CHAPITRE 12
Probabilités sur un univers dénombrable

Plan du chapitre

Introduction

Dans le chapitre précédent portant sur les probabilités, nous avons étudié les phénomènes aléatoires admettant un nombre fini d'issues possibles. Dans ce nouveau chapitre, nous allons nous intéresser à certaines expériences aléatoires admettant une infinité d'issues possibles. Pour éviter une complexité théorique trop importante, nous nous limiterons à des cas où l'univers est dénombrable, ce qui signifie intuitivement que l'on peut énumérer ses éléments.

Les expériences aléatoires admettant une infinité d'issues possibles apparaissent naturellement en science. On peut par exemple étudier la longévité en année d'une ampoule ($\Omega = \mathbb{N}$) ou la durée de vie d'un atome radioactif ($\Omega = \mathbb{R}_+$). En théorie des jeux, on peut également s'intéresser au nombre de lancers nécessaire d'une pièce jusqu'à obtenir pour la première fois pile ($\Omega = \mathbb{N}^*$).

Dans un premier temps, nous définirons la notion d'ensemble dénombrable, puis nous étendrons la définition d'espace probabilisé ainsi que ses propriétés à ce nouveau cadre. Dans un second temps, nous étudierons les variables aléatoires dans ce nouveau contexte. Nous verrons notamment qu'une variable aléatoire n'admet plus nécessairement une espérance et une variance.

I - Espaces probabilisés dénombrables

I.A - Ensembles dénombrables

Définition (Ensemble dénombrable) : Un ensemble Ω est dit dénombrable s'il existe une suite $(\omega_n)_{n\in\mathbb{N}}$ d'éléments de Ω telle que $\Omega = \{\omega_n \mid n \in \mathbb{N}\}$.

Remarque 1 : Un ensemble est dénombrable si on peut « énumérer » ses éléments.

Exemples 1 :

a) Les ensembles finis sont dénombrables.

b) Les ensembles $\mathbb{N}, \mathbb{Z}$ et $\mathbb{Q}$ sont dénombrables.

c) L'ensemble $\mathbb{R}$ n'est pas dénombrable.

I.B - Probabilités sur un univers dénombrable

Jusqu'à la fin du chapitre, on fixe un univers Ω **dénombrable** modélisant une expérience aléatoire. Nous allons étendre les définitions et les résultats vus dans le cas où Ω est fini.

Définition (Probabilité) : On appelle probabilité sur l'univers Ω toute application $P : \mathscr{P}(\Omega) \to [0,1]$ vérifiant

(i) $P(\Omega) = 1$,

(ii) Pour toute suite $(A_n)_{n\in\mathbb{N}}$ d'événements deux à deux incompatibles

$$P\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n=0}^{+\infty} P(A_n).$$

Remarque 2 : Dans le cas où Ω est fini, la définition ci-dessus est équivalente à celle vue dans le chapitre précédent portant sur les probabilités.

Définition (Espace probabilisé) : Un espace probabilisé (dénombrable) est un couple (Ω, P) où Ω est un univers fini et P une probabilité sur Ω.

Proposition 1 : Soit $P : \Omega \to [0,1]$ une probabilité sur l'univers Ω. Alors

(i) $P(\varnothing) = 0$,

(ii) Si $A_1, \ldots, A_n$ sont des évènements deux à deux incompatibles, alors

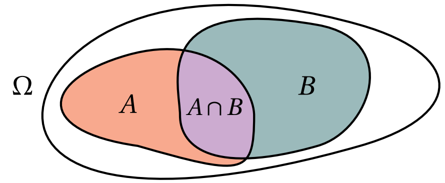
$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i),$$

(iii) $\forall A \in \mathscr{P}(\Omega)$, $P(\bar{A}) = 1 - P(A)$,

(iv) $\forall (A,B) \in \mathscr{P}(\Omega)^2$, $A \subset B \Rightarrow P(A) \leqslant P(B)$ et $P(B \setminus A) = P(B) - P(A)$,

(v) $\forall (A,B) \in \mathscr{P}(\Omega)^2$, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Illustration : Le schéma ci-dessous permet de comprendre l'origine de la formule du point (v) de la proposition.



Proposition 2 : Si $p : \Omega \to [0,1]$ est une application vérifiant $\sum_{\omega\in\Omega} p(\omega) = 1$, alors il existe une et une seule probabilité $P : \mathscr{P}(\Omega) \to [0,1]$ vérifiant

$$\forall \omega \in \Omega, \quad P(\{\omega\}) = p(\omega).$$

Remarque 3 : Autrement dit, pour définir une probabilité P sur un univers dénombrable Ω, il suffit de définir P sur chaque singleton de Ω de sorte que la somme donne 1. La formule générale pour P est alors simplement donnée par

$$\forall A \in \mathscr{P}(\Omega), \quad P(A) = \sum_{\omega\in A} P(\{\omega\}).$$

Exemple 2 : Comme on a la relation avec la somme géométrique

$$\sum_{k=0}^{+\infty} \frac{1}{2^{k+1}} = \frac{1}{2}\sum_{k=0}^{+\infty} \frac{1}{2^k} = \frac{1}{2}\frac{1}{1-1/2} = 1,$$

on en déduit d'après la proposition précédente qu'il existe une unique probabilité sur l'univers $\Omega = \mathbb{N}$ vérifiant

$$\forall k \in \Omega, \quad P(\{k\}) = \frac{1}{2^{k+1}}.$$

I.C - Conditionnement par un évènement

Définition (Probabilité conditionnelle) : Soient A et B deux évènements telles que $P(B) > 0$. On appelle probabilité conditionnelle de A sachant B le nombre

$$P(A \mid B) = P_B(A) = \frac{P(A \cap B)}{P(B)}.$$

Exemple 3 : On jette un dé équilibré. On désigne par A l'évènement le chiffre du dé est 6 et par B l'évènement le chiffre du dé est pair. On a

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{1/6}{1/2} = \frac{1}{3}.$$

Le résultat est intuitif : si on sait que le dé affiche une chiffre pair, on a une chance sur trois que le chiffre soit six.

Proposition 3 : Si B est un évènement de probabilité non nulle, alors l'application $P_B : \mathscr{P}(\Omega) \to [0,1]$ est une probabilité.

Remarque 4 : En particulier, on peut utiliser toutes les propriétés de la proposition 1 avec la probabilité P_B. Par exemple, on a $P_B(\bar{A}) = 1 - P_B(A)$.

Formule des probabilités composées : Si $A_1, \dots, A_m$ sont des événements tels que $P(A_1 \cap \dots \cap A_m) \neq 0$, alors

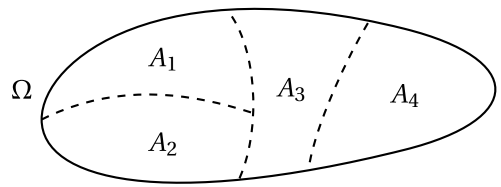
$$P\left(\bigcap_{i=1}^{m} A_i\right) = P(A_1)P(A_2|A_1)P(A_3|A_2 \cap A_1) \cdots P(A_m|A_{m-1} \cap \dots \cap A_1).$$

Définition (Système complet d'évènements) : On dit qu'une famille dénombrable d'événements $(A_i)_{i \in I}$ est un système complet d'événements si

(i) les événements $(A_i)_{i \in I}$ sont deux à deux incompatibles,

(ii) les événements $(A_i)_{i \in I}$ recouvrent Ω, i.e. $\Omega = \bigcup_{i \in I} A_i$.

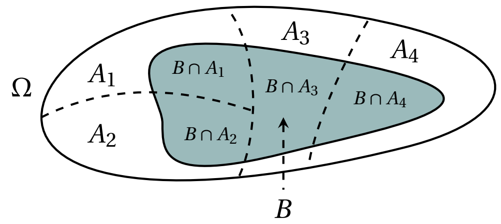
Illustration : On peut représenter la notion de système complet d'évènement par le schéma ci-dessous.



Formule des probabilités totales : Soit B un événement. Si $(A_n)_{n \in \mathbb{N}}$ est un système complet d'évènements de probabilités non nulles, alors la série $\sum P(B \cap A_n)$ converge et

$$P(B) = \sum_{n=0}^{+\infty} P(B \cap A_n) = \sum_{n=0}^{+\infty} P(B \mid A_n) P(A_n).$$

Illustration : En reprenant l'illustration précédente, le schéma ci-dessous permet de comprendre la formule des probabilités totales.



Remarque 5 : La formule ci-dessus reste valable lorsque $(A_n)_{n \in \mathbb{N}}$ n'est pas un système complet d'évènements mais s'ils sont deux à deux incompatibles et vérifient

$$P\left(\bigcup_{n \in \mathbb{N}} A_n\right) = 1.$$

Dans ce cas, on dit que $(A_n)_{n \in \mathbb{N}}$ est un système quasi-complet d'événements.

Formule de Bayes : Si A et B sont deux événements de probabilités non nulles, alors

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}.$$

I.D - Indépendance d'évènements

Définition (Évènements indépendants) : Deux évènements A et B sont dits indépendants si $P(A \cap B) = P(A)P(B)$.

Remarque 6 : Si $P(B) > 0$, l'indépendance de A et B équivaut à $P(A \mid B) = P(A)$. Autrement dit, avoir des informations sur la réalisation de B ne renseigne pas la réalisation de A.

Définition (Évènements mutuellement indépendants) : On dit que des évènements $A_1, \dots, A_n$ sont dits mutuellement indépendants si

$$\forall I \subset [\![1,n]\!], \quad P\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} P(A_i).$$

Remarque 7 : Si n évènements sont mutuellement indépendants, alors ils sont indépendants deux à deux. La réciproque est fausse comme le montre l'exemple suivant.

Exemple 4 : On considère $\Omega = \{1,2,3,4\}$ muni de la probabilité uniforme et on note

$$A = \{1,2\}, \quad B = \{1,3\}, \quad C = \{1,4\}.$$

On vérifie facilement que ces évènements sont indépendants deux à deux, mais ils ne sont pas mutuellement indépendants car

$$P(A \cap B \cap C) = \frac{1}{4} \neq \frac{1}{8} = P(A)P(B)P(C).$$

II - Variables aléatoires discrètes

II.A - Variables aléatoires réelles

Définition (Variable aléatoire réelle) : Une variable aléatoire réelle est une application $X : \Omega \to \mathbb{R}$.

Jusqu'à la fin du chapitre, on fixe une variable aléatoire réelle X sur Ω.

Notation : Si $U \subset \mathbb{R}$, on note l'évènement

$$(X \in U) = X^{-1}(U) = \{\omega \in \Omega \mid X(\omega) \in U\}$$

En particulier, si $x \in \mathbb{R}$, on a

$$(X = x) = \{\omega \in \Omega \mid X(\omega) = x\} \quad \text{et} \quad (X \leqslant x) = \{\omega \in \Omega \mid X(\omega) \leqslant x\}.$$

Définition (Image d'une variable aléatoire par une fonction) : Soit $f : \mathbb{R} \to \mathbb{R}$ une application. On définit la variable aléatoire $f(X)$ comme la composée $f \circ X$, i.e.

$$\forall \omega \in \Omega, \quad f(X)(\omega) = f(X(\omega)).$$

II.B - Loi d'une variable aléatoire réelle

Proposition 4 : L'application $P_X : \mathscr{P}(X(\Omega)) \to [0,1]$, définie par $P_X(U) = P(X \in U)$ est une probabilité sur l'ensemble $X(\Omega)$.

Remarque 8 : L'ensemble $X(\Omega)$ est l'ensemble des valeurs que prend X. Par exemple, si X donne le résultat du lancer d'un dé classique, alors $X(\Omega) = [\![1,6]\!]$.

Définition (Loi d'une variable aléatoire) : L'application P_X est appelée la loi de la variable aléatoire X.

Proposition 5 : La loi P_X est entièrement déterminé par les nombres $P(X = x)$ pour $x \in X(\Omega)$.

Remarque 9 : Pour donner la loi d'une variable aléatoire réelle, on en déduit qu'il suffit de donner les $P(X = x)$ pour tout $x \in X(\Omega)$.

II.C - Fonction de répartition d'une variable aléatoire réelle

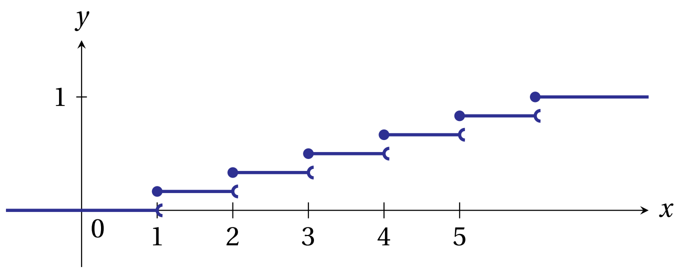
Définition (Fonction de répartition d'une variable aléatoire) : La fonction de répartition de X est l'application $F_X : \mathbb{R} \to \mathbb{R}$ définie par $F_X(x) = P(X \leqslant x)$.

Proposition 6 : On a les propriétés suivantes.

(i) La fonction F_X est croissante et continue à droite en tout point de $\mathbb{R}$.

(ii) La fonction F_X vérifie $\lim\limits_{x \to -\infty} F_X(x) = 0$ et $\lim\limits_{x \to +\infty} F_X(x) = 1$.

Exemple 5 : Si X désigne le résultat du lancer d'un dé classique équilibré, alors le graphe de la fonction de répartition F_X est le suivant.



Proposition 7 : La loi P_X est entièrement déterminée par la fonction F_X.

Remarque 10 : De manière générale, on a pour tout $x \in \mathbb{R}$ la relation

$$P(X = x) = P(X \leqslant x) - P(X < x) = F_X(x) - \lim_{t \to x^-} F_X(t).$$

Dans le cas où X prend ces valeurs dans $\mathbb{N}$, on a simplement

$$\forall n \in \mathbb{N}, \quad P(X = n) = P(X \leqslant n) - P(X \leqslant n - 1) = F_X(n) - F_X(n-1).$$

II.D - Espérance d'une variable aléatoire réelle

En première année, nous avons vu la définition de l'espérance d'une variable aléatoire X lorsque $X(\Omega)$ est fini. Rappelons que dans ce cas, on énumère les éléments de $X(\Omega)$ en écrivant $X(\Omega) = \{x_1, \dots, x_n\}$ et on définit l'espérance de X par

$$E(X) = \sum_{x \in X(\Omega)} x P(X = x) = \sum_{k=1}^{n} x_k P(X = x_k).$$

Dans la suite, on s'intéresse au cas où $X(\Omega)$ est un ensemble infini dénombrable. Il est naturel de chercher à étendre la définition précédente en énumérant les éléments de l'ensemble $X(\Omega)$ en écrivant $X(\Omega) = \{x_k \mid k \in \mathbb{N}\}$, puis en posant

$$E(X) = \sum_{k=0}^{+\infty} x_k P(X = x_k).$$

Cependant, nous sommes confrontés à deux difficultés avec cette approche.

1) La série $\sum\limits_{k \geqslant 0} x_k P(X = x_k)$ est-elle nécessairement convergente ?

2) La somme est-elle indépendante du choix d'énumération de l'ensemble $X(\Omega)$, autrement dit de l'ordre de sommation ?

Définition (Espérance d'une variable aléatoire) : On suppose que $X(\Omega)$ est infini et on fixe une énumération $X(\Omega) = \{x_k \mid k \in \mathbb{N}\}$. On dit que X admet une espérance si la série

$$\sum_{k \geqslant 0} x_k P(X = x_k)$$

converge absolument. Dans ce cas, on appelle espérance de X et on note $\mathrm{E}(X)$ le nombre réel

$$\mathrm{E}(X) = \sum_{k=0}^{+\infty} x_k P(X = x_k).$$

Remarques 11 :

a) Dans la définition, on impose que la série converge absolument, car dans ce cas, on peut vérifier (mais c'est difficile) que la somme de la série ne dépend pas de l'ordre de sommation.

b) L'espérance correspond à la moyenne théorique des valeurs prises par X.

Exemples 6 :

a) Si X est une variable aléatoire dont la loi est donnée par

$$X(\Omega) = \mathbb{N}^* \quad \text{et} \quad \forall k \in \mathbb{N}^*, \quad P(X = k) = \frac{6}{\pi^2 k^2},$$

alors X n'admet pas d'espérance, car la série ci-dessous n'est pas absolument convergente.

$$\sum_{k \geqslant 1} kP(X = k) = \sum_{k \geqslant 1} \frac{6}{\pi^2 k}.$$

b) Si X est une variable aléatoire dont la loi est donnée par

$$X(\Omega) = \mathbb{N} \quad \text{et} \quad \forall k \in \mathbb{N}, \quad P(X = k) = \frac{1}{2^{k+1}},$$

alors X admet une espérance, car la série

$$\sum_{k \geqslant 0} kP(X = k) = \sum_{k \geqslant 0} \frac{k}{2^{k+1}}$$

converge absolument par la règle de d'Alembert. De plus, on peut vérifier que

$$\mathrm{E}(X) = \sum_{k=0}^{+\infty} kP(X = k) = \sum_{k=0}^{+\infty} \frac{k}{2^{k+1}} = 1.$$

Théorème du transfert : On suppose que $X(\Omega)$ est infini et on fixe une énumération $X(\Omega) = \{x_k \mid k \in \mathbb{N}\}$. Si $f : \mathbb{R} \to \mathbb{R}$ est une application, alors la variable aléatoire $f(X)$ admet une espérance si et seulement si la série

$$\sum_{k \geqslant 0} f(x_k) P(X = x_k)$$

converge absolument. Dans ce cas, on a l'égalité

$$\mathrm{E}(f(X)) = \sum_{k=0}^{+\infty} f(x_k) P(X = x_k).$$

Exemple 7 : En reprenant l'exemple précédent, on trouve que X^2 admet une espérance, car la série ci-dessous converge absolument. De plus,

$$\mathrm{E}\left(X^2\right) = \sum_{k=0}^{+\infty} k^2 P(X = k) = \sum_{k=0}^{+\infty} \frac{k^2}{2^{k+1}} = 3.$$

Proposition (Linéarité de l'espérance) : Soient Y une variable aléatoire réelle sur Ω et $(a, b) \in \mathbb{R}^2$. Si X et Y admettent une espérance, alors il en est de même de la variable aléatoire $aX + bY$, et on a

$$\mathrm{E}(aX + bY) = a\mathrm{E}(X) + b\mathrm{E}(Y).$$

Définition (Variable aléatoire centrée) : On dit que X est centrée si X admet une espérance et $\mathrm{E}(X) = 0$.

II.E - Variance et écart type d'une variable aléatoire réelle

Lemme 1 : Si X^2 admet une espérance, alors X admet une espérance.

Définition (Variance d'une variable aléatoire) : Si X^2 admet une espérance, on appelle variance de X le réel

$$\mathrm{V}(X) = \mathrm{E}\left((X - \mathrm{E}(X))^2\right).$$

Définition (Écart type d'une variable aléatoire) : Si X^2 admet une espérance, on appelle écart type de X le réel $\sigma(X) = \sqrt{\mathrm{V}(X)}$.

Remarque 12 : L'écart type est aussi appelée l'écart quadratique moyen à la moyenne. Il permet d'évaluer la dispersion de X autour de sa moyenne $\mathrm{E}(X)$.

Théorème de Koenig-Huyghens : Si X^2 admet une espérance, on a

$$\mathrm{V}(X) = \mathrm{E}\left(X^2\right) - \mathrm{E}(X)^2.$$

Exemple 8 : En reprenant l'exemple précédent, on trouve

$$\mathrm{V}(X) = \mathrm{E}\left(X^2\right) - \mathrm{E}(X)^2 = 3 - 1^2 = 2.$$

Proposition 8 : Soit $(a,b) \in \mathbb{R}^2$. Si X^2 admet une espérance, alors il en est de même de la variable aléatoire $(aX+b)^2$, et on a

$$\mathrm{V}(aX+b) = a^2\,\mathrm{V}(X).$$

ATTENTION : En général si X et Y sont des variables aléatoires réelles, on n'a pas la relation $\mathrm{V}(X+Y) = \mathrm{V}(X) + \mathrm{V}(Y)$.

Inégalité de Bienaymé-Tchebychev : Si X^2 admet une espérance, alors pour tout réel $\varepsilon > 0$, on a l'inégalité

$$P\left(|X - \mathrm{E}(X)| \geqslant \varepsilon\right) \leqslant \frac{\mathrm{V}(X)}{\varepsilon^2}.$$

Définition (Variable aléatoire réduite) : On dit que X est réduite si X^2 admet une espérance et $\mathrm{V}(X) = 1$.

II.F - Lois usuelles

II.F.1 - Loi géométrique

Définition (Loi géométrique) : Soit $p \in]0,1[$. On dit que X suit la loi géométrique de paramètre p, que l'on note $\mathscr{G}(p)$, si

$$X(\Omega) = \mathbb{N}^* \quad \text{et} \quad \forall k \in \mathbb{N}^*, \quad P(X=k) = p(1-p)^{k-1}.$$

Remarque 13 : Si X désigne le rang du premier succès dans une succession indépendante d'épreuve de Bernoulli de paramètre p, alors X suit une loi géométrique de paramètre p.

Proposition 9 : Si X suit la loi $\mathscr{G}(p)$ avec $p \in]0,1[$, alors X^2 admet une espérance et on a

$$\mathrm{E}(X) = \frac{1}{p} \quad \text{et} \quad \mathrm{V}(X) = \frac{1-p}{p^2}.$$

II.F.2 - Loi de Poisson

Définition (Loi de Poisson) : Soit $\lambda \in \mathbb{R}_+^*$. On dit que X suit la loi de Poisson de paramètre λ, que l'on note $\mathscr{P}(\lambda)$, si

$$X(\Omega) = \mathbb{N} \quad \text{et} \quad \forall k \in \mathbb{N}, \quad P(X=k) = \frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}.$$

Exemples 9 :
La loi de Poisson peut être utiliser pour modéliser de nombreuses situations :

a) compter le nombre de communications dans un intervalle de temps donné,

b) modéliser le risque de défaut d'un crédit,

c) etc...

Proposition 10 : Si X suit la loi $\mathscr{P}(\lambda)$ avec $\lambda \in \mathbb{R}_+^*$, alors X^2 admet une espérance et on a

$$\mathrm{E}(X) = \lambda \quad \text{et} \quad \mathrm{V}(X) = \lambda.$$

Théorème d'approximation de la loi binomiale par la loi de Poisson : On considère une suite $(X_n)_{n\in\mathbb{N}}$ de variables aléatoires telle que X_n suit une loi binomiale de paramètre (n, p_n) pour tout $n \in \mathbb{N}$. Si $\lim_{n\to+\infty} np_n = \lambda$, alors

$$\forall k \in \mathbb{N}, \quad \lim_{n\to+\infty} P(X_n = k) = \frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}.$$

Remarque 14 : Intuitivement, le résultat précédent signifie que X_n suit approximativement une loi de Poison de paramètre λ si n est suffisamment grand.

II.F.3 - Résumé des lois usuelles

Nom	Paramètres	Notation	Valeurs	Loi	Espérance	Variance
Uniforme	$n \in \mathbb{N}^*$	$\mathscr{U}(\llbracket 1, n \rrbracket)$	$\llbracket 1, n \rrbracket$	$P(X=k) = \dfrac{1}{n}$	$\dfrac{n+1}{2}$	$\dfrac{n^2-1}{12}$
Bernoulli	$p \in [0,1]$	$\mathscr{B}(p)$	$\{0,1\}$	$P(X=0) = 1-p$, $P(X=1) = p$	p	$p(1-p)$
Binomiale	$n \in \mathbb{N}^*$, $p \in [0,1]$	$\mathscr{B}(n,p)$	$\llbracket 0, n \rrbracket$	$P(X=k) = \dbinom{n}{k} p^k (1-p)^{n-k}$	np	$np(1-p)$
Géométrique	$p \in]0,1[$	$\mathscr{G}(p)$	$\mathbb{N}^*$	$P(X=k) = p(1-p)^{k-1}$	$\dfrac{1}{p}$	$\dfrac{1-p}{p^2}$
Poisson	$\lambda \in \mathbb{R}_+^*$	$\mathscr{P}(\lambda)$	$\mathbb{N}$	$P(X=k) = \dfrac{\lambda^k}{k!} \mathrm{e}^{-\lambda}$	λ	λ