Apportons maintenant une petite modification aux conditions de l'addition: supposons que $p^{(i)}_{\cdot j}$ désigne la probabilité conditionnelle pour que la variable Y prenne la valeur Y_j, à condition que *ce ne soit pas* X qui ait pris la valeur X_i, mais que ξ ait pris la valeur ξ_i; de la même manière, $p_{i\cdot}$ désignera la probabilité pour que ξ ait pris la valeur ξ_i.

Désignons le nouveau rapport — mesure, obtenu dans ces couditions, par

$$\eta_{y/\xi}$$

Si dans la formule relative à $\mu^{(i)}_{\cdot/1}$ nous remplaçons Y_j par $f(\xi_i) + \varepsilon_j$ et si nous notons que

$$m^{(i)}_{\cdot/1} = \sum_j p^{(i)}_{\cdot/j}\,[f(\xi_i) + \varepsilon_j) = f(\xi_i) \sum_j p^{(i)}_{\cdot/j} + \\ + \sum_j p^{(i)}_{\cdot/j}\,\varepsilon_j = f(\xi_i) + \sum_j p^{(i)}_{\cdot/j}\,\varepsilon_j,$$

nous obtenons:

$$\mu^{(i)}_{\cdot/2} = \sum_j p^{(i)}_{\cdot/j}\left[\varepsilon_j - \sum_j p^{(i)}_{\cdot/j}\,\varepsilon_j\right]^2$$

Si la série des composantes ε est homogène et stochastiquement indépendante de ξ, on a:

$\mu^{(i)}_{\cdot/2} = \mu(\varepsilon)_2 = \sigma_\varepsilon{}^2$ pour chaque i, et, par conséquent:

$$\sum_j p_{i/\cdot}\,\mu^{(i)}_{\cdot/2} = \mu(\varepsilon)_2 \sum_j p_{i/\cdot} = \mu(\varepsilon)_2 = \sigma_\varepsilon{}^2$$

De cette façon on obtient définitivement:

$$\eta^2_{y/\xi} = 1 - \frac{\sigma_\varepsilon{}^2}{\sigma_y{}^2} = q_2{}^2$$

En appliquant des raisonnements analogiques on obtient:

$$\eta^2_{x/\psi} = 1 - \frac{\sigma_e{}^2}{\sigma_x{}^2} = q_1{}^2;\ H = \eta_{x/\psi}\,\eta_{y/\xi}$$

Si dans la série X la composante e manque, de sorte que $X_i = \xi_i$, il est évident que $q_1 = 1$, par conséquent $H = q_2 = \eta_{x/y}$, c.-à.-d. H est exactement égal au rapport de corrélation à priori de Karl Pearson.

De la même manière, si la composante ε manque, on a

$$H = q_1 = \eta_{x/y}.$$

Les grandeurs q_1 et q_2 sont des grandeurs à priori. Nous sommes obligés de prendre, comme leurs valeurs empiriques approchées, les mêmes $\eta'_{y/x}$ et $\eta'_{x/y}$ (comparer Tchouproff, page 69 et Darmois, page 214) qui sont considérés comme approximations empiriques des rapports $\eta_{x/y}$ et $\eta_{y/x}$, et qui, d'après les notations adoptées par les statisticiens anglais, s'expriment par les formules suivantes:

$$[\eta'_{x/y}]^2 = \frac{\Sigma\{n_y(\bar{x} - \bar{x}_y)^2\}}{N\,\sigma_x{}^2};\ [\eta'_{y/x}]^2 = \frac{\Sigma\{n_x(\bar{y} - \bar{y}_x)^2\}}{N\,\sigma_y{}^2}$$

(Comparer W. Palin Elderton, Fraquency Curves and Correlation, 2[d] edition, London 1927, page 195). Ici n_y est le nombre des cas dans la série „y" du tableau de corrélations; $\bar{x}$ — la moyenne arithmétique de tous les x; $\bar{x}_y$ — la moyenne arithmétique de tous les x se trouvant dans la série „y"; N — le nombre total des cas; σ_x^2 — l'écart-type de la série x; pour $\eta'_{y/x}$ on a des notations analogiques).

La fait est que nous sommes obligés de considérer les grandeurs X_i et Y_i comme approximations empiriques de ξ_i et ψ_i. Quand $E_e = 0$ et $E_\varepsilon = 0$, cela ne provoque pas d'incertitude. Toujours est-il que nous apportons par là une erreur systématique.

Il est nécessaire de souligner encore qu'on doit manipuler les grandeurs $\eta'_{x/y}$ et $\eta'_{y/x}$ avec la plus grande circonspection, vu les limites très larges de leurs erreurs (voir Tchouproff, Grundbegriffe, page 102). De plus, le rapport de corrélation empirique a encore une qualite désavantageuse, à savoir: si le nombre des couples de valeurs corréliées X_i Y_i est égal au nombre des différentes valeurs que prend la grandeur X dans les limites de nos observations, de sorte qu'à chaque valeur X ne correspond qu'une valeur Y, $\eta'_{y/x}$ étant toujours égal à 1. Le même mutatis mutandis est valable aussi pour $\eta'_{x/y}$ (Comparer Tchouproff page 70).

Examinons à présent le cas où ψ_i et ξ_i sont liés entre eux par une dépendance linéaire, de sorte que pour chaque i

$$\psi_i = a + b\xi_i$$

Alors

$$\sigma^2_\psi = E[a + b\xi_i - E(a + b\xi_i)]^2 = E[b(\xi_i - E\xi)]^2 = b^2\sigma^2_\xi$$

D'autre part:

$$E[(X_i - EX)(Y_i - EY)] = E\left\{[(\xi_i - E\xi) + \right. \\ + (e_i - E_e)]\,[(\psi_i - E\psi) + (\varepsilon_i - E\varepsilon)]\left.\right\} = \\ = E[(\xi_i - E\xi)(\psi_i - E\psi)] + E[(e_i - Ee)(\psi_i - E\psi)] + \\ + E[(\xi_i - E\xi)(\varepsilon_i - E\varepsilon)] + E[(e_i - Ee)(\varepsilon_i - E\varepsilon)]$$

Si „e" et ψ, ξ et ε, „e" et ε sont des grandeurs stochastiques réciproquement indépendantes, les derniers trois termes de la partie droite disparaissent et on obtient:

$$E[(X_i - EX)(Y_i - EY)] = E[(\xi_i - E\xi)(\psi_i - E\psi)]$$

Ensuite:

$$E[(\xi_i - E\xi)(\psi_i - E\psi)] = E\{(\xi_i - E\xi)\,b\,(\xi_i - E\xi)\} = \\ = b\,\sigma_\xi{}^2 = \sigma_\xi\,.\,b\,\sigma_\xi$$

Mais puisque, comme on vient de le dire plus haut,

$$\sigma_\psi = |b|\,\sigma_\xi$$

(les lignes verticales autour de b désignent que cette grandeur est prise ici toujours avec le signe +), on obtient définitivement:

$$E[(X_i - EX)(Y_i - EY)] = \pm\,\sigma_\xi\,\sigma_\psi$$

où le signe (+) ou (—) est déterminé par le signe du coefficient b.

Comme on sait, la formule relative au coefficient de corrélation à priori est:

$$r_{12} = \frac{E[(X_i - E\,X)\,(Y_i - E\,Y)]}{\sqrt{E\,(X_i - E\,X)^2\,E\,(Y_i - E\,Y)^2}}$$

En y introduisant les valeurs obtenues, on a:

$$r_{12} = \pm\,\frac{\sigma_\xi}{\sigma_x}\cdot\frac{\sigma_\psi}{\sigma_y} = \pm\,q_1\,q_2 = \pm\,H$$

Nous allons signaler encore les rapports intéressants ci-après.