

programme des deux dernières interrogations avant les écrits.

Première semaine.

Probabilité : tout le programme précédent + les variables aléatoires discrètes i.e.

- a) *Espaces probabilisés*

Tribu sur un ensemble Ω . On se borne à la définition et à la stabilité par les opérations ensemblistes finies ou dénombrables. Événements. Généralisation du vocabulaire relatif aux événements introduit en première année.

Espace probabilisable (Ω, \mathcal{A}) .

Si \mathcal{A} est une tribu sur Ω , une probabilité sur (Ω, \mathcal{A}) est une application P définie sur \mathcal{A} , à valeurs dans $[0, 1]$, telle que $P(\Omega) = 1$ et, pour toute suite $(A_n)_{n \geq 0}$ d'événements deux à deux disjoints, on ait :

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

Espace probabilisé (Ω, \mathcal{A}, P) .

Si Ω est fini ou dénombrable et si $\mathcal{A} = \mathcal{P}(\Omega)$, une probabilité P sur (Ω, \mathcal{A}) s'identifie, via la formule

$$P(\{\omega\}) = p_\omega,$$

à une famille $(p_\omega)_{\omega \in \Omega}$ de réels positifs sommable de somme 1.

- b) *Propriétés élémentaires des probabilités*

Continuité croissante : si $(A_n)_{n \geq 0}$ est une suite d'événements croissante pour l'inclusion, alors :

$$P(A_n) \xrightarrow{n \rightarrow +\infty} P\left(\bigcup_{n=0}^{+\infty} A_n\right).$$

Continuité décroissante : si $(A_n)_{n \geq 0}$ est une suite d'événements décroissante pour l'inclusion, alors :

$$P(A_n) \xrightarrow{n \rightarrow +\infty} P\left(\bigcap_{n=0}^{+\infty} A_n\right).$$

Si $(A_n)_{n \geq 0}$ est une suite d'événements, alors :

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

Événements négligeables, événements presque sûrs. Une réunion finie ou dénombrable d'événements négligeables est négligeable. Propriétés presque sûres.

Tout développement sur ces notions est hors programme.

- c) *Probabilités conditionnelles et indépendance*

Extension des résultats vus en première année dans le cadre des univers finis : probabilité conditionnelle, formule des probabilités composées, formule des probabilités totales, formules de Bayes. Notations $P_B(A)$, $P(A|B)$.

Couple d'événements indépendants. Famille quelconque d'événements mutuellement indépendants.

- d) *Variables aléatoires discrètes*

Étant donné un ensemble E et un espace probabilisé (Ω, \mathcal{A}, P) , une variable aléatoire discrète définie sur Ω est une application X de Ω dans E telle que $X(\Omega)$ soit fini ou dénombrable et que, pour tout x de $X(\Omega)$, $X^{-1}(\{x\}) \in \mathcal{A}$. Lorsque $E \subset \mathbb{R}$, la variable aléatoire est dite réelle.

Loi P_X de la variable aléatoire X . Notations $X \sim Y$, $X \sim \mathcal{L}$.

Notations $(X \geq x)$, $(X \leq x)$, $(X < x)$, $(X > x)$ pour une variable aléatoire réelle X .

- e) *Couples de variables aléatoires, variables aléatoires indépendantes*

Couple de variables aléatoires. Loi conjointe, lois marginales.

Loi conditionnelle de Y sachant $X = x$. Extension au conditionnement par $X > x$ ou autres inégalités.

Extension aux n -uplets de variables aléatoires. Vecteurs aléatoires discrets.

Couple de variables aléatoires indépendantes. Extension des résultats vus en première année.

Famille de variables aléatoires mutuellement indépendantes. Si X_1, X_2, \dots, X_n sont des variables aléatoires mutuellement indépendantes, alors pour tout m compris entre 1 et $n-1$, et toutes fonctions f et g , les variables $f(X_1, \dots, X_m)$ et $g(X_{m+1}, \dots, X_n)$ sont indépendantes. Démonstration non exigible.

Existence d'espaces probabilisés portant une suite de variables indépendantes de lois discrètes données. La démonstration est hors programme.

Modélisation du jeu de pile ou face infini : suite d'épreuves de Bernoulli mutuellement indépendantes.

- f) *Lois usuelles*

Pour p dans $]0, 1[$, loi géométrique de paramètre p . Notation $\mathcal{G}(p)$.

La variable aléatoire X suit la loi $\mathcal{G}(p)$ si

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

Interprétation comme rang du premier succès dans une suite d'épreuves de Bernoulli mutuellement indépendantes de paramètre p .

Caractérisation comme loi sans mémoire :

$$P(X > n+k \mid X > n) = P(X > k).$$

Pour λ dans \mathbb{R}_+^* , loi de Poisson de paramètre λ . Notation $\mathcal{P}(\lambda)$.

Approximation de la loi binomiale par la loi de Poisson : si, pour tout n , $X_n \sim \mathcal{B}(n, p_n)$ et si (np_n) converge vers λ , alors :

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

Interprétation de la loi de Poisson comme loi des événements rares.

- g) *Espérance*

Si X est une variable aléatoire à valeurs dans \mathbb{R}^+ , l'espérance de X est la somme, dans $[0, +\infty[$, de la famille $(P(X = x) \cdot x)_{x \in X(\Omega)}$. Notation $E(X)$.

Si X est une variable aléatoire réelle, la variable aléatoire X est dite d'espérance finie si la famille $(x P(X = x))_{x \in X(\Omega)}$ est sommable ; dans ce cas, la somme de cette famille est l'espérance de X . Notation $E(X)$.

Variables centrées.

Espérance d'une variable aléatoire suivant une loi géométrique, d'une variable aléatoire suivant une loi de Poisson.

Linéarité, positivité et croissance de l'espérance sur l'espace des variables aléatoires d'espérance finie définies sur Ω . Si $|X| \leq Y$ et si Y est d'espérance finie, alors X est d'espérance finie.

Formule de transfert : soit X une variable aléatoire discrète, f une fonction définie sur $X(\Omega)$ à valeurs dans \mathbb{R} ; alors $f(X)$ est d'espérance finie si et seulement si la famille $(P(X = x) \cdot f(x))$ est sommable ; si tel est le cas :

$$E(f(X)) = \sum_{x \in X(\Omega)} f(x) P(X = x).$$

Démonstration non exigible.

Inégalité de Markov.

Si X et Y sont deux variables aléatoires indépendantes d'espérances finies, alors :

$$E(XY) = E(X) E(Y).$$

- h) *Variance, écart type et covariance*

Moments.

Si une variable aléatoire admet un moment d'ordre 2, elle est d'espérance finie.

Inégalité de Cauchy-Schwarz : si X et Y admettent chacune un moment d'ordre 2, alors XY est d'espérance finie et $E(XY)^2 \leq E(X^2) E(Y^2)$.

Espace des variables aléatoires définies sur Ω admettant un moment d'ordre 2.

Variance, écart type. Notations $V(X), \sigma(X)$.

Variables réduites.

Relation $V(X) = E(X^2) - E(X)^2$.

Relation $V(aX + b) = a^2 V(X)$. Si $\sigma(X) > 0$, la variable aléatoire $\frac{X - E(X)}{\sigma(X)}$ est centrée réduite.

Variance d'une variable aléatoire géométrique, d'une variable aléatoire de Poisson.

Inégalité de Bienaymé-Tchebychev.

Covariance de deux variables aléatoires.

Relation $\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$. Cas de variables indépendantes.

Variance d'une somme finie, cas de variables deux à deux indépendantes.

- i) *Loi faible des grands nombres*

Si $(X_n)_{n \geq 0}$ est une suite de variables aléatoires deux à deux indépendantes, de même loi et admettant un moment d'ordre

2, alors, si $S_n = \sum_{k=1}^n X_k$ et $m = E(X_1)$, on a,

$$P\left(\left|\frac{S_n}{n} - m\right| \geq \varepsilon\right) \xrightarrow{n \rightarrow +\infty} 0.$$

Les étudiants doivent savoir retrouver, pour $\varepsilon > 0$, l'inégalité :

$$P\left(\left|\frac{S_n}{n} - m\right| \geq \varepsilon\right) \leq \frac{\sigma^2}{n\varepsilon^2}$$

où σ est la variance commune des X_k .

- j) *Fonctions génératrices*

Fonction génératrice de la variable aléatoire X à valeurs dans \mathbb{N} :

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

La série entière définissant G_X est de rayon supérieur ou égal à 1 et converge normalement sur le disque fermé de centre 0 et de rayon 1. Continuité de G_X .

Détermination de la loi de X par G_X . Utilisation de G_X pour calculer les moments de X .

La variable aléatoire X est d'espérance finie si et seulement si G_X est dérivable en 1 ; dans ce cas $E(X) = G_X'(1)$. La variable aléatoire X admet un second moment si et seulement si G_X est deux fois dérivable en 1. Les étudiants doivent savoir retrouver l'expression de la variance de X à l'aide de $G_X'(1)$ et $G_X''(1)$.

Les étudiants doivent savoir calculer la fonction génératrice d'une variable aléatoire de Bernoulli, binomiale, géométrique, de Poisson.

Fonction génératrice d'une somme finie de variables aléatoires indépendantes à valeurs dans \mathbb{N} .

Deuxième semaine.

Calcul différentiel et optimisation

Les fonctions considérées dans cette section sont définies sur un ouvert d'un \mathbb{R} -espace vectoriel normé E de dimension finie et à valeurs dans un \mathbb{R} -espace vectoriel normé F de dimension finie.

Le choix d'une base de l'espace d'arrivée permet de se ramener au cas des fonctions à valeurs réelles.

- a) *Dérivée selon un vecteur, dérivées partielles*

Dérivée de l'application f au point a selon le vecteur v . Notations $D_v f(a)$, $D_v f$.

Dérivées partielles dans une base. Notations $\frac{\partial f}{\partial x_i}(a)$, $\partial_i f(a)$.

Lorsqu'une base de E est fixée, identification entre $f(x)$ et $f(x_1, \dots, x_n)$.

- b) *Différentielle*

Application différentiable au point a . Notation $o(h)$. Développement limité à l'ordre 1.

Lorsque $f = (f_1, \dots, f_p)$, f est différentiable en a si et seulement si toutes les f_i le sont.

Si f est différentiable en a , alors f est continue en a et dérivable en a selon tout vecteur.

Différentielle de f en a , encore appelée application linéaire tangente à f en a . Unicité de la différentielle et relation $f'(a) \cdot v = D_v f(a)$. Notation $df(a)$.

Application différentiable sur un ouvert Ω . Différentielle sur Ω . Notation df .

Cas particuliers : application constante, application linéaire.

Lien entre différentielle et dérivées partielles. Si Ω est un ouvert de \mathbb{R}^n et si f est à valeurs dans \mathbb{R}^m , la matrice jacobienne de f en a est la matrice de $df(a)$ dans les bases canoniques.

Cas des fonctions d'une variable : si Ω est un intervalle ouvert de \mathbb{R} , la différentiabilité de f en a équivaut à la dérivabilité de f en a ; relation $f'(a) = df(a) \cdot 1$.

Si l'espace E est euclidien, gradient en a d'une application numérique différentiable en a . Expression du gradient en base orthonormée. Notation $\nabla f(a)$.

Interprétation géométrique : si $\nabla f(a) \neq 0$, $\nabla f(a)$ est positivement colinéaire au vecteur unitaire selon lequel la dérivée de f en a est maximale.

- c) *Opérations sur les applications différentiables*

Différentielle d'une combinaison linéaire d'applications différentiables, de $M(f_1, \dots, f_p)$ où M est multilinéaire et où f_1, \dots, f_p sont des applications différentiables.

Règle de la chaîne : différentielle d'une composée d'applications différentiables.

Dérivée le long d'un arc : si γ est une application définie sur l'intervalle I de \mathbb{R} , dérivable en t , si f est différentiable en $\gamma(t)$, alors $(f \circ \gamma)'(t) = df(\gamma(t)) \cdot \gamma'(t)$. Interprétation géométrique en termes de tangentes.

Cas particulier fondamental : $\gamma(t) = x + tv$.

Dérivation de $t \mapsto f(x_1(t), \dots, x_n(t))$.

Dérivées partielles d'une composée d'applications différentiables. Dérivées partielles de

$$(u_1, \dots, u_m) \mapsto f(x_1(u_1, \dots, u_m), \dots, x_n(u_1, \dots, u_m)).$$

- d) Applications de classe \mathcal{C}^1

Une application f est dite de classe \mathcal{C}^1 sur un ouvert Ω si elle est différentiable sur Ω et si df est continue sur Ω .

L'application f est de classe \mathcal{C}^1 sur Ω si et seulement si les dérivées partielles relativement à une base de E existent en tout point de Ω et sont continues sur Ω . La démonstration n'est pas exigible.

Opérations algébriques sur les applications de classe \mathcal{C}^1 .

Si f est une application de classe \mathcal{C}^1 de Ω dans F , si γ est une application de classe \mathcal{C}^1 de $[0, 1]$ dans Ω , si $\gamma(0) = a$ et $\gamma(1) = b$, alors :

$$f(b) - f(a) = \int_0^1 df(\gamma(t)) \cdot \gamma'(t) dt.$$

Cas particulier $\gamma(t) = a + tv$ pour tout $t \in [0, 1]$.

Si Ω est connexe par arcs, caractérisation des fonctions constantes sur Ω . Démonstration pour Ω convexe.

- e) Vecteurs tangents à une partie d'un espace normé de dimension finie

Si X est une partie de E et x un point de X , un vecteur v de E est tangent à X en x s'il existe $\varepsilon > 0$ et un arc γ défini sur $] - \varepsilon, \varepsilon[$, à valeurs dans X , dérivable en 0, tel que $\gamma(0) = x, \gamma'(0) = v$. Notation $T_x X$ pour l'ensemble des vecteurs tangents à X en x .

Exemples : sous-espace affine, sphère d'un espace euclidien, graphe d'une fonction numérique définie sur un ouvert de \mathbb{R}^2 .

Si g est une fonction numérique définie et de classe \mathcal{C}^1 sur l'ouvert Ω de E , si $x \in X$ et $dg(x) \neq 0$, alors $T_x X$ est égal au noyau de $dg(x)$. La démonstration de cet énoncé et le théorème des fonctions implicites sont hors programme.

Traduction en termes de gradient si E est euclidien, en particulier pour $E = \mathbb{R}^n$ muni de sa structure euclidienne canonique.

Exemple : plan tangent à une surface de \mathbb{R}^3 définie par une équation.

- f) Optimisation : étude au premier ordre

Point critique d'une application différentiable.

Condition nécessaire d'existence d'un extremum local en un point intérieur. Exemples de recherches d'extrema globaux.

Si f est une fonction numérique définie sur l'ouvert Ω , si X est une partie de Ω , si la restriction de f à X admet un extremum local en x et si f est différentiable en x , alors $df(x)$ s'annule en tout vecteur tangent à X en x .

Théorème d'optimisation sous une contrainte : si f et g sont des fonctions numériques définies et de classe \mathcal{C}^1 sur l'ouvert Ω de E , si X est l'ensemble des zéros de g , si $x \in X$ et $dg(x) \neq 0$ et si la restriction de f à X admet un extremum local en x , alors $df(x)$ est colinéaire à $dg(x)$. Si E est euclidien, traduction en termes de gradient.

Exemples de recherches d'extrema sous contrainte.

- g) Applications de classe \mathcal{C}^k

Dérivées partielles d'ordre k d'une fonction définie sur un ouvert de \mathbb{R}^n . Notations $\frac{\partial^k f}{\partial x_{j_k} \dots \partial x_{j_1}}, \partial_{j_k} \dots \partial_{j_1} f, \partial_{j_1, \dots, j_k} f$.

Une application est dite de classe \mathcal{C}^k sur un ouvert Ω de \mathbb{R}^n si ses dérivées partielles d'ordre k existent et sont continues sur Ω . La notion de différentielle seconde est hors programme.

Théorème de Schwarz. La démonstration n'est pas exigible.

Opérations algébriques sur les applications de classe \mathcal{C}^k . Composition d'applications de classe \mathcal{C}^k . Les démonstrations ne sont pas exigibles.

Exemples simples d'équations aux dérivées partielles du premier et du second ordre.

- h) Optimisation : étude au second ordre

Matrice hessienne en un point d'une fonction de classe \mathcal{C}^2 sur un ouvert de \mathbb{R}^n , à valeurs réelles. Notation $H_f(x)$.

Formule de Taylor-Young à l'ordre 2 :

$$f(x+h) \underset{h \rightarrow 0}{=} f(x) + \langle \nabla f(x), h \rangle + \frac{1}{2} \langle H_f(x) \cdot h, h \rangle + o(\|h\|^2),$$

$$f(x+h) \underset{h \rightarrow 0}{=} f(x) + \nabla f(x) \bar{h} + \frac{1}{2} h^\top H_f(x) h + o(\|h\|^2).$$

La démonstration n'est pas exigible.

Si f est une fonction de classe \mathcal{C}^2 sur un ouvert de \mathbb{R}^n et si f admet un minimum local en x , alors x est point critique de f et $H_f(x) \in \mathcal{S}_n^+(\mathbb{R})$.

Adaptation au cas d'un maximum local.

Si f est une fonction de classe \mathcal{C}^2 sur un ouvert de \mathbb{R}^n , si x est point critique de f et si $H_f(x) \in \mathcal{S}_n^{++}(\mathbb{R})$, alors f atteint un minimum local strict en x .

Adaptation au cas d'un maximum local.

Explicitation pour $n = 2$ (trace et déterminant).

