CORRIGÉ

Par Francis Raccaglia, professeur au lycée Thiers à Marseille.

Partie I - Lois de Bernoulli généralisées

1.a) Par définition $M = \begin{pmatrix} a_1 \\ \vdots \\ a_k \end{pmatrix} (1 \dots 1)$. D'où $\boxed{M \text{ est la matrice carrée d'ordre } k \text{ d'élément générique } m_{i,j} = a_i}$. Toutes les colonnes de M sont donc égales à a ce qui implique que $\text{rg}(M) = 1$.

---

b) $Ma = (a\,{}^t u)a = a({}^t ua)$. Or ${}^t ua = \sum_{i=1}^{k} a_i$, d'où $Ma = \left(\sum_{i=1}^{k} a_i\right) a$. On a donc $Ma = \alpha a$ et a n'étant pas la matrice nulle cela montre que $\boxed{\alpha \text{ est une valeur propre de } M}$ et a un vecteur propre associé à cette valeur propre.

---

c) Par définition $M^2 = (a\,{}^t u)(a\,{}^t u) = a({}^t ua)\,{}^t u$. ${}^t ua$ étant un réel on peut le permuter avec la matrice a :

$$M^2 = \left(\sum_{i=1}^{k} a_i\right) {}^t ua = \alpha M.$$

Le polynôme $X^2 - \alpha X$ est un polynôme annulateur de M, il a α et 0 comme racines, donc d'après le cours

$$\boxed{\text{si } \lambda \text{ est une valeur propre de } M \text{ alors } \lambda = \alpha \text{ ou } \lambda = 0}.$$

---

d) Si $\alpha = 0$, M possède 0 comme unique valeur propre, étant de rang 1, M n'est pas la matrice nulle donc M n'est pas diagonalisable.

Si $\alpha \neq 0$, puisque $\text{rg}(M) = 1$, le noyau de M est de dimension $k-1$ (théorème du rang) donc 0 est une valeur propre de M et le sous espace propre associé est de dimension $k-1$. α est une autre valeur propre de M dont le sous espace propre associé est au moins de dimension 1. On a donc d'après les résultats du cours :

$$k \leq \dim(E_0(M)) + \dim(E_\alpha(M)) \leq k$$

d'où l'égalité, ce qui prouve que $\boxed{M \text{ est diagonalisable}}$.

---

e) $I_k - M$ est inversible si et seulement si 1 n'est pas valeur propre de M i.e. si et seulement si $\alpha \neq 1$.

---

f) Si $\alpha = 1$, M vérifie $M^2 = M$ et ainsi on peut affirmer que M est la matrice canonique associée à un projecteur de $\mathbb{R}^k$. D'après l'écriture de M,

$$\boxed{\text{l'image de ce projecteur est engendrée par le vecteur } (a_1, \dots, a_k).}$$

De plus si X est la matrice colonne associée à un vecteur $x = (x_1, ..., x_k)$ de $\mathbb{R}^k$ dans la base canonique, $MX = 0$ si et seulement si, pour tout $i \in [\![1, k]\!]$, $a_i \sum_{j=1}^{k} x_j = 0$. Les a_i n'étant pas tous nuls, cela équivaut à $\sum_{j=1}^{k} x_j = 0$.

$$\boxed{\text{Le noyau est donc composé des } x = (x_1, ..., x_k) \text{ tels que } \sum_{j=1}^{k} x_j = 0.}$$

Ce projecteur est orthogonal si est seulement si l'image et le noyau sont orthogonaux. Or l'orthogonal du noyau est la droite engendrée par le vecteur dont toutes les composantes valent 1. D'où ce projecteur sera orthogonal uniquement lorsque tous les a_i seront égaux.

2.a) Il est évident que si $[X = e_i]$ est réalisé alors $[X_i = 1]$ l'est aussi. Par contre on n'a pas forcément $[X_i = 1] \subset [X = e_i]$. En effet d'après l'énoncé, le vecteur aléatoire X est presque surement à valeurs dans $\{e_1, ..., e_k\}$, ce qui ne l'empêche pas de prendre éventuellement la valeur $e_1 + ... + e_k$ avec la probabilité 0.

Ainsi on peut affirmer que : $p_i \leq P([X_i = 1])$ (1). Et aussi que :$\bigcup_{j \neq i} [X = e_j] \subset [X_i = 0]$. En passant aux probabilités, $1 - p_i \leq P([X_i = 0])$ (2).

On obtient alors que : $1 \leq P([X_i = 0]) + P([X_i = 1])$. Cette somme étant majorée par 1 par incompatibilité, elle vaut 1 et les inégalités (1) et (2) sont des égalités. $\boxed{X_i \text{ suit bien une loi de Bernoulli de paramètre } p_i.}$

D'où $\boxed{\mathcal{E}(X) = \begin{pmatrix} p_1 \\ \vdots \\ p_k \end{pmatrix}.}$

Pour éviter de compliquer les raisonnements qui suivent, on considèrera que $X(\Omega) = \{e_1, ..., e_k\}$, ce qui en termes probabilistes ne change rien à la suite de cette partie.

---

b) A priori, $X_1 + X_2$ peut prendre les valeurs $0, 1, 2$. Mais les événements $[X_1 = 1]$ et $[X_2 = 1]$ sont incompatibles, donc $X_1 + X_2$ ne peut prendre que les valeurs 0 et 1. Ainsi :

$$\boxed{X_1 + X_2 \text{ suit une loi de Bernoulli de paramètre } E(X_1 + X_2) = p_1 + p_2.}$$

---

c) Posons pour tout $i \in [\![1, k]\!]$, $q_i = 1 - p_i$. On a $V(X_1) = p_1 q_1$, $V(X_2) = p_2 q_2$. D'où $\text{Cov}(X_1, X_2)$ existe et $\text{Cov}(X_1, X_2) = \frac{1}{2}(V(X_1 + X_2) - V(X_1) - V(X_2)) = \frac{1}{2}((p1 + p2)(1 - p_1 - p2) - p_1(1 - p_1) - p_2(1 - p_2))$. Après développements et simplifications, il ne reste que $-p_1 p_2$ i.e. $\boxed{\text{Cov}(X_1, X_2) = -p_1 p_2}$.

d) Par analogie, on a pour tout couple $(i, j) \in [\![1, k]\!]^2$ tel que $i \neq j$, $\text{Cov}(X_i, X_j) = -p_i p_j$. D'où

$$\boxed{\mathcal{V}(X) = \begin{pmatrix} p_1q_1 & -p_1p_2 & \dots & \dots & -p_1p_k \\ -p_2p_1 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & -p_{k-1}p_k \\ -p_kp_1 & \dots & \dots & -p_kp_{k-1} & p_kq_k \end{pmatrix}}$$

3.a) De manière générale, si A est une matrice, notons $A_{i,j}$ le coefficient se trouvant sur la ligne i et la colonne j. On notera dans cette question plus simplement M à la place de $M(p)$.

D'après la définition du produit matriciel, $((I_k - M)\mathrm{Diag}(p))_{i,j} = \displaystyle\sum_{r=1}^{k}(I_k - M)_{i,r}\mathrm{Diag}(p)_{r,j} = ((I_k - M)_{i,j}p_j$. Or $M_{i,j} = p_i$, d'où $((I_k - M)\mathrm{Diag}(p))_{i,j} = \begin{cases}(1-p_i)p_i & \text{si } i = j \\ -p_ip_j & \text{sinon.}\end{cases}$

On a bien l'égalité annoncée.

---

b) Puisque les p_i sont non nuls, $\mathrm{Diag}(p)$ est inversible. D'où le rang de $\mathcal{V}(X)$ est égal au rang de $I_k - M(p)$. Or d'après la question 1, $M(p)$ est la matrice d'un projecteur de rang 1, d'où le rang de $I_k - M(p)$ vaut $k-1$ puisque son noyau est l'image de $M(p)$.

Finalement, on a bien $\boxed{\mathrm{rg}(\mathcal{V}(X)) = k-1}$.

---

c) Soit v l'endomorphisme de $\mathbb{R}^k$ associé à $\mathcal{V}(X)$ dans la base canonique. Considérons la base $(e'_1, ..., e'_k)$ avec pour tout i, $e'_i = e_{\sigma(i)}$. On a alors : $v(e'_i) = v(e_{\sigma(i)}) = p_{\sigma(i)}q_{\sigma(i)}e_{\sigma(i)} - \displaystyle\sum_{r=1, r\neq\sigma(i)}^{k} p_rp_{\sigma(i)}e_r$.

Or $\displaystyle\sum_{r=1, r\neq\sigma(i)}^{k} p_rp_{\sigma(i)}e_r = \sum_{s=1, s\neq i}^{k} p_{\sigma(s)}p_{\sigma(i)}e_{\sigma(s)} = p_{\sigma(i)}q_{\sigma(i)}e'_i - \sum_{s=1, s\neq i}^{k} p_{\sigma(s)}p_{\sigma(i)}e'_s$.

Ceci démontre que la matrice de v dans la base $(e'_1, ..., e'_k)$ a pour éléments diagonaux les réels $p_{\sigma(i)}q_{\sigma(i)}$ et pour éléments non diagonaux les réels $-p_{\sigma(i)}p_{\sigma(j)}$, c'est à dire que la matrice de v dans la base $(e'_1, ..., e'_k)$ est $(I_k - p_\sigma @@,^t u)\mathrm{Diag}(p_\sigma)$.

Les matrices $\mathcal{V}(X)$ et $(I_k - p_\sigma {}^tu)\mathrm{Diag}(p_\sigma)$ représentent le même endomorphisme de $\mathbb{R}^k$, elles sont donc semblables.

---

d) Soit r le nombre de p_i non nuls. Soit σ une permutation de $[\![1, k]\!]$ telle que $p_{\sigma(1)} \neq 0, ..., p_{\sigma(r)} \neq 0, p_{\sigma(r+1)} = 0, ..., p_{\sigma(k)} = 0$.

La matrice $(I_k - p_\sigma {}^tu)\mathrm{Diag}(p_\sigma)$ n'a que des coefficients nuls sur les lignes et les colonnes d'indices strictement supérieur à r. De plus le bloc formé par les coefficients dont les indices de ligne et de colonne sont inférieurs à r correspond à une matrice de la forme $(I_r - a\,{}^tv)\mathrm{Diag}(a)$ où a est la matrice colonne de coefficients $p_{\sigma(1)}, ..., p_{\sigma(r)}$ et v la matrice colonne d'ordre r dont tous les coefficients valent 1. D'après les résultats des questions qui précédent, on peut dire que son rang est $r-1$.

On peut donc en conclure que $\boxed{\text{le rang de } \mathcal{V}(X) \text{ vaut } r-1 \text{ où } r \text{ est le nombre de } i \in [\![1,k]\!] \text{ pour lesquels } p_i \neq 0.}$

Partie II - Tirages avec remise dans une population stratifiée

4.a) Posons $s = \begin{pmatrix} s_1 \\ \vdots \\ s_k \end{pmatrix}$. $[S^{(n)} = s]$ est réalisé si et seulement si, pour tout $i \in [\![1, k]\!]$, s_i tirages exactement ont donné un individu de la i-ème catégorie. Donc, cet événement est réalisable si et seulement si, pour tout i, $s_i \in \mathbb{N}$ et $\sum_{i=1}^{k} s_i = n$.

De plus, grâce à l'indépendance, si s vérifie ces conditions, la probabilité est non nulle. En effet on peut par exemple la minorer par $p_1^{s_1}...p_k^{s_k}$ qui correspond à la probabilité de tirer s_1 fois un individu de la catégorie 1, puis s_2 fois un individu de la catégorie 2,..., et enfin s_k fois un individu de la catégorie k. D'où :

$\boxed{P([S^{(n)} = s]) > 0 \text{ si et seulement si les coefficients de } s \text{ sont des entiers naturels dont la somme vaut } n.}$

---

b) $S_1^{(n)}$ est la variable aléatoire réelle égale au nombre de fois où l'on a tiré un individu de la catégorie 1. Classiquement, $\boxed{S_1^{(n)} \text{ suit la loi binomiale } \mathcal{B}(n, p_1)}$.

Pour $S_1^{(n)} + S_2^{(n)}$ on raisonne de même en « fusionnant les catégories 1 et 2 » .

D'où $\boxed{S_1^{(n)} + S_2^{(n)} \text{ suit la loi binomiale } \mathcal{B}(n, p_1 + p_2)}$.

Ces variables ne sont pas indépendantes puisque par exemple $P_{[S_1=1]}([S_1^{(n)} + S_2^{(n)} = 0]) = 0$ et $P([S_1^{(n)} + S_2^{(n)} = 0]) \neq 0$.

---

c) On généralise les résultats précédents sans difficulté : $S_i^{(n)}$ suit la loi binomiale $\mathcal{B}(n, p_i)$ et si $i \neq j$, $S_i^{(n)} + S_j^{(n)}$ suit la loi binomiale $\mathcal{B}(n, p_i + p_j)$. D'où,

$$V(S_i^{(n)}) = np_i(1 - p_i) \text{ et } \mathrm{Cov}(S_i^{(n)}, S_j^{(n)}) = \frac{1}{2}\left[V(S_i^{(n)} + S_j^{(n)}) - V(S_i^{(n)}) - V(S_j^{(n)})\right]$$

Or

$$V(S_i^{(n)} + S_j^{(n)}) = n(p_i + p_j)(1 - p_i - p_j)\,,\ V(S_i^{(n)}) = np_i(1 - p_i)\,,\ V(S_j^{(n)}) = np_j(1 - p_j)$$

ce qui donne $\mathrm{Cov}(S_i^{(n)}, S_j^{(n)}) = -np_ip_j$. On obtient bien $\boxed{\mathcal{V}(S^{(n)}) = n\mathcal{V}(X^{(1)})}$.

5.a) Notons $W(\Omega) = \{w_i / i \in I\}$ où I est un ensemble fini ou indexable par les entiers naturels.

Pour tout $i \in I$, $P_H([W = w_i]) = \dfrac{P([W_i = w_i] \cap H)}{P(H)}$, d'où $P_H([W = w_i]) \leq \dfrac{P([W = w_i])}{P(H)}$. On en déduit que $0 \leq w_i^2 P_H([W = w_i]) \leq w_i^2 P([W = w_i])$.

Or $E(W^2)$ existe, donc le théorème de transfert montre que $\sum w_i^2 P([W = w_i])$ converge et par le théorème de convergence des séries à termes positifs par comparaison, $\sum w_i^2 P_H([W = w_i])$ converge. $\boxed{\text{Ceci prouve l'existence de } E(W^2|H).}$

---

b) Les propriétés de la variance sont bien entendu vraies pour la variance associée à la probabilité conditionnelle P_H !

Notons m l'espérance de W. On a donc $V(W) = E((W - m)^2)$. On applique la formule de l'espérance totale (les « existences » étant assurées par le résultat de la question précédente) :

$$V(W) = E((W - m)^2|H)P(H) + E((W - m)^2|\overline{H})P(\overline{H})$$

d'où

$$V(W) \geq E((W-m)^2|H)P(H)$$

La formule de Huyghens montre que :

$$E((W-m)^2|H) = V(W-m|H) + (E(W-m|H))^2 = V(W|H) + (E(W-m|H))^2$$

d'où $E((W-m)^2|H) \geq V(W|H)$ puis l'inégalité demandée : $\boxed{V(W) \geq V(W|H)P(H)}$.

6.a) On cherche la probabilité de l'événement $A = \bigcap_{k=1}^{+\infty}[T_i \geq k]$. Or la suite d'événements associée à l'intersection précédente est décroissante d'où d'après le cours $P(A) = \lim_{k\to+\infty} P([T_i \geq k])$.

De plus par indépendance, $P([T_i \geq k]) = (1-p_i)^{k-1}$ et $0 < 1-p_i < 1$, d'où $\lim_{k\to+\infty}(1-p_i)^{k-1} = 0$.

Classiquement, $\boxed{T_i \text{ suit la loi géométrique de paramètre } p_i}$ puisque les tirages s'effectuent avec remise.

---

b) On a : $H_k = \bigcap_{i=1}^{k-1}[X_i^{(i)} = 1]$. Par hypothèse $X_1^{(1)}, ..., X_{k-1}^{(k-1)}$ sont indépendantes, d'où $P(H_k) = \prod_{i=1}^{k-1} P([X_i^{(i)} = 1])$ et

ainsi $\boxed{P(H_k) = \prod_{i=1}^{k-1} p_i}$.

Si l'on pose que H_k est réalisé, $T_k - (k-1)$ est la variable aléatoire égale au nombre de tirages qu'il a fallu réaliser en plus des $k-1$ premiers pour obtenir un individu de la k-ième catégorie.

$\boxed{\text{La loi conditionnelle de cette variable est la loi géométrique de paramètre } p_k.}$

D'où, d'après le cours, $E(T_k - (k-1)|H_k) = \dfrac{1}{p_k}$ et $V(T_k - (k-1)|H_k) = \dfrac{1-p_k}{p_k^2}$, ce qui induit :

$$\boxed{E(T_k|H_k) = k-1+\frac{1}{p_k} \text{ et } V(T_k|H_k) = \frac{1-p_k}{p_k^2}} .$$

---

c) Posons $W = \sum_{i=1}^{k} v_i T_i$. Les T_i admettant une variance, W aussi et on peut appliquer l'inégalité du 5.b) avec l'événement H_k qui vérifie bien les hypothèses : $V(W) \geq V(W|H_k)\prod_{i=1}^{k-1} p_i$.

Or $V(W|H_k) = V(v_k T_k + \sum_{i=1}^{k-1} i v_i|H_k) = V(v_k T_k|H_k) = v_k^2 V(T_k|H_k)$ d'après les propriétés de la variance. On n'a plus qu'à remplacer $V(T_k|H_k)$ par $\dfrac{1-p_k}{p_k^2}$:

$$\boxed{V\left(\sum_{i=1}^{k} v_i T_i\right) \geq v_k^2 \frac{1-p_k}{p_k^2}\prod_{i=1}^{k-1} p_i}$$

d) On raisonne de même, la catégorie j jouant le rôle que joue la catégorie k dans les questions précédentes.

Partie III - Support et rang stochastique d'un vecteur aléatoire

7.a) $\mathcal{S}(Y)$ contient $\mathbb{R}^k$. Notons : $\mathcal{D}(Y)$ l'ensemble des dimensions des supports de Y. Cet ensemble est un sous ensemble non vide ($k \in \mathcal{D}(Y)$) de $\mathbb{N}$, $\boxed{\text{il admet donc un plus petit élément.}}$

b) Le rang stochastique est nul si et seulement si le sous-espace réduit au vecteur nul est un support de Y i.e. si et seulement si $P(Y - E(Y) = 0) = 1$, c'est à dire lorsque $\boxed{Y \text{ est quasi-constante}}$.

c) Soit F_1 et F_2 deux supports. $[Y - E(Y) \in F_1 \cap F_2] = [Y - E(Y) \in F_1] \cap [Y - E(Y) \in F_2]$ d'où d'après la formule du crible au rang 2 :

$$P([Y - E(Y) \in F_1 \cap F_2]) = \underbrace{P([Y - E(Y) \in F_1]) + P([Y - E(Y) \in F_2])}_{=2} - \underbrace{P([Y - E(Y) \in F_1] \cup [Y - E(Y) \in F_2])}_{\leq 1}$$

d'où $P([Y - E(Y) \in F_1 \cap F_2]) \geq 1$ et comme c'est une probabilité elle vaut 1 ce qui prouve que $\boxed{F_1 \cap F_2 \text{ est un support.}}$

d) Supposons que F_1 et F_2 sont des supports de dimension $R_s(Y)$. Alors $F_1 \cap F_2$ est aussi un support de Y d'où sa dimension est supérieure à $R_s(Y)$, étant un sous espace vectoriel de F_1 sa dimension est alors égale à $R_s(Y)$. On a donc $\dim(F_1 \cap F_2) = dim(F_1) = dim(F_2)$ et $F_1 \cap F_2$ étant un sous espace vectoriel de F_1 et de F_2, on a les égalités $F_1 \cap F_2 = F_1$ et $F_1 \cap F_2 = F_2$ d'où $F_1 = F_2$.

$\boxed{\text{Il existe bien un unique } F \text{ de } \mathcal{S}(Y) \text{ tel que } \dim(F) = R_s(Y) \text{ .}}$

8.a) On sait d'après le cours qu'une variable aléatoire réelle combinaison linéaire de variables aléatoires admettant une variance admet une variance. De plus :

$$V\left(\sum_{i=1}^k u_i Y_i\right) = \text{Cov}\left(\sum_{i=1}^k u_i Y_i, \sum_{j=1}^k u_j Y_j\right)$$

Or par bilinéarité de la covariance $\text{Cov}\left(\sum_{i=1}^k u_i Y_i, \sum_{j=1}^k u_j Y_j\right) = \sum_{i=1}^k u_i \left(\sum_{j=1}^k \text{Cov}(Y_i, Y_j) u_j\right) = {}^t u \begin{pmatrix} \sum_{j=1}^k \text{Cov}(Y_1, Y_j) u_j \\ \vdots \\ \sum_{j=1}^k \text{Cov}(Y_k, Y_j) u_j \end{pmatrix}$

et $\begin{pmatrix} \sum_{j=1}^{k} \operatorname{Cov}(Y_1, Y_j) u_j \\ \vdots \\ \sum_{j=1}^{k} \operatorname{Cov}(Y_k, Y_j) u_j \end{pmatrix} = \mathcal{V}(X)u$ ce qui permet de conclure : $\boxed{V\left(\sum_{i=1}^{k} u_i Y_i\right) = {}^t u \mathcal{V}(X) u}$.

---

b) On sait que la matrice de variance-covariance $\mathcal{V}(Y)$ est symétrique réelle donc diagonalisable dans une base orthonormée. De plus ses valeurs propres sont positives.

En effet, si u est un vecteur propre de composantes $u_1, ..., u_k$, associé à la valeur propre λ, ${}^t u \mathcal{V}(X) u = \lambda \|u\|^2$ d'où $\lambda = \dfrac{V\left(\sum_{i=1}^{k} u_i Y_i\right)}{\|u\|^2}$. Les valeurs propres de la matrice de variance-covariance de Y sont bien positives.

Pour l'existence, il suffit de prendre une base de vecteurs propres de l'endomorphisme v canoniquement associées à $\mathcal{V}(X)$ et de réordonner si nécessaire cette base dans l'ordre décroissant des valeurs propres associées, en une nouvelle base. La matrice diagonale dans cette base de v vérifie la propriété demandée et est semblable à $\mathcal{V}(X)$.

Montrons l'unicité. Soit $(\lambda_1, ..., \lambda_k)$ et $(\mu_1, ..., \mu_k)$ deux vecteurs qui conviennent. Il sont associés à deux bases de vecteurs propres de v, $(e'_1, ..., e'_k)$ et $(e"_1, ..., e"_k)$. On note $E_\lambda(v)$ le sous espace propre de v associé à λ et $Sp(v)$ son spectre.

Supposons par l'absurde qu'il existe j tel que $\lambda_j > \mu_j$. Alors d'une part,

$$\dim\left(\bigoplus_{\lambda \in Sp(v)/\lambda \geq \lambda_j} E_\lambda(v)\right) \geq \dim(\operatorname{Vect}(e'_1, ..., e'_j)) \text{ i.e. } \dim\left(\bigoplus_{\lambda \in Sp(v)/\lambda \geq \lambda_j} E_\lambda(v)\right) \geq j$$

et d'autre part,

$$\dim\left(\bigoplus_{\lambda \in Sp(v)/\lambda < \lambda_j} E_\lambda(v)\right) \geq \dim(\operatorname{Vect}(e"_j, ..., e"_k)) \text{ i.e. } \dim\left(\bigoplus_{\lambda \in Sp(v)/\lambda < \lambda_j} E_\lambda(v)\right) \geq k - j + 1$$

Or $\dim\left(\bigoplus_{\lambda \in Sp(v)} E_\lambda(v)\right) = k$ d'où puisque les sous espaces propres sont en somme directe :

$$\dim\left(\bigoplus_{\lambda \in Sp(v)/\lambda \geq \lambda_j} E_\lambda(v)\right) + \dim\left(\bigoplus_{\lambda \in Sp(v)/\lambda < \lambda_j} E_\lambda(v)\right) = k$$

ce qui est en contradiction avec les deux inégalités qui précédent. Il y a donc bien $\boxed{\text{unicité d'un tel vecteur}}$.

---

c) On a l'existence de cette espérance par linéarité de l'espérance et le fait que les variables aléatoires Y_i admettent des variances. De plus $E(\|Y - E(Y)\|^2) = \sum_{i=1}^{k} V(Y_i)$. On obtient la somme des éléments diagonaux de $\mathcal{V}(X)$ qui est la même pour deux matrices semblables d'après les propriétés de la trace. D'où, on a bien : $\boxed{E(\|Y - E(Y)\|^2) = \sum_{i=1}^{k} \lambda_i.}$

Remarque. Il me semble que l'on dépasse ici le cadre du programme mais il y a peut-être une idée lumineuse qui évite le recours aux propriétés de la trace !

9.a) $\{\|Y(\omega)-\mathcal{E}(Y)-x\|^2; x\in\mathbb{R}\}$ est un sous ensemble non vide de $\mathbb{R}$, il est minoré par 0 et ainsi il possède une borne inférieure. Posons $y=Y(\omega)-\mathcal{E}(Y)$. On sait d'après le cours d'algèbre bilinéaire que le minimum de $\|y-x\|$ lorsque x décrit le sous espace vectoriel F est atteint lorsque x est le projeté orthogonal de y sur F. Notons alors z ce projeté orthogonal, $y-z$ appartient à $F^{\perp}$. On a donc en utilisant Pythagore : $\|y\|^2=\|y-z\|^2+\|z\|^2$. Or puisque $(f^{(1)},...,f^{(q)})$ est une base orthonormale, $z=\sum_{j=1}^{k}\langle y,f^{(j)}\rangle f^{(j)}$ d'où $\|z\|^2=\sum_{j=1}^{k}\langle y,f^{(j)}\rangle^2$. On a donc bien :

$$\boxed{\|Y(\omega)-\mathcal{E}(Y)\|^2=Q_F(\omega)+\sum_{j=1}^{k}\langle Y(\omega)-\mathcal{E}(Y),f^{(j)}\rangle^2}$$

---

b) Par définition, pour tout $j\in[\![1,k]\!]$, $\langle Y(\omega)-\mathcal{E}(Y),f^{(j)}\rangle^2={}^tf^{(j)}(Y(\omega)-\mathcal{E}(Y))\,{}^t(Y(\omega)-\mathcal{E}(Y))f^{(j)}$. D'après les résultats de la première question de la partie I, la matrice $(Y(\omega)-\mathcal{E}(Y))\,{}^t(Y(\omega)-\mathcal{E}(Y))$ a pour élément générique $(Y_i(\omega)-\mathcal{E}(Y_i))(Y_j(\omega)-\mathcal{E}(Y_j))$.

D'où $E\left((Y-\mathcal{E}(Y))\,{}^t(Y-\mathcal{E}(Y)\right)=\mathcal{V}(Y)$.

On remarque par linéarité de l'espérance, les $f^{(j)}$ ne dépendant pas de ω, que $E\left({}^tf^{(j)}(Y-\mathcal{E}(Y))\,{}^t(Y-\mathcal{E}(Y))f^{(j)}\right)$ existe et :

$$E\left({}^tf^{(j)}(Y-\mathcal{E}(Y))\,{}^t(Y-\mathcal{E}(Y))f^{(j)}\right)={}^tf^{(j)}E\left((Y-\mathcal{E}(Y))\,{}^t(Y-\mathcal{E}(Y)\right)f^{(j)}={}^tf^{(j)}\mathcal{V}(Y)f^{(j)}$$

Or pour tout ω, $Q_F(\omega)=\|Y(\omega)-\mathcal{E}(Y)\|^2-\sum_{j=1}^{q}\langle Y(\omega)-\mathcal{E}(Y),f^{(j)}\rangle^2$, d'où par linéarité de l'espérance $E(Q_F)$ existe et :

$$\boxed{E(Q_F)=E(\|Y-\mathcal{E}(Y)\|^2)-\sum_{i=1}^{q}{}^tf^{(j)}\mathcal{V}(Y)f^{(j)}=\sum_{i=1}^{k}\lambda_i-\sum_{i=1}^{q}{}^tf^{(j)}\mathcal{V}(Y)f^{(j)}}$$

---

c) Si $F=\mathcal{M}_{k,1}(\mathbb{R})$, $E(Q_F)$ vaut 0, d'où $\sum_{i=1}^{k}\lambda_i=\sum_{i=1}^{k}{}^tf^{(j)}\mathcal{V}(Y)f^{(j)}$.

10.a) Décomposons f dans une base $(e'_1,..,e'_k)$ associée au valeurs propres $\lambda_1,...,\lambda_k$. $f=\sum_{i=1}^{k}x_ie'_i$.

On a alors $\mathcal{V}(Y)f=\sum_{i=1}^{k}x_i\mathcal{V}(Y)e'_i=\sum_{i=1}^{k}x_i\lambda_ie'_i$ et puisque la base choisie est orthonormée, ${}^tf\mathcal{V}(Y)f=\sum_{i=1}^{k}\lambda_ix_i^2$.

Or pour tout i, $\lambda_ix_i^2\leq\lambda_1x_i^2$ d'où ${}^tf\mathcal{V}(Y)f\leq\lambda_1\sum_{i=1}^{k}x_i^2$ et puisque $\sum_{i=1}^{k}x_i^2=\|f\|^2$, on obtient l'inégalité : $\boxed{{}^tf\mathcal{V}(Y)f\leq\lambda_1\ .}$

---

b) Si F est une droite vectorielle, choisissons un vecteur normé f de cette droite. D'après le 9.b) :

$$E(Q_F)=\sum_{i=1}^{k}\lambda_i-{}^tf\mathcal{V}(Y)f$$

d'où $\boxed{E(Q_F) \geq \sum_{i=2}^{k} \lambda_i \text{ et cette valeur est atteinte lorsque } f \text{ est un vecteur propre associé à } \lambda_1}$.

---

c) Dans le cas proposé $\mathcal{V}(Y) = \frac{1}{k} I_k - \frac{1}{k^2} J$ où J est la matrice carrée d'ordre k dont tous les coefficients valent 1. On montre facilement que les valeurs propres de J sont 0 et k d'où l'on déduit que les valeurs propres de $\mathcal{V}(Y)$ sont $\frac{1}{k}$ et 0.

Dans ce cas $\sum_{i=1}^{k} \lambda_i = \frac{k-1}{k}$ et on a alors : $\boxed{E(Q_F) = \frac{k-2}{k}}$.

Les vecteurs propres associés à la valeur propre $\frac{1}{k}$ engendrent les droites qui réalisent le minimum. On voit que les vecteurs de ce sous espace propre sont les vecteurs $(x_1, ..., x_k)$ qui vérifient $x_1 + ... + x_k = 0$.

11.a) En notant q la dimension de F_0 et $(f^{(1)}, .., f^{(q)})$ une base orthonormée de ce sous espace vectoriel formé de vecteurs propres de $\mathcal{V}(Y)$, on a alors $\sum_{i=1}^{q} {}^t f^{(j)} \mathcal{V}(Y) f^{(j)} = \sum_{i=1}^{k} \lambda_i$ i.e. $\boxed{E(Q_{F_0}) = 0}$. Or la variable aléatoire Q_{F_0} est à valeur positives et d'espérance nulle est donc presque surement nulle. Mais $Q_{F_0}(\omega) = 0$ si et seulement si $Y(\omega) - E(Y)$ appartient à F_0. On peut donc en conclure que $\boxed{P(Y - E(Y) \in F_0) = 1}$, cqfd.

---

b) On peut, par exemple, compléter une base orthonormée de F en une base orthonormée de F_0. Alors tout vecteur que l'on rajoute satisfait aux conditions imposées.

---

c) $f^{(r)}$ est un vecteur non nul combinaison linéaire de vecteurs propres associées à des valeurs propres strictement positives, d'où en décomposant $f^{(r)}$ dans une base orthonormée de F_0 formées de vecteurs propres, on voit que ${}^t f^{(r)} \mathcal{V}(Y) f^{(r)} > 0$.

Soit alors s la dimension de F et $(f^{(1)}, .., f^{(s)})$ une base orthonormée de F. On pose $G = \text{Vect}(f^{(1)}, .., f^{(s)}, f^{(r)})$ alors $E(Q_G) \geq 0$ et $E(Q_G) = E(Q_F) - {}^t f^{(r)} \mathcal{V}(Y) f^{(r)}$, d'où $\boxed{E(Q_F) > 0}$.

---

d) F_0 est un support vectoriel de Y d'où $R_s(Y) \geq r$.

Supposons par l'absurde que $R_s(Y) < r$ et soit F un support vectoriel de Y de dimension $R_s(Y)$.

On a forcément $F \subset F_0$ car sinon, $F \cap F_0$ est un support vectoriel de Y de dimension strictement inférieure à $R_s(Y)$ ce qui est absurde.

Donc, $F \subset F_0$ et $F \neq F_0$, d'où $E(Q_F) \neq 0$. Ceci prouve que Q_F n'est pas presque surement nulle, ce qui est toujours le cas lorsque F est un support vectoriel de Y. On a donc établi une contradiction.

Finalement, on peut affirmer que $\boxed{R_s(Y) = r}$.

12.a) Raisonnons par l'absurde.

Supposons que $R_s(T) < k$. Alors il existe un sous espace vectoriel F de $\mathbb{R}^k$ de dimension strictement inférieure à k tel que $P(T - E(T) \in F) = 1$.

Soit alors $(v_1, ..., v_k)$ un vecteur non nul orthogonal à F :