

Zeitschrift: Archives des sciences [1948-1980]
Herausgeber: Société de Physique et d'Histoire Naturelle de Genève
Band: 15 (1962)
Heft: 1

Artikel: Corrélation statistique de deux caractères relatifs à la pluie
Autor: Bouvier, P.
DOI: <https://doi.org/10.5169/seals-738655>

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. [Siehe Rechtliche Hinweise.](#)

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. [Voir Informations légales.](#)

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. [See Legal notice.](#)

Download PDF: 02.02.2025

ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>

CORRÉLATION STATISTIQUE DE DEUX CARACTÈRES RELATIFS A LA PLUIE

par P. BOUVIER

Nous avons examiné la corrélation statistique présumée entre la *hauteur d'eau* de pluie tombée en une station donnée et la *durée de précipitation* en cette même station.

Le matériel d'observation choisi était constitué par les mesures mensuelles relevées à l'Observatoire de Genève de 1911 à 1960; nous disposons ainsi de 600 valeurs pour chacune des deux variables x , hauteur d'eau en mm et y durée de précipitation en h.

Rangeons ces valeurs dans des intervalles partiels de largeur respective $\Delta x = 10$ mm et $\Delta y = 10$ h. et notons n_{ik} le nombre des valeurs y comprises entre y_{i-1} et y_i , lorsque x est compris entre x_{k-1} et x_k . Nous aurons ainsi comme moyenne liée à (x_{k-1}, x_k) :

$$\bar{y}_k = \frac{\sum_i n_{ik} y_i}{\sum_i n_{ik}} = Y_k$$

De même la moyenne des x liée à l'intervalle (y_{i-1}, y_i) vaudra

$$\bar{x}_i = \frac{\sum_k n_{ik} x_k}{\sum_k n_{ik}} = X_i$$

Ces grandeurs moyennes \bar{x}_i et \bar{y}_k dépendent de la largeur des intervalles Δx et Δy , mais cette dernière a été choisie à la fois assez réduite et assez riche en points représentatifs des mesures pour que l'on puisse considérer \bar{x}_i et \bar{y}_k comme de bonnes estimations des moyennes liées $\bar{y}(x)$ et $\bar{x}(y)$ respectivement.

Nous obtenons donc une succession de points dans le plan (x, y) , donnant l'allure des courbes de régression (v. fig. 1).

La courbe $\bar{y} = Y(x)$ se révèle presque linéaire et il en va de même de l'autre courbe de régression $\bar{x} = X(y)$ tant que x ne dépasse pas

110 mm. De toute façon, ces courbes n'ont plus de sens pour les valeurs extrêmement faibles ($x < 10$ mm) ou notoirement élevées ($x > 120$ mm) vu le nombre restreint des mesures relatives à ces cas-là.

Calculons ensuite les grandeurs

$$s_k^2 = \frac{1}{n_k - 2} \sum_i n_{ik} (y_i - \bar{y}_k)^2 \quad (n_k = \sum_i n_{ik})$$

que nous adoptons comme estimations des variabilités liées $\sigma(y|x)$; nous trouvons des valeurs qui, au début surtout, croissent avec x , quoique de manière irrégulière; la dépendance stochastique de y avec x apparaît *hétéroscédastique*. Les coefficients a , b de la droite de régression

$$Y(x) = a + b(x - \xi)$$

sont donc à déterminer de manière à rendre minimum la fonction

$$S = \sum_k p_k (y_k - Y)^2$$

où les poids p_k sont inversement proportionnels aux variances s_k^2

$$p_k s_k^2 = \text{const.}$$

Choisisant l'abscisse arbitraire ξ égale à

$$\xi = \frac{\sum_k p_k x_k}{\sum_k p_k}$$

les conditions d'extremum

$$\frac{\partial S}{\partial a} = 0, \quad \frac{\partial S}{\partial b} = 0$$

nous fournissent alors

$$a = \frac{\sum_k p_k y_k}{\sum_k p_k} = \eta, \quad b = \frac{\sum_k p_k (x_k - \xi) (y_k - \eta)}{\sum_k p_k (x_k - \xi)^2}$$

Dans ces formules, les sommes devraient en principe être effectuées sur tous les couples (au nombre de 600) de mesures x_k, y_k . En procédant de manière plus expéditive à partir de la subdivision en inter-

valles partiels, c'est-à-dire en prenant pour x_k les centres des intervalles successifs de cette subdivision ($x_1 = 5, x_2 = 15, x_3 = 25$ etc.) et en remplaçant les y_k par les moyennes liées correspondantes Y_k , M^{lle} Y. Peccoux obtient les valeurs suivantes

$$\begin{aligned}\xi &= 31,3 \text{ mm} \\ a = \eta &= 35,8 \text{ h} \\ b &= 0,72\end{aligned}$$

Ces nombres sont évalués à environ 4 ou 5% près; l'on ne peut guère exiger plus car d'une part la précision des durées de précipitation est parfois discutable et d'autre part les coefficients a,b sont sensibles aux choix des p_k , lesquels dépendent des nombres $n_k = \sum_i n_{ik}$ de mesures à k donné, nombres souvent assez différents d'une valeur de k à l'autre.

Adoptons finalement pour équations des droites de régression

$$Y = 0,7 x + 15 \quad (10 \leq x \leq 150)$$

$$X = y + 5 \quad (15 \leq y \leq 110)$$

La légère inclinaison mutuelle de ces droites traduit un coefficient de corrélation de l'ordre de $\sqrt{0,7} = 0,84$. Nous avons dessiné sur la fig. 1 deux lignes pointillées fines qui délimitent une bande relativement étroite du plan x, y et contenant les deux droites de régression; tout point représentatif tombant à l'intérieur de cette « bande de régression » correspond à des valeurs mensuelles ayant entre elles une corrélation proche de celle qui régnait en moyenne sur la période 1911-1960.

Dans le but de conférer une certaine utilité pratique à la fig. 1 précisons le sens de quelques termes d'usage courant. L'affirmation que tel mois a été *sec* ou *pluvieux* repose exclusivement sur la hauteur d'eau tombée (variable x), tandis qu'une pluie dite *forte* ou *faible* implique une appréciation du débit moyen (variables x et y); quant aux termes *intermittent*, *continu*, ils se réfèrent à la durée seulement (variable y).

Or le débit moyen d'un mois caractérisé par les valeurs x,y est donné par le quotient x/y; en particulier l'inclinaison de la bande de régression mesure ce que l'on peut appeler un débit moyen *normal*

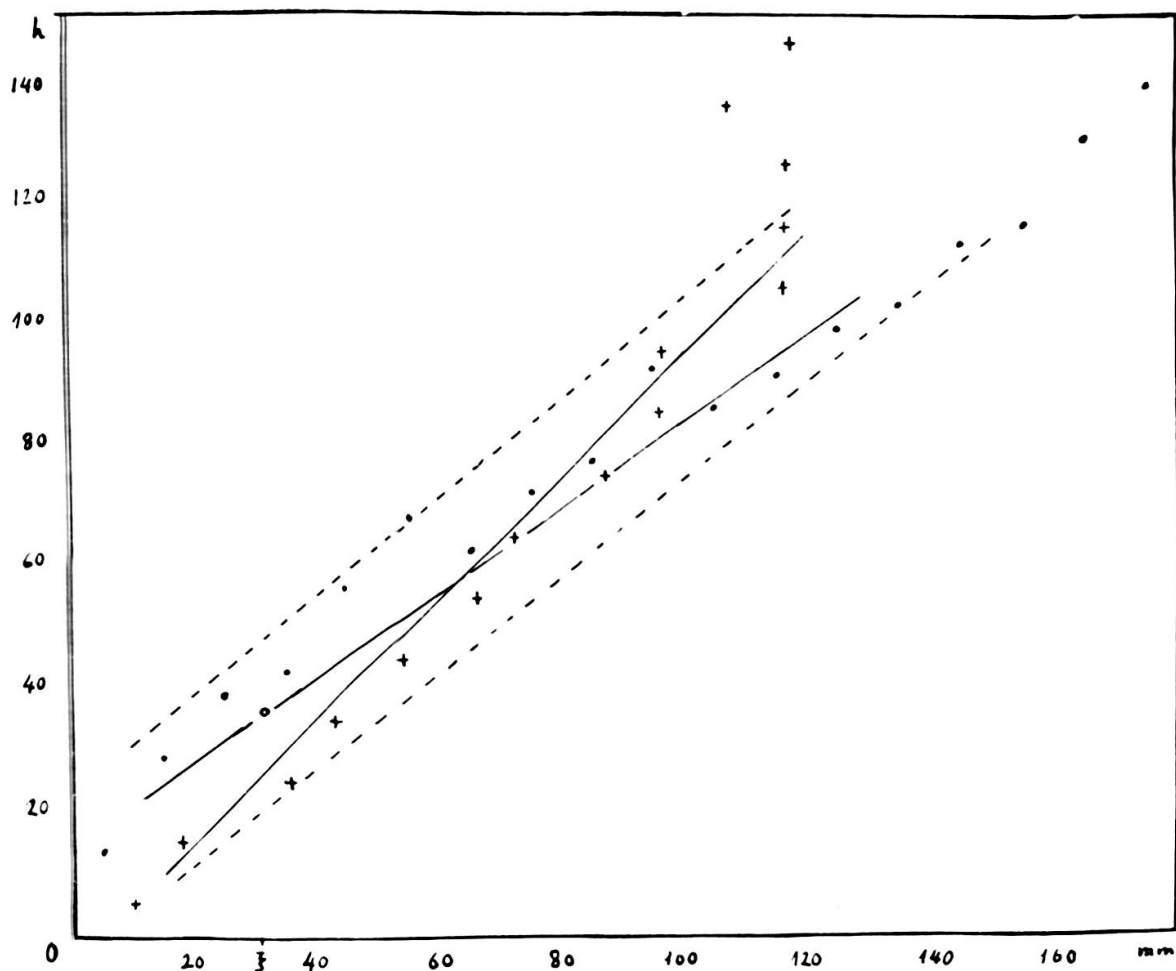


Fig. 1.

et nous sommes en mesure de dégager ici les règles suivantes:
Un mois est d'autant plus sec que son point représentatif est à gauche dans le plan x,y .

Tout point se trouvant au-dessus (ou à gauche) de la bande de régression correspond à un mois à précipitations anormalement faibles; inversement si un point se situe au-dessous (ou à droite) de la bande de régression, le mois correspondant se distingue par des précipitations anormalement fortes.

*Observatoire de Genève
 Avril 1962.*

