**Pour les vacances :** tous les exercices de proba de la banque CCINP ($\simeq$ un par jour !)

**Exercice 1** (Inégalité de Jensen en proba)**.** Soit $X$ une v.a. réelle admettant une espérance et $f$ sur une fonction concave sur un intervalle $I$ contenant $X(\Omega)$. On suppose aussi que $f(X)$ admet une espérance.

Montrer alors que $\mathbb{E}(f(X)) \leq f(\mathbb{E}(X))$.

**Remarque :** si $X$ a un nombre *fini* de valeurs, vous connaissez déjà ce résultat, pourquoi ?

*Indication pour le cas général* – Si $x_0$ est un réel dans $I$, expliquer qu'il existe une fonction affine $g$ telle que $g(x_0) = f(x_0)$ et telle que $\forall\, x \in X(\Omega)$, $g(x) \geq f(x)$.

En écrivant $g(x) = a(x - x_0) + f(x_0)$ et en choisissant $x_0 = \mathbb{E}(X)$, conclure, en appliquant les propriétés de l'espérance.

**Exercice 2** (La géométrie cachée des variances...)**.** Soit $\mathcal{E}$ un ensemble de v.a.discrète centrées (i.e. d'espérance nulle) définies sur un même espace probabilisé et admettant une variance.

On suppose que pour tout couple $(X_1, X_2) \in \mathcal{E}^2$, on a : $\frac{X_1+X_2}{2} \in \mathcal{E}$.

On suppose aussi que $\{\mathbb{V}(X),\ X\ \in \mathcal{E}\}$ admet un minimum, qu'on note $\mathbb{V}_0$.

Soient $X_1$ et $X_2$ deux v.a. de $\mathcal{E}$ vérifiant $\mathbb{V}(X_1) = \mathbb{V}(X_2) = \mathbb{V}_0$.

Montrer que $X_1 = X_2$ presque sûrement.

*Indication 1 :* Comment montrer la conclusion avec des variances et des espérances ?

*Indication 2 :* Connaissez-vous l'égalité du parallélogramme en géométrie ?

**Exercice 3** (Droite de régression linéaire)**.** Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé.

1) Justifier que l'application $N\ :\ \mathscr{L}^2(\Omega, \mathbb{R}) \to \mathbb{R}^+,\ X \mapsto N(X) := \sqrt{\mathbb{E}(X^2)}$ est une semi-norme sur $\mathscr{L}^2(\Omega, \mathbb{R})$.

2) On fixe arbitrairement deux v.a.d. $X, Y \in \mathscr{L}^2(\Omega, \mathbb{R})$. Nous souhaitons pratiquer une régression linéaire, c'est-à-dire approcher, au sens de la semi-norme du 1), une variable aléatoire $Y$ par une fonction affine de la variable aléatoire $X$. Autrement dit, nous cherchons à minimiser l'application $F_{X,Y} : \mathbb{R}^2 \to \mathbb{R}_+$définie par :

$$\forall (a,b) \in \mathbb{R}^2, \quad F_{X,Y}(a,b) = N^2(Y - (aX + b)).$$

Pour deux v.a. de variances non nulles, on note $\rho(X,Y) = \dfrac{\mathrm{Cov}(X,Y)}{\sigma(X).\sigma(Y)}$ leur coefficient de corrélation.

a) On se place dans le cas particulier où les variables $X$ et $Y$ sont centrées et réduites. On a alors $\rho(X,Y) = \mathrm{Cov}(X,Y)$.

Montrer que l'application $F_{X,Y}$ atteint son minimum en l'unique point $(a_0, b_0) = (\rho(X,Y), 0)$ où elle prend la valeur $1 - \rho^2(X,Y)$.

b) On revient au cas général où on suppose seulement $\sigma(X)$ et $\sigma(Y)$ non nuls. Montrer que l'application $F_{X,Y}$ atteint son minimum en un unique point $(a_0, b_0)$, donné par :

$$a_0 = \frac{\rho(X,Y)\sigma(Y)}{\sigma(X)} \text{ et } b_0 = E[Y] - a_0 E[X],$$

où elle prend la valeur $\sigma^2(Y) \times \big(1 - \rho^2(X,Y)\big)$.

*Indication* – On pourra commencer par le cas centré et se ramener au a), puis se ramener à ce cas centré.

**Exercice 4.** Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires réelles indépendantes de loi de Bernoulli de paramètre $p$. On pose $Y_n = X_n X_{n+1}$.

a) Calculer l'espérance de $S_n = \sum_{i=1}^n Y_i$.

b) Montrer que $\mathbb{E}\left(S_n^2\right) = np^2 + 2(n-1)p^3 + p^4(n-1)(n-2)$.

c) Que dire de $\mathbb{P}\left(\left|\frac{S_n}{n} - m\right| \geqslant \varepsilon\right)$ ?

**Exercice 5.** Une action vaut initialement 1 euro. À chaque instant $n \geqslant 1$, sa valeur est multipliée par une quantité aléatoire $Z_n$. On suppose que les variables $Z_n$ sont indépendantes et de même loi, telles que :

$$\mathbb{P}(Z_n = 1 + a) = \mathbb{P}(Z_n = 1 - a) = \frac{1}{2} \quad \text{avec} \quad a \in ]0,1[.$$

On note $X_n$ la valeur de l'action à l'instant $n$ et l'on pose $Y_n = \ln(Z_n)$. On définit, pour tout entier naturel non nul $n$, la variable :

$$\widehat{Y}_n = \frac{Y_1 + \cdots + Y_n}{n}.$$

a) Calculer, pour tout $n \geqslant 0$, l'espérance et la variance de $X_n$. Déterminer la limite de $\mathbb{V}(X_n)$ quand $n$ tend vers l'infini.

b) Montrer qu'il existe $\delta > 0$ tel que :

$$\lim_{n\to+\infty} \mathbb{P}\left(\widehat{Y}_n > -\delta\right) = 0$$

c) En déduire que pour tout $\varepsilon > 0$ :

$$\lim_{n\to+\infty} \mathbb{P}\left(X_n > \varepsilon\right) = 0$$

**Exercice 6** (Formule de Wald pour les fonctions génératrices de sommes aléatoires de variables aléatoires)**.** Soit $(X_n)$ une suite de v.a.d. à valeurs dans $\mathbb{N}$, mutuellement indépendantes de même loi qu'une v.a. $X$ définies sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ . Soit $T$ une v.a.d. à valeurs dans $\mathbb{N}$, définie sur le même espace et telle que $(T, (X_n)_{n\in\mathbb{N}}$ soit mutuellement indépendantes. On note $G_X$ (resp. $G_T$) la fonction génératrice de $X$ (et donc de chaque $X_i$) (resp. de $T$).

Pour chaque $n \in \mathbb{N}$, on note $S_n = X_1 + \cdots + X_n$ et on note $S_T$ la fonction définie par $\forall \omega \in \Omega$, $S_T(\omega) = S_{T(\omega)}(\omega)$.

a) Montrer que $S_T$ est une v.a.d.

b) Montrer que $G_{S_T} = G_T \circ G_X$

**Exercice 7** (La loi forte des grands nombres dans l'hyp. $L^2$)**.** Soit $(X_n)$ une suite de v.a. réelles indépendantes de même loi. On suppose que $E(X_1^2) < +\infty$ et donc a fortiori $E(X_1) < +\infty$.

On pose $\mu = E(X_1)$ et $S_n = X_1 + \cdots + X_n$.

a) Soit $\varepsilon > 0$, majorer $P\left(\left|\frac{S_n}{n} - \mu\right| > \varepsilon\right)$ en fonction de $V(X_1)$ et en déduire la loi faible des grands nombres :

$$\forall\, \varepsilon > 0, \quad P\left(\left|\frac{S_n}{n} - \mu\right| > \varepsilon\right) \underset{n\to+\infty}{\longrightarrow} 0.$$

b) Montrer que la série $\sum_m P\left(\left|\frac{S_{m^2}}{m^2} - \mu\right| > \varepsilon\right)$ converge.

c) En déduire, à l'aide du premier lemme de Borel-Cantelli, que $\frac{S_{m^2}}{m^2} \underset{m\to+\infty}{\longrightarrow} \mu$ presque sûrement.

d) On suppose dans cette question que les $(X_n)$ sont positifs. Déduire de la question précédente, à l'aide d'un encadrement, que $\frac{S_n}{n} \underset{m\to+\infty}{\longrightarrow} \mu$ presque sûrement.

e) A l'aide de $X_n^+ = \max(X_n, 0)$ et $X_n^- = \max(-X_n, 0)$ généraliser le résultat du d) sans l'hyp. que les v.a. $(X_n)$ sont positives.

**Exercice 8** (Graphes aléatoires : bien pour les écrits)**.** Soit $G$ un graphe d'Erdos Renyi de paramètre $\mathcal{G}(n,p)$ i.e. un graphe à $n$ sommets et pour chaque des $\binom{n}{2}$ arêtes possibles on tire suivant une loi de Bernoulli de paramètre $p$ le fait que cette arête existe ou pas, les tirages étant indépendants pour les différentes arêtes.

a) Montrer que le nombre moyen d'arêtes d'un tel graphe est $\binom{n}{2}p$.

b) Montrer que le nombre moyen de sommets isolés d'un tel graphe est $n(1-p)^{n-1}$.

c) (Arrive la méthode du premier moment cf. ex. du cours))

On considère désormais pour chaque $n \in \mathbb{N}$ un graphe d'Erdos Renyi de paramètre $\mathcal{G}(n, p_n)$ avec $p_n \underset{n\to+\infty}{\longrightarrow} 0$ et $\frac{\ln(n)}{n} = o(p_n)$.

On note $N_n$ la v.a. renvoyant le nombre de sommets isolés dans $G_n$.

Montrer que $\mathbb{P}(N_n = 0) \underset{n\to+\infty}{\longrightarrow} 1$

d) Reprendre la question c) avec le même résultat si $p_n \underset{n\to+\infty}{\sim} c\frac{\ln(n)}{n}$ avec $c > 1$.

**Exercice 9** (Suite de l'exercice 8, avec la méthode du second moment)**.** On prend les mêmes hyp. que dans l'exercice 8 sauf que cette fois :

$$p_n \underset{n\to+\infty}{\sim} c\frac{\ln(n)}{n}$$

avec cette fois $c \in ]0,1[$.

Montrer qu'alors $\mathbb{P}(N_n = 0) \underset{n\to+\infty}{\longrightarrow} 0$.