Les Statistiques portent-elles atteinte à la liberté?

Pour un regard critique sur les informations chiffrées

François Sauvageot
Maître de conférences à l'Université Paris 7
Animateur à l'IREM-Paris 7

A propos des sondages

- Arnold Schwarzenegger, pourtant crédité de 26% des intentions de vote, vient d'être élu gouverneur de la Californie avec 49% des voix.
- Tous les sondages donnaient un classement J. Chirac, L. Jospin, J-M. Le Pen au premier tour en 2002.
- F. Bayrou aurait gagné 1 point dans les sondages grâce à une gifle …

Première idée forte : les intervalles

- En statistiques, on travaille avec des intervalles, pas avec des nombres. On dit par exemple : entre 14% et 20%.
- Attention, ça ne veut certainement pas dire 17% ! Il se peut que le « pic » de probabilité soit à 16%, autrement dit que les chances **ne soient pas uniformes** dans l'intervalle donné.
- **Aucun** sondage donnant des nombres bruts n'est crédible. Aucun ! En particulier F. Bayrou n'a certainement rien gagné avec sa gifle : statistiquement, il ne s'est **rien** passé.

Seconde idée forte : des groupes, pas des classements !

- La conséquence de ce qui précède est qu'on peut prédire des tendances et des groupements, mais pas prédire un classement. Pas même avec de grosses chances : c'est impossible !
- Par exemple, lors des élections présidentielles de 2002, les sondages ont en fait **<u>toujours</u>** donné trois groupes : un trio de tête, cinq outsiders, un groupe de queue.
- Il n'a **<u>jamais</u>** été possible d'écarter la possibilité que J-M. Le Pen arrive au second tour, ni que J. Chirac n'y parvienne pas. Alors que tous les médias ont parlé du troisième homme (en fait, un des cinq outsiders), ils ont omis de voir qu'il y avait déjà trois prétendants !

Loi des grands nombres

- Soit une suite (X_n) de variables aléatoires **indépendantes**, de mêmes moyenne μ (espérance) et variance, alors (X_n) converge en moyenne vers μ, au sens des probabilités :

(En fait on a même convergence presque sûre.)

- En particulier, le théorème s'applique lorsque les X_n suivent la même loi (on dit qu'elles sont indépendantes et identiquement distribuées : i.i.d.)
- On prendra garde à ne pas croire que, pour autant, les erreurs se compensent : la probabilité que la moyenne des X_n vaille μ ne tend pas vers 1 !

Théorème Central Limite

- Ce théorème précise la vitesse de convergence dans la loi des grands nombres : $1/\sqrt{n}$.
- Soit une suite (X_n) de variables aléatoires **indépendantes**, de mêmes moyenne μ (espérance) et variance σ^2, alors la moyenne des (X_n), centrée et réduite, admet une limite qui suit une loi normale $\mathcal{N}(0,1)$, centrée réduite, i.e.
- En particulier, ce théorème est vrai lorsque les (X_n) sont i.i.d. et admettent des moments d'ordre 1 et 2 (espérance et variance).

Approximations

- Le Théorème Central Limite est un théorème de convergence (en loi), pas d'approximation. Néanmoins on l'utilise comme tel.
- Pour $n \geq 30$, on peut en pratique faire les approximations :
- Et même pour $n \geq 10$ lorsque la loi initiale est continue, régulière et symétrique.

Intervalles de confiance

- Soit X une variable aléatoire suivant une loi normale centrée réduite et α un nombre compris entre 0 et 1 (typiquement 95%). Il existe deux réels t et u tels que

$$P(-t \leq X \leq t)=\alpha \quad \text{et} \quad P(X \leq u)=P(X \geq -u)=\alpha$$

- Ces quantités se trouvent sur des tables de la loi de Gauss, et servent de base à la construction d'intervalles de confiances bilatéraux ou unilatéraux.
- On a par exemple

α	90%	95%	99%
t_α	1.64	1.96	2.57
u_α	1.28	1.64	2.33

Estimation d'une proportion

- On modélise le problème ainsi : X est la variable aléatoire qui à un individu associe 1 s'il possède le caractère qualitatif étudié et 0 sinon. Si la proportion d'individus ayant ce caractère est *p*, alors X suit une loi de Bernoulli de paramètre *p*, notée $\mathcal{B}(p)$

$$P(X=0)=1-p \quad \text{et} \quad P(X=1)=p$$

- On a $\mu=p$ et $\sigma^2=p(1-p)$. Si (X_n) est une suite de variables aléatoires i.i.d. selon $\mathcal{B}(p)$, et *f* la moyenne des X_n :
- D'après la loi des grands nombres f_n, la « fréquence empirique », tend vers *p* lorsque *n* tend vers l'infini. Attention ! f_n est une variable aléatoire …
- De même on peut démontrer que tend vers σ^2. Tout comme f_n, est une variable aléatoire.
- Un intervalle de confiance pour *p* au niveau de confiance α est donné par

Estimateurs sans biais

- La fréquence empirique f_n est un estimateur de p en ce sens que f_n tend vers p en loi. Mais c'est aussi un estimateur sans biais, en ce sens que, pour tout n, l'espérance de f_n est p :
- C'est l'explication du *n-1* dans la formule de l'estimateur de la variance. On parle également de « correction pour les petites populations ».
- Cette correction s'interprète en termes de degrés de liberté : il en faut *1* pour estimer μ et *n-1* pour estimer σ^2.

Application numérique

- Dans un sondage d'opinion, on enregistre 14% des intentions de vote pour A et 17% pour B. En fait il faudrait dire (avec 850 sondés)

  « Avec une certitude de 95%, les intentions de vote sont,
  entre 11.7% et 16.3% pour A
  et
  entre 14.5% et 19.5% pour B. »
- On ne peut donc pas exclure que A récolte plus de suffrages que B, même si les électeurs ne changent pas d'opinion d'ici les élections …
- Avec une confiance de 90%, les intervalles deviennent [12%,16%] et [14.9%,19.1%], ce qui laisse encore de la marge. Il faut arriver à 77% de confiance pour que les intervalles soient disjoints : [13.6%,15.4%] et [15.4%,18.6%].

Quel enseignement en tirer ?

- Vouloir ordonner totalement est sans objet.
- En particulier le classement des lycées donné par les médias est **ridicule** : on peut au mieux donner, disons cinq groupes de lycées et dans chaque groupement les lycées sont **indiscernables** statistiquement parlant.
- Il en va de même du classement des hôpitaux.

Du vrai, du faux et de la manipulation

- C'est un phénomène bien connu des physiciens, toute mesure perturbe le phénomène observé. Autrement dit tout sondage d'opinion est faux au moment même où il est rendu public !
- Vouloir informer à tout prix, c'est très souvent, dans ce domaine, transmettre des mensonges à très grande échelle. Car une information erronée ou partielle n'est pas un début de vérité, mais un mensonge.
- Un exemple : Google, devenu en trois ans une référence (53% des requêtes mondiales, soit 200 millions par jour), est fondé sur l'algorithme suivant : une page est d'autant plus pertinente qu'elle est fréquemment citée par d'autres.
- En conséquence, pour devenir LA référence du web sur un sujet il suffit d'être cité par de nombreux sites.

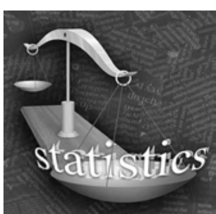
Exemples de nouvelles références créées par Google

- Ainsi en cherchant sur google « massive destruction weapons » on tombait sur une page humoristique ironisant « nous n'en avons pas trouvé dans votre pays, il doit y avoir une erreur, veuillez recommencer la recherche » …
- Un slogan anti-guerre a été reformulé et effacé d'Internet par un expert en stratégie économique de Harvard en 42 jours. Il a juste proposé une autre définition que celle utilisée par Kofi Annan (expliquant qu'il y a deux superpuissances sur la planète : les Etats-Unis et l'opinion publique) et est devenu LA référence d'Internet, au point que la déclaration de Kofi Annan a quasiment disparu. (http://www.uzine.net/article1960.html)

Quelques questions

- Quelles autorités peuvent avoir d'informations de nature statistique?
- Peut-on se servir d'une probabilité comme élément de preuve lors d'un procès?
- Des indicateurs statistiques peuvent-ils servir de fondement à une décision politique?
- Comment peut-on se servir de statistiques en médecine ou en biologie ?

Statistiques et justice



L'affaire Sally Clark

- Le 9 novembre 1999, au Royaume-Uni, Sally Clark est accusée d'avoir tué ses deux enfants :
  - Christopher, âgé de 11 semaines, le 13 décembre 1996 à 21h35.
  - Harry, âgé de 8 semaines, le 26 janvier 1998 à 21h27
- Le 2 octobre 2000, son appel est rejeté.
- Le 29 janvier 2003, elle est libérée en second appel.

Christopher : les faits médicaux

- Christopher est d'abord reconnu comme étant mort de mort naturelle, due à une infection des voies respiratoires inférieures.
- Lors du procès le médecin ayant pratiqué l'autopsie revient sur son diagnostic et assure qu'il n'y a jamais eu aucun signe d'infection et que le bébé a été étouffé.
- Il se base sur des tests sanguins qu'il a menés à l'époque de la mort de Christopher mais dont il n'a pas fait référence dans son rapport post-mortem.

Christopher : les conclusions des experts du procureur

- Dr Williams : étouffement
- Pr Meadow : ni une infection des voies respiratoires inférieures, ni une « Mort Subite du Nourrisson » mais pas non plus une mort naturelle.
- Pr Greene & Dr Keeling : raison inconnue, mais cela pourrait être non naturel.

Christopher : les conclusions des experts de la défense

- Pr Berry & Dr Rushton : raison inconnue, c'est-à-dire une maladie non détectée, une mort non naturelle ou une « Mort Subite du Nourrisson ». Le rapport d'autopsie est trop incomplet pour donner lieu à des certitudes.
- Pr David : hémosiderosis pulmonaire idiopathique, mais une suffocation est possible.

Harry : les faits médicaux

- Des hémorragies (yeux, paupières et moelle épinière)
- Des lésions dans le cerveau
- Dislocation d'une côte
- Le rapport d'autopsie envisage que la mort a été provoquée par de violentes secousses, répétées plusieurs fois.
- Néanmoins la quasi-totalité des éléments énoncés dans le rapport d'autopsie sont sujet à caution, certains étant démontrés comme des artefacts (de la même façon que des **centaines** de jugements basés sur une analyse ADN sont en train d'être cassés car ils sont entachés d'erreurs de manipulation d'un laboratoire du FBI qui rend les tests totalement inutilisables).

Harry : les conclusions des experts du procureur.

- Dr Williams : secousses violentes et répétées.
- Pr Meadow : ni une « Mort Subite du Nourrisson », ni une mort naturelle, probablement un étouffement.
- Pr Greene & Dr Keeling : raison inconnue, mais cela pourrait être non naturel.

Harry : les conclusions des experts de la défense

- Pr Berry : si les preuves étaient sûres, la mort serait due à un traumatisme, mais ça pourrait être un accident.
- Dr Whitwell & Dr Rushton : raison inconnue, mais pas une « Mort Subite du Nourrisson ».
- Pr David : rapport d'autopsie trop incomplet pour donner lieu à des certitudes.

L'argument du Pr Meadow

« La probabilité que les deux nourrissons soient morts d'une Mort Subite du Nourrisson est très très faible, 1 chance sur 73 millions. C'est comme si un outsider côté à 80 contre 1 gagnait 4 années de suite le Grand Prix National »

(Le Grand Prix National est un concours hippique.)

Les bases de l'argument

- Le rapport du CESDI, intitulé SUDI, sur les morts inexpliquées chez les nourrissons, incluant les MSN, donne la proportion de nourrissons mourant de MSN dans une famille de non-fumeurs, ayant un revenu et dont la mère est âgée de plus de 26 ans, à savoir 1/8 543.
- Au Royaume-Uni il naît environ 650 000 enfants chaque année.
- Comme (1/8543)*(1/8543)≈1/73 000 000, cela fait qu'une famille environ par siècle ait à subir deux MSN, au Royaume-Uni.

Les conclusions du procès et du premier appel

- Les faits accablent l'accusée.
- Les bébés sont morts au même âge, dans des circonstances similaires (dans un relax, juste après avoir mangé, vers 21h30, lorsque la mère était seule avec l'enfant et que le père était au travail ou allait s'absenter).
- Il y avait des signes de maltraitance : tentative d'étouffement pour Christopher et fracture d'une côte pour Harry.
- Il y avait des signes de blessures récentes.
- La rareté de deux telles morts naturelles et la coïncidence extraordinaire que ce serait d'observer simultanément des maltraitances récentes et anciennes.

Le « sophisme du procureur »

- C'est un paradoxe très largement débattu et dont il faut se méfier dès que l'on manipule des probabilités.
- L'argument fallacieux consiste à confondre « la probabilité qu'un évènement survienne » avec « la probabilité qu'un évènement soit survenu dans un cas étudié ».
- Ici c'est confondre « il y a une chance sur 73 millions qu'un double décès naturel survienne » avec « dans le cas d'un double décès, il y a une chance sur 73 millions qu'il soit naturel » …
- Être innocent d'un crime, c'est se trouver dans la situation où 1° une mort est survenue et 2° un évènement extérieur (par exemple une mort naturelle inexpliquée) s'est produit.

Les probabilités conditionnelles

- Par conséquent pour évaluer la probabilité d’innocence (I) de Sally Clark, il faut évaluer la chance qu’il y ait un double décès inexpliqué (A) **sachant qu’il y a un double décès** (D).
- Autrement dit chercher **le nombre de fois qu’un évènement rarissime se produit dans une population très restreinte** (celle de ceux qui ont subi un double décès) et non <u>le nombre de fois qu’il se produit au sein de la population totale</u>.

Du rare au très rare

- Quand on analyse des évènements rares l’important est de comparer le rare au très rare. Et pareillement avec les évènements très probables.
- Une chose évidente qui n’a pas été faite était de calculer de la même façon le pourcentage de mères infanticides et récidivistes !
- Ainsi au Royaume-Uni environ 30 enfants sont tués par leur mère chaque année, un nombre comparable au nombre de MSN par an, ce qui fait que les deux hypothèses sont en fait à peu près aussi probables …

Un calcul sordide

- Nous prenons deux hypothèses : (A) les enfants de Sally Clark sont morts par accident, (M) Sally Clark les a tués ; et nous négligeons les autres possibilités, autrement dit :
- En particulier
- En utilisant la définition des probabilités conditionnelles
- D’où
- Nous prenons ici : 1/1000 comme probabilité pour une MSN et 1/100 pour une deuxième MSN, soit 1/100000 pour A. Avec 30 infanticides par an et 650000 naissances par an, on arrive à 1/20000 comme probabilité pour un infanticide et on prend 1/200000 pour un double infanticide. Ce qui donne

P(I)=1/(1+1/2)=2/3.

A-t-on le droit de transformer la question posée « Sally Clark est-elle coupable ? » en une question quantifiée ?

- Assurément, non. L’indication statistique n’a pas d’autre vertu que de donner à réfléchir, stigmatiser les problèmes.
- Elle n’en résout aucun : c’est à l’être humain, par son analyse, de chercher les réponses.
- Idéalement, c’est en cherchant à comprendre la statistique que l’on peut découvrir des clefs que l’on avait omises …

Quel était l’objet du rapport SUDI dont s’est servi le Pr Meadow ?

- Il s’agit d’un rapport en vue de la prévention du risque de MSN, notamment d’un deuxième tel évènement dans une même famille.
- La probabilité 1/8 543 est donnée en comparaison avec la probabilité 1/ 214 obtenue avec des parents fumeurs, sans revenu et avec une mère âgée de moins de 26 ans.
- Notamment, le rapport explicite : « Ainsi, pour les familles comportant plusieurs facteurs de risques pour la MSN, une seconde MSN, bien que très rare, est 1 600 fois plus probable que dans les familles dans ces facteurs. »

Les erreurs

- De façon assez remarquable, personne n’a jamais prétendu que les morts de Christopher et Harry étaient des MSN, puisqu’à chaque fois on a cherché à trouver des explications et qu’on a trouvé des éléments de réponse (par exemple Harry est mort 4 heures après avoir été vacciné). Le calcul même du Pr Meadow est sans objet !
- Le rapport 1/73 000 000 est erroné : les deux événements (1ère MSN et 2ème MSN) n’ont aucune raison d’être indépendants, et ont plutôt de nombreuses raisons de ne pas l’être !
- L’image du concours hippique est erronée car il n’y a pas 650 000 Grand Prix par an, mais un seul et on ne peut pas faire des statistiques sur un seul évènement !

Pourquoi Sally Clark a-t-elle été libérée ?

- Contrairement à ce qu’on pourrait penser ou espérer, ce n’est pas grâce aux arguments précédents.
- Il a en fait été prouvé que les autopsies ont été pratiquées en dehors des normes de rigueur, notamment quant à la transmission des faits observés.
- Des analyses sanguines demandées pour Harry lors de l’autopsie avaient des résultats alarmants qui n’ont pas été communiqués. Seule la persévérance du mari et du père de Sally Clark a permis de les découvrir. On peut penser à partir de ces analyses qu’Harry est mort de méningite ou d’un autre type de maladie fulgurante liée au staphylocoque doré.
- **Néanmoins suite à un nouveau cas du même genre en juin dernier, le procureur général, Lors Goldsmith, a annoncé la révision de 258 procès intentés pour infanticide, dont ceux de 54 parents encore emprisonnés. Le professeur Meadow, ancien président du Collège royal de pédiatrie, est à l’heure actuelle l’objet d’une enquête de l’Ordre britannique des médecins, pour faute professionnelle.** (Libération, daté du 22 janvier 2004).

Troisième idée forte : mettre en balance les hypothèses

- **La statistique du rapport SUDI n’a de pertinence que si elle est utilisée pour comparer deux évènements** et c’est ainsi qu’elle est utilisée dans ce rapport.
- Et pourtant le Pr Meadow a marqué l’opinion publique en faisant croire
  - d’une part qu’il y avait 1 chance sur 73 millions pour que Sally Clark soit innocente,
  - d’autre part qu’en prouvant que les morts n’étaient pas des MSN, elles n’étaient pas naturelles.

Quatrième idée forte : la taille de la population testée ou questionnée

- Il n’est pas rare qu’en médecine les tests sur la foi desquels sont prises des décisions soient effectués sur des populations aussi petites que 5 individus.
- Par exemple l’étude SUDI analysait les doubles décès à travers l’étude de 24 familles.

Intervalles de confiance et Tests d'hypothèses

- Les intervalles de confiance permettent de quantifier le crédit que l'on apporte à une hypothèse ; ils sont d'autant plus larges que la population étudiée est petite.
- Dans l'exemple électoral, avec 24 personnes testées, on obtient des intervalles comme [0%,28.2%] …!
- On aimerait également pouvoir prédire si une quantité est plus grande qu'un autre avec le minimum de risque : donner deux intervalles de confiance qui se recouvrent est alors insuffisant.

Risque α

- On veut tester à partir d'observations $(x_1,\dots,x_n)$ d'une variable aléatoire X, d'espérance μ, si l'hypothèse H_0 « $\mu=\mu_0$ » est raisonnable.
- On se donne un intervalle I tel que la moyenne empirique $\chi=(x_1+\dots+x_n)/n$ appartient à I avec une probabilité $1-\alpha$, si H_0 est vraie. La quantité α est appelée « risque de 1ère espèce ».
- Cas bilatéral : si H_0 est l'hypothèse «$\mu=\mu_0$ », on cherche d tel que $P(|\chi-\mu_0|\leq\delta)=1-\alpha$.
- Cas unilatéral : si H_0 est l'hypothèse «$\mu>\mu_0$ », on cherche d tel que $P(\chi\geq\mu_0-\delta)=1-\alpha$.
- Ensuite on observe si χ appartient à I ou non : dans le premier cas, on accepte H_0 avec un risque α ; dans le second, soit on effectue un autre test, soit on rejette H_0.

Exemple

- On veut tester si A récoltera plus de voix que B lors d'une élection. Après avoir interrogé *n* personnes, on en a trouvé respectivement $p_A=n_A/n$ et $p_B=n_B/n$ déclarant vouloir voter pour A et pour B.
- On note X la variable aléatoire qui prend la valeur +1, -1 ou 0 selon qu'un individu dit vouloir voter pour A, pour B ou faire autre chose. On s'intéresse à l'hypothèse

  H_0 : « E(X)=0 » ou encore «$p_A=p_B$ »

  autrement dit : rien ne permet de dire si A ou B va l'emporter …
- La moyenne empirique est p_A-p_B.
- L'estimateur de la variance est :
- Au risque α, on peut donc accepter l'hypothèse $p_A=p_B$, si
- Par exemple, si p_A=14%, p_B=17%, n=850 et α=90%, la condition est

  $$0.03\leq1.64\text{x}((0.31\text{x}0.69+4\text{x}0.14\text{x}0.17)/849)^{½} = 0.031$$
- Par conséquent on accepte H_0, i.e. statistiquement la différence entre A et B n'est pas significative …

Hypothèse alternative

- Quand on teste une hypothèse H_0, dite hypothèse nulle, on doit en général se placer dans une alternative, et définir une hypothèse H_1. Implicitement H_1 est prise comme la négation de H_0.
- Dans l'exemple précédent, H_1 était « p_A est distinct de p_B ».
- Si on avait l'idée que B doit l'emporter, on aurait testé avec, pour H_1, « $p_A < p_B$ ». Dans ce cas on cherche un intervalle unilatéral
- Soit par exemple, avec α=90% : 14%≤17%-2.4%, et, finalement, il est plus probable que B l'emporte. Par contre avec α=95%, le test devient : 14%≤17%-3.1%, et donc le risque que l'on a à parier que B l'emporte est supérieur à 5%, ce qui laisse un doute réel planer.

Risque β

- Le risque α est tout simplement le seuil du test. C'est la probabilité de rejeter H_0 alors qu'elle est vraie.
- Maintenant supposons que H_0 soit fausse et, plus précisément, que H_1 soit vraie. Le risque β, ou risque de seconde espèce, est la probabilité d'admettre H_0 alors que H_1 est vraie.

Décision	Acceptation de H_0	Rejet de H_0
Si H_0 est vraie	Seuil de confiance $1-\alpha$	Seuil du test α
Si H_1 est vraie	Erreur de 2nde espèce β	Puissance du test $1-\beta$

Interdépendance des risques

- Prenons l'exemple de l'estimation d'une proportion *p* à partir d'une fréquence empirique *f*, avec H_0 : « $p=p_0$ » et, pour un certain $\delta>0$, H_1 : « $p=p_1=p_0+\delta$ ».
- Sous H_0, *f* suit une loi $\mathcal{N}(p_0,\sigma)$ et sous H_1, *f* suit une loi $\mathcal{N}(p_1,\sigma)$, avec σ^2 l'estimateur de la variance.
- Au risque α correspond un seuil t_α de sorte que l'on rejette H_0 si $|p-p_0|>t_\alpha\sigma$.
- Le risque β correspond au cas où $|p-p_0|<t_\alpha\sigma$ mais c'est H_1 qui est vraie. On a donc
- Par conséquent si α diminue, t_α augmente et donc β augmente.
- Par contre si n augmente, σ diminue, l'intégration se fait donc plus sur la « queue » de la gaussienne, et donc β diminue.
- La courbe de β en fonction de δ s'appelle Courbe des Caractéristiques Opérationnelles. C'est une courbe décroissante de 1 vers 0, passant par le point $(p_0+t_\alpha\sigma,1/2)$.

Interprétations

- Prendre moins de risque à réfuter H_0, conduit à prendre plus de risque à ne pas voir que c'est H_1 qui est vraie : augmenter α, c'est diminuer β.
- Pour pouvoir diminuer à la fois α et β, il faut recueillir plus d'informations : il faut augmenter *n*.
- On voit que les hypothèses H_0 et H_1 ne sont pas symétriques. Quand on effectue un test, on pense que H_0 est vraie et on se donne une alternative. Ou alors on pense que H_1 est vraie et on veut la mettre à l'épreuve avant de l'accepter.

Exemples

- Un exemple de risque α très important est donné par les tests en médecine : ne pas détecter une maladie mortelle est un risque que l'on peut difficilement prendre.
- Un autre exemple est fourni par les procédés industriels. Si un nouveau procédé est meilleur que le précédent, mais que sa mise en œuvre est excessivement coûteuse, on désire s'assurer que l'on ne fait pas tout ça pour trop peu. Et donc on prend α petit.
- Mais les risques β ne sont pas moins importants ! En médecine les conséquences du test peuvent ne pas être anodines, comme par exemple pour le HT21 …

Tests en médecine

Le dosage HT21 ou « Triple test »

- Le dosage « Human Trisomy 21 » est un dosage sanguin de deux ou trois marqueurs pris dans le sang de la mère (entre 15 et 18 semaines d'aménorrhée).
- A partir de ces dosages est calculé un risque que l'enfant soit affecté d'une anomalie du chromosome 21.

Triple-test et Amniocentèse

- D'une part le triple-test ne détecte pas tous les fœtus atteint de trisomie 21, d'autre part, très souvent, il donne un résultat positif alors que le fœtus est normal : les modifications biologiques entraînées par une trisomie 21 ne sont ni systématiques, ni exclusives de cette pathologie.
- Lorsque le risque est supérieur à 1/250, une confirmation est fortement conseillée, par amniocentèse.

Quelques données

860 000 tests en 1997 et 1998

799 700 enfants normaux (Test HT21 négatif)

300 trisomies 21 non détectées (Test HT21 faussement négatif)

59 300 enfants normaux faussement détectés (Test HT21 faussement positif)

700 trisomies 21 détectées par amniocentèse (Test HT21 positif)

Tests positifs 7 %	dont 98.8% faussement	soit 6.9% des femmes testées
	dont 1.2% confirmés	soit 0.08% des femmes testées soit 70% des trisomies
Tests négatifs 93 %	dont 0.04% faussement	soit 0.03% des femmes testées soit 30% des trisomies
	dont 99.96% vraiment	soit 93% des femmes testées

Amniocentèse

- L'amniocentèse comporte des risques de fausse couche évalués à 1% environ.
- Elle est proposée aux femmes de plus de 38 ans et aux femmes dont le risque calculé est supérieur à 1/250. Elle est alors remboursée.
- Les autres femmes peuvent prendre à leur charge une amniocentèse, avec le coût et le risque de fausse couche que cela représente.

Les risques α et β

- On pratique un dosage HT21. Soit H_0 l'hypothèse « le risque est inférieur à 1/250 » et H_1 sa négation. Soit D_0 la décision « le fœtus est normal » et D_1 « le fœtus est atteint de trisomie 21 ».
- Avec les données précédentes, on a donc

$$\alpha=P(D_1|H_0)=0.04\%$$

et

$$\beta=P(D_0|H_1)=98.8\%$$

- On a donc réduit à l'extrême α, et par conséquent β est très élevé.

Théorie Bayesienne de la décision

- Plutôt que de fixer arbitrairement a, on peut introduire une notion de « coût » pour les risques α et β. Par exemple on peut évaluer le coût d'avoir accepté un lot de marchandises trop défectueuses (donc invendables) et celui de perdre un fournisseur alors qu'il avait respecté le cahier des charges.
- Si on observe une moyenne de χ, on donne un coût au rejet à tort de H_0 en fonction de χ et α, et un coût à l'acceptation à tort de H_0 en fonction de χ et β (i.e. de χ et α). Ce coût devient alors une variable aléatoire et on cherche à en minimiser l'espérance.
- Bien entendu dans le cas du test HT21, de tels coûts n'ont aucun sens …mais il reste des questions !

Quelques remarques polémiques

- Si l'on croit au risque dû à l'amniocentèse, il y a donc, sur 860 000 grossesses testées, 700 trisomies 21 détectées, 300 trisomies 21 non détectées et 600 fausses couches ayant entraîné la mort d'un enfant sain.
- Calculons donc la probabilité de chacun de ces événements sachant que la grossesse a été pathologique.

Trisomie 21 détectée	Trisomie 21 non détectée	Mort d'un enfant sain
44%±2.4%	19% ±1.9%	37% ±2.4%

On est loin du 70% qui donne confiance et on est un peu effrayé par le 37% de morts « inutiles ».

Le risque dû à l'amniocentèse

- Le risque de fausse couche dû à l'amniocentèse n'est pas uniforme : il y a des populations à risque. Par exemple lorsque l'utérus a développé un fibrome, le risque est plus élevé.
- En pratique le risque 1/100 ne représente pas grand-chose : il y a des femmes pour qui le risque est bien plus important alors que pour les autres le risque est réellement minime.
- Il faudrait donc approfondir cette étude.

Bilan : les réflexes à acquérir

- Pour comparer des pourcentages, il faut les ramener à un « tout » commun.
- Les statistiques donnent des intervalles, pas des nombres.
- On ne peut pas donner un classement individuel statistique, mais un classement par groupes.
- On peut opposer deux hypothèses pour décider laquelle est la plus probable, mais la notion de probabilité « absolue » est rarement pertinente.
- Une statistique n'a de pertinence que si elle s'est opérée sur un grand nombre de cas : dans le cas contraire, l'intervalle obtenu statistiquement est tout simplement tellement gigantesque qu'on ne peut rien en conclure.