U.D.L Sidi Bel Abbès — **Module : Analyse des données**
Faculté des Sciences Exactes — **Responsable : M. HAMMAD**
Département : Probabilités-Statistique — **Jeudi 25/03/2021**
Master 2 : Statistique/Probabilités et ses Applications — **Durée :1h30**

EXAMEN FINAL

Exercice 1 **(06 points).**

1. Étant donné le tableau $X = (x_{ji})_{1\leq i\leq n,\, 1\leq j\leq p}$ de données quantitatives sur n individus décrits par p variables. Démontrer que :

$$R = D_{1/\sigma} \cdot V \cdot D_{1/\sigma} = Z \cdot D_p \cdot Z^{'},$$

où Z est le tableau centré réduit de X, V est la matrice des variances-covariances, D_p est la matrice diagonale des poids et $D_{1/\sigma}$ est la matrice des inverses des écartes types.

2. Supposons maintenant que le tableau de données X est centré. Vérifier que :
   – $Var(x^j) = \parallel x^j \parallel^2_{D_p} = D_p(x^j, x^j)$ et $\sigma_j = \parallel x^j \parallel_{D_p}$.
   – $Cov(x^j, x^k) = D_p(x^j, x^k)$ et $r(x^j, x^k) = \dfrac{D_p(x^j, x^k)}{\parallel x^j \parallel_{D_p} \cdot \parallel x^k \parallel_{D_p}}$, pour $j \neq k$.

Exercice 2 **(14 points).**

On considère le tableau de données de type $(3, 2)$ suivant :

$$X = \begin{pmatrix} 2 & 3 \\ 4 & 5 \\ 6 & 1 \end{pmatrix}$$

– Donner le tableau Z de données centrés et réduites(normés).
– Déterminer la matrice R des corrélations.
– Diagonaliser la matrice R. On note λ_1 et λ_2 ses valeurs prroppres avec $\lambda_1 > \lambda_2$.
– Déterminer F_j les axes factoriels (les sous espaces propres). Préciser le vecteur unitaire $\sqcup_j$ de chaque axe F_j et vérifier que ces axes sont perpendiculaires.
– Quelle est la qualité de la représentation avec un ou deux axes ?.
– Calculer les composantes principales et représenter les individus dans le plan factoriel.