Ex

Test d'indépendance

On veut étudier la liaison entre les caractères : «être fumeur» (plus de 20 cigarettes par jour, pendant 10 ans) et «avoir un cancer de la gorge», sur une population de 1000 personnes, dont 500 sont atteintes d'un cancer de la gorge. Voici les résultats observés :

Observé	cancer	non cancer	TOTAL
fumeur	342	258	600
non fumeur	158	242	400
TOTAL	500	500	1000

Faire un test d'indépendance pour établir la liaison entre ces caractères.

✓

Mise en oeuvre du test :

1. On définit un risque : 5%. Pour étudier la dépendance de ces caractères faisons l'hypothèse H_0 : «les deux caractères sont indépendants » et voyons ce qui se passerait sous cette hypothèse. Notons les événements :
   — C : «avoir un cancer dans la population observée»
   — F : «être fumeur dans la population observée»
   Si les événements F et C sont indépendants, alors : $P(F \cap C) = P(F) \cdot P(C)$ et de même pour les trois autres possibilités : $P(\overline{C} \cap F)$, $P(\overline{C} \cap \overline{F})$, $P(C \cap \overline{F})$, quantités que l'on peut donc calculer sous H_0 :
   $P(F) = \frac{600}{1000}$, $P(C) = \frac{500}{1000}$, $P(F) \cdot P(C) = \frac{3}{10}$, alors l'effectif théorique correspondant à la catégorie «fumeur et cancéreux» est de 300.
2. On en déduit le tableau théorique sous H_0 :

Théorique	cancer	non cancer	marge
fumeur	300	300	600
non fumeur	200	200	400
marge	500	500	1000

3. On calcule alors la valeur de $s = \sum_{i=1}^{i=4} \frac{(O_i - T_i)^2}{T_i}$: on obtient : $s = 34.73$. On a précisé le risque de %, mais pour $\alpha = 0,001$, on lit dans la table du khi-deux à un degré de liberté : $P[\chi^2 \geqslant 10.83] = 0.001$ et le χ^2 calculé est 34.73 !
4. On décide de rejeter H_0. Ainsi, en rejetant l'hypothèse de l'indépendance des caractères «être fumeur» et «avoir un cancer de la gorge», on a moins de une chance sur 1000 de se tromper, puisque moins de un tableau possible sur mille conduit à un calcul de χ^2 plus grand que 10.83 ; beaucoup moins sans doute, conduiraient à un calcul de χ^2 plus grand que 34.73.