le nombre de variables X_i qui tombent dans E_j (on a donc $N_1^n + \cdots + N_q^n = n$) et on pose

$$T_n = \sum_{j=1}^{q} \frac{1}{n\mu_j}(N_j^n - n\mu_j)^2.$$

N_j^n est appelée variable de comptage. L'idée du test du χ^2 d'adéquation est la suivante. Si $p = p_0$, on a convergence presque sûre de N_j^n/n vers μ_j par la loi des grands nombres. On a vu au chapitre 12 que T_n converge en loi vers une variable aléatoire finie. En revanche, si $p \neq p_0$, on a $\lim_{n\to+\infty} N_j^n/n = p(E_j)$ p.s. Par suite, si $p(E_j) \neq \mu_j$ pour au moins une valeur de j, alors les variables T_n tendent vers l'infini.

Proposition 13.4.1. *Soit p une loi de probabilité sur E. Alors, si p est la loi commune des variables indépendantes X_i définies plus haut, la quantité associée T_n vérifie*

$$\frac{T_n}{n} = \sum_{j=1}^{q} \frac{1}{n^2\mu_j}(N_j^n - n\mu_j)^2 \underset{n\to+\infty}{\rightarrow} \sum_{j=1}^{q} \frac{(\mu_j - p(E_j))^2}{\mu_j} \quad p.s.$$

En particulier, on retrouve que T_n tend vers l'infini si $p \neq p_0$.

Démonstration. En effet, $\displaystyle\frac{T_n}{n} = \sum_{j=1}^{q} \frac{1}{\mu_j}\left(\frac{N_j^n}{n} - \mu_j\right)^2$. D'après la loi forte des grands nombres, $\frac{N_j^n}{n}$ converge presque sûrement vers $p(E_j)$, d'où le résultat voulu. $\square$

Cela conduit à choisir une région critique de la forme $D = \{T_n \geq a\}$.

Il nous reste à déterminer le niveau du test. Cela est essentiel. Ce niveau est $\mathbb{P}_{p_0}(D)$ où $\mathbb{P}_{p_0}$ est la probabilité sur notre espace d'états lorsque la loi des X_i est p_0. Comme p_0 est connu, on peut calculer ce niveau : sous $\mathbb{P}_{p_0}$, la variable $S_n = (N_1^n, \ldots, N_q^n)$ suit une loi multinomiale donnée par

$$\mathbb{P}_{p_0}(N_1^n = i_1, \ldots, N_q^n = i_q) = n! \prod_{j=1}^{q} \frac{(\mu_j)^{i_j}}{i_j!} \quad \text{si} \quad i_1 + \cdots + i_q = n.$$

On peut alors en déduire la loi de T_n et calculer $\mathbb{P}_{p_0}(T_n \geq a)$. Mais les calculs sont compliqués et numériquement impossibles dès que n est grand. On rappelle donc le théorème 12.8.4.

Théorème 13.4.2. *Si $p = p_0$, alors la suite $(T_n)_{n\geq 1}$ converge en loi vers une variable aléatoire de loi χ^2_{q-1} à $q-1$ degrés de liberté.*