À propos de l'article de Chebyshev, *Démonstration élémentaire d'une proposition générale de la théorie des probabilités*, 1846

Didier Trotoux

avril 2013

Cet article a été publié par P. L. Chebyshev dans le Journal de Crelle en 1846 (B. 33, p. 259-267). Comme il l'indique dès le premier paragraphe, Chebyshev propose de démontrer la proposition : ≪ On peut toujours assigner un nombre d'épreuves tel que la probabilité de ce que le rapport des répétitions de l'événement *E* à celui des épreuves ne s'écartera pas de la moyenne des chances de *E* au delà des limites données, quelques resserrées que soient ces limites, s'approchera autant qu'on le voudra de la certitude ≫. S. D. Poisson avait démontré cette proposition qu'il avait appelée *Loi des grands nombres*, dans ses *Recherches sur la probabilité des jugements* (Bachelier, Paris, 1837), en la déduisant d'une formule qu'il avait obtenue par l'approximation de la valeur d'une intégrale définie assez compliquée. Chebyshev trouvant cette démonstration incomplète car ne fournissant pas l'erreur que comportait cette approximation et manquant de rigueur à cause de cette incertitude, propose d'en donner une démonstration rigoureuse en utilisant des considérations élémentaires.
L'objet de cet article est de présenter cette démonstration en la commentant.

Dans le deuxième paragraphe, Chebyshev s'intéresse à la réalisation d'un événement *E* lors de la répétition de μ épreuves consécutives. p_i est la probabilité que *E* se réalise lors de l'épreuve i et P_m la probabilité que E arrivera au moins m fois au cours de ces μ épreuves. Il indique que P_m s'obtient en prenant la somme des coefficients de $t^m, t^{m+1}, \dots t^\mu$ dans le développement du produit $(p_1t+1-p_1)(p_2t+1-p_2)\cdots(p_\mu t+1-p_\mu)$.

Commentaire :
Notons X_i la variable qui vaut 1 si l'événement *E* se produit lors de l'épreuve i et 0 dans le cas contraire. La variable X_i suit une loi de Bernoulli $B(p_i)$. La variable égale au nombre total de succès au cours des μ épreuves vaut $X_1+X_2+\cdots+X_\mu$ et la probabilité P_m s'écrit $P_m = \mathbb{P}(X_1+X_2+\cdots+X_\mu \geq m)$.
Rappelons que la fonction génératrice d'une variable aléatoire X est définie par $g_X(t) = \sum_{k=0}^{k=+\infty} \mathbb{P}(X=k)\cdot t^k$. Pour une variable X suivant une loi de Bernoulli de paramètre p, $g_X(t) = (1-p)t^0 + pt = pt+1-p$ et comme on a, pour deux variables indépendantes X et Y, $g_{X+Y}(t) = g_X(t)\times g_Y(t)$, on obtient $g_{X_1+X_2+\cdots+X_\mu}(t) = (p_1t+1-p_1)(p_2t+1-p_2)\cdots(p_\mu t+1-p_\mu)$ et le résultat énoncé.

L'observation de cette expression lui permet de faire les deux remarques suivantes :

1. P_m est une fonction du premier degré des variables $p_1, p_2, \dots p_\mu$
2. P_m est une fonction symétrique des variables $p_1, p_2, \dots p_\mu$

ce qui lui permet d'écrire P_m sous la forme $P_m = U + Vp_1 + V_1p_2 + Wp_1p_2$ où U, V, V_1, W ne dépendent pas de p_1 et p_2 (d'après 1.) puis, $P_m = U + V(p_1+p_2) + Wp_1p_2$ où U, V, W ne dépendent pas de p_1 et p_2 (d'après 2.).

Il peut alors énoncer un premier théorème :

Théorème 1
Si p_1 et p_2 ne sont pas égales, on peut sans changer les valeurs de $p_1+p_2, p_3, \dots p_\mu$, augmenter celle de P_m en prenant $p_1 = p_2$; ou on peut parvenir à une des équations suivantes : $p_1 = 0, p_1 = 1$ sans diminuer la valeur de P_m.

Commentaire :
$P_m(p_1, p_2, p_3, \cdots p_\mu) = U + V(p_1+p_2) + Wp_1p_2$ est une expression algébrique où l'on peut, suivant le signe de W, modifier les valeurs de p_1 et p_2 tout en gardant p_1+p_2 constant sans la diminuer.

1. Si $W > 0$, $P_m(\frac{1}{2}(p_1+p_2), \frac{1}{2}(p_1+p_2), p_3, \cdots p_\mu) - P_m(p_1, p_2, p_3, \cdots p_\mu) = \frac{1}{4}W(p_1-p_2)^2$.
2. Si $W \leq 0$ et $p_1+p_2 \leq 1$, $P_m(0, p_1+p_2, p_3, \cdots p_\mu) - P_m(p_1, p_2, p_3, \cdots p_\mu) = -Wp_1p_2$
3. Si $W \leq 0$ et $p_1+p_2 \geq 1$, $P_m(1, p_1+p_2-1, p_3, \cdots p_\mu) - P_m(p_1, p_2, p_3, \cdots p_\mu) = -W(1-p_1)(1-p_2)$

Il énonce ensuite un deuxième théorème :

Théorème 2
La plus grande valeur que P_m peut avoir dans le cas où $p_1 + p_2 + p_3 + \cdots + p_\mu = S$, correspond aux valeurs de $p_1, p_2, p_3, \cdots p_\mu$ données par les équations

$$p_1 = 0, p_2 = 0, \cdots p_\rho = 0, p_{\rho+1} = 1, p_{\rho+2} = 1, p_{\rho+\sigma} = 1$$
$$p_{\rho+\sigma+1} = \frac{S-\sigma}{\mu-\rho-\sigma}, p_{\rho+\sigma+2} = \frac{S-\sigma}{\mu-\rho-\sigma}, \cdots p_\mu = \frac{S-\sigma}{\mu-\rho-\sigma},$$

où ρ, σ désignent certains nombres.

Commentaire :
Chebyshev introduit $\pi_1, \pi_2, \pi_3, \cdots \pi_\mu$ les valeurs de $p_1, p_2, p_3, \cdots p_\mu$ maximisant P_m sous la condition $p_1 + p_2 + p_3 + \cdots + p_\mu = S$ et ayant le plus grand nombre de valeurs égales à 0 et à 1. En supposant que $\pi_1 = \pi_2 = \cdots = \pi_\rho = 0$ et $\pi_{\rho+1} = \pi_{\rho+2} = \cdots = \pi_{\rho+\sigma} = 1$, il montre, par un raisonnement par l'absurde utilisant le premier théorème, que les valeurs $\pi_{\rho+\sigma+1}, \pi_{\rho+\sigma+2}, \cdots, \pi_\mu$, supposées différentes de 0 et 1 doivent être choisies égales entre elles. Il y a donc ρ valeurs nulles, σ valeurs égales à 1 et $\mu-\rho-\sigma$ valeurs égales à une constante π. Le total de ces valeurs étant égal à S, il obtient $\pi = \dfrac{S-\sigma}{\mu-\rho-\sigma}$.

En conséquence, sous la condition $p_1 + p_2 + p_3 + \cdots + p_\mu = S$, P_m est la somme des coefficients de t^k pour $k \geq m$ dans le développement du produit

$$t^\sigma \left(\frac{S-\sigma}{\mu-\rho-\sigma}t + \frac{\mu-S-\rho}{\mu-\rho-\sigma}\right)^{\mu-\rho-\sigma} = t^\sigma \sum_{i=0}^{\mu-\rho-\sigma} \binom{\mu-\rho-\sigma}{i} \left(\frac{S-\sigma}{\mu-\rho-\sigma}t\right)^{\mu-\rho-\sigma-i} \left(\frac{\mu-S-\rho}{\mu-\rho-\sigma}\right)^i$$
$$= \sum_{i=0}^{\mu-\rho-\sigma} \binom{\mu-\rho-\sigma}{i} \left(\frac{S-\sigma}{\mu-\rho-\sigma}\right)^{\mu-\rho-\sigma-i} \left(\frac{\mu-S-\rho}{\mu-\rho-\sigma}\right)^i \cdot t^{\mu-\rho-i}$$

On ne garde que les coefficients de $t^{\mu-\rho-i}$ pour lesquels $\mu-\rho-i \geq m$, c'est-à-dire $i \leq \mu-m-\rho$.

$$P_m = \sum_{i=0}^{\mu-m-\rho} \binom{\mu-\rho-\sigma}{i} \left(\frac{S-\sigma}{\mu-\rho-\sigma}\right)^{\mu-\rho-\sigma-i} \left(\frac{\mu-S-\rho}{\mu-\rho-\sigma}\right)^i$$

et en faisant le changement de variable $k = \mu-m-\rho-i$, il vient :

$$P_m = \sum_{k=0}^{\mu-m-\rho} \binom{\mu-\rho-\sigma}{\mu-m-\rho-k} \left(\frac{S-\sigma}{\mu-\rho-\sigma}\right)^{m-\sigma+k} \left(\frac{\mu-S-\rho}{\mu-\rho-\sigma}\right)^{\mu-m-\rho-k}$$

$$P_m = \binom{\mu-\rho-\sigma}{\mu-m-\rho} \left(\frac{S-\sigma}{\mu-\rho-\sigma}\right)^{m-\sigma} \left(\frac{\mu-S-\rho}{\mu-\rho-\sigma}\right)^{\mu-m-\rho} \cdot \left[\sum_{k=0}^{\mu-m-\rho} \frac{\binom{\mu-\rho-\sigma}{\mu-m-\rho-k}}{\binom{\mu-\rho-\sigma}{\mu-m-\rho}} \left(\frac{S-\sigma}{\mu-S-\rho}\right)^k\right]$$

La quantité entre crochets s'écrit :

$$A = 1 + \frac{\mu-m-\rho}{m-\sigma+1} \cdot \frac{S-\sigma}{\mu-S-\rho} + \frac{\mu-m-\rho}{m-\sigma+1} \cdot \frac{\mu-m-\rho-1}{m-\sigma+2} \cdot \left(\frac{S-\sigma}{\mu-S-\rho}\right)^2$$
$$+ \cdots + \frac{\mu-m-\rho}{m-\sigma+1} \cdot \frac{\mu-m-\rho-1}{m-\sigma+2} \cdot \ldots \cdot \frac{1}{\mu-\rho-\sigma} \left(\frac{S-\sigma}{\mu-S-\rho}\right)^{\mu-m-\rho}$$

et est majorée par la série géométrique :

$$\sum_{k=0}^{\mu-m-\rho} \left(\frac{\mu-m-\rho}{m-\sigma} \cdot \frac{S-\sigma}{\mu-S-\rho}\right)^k = \frac{1 - \left(\dfrac{\mu-m-\rho}{m-\sigma} \cdot \dfrac{S-\sigma}{\mu-S-\rho}\right)^{\mu-m-\rho+1}}{1 - \dfrac{\mu-m-\rho}{m-\sigma} \cdot \dfrac{S-\sigma}{\mu-S-\rho}}$$

ou bien :

$$\frac{(m-\sigma)(\mu-S-\rho)}{(m-S)(\mu-\rho-\sigma)} \left[1 - \left(\frac{\mu-m-\rho}{m-\sigma} \cdot \frac{S-\sigma}{\mu-S-\rho}\right)^{\mu-m-\rho+1}\right].$$

À l'aide de cette majoration, Chebyshev peut alors énoncer un troisième théorème :

Théorème 3
Pour certains nombres entiers et positifs ρ et σ, la valeur de l'expression

$$\binom{\mu-\rho-\sigma}{\mu-m-\rho}\left(\frac{S-\sigma}{\mu-\rho-\sigma}\right)^{m-\sigma}\left(\frac{\mu-S-\rho}{\mu-\rho-\sigma}\right)^{\mu-m-\rho+1}\left(\frac{m-\sigma}{m-S}\right)\left[1-\left(\frac{\mu-m-\rho}{m-\sigma}\cdot\frac{S-\sigma}{\mu-S-\sigma}\right)^{\mu-m-\rho+1}\right]$$

surpasse la valeur P_m de la probabilité que dans μ épreuves l'événement E, ayant les chances $p_1, p_2, p_3, \cdots p_\mu$, arrivera au moins m fois, où S est la somme $p_1 + p_2 + p_3 + \cdots + p_\mu$.

Il déduit de ce troisième théorème que :

$$P_m < \binom{\mu-\rho-\sigma}{\mu-m-\rho}\left(\frac{S-\sigma}{\mu-\rho-\sigma}\right)^{m-\sigma}\left(\frac{\mu-S-\rho}{\mu-\rho-\sigma}\right)^{\mu-m-\rho+1}\left(\frac{m-\sigma}{m-S}\right) \quad (1)$$

et justifie que, sous l'hypothèse $m > S+1$, cette dernière expression, que nous noterons $f(\sigma,\rho)$, augmente quand σ ou ρ diminue.

En effet, le rapport $\dfrac{f(\sigma-1,\rho)}{f(\sigma,\rho)}$ peut être écrit sous la forme :

$$\frac{1}{1+\dfrac{m-S-1}{S-\sigma+1}}\exp\left(-(m-\sigma)\ln\left(1-\frac{1}{s-\sigma+1}\right)+(\mu-\rho-\sigma+1)\ln\left(1-\frac{1}{\mu-\rho-\sigma+1}\right)\right)$$

$$=\frac{1}{1+\dfrac{m-S-1}{S-\sigma+1}}\exp\left(\frac{m-S-1}{S-\sigma+1}\right)\exp\left(\sum_{k=1}^{+\infty}\frac{1}{k+1}\left(\frac{m-\sigma}{(S-\sigma+1)^{k+1}}-\frac{1}{(\mu-\rho-\sigma+1)^k}\right)\right)$$

D'une part, l'expression $\dfrac{1}{1+\dfrac{m-S-1}{S-\sigma+1}}\exp\left(\dfrac{m-S-1}{S-\sigma+1}\right)$, de la forme $\dfrac{e^u}{1+u}$ est supérieure à 1, pour $u \geq 0$.

D'autre part, $\dfrac{m-\sigma}{(S-\sigma+1)^{k+1}}-\dfrac{1}{(\mu-\rho-\sigma+1)^k}=\dfrac{1}{(\mu-\rho-\sigma+1)^k}\left[\dfrac{m-\sigma}{S-\sigma+1}\left(\dfrac{\mu-\rho-\sigma+1}{S-\sigma+1}\right)^k-1\right]$ est une quantité positive car l'hypothèse $m > S+1$ entraîne $\dfrac{m-\sigma}{S-\sigma+1} > 1$ et qu'il résulte du fait que la quantité $\dfrac{S-\sigma}{\mu-\rho-\sigma}$ est une probabilité (*cf* commentaire du *Théorème* 2) que $\dfrac{\mu-\rho-\sigma+1}{S-\sigma+1} \geq 1$.

Le rapport $\dfrac{f(\sigma-1,\rho)}{f(\sigma,\rho)}$ étant supérieur à 1, $f(\sigma-1,\rho) < f(\sigma,\rho)$ et on peut obtenir un résultat analogue lorsque l'on diminue ρ.
Chebyshev en conclut que, sous l'hypothèse $m > S+1$, P_m est nécessairement inférieur à la valeur $f(0,0)$ obtenue en remplaçant ρ par 0 et σ par 0 :

$$P_m < \binom{\mu}{\mu-m}\left(\frac{S}{\mu}\right)^m\left(\frac{\mu-S}{\mu}\right)^{\mu-m+1}\frac{m}{m-S} \quad (2)$$

Il lui reste à majorer le coefficient binomial $\binom{\mu}{\mu-m}$.
Pour ce faire, il utilise l'encadrement $2,50\, x^{x+\frac{1}{2}}e^{-x} < x! < 2,53\, x^{x+\frac{1}{2}}e^{-x+\frac{1}{12x}}$, qu'il justifie dans une note de bas de page.

À l'aide de cet encadrement, il obtient $\binom{\mu}{\mu-m} < \dfrac{2,53\, e^{\frac{1}{12\mu}}}{(2,50)^2}\dfrac{\mu^{\mu+\frac{1}{2}}}{m^{m+\frac{1}{2}}(\mu-m)^{\mu-m+\frac{1}{2}}}$.

Comme on a $\dfrac{2,53\, e^{\frac{1}{12\mu}}}{(2,50)^2} < \dfrac{2,53\, e^{\frac{1}{12}}}{(2,50)^2} < \dfrac{1}{2}$, il en déduit :

$$P_m < \frac{\frac{1}{2}\mu^{\mu+\frac{1}{2}}}{m^{m+\frac{1}{2}}(\mu-m)^{\mu-m+\frac{1}{2}}}\left(\frac{S}{\mu}\right)^m\left(\frac{\mu-S}{\mu}\right)^{\mu-m+1}\frac{m}{m-S}$$

ou de manière équivalente :

$$P_m < \frac{1}{2(m-S)}\sqrt{\frac{m(\mu-m)}{\mu}}\left(\frac{S}{m}\right)^m\left(\frac{\mu-S}{\mu-m}\right)^{\mu-m+1}$$

Il peut énoncer alors le théorème suivant :

Théorème 4
≪ Si les chances de l'événement E dans μ épreuves consécutives sont $p_1, p_2, p_3, \cdots p_\mu$, et que leur somme est S, la valeur de l'expression

$$\frac{1}{2(m-S)}\sqrt{\frac{m(\mu-m)}{\mu}}\left(\frac{S}{m}\right)^m\left(\frac{\mu-S}{\mu-m}\right)^{\mu-m+1}$$

pour m plus grand que $S+1$, surpasse toujours la probabilité que E arrivera au moins m fois dans ces μ épreuves. ≫

En changeant ensuite $m, p_1, p_2, p_3, \cdots p_\mu, S$ en $\mu-n, 1-p_1, 1-p_2, 1-p_3, \cdots 1-p_\mu, \mu-S$, il déduit de ce théorème que, si la somme $1-p_1+1-p_2+1-p_3+\cdots+1-p_\mu=\mu-S$, la valeur de l'expression

$$\frac{1}{2(S-n)}\sqrt{\frac{n(\mu-n)}{\mu}}\left(\frac{\mu-S}{\mu-n}\right)^{\mu-n}\left(\frac{S}{n}\right)^{n+1}$$

pour $\mu-n$ plus grand que $\mu-S+1$, surpasse celle de la probabilité que l'événement contraire à E arrivera au moins $\mu-n$ fois dans ces μ épreuves où $p_1, p_2, p_3, \cdots p_\mu$ sont les chances de E.

Mais comme les conditions $1-p_1+1-p_2+1-p_3+\cdots+1-p_\mu=\mu-S$ et $\mu-n>\mu-S+1$ se ramènent respectivement à $p_1+p_2+p_3+\cdots+p_\mu=S$ et $n<S-1$ et que dire que l'événement contraire à E arrivera au moins $\mu-n$ fois dans μ épreuves équivaut à dire que l'événement E se produira au plus n fois, il peut énoncer le théorème suivant :

Théorème 5
Si les chances de l'événement E dans μ épreuves consécutives sont $p_1, p_2, p_3, \cdots p_\mu$ et que leur somme est S, la valeur de l'expression

$$\frac{1}{2(S-n)}\sqrt{\frac{n(\mu-n)}{\mu}}\left(\frac{\mu-S}{\mu-n}\right)^{\mu-n}\left(\frac{S}{n}\right)^{n+1}$$

pour n plus petit que $S-1$, surpasse la probabilité que E n'arrivera pas plus de n fois dans ces μ épreuves.

Comme conséquence de ces deux derniers théorèmes, résulte le suivant :

Théorème 6
Si les chances de l'événement E dans μ épreuves consécutives sont $p_1, p_2, p_3, \cdots p_\mu$ et que leur somme est S, la probabilité que le nombre de répétitions de l'événement E dans ces μ épreuves sera moindre que m et plus grand que n, surpassera, pour m plus grand que $S+1$ et pour n plus petit que $S-1$, la valeur de l'expression

$$1-\frac{1}{2(m-S)}\sqrt{\frac{m(\mu-m)}{\mu}}\left(\frac{S}{m}\right)^m\left(\frac{\mu-S}{\mu-m}\right)^{\mu-m+1}-\frac{1}{2(S-n)}\sqrt{\frac{n(\mu-n)}{\mu}}\left(\frac{\mu-S}{\mu-n}\right)^{\mu-n}\left(\frac{S}{n}\right)^{n+1}.$$

Ce dernier théorème lui permet d'atteindre le résultat qu'il a annoncé au début de son article.

En effet, pour $n=S-\mu z$, $m=S+\mu z$ et $\mu z>1$, les conditions $n<S-1$ et $m>S+1$ sont vérifiées et, par conséquent, la probabilité que le nombre de répétitions de l'événement E dans ces μ épreuves sera moindre que m et plus grand que n, surpasse la valeur de l'expression

$$1-\frac{1}{2\mu z}\sqrt{\frac{(S-\mu z)(\mu-S-\mu z)}{\mu}}\left(\frac{S}{S+\mu z}\right)^{S+\mu z}\left(\frac{\mu-S}{\mu-S-\mu z}\right)^{\mu-S-\mu z+1}$$
$$-\frac{1}{2\mu z}\sqrt{\frac{(S-\mu z)(\mu-S+\mu z)}{\mu}}\left(\frac{\mu-S}{\mu-S+\mu z}\right)^{\mu-S+\mu z}\left(\frac{S}{S-\mu z}\right)^{S-\mu z+1}$$

expression qui, en posant $p=\dfrac{S}{\mu}$, peut être simplifiée en

$$1-\frac{1-p}{2z\sqrt{\mu}}\sqrt{\frac{p+z}{1-p-z}}\,(H_p(z))^\mu-\frac{p}{2z\sqrt{\mu}}\sqrt{\frac{1-p+z}{p-z}}\,(H_{1-p}(z))^\mu \text{ où } H_p(z)=\left(\frac{p}{p+z}\right)^{p+z}\left(\frac{1-p}{1-p-z}\right)^{1-p-z}.$$

Il développe en série entière $\ln(H_p(z))$.

$$\ln(H_p(z)) = -(p+z)\ln\left(1+\frac{z}{p}\right) - (1-p-z)\ln\left(1-\frac{z}{1-p}\right) = -(p+z)\sum_{k\geq 1}\frac{(-1)^{k-1}}{k}\frac{z^k}{p^k} + (1-p-z)\sum_{k\geq 1}\frac{1}{k}\frac{z^k}{(1-p)^k}.$$

Le premier morceau de cette expression vaut :

$$-(p+z)\sum_{k\geq 1}\frac{(-1)^{k-1}}{k}\frac{z^k}{p^k} = -\sum_{k\geq 1}\frac{(-1)^{k-1}}{k}\frac{z^k}{p^{k-1}} - \sum_{k\geq 1}\frac{(-1)^{k-1}}{k}\frac{z^{k+1}}{p^k} = -z - \sum_{k\geq 2}\frac{(-1)^{k-1}}{k}\frac{z^k}{p^{k-1}} - \sum_{k\geq 2}\frac{(-1)^{k-2}}{k-1}\frac{z^k}{p^{k-1}}$$

$$= -z - \sum_{k\geq 2}(-1)^{k-1}\left(\frac{1}{k}-\frac{1}{k-1}\right)\frac{z^k}{p^{k-1}} = -z + \sum_{k\geq 2}\frac{(-1)^{k-1}}{k(k-1)}\frac{z^k}{p^{k-1}} = -z - \sum_{k\geq 1}\left(1-\frac{2k-1}{2k+1}\frac{z}{p}\right)\frac{1}{2k(2k-1)}\frac{z^{2k}}{p^{2k-1}}.$$

Le deuxième morceau de cette expression vaut :

$$(1-p-z)\sum_{k\geq 1}\frac{1}{k}\frac{z^k}{(1-p)^k} = \sum_{k\geq 1}\frac{1}{k}\frac{z^k}{(1-p)^{k-1}} - \sum_{k\geq 1}\frac{1}{k}\frac{z^{k+1}}{(1-p)^k} = z + \sum_{k\geq 2}\frac{1}{k}\frac{z^k}{(1-p)^{k-1}} - \sum_{k\geq 2}\frac{1}{k-1}\frac{z^k}{(1-p)^{k-1}}$$

$$= z + \sum_{k\geq 2}\left(\frac{1}{k}-\frac{1}{k-1}\right)\frac{z^k}{(1-p)^{k-1}} = z - \sum_{k\geq 2}\frac{1}{k(k-1)}\frac{z^k}{(1-p)^{k-1}}.$$

ce qui permet d'obtenir :

$$\ln(H_p(z)) = -\left[\sum_{k\geq 1}\left(1-\frac{2k-1}{2k+1}\frac{z}{p}\right)\frac{1}{2k(2k-1)}\frac{z^{2k}}{p^{2k-1}} + \sum_{k\geq 2}\frac{1}{k(k-1)}\frac{z^k}{(1-p)^{k-1}}\right].$$

et de constater que les quantités $H_p(z)$ et $H_{1-p}(z)$, ayant un logarithme négatif, ont des valeurs inférieures à 1.

Chebyshev peut alors conclure que

$$\lim_{\mu\to+\infty}\left(1-\frac{1-p}{2z\sqrt{\mu}}\sqrt{\frac{p+z}{1-p-z}}\,(H_p(z))^\mu - \frac{p}{2z\sqrt{\mu}}\sqrt{\frac{1-p+z}{p-z}}\,(H_{1-p}(z))^\mu\right) = 1$$

et précise que pour pour rendre la différence entre cette expression et 1 inférieure à ϵ, il suffit de choisir pour μ, un nombre quelconque supérieur à

$$\max\left(\frac{\ln\left(\epsilon\cdot\frac{z}{1-p}\sqrt{\frac{1-p-z}{p+z}}\right)}{\ln(H_p(z))}, \frac{\ln\left(\epsilon\cdot\frac{z}{p}\sqrt{\frac{p-z}{1-p+z}}\right)}{\ln(H_{1-p}(z))}\right).$$

Commentaire final :

Si les outils mathématiques utilisés par Chebyshev sont relativement élémentaires (fonctions génératrices, développement du binôme, encadrement de la fonction factorielle, développements en série), ils nécessitent cependant une bonne aptitude au calcul algébrique et une certaine virtuosité technique en analyse pour obtenir les encadrements souhaités. Chebyshev a le grand mérite de fournir une preuve nettement plus simple que celle de Poisson et qui n'est pas contestable.