Programme de l'épreuve de probabilité et statistique Recrutement international des écoles d'IP-Paris dans les cycles ingénieurs

Les notions du programme de l'épreuve de mathématique (algèbre, analyse réelle) sont supposées être connues. La théorie de la mesure et les théories générales de l'intégration ne sont pas au programme (en particulier toutes les questions relatives à la mesurabilité des fonctions).

1 Probabilité

1.1 Événements Aléatoires

1. Univers Ω, ensemble $\mathcal{E}$ des évènements comme ensemble non vide de parties de Ω stable par passage au complémentaire et union dénombrable
2. Opérations ensemblistes sur les événements
3. Probabilité $\mathbb{P}$: fonction de $\mathcal{E}$ dans $[0,1]$ tel que $\mathbb{P}(\Omega)=1$, $\mathbb{P}(\emptyset)=0$, $\mathbb{P}(A)\leq\mathbb{P}(B)$ si $A\subset B$, $\mathbb{P}(A\cup B)=\mathbb{P}(A)+\mathbb{P}(B)$ si $A\cap B=\emptyset$ et $\mathbb{P}(\cup_i A_i)=\sum_i \mathbb{P}(A_i)$ si A_i est une suite d'événements deux à deux disjoints.
4. Probabilité conditionnelle à un événement B de probabilité non nulle: $\mathbb{P}(A|B)=\frac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)}$
5. $\mathbb{P}(\cap_{i=1}^n A_i)=\mathbb{P}(A_1)\prod_{k=2}^n \mathbb{P}(A_k|A_1\cap...\cap A_{k-1})$
6. Probabilité totale : $\mathbb{P}(A)=\sum_i \mathbb{P}(A|B_i)\mathbb{P}(B_i)$ si $\mathbb{P}(\cup_i B_i)=1$ et $\mathbb{P}(B_i\cap B_j)=0$ pour $i\neq j$
7. Bayes : $\mathbb{P}(A|B)=\frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B)}$ et $\mathbb{P}(A_k|B)=\frac{\mathbb{P}(B|A_k)\mathbb{P}(A_k)}{\sum_i \mathbb{P}(B|A_i)\mathbb{P}(A_i)}$ si $\mathbb{P}(\cup_i A_i)=1$ et $\mathbb{P}(A_i\cap A_j)=0$ pour $i\neq j$
8. Événements indépendants deux à deux et dans leur ensemble
9. Lemme de Borel-Cantelli : si $(A_n)_{n\geq 0}$ est une suite d'éléments de $\mathcal{E}$ telle que $\sum_{n\geq 0}\mathbb{P}(A_n)<\infty$ alors $\mathbb{P}(\limsup_n A_n)=0$.

1.2 Variables aléatoires à valeurs réelles, vecteurs aléatoires

1. Définition : Fonction X de Ω dans $\mathbb{R}$ telle que $\{\omega\in\Omega : X(\omega)\leq a\}\in\mathcal{E}$ pour tout a.

2. Notation $\mathbb{P}(X \in A)$, $\mathbb{P}(X > b)$, ... pour $\mathbb{P}(\{\omega \in \Omega : X(\omega) \in A\})$, $\mathbb{P}(\{\omega \in \Omega : X(\omega) > b\})$, ...

3. Fonction de répartition : $F_X(a) = \mathbb{P}(X^{-1}(]-\infty; a])) = \mathbb{P}(X \leq a)$

4. Support $\mathcal{S}(X)$: plus petit ensemble fermé F de $\mathbb{R}$ tel que $\mathbb{P}(X \in F) = 1$.

5. Variable aléatoire discrète : variable dont le support est dénombrable (ou fini)

6. Variable aléatoire à densité : variable X dont la fonction de répartition admet la représentation intégrale $F_X(a) = \int_{-\infty}^{a} f_X(x)dx$, f_X est la densité de X

7. Vecteur de variables aléatoires $X = (X_i)_{i=1,...,n}$, fonction de répartition $F_X(x_1, ..., x_n) = \mathbb{P}(\cap_{i=1}^{n}\{X_i \leq x_i\})$, support $\mathcal{S}(X)$ (plus petit fermé F de R^n tel que $\mathbb{P}(X \in F) = 1$), cas discret ($\mathcal{S}(X)$ fini ou dénombrable), cas "à densité" si F_X admet une représentation intégrale $F_X(x_1, ..., x_n) = \int_{-\infty}^{x_1} ... \int_{-\infty}^{x_n} f_X(x_1, ..., x_n)dx_1...dx_n$

8. Variable aléatoire positive : variable aléatoire X telle que $F_X(a) = 0$ pour tout $a < 0$

9. Définition de l'espérance de $g(X)$ pour g une fonction positive sur $\mathcal{S}(X)$ pour X une variable ou un vecteur aléatoire: $\mathbb{E}(g(X)) = \sum_{x \in \mathcal{S}(X)} g(x)\mathbb{P}(X = x)$ si X est discrète, $\mathbb{E}(g(X)) = \int_{\mathcal{S}(X)} g(x)f_X(x)dx$ si X est continue. Pour les fonctions g de signe non constants sur $\mathcal{S}(X)$, l'espérance de $g(X)$ est définie si $\mathbb{E}(|g(X)|) < \infty$ par $E(g(X)) = \sum_{x \in \mathcal{S}(X)} g(x)\mathbb{P}(X = x)$ si X est discrète, $\mathbb{E}(g(X)) = \int_{\mathcal{S}(X)} g(x)f_X(x)dx$ si X admet une densité.

10. L'espérance d'un vecteur aléatoire est le vecteur des espérances des composantes du vecteur.

11. Variance, covariance des variables aléatoires : $\mathbb{V}(X) = \mathbb{E}\left((X - \mathbb{E}(X))^2\right) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$ et $\mathbb{C}ov(X, Y) = \mathbb{E}\left((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))\right) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$. Ecart-type d'une variable aléatoire : $\sqrt{\mathbb{V}(X)}$.

12. Matrice de variance-covariance d'un vecteur de variables aléatoires $(X_1, ..., X_n)$: matrice carrée de taille $n \times n$ de terme générique $\mathbb{C}ov(X_i, X_j)$

13. Inégalité de Markov : $\mathbb{P}(|X| \geq a) \leq \frac{\mathbb{E}(|X|)}{a}$

14. Inégalité de Bienaymé-Tchebychev : $\mathbb{P}(|X - \mathbb{E}(X)| \geq a) \leq \frac{\mathbb{V}(X)}{a^2}$

15. Inégalité triangulaire : $|\mathbb{E}(X)| \leq \mathbb{E}(|X|)$

16. Indépendance dans leur ensemble des variables aléatoires $(X_i)_{i=1,...,n}$: On admettra l'équivalence des définitions suivantes

    (a) $F_{(X_1,...,X_n)}(x_1, ..., x_n) = \prod_{i=1}^{n} F_{X_i}(x_i)$

    (b) Pour toutes fonctions positives g_i : $\mathbb{E}\left(\prod_{i=1}^{n} g_i(X_i)\right) = \prod_{i=1}^{n} \mathbb{E}\left(g_i(X_i)\right)$

    (c) Pour toutes fonctions g_i bornées : $\mathbb{E}\left(\prod_{i=1}^{n} g_i(X_i)\right) = \prod_{i=1}^{n} \mathbb{E}\left(g_i(X_i)\right)$

    (d) Pour toutes fonctions g_i telles que $\mathbb{E}\left(\prod_{i=1}^{n} |g_i(X_i)|\right) < \infty$ et $\mathbb{E}\left(|g_i(X_i)|\right) < \infty$: $\mathbb{E}\left(\prod_{i=1}^{n} g_i(X_i)\right) = \prod_{i=1}^{n} \mathbb{E}\left(g_i(X_i)\right)$

Et dans le cas où $(X_i)_{i,...,n}$ est un vecteur aléatoire admettant une densité : $f_{(X_1,...,X_n)}(x_1,...,x_n) = \prod_{i=1}^n f_{X_i}(x_i)$.

17. Les variables aléatoires $(X_i)_{i=1,...,n}$ sont identiquement distribuées si elles ont la même fonction de répartition.

18. Les variables aléatoires $(X_i)_{i=1,...,n}$ sont dites iid si elles sont indépendantes et identiquement distribuées

19. Suite $(X_i)_{i\geq 1}$ de variables aléatoires indépendantes (respectivement identiquement distribuées) : si pour tout $n \geq 2$, les variables $(X_i)_{1\leq i\leq n}$ sont indépendantes (respectivement identiquement distribuées)

20. Savoir calculer les moments des lois usuelles : uniforme, normale, de Poisson, de Bernoulli, binomiale, exponentielle, Gamma, ...

21. Exemples simples de transfert : à partir de la densité de X, les candidats doivent savoir retrouver la densité de $aX + b$, X^2, $\exp(X)$,...

1.3 Convergence des variables aléatoires

1. Convergence presque-sure de X_n vers X: $\mathbb{P}(\{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\}) = 1$

2. Convergence dans L^p (pour $p \geq 1$): Soit une suite de variables aléatoires X_n telle que $\mathbb{E}(|X_n|^p) < \infty$ pour tout n et X une variable aléatoire telle que $\mathbb{E}(|X|^p) < \infty$, X_n converge vers X dans L^p si $\mathbb{E}(|X_n - X|^p)$ tend vers 0 quand n tend vers ∞.

3. Convergence en probabilité : la suite de variables aléatoires X_n converge vers la variable aléatoire X si pour tout $\epsilon > 0$ $\mathbb{P}(|X_n - X| > \varepsilon)$ tend vers 0 quand n tend vers ∞.

4. Convergence en loi de X_n vers X: En tout point de continuité x de F_X, $F_{X_n}(x)$ converge vers $F_X(x)$. Autres caractérisations :$\mathbb{E}(f(X_n))$ converge vers $\mathbb{E}(f(X))$ pour toute f continue et bornée; ou pour toute f Lipschitz bornée; ou $\mathbb{E}(\exp(itX_n))$ converge vers $\mathbb{E}(\exp(itX))$ pour tout $t \in \mathbb{R}$ (Théorème de Lévy).

5. Relation entre les différents modes de convergence : la convergence dans L^p implique la convergence dans L^q pour $p \geq q \geq 1$, la convergence dans L^p implique la convergence en probabilité, la convergence presque-sure implique la convergence en probabilité, la convergence en probabilité implique la convergence en loi. Théorème de convergence dominée pour passer de la convergence presque-sure à la convergence dans L^p.

6. Théorème de continuité pour la convergence en probabilité, presque-sure, en loi. Si X_n converge vers X en probabilité, (respectivement en loi, respectivement presque-surement) et si g est une fonction continue en tout point du support de X alors $g(X_n)$ converge vers $g(X)$ en probabilité (respectivement en loi, respectivement presque-surement).

7. Lemme de Slutsky : si X_n converge en loi vers X et Y_n converge en loi vers une constante c, alors $X_n + Y_n$ converge en loi vers $X + c$ et Y_nX_n converge en loi vers cX.

1.4 Statistique

1. Définition d'un modèle statistique paramétrique paramétré par $\Theta \subset \mathbb{R}^K$: $\{(\mathbb{P}_\theta)_{\theta\in\Theta}, \Theta\}$, comme famille de distribution indicée par Θ.

2. Identification de $\theta \in \Theta$ dans le modèle statistique : θ est identifiable si $\mathbb{P}_\theta \neq \mathbb{P}_{\theta'}$ dès que $\theta \neq \theta'$

3. Loi faible des grands nombres pour des variables de variance finie : si les X_i sont des variables iid, telles que $\mathbb{E}(X_1^2) < \infty$ alors $\frac{1}{n}\sum_{i=1}^n X_i$ converge en probabilité vers $\mathbb{E}(X_1)$. Savoir la démontrer simplement par l'inégalité de Bienaymé-Tchebychev.

4. Loi forte des grands nombres pour des variables d'espérance finie : si les X_i sont des variables iid, telles que $\mathbb{E}(|X_1|) < \infty$ alors $\frac{1}{n}\sum_{i=1}^n X_i$ converge presque-surement, dans L^1 et en probabilité vers $\mathbb{E}(X_1)$.

5. Théorème Central Limite : si les X_i sont des variables iid, telles que $\mathbb{E}(X_1^2) < \infty$ alors $\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^n X_i - \mathbb{E}(X_1)\right)/\sqrt{\mathbb{V}(X_1)}$ converge en loi vers une loi normale centrée réduite, en particulier :
$$\mathbb{P}\left(a\sqrt{\mathbb{V}(X_1)} \leq \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^n X_i - \mathbb{E}(X_1)\right) \leq b\sqrt{\mathbb{V}(X_1)}\right)$$
converge vers
$$\frac{1}{\sqrt{2\pi}}\int_a^b \exp\left(-\frac{t^2}{2}\right)dt.$$

6. Définition d'un estimateur : Un estimateur de $g(\theta) \in \mathbb{R}^k$ est une variable (ou un vecteur si $k > 1$) aléatoire de la forme $T_n = \varphi_n(X_1, ..., X_n)$ où $X_1, ..., X_n$ sont n variables aléatoires dont les réalisations sont observées. La fonction φ_n ne dépend pas de θ

7. Le biais d'un estimateur est $B = \mathbb{E}(T_n) - g(\theta)$, un estimateur sans biais de $g(\theta)$ est un estimateur T_n tel que $\mathbb{E}(T_n) = g(\theta)$

8. Risque quadratique d'un estimateur T_n (pour $g(\theta) \in \mathbb{R}$) : $R = \mathbb{E}\left(|T_n - g(\theta)|^2\right)$

9. Décomposition du risque quadratique T_n (pour $g(\theta) \in \mathbb{R}$) : $R = B^2 + V$, où V est la variance de T_n.

10. Un estimateur consistant est un estimateur T_n qui converge en probabilité vers $g(\theta)$. Un estimateur fortement consistant est un estimateur T_n qui converge presque-surement vers $g(\theta)$.

11. Un intervalle de confiance de $g(\theta) \in \mathbb{R}$ de niveau de confiance $1-\alpha$ ($\alpha \in [0,1]$) de $g(\theta)$ est un intervalle $[U_n, V_n]$ tel que :

    - U_n et V_n sont des fonctions des observations $(X_1, ..., X_n)$ et de n mais pas de θ
    - $\mathbb{P}(U_n \leq V_n) = 1$
    - $\mathbb{P}(U_n \leq g(\theta) \leq V_n) \geq 1 - \alpha$

12. Un intervalle de confiance asymptotique de $g(\theta) \in \mathbb{R}$ de niveau de confiance $1-\alpha$ ($\alpha \in [0,1]$) de $g(\theta)$ est une suite d'intervalle $[U_n, V_n]$ tel que :

    - U_n et V_n sont des fonctions des observations $(X_1, ..., X_n)$ et de n mais pas de θ,
    - $\mathbb{P}(U_n \leq V_n) = 1$,
    - $\mathbb{P}(U_n \leq g(\theta) \leq V_n) \geq 1 - \alpha_n$ pour une suite α_n qui tend vers α.

All these notions are