Les statistiques chez Wikipedia

Joseph Saint Pierre

Université de Toulouse

rencontre ingénieurs statisticiens 11 juillet 2013

Une anecdote

Table de contingence :
Comment traiter un tableau croisé 2 par 2 ?

Test de χ^2 de Pearson - Mais pas optimal car manque d'observations

Alternative classique : **test exact de Fisher**

Explications sur ce test sur Internet...

Wikipedia

Première page trouvée avec google :
http://en.wikipedia.org/wiki/Fisher%27s_exact_test
en anglais

mais aussi
http://de.wikipedia.org/wiki/Exakter_Test_nach_Fisher
en allemand
et
http://it.wikipedia.org/wiki/Test_esatto_di_Fisher
en italien
ou
http://eu.wikipedia.org/wiki/Fisherren_froga_zehatz
en basque
etc..

Page wikipedia en français sur le test exact de Fisher

Traduction sommaire, rapide de l'anglais.

Lecture de la page allemande puis de l'italienne.

Ajout de calculs.

Liens sur la page :
`http://fr.wikipedia.org/wiki/Test_du_khi2`
Liens vers d'autres pages :
`http://fr.wikipedia.org/wiki/Loi_hyperg%C3%A9ometrique`

Difficultés

Liens douteux.

Tentatives de corrections.

Regards sur les pages sur les statistiques de wikipedia en français.

Puis sur les probabilités, les mathématiques.

Comparaisons avec les pages en anglais.

Pages en français plus récentes, moins nombreuses, moins consultées.
Pages sur les statistiques en français moins riches et moins structurées que les pages sur les probabilités et le reste des mathématiques.

Encyclopédie libre - logiciels libres etc.

Contact avec **PSPP**.

Analyse sur l'attitude face à Internet.

Remise en cause hiérarchie, centralité.

Difficultés de régulation.

Particularités du français et de la France.

Conclusions provisoires

L'encyclopédie est utile mais imparfaite.

Elle est jeune (2001) et croit rapidement.

Dfauts non dus à l'informatique ni à la liberté.
Découpage en articles.

Qualités essentielles : plusieurs langues.

Une forme de régulation par le grand nombre...