

REVUE DE STATISTIQUE APPLIQUÉE

YVES ESCOUFIER

Calculs de probabilités par une méthode de Monte-Carlo pour une variable p -normale

Revue de statistique appliquée, tome 15, n° 4 (1967), p. 5-15

http://www.numdam.org/item?id=RSA_1967__15_4_5_0

© Société française de statistique, 1967, tous droits réservés.

L'accès aux archives de la revue « *Revue de statistique appliquée* » (<http://www.sfds.asso.fr/publicat/rsa.htm>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

**CALCULS DE PROBABILITÉS
PAR UNE MÉTHODE DE MONTE-CARLO
POUR UNE VARIABLE p -NORMALE**

Yves ESCOUFIER

M. A. à l'Institut Universitaire de Technologie de Montpellier

Soit $X = (X_1, X_2, \dots, X_p)$ une variable vectorielle distribuée selon la loi $\mathcal{N}(M, A)$. Notre but est de calculer $\text{Prob}(X \in \mathcal{D})$ où \mathcal{D} est un domaine défini par les inégalités de la forme

$$\sum_{i=1}^p a_{ki} X_i < C_k$$

en nombre quelconque. Le problème est posé à partir de deux exemples puis développé de façon théorique.

Les solutions numériques des deux exemples sont fournies.

I - Considérons un bassin fluvial équipé sur la rivière principale et sur ses affluents de stations de jaugeage où sont relevées, en particulier, les valeurs des étiages annuels, c'est-à-dire des plus faibles débits de l'année. Il est important dès que l'on veut utiliser l'eau disponible dans le bassin, à des fins industrielles ou agricoles, de pouvoir donner la probabilité de la concomitance de plusieurs étiages, en ces différentes stations.

Intéressons-nous maintenant à une seule station de jaugeage. Nous y notons les débits mensuels moyens et nous pouvons pour différentes applications avoir besoin de la probabilité de l'évènement suivant : les 12 débits mensuels moyens sont inférieurs à une valeur donnée.

L'intérêt de ces calculs, (mais aussi leur difficulté) vient de ce que les variables aléatoires de ces exemples ne sont certainement pas indépendantes : il en résulte que des déficits en eau peuvent apparaître avec des probabilités non négligeables en plusieurs points ou sur plusieurs mois successifs.

Nous allons exposer, comment sous certaines hypothèses, nous avons résolu ces deux problèmes.

II - La première parenté de ces deux problèmes, réside dans le fait que la variable aléatoire étudiée est vectorielle.

Dans le premier problème, en retenant par exemple 8 stations, nous devons trouver la probabilité avec laquelle la variable $X = (X_1, X_2, \dots, X_8)$ appartient à un domaine défini par des inégalités de la forme $X_i < C_i$ ($i = 1, 2, \dots, 8$) où X_i est la valeur de l'étiage à la station i et C_i une valeur que peut prendre cet étiage.

Dans le second problème, si nous appelons X_i ($i = 1, 2, \dots, 12$) le débit mensuel moyen du mois i , nous avons à calculer la probabilité avec laquelle le vecteur $X = (X_1, X_2, \dots, X_{12})$ appartient au domaine défini par des inégalités de la forme $X_i < C$ où C est une valeur donnée pour le débit mensuel moyen.

La deuxième parenté de ces deux problèmes, vient de ce qu'il est possible de passer du vecteur X des observations à un vecteur \hat{X} dont la loi de distribution soit une loi p-Normale. En effet, des études précédentes ont permis de justifier l'hypothèse selon laquelle les variables "étiage annuel en une station" et "débit moyen pour un mois donné" sont distribuées suivant une loi de Galton ou loi lognormale. Nous ferons des hypothèses plus fortes en disant que le vecteur constitué par les logarithmes des huit étiages annuels suit une loi normale de dimension huit et que le vecteur constitué par les logarithmes des douze débits mensuels moyens, suit une loi normale de dimension douze.

Le principe de la méthode est alors simple : à partir des données disponibles nous estimons la matrice des covariances A et le vecteur moyen M de la distribution. En générant ensuite par une méthode appropriée des vecteurs distribués selon la loi $\mathcal{N}(M, A)$, nous obtenons une estimation de la probabilité attachée aux domaines qui nous intéressent. Une remarque sur cette manière de faire permettra d'obtenir le résultat de façon plus économique.

III - RAPPELS

Soit $f(X_1, X_2, \dots, X_p)$ la fonction de densité du vecteur colonne X . Si f est la fonction de densité d'une loi gaussienne :

$$f(X_1, X_2, \dots, X_p) = \frac{1}{(2\pi)^{\frac{p}{2}} \sqrt{|A|}} \exp \left[-\frac{1}{2} (X - M)^t A^{-1} (X - M) \right]$$

où M est le vecteur espérance mathématique de X , A la matrice des covariances, $|A|$ le déterminant de A . Nous supposons dans ce qui suit que la distribution de X n'est pas singulière.

Soit V une matrice $p \times p$ orthogonale. On sait que tout vecteur Y , tel que $Y = VX$, suit une loi p-normale de moyenne VM et de matrice des covariances VAV^t . (voir par exemple [2]). Toute matrice symétrique étant congrue à une matrice diagonale, V peut être choisie de telle sorte que VAV^t soit diagonale. Alors Y est un vecteur p-normal tel que ses composantes soient deux à deux indépendantes. Nous choisirons la matrice V de telle sorte que Y_1 soit la combinaison linéaire normée des X_i qui ait la plus grande variance et Y_k ($k = 2, \dots, p$) la combinaison linéaire normée des X_i qui ait la plus grande variance sous réserve d'être orthogonale à tous les Y_j ($j < k$). Les Y_k qui satisfont à ses conditions sont les vecteurs propres de A et Y_k a pour variance λ_k , valeur propre de rang k dans la suite des valeurs propres de A supposées rangées par ordre décroissant. Les Y_k sont couramment appelés composantes principales [1].

IV - PRINCIPE DE LA METHODE DE CALCUL

Soit à calculer $\text{Prob}(X \in \mathcal{O}) = \int_{\mathcal{O}} \dots \int f(X_1, X_2, \dots, X_p) dX_1 \dots dX_p$ où \mathcal{O} est un domaine défini par m inégalités

$$\sum_{i=1}^p a_{ki} X_i \leq C_k \quad (k = 1, \dots, m)$$

m étant quelconque et $f(X_1, \dots, X_p)$ étant la fonction de densité de la loi $\mathcal{P}(M, A)$.

Nous nous proposons de générer n vecteurs X^i selon la loi $\mathcal{P}(M, A)$. Soit P_i le point image de X^i dans l'espace de dimension p et g_i une variable prenant la valeur 1 si $P_i \notin \mathcal{O}$, la valeur 0 si $P_i \in \mathcal{O}$. $g = \sum_{i=1}^p g_i$ est le nombre de points P éléments de \mathcal{O} et $\frac{g}{n}$ est connu pour être un estimateur non biaisé de $\text{Prob}(X \in \mathcal{O})$, de variance inférieure à $\frac{1}{4n}$ [3].

Soit Z un vecteur dont les composantes sont indépendantes, normales, centrées et réduites. Un tel vecteur peut-être généré par un procédé numérique approprié ou par consultation de tables.

Si nous conservons à Y, V, M et λ_i leur signification du paragraphe précédent et si nous appelons B la matrice diagonale telle que $b_{ii} = \sqrt{\lambda_i}$, alors :

$$Y = BZ + VM.$$

Ceci permet de générer facilement Y et donc $X = V^t Y$.

En résumé, pour calculer $\frac{g}{n}$, nous avons donc à générer n vecteurs Z^i , à calculer les Y^i et les X^i correspondants, à donner à g_i la valeur 1 si les m inégalités qui définissent \mathcal{O} sont vérifiées, et la valeur 0 si une au moins des inégalités n'est pas vérifiée.

V₁ - METHODE DE CALCUL SIMPLIFIEE

La remarque suivante permet de gagner du temps de calcul sans pour autant perdre trop de précisions.

Nous avons utilisé la propriété d'indépendance des Y_k ; utilisons maintenant le fait que les variances des premiers Y_k sont fortes, alors que celles des derniers sont faibles.

V₁ - Supposons que nous ayons à calculer la probabilité attachée à un domaine \mathcal{O} défini par une seule inégalité

$$\sum_{i=1}^p a_i X_i \leq C$$

Appelons v_{ij} les éléments de la matrice V ; de $X = V^t Y$ nous tirons :

$$X_i = \sum_{j=1}^p v_{ji} Y_j$$

et donc
$$Z = \sum_{i=1}^p a_i X_i = \sum_{i=1}^p \sum_{j=1}^p a_i v_{ji} Y_j$$

On a immédiatement :

$$\text{moy}(Z) = m = \sum_{i=1}^p \sum_{j=1}^p a_i v_{ji} \bar{Y}_j$$

$$\text{Var}(Z) = \sigma^2 = \sum_{i=1}^p \sum_{j=1}^p a_i^2 v_{ji}^2 \lambda_j$$

Supposons que les $(p-l)$ dernières composantes principales aient une variance très petite et calculons l'erreur commise, en admettant que ces composantes restent égales à leur moyenne.

Au lieu de générer les X_i , nous générons :

$$X_i^* = \sum_{j=1}^l v_{ji} Y_j + \sum_{j=l+1}^p v_{ji} \bar{Y}_j$$

On constate aisément que :

$$\text{moy}(Z^*) = m^* = m$$

$$\text{Var}(Z^*) = \sigma^{*2} = \sum_{i=1}^p \sum_{j=1}^l a_i^2 v_{ji}^2 \lambda_j = \sigma^2 - \sum_{i=1}^p \sum_{j=l+1}^p a_i^2 v_{ji}^2 \lambda_j$$

Au lieu d'un estimateur non biaisé de :

$$I_1 = \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^c \exp \left[-\frac{1}{2} \left(\frac{Z-m}{\sigma} \right)^2 \right] dZ$$

Nous allons donc calculer un estimateur non biaisé de :

$$I_2 = \frac{1}{\sigma^* \sqrt{2\pi}} \int_{-\infty}^c \exp \left[-\frac{1}{2} \left(\frac{Z-m}{\sigma^*} \right)^2 \right] dZ$$

Si la différence entre I_2 et I_1 est négligeable, nous pourrions considérer l'estimateur de I_2 comme une estimation de I_1 .

Posons :

$$\eta = \frac{\sigma^2 - \sigma^{*2}}{\sigma^2}$$

alors :

$$I_2 = \frac{1}{\sqrt{1-\eta}} \times \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^c \exp \left[-\frac{1}{2} \frac{(Z-m)^2}{\sigma^2(1-\eta)} \right] dZ$$

Or :

$$0 < \eta < 1 \text{ implique } \exp \left[-\frac{1}{2} \frac{(Z-m)^2}{\sigma^2(1-\eta)} \right] < \exp \left[-\frac{1}{2} \left(\frac{Z-m}{\sigma} \right)^2 \right]$$

d'où

$$I_2 < \frac{I_1}{\sqrt{1-\eta}} \quad \text{et} \quad I_2 - I_1 < I_1 \left(\frac{1}{\sqrt{1-\eta}} - 1 \right)$$

En raisonnant sur le complémentaire du domaine qui nous intéresse, on obtiendrait de même :

$$(1-I_2) - (1-I_1) < (1-I_1) \left(\frac{1}{\sqrt{1-\eta}} - 1 \right)$$

si bien, qu'en définitive :

$$V_{1-1} - (1-I_1) \left(1 - \frac{1}{\sqrt{1-\eta}} \right) < I_2 - I_1 < I_1 \left(\frac{1}{\sqrt{1-\eta}} - 1 \right)$$

V_2 - Pour généraliser la démonstration au cas d'un domaine défini par un nombre quelconque d'inégalités, il faudrait comparer à $\text{Prob}(X \in \mathcal{D})$ la quantité $\text{Prob}(X^* \in \mathcal{D})$ où X^* est le vecteur de composantes X_1^* . Mais par construction X^* suit une loi normale singulière et l'extension ne peut pas être faite directement.

Considérons le domaine E défini par

$$\sum_{i=1}^p (Y_i - \bar{Y}_i)^2 < R^2$$

Soit \hat{E} la projection de E dans l'espace de dimension p-m rapporté aux axes $Y_{m+1}, Y_{m+2}, \dots, Y_p$ et \hat{Y} la projection de Y dans cet espace. Il découle des définitions que nous venons de poser, que :

$$\text{Prob}(Y \in E) = \text{Prob}(\hat{Y} \in \hat{E})$$

Appelons $\lambda_{m+1}, \lambda_{m+2}, \dots, \lambda_p$, les valeurs propres associées aux composantes $Y_{m+1}, Y_{m+2}, \dots, Y_p$ (c'est-à-dire les variances de ces composantes). En utilisant un résultat de Guelfand et Vilenkin sur la mesure de la boule de rayon R dans un espace vectoriel topologique muni d'une mesure gaussienne [4], on obtient :

$$V_{2-1} - \text{Prob}(\hat{Y} \in \hat{E}) \geq 1 - \frac{\lambda_{m+1} + \lambda_{m+2} + \dots + \lambda_p}{R^2}$$

Proposons-nous alors de calculer $I_2 = \text{Prob}(X \in \mathcal{D} / Y \in E)$ en fonction de $I_1 = \text{Prob}(X \in \mathcal{D})$

$$\text{Prob}(X \in \mathcal{D} / Y \in E) \times \text{Prob}(Y \in E) = \text{Prob}(X \in \mathcal{D} \quad \text{et} \quad Y \in E) < \text{Prob}(X \in \mathcal{D})$$

d'où :

$$I_2 < \frac{I_1}{\text{Prob}(Y \in E)} \quad \text{et} \quad I_2 - I_1 < I_1 \left(\frac{1}{\text{Prob}(Y \in E)} - 1 \right)$$

En utilisant l'inégalité V_{2-1} , on a :

$$I_2 - I_1 < I_1 \left(\frac{1}{1 - \frac{\lambda_{m+1} + \dots + \lambda_p}{R^2}} - 1 \right)$$

et si, comme précédemment, nous raisonnons sur le complémentaire de \mathcal{O} nous avons :

$$V_{2-2} - (1 - I_1) \left(1 - \frac{1}{1 - \frac{\lambda_{m+1} + \dots + \lambda_p}{R^2}} \right) < I_2 - I_1 < I_1 \left(\frac{1}{1 - \frac{\lambda_{m+1} + \dots + \lambda_p}{R^2}} - 1 \right)$$

Dans l'application sur ordinateur, nous avons imposé aux composantes principales de faible variance de rester égales à leur valeur moyenne. Ceci ne revient pas à considérer que R^2 est nul mais que les quantités $|Y_1 - \bar{Y}_1|$ sont inférieures au dernier chiffre significatif de \bar{Y}_1 .

VI - Reprenons les deux problèmes proposés au paragraphe I.

La résolution du premier, c'est-à-dire le calcul de la probabilité de la concomitance de huit étiages donnés, sera l'occasion de comparer les résultats obtenus suivant qu'on utilise ou non la simplification proposée.

Le vecteur moyen M et la matrice A des covariances sont estimés à partir des données disponibles (tableau 1). Rappelons que nous travaillons sur les logarithmes népériens des valeurs observées.

TABLEAU I

Les données sont les logarithmes des étiages

Demi-matrice des variances et covariances estimées à partir de l'échantillon

| | | | | | | | | | |
|---|--|-------|-------------|-------------------|-------------------------|-------------------------------|-------------------------------------|---|---|
| A = | <table style="border-collapse: collapse; width: 100%; text-align: center;"> <tr><td>0,327</td></tr> <tr><td>0,228 0,373</td></tr> <tr><td>0,220 0,283 0,693</td></tr> <tr><td>0,245 0,281 0,507 0,648</td></tr> <tr><td>0,227 0,276 0,535 0,574 0,633</td></tr> <tr><td>0,172 0,208 0,411 0,439 0,424 0,393</td></tr> <tr><td>0,172 0,179 0,313 0,288 0,327 0,250 0,251</td></tr> <tr><td>0,079 0,115 0,260 0,242 0,279 0,205 0,187 0,201</td></tr> </table> | 0,327 | 0,228 0,373 | 0,220 0,283 0,693 | 0,245 0,281 0,507 0,648 | 0,227 0,276 0,535 0,574 0,633 | 0,172 0,208 0,411 0,439 0,424 0,393 | 0,172 0,179 0,313 0,288 0,327 0,250 0,251 | 0,079 0,115 0,260 0,242 0,279 0,205 0,187 0,201 |
| 0,327 | | | | | | | | | |
| 0,228 0,373 | | | | | | | | | |
| 0,220 0,283 0,693 | | | | | | | | | |
| 0,245 0,281 0,507 0,648 | | | | | | | | | |
| 0,227 0,276 0,535 0,574 0,633 | | | | | | | | | |
| 0,172 0,208 0,411 0,439 0,424 0,393 | | | | | | | | | |
| 0,172 0,179 0,313 0,288 0,327 0,250 0,251 | | | | | | | | | |
| 0,079 0,115 0,260 0,242 0,279 0,205 0,187 0,201 | | | | | | | | | |

Vecteur moyenne estimé à partir de l'échantillon

| | | | | | | | | | |
|---------|--|-------|---------|---------|-------|-------|-------|-------|---------|
| M = | <table style="border-collapse: collapse; width: 100%; text-align: center;"> <tr><td>1,010</td></tr> <tr><td>- 0,807</td></tr> <tr><td>- 1,576</td></tr> <tr><td>0,385</td></tr> <tr><td>0,638</td></tr> <tr><td>0,973</td></tr> <tr><td>0,035</td></tr> <tr><td>- 0,346</td></tr> </table> | 1,010 | - 0,807 | - 1,576 | 0,385 | 0,638 | 0,973 | 0,035 | - 0,346 |
| 1,010 | | | | | | | | | |
| - 0,807 | | | | | | | | | |
| - 1,576 | | | | | | | | | |
| 0,385 | | | | | | | | | |
| 0,638 | | | | | | | | | |
| 0,973 | | | | | | | | | |
| 0,035 | | | | | | | | | |
| - 0,346 | | | | | | | | | |

Le tableau 2 fournit les valeurs propres et les vecteurs propres de la matrice A.

Tableau 2

Valeurs propres et vecteurs propres de la matrice A

| | λ_1 | λ_2 | λ_3 | λ_4 | λ_5 | λ_6 | λ_7 | λ_8 |
|------------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| Valeurs propres | 2,630 | 0,325 | 0,183 | 0,129 | 0,119 | 0,072 | 0,040 | 0,021 |
| | Y_1 | Y_2 | Y_3 | Y_4 | Y_5 | Y_6 | Y_7 | Y_8 |
| Vecteurs propres | 0,218 | -0,671 | 0,011 | 0,278 | 0,591 | -0,049 | 0,025 | 0,244 |
| | 0,259 | -0,634 | -0,163 | -0,278 | -0,653 | -0,032 | -0,002 | 0,002 |
| | 0,460 | 0,185 | -0,756 | -0,282 | 0,309 | -0,051 | 0,069 | -0,005 |
| | 0,460 | 0,106 | 0,545 | -0,396 | 0,110 | -0,009 | 0,506 | -0,234 |
| | 0,466 | 0,192 | 0,224 | 0,115 | -0,135 | -0,602 | -0,532 | 0,140 |
| | 0,354 | 0,148 | 0,157 | 0,005 | -0,004 | 0,786 | -0,399 | 0,226 |
| | 0,270 | -0,349 | -0,099 | 0,585 | -0,080 | 0,115 | 0,002 | -0,744 |
| | 0,218 | 0,024 | -0,083 | 0,503 | -0,304 | 0,005 | 0,545 | 0,511 |

Le programme de calcul permet d'obtenir simultanément les probabilités attachées à cinq domaines définis par des inégalités de la forme $X_1 < C_1$. Le tableau 3 fournit les valeurs des constantes C_1 définissant cinq domaines particuliers.

Tableau 3

Valeurs des constantes définissant les différents domaines

| | C_1 | C_2 | C_3 | C_4 | C_5 | C_6 | C_7 | C_8 |
|--------------------------|-------|-------|-------|-------|-------|-------|-------|-------|
| 1 ^{er} Domaine | 1,4 | 0,22 | 0,07 | 0,59 | 0,93 | 1,3 | 0,57 | 0,38 |
| 2 ^{ème} Domaine | 1,92 | 0,31 | 0,12 | 0,92 | 1,3 | 1,92 | 0,76 | 0,5 |
| 3 ^{ème} Domaine | 2,75 | 0,45 | 0,21 | 1,52 | 1,9 | 2,8 | 1,05 | 0,68 |
| 4 ^{ème} Domaine | 3,9 | 0,72 | 0,32 | 2,5 | 2,8 | 4,15 | 1,5 | 0,92 |
| 5 ^{ème} Domaine | 5,4 | 1,3 | 0,60 | 3,9 | 4,5 | 6 | 2 | 1,2 |

Les calculs ont été faits pour chacune des quatre méthodes suivantes :

- 1/ en générant les huit composantes.
- 2/ en gardant la huitième composante égale à sa valeur moyenne.
- 3/ en gardant la huitième et la septième composantes égales à leurs valeurs moyennes.
- 4/ en gardant les trois dernières composantes (8, 7, 6) égales à leurs valeurs moyennes.

Chaque fois nous avons généré 625 vecteurs aléatoires à partir des tables fournies par [5]. On trouvera dans le tableau 4 les temps de calcul nécessaires à la détermination des probabilités attachées aux cinq

domaines pour chacune des méthodes et dans le tableau 5 les résultats obtenus pour chacun des domaines. Notons que les calculs ont été effectués à partir de programmes écrits en FORTRAN I sur l'ordinateur I.B.M. 1620 à cartes, de la faculté des Sciences de Montpellier.

Tableau 4

Durée du calcul

| | | |
|-------------------|---|-------|
| Première méthode | | 40 mm |
| Deuxième méthode | la 8° composante est prise à sa valeur moyenne. | 36 mm |
| Troisième méthode | la 7° et la 8° composante sont prises égales à leur valeur moyenne. | 32 mm |
| Quatrième méthode | les 6°, 7° et 8° composantes sont prises égales à leur valeur moyenne | 28 mm |

Tableau 5

| | 1 ^{ère} méthode | 2 ^{ème} méthode | 3 ^{ème} méthode | 4 ^{ème} méthode |
|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 1 ^{er} Domaine | 0,0080 | 0,0064 | 0,0064 | 0,0080 |
| 2 ^{ème} Domaine | 0,0384 | 0,0416 | 0,0416 | 0,0432 |
| 3 ^{ème} Domaine | 0,1696 | 0,1680 | 0,1936 | 0,1952 |
| 4 ^{ème} Domaine | 0,4336 | 0,4336 | 0,4464 | 0,4384 |
| 5 ^{ème} Domaine | 0,7120 | 0,6976 | 0,7040 | 0,7200 |

Sur la base du test χ^2 pour comparer deux probabilités [6] nous avons testé à partir des fréquences données par les différentes méthodes, l'égalité des probabilités attachées à chacun des domaines. En aucun cas l'hypothèse de l'égalité des probabilités n'a été rejetée. Certes un plus grand nombre de vecteurs aléatoires générés conduirait peut-être à séparer ces probabilités. On gardera cependant de cette application que les différentes méthodes donnent des résultats comparables si la précision demandée est de l'ordre de 1 %.

VII - Abordons le calcul de la probabilité avec laquelle les douze débits mensuels moyens sont inférieurs à une valeur donnée. Remarquons que le problème revient à la détermination de la loi du maximum annuel des douze débits moyens. Les tableaux 6 et 7 contiennent la matrice des corrélations, les écarts-types et le vecteur moyen estimés à partir des logarithmes népériens des données disponibles. A condition de réduire au préalable les variables, les calculs peuvent être menés avec la matrice des corrélations exactement de la même manière qu'avec la matrice des covariances. 500 vecteurs aléatoires ont été générés et les onzième et douzième composantes principales ont été conservées égales à leur valeur moyenne.

Tableau 6

| | | | | |
|-----|-------|-------------|-----------|-------|
| M = | 1.350 | Ecart types | Janvier | 0,229 |
| | 1.360 | | Février | 0,232 |
| | 1.326 | | Mars | 0,238 |
| | 1.262 | | Avril | 0,285 |
| | 1.177 | | Mai | 0,309 |
| | 1.017 | | Juin | 0,299 |
| | 0.769 | | Juillet | 0,344 |
| | 0.648 | | Août | 0,331 |
| | 0.743 | | Septembre | 0,325 |
| | 0.901 | | Octobre | 0,369 |
| | 1.166 | | Novembre | 0,325 |
| | 1.277 | | Décembre | 0,288 |

Tableau 7

Matrice des corrélations

| | | | | | | | | | | | | | | | | | | | | |
|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|---|--|--|--|--|--|--|--|--|--|
| 1 | | | | | | | | | | | | | | | | | | | | |
| 0,3207 | 1 | | | | | | | | | | | | | | | | | | | |
| 0,1355 | 0,3007 | 1 | | | | | | | | | | | | | | | | | | |
| 0,1059 | 0,0750 | 0,4843 | 1 | | | | | | | | | | | | | | | | | |
| 0,0925 | 0,0929 | 0,2386 | 0,7401 | 1 | | | | | | | | | | | | | | | | |
| 0,0364 | 0,0253 | 0,1474 | 0,4971 | 0,7500 | 1 | | | | | | | | | | | | | | | |
| 0,1668 | 0,0118 | 0,3173 | 0,4401 | 0,5646 | 0,6858 | 1 | | | | | | | | | | | | | | |
| 0,2186 | 0,0970 | 0,2919 | 0,3903 | 0,4295 | 0,4574 | 0,7295 | 1 | | | | | | | | | | | | | |
| 0,0000 | 0,0024 | 0,3196 | 0,4055 | 0,4233 | 0,3228 | 0,5763 | 0,6997 | 1 | | | | | | | | | | | | |
| 0,0710 | 0,0494 | 0,3653 | 0,3606 | 0,2878 | 0,2054 | 0,4623 | 0,4851 | 0,7277 | 1 | | | | | | | | | | | |
| 0,1512 | 0,1516 | 0,3845 | 0,4453 | 0,3794 | 0,1597 | 0,2364 | 0,2903 | 0,5095 | 0,7251 | 1 | | | | | | | | | | |
| 0,0496 | 0,0845 | 0,3308 | 0,4241 | 0,3815 | 0,0128 | 0,0926 | 0,0670 | 0,3654 | 0,5255 | 0,6967 | 1 | | | | | | | | | |

Les résultats obtenus sont consignés dans le tableau 8, sous forme des probabilités attachées aux domaines définis par les inégalités $X_i < C$ ($i = 1, 2, \dots, 12$) pour des valeurs de C allant de $\text{Log } C = 1, 2$ à $\text{Log } C = 2, 15$ avec un pas de 0,05 pour le logarithme.

On trouvera dans la figure 9 la représentation de la courbe

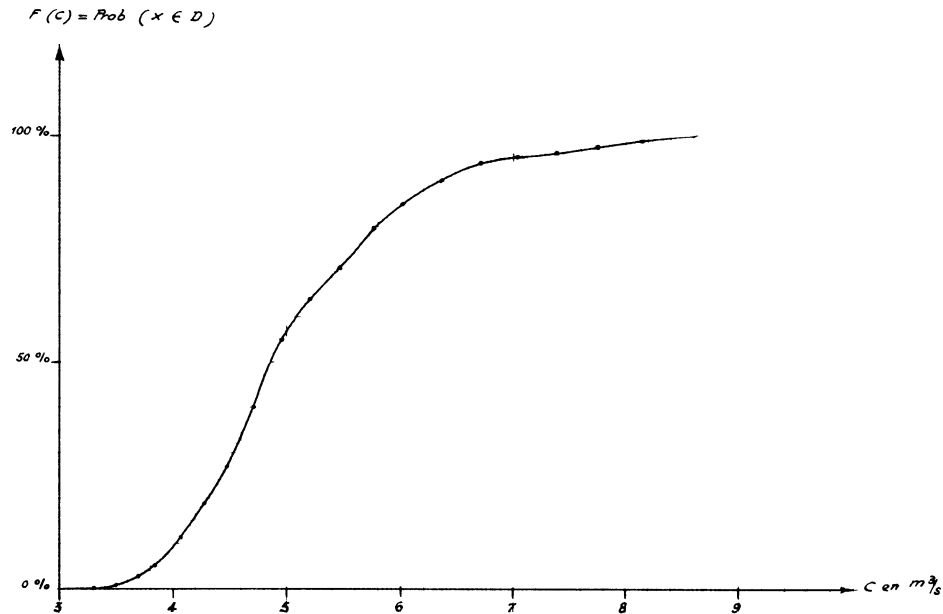
$$f(C) = \text{Prob}(X \in \mathcal{D}) \quad \text{avec} \quad \mathcal{D} = \{X : X_i < C, i = 1, 2, \dots, 12\}.$$

Je tiens à remercier Monsieur BERNIER, C.R.E.C. de l'E.D.F. à Chatou qui m'a fourni les données relatives au second problème et s'est intéressé à mon travail ; Monsieur le Professeur MARTINEAU, qui a toujours été disponible et prêt à répondre à mes questions ; tous

les membres du bureau hydrométéorologique de l'E.D.F. à Montpellier et en particulier Monsieur LOBERT, qui a bien voulu lire ces pages et qui m'a fait des remarques très utiles.

Tableau 8

| Log C | C en m ³ /s | Probabilité attachée au domaine $X_1 < C(i = 1, 2, \dots, 12)$ |
|-------|------------------------|--|
| 1,2 | 3,32 | 0,002 |
| 1,25 | 3,49 | 0,014 |
| 1,3 | 3,67 | 0,030 |
| 1,35 | 3,86 | 0,052 |
| 1,4 | 4,06 | 0,116 |
| 1,45 | 4,26 | 0,190 |
| 1,5 | 4,48 | 0,268 |
| 1,55 | 4,71 | 0,398 |
| 1,6 | 4,95 | 0,552 |
| 1,65 | 5,21 | 0,642 |
| 1,7 | 5,47 | 0,714 |
| 1,75 | 5,75 | 0,804 |
| 1,8 | 6,05 | 0,852 |
| 1,85 | 6,36 | 0,898 |
| 1,90 | 6,69 | 0,942 |
| 1,95 | 7,03 | 0,952 |
| 2,0 | 7,39 | 0,964 |
| 2,05 | 7,77 | 0,976 |
| 2,10 | 8,17 | 0,988 |
| 2,15 | 8,58 | 0,994 |



BIBLIOGRAPHIE

- [1] T. W. ANDERSON - An Introduction to Multivariate statistical Analysis. New-York, John Wiley and sons, 1958.
- [2] H. CRAMER - Mathematical methods of statistics. Princeton University Press, 1946.
- [3] J. M. HAMMERSLEY and D. C. HANDSCOMB - Monte-Carlo Methods - Menhuen's monographs on applied probability and statistics, 1964.
- [4] L. M. GUELFAND and N. Y. VILENKIN - Les distributions. Tome 4 : applications de l'analyse harmonique. Dunod 1967, p. 332.
- [5] Handbook of tables for probability and statistics - William H. Beyer. The Chemical Rubber Co, 1966.
- [6] B. L. VAN DER WEARDEN - Statique mathématique. Dunod, 1967, p. 38.