Mathématiques 1

TSI **2016**

4 heures Calculatrices autorisées

Nombres de Stirling et problème du collectionneur

Ce problème est consacré à l'étude de nombre introduits au XVIII^e siècle par James Stirling et intervenant en particulier en théorie des probabilités et dans l'étude de la fiabilité de certains circuits électroniques.

Les parties du problème sont très largement indépendantes. Tout candidat peut admettre un résultat précédemment donné dans le texte pour aborder les questions suivantes, à condition de l'indiquer clairement sur sa copie.

On rappelle les conventions $0! = 1$ et $\forall a \in \mathbb{R}, a^0 = 1$ et on note

- $C^1(\mathbb{R},\mathbb{R})$ l'ensemble des fonctions de $\mathbb{R}$ dans $\mathbb{R}$, dérivables sur $\mathbb{R}$ et dont la dérivée est continue sur $\mathbb{R}$;
- $\mathcal{M}_n(\mathbb{R})$ l'ensemble des matrices carrées d'ordre n à coefficients réels.

I Généralités sur les nombres de Stirling

Si E est un ensemble et k un entier naturel non nul, une partition de E est un ensemble $\{A_1, \dots, A_k\}$ de parties de E non vides, deux à deux disjointes, et dont la réunion est égale à E :

$$\forall i \in \{1,\dots,k\},\ A_i \neq \emptyset \qquad \forall (i,j) \in \{1,\dots,k\}^2,\ i \neq j \implies A_i \cap A_j = \emptyset \qquad \bigcup_{i=1}^{k} A_i = E$$

Par exemple, $\{\{1\},\{2,3\},\{4,5\}\}$ est une partition de $E = \{1,2,3,4,5\}$ en trois parties. On notera en particulier que l'ordre dans lequel interviennent les parties $A_1, \dots, A_k$ n'a pas d'incidence sur la définition de la partition.

Pour $n \in \mathbb{N}^*$ et $k \in \{1,\dots,n\}$, on note $S_{n,k}$ le nombre de partitions d'un ensemble à n éléments en k parties. On pose également $S_{0,0} = 1$ et $S_{n,0} = 0$ lorsque $n \in \mathbb{N}^*$.

Les nombres $S_{n,k}$ sont appelés nombres de Stirling de deuxième espèce.

I.A – Premières propriétés des nombres de Stirling

I.A.1)

a) Déterminer la valeur de $S_{3,2}$.

b) Soit $n \in \mathbb{N}^*$. Que valent $S_{n,1}$ et $S_{n,n}$?

I.A.2) Soit n un entier, $n \geqslant 2$, k un entier compris entre 1 et $n-1$ et $E = \{x_1, \dots, x_n\}$ un ensemble à n éléments. On souhaite établir une relation entre $S_{n,k}$, $S_{n-1,k-1}$ et $S_{n-1,k}$.

a) Dans cette question, on étudie l'exemple $E = \{1,2,3,4\}$ $(n = 4)$ et $k = 2$.

i. Expliciter les partitions de E en deux parties, dont l'une est le singleton $\{4\}$.
ii. Expliciter les partitions de E en deux parties, dont l'une contient 4 tout en étant différente du singleton $\{4\}$.
iii. Vérifier, pour l'exemple traité, la relation $S_{n,k} = S_{n-1,k-1} + kS_{n-1,k}$.

b) On revient au cas général présenté en début de question I.A.2.

i. Quel est le nombre de partitions de E en k parties, dont l'une est $\{x_n\}$? On exprimera le résultat à l'aide d'un nombre de Stirling.
ii. Quel est le nombre de partitions de E en k parties, dont l'une contient x_n tout en étant différente du singleton $\{x_n\}$? On exprimera le résultat à l'aide d'un nombre de Stirling.
iii. En déduire que $S_{n,k} = S_{n-1,k-1} + kS_{n-1,k}$.

I.A.3) En déduire que, pour tout entier $n \geqslant 2$, $S_{n,n-1} = \dfrac{n(n-1)}{2}$. On pourra par exemple procéder par récurrence.

I.A.4) Montrer que, pour tout entier $n \geqslant 3$, $S_{n,2} + 1 = 2(S_{n-1,2} + 1)$ et en déduire la valeur de $S_{n,2}$ pour tout entier $n \geqslant 2$.

I.A.5) Soient $n \in \mathbb{N}^*$, $k \in \mathbb{N}^*$ et deux ensembles $E = \{x_1, \dots, x_n\}$ et $F = \{y_1, \dots, y_k\}$. On note $\sigma_{n,k}$ le nombre d'applications surjectives de E dans F.

a) Que vaut $\sigma_{n,k}$ si $n < k$?

b) Que vaut $\sigma_{n,k}$ si $n = k$?

c) Expliciter les applications surjectives de E dans F lorsque $n = 3$ et $k = 2$, et préciser la valeur de $\sigma_{3,2}$.

d) Dans cette question, on souhaite obtenir une relation entre $\sigma_{n,k}$ et $S_{n,k}$.

Si f est une application de E dans F et j un entier compris entre 1 et k, on note $A_j = f^{-1}(\{y_j\})$ l'ensemble des antécédents par f de y_j.

i. Étant donnée une application f surjective de E dans F, montrer que $\mathcal{A} = \{A_1, ..., A_k\}$ est une partition de E en k parties. On note $\Pi(f)$ cette partition.

ii. Étant donnée $\mathcal{A}' = \{A'_1, ..., A_k{}'\}$ une partition de E en k parties, combien y a-t-il d'applications f surjectives de E dans F telles que $\Pi(f) = \mathcal{A}'$?

iii. En déduire la relation $\sigma_{n,k} = k!\, S_{n,k}$.

I.A.6) Montrer, par exemple par récurrence, que, pour tout $n \in \mathbb{N}$, $\forall k \in \{0, ..., n\},\ S_{n,k} \leqslant (2k)^n$.

I.B – Utilisation des nombres de Stirling en algèbre linéaire

Soit $N \in \mathbb{N}^*$; on note $\mathbb{R}_N[X]$ l'ensemble des polynômes à coefficients réels de degré inférieur ou égal à N et $\mathcal{B}$ la base canonique de $\mathbb{R}_N[X]$. On définit la famille de polynômes $(P_k)_{0 \leqslant k \leqslant N}$ par

$$\begin{cases} P_0 = 1 \\ P_k = \displaystyle\prod_{j=0}^{k-1} (X - j) \quad \text{pour tout } k \in \{1, ..., N\} \end{cases}$$

I.B.1) Montrer que la famille $\{P_0, ..., P_N\}$ est une base de $\mathbb{R}_N[X]$, que l'on notera $\mathcal{B}_0$.

I.B.2)

a) Pour tout entier $k \in \{0, ..., N-1\}$, démontrer l'égalité $XP_k = P_{k+1} + kP_k$.

b) En déduire par récurrence que, pour tout entier $n \in \{0, ..., N\}$, $X^n = \displaystyle\sum_{k=0}^{n} S_{n,k} P_k$. On pourra utiliser la relation démontrée en I.A.2.

I.B.3) Donner la matrice M de passage de la base $\mathcal{B}_0$ à la base $\mathcal{B}$, en précisant en particulier ses coefficients diagonaux.

I.B.4)

a) Quel est le polynôme caractéristique de la matrice M ?

b) Montrer que la matrice M n'est pas diagonalisable dans $\mathcal{M}_{N+1}(\mathbb{R})$ lorsque $N \geqslant 2$. Qu'en est-il si $N = 1$?

I.C – Un lien entre nombres de Stirling et loi de Poisson

Pour $n \in \mathbb{N}$ et pour les réels x pour lesquels cette somme a un sens, on pose $f_n(x) = \displaystyle\sum_{i=0}^{+\infty} \frac{i^n}{i!} x^i$.

I.C.1) Déterminer le rayon de convergence de la série entière $\displaystyle\sum_{i=0}^{+\infty} \frac{i^n}{i!} x^i$ et en déduire le domaine de définition de la fonction f_n.

I.C.2) Déterminer les valeurs de $f_0(x)$, $f_1(x)$ et $f_2(x)$.

I.C.3) Soit $k \in \{0, ..., n\}$ et $i \in \mathbb{N}$. Préciser la valeur de $P_k(i)$ en distinguant les cas $0 \leqslant i \leqslant k-1$ (lorsque k est non nul) et $i \geqslant k$.

I.C.4) À l'aide de la relation montrée en I.B.2, vérifier que pour tout couple (n, i) d'entiers naturels

$$i^n = \sum_{k=0}^{n} S_{n,k} P_k(i)$$

I.C.5) Montrer que $f_n(x) = \displaystyle\sum_{k=0}^{n} S_{n,k} \sum_{i=k}^{+\infty} \frac{x^i}{(i-k)!}$ et en déduire l'égalité $f_n(x) = e^x \displaystyle\sum_{k=0}^{n} S_{n,k} x^k$.

I.C.6) Soit Y une variable aléatoire suivant une loi de Poisson de paramètre 1.

Montrer que pour tout entier $n \in \mathbb{N}^*$, la variable aléatoire Y^n admet une espérance $E(Y^n)$ et donner une expression de $E(Y^n)$ en fonction des nombres de Stirling. Confronter les valeurs trouvées pour $E(Y)$ et $E(Y^2)$ avec les résultats du cours sur l'espérance et la variance d'une variable aléatoire suivant une loi de Poisson de paramètre λ.

I.D – Comportement asymptotique des nombres de Stirling

I.D.1) Pour $k \in \mathbb{N}$, on considère l'équation différentielle

$$y'(x) = (k+1)y(x) + \frac{1}{k!}(e^x - 1)^k \tag{I.1}$$

d'inconnue $y \in C^1(\mathbb{R}, \mathbb{R})$.

a) Montrer que la fonction $x \mapsto \dfrac{1}{(k+1)!}(e^x-1)^{k+1}$ est solution de (I.1).

b) Résoudre l'équation différentielle (I.1).

I.D.2) On se propose de démontrer par récurrence sur $k \in \mathbb{N}$ la proposition

$$\forall x \in \mathbb{R},\ \sum_{n=k}^{+\infty} \frac{S_{n,k}}{n!} x^n = \frac{1}{k!}(e^x-1)^k \tag{I.2}$$

a) Soit $k \in \mathbb{N}$ fixé. En utilisant par exemple l'inégalité établie en I.A.6, montrer que le rayon de convergence de la série entière $\displaystyle\sum_{n=k}^{+\infty} \frac{S_{n,k}}{n!} x^n$ est infini.

b) Montrer que la proposition (I.2) est vraie pour $k=0$.

c) Soit k un entier naturel fixé tel que la proposition (I.2) soit vraie. Montrer alors, en utilisant le résultat de la question I.A.2, que la fonction $x \mapsto \displaystyle\sum_{n=k+1}^{+\infty} \frac{S_{n,k+1}}{n!} x^n$ est solution de l'équation différentielle (I.1).

d) Conclure.

I.D.3) En déduire, pour tous $n \in \mathbb{N}$ et $k \in \{0, ..., n\}$, l'égalité $S_{n,k} = \dfrac{1}{k!} \displaystyle\sum_{j=0}^{k} (-1)^{k-j} \binom{k}{j} j^n$. On pourra commencer par développer $(e^x-1)^k$ à l'aide de la formule du binôme de Newton.

I.D.4) On fixe un entier $k \in \mathbb{N}^*$. Déterminer $\displaystyle\lim_{n\to+\infty} \frac{S_{n,k}}{k^n}$ et en déduire un équivalent de $S_{n,k}$ lorsque $n \to +\infty$.

II Nombres de Stirling et problème du collectionneur

Le problème du collectionneur (« *coupon collector's problem* ») est un problème de probabilités classique. Un fabricant de tablettes de chocolat propose à ses acheteurs de collectionner des images. Chaque tablette vendue contient une image de la collection, que l'on découvre à l'ouverture de la tablette. Chaque image peut être collée dans un album contenant k emplacements (k est un entier supérieur ou égal à deux) correspondant au nombre d'images distinctes de la collection. La probabilité pour un acheteur de découvrir dans une tablette une image donnée est égale à $1/k$.

On se propose de déterminer le nombre moyen d'achats nécessaires pour constituer la collection complète des k images et d'étudier comment les nombres de Stirling interviennent dans le problème du collectionneur.

Pour i entier compris entre 1 et k, on note X_i le nombre minimum d'achats nécessaires pour obtenir i vignettes différentes et Z_i le nombre minimum d'achats nécessaires pour qu'un collectionneur possédant déjà i vignettes distinctes puisse enrichir sa collection d'une vignette autre que celles qu'il possède déjà.

II.A – Équivalent de $E(X_k)$ lorsque $k \to +\infty$

II.A.1) Préciser la loi de la variable aléatoire X_1 et donner $E(X_1)$, son espérance.

II.A.2) Pour i compris entre 1 et $k-1$, vérifier que la variable aléatoire Z_i suit la loi géométrique de paramètre $\dfrac{k-i}{k}$ et donner $E(Z_i)$, son espérance.

II.A.3) Pour i compris entre 1 et $k-1$, comparer $X_{i+1}-X_i$ et Z_i.

II.A.4) En remarquant que $X_k = X_1 + \displaystyle\sum_{i=1}^{k-1}(X_{i+1}-X_i)$, donner une expression de $E(X_k)$, l'espérance de X_k, sous la forme d'une somme.

II.A.5)

a) Pour $n \in \mathbb{N}^*$, on pose $u_n = \left(\displaystyle\sum_{i=1}^{n} \frac{1}{i}\right) - \ln(n)$. Montrer que la série de terme général $u_n - u_{n-1}$ ($n \geqslant 2$) est convergente. Qu'en déduit-on pour la suite $(u_n)_{n\in\mathbb{N}^*}$?

b) En déduire un équivalent de $E(X_k)$ lorsque k tend vers $+\infty$ et interpréter ce résultat.

II.B – Équivalent de $P(X_k = n)$ lorsque $n \to +\infty$

Dans cette sous-partie, on fixe l'entier k.

II.B.1) Quelle est la valeur de $P(X_k = n)$ pour n compris entre 1 et $k-1$?

II.B.2) Montrer que si n est un entier supérieur ou égal à k, alors $P(X_k = n) = \dfrac{k(k-1)!\,S_{n-1,k-1}}{k^n}$.

II.B.3) En déduire un équivalent et la limite de $P(X_k = n)$ lorsque $n \to +\infty$, et interpréter ce résultat.

• • • FIN • • •
