Longeur de clef de Vigenère par Friedman *

m : Longueur de clef $K = K_1 \dots K_m$; $n = t \cdot m$: Longueur du clair (et du chiffré).

Occurences des lettres dans le clair :$n_A \dots n_Z$.

On a l'indice de coïncidence induit (en supposant que les lettres sont réparties uniformément dans le message avec les occurences $n_A \dots n_Z$, l'IC est la probabilité que deux lettres prises au hasard dans le message soient identiques) :

$$I_c = \sum_{\lambda=A}^{Z} \frac{n_\lambda(n_\lambda - 1)}{n(n-1)} = \sum_{\lambda=A}^{Z} f_z \left(f_z - \frac{1 - f_z}{n-1} \right)$$

On note également I_v l'indice de coïncidence induit par le chiffré.

On découpe le chiffré en m lignes associées à chacune des lettres de la clef. chaque ligne comporte donc t lettres.

Prenons 2 lettres au hasard parmi les n lettres, deux cas possibles :

1. Deux lettres sont choisies dans la même ligne : m lignes, $\frac{t(t-1)}{2}$ paires de lettres dans une ligne. Deux telles lettres sont égales avec la probabilité induite par l'indice de coïncidence du message en clair $I_c(Messageclair)$, qui peut être approché, si le clair est suffisamment long, par l'indice de coïncidence de la Langue L considérée I_L. Le nombre de coïncidences dans ce premier cas est donc :
$$N_1 = m\frac{t(t-1)}{2} I_c \approx m\frac{t(t-1)}{2} I_L$$
2. Deux lettres sont dans des lignes distinctes : $\frac{m(m-1)}{2}$ paires de lignes, t^2 paires de lettres. Une fois les positions lignes de lettre choisies i et j, étant donnée la lettre associée α à la première position, une seule lettre permet l'égalité : $\beta = \alpha + K_i - K_j$.
Hypothèse : *on suppose que les lettres du clair ont été placées aléatoirement (uniformément) dans les cases de la matrice.*
Dans ce cas, sachant que $\alpha \neq \beta$ a déjà été choisie, il y a $\frac{n_\beta}{n-1}$ chances pour que la lettre β se trouve à cette position. Le nombre de coïncidences est donc
$$N_2 = t^2 \sum_{i \neq j} \sum_{\alpha=A}^{Z} \frac{n_\alpha}{n} \frac{n_{\beta(\alpha,i,j)}}{n-1}$$

*Jean-Guillaume Dumas 29 janvier 2016

ou encore

$$N_2 = \frac{t^2}{n(n-1)} \sum_{\alpha=A}^{Z} n_\alpha \left(\sum_{i \neq j} n_{\beta(\alpha,i,j)} \right)$$

Hypothèse : *on suppose alors que les écarts (i, j) couvrent chacune des lettres de l'alphabet en nombre quasiment égal, soit $\frac{m(m-1)}{2}\frac{1}{26}$ fois chacune.*

$$\sum_{i \neq j} n_{\beta(\alpha,i,j)} \approx \frac{m(m-1)}{2}\frac{1}{26} \sum_{\lambda=A}^{Z} n_\lambda = \frac{m(m-1)}{2}\frac{n}{26}$$

Et donc le nombre de coïncidences dans ce deuxième cas est approché par :

$$N_2 \approx \frac{t^2}{n(n-1)} \frac{m(m-1)}{2}\frac{n}{26} \sum_{\alpha=A}^{Z} n_\alpha = \frac{m(m-1)t^2}{2}\frac{n}{26(n-1)}$$

Au total le nombre de coïncidences sur le nombre total de paires vérifie donc :

$$I_v = \frac{N_1 + N_2}{\frac{n(n-1)}{2}} \approx \frac{\frac{mt(t-1)}{2} I_L + \frac{m(m-1)t^2}{2}\frac{n}{26(n-1)}}{\frac{n(n-1)}{2}}$$

Soit

$$I_v \approx \frac{(t-1)I_L + (m-1)t\frac{n}{26(n-1)}}{n-1}$$

On note

$$p = \frac{t-1}{n-1} = \frac{1}{m}\left(1 - \frac{m-1}{n-1}\right) \tag{1}$$

Alors

$$I_v \approx p \cdot I_L + (1-p) \cdot \frac{1}{26} \cdot \frac{n}{n-1} \tag{2}$$

Résolvant (2) pour p, on obtient

$$p \approx \frac{(n-1)\left(I_v - \frac{1}{26}\right) - 1}{(n-1)\left(I_L - \frac{1}{26}\right) - 1} \tag{3}$$

Résolvant (1) pour $\frac{1}{m}$, on obtient

$$\frac{1}{m} = p + \frac{1-p}{n} \tag{4}$$

Maintenant si n est grand alors on a

$$\boxed{\frac{I_v - \frac{1}{26}}{I_L - \frac{1}{26}} \approx p \approx \frac{1}{m}}$$