

✘ Notre travail est consacré à l'étude des tests de normalité vectorielle.

La loi normale vectorielle est utilisée depuis longtemps pour modéliser des données multidimensionnelles, en particulier en biométrie où Galton et son école l'introduisirent dès la fin du 19^{ème} siècle, en géodésie avec les travaux de Bravais des années 1830, en artillerie à partir de 1870, etc...

Cependant l'étude de la validité des modèles normaux multidimensionnels est très peu traitée dans la littérature statistique et ce n'est que ces dernières années que l'on s'est intéressé à ces problèmes. Les raisons en sont multiples, l'une vient évidemment de la difficulté des calculs mais une autre, plus fondamentale, résulte évidemment du grand nombre de paramètres à estimer dès que la dimension augmente, 5 en dimension 2, 9 en dimension 3, ..., 65 en dimension 10 et des incertitudes diverses qui en résultent.

C'est pourquoi les premières études ont développé des tests graphiques qui ont l'avantage d'être simple et efficace en dépit de leur grossièreté. D'ailleurs le premier test visuel de normalité bidimensionnelle est celui de Galton lui même qui observa directement, comme d'autres avant lui, les ellipses d'égale fréquence et les droites de régression empiriques.

D'autres tests graphiques, plus simples, ont été mis en oeuvre depuis et nous en exposons certains au chapitre 3 de cette thèse. Nous introduisons notamment un test très élégant basé sur l'uniformité de la loi des statistiques de Lhoste en toute dimension.

Mais l'essentiel de notre travail a consisté à étudier un test d'ajustement à une loi normale vectorielle du type du chi-deux.

Comme on le sait le test du χ^2 est bien adapté à l'ajustement à une loi discrète, dès que la loi est continue il faut découper l'ensemble des observations en classes dont le nombre et l'amplitude posent problème.

De plus lorsque la loi dépend de paramètres inconnus qui doivent donc être estimés à partir des données, il se pose un nouveau problème. En effet, si l'on estime les paramètres directement à partir des données, comme cela devrait être si l'on voulait tirer de l'observation le plus d'information possible, la statistique du X^2 ne suit plus asymptotiquement une loi du chi-deux, sa loi dépend même en général des paramètres à estimer. Pour que la statistique X^2 de Pearson suive asymptotiquement une loi du chi-deux après estimation des paramètres, il faut que les estimateurs soient basés sur les données regroupées en classes, c'est-à-dire arrondies plus ou moins arbitrairement.

Ce phénomène observé en 1954 par Chernoff et Lehmann [7], a conduit dans les années 70, divers auteurs (Nikulin [19], [20], Dzapharidze et Nikulin [10], Moore et Spruil [18] ; voir aussi Drost [8], [9] et Nikulin et Greenwood [22]) à s'intéresser à des modifications de la statistique du X^2 de Pearson qui suivent asymptotiquement une loi du Chi-deux après estimation des paramètres à partir des données non regroupées. X