Université de Nice Sophia–Antipolis

Licence MASS 2000-2001

(Re-)Mise à niveau en Probabilités

Feuilles de 1 à 7

Ces feuilles avec 25 exercices et quelques rappels historiques furent distribuées à des étudiants de troisième année, dans le cadre d'un cours intensif sur deux semaines, en début d'année, pour une brève « remise à niveau » en théorie des probabilités.

Université de Nice Sophia–Antipolis

Licence MASS 2000-2001

(Re-)Mise à niveau en Probabilités

Feuille numéro 1

1a. On répartit 40 étudiants sur deux salles d'examen, la salle A et la salle B, avec 20 étudiants dans chacune des deux salles. De combien de façons distinctes cela peut-il se faire ?

1b. On divise un groupe de 40 étudiants en deux groupes de chacun 20 étudiants pour les TDs. De combien de façons distinctes cela peut-il se faire ?

2. On jette deux dés. Quelle est la probabilité que la somme des points soit paire ?

3. Quelle est la probabilité d'avoir (au moins) une paire dans une main de poker ? (main de 5 cartes, jeu de 52 cartes).

4. On divise un groupe de 30 étudiants en 3 groupes de 10 étudiants. Quelle est la probabilité que Frénégonde, Hubert et Archibald se retrouvent dans le même groupe ? Quelle est la probabilité qu'ils se retrouvent totalement séparés ?

Université de Nice Sophia–Antipolis

Licence MASS 2000-2001

(Re-)Mise à niveau en Probabilités

Feuille numéro 2

1. Un problème d'examen consiste à déterminer la probabilité d'avoir l'As de Coeur dans une main de poker. L'étudiant répond $1/52 + 1/51 + 1/50 + 1/49 + 1/48$ avec l'argument suivant : « *lors du tirage de la main, il y a une chance sur cinquante-deux que l'As de Coeur sorte le premier, puis une chance sur cinquante-et-une en deuxième car il reste alors cinquante-et-une cartes, puis etc. . .* ». Cet argument est incorrect, où est l'erreur ? Montrer que le résultat correct est $5/52$.

2. On a vu dans l'exercice précédent que la probabilité de la présence d'une carte donnée (par exemple l'As de Coeur) est $5/52$. Il y a 52 cartes, chacune ayant la probabilité $5/52$ d'apparaître. Cela fait donc une probabilité de $52 \times 5/52 = 5$ d'avoir dans sa main une carte quelconque non spécifiée. Donc $5 = 1$. Qu'en pensez-vous ?

3. Soit A l'évènement : « ma main de poker comporte un As de Coeur », et soit B l'évènement : « ma main de poker comporte un As de Pique ». Quelle est la probabilité conditionnelle de B sachant A ? Les évènements A et B sont-ils mutuellement indépendants ?

4. Montrer la relation générale $P(A \cdot B) = 1 - P(\overline{A}) - P(\overline{B}) + P(\overline{A} \cdot \overline{B})$.

5. Soit A l'évènement : « ma main de poker comporte (au moins) un Coeur », et soit B l'évènement : « ma main de poker comporte (au moins) un As ». Calculer les probabilités de A, de B et de $A \cdot B$. Les évènements A et B sont-ils mutuellement indépendants ? (pour cet exercice, il est conseillé de raisonner avec les évènements complémentaires ; calculer les probabilités de $\overline{A}$, $\overline{B}$ et $\overline{A} \cdot \overline{B}$).

Université de Nice Sophia–Antipolis

Licence MASS 2000-2001

(Re-)Mise à niveau en Probabilités

Feuille numéro 3

1. Soient A et B deux évènements. On suppose $0 < P(A) < 1$. Montrer que A et B sont indépendants si et seulement si $P(B|\ A) = P(B|\ \overline{A})$.

2. On tire une carte d'un jeu de 52. Soit A l'évènement « la carte est Coeur ou Carreau », soit B l'évènement « la carte est Coeur ou Pique », et soit C l'évènement « la carte est Coeur ou Trèfle ». Montrer que A, B, C sont deux-à-deux indépendants mais ne sont pas indépendants dans leur ensemble.

3. On considère une population pour laquelle il y a une chance sur vingt pour un homme d'être daltonien, et une chance sur quatre-cents pour une femme d'être daltonienne. Par ailleurs il y a 49% d'hommes et 51% de femmes. Quelle est la proportion de Daltoniens dans cette population ? Quelle est la probabilité conditionnelle d'être un homme sachant que l'on est daltonien ?

4. Une certaine maladie est caractérisée partiellement par certains symptômes : on estime que 60% des personnes présentant ces symptômes sont atteints. On applique sur cette population présentant ces symptômes un test sanguin qui donne un résultat positif sur 70% des personnes malades, et aussi sur 10% des personnes non-malades. Quelle est la probabilité conditionnelle d'être malade si le test est positif ? On applique un deuxième test indépendant du premier qui donne un résultat positif sur 80% des malades et 15% des non-malades. Quelle est la probabilité conditionnelle d'être malade si les deux tests sont positifs ? Quelle est la probabilité conditionnelle de ne pas être malade si les deux tests sont positifs ? Quelle est la probabilité conditionnelle pour le deuxième test d'être positif si le premier l'est ? Quelle est la probabilité absolue que les tests donnent des résultats contradictoires ?

(L'énoncé est imprécis (incomplet) sur un point. Quelle précision faut-il apporter ?)

Université de Nice Sophia–Antipolis

Licence MASS 2000-2001

(Re-)Mise à niveau en Probabilités

Feuille numéro 4

1. Soient A, B, C trois évènements. Montrer :

$$\mathbf{1}_{A\cup B\cup C} = 1 - (1-\mathbf{1}_A)(1-\mathbf{1}_B)(1-\mathbf{1}_C)$$

En déduire la formule (dite d'« inclusion-exclusion ») :

$$\begin{aligned} P(A\cup B\cup C) &= P(A)+P(B)+P(C) \\ &- P(A\cdot B) - P(A\cdot C) - P(B\cdot C) \\ &+ P(A\cdot B\cdot C) \end{aligned}$$

2. Une regrettable erreur informatique au services des Impôts résulte en un mauvais étiquetage des envois de 5 avis d'imposition à des contribuables : les 5 adresses ont été aléatoirement échangées lors de l'impression sur les enveloppes. Soit X la variable aléatoire égale au nombre de lettres sur les 5 qui ont été correctement expédiées. Quelle est l'espérance de X ?

3. Cinq personnes se retrouvent à une soirée. Soit X le nombre total de jours de la semaine qui sont des jours de naissance pour une au moins des cinq personnes. Quelle est l'espérance de X ?

4. Un débutant au basket a une chance sur trois de marquer un panier de la ligne des coup-francs. Soit X la variable aléatoire égale au nombre de tentatives nécessaires pour marquer son premier panier. Déterminer la loi de X et son espérance.

Licence MASS 2000-2001

(Re-)Mise à niveau en Probabilités

Feuille numéro 5

1. Soit X une variable aléatoire. Notons $a_0 = \mathbf{E}X$. Soit $v(a)$ la fonction $a \mapsto \mathbf{E}((X-a)^2)$. Montrer que la fonction $v(a)$ atteint en a_0 son minimum global, et que pour tout $a \neq a_0$ on a une inégalité stricte $v(a) > v(a_0)$. Vérifier que $v(a_0)$ que l'on appelle « variance de X » s'exprime identiquement comme $\mathbf{V}X = \mathbf{E}(X^2) - (\mathbf{E}X)^2$ ou comme $\mathbf{V}X = \mathbf{E}((X - \mathbf{E}X)^2)$. Montrer que X est (presque sûrement) constante si et seulement si $\mathbf{V}X = 0$.

2. Soient X et Y deux variables aléatoires. On suppose que $\mathbf{V}Y > 0$. Soit $v(a,b)$ la fonction $(a,b) \mapsto \mathbf{E}((X - a - bY)^2)$. Montrer que le minimum de $v(a,b)$ est obtenu en un unique couple (a_0, b_0), qui satisfait le système d'équations linéaires :

$$\begin{aligned} a_0 + \quad \mathbf{E}Y \cdot b_0 &= \mathbf{E}X \\ \mathbf{E}Y \cdot a_0 + \mathbf{E}(Y^2) \cdot b_0 &= \mathbf{E}(YX) \end{aligned}$$

Soit $\mathbf{Cov}(X,Y) = \mathbf{E}(XY) - \mathbf{E}X \cdot \mathbf{E}Y$. Montrer :

$$b_0 = \frac{\mathbf{Cov}(X,Y)}{\mathbf{V}Y}$$

3. Montrer :

$$\mathbf{Cov}(X,Y) = \mathbf{E}((X - \mathbf{E}X) \cdot (Y - \mathbf{E}Y)) \qquad \mathbf{Cov}(X,X) = \mathbf{V}X$$

$$\mathbf{V}(\alpha X + \beta Y) = \alpha^2 \cdot \mathbf{V}X + 2\alpha\beta \cdot \mathbf{Cov}(X,Y) + \beta^2 \cdot \mathbf{V}Y$$

4. Montrer $|\mathbf{Cov}(X,Y)| \leq \sqrt{\mathbf{V}X \cdot \mathbf{V}Y}$ avec égalité si et seulement si X et Y sont liés par une relation affine.

Université de Nice Sophia–Antipolis

Licence MASS 2000-2001

(Re-)Mise à niveau en Probabilités

Feuille numéro 6

Cette session est consacrée à la **loi des grands nombres** de JACOB BERNOULLI (1654-1705), suivant la démonstration de TCHEBYCHEV (1821-1894).

1. Montrer que deux variables aléatoires indépendantes ne sont pas corrélées.

2. Montrer que la variance d'une somme de variables aléatoires deux-à-deux non corrélées est égale à la somme des variances de ces variables aléatoires.

3. Soit X une variable aléatoire d'espérance α et d'écart-type σ. Montrer pour tout $\delta > 0$:

$$P(|X - \alpha| > \delta) \leq \frac{\sigma^2}{\delta^2}$$

4. On procède à un nombre infini de répétitions indépendantes d'une même épreuve qui a la probabilité p de succès. Au k^{e} essai on pose $X_k = 1$ en cas de succès et $X_k = 0$ en cas d'échec. Soit S_n la somme de X_1, ..., X_n, qui compte donc le nombre des succès sur les n premiers essais. Quelle est la variance de S_n ? Quelle est la variance de S_n/n (qui compte la fréquence empirique de succès) ? Quelle est l'espérance de S_n/n ? Montrer :

$$P(|S_n/n - p| > \delta) \leq \frac{1}{n} \cdot \frac{p(1-p)}{\delta^2}$$

En déduire la **loi des grands nombres** : quel que soit le choix de $\delta > 0$ on a $\lim_{n\to\infty} P(|S_n/n - p| > \delta) = 0$. Un résultat plus fort a été démontré par BOREL (1871-1956) en 1909 : « lors d'une infinité de répétitions indépendantes d'une épreuve de probabilité p de succès on a avec probabilité 1 la convergence de la fréquence empirique de succès vers p : $\lim \frac{S_n}{n} = p$ ».

Licence MASS 2000-2001

(Re-)Mise à niveau en Probabilités

Feuille numéro 7

La loi des grands nombres de BERNOULLI pour la répétition d'épreuves indépendantes admet un raffinement, le **Théorème de la Limite Centrale** (TLC) démontré par DE MOIVRE (1667-1754) en 1733. Il utilisa pour cela une formule ($n! \sim \sqrt{2\pi n}(n/e)^n$) attribuée depuis par erreur uniquement à STIRLING (1692-1770). DE MOIVRE n'avait considéré que le cas $p = 0.5$ c'est apparemment LAPLACE (1749-1827) dans un livre paru en 1812 qui publia le premier le cas général sans restriction sur p. L'énoncé est le suivant : soit, comme dans la **feuille 6**, S_n le nombre total de succès lors de n épreuves indépendantes ayant chacune probabilité p de succès. Soient $A \leq B$ deux nombres quelconques, alors :

$$\lim_{n\to\infty} P(A \cdot \sqrt{p(1-p)} \leq \frac{S_n - np}{\sqrt{n}} \leq B \cdot \sqrt{p(1-p)}) = \frac{1}{\sqrt{2\pi}} \int_A^B e^{-x^2/2}\,dx$$

On sait pour $A = -\infty$ et $B = +\infty$ que l'intégrale vaut $\sqrt{2\pi}$. Il n'y a pas d'expression générale en fonction de A et B, mais des tables de valeurs ont été calculées.

BERNOULLI avait aussi considéré des sommes $S_n = X_1 + \ldots + X_n$ avec des variables X_k non-restreintes à ne prendre que les valeurs 0 ou 1, mais qui restaient des copies indépendantes d'une même variable X d'espérance α. Il avait montré dans plusieurs cas que $\delta > 0 \Rightarrow \lim_{n\to\infty} P(|\frac{S_n}{n} - \alpha| > \delta) = 0$.

Après de nombreux travaux aboutissant dans les années 1920s aux résultats de LINDEBERG (1876-1932), et LÉVY (1886-1971), on a maintenant une version générale du **Théorème de la Limite Centrale**, qui affine la **Loi des grands nombres**. En notant α l'espérance et σ l'écart-type de X :

$$\lim_{n\to\infty} P(A \cdot \sigma \leq \frac{S_n - n\alpha}{\sqrt{n}} \leq B \cdot \sigma) = \frac{1}{\sqrt{2\pi}} \int_A^B e^{-x^2/2}\,dx$$

Le TLC nous informe que en cas de somme de n quantités indépendantes, les fluctuations autour de la moyenne sont de l'ordre de $\sqrt{n}$ fois l'écart-type σ. Lorsque σ est infini, le TLC ne s'applique pas, mais tant que X a une espérance finie $\alpha = \mathbf{E}X$ la loi de Bernoulli s'applique.