UNIVERSITÉ DE TOURS **UFR SCIENCES ET TECHNIQUES**

MASTER DE MATHEMATIQUES

FORMULAIRE du cours PROBABILITÉS 1 (2011-2012)

I) Généralités du calcul des probabilités : Soit Ω un ensemble (univers des possibles)

<u>Espace probabilisé</u> $(\Omega, \mathcal{F}, \mathbb{P})$: $\mathcal{F}$ est une tribu sur Ω c'est à dire $\mathcal{F} \subset \mathcal{P}(\Omega)$ vérifie

1)$\Omega \in \mathcal{F}$, 2) $(A_n)_{n\in N^*} \in \mathcal{F} \Rightarrow \bigcup_{n=1}^{+\infty} A_n \in \mathcal{F}$, 3) $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$,

et $\mathbb{P} : \mathcal{F} \to [0,1]$ (la probabilité) est telle que :

1)$\mathbb{P}(\Omega) = 1$, 2)$\mathbb{P}\left(\bigcup_{n=0}^{\infty} A_n\right) = \sum_{n=0}^{+\infty} \mathbb{P}(A_n)$ pour tous $(A_n) \in \mathcal{F}$, tels que $A_i \cap A_j = \emptyset$ si $i \neq j$.

<u>Limite sup et Limite inf</u> : Si $(A_n)_{n\in N^*} \in \mathcal{F}$, on note

1) $\liminf A_n = \bigcup_{n=1}^{+\infty}\left(\bigcap_{m=n}^{+\infty} A_m\right)$, 2) $\limsup A_n = \bigcap_{n=1}^{+\infty}\left(\bigcup_{m=n}^{+\infty} A_m\right)$. On a alors :

1) $\liminf A_n \subset \limsup A_n$, 2) $(\liminf A_n)^c = \limsup A_n^c$, 3) $(\limsup A_n)^c = \liminf A_n^c$.

<u>Continuité de $\mathbb{P}$</u> : Pour toute suite <u>monotone</u> $(A_n) \in \mathcal{F}$, on a $\mathbb{P}(\lim A_n) = \lim_{n\to+\infty} \mathbb{P}(A_n)$ (où $\lim A_n = \bigcup_n A_n$ si $A_n \nearrow$ et $\lim A_n = \bigcap_n A_n$ si $A_n \searrow$).

<u>Probabilité conditionnelle</u> : Si $B \in \mathcal{F}$ (avec $\mathbb{P}(B) > 0$) est un événement fixé, pour tout $A \in \mathcal{F}$, la quantité $P_B(A) = \mathbb{P}(A|B) = \mathbb{P}(A \cap B)/\mathbb{P}(B)$ est la probabilité conditionnelle de A sachant B. L'application $\mathbb{P}_B : \mathcal{F} \to [0,1]$ ainsi définie est la probabilité conditionnelle sachant B.

<u>Formule de l'intersection</u> : Si $A_1, A_2, \ldots, A_k \in \mathcal{F}$ et $\mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_{k-1}) > 0$,

$$\mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_k) = \mathbb{P}(A_1)\mathbb{P}(A_2|A_1)\mathbb{P}(A_3|A_1 \cap A_2)\ldots\mathbb{P}(A_k|A_1 \cap \ldots \cap A_{k-1}).$$

<u>Formule de la probabilité totale</u> : Soit (A_n) un système complet d'événements. Alors :
$\forall A \in \mathcal{F}$, $\mathbb{P}(A) = \sum_n \mathbb{P}(A|A_n)\mathbb{P}(A_n)$.

<u>Formule de Bayes</u> : $\mathbb{P}(A_n|A) = \dfrac{\mathbb{P}(A|A_n)\mathbb{P}(A_n)}{\sum_k \mathbb{P}(A|A_k)\mathbb{P}(A_k)}$ $(A \in \mathcal{F})$ si (A_n) est un système complet.

<u>Événements indépendants</u> : Ce sont des événements $A_1, \ldots, A_n$ tels que $\mathbb{P}(A_1 \cap \ldots \cap A_n) = \mathbb{P}(A_1)\ldots\mathbb{P}(A_n)$.

<u>Espace probabilisé produit</u> : Si $(\Omega_i, \mathcal{F}_i, \mathbb{P}_i)$ $(1 \leq i \leq n)$ sont des espaces probabilisés, leur espace produit est $(\Omega, \mathcal{F}, \mathbb{P})$ où $\Omega = \Omega_1 \times \cdots \times \Omega_n$, $\mathcal{F} = \mathcal{F}_1 \otimes \cdots \otimes \mathcal{F}_n$ et $\mathbb{P} = \mathbb{P}_1 \otimes \cdots \otimes \mathbb{P}_n$ où $\mathcal{F}$ est engendrée par les rectangles $A_1 \times \cdots \times A_n$, $A_i \in \mathcal{F}_i$, et $\mathbb{P}$ est telle que $\mathbb{P}(A_1 \times \cdots \times A_n) = \prod_{i=1}^n \mathbb{P}_i(A_i)$.

II) Variables aléatoires discrètes : Ce sont les applications $X : \Omega \to \mathbb{R}$ telles que $X(\Omega) = \{x_k; k \in D\}$ est fini où dénombrable ($D = \{1, \dots, N\}$ ou $D = \mathbb{N}$) et vérifiant la condition de mesurabilité : $\forall k \in D$, $[X = x_k] \in \mathcal{F}$.

Distribution de probabilité : C'est la suite des nombres $p_k = \mathbb{P}(X = x_k)$ $(k \in D)$. Ce sont des nombres $0 \leq p_k \leq 1$ tels que $\sum_{k \in D} p_k = 1$.

Moments : La v.a. X a un moment d'ordre n ($n \in \mathbb{N}^*$) si $\sum_{k \in D} p_k |x_k|^n < +\infty$ et le moment d'ordre n est $\mathbb{E}(X^n) = \sum_{k \in D} p_k x_k^n$ (c'est l'espérance si $n = 1$). Si $\mathbb{E}(X^2)$ existe, la variance est $Var(X) = \mathbb{E}((X - \mathbb{E}(X))^2) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$ et $\sigma_X = \sqrt{Var(X)}$ est l'écart type.

Fonction génératrice : Si X est à valeurs dans $\mathbb{N}$, c'est la série entière $G_X(t) = \sum_{k=0}^{\infty} p_k t^k$. Si X et Y sont indépendantes, $G_{X+Y}(t) = G_X(t) G_Y(t)$. Si X a un moment d'ordre 2, G_X est deux fois dérivable, $\mathbb{E}(X) = G'_X(1)$ et $Var(X) = G''_X(1) + G'_X(1) - (G'_X(1))^2$.

v.a. binomiale $\mathcal{B}(n,p)$: Toute v.a. X telle que $X(\Omega) = \{0, 1, \dots, n\}$ et pour tout $0 \leq k \leq n$ $\mathbb{P}(X = k) = C_n^k p^k (1-p)^{n-k}$, (où $n \geq 1$ et $p \in [0,1]$ sont fixés). Si $n = 1$, on dit que X est une v.a. de Bernoulli. On a $\mathbb{E}(X) = np$ et $Var(X) = np(1-p)$.

v.a. de Poisson de paramètre $\lambda > 0$: Toute v.a. X telle que $X(\Omega) = \mathbb{N}$ et $\mathbb{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$ $(k \in \mathbb{N})$. On a $\mathbb{E}(X) = Var(X) = \lambda$.

v.a. de Pascal de paramètre $p \in]0,1[$: Toute v.a. telle que $X(\Omega) = \mathbb{N}^*$ et $\mathbb{P}(X = k) = p(1-p)^{k-1}$ (k entier ≥ 1) (on l'appelle aussi v.a. géométrique, ou v.a. instant du premier succès).

III) Variables aléatoires quelconques, lois de probabilité : Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé.

Vecteur aléatoire (v.a.) (ou variable aléatoire si $d = 1$) : C'est une application mesurable $X : (\Omega, \mathcal{F}) \longrightarrow (\mathbb{R}^d, \mathcal{B}_d)$ où $\mathcal{B}_d$ est la tribu borélienne de $\mathbb{R}^d$, c'est à dire :
$\forall B \in \mathcal{B}_d, [X \in B] = \{\omega \in \Omega \ ; \ X(\omega) \in B\} \in \mathcal{F}$ (condition de mesurabilité).

Condition suffisante de mesurabilité : Soit $\mathcal{C}$ une partie génératrice de $\mathcal{B}_d$. Si $X : \Omega \to \mathbb{R}^d$ est telle que $\forall C \in \mathcal{C}, \ [X \in C] \in \mathcal{F}$, alors X est un v.a.

Transformée déterministe d'un(e) v.a. : Si $X : (\Omega, \mathcal{F}) \to (\mathbb{R}^p, \mathcal{B}_p)$ est un v.a. de $\mathbb{R}^p$ et $f : (\mathbb{R}^p, \mathcal{B}_p) \to (\mathbb{R}^q, \mathcal{B}_q)$ est mesurable, alors $f \circ X = f(X)$ est un vecteur aléatoire de $\mathbb{R}^q$.

Limite d'une suite de v.a. : Si (X_n) est une suite de v.a. telle que $\forall \omega \in \Omega, \lim\limits_{n \to \infty} X_n(\omega) = X(\omega) \in \mathbb{R}$ existe, alors X est une v.a..

v.a. indicatrice et v.a. simple : Soit $A \in \mathcal{F}$. La v.a. indicatrice de A est donnée par $\mathbf{1}_A(\omega) = 0$ si $\omega \notin A$ et $= 1$ si $\omega \in A$. Toute v.a. $X = \sum\limits_{finie} \alpha_i \mathbf{1}_{A_i}$ ($\alpha_i \in \mathbb{R}$ et $A_i \in \mathcal{F}$) est dite v.a. simple.

Théorème de structure : Toute v.a. $X : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}_1)$ (resp. toute v.a.≥ 0) est limite simple d'une suite de v.a. simples (resp. de v.a. simples≥ 0).

Loi de probabilité d'un v.a. X de $\mathbb{R}^d$: C'est la mesure de probabilité μ_X sur $\mathcal{B}_d$ définie par :
$\forall B \in \mathcal{B}_d \ , \ \mu_X(B) = \mathbb{P}(X \in B)$.

Fonction de répartition d'une v.a. X : C'est la fonction $F : \mathbb{R} \to \mathbb{R}$ telle que : $\forall t \in \mathbb{R}, F(t) = \mathbb{P}(X \le t)$. Elle est telle que : 1) $0 \le F(t) \le 1$; 2) F est croissante (au sens large), continue à droite en chaque $t \in \mathbb{R}$; 3) $\lim\limits_{t \to -\infty} F(t) = 0$ et $\lim\limits_{t \to +\infty} F(t) = 1$.

Densité d'un v.a. : Le v.a. X de $\mathbb{R}^d$ a une densité f si $f : \mathbb{R}^d \to \mathbb{R}_+$ est intégrable avec $\int_{\mathbb{R}^d} f(x_1, \ldots, x_d) dx_1 \ldots dx_d = 1$ et si $\forall B \in \mathcal{B}_d$, $\mu_X(B) = \int_B f(x_1, \ldots, x_d) dx_1 \ldots dx_d$.

Densité normale $N(0,1)$: C'est la fonction $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ $(x \in \mathbb{R})$.

Densité normale $N(m, \sigma^2)$: C'est la fonction $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}(\frac{x-m}{\sigma})^2\right)$ $(x \in \mathbb{R})$.

Densité uniforme sur $[a,b]$: C'est la fonction $f(x) = \frac{1}{b-a} \mathbf{1}_{[a,b]}(x)$ $(x \in \mathbb{R})$.

Densité exponentielle de paramètre $\lambda > 0$: C'est la fonction $f(x) = \lambda e^{-\lambda x}\ \mathbf{1}_{\mathbb{R}_+}(x)$.

Densités marginales : Si (X, Y) est un couple de v.a. ayant une densité $f(x, y)$ sur $\mathbb{R}^2$, les densités marginales sont données par $f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy$ et $f_Y(y) = \int_{-\infty}^{\infty} f(x, y) dx$.

v.a. indépendantes : Les v.a. $X_1, \ldots, X_k$ sont dites (mutuellement) indépendantes si
$\forall B_1, \ldots, B_k \in \mathcal{B}_1$, $\mathbb{P}(X_1 \in B_1, \ldots, X_k \in B_k) = \mathbb{P}(X_1 \in B_1) \ldots \mathbb{P}(X_k \in B_k)$.

CNS d'indépendance : Des v.a. X et Y de lois respectives μ_X et μ_Y, sont indépendantes si et seulement si $\mu_{(X,Y)} = \mu_X \otimes \mu_Y$ (où $\mu_{(X,Y)}$ est la loi dans $\mathbb{R}^2$ du vecteur (X, Y)). Si (X, Y) a une densité $f(x, y)$ ceci équivaut à : $f(x, y) = f_X(x) f_Y(y)$ (produit des densités marginales).

IV) Espérance, variance et moments des v.a. : Soit $(\Omega, \mathcal{F}, P)$ un espace probabilisé.

Espace $\mathcal{L}^1$: Une v.a. X est dans $\mathcal{L}^1$ si $\int_\Omega |X| d\mathbb{P} < +\infty$. On pose alors $\mathbb{E}(X) = \int_\Omega X d\mathbb{P}$ et on l'appelle l'espérance mathématique de X.

Formule de l'espérance totale : Soit X un v.a. de $\mathcal{L}^1$ et (A_n) un système complet d'événements. Alors :
$\mathbb{E}(X) = \sum\limits_n \mathbb{E}(X|A_n)\mathbb{P}(A_n)$,
où $\mathbb{E}(X|A_n) = \int_\Omega X d\mathbb{P}_{A_n}$ et $\mathbb{P}_{A_n}$ est la probabilité conditionnelle sachant A_n.

Théorème (ou Formule) du transfert : Soit X un v.a. de $\mathbb{R}^d$ et μ_X sa loi de probabilité sur $\mathcal{B}_d$. Soit $\varphi : \mathbb{R}^d \to \mathbb{R}$ borélienne. Alors $\varphi \circ X \in \mathcal{L}^1(\Omega, \mathcal{F}, P) \Leftrightarrow \varphi \in \mathcal{L}^1(\mathbb{R}^d, \mathcal{B}_d, \mu_X)$ et on a $\mathbb{E}(\varphi \circ X) = \int_\Omega \varphi \circ X d\mathbb{P} = \int_{\mathbb{R}^d} \varphi(x) d\mu_X(x)$. Ainsi si X est réelle et dans $\mathcal{L}^1$ on a : $\mathbb{E}(X) = \int_\mathbb{R} x d\mu_X(x)$.

Espérance d'un produit : Si $X_1, \ldots, X_n$ sont dans $\mathcal{L}^1$ et sont indépendantes : $\mathbb{E}(\prod_{i=1}^n X_i) = \prod_{i=1}^n \mathbb{E}(X_i)$.

Espace $\mathcal{L}^2$ et variance : Une v.a. X est dans $\mathcal{L}^2$ si $\mathbb{E}(X^2) < +\infty$. On a $\mathcal{L}^2 \subset \mathcal{L}^1$. $\mathbb{E}(X^2)$ est le moment d'ordre 2 de X. La variance de X est le nombre $VarX = \mathbb{E}\left((X - \mathbb{E}(X))^2\right) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$ et l'écart type est $\sigma_X = \sqrt{VarX}$.

Espace de Hilbert L^2 : Si dans $\mathcal{L}^2$, on considère comme égales deux v.a qui sont égales $\mathbb{P}$-presque sûrement, on obtient l'espace L^2. Avec la norme $||X||_2 = \sqrt{\mathbb{E}(X^2)}$, l'espace $(L^2, ||.||_2)$ est un espace de Hilbert.

Covariance : Pour X et Y dans $\mathcal{L}^2$ c'est le nombre $cov(X,Y) = \mathbb{E}\big((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))\big)$. Si X et Y sont indépendantes alors $cov(X,Y) = 0$ (la réciproque est fausse).

Variance d'une somme : Soient $X_1, \ldots, X_n$ des v.a. de $\mathcal{L}^2$ deux à deux de covariance nulle. Alors $Var(X_1 + \cdots + X_n) = VarX_1 + \cdots + VarX_n$.

Inégalité de Markov : Si $X \in \mathcal{L}^1$, alors $\forall a > 0$, on a $\mathbb{P}(|X| \geq a) \leq \frac{\mathbb{E}(|X|)}{a}$.

Inégalité de Bienaymé-Tchebychev : Si $X \in \mathcal{L}^2, \forall a > 0, \mathbb{P}\Big(|\frac{X-\mathbb{E}(X)}{\sigma_X}| \geq a\Big) \leq \frac{1}{a^2}$.

Espace $\mathcal{L}^p$: Soit $p \geq 1$ un entier. Une v.a. X est dans $\mathcal{L}^p$ si $\mathbb{E}(|X|^p) < +\infty$. La quantité $\mathbb{E}(X^p)$ est le moment d'ordre p de X. Si X est dans $\mathcal{L}^p$ alors X est dans $\mathcal{L}^{p'}$ pour tout $p' \leq p$.

V) Convergence des v.a., loi forte des grands nombres et théorème limite central

Convergence p.s. et en probabilité : La suite de v.a. (X_k) converge p.s. (resp. en probabilité) vers la v.a. X si $\mathbb{P}([\lim_{k\to\infty} X_k = X]) = 1$ (resp. $\forall \delta > 0$, $\lim\limits_{n\to+\infty} P(|X_n - X| \geq \delta) = 0$). La convergence p.s. implique la convergence en probabilité.

CS de convergence p.s. : Si $\forall \varepsilon > 0$, $\sum_{k=1}^{\infty} P(|X_k| > \varepsilon) < +\infty$ alors $X_n \to 0$ p.s. En particulier si $(X_n) \in \mathcal{L}^p$ et si $\sum_{n=1}^{+\infty} E(|X_n|^p) < +\infty$ alors $X_n \to 0$ p.s.

Lemme de Borel Cantelli : Si $(A_n) \in \mathcal{F}$ et $\displaystyle\sum_{n=1}^{\infty} \mathbb{P}(A_n) < +\infty$. Alors $\mathbb{P}(\limsup A_n) = 0$.

Loi forte des grands nombres : Si (X_k) est une suite de v.a. de $\mathcal{L}^2$, de même espérance m, de variances bornées et 2 à 2 de covariance nulle alors $\frac{1}{n}\sum_{k=1}^{n} X_k \to m$ p.s. $(n \to +\infty)$.

Loi forte des grands nombres de Kolmogorov : Si (X_k) est une suite de v.a. de L^1 indépendantes et de même loi alors $\lim_{n\to+\infty} \frac{1}{n}\sum_{k=1}^{n} X_k = m$ p.s. $(m = \mathbb{E}(X_1))$.

La convergence en loi : La suite de v.a. (X_n) (de loi μ_n) converge en loi vers la v.a. X (de loi μ) si $\forall f \in C_b(\mathbb{R})$, $\lim_{n\to+\infty} < \mu_n, f > = < \mu, f >$. Si F_n (resp. F) est la fonction de répartition de X_n (resp. X) ceci équivaut à : $\lim_{n\to+\infty} F_n(t) = F(t)$ en tout point de continuité t de F.

Théorème limite central : Si $X_n (n \geq 1)$ sont des v.a. de $\mathcal{L}^2$, indépendantes et de même loi d'espérance m et de variance σ^2 , alors $\lim_{n\to+\infty} \frac{1}{\sigma\sqrt{n}}\big((\sum_{i=1}^{n} X_i) - nm\big) = N(0,1)$ en loi .