Reprenons nos deux variables aléatoires $x = u + w$ et $y = v + w$ et considérons, par exemple, que $< u >=< v >=< w >= 0$ ainsi que $\sigma_u = \sigma_v = \sigma_w = 1$. Nous avons alors $\sigma_x = \sqrt{2}$ et $\sigma_y = \sqrt{2}$ d'où :

$$\begin{aligned} r &= \frac{< xy >}{\sqrt{2}\sqrt{2}} \\ &= \frac{< (u + w)(v + w) >}{2} \\ &= \frac{< (u + w)(v + w) >}{2} \\ &= \frac{< uv + uw + wv + w^2 >}{2} \end{aligned}$$

Les variables u, v et w sont indépendantes, nous avons donc $< uv >=< u >< v >= 0$ et, de même, $< uw >= 0$ et $< wv >= 0$. Par ailleurs, nous avons $\sigma_w^2 =< w^2 >$ puisque $< w >= 0$ donc :

$$r = \frac{1}{2}$$

1.4 Dénombrement

Lors du calcul de probabilités, il est important de pouvoir dénombrer les différentes configurations qui mènent au même événement. Il existe quatre modèles fondamentaux de dénombrements. Il s'agit dans tous les cas de choisir p éléments parmi n. On utilise une boîte qui contient n boules discernables, nous pouvons par exemple imaginer que toutes les boules sont numérotées.

- **Tirage de p éléments ordonnées parmi n sans remise :** l'ordre est important, ainsi tirer les boules {1,5,3} et {5,1,3} sont des événements différents. Il y a n possibilités pour choisir la première boule, puis $n-1$ possibilités pour choisir la deuxième boule, ... , puis $n-p+1$

possibilités pour choisir la p-ième boule. Le nombre de configurations possibles est donc :

$$\boxed{A_n^p = n(n-1)(n-2)...(n-p+1) = \frac{n!}{(n-p)!}} \qquad (1.13)$$

où A_n^p est le nombre d'arrangements possibles.

- **Tirage de p éléments ordonnées parmi n avec remise :** L'ordre est important mais il y a toujours n possibilités à chaque tirage puisque la boule choisie est remise. Il y a donc n^p configurations possibles.
- **Tirage de p éléments non ordonnées parmi n sans remise :** l'ordre est sans importance cette fois-ci. le début du raisonnement est le-même que dans le cas où l'ordre a de l'importance. Il y a n possibilités pour choisir la première boule, puis $n-1$ possibilités pour choisir la deuxième boule, ... , puis $n-p+1$ possibilités pour choisir la p-ième boule. Le nombre d'arrangements correspondants doit être ensuite divisé par le nombre de permutations des p éléments puisque l'ordre n'a pas d'importance. Le nombre de configurations possibles est donc :

$$\boxed{\binom{n}{p} = \frac{n!}{(n-p)!p!}} \qquad (1.14)$$

- **Tirage de p éléments non ordonnées parmi n avec remise :** le calcul est plus complexe dans ce cas. Les n boules sont toujours présentes dans la boîte dans ce cas, nous pouvons ainsi considérer qu'il y a $n-1$ séparateurs entre les boules. Tirer p éléments revient donc à choisir une configuration parmi $n-1+p$ emplacements. Le nombre de permutations de ces $n-1+p$ emplacements est $(n-1+p)!$. Les séparateurs sont indiscernables, l'ordre des séparateurs et des boules n'a donc pas d'importance et nous devons donc diviser le nombre de configurations précédentes par $(n-1)!p!$. Le nombre de configurations recherchées est donc égale à $\frac{(n-1+p)!}{(n-1)!p!}$.