

Enquetes Statistiques en Sciences Sociales et paradoxe de Statistique Inverse	:العنوان
آداب وإنسانيات	:المصدر
الجمعية التونسية للدراسات الأدبية والإنسانية	:الناشر
Marco, Ayoub	:المؤلف الرئيسي
3,4ع	:المجلد/العدد
نعم	:محكمة
2017	:التاريخ الميلادي
99 - 109	:الصفحات
926276	:رقم MD
بحوث ومقالات	:نوع المحتوى
French	:اللغة
AraBase, HumanIndex	:قواعد المعلومات
العلوم الاجتماعية، الدراسات الإحصائية، التحليل الإحصائي	:مواضيع
<a href="http://search.mandumah.com/Record/926276">http://search.mandumah.com/Record/926276</a>	:رابط

# **Enquêtes statistiques en sciences sociales et paradoxe de statistique inversé<sup>1</sup>**

## **Résumé:**

L'utilisation déviée qui peut être faite des statistiques lors de la définition des variables, de la collecte des données, des méthodes de traitement ou de l'interprétation des résultats sont fréquentes aussi que les risques liés à des erreurs, oublis, imprécisions rendent les choses plus difficiles. C'est justement pourquoi on croit en avoir trouvé la conclusion d'une étude statistique, une remarque peut nous faire découvrir que rien n'est résolu. Les paradoxes statistiques qui sont des pièges sur lesquels on tombe bien plus souvent. Mais le plus élémentaire de tous est le paradoxe de statistique inversé dont on imagine des solutions... qui conduisent à d'autres paradoxes!

Sans cesse, des scientifiques que se soit en sciences sociales notamment en sociologie ou en d'autres disciplines académiques tombent dans les pièges qu'il tend. Rarement en sciences sociales paraissent des articles qui essaient de déterminer son sens profond et la façon dont on doit le traiter. Il n'est pas certain ainsi que l'on détienne une solution entièrement satisfaisante pour se libérer de ces conclusions erronées.

## **Mots-clés:**

Paradoxe de Simpson, facteur de confusion, lien cause à effet.

## **Introduction:**

Nous visons à attirer l'attention des jeunes chercheurs sur un des paradoxes de statistique que malheureusement le plus souvent leurs y échappe complètement parce qu'il est non reconnu que par les praticiens du domaine de la statistique. Ce paradoxe est le paradoxe des statistiques inversées dit aussi paradoxe de Simpson ou effet de Yule-Simpson, a été

---

<sup>1</sup> Ayoub MARCO.

décrit par Edward H. Simpson en 1951<sup>1</sup>. IL est souvent rencontré dans la réalité en particulier dans des statistiques de sciences sociales notamment en sociologie et d'autres disciplines académiques. Il rend compte du changement de sens d'une relation entre deux variables lors de l'intervention d'une troisième variable<sup>2</sup>. De plus, Il qualifie une situation où des résultats contradictoires sont obtenus selon que l'on considère le groupe global ou des sous-groupes des participants à une étude. Un bon nombre d'études scientifiques qui se justifient par le moyen d'enquêtes statistiques révèlent souvent l'existence de ce paradoxe comme le justifie Nicholas Falletta<sup>3</sup> ou Gauvrit Nicolas<sup>4</sup>, cela prouve la difficulté d'y penser.

Nous allons présenter dans ce que suit des exemples explicatifs et détaillés afin d'étudier les conditions que les données statistiques doivent être vérifiées pour que ce paradoxe apparaisse, puis examiné les effets de ces conditions pour chaque exemple, nous évaluons par la suite l'incidence de ce paradoxe sur l'interprétation des différentes conclusions et enfin, nous proposons les moyens de prémunition contre ce paradoxe.

### **I- Conditionnement statistique:**

Pour bien expliquer le paradoxe de statistique inversé et comment se fait le conditionnement au niveau des résultats obtenus, on propose l'exemple suivant: On est à l'hôpital entrain de passer des examens et puis les diagnostiques révèlent l'existence d'une tumeur maligne... Le médecin traitant nous explique qu'il faut soigner la tumeur et qu'il existe deux types de traitement: le premier est l'utilisation de la chimiothérapie (les médicaments) et le second traitement est de faire opérer pour enlever la tumeur. La situation est embarrassante, nous laisse perplexes. Quel traitement faut-il choisir? La chirurgie ou les médicaments? On demande l'avis du médecin, il nous montre des statistiques d'une étude qui a été faite à l'échelle nationale<sup>5</sup> et qui recense

---

<sup>1</sup> SIMPSON H. Edward, "The interpretation of interaction in contingency tables", dans Journal of the Royal Statistical Society, Ser. B, 13, p. p. 238-241.

<sup>2</sup> CADOT Martine, *Extraire et valider les relations complexes en sciences humaines: statistiques, motifs et règles d'association*, Université de Franche-Comté, 2006, p. 66.

<sup>3</sup> FALLETTA Nicholas, *Le livre des paradoxes*, Diderot multimédia, 1998.

<sup>4</sup> NICOLAS Gauvrit, *Statistiques, méfiez-vous!*, Ellipses Paris, 2007.

<sup>5</sup> BEN AHMED Saad et al., «Pronostic du cancer du sein chez les femmes tunisiennes: analyse d'une série hospitalière de 729 patientes», Dans Santé Publique, Vol. 14, 2002, p.231-241.

les taux des guérisons des derniers 729 patients de tranche d'âge confondue et qui ont été traitées avec chacune des deux techniques suivantes :

Taux des guérisons

Médicament	Chirurgie
73,5 % (536/729*100)	65,8 % (480/729*100)

L'affaire paraît rassurante, les médicaments sont plus efficaces, on choisira de partir sur le traitement à base de médicaments.

Pour être plus confiant, nous consultons un autre médecin spécialiste. Ce dernier énonce que le choix de la méthode de guérison des tumeurs fait référence non à la technique de traitement (médicament ou chirurgie) mais à la taille tumorale comme le justifie une étude à l'appui qui montre des taux des guérisons séparés par la taille de la tumeur, voici les chiffres:

**Taux des guérisons selon que la taille de la tumeur**

	Médicament	Chirurgie
Grosse tumeur > 2cm	49,1% (60/122*100)	63 % (314/498*100)
Petite tumeur < 2cm	78 % (474/607*100)	89,5 % (166/185*100)

Il est évident de cette enquête que la chirurgie est efficace pour les petites et les grosses tumeurs.

La question reste indécise, une première étude indiquant que les médicaments sont les plus sollicités que l'intervention chirurgicale et une deuxième étude prouve le contraire, c'est la chirurgie qui marche le mieux pour les petites et les grosses tumeurs. En comparant les deux études tout en en se fixant les statistiques pour arriver à se décider du choix de traitement à suivre, on réalise une découverte étonnante: ce sont exactement les mêmes chiffres, ils ne forment qu'une seule et même étude! Démonstration:

**Taux des guérisons selon la taille de la tumeur**

	Médicament	Chirurgie
Grosses tumeurs	49% (60/122*100)	63% (314/498*100)
Petites tumeurs	78,4% (476/607*100)	89,5% (166/185*100)

---

TOTAL	73,5% (536/729*100)	65,8% (480/729*100)
-------	---------------------	---------------------

---

On voit bien au global les médicaments ont un taux de guérison meilleur, mais si on sépare par type de tumeur c'est la chirurgie qui a toujours un taux de guérison meilleur.

Alors nous demandons comment c'est possible? Comment partons-nous des mêmes chiffres, on peut au global conclure que les médicaments sont plus fiables mais sur chaque type de tumeur prise séparément on arrive à conclure que c'est la chirurgie qui est plus efficace?

Ce que nous voyons c'est juste des vrai chiffres issus d'une même enquête statistique et nous sommes face à un paradoxe de statistique qu'on appelle le paradoxe de statistique inversée et c'est un paradoxe qui prouve que certaines données statistiques se conditionnent suivant la manière dont on les analyse, on peut aboutir ainsi à une conclusion ou à la conclusion opposée. Donc là le problème, c'est qu'on vraiment envie de savoir quel traitement marche le mieux, est-ce les médicaments ou la chirurgie? Est-ce qu'il faut croire à la première manière de présenter les chiffres ou plutôt la deuxième?

En faite, la vérité c'est la chirurgie est la plus efficace dans ce genre de dilemme, le fait que les médicaments montrent des taux de guérison plus élevés en moyenne, c'est justement une apparence. Pour déchiffrer ça, il faut retourner dans les détails des chiffres. Primordialement, on doit constater qu'il y a deux constatations importantes à dégager; la première, les grosses tumeurs ont des taux de guérison plus faibles que les petites tumeurs (constatation évidente). La deuxième, c'est qu'on a une grosse tumeur, on utilise beaucoup plus souvent la chirurgie que les médicaments. Comment surpasser cet embrouillement? La réponse se situe dans le savoir faire médical; Principalement la chirurgie fonctionne mieux sauf que méthodologiquement dès qu'on a un cas plus au moins difficile on utilise plus souvent la chirurgie que les médicaments donc cela va se réfléchir sur les données statistiques. Pour cette raison que globalement le taux de guérison grâce à la chirurgie est moins élevé de celui de la guérison par les médicaments parce qu'on lui a réservé les cas difficile.

## **2- Le facteur de confusion:**

Pour comprendre comment se produit ce genre de situation au niveau statistique, il faut faire appelle à une notion qu'on appelle le facteur de

confusion, c'est-à-dire un facteur non causal mais statistiquement lié à un facteur réellement causal non identifié<sup>1</sup>. Le but et d'essayer d'étudier un lien de cause à effet entre le choix du traitement et la guérison, alors la probabilité de guérison elle dépend du choix de traitement et dépend bien sûr sur un tas d'autres facteurs et notamment la taille de la tumeur, sauf que dans les données relatives à la taille de la tumeur ont aussi joué sur le choix des traitements (chaque fois quand on a une grosse tumeur on a plutôt tendance à utiliser la chirurgie que les médicaments) et donc on a un facteur externe : la taille de la tumeur joue à la fois sur la conséquence mais aussi sur la cause qu'on veut étudier et donc c'est ça qu'on appelle facteur de confusion et c'est qu'on est dans ce genre de situation que le paradoxe de Simpson peut se produire.

Pour porter plus d'éclairage sur ce paradoxe, on va prendre un deuxième exemple intéressant qui est le redoublement scolaire. Pratiqué en Tunisie dès l'école primaire, le redoublement scolaire alimente régulièrement les discussions au sein du cadre éducatif. Le redoublement pour certains professeurs est synonyme d'une nécessaire consolidation des connaissances de l'élève, tandis que pour les autres il est vu comme une perte de temps. Parmi ces protagonistes l'un disait la chose suivante: «... Le redoublement en Tunisie ça sert à rien, la preuve les élèves qui ont redoublé pendant leur scolarité<sup>2</sup>, au final au Baccalauréat, ils ont des notes moins bonnes que les autres... Le redoublement n'est pas efficace, les pays dont la France par exemple qui font le plus largement redoubler leurs élèves, sont aussi ceux qui ont les moins bons résultats scolaires<sup>3</sup>.

Alors on veut bien voir que les questions des bienfaits du redoublement soit compliquée mais on est sûre que cet argument est erroné parce que les élèves qu'on fait redoubler en générale sont ceux qui ont une difficulté scolaire et ce n'est pas difficile de constater qu'à la fin ces élèves ont des notes moins bonnes que les autres aux épreuves de Baccalauréat, et si on regarde finement et attentivement la question du redoublement, on voit qu'on est aussi

---

<sup>1</sup> DODGE Yodolah, *Statistique, dictionnaire scientifique*, Springer Science & Business Media, 2007, p. 68.

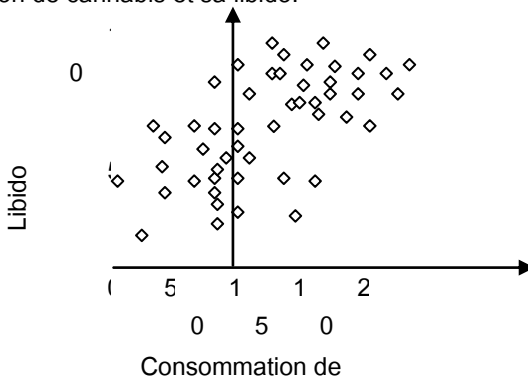
<sup>2</sup> Les taux moyens de redoublement entre l'année scolaire 2001 et 2012 sont respectivement de 7,09% au cycle primaire, 16,97% au cycle préparatoire et 15,46% au cycle secondaire selon YAAKOUBI Mahmoud, «Promotion de l'emploi productif et le travail décent des jeunes en Tunisie» 2013, in [http://www.adapt.it/adapt-indice-a-z/wp-content/uploads/2014/09/oit\\_analyse\\_syst%C3%A9me\\_%C3%A9ducatif\\_tunisien\\_2013.pdf](http://www.adapt.it/adapt-indice-a-z/wp-content/uploads/2014/09/oit_analyse_syst%C3%A9me_%C3%A9ducatif_tunisien_2013.pdf).

<sup>3</sup> C'est ce que rappelle l'OCDE en juillet 2011 dans un focus, établi à partir de l'enquête PISA 2009 (publiée en 2010)

dans une situation où on a un facteur de confusion. C'est qu'on essaye de faire, c'est étudier le lien de cause à effet entre la faite de redoubler ou pas et les notes au Bac. En effet, On a un facteur qui est le niveau scolaire de l'élève qui va jouer sur ses notes au Bac mais qui va aussi jouer sur le faite qu'on choisi de faire redoubler ou pas, donc le niveau scolaire de l'élève est un facteur de confusion et si on l'oublie, on aboutit à une conclusion qui est erronée au plan statistique qui est que faire redoubler les élèves ça fait baisser leurs notes au Bac!

Pour bien montrer que le facteur de confusion et donc le paradoxe de statistique inversée peuvent se cachés un peu partout dans les statistiques, nous passons au troisième exemple inspiré de la publication du Groupe Pompidou<sup>1</sup> qui décrit plusieurs enquête sur la consommation de la drogue en Tunisie entre 2012 et 2014. Imaginons qu'un chercheur étudie le lien qui existe entre la consommation du cannabis des jeunes dans les banlieues et leurs libidos<sup>2</sup>.

Le chercheur va prendre un échantillon représentatif de la population, fixe la tranche d'âge cible et puis il va soumettre un questionnaire aux personnes choisies où il doit évaluer leurs consommation mensuelle de cannabis (Zatla) et puis noter leur libido sur une échelle de 1 à 10. Après avoir collecté les réponses, il analyse les données et puis il décide de les faire présentées sur un graphique qui va montrer la libido en fonction de la consommation de cannabis (nombre des joints). Alors, sur ce graphique ci-dessous chaque point représente une des personnes qui a été interrogée et montre sa consommation de cannabis et sa libido:

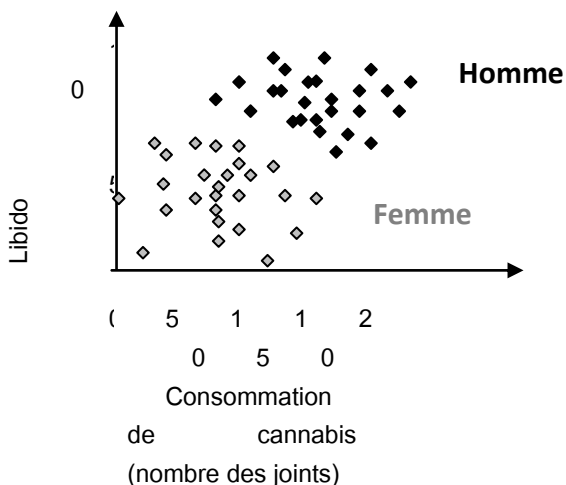


<sup>1</sup> AOUNALLAH Hajer et al., «Situation et politique en matière de drogues», Groupe Pompidou, Volume 12, 2014. 38 page

<sup>2</sup> Toute forme d'énergie psychique, quel que soit son objet. Ici on s'intéresse aux libidos sexuelles.

### cannabis (nombre des joints)

On constate qu'il semble avoir une belle corrélation entre les deux variables, c'est-à-dire que plus on consomme de cannabis plus on a une libido élevée, c'est parfait, le chercheur est satisfait de sa conclusion. Soudain on lui demande qu'est-ce qu'il se passe si on regarde qu'elles sont les hommes et les femmes sur ce graphique? Il reprenne ces données et puis il retrace un deuxième graphique en représentant différemment les hommes et les femmes et voici le résultat:



En effet, chez les femmes la consommation de cannabis n'a pas franchement une influence sur la libido même plutôt une influence négative et puis chez les hommes il n'y a pas vraiment d'influence non plus.

On voit clairement que si on différencie par sexe, on n'aura pas du tout une corrélation entre la consommation de cannabis et la libido et pourtant si on oublie de différencier l'échantillon par sexe on a l'impression qu'on a une corrélation et donc on est encore dans le même genre de situation. C'est-à-dire on essaye d'étudier le lien cause à effet entre consommation de cannabis et libido et on a un facteur de confusion qui est le sexe des enquêtés qui joue à la fois sur la consommation de cannabis puisque les hommes ont tendance de fumer plus<sup>1</sup> et sur leurs libidos parce que les hommes ont tendance d'avoirs

<sup>1</sup> Le taux des fumeurs en Tunisie est de 61.9% des hommes contre 7,7% des femmes selon le rapport de la République Tunisienne «Etude du secteur de la santé»,



des libidos plus élevées<sup>1</sup>. L'intervention d'une troisième variable, ici le "sexe", contredit donc la conclusion tirée précédemment sur la relation entre le cannabis et la libido. Les chercheurs en sciences humaines doivent connaître ce problème qui fait que toute conclusion sur des relations entre variables peut être remise en cause par l'intervention d'une variable "oubliée".

### **3- La manière de se prémunir contre le paradoxe de statistique inversé:**

Pour se prémunir contre ce paradoxe deux solutions sont envisageables: la première solution, c'est il faut savoir qu'il y a toujours un facteur de confusion lorsqu'on manipule des variable statistique. Il faut y avoir pensé et il faut analyser les données en conséquence. C'est comme nous avons fait dans le cas de la tumeur, on a analysé les données par taille de la tumeur et c'est ce qu'on a fait dans le cas de la consommation de cannabis et de libido, on a différencié l'analyse par sexe. Mais ce n'est pas toujours évident, parce qu'il faut connaître le facteur de confusion, par exemple dans le cas de redoublement c'était un peu évident parce que le piège est quasiment clair, dans le cas de cannabis et de libido s'était beaucoup moins évident de penser au facteur de confusion et puis dans le cas des tumeurs si le chercheur n'est pas un spécialiste du sujet, il est impossible pour y penser.

A l'heure actuelle des «big data<sup>2</sup>», de «l'open data<sup>3</sup>», de cette formidable disponibilité des données quantitatives et l'accès aux outils d'analyse, il n'y a plus de raisons qui justifient un rejet, une ignorance ou une méfiance des méthodes quantitatives. Elles doivent devenir, plus qu'avant, une composante

---

Département du développement humain, région Moyen-Orient et Afrique du Nord, La Banque mondiale Washington, DC. MAI 2006.

<sup>1</sup> Les études les plus récentes nous montrent par exemple que les hommes pensent 284% de fois plus souvent à la sexualité que les femmes (Laumann, 1994), qu'ils se masturbent 351% davantage qu'elles (Laumann, 1994), et qu'ils prennent l'initiative de la relation sexuelle à peu près 200% plus souvent (O'Sullivan, Byers, 1992). Ils souhaiteraient avoir davantage de partenaires 400% plus souvent que les femmes (Mac Cabe, 1987), et recherchent la performance deux fois plus souvent qu'elles (Colson, 2006), tout en avouant être excités par un film érotique trois fois plus facilement (Colson, 2006).

<sup>2</sup> Ensembles de données qui deviennent tellement volumineux qu'ils en deviennent difficiles à travailler avec des outils classiques de gestion de base de données ou de gestion de l'information.

<sup>3</sup> L'open data ou donnée ouverte est une donnée numérique d'origine publique ou privée.

essentielle de l'analyse en sciences sociales. Mais le problème lorsque des chercheurs initiés analysent et traitent ces données tout en oubliant l'existence des paradoxes statistiques sans parler de ceux qui sortent régulièrement du périmètre des «chiffres» pour entrer dans des considérations théoriques et qualitatives afin de redonner ensuite un sens aux chiffres. Ceux-ci nous montrent que lorsqu'on regarde des données ou des variables statistiques, qui relèvent des sciences sociales, on ne peut pas se passer de s'interroger sur la manière dont les conclusions ont été déduites, que l'auteur de ses statistiques sait de quoi il parle et qui connaît vraiment le domaine dont il est question.

L'autre solution pour qu'on soit sûr de ne pas tomber dans le paradoxe de Simpson, c'est de faire une expérience. Dans tous les exemples que nous avons étalés précédemment, à chaque fois on utilise des données qui existent déjà, c'est qu'on appelle une étude rétrospective. Dans ce type d'étude, on interroge des individus sur ce qu'ils ont fait ou vécu au cours des dernières semaines ou année ou, encore, tout au long de vie écoulée depuis un âge donné ou événement origine<sup>1</sup>, c'est le regard scientifique sur le passé. Alors que si on craie de tout pièce une expérience dont on va analyser les données, on fait ce qu'on appelle une étude prospective. C'est-à-dire on suit plus au moins longtemps un échantillon de population en le visitant à intervalle régulier pour enregistrer les changements survenus depuis la visite précédente<sup>2</sup>; c'est un peu le regard scientifique sur l'avenir. Rappelons que le facteur de confusion pose problème parce qu'il influe à la fois sur la cause et la conséquence d'un phénomène sociale qu'on souhaite l'étudier. La solution c'est d'éliminer le lien entre le facteur de confusion et la cause. Prenons l'exemple suivant: si on revient au cas du redoublement, qu'est-ce qu'il faudrait faire pour prouvez à coup sûr si le redoublement c'est efficace ou pas? C'est prendre une classe, la coupée en deux au hasard et dire la première moitié on la faire doublée et la deuxième moitié on la fait passée. Si on fait ça, dans chacun des deux groupes on aura des élèves de tous les niveaux et donc à la fin en comparant les notes au Bac de ces deux groupes, on pourra savoir s'il y a un impact positive du redoublement sur les notes au Bac ou pas. Le problème de ce type de procédé c'est que faire une expérience de telle sorte va être très compliquée, long et peut poser des problèmes d'éthiques mais pourtant c'est le seul moyen de porter à coup sûr un lien entre une cause et

---

<sup>1</sup> CASELLI Graziella, VALLIN Jacques, *Observation, méthodes auxiliaires, enseignement et recherche*, Guillaume Wunsch INED, 2006, p. 42.

<sup>2</sup> *Idem*.

un effet. D'ailleurs, c'est pour ça lorsqu'on développe un médicament, à un moment donnée pour prouver l'efficacité du médicament on doit toujours faire une étude prospective et c'est pour ça que développer un médicament ça coute cher et ça prend du temps parce qu'il y a toujours des tests très long à faire mais qui sont absolument nécessaire pour être sûre de se prémunir de tous les effets des facteurs de confusion.

Le chercheur en sciences sociales doit se méfier des statistiques d'une manière intelligente c'est-à-dire quand on montre des chiffres, il faut toujours demander si ses chiffres proviennent d'une étude rétrospective, c'est-à-dire qu'on a analysé des chiffres existants ou s'ils viennent d'une étude prospective, on a fait l'expérience. S'ils viennent d'une étude rétrospective, il faut faire attention, il peut y avoir des facteurs de confusion, on les connaît pas forcément et on peut conclure des choses qui sont complètement fausses. Par contre si c'est une étude prospective alors là il y a un peu plus de chance que les résultats sont fiables. A titre d'exercice on peut faire une étude prospective dans le cas de la consommation de cannabis et de libido. Justement, il faudrait prendre un groupe de personne (échantillon) et puis on attribuerait au hasard pour chaque personne un niveau de consommation de joint de cannabis compris par exemple entre 0 et 20 joints par mois, les personnes y tiennent tout au long de se mois et à sa fin ils notent leur libido sur un échèle de 1 à 10. On est certain qu'on a éliminé ainsi tout facteur de confusion et on sera vraiment si la consommation de cannabis à un effet sur la libido.

## **Liste bibliographique:**

- SIMPSON H. Edward, "The interpretation of interaction in contingency tables", dans Journal of the Royal Statistical Society, Ser. B, 13, p. p. 238-241.
- GAUVRIT Nicolas, Statistiques, méfiez-vous !, Ellipses Paris, 2007, 204 pages
- CASELLI Graziella, Jacques Vallin, Observation, méthodes auxiliaires, enseignement et recherche, Guillaume Wunsch INED, 2006, 779 pages.
- ADUNALLAH Hajer et al., «Situation et politique en matière de drogues», Groupe Pompidou, Volume 12, 2014, 38 pages.
- CADOT Martine, Extraire et valider les relations complexes en sciences humaines : statistiques, motifs et règles d'association, Université de Franche-Comté, 2006, 299 pages.
- FALLETTA Nicholas, Le livre des paradoxes, Diderot multimédia, 1998, 200 pages.
- Rapport de la République Tunisienne «Etude du secteur de la santé», Département du développement humain, région Moyen-Orient et Afrique du Nord. La Banque mondiale Washington, DC, MAI 2006, in <http://ettajdid.org/IMG/pdf/Tunisia-Sante-BM-FR-May2006.pdf>.
- YAAKOUBI mahmoud, «Promotion de l'emploi productif et le travail décent des jeunes en Tunisie» 2013, in [http://adapt.it/adapt-indice-a-z/wp-content/uploads/2014/09/oit\\_analyse\\_syst%C3%A9me\\_%C3%A9ducatif\\_tunisien\\_2013.pdf](http://adapt.it/adapt-indice-a-z/wp-content/uploads/2014/09/oit_analyse_syst%C3%A9me_%C3%A9ducatif_tunisien_2013.pdf).
- BEN AHMED Saad et al., «Pronostic du cancer du sein chez les femmes tunisiennes : analyse d'une série hospitalière de 729 patientes», Dans Santé Publique, Vol. 14, 2002, p. p. 231-241.
- DODGE Yodolah, Statistique, dictionnaire scientifique, Springer Science & Business Media, 2007, 662 pages.