

Résumé conférence "Jugement critique et statistique: la statistique au service de l'émission de jugements critiques"

par Gilles Stoltz, Université Paris Sud - CNRS - HEC Paris

<https://youtu.be/B47CRs6iPL4>

Dans sa conférence, Gille Stoltz mène une réflexion autour de l'utilisation d'éléments de statistique pour la construction d'un esprit critique nécessaire aux citoyens d'aujourd'hui en proposant quelques exemples de la vie quotidienne. Sa présentation repose sur l'explicitation et l'illustration de deux concepts clefs du traitement statistique :

- **l'intervalle de confiance** (notion étudiée au programme de lycée en mathématiques) permet notamment de définir la précision d'un sondage. En effet, puisqu'un relevé exhaustif de la population n'est pas toujours praticable pour différentes raisons, des sondages sont réalisés sur un échantillon réduit : cela induit l'existence d'une marge d'erreur qu'il est nécessaire de prendre en compte. Elle est exprimée par un intervalle de confiance.

Le conférencier illustre cela à travers l'exemple de "*Twitter audit*" qui indique en temps réel le nombre de comptes réels et le nombre de faux comptes qui suivent, par exemple une personnalité politique sur Twitter.

Exemple du *Twitter Audit Report* de Donald J. Trump, extrait de <https://www.twitteraudit.com>



Capture effectuée le 4 octobre 2017 à 9h 20

Le questionnement porte sur le calcul réalisé et la précision des données est confrontée à l'intervalle de confiance qui y est lié : quelle est la marge d'incertitude dont il aurait fallu prendre compte afin de présenter des données précises ? L'étude de *Twitter Audit* porte sur environ 5 000 followers. La marge d'erreur d'un intervalle de confiance au niveau de confiance de 95 % est $1/\text{racine}(5\ 000)$ c'est-à-dire 1,4 %. Pour une fréquence observée de vrais comptes égale à 58,9 % cela donne un intervalle de confiance de la proportion de vrais comptes égal à 58,9 % plus ou moins 1,4 %. Soit un nombre de "vrais" followers compris entre 23,3 et 24,5 millions au niveau de confiance de 95 %. Il apparaît alors que l'intervalle de confiance ne permet pas de proposer une précision aussi fine que celle que *Twitter audit* affiche (23 859 369 "vrais" followers) : la question de la fausse précision des données est alors posée. Cela illustre la citation de l'économiste français Georges Elgozy (1909-1989), "*dans toutes*

statistiques, l'inexactitude du nombre est compensée par la précision des décimales".

Le second exemple porte sur les sondages qui, dans les médias français, sont rarement accompagnés de la marge d'incertitude. Lors d'un sondage, des biais d'échantillonnage et les marges d'incertitude doivent être pris en compte, ce qui est rarement le cas. La plupart des journalistes français ne la comprennent pas, donc ne la reprennent pas et proposent à la place : « nous vous rappelons que les sondages ne sont qu'un instantané de l'opinion et ne préjugent pas de l'élection à venir », ce qui ne quantifie en rien l'incertitude. C'est ce qui explique qu'à un même moment, dans un certain sondage la cote de popularité d'une personne politique aura augmenté alors qu'elle aura diminué dans un autre sondage, comme dans l'exemple ci-dessous dont s'étonne, à tort, le Canard.



Source : Canard enchaîné, août 2010 : Pan sur le bec ! (variations non significatives)

Il existe pourtant des pistes de réflexion comme l'utilisation d'abaques qui permettent de questionner de manière simple la précision des résultats en fonction de l'intervalle de confiance.

- **le test d'hypothèse** (qui n'est pas formalisé au lycée en mathématiques, mais correspond à la notion de prise de décision à l'aide d'un intervalle de fluctuation dès le programme de seconde) permet en statistique d'indiquer si une affirmation concernant une population doit être acceptée ou rejetée avec un certain niveau de risque. En effet, dans une étude statistique, il est nécessaire d'identifier la part du hasard afin de pouvoir conclure sur les résultats de l'étude. A travers deux exemples, le conférencier questionne la part du hasard dans des résultats statistiques : les résultats obtenus sont-ils uniquement imputables au hasard ou est-ce qu'ils sont révélateurs de l'effet étudié ? Ces tests (test de Fisher, test du Chi2...) permettent de calculer la *p-valeur* qui permettra alors de rejeter ou d'accepter l'hypothèse nulle H_0 , hypothèse selon laquelle les résultats observés ne permettent pas de mettre en évidence une différence

significative. Ainsi, si la *p-valeur* obtenue est inférieure à 5%, il sera alors possible de rejeter l'hypothèse nulle et amènera le statisticien à conclure sur l'existence d'une différence significative dans les données qui n'est pas liée au hasard (la différence observée a moins de 5 % de chances d'être due au hasard). Cela est illustré par différents exemples :

- *The Lady tasting tea*, (1935) : Mme Fisher affirmait qu'au goût, elle pouvait indiquer en buvant une tasse de thé au lait, ce qui du thé ou du lait avait été versé en premier dans la tasse ;
- une suspicion de fraude aux élections présidentielles iraniennes en 2009 ;
- les résultats des expériences de Mendel en génétique ;
- l'identification des tricheries, grâce à loi de Benford, dans les bilans comptables d'entreprises suspectées de fraude.

A travers cette présentation, le conférencier a montré que les chiffres ne parlent pas d'eux-mêmes : il faut les faire parler, et la statistique le permet. L'utilisation des intervalles de confiance permet d'expliquer quelles généralisations sont possibles de l'échantillon à la population et de discuter la précision des données d'échantillon.

Les tests d'hypothèses permettent de savoir si l'on peut considérer qu'un fait est établi ou non à partir des données : soit ils le montrent avec un faible risque d'erreur (on rejette l'hypothèse nulle de différence non significative), soit que ce n'est pas le cas car les résultats obtenus peuvent être imputables au hasard (l'hypothèse nulle est tenue pour vraie jusqu'à preuve du contraire). C'est la *p-valeur* calculée par le test qui indique la force de conclusion possible ou de l'absence de conclusion possible.

Souvent, nous pensons que ce sont les chiffres qui dictent la décision à prendre mais ce n'est pas le cas puisque la mission du statisticien est d'éclairer la décision en quantifiant le risque et les marges d'incertitude : c'est au décisionnaire de prendre la décision, d'endosser la responsabilité de l'action et de l'assumer politiquement.