Lorsque $\mathcal{K}$ est fini on définit la **distance en variation** entre les probabilités $P$ et $Q$ par

$$D(P,Q) = \frac{1}{2}\sum_{k\in\mathcal{K}} |p_k - q_k| \,. \tag{i}$$

I. 1) Lorsque $\mathcal{K}=\{0,1\}$, exprimer $D(P,Q)$ en fonction de $p_1$ et $q_1$.
I. 2) Lorsque $\mathcal{K}=\mathbb{N}$, vérifier que la série de terme général $(|p_k - q_k|)_{k\in\mathbb{N}}$ est convergente.
On étend donc la définition de la distance en variation donnée par (i) au cas où $\mathcal{K}=\mathbb{N}$.
I. 3) Vérifier que $|P(A) - Q(A)| \in [0,1]$ pour tout $A \in \mathcal{A}$.
I. 4) Montrer que pour tout $A \in \mathcal{A}$

$$2\,|P(A) - Q(A)| = \left|\sum_{k\in A}(p_k - q_k)\right| + \left|\sum_{k\in \overline{A}}(p_k - q_k)\right| .$$

I. 5) En déduire que pour tout $A \in \mathcal{A}$

$$|P(A) - Q(A)| \leq D(P,Q) \,. \tag{ii}$$

I. 6) Montrer que la partie $A_0 = \{k \in \mathcal{K} : q_k \geq p_k\}$ réalise l'égalité dans (ii), c'est à dire que

$$|P(A_0) - Q(A_0)| = D(P,Q) \,.$$

I. 7) Démontrer la formule

$$D(P,Q) = 1 - \sum_{k\in\mathcal{K}} (p_k \wedge q_k) \,.$$

I. 8) On considère un couple de variables aléatoires $(X,Y)$ tel que $X$ soit de loi $P$ et $Y$ soit de loi $Q$. Autrement dit, pour tout $k \in \mathcal{K}$

$$\mathbb{P}(X = k) = p_k \text{ et } \mathbb{P}(Y = k) = q_k.$$

Montrer que $D(P,Q) \leq \mathbb{P}(X \neq Y)$.

## Partie II (Couplage binomiale-Poisson)

Soit $n$ un entier strictement positif et $\lambda$ un réel strictement positif, strictement plus petit que $n$. L'objet de cette deuxième partie est d'étudier un exemple: l'approximation de la loi binomiale par la loi de Poisson en terme de **distance en variation**. Plus précisément, si d'une part $\mathcal{B}(n,\lambda/n)$ désigne la loi binomiale de paramètres $n$ et $\lambda/n$ et si d'autre part on note $\mathcal{P}(\lambda)$ la loi de Poisson de paramètre $\lambda$, le but est de prouver la majoration suivante:

$$D(\mathcal{B}(n,\lambda/n), \mathcal{P}(\lambda)) \leq \frac{\lambda^2}{n} \tag{iv}$$

où $D$ est définie au (i).
II. 1) Soit $Y_1, ..., Y_n$ $n$ variables aléatoires indépendantes et de même loi de Poisson de paramètre $\lambda/n$, donner sans démonstration la loi de $\sum_{i=1}^n Y_i$.

II. 2) Vérifier que pour tout $x \in [0,1]$

$$f(x) = 1 - (1-x)\exp(x)$$

appartient à $[0,1]$.

Soit $U_1, ..., U_n$ $n$ variables aléatoires indépendantes de même loi de Bernoulli de paramètre $f(\lambda/n)$. On suppose que les variables $U_1, ..., U_n$ sont indépendantes des variables $Y_1, ..., Y_n$ de la question II. 1). Pour $i \in \{1, ..., n\}$, on pose $X_i = 0$ si $U_i = Y_i = 0$ et $X_i = 1$ sinon.

II. 3) Vérifier que pour tout $i \in \{1, ..., n\}$, $X_i$ suit une loi de Bernoulli de paramètre $\lambda/n$ et donner la loi de $\sum_{i=1}^n X_i$.

II. 4) Montrer que pour tout $i \in \{1, ..., n\}$

$$\mathbb{P}(X_i \neq Y_i) \leq \frac{\lambda^2}{n^2}.$$

(On pourra établir que pour tout $x$ réel $1 + x \leq \exp(x)$).

II. 5) Montrer que

$$\mathbb{P}\left(\sum_{i=1}^n X_i \neq \sum_{i=1}^n Y_i\right) \leq P\left(\bigcup_{i=1}^n \{X_i \neq Y_i\}\right).$$

II. 6) En déduire que

$$\mathbb{P}\left(\sum_{i=1}^n X_i \neq \sum_{i=1}^n Y_i\right) \leq \frac{\lambda^2}{n},$$

puis conclure quant à (iv).

II. 7) Quel résultat connu peut-on déduire de (iv) lorsque $n$ tend vers l'infini?

## Problème 2 (Couplage exponentielle-normale)

Dans ce problème $X$ désigne une variable aléatoire de loi normale centrée et réduite, $\varphi$ sa densité de probabilité et $\Phi$ sa fonction de répartition. On note par ailleurs $f$ la densité de la loi exponentielle de paramètre égal à 1. On définit également pour tout nombre réel $x$, $g(x) = -\ln(1 - \Phi(x))$ puis $Y = g(X)$.

On admettra que $X$ admet des moments de tout ordre, ce qui signifie que pour tout entier naturel $k$, l'intégrale $\int_{-\infty}^{+\infty} |x|^k \varphi(x)\,dx$ converge.

## Partie I (Quantiles gaussiens)

On démontre dans cette partie des résultats utiles pour la partie II.

I. 1) Montrer que $\Phi$ réalise une bijection de $\mathbb{R}$ sur $]0,1[$ dont on notera $\Phi^{-1}$ l'application réciproque.

I. 2) Calculer la fonction de répartition de $Y$ puis constater que $Y$ suit la loi exponentielle de paramètre 1.

I. 3) *a*) Vérifier la validité de l'identité suivante

$$1 - \Phi(x) = \frac{\varphi(x)}{x} - \int_x^{+\infty} \frac{\varphi(t)}{t^2}dt \text{ pour tout } x > 0.$$

*b*) En déduire l'encadrement

$$1-\frac{1}{x^2} \leq \frac{x\,(1-\Phi\,(x))}{\varphi\,(x)} \leq 1 \text{ pour tout } x>0. \tag{E}$$

*Indication pour la minoration:* On pourra montrer tout d'abord que $\int_x^{+\infty} t^{-2}\varphi\,(t)\,dt \leq x^{-3}\int_x^{+\infty} t\varphi\,(t)\,dt$.

*c*) Montrer l'équivalence

$$1-\Phi\,(x) \sim_{x\to+\infty} \frac{\varphi\,(x)}{x}\,.$$

*d*) En utilisant (E) énoncée à la question I. 3)*b*), montrer que pour tout $x>1$

$$\ln\left(1-\frac{1}{x^2}\right) \leq \ln\,(x) - g\,(x) + \frac{1}{2}\ln\,(2\pi) + \frac{x^2}{2} \leq 0$$

et en déduire l'équivalence

$$g\,(x) \sim_{x\to+\infty} \frac{x^2}{2}.$$

## Partie II (Inégalité de transport)

On définit une application $h$ sur $[0,+\infty[$ par: $h\,(t) = t\ln\,(t) - t + 1$ pour $t>0$ et $h\,(0) = 1$.

II. 1) Vérifier que $h$ est une application continue de $[0,+\infty[$ vers $[0,+\infty[$.

Sous réserve qu'elle converge, on note $K\,(f,\varphi)$ la valeur de l'intégrale $\int_{-\infty}^{+\infty} \varphi\,(x)\,h\left(\frac{f(x)}{\varphi(x)}\right)dx$. On désire vérifier l'inégalité (dite de transport) suivante

$$E\left((X-Y)^2\right) \leq 2K\,(f,\varphi)\,. \tag{v}$$

II. 2) Montrer que $g$ est une application dérivable sur $\mathbb{R}$. Pour tout $x$ réel calculer $g'\,(x)$ et vérifier l'identité $g'\,(x)\,f\,(g\,(x)) = \varphi\,(x)$.

II. 3) Vérifier que l'intégrale définissant $K\,(f,\varphi)$ converge et montrer que

$$K\,(f,\varphi) = \int_{-\infty}^{+\infty} \varphi\,(x)\ln\left[f\,(g\,(x))\,/\varphi\,(g\,(x))\right]dx.$$

II. 4) Montrer que $\int_{-\infty}^{+\infty} \varphi\,(x)\,(x-g\,(x))^2\,dx$ converge et justifier l'égalité suivante:

$$E\left((X-Y)^2\right) = \int_{-\infty}^{+\infty} \varphi\,(x)\,(x-g\,(x))^2\,dx.$$

II. 5) Montrer que l'intégrale $\int_{-\infty}^{+\infty} \varphi\,(x)\,(1-g'\,(x))\,dx$ converge.

II. 6) Démontrer que

$$K\,(f,\varphi) \geq \int_{-\infty}^{+\infty} \varphi\,(x)\ln\left[\varphi\,(x)\,/\varphi\,(g\,(x))\right]dx + \int_{-\infty}^{+\infty} \varphi\,(x)\,(-g'\,(x)+1)\,dx.$$

(On pourra utiliser en la justifiant l'inégalité $\ln\,(u) \leq u-1$, pour tout $u$ réel strictement positif).

II. 7) Conclure grâce à une intégration par parties que l'on justifiera soigneusement.