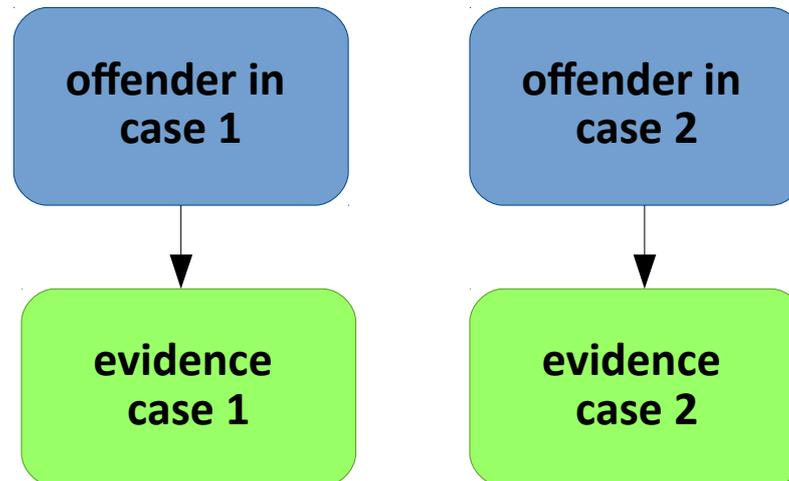




En quête du raisonnement bayésien en criminalistique

ou comment évaluer la preuve en criminalistique ?



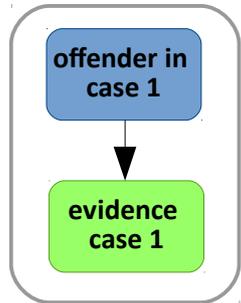
Dr. Guillaume Boudarham



Plan



Introduction à l'approche bayésienne



L'approche bayésienne en criminalistique



Évaluation d'un indice selon deux propositions alternatives : le rapport de vraisemblance

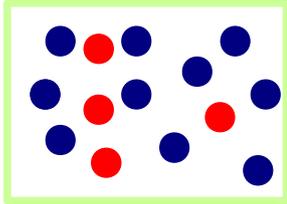


Conclusion : avantages et inconvénients de l'approche bayésienne en criminalistique



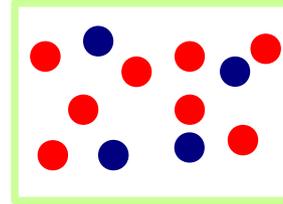
Introduction : problème des urnes

Urne 1



boules bleues = 10
boules rouges = 4

Urne 2



boules bleues = 4
boules rouges = 8

Problème : une personne qui a les yeux bandés tire une boule au hasard
→ on lui demande alors de dire de quelle urne provient la boule tirée

Puisqu'elle a les yeux bandés, elle ne peut donner que la *probabilité* P que la boule tirée provienne de l'urne 1 (ou de l'urne 2) → pas de certitude

$$P(\text{boule vient de l'urne 1}) = 1/2 = 50 \%$$

$$P(\text{boule vient de l'urne 2}) = 1/2 = 50 \%$$



problème simple
et intuitif



P se lit « probabilité »



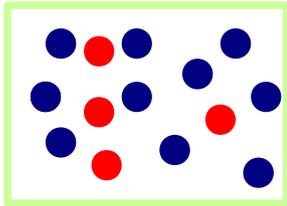
« événement » dont on cherche la probabilité

→ il y a donc 50 % de chance que la boule tirée au hasard provienne de l'urne 1 et 50 % de chance qu'elle provienne de l'urne 2



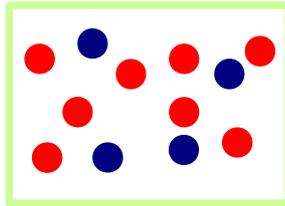
Introduction : problème des urnes

Urne 1



boules bleues = 10
boules rouges = 4

Urne 2



boules bleues = 4
boules rouges = 8

$$P(\text{boule vient de l'urne 1}) = 1/2 = 0.5$$

$$P(\text{boule rouge}) = 0.3 \times 0.5 + 0.7 \times 0.5 = 0.5^*$$

$$P(\text{boule rouge} \mid \text{boule vient de l'urne 1}) = 4/14 \approx 0.3$$

se lit « probabilité que la boule tirée soit rouge sachant qu'elle provient de l'urne 1 »

On lui enlève ensuite le bandeau ce qui lui permet de voir la couleur de la boule : la boule tirée est rouge. Connaissant la composition des urnes, on lui demande alors de réévaluer la probabilité que la boule provienne de l'urne 1 → résolution par le théorème de Bayes

proposition

observation

$$P(\text{boule vient de l'urne 1} \mid \text{boule rouge}) = \frac{P(\text{boule rouge} \mid \text{boule vient de l'urne 1}) \times P(\text{boule vient de l'urne 1})}{P(\text{boule rouge})}$$

| se lit « sachant que »

$$* P(\text{boule rouge}) = P(\text{boule rouge} \mid \text{boule vient de l'urne 1}) \times P(\text{boule vient de l'urne 1}) + P(\text{boule rouge} \mid \text{boule vient de l'urne 2}) \times P(\text{boule vient de l'urne 2}) = 0.3 \times 0.5 + 0.7 \times 0.5 = 0.5$$

(formule des probabilités totales)

Conclusion

$$P(\text{boule vient de l'urne 1} \mid \text{boule rouge}) = \frac{0.3 \times 0.5}{0.5} = 0.3 = \underline{30\%}$$

→ le fait de savoir que la boule tirée est rouge lui a permis de réévaluer à la baisse la probabilité qu'elle provienne de l'urne 1 (30 % au lieu de 50 %)



Le théorème de Bayes

Étant donnés deux événements **A** et **B**, le *théorème de Bayes* permet de déterminer la probabilité de **A** sachant **B**, si l'on connaît les probabilités de **A**, **B** et de « **B** sachant **A** »

à partir d'une (ou plusieurs) observation(s), on cherche à évaluer une proposition → pas de certitude

proposition

observation(s)

$$P(\mathbf{A} \mid \mathbf{B}) = \frac{P(\mathbf{B} \mid \mathbf{A}) \times P(\mathbf{A})}{P(\mathbf{B})}$$



Thomas Bayes (1702-1761)

→ formalisme adapté au contexte judiciaire où l'expert doit évaluer ses observations sachant différentes propositions (« évaluation d'une trace ») et où le tribunal doit prendre une décision en évaluant une proposition compte-tenu de diverses observations (voir plus loin)

$P(\mathbf{A})$ est la probabilité *a priori* de **A**, c'est-à-dire avant de connaître le fait **B**

$P(\mathbf{A} \mid \mathbf{B})$ est la probabilité *a posteriori* de **A** sachant **B**, c'est-à-dire la probabilité de **A** après avoir pris connaissance de **B**

Remarque $P(\mathbf{B})$ peut être obtenue à partir de la formule des « *probabilités totales* »

$$P(\mathbf{B}) = P(\mathbf{B} \mid \mathbf{A}) \times P(\mathbf{A}) + P(\mathbf{B} \mid \sim \mathbf{A}) \times P(\sim \mathbf{A}), \text{ où } P(\sim \mathbf{A}) = 1 - P(\mathbf{A}) \quad (\sim \mathbf{A} : \text{complémentaire de } \mathbf{A})$$

Exemples **A** : la boule tirée provient de l'urne 1 ou le suspect est la source de la trace

$\sim \mathbf{A}$: la boule tirée provient de l'urne 2 ou le suspect n'est pas la source de la trace

B : la boule tirée est rouge ou les profils ADN de l'agresseur et du suspect sont identiques

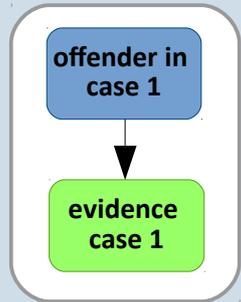


Plan

Introduction à l'approche bayésienne



L'approche bayésienne en criminalistique



Évaluation d'un indice selon deux propositions alternatives : le rapport de vraisemblance



Conclusion : avantages et inconvénients de l'approche bayésienne en criminalistique





Exemple d'erreur d'interprétation

L'inversion du conditionnel

Un homicide a été commis à Poitiers et une trace de sang appartenant à l'agresseur a été retrouvée sur la victime. Selon les enquêteurs, l'agresseur est un habitant de Poitiers. Un suspect est arrêté et son profil génétique est identique à celui trouvé sur la victime

→ L'expert : « *il y a 1 chance sur 1 milliard* pour qu'une personne choisie au hasard dans la population de Poitiers et n'étant pas la source de la trace présente un profil génétique identique à celui trouvé sur la victime* »

→ Interprétation : « *il y a 1 chance sur 1 milliard que le suspect, s'il n'est pas la source de la trace, possède un profil génétique identique à celui trouvé sur la victime. Donc, il y a 1 chance sur 1 milliard qu'il ne soit pas la source de la trace et 99.9999999 % de chance qu'il le soit* » → **interprétation fallacieuse**

« $P(\text{profils identiques} \mid \text{le suspect n'est pas la source de la trace}) = 1/1 \text{ milliard}$ »

égalité fallacieuse → $= P(\text{le suspect n'est pas la source de la trace} \mid \text{profils identiques})$

donc : $P(\text{le suspect est la source de la trace} \mid \text{profils identiques}) = 1 - 1/1 \text{ milliard} = 99.9999999 \% \text{ »}$

* fréquence d'apparition ou probabilité de coïncidence fortuite dudit profil dans la population de référence considérée (Poitiers). Idéalement, il faudrait calculer des fréquences alléliques pour la population la plus proche de celle de l'agresseur. A défaut de telles informations, les calculs doivent être basés sur les fréquences apparaissant dans la population générale.

** voir remarque slide 10 à propos de cette valeur choisie pour illustrer notre propos.



Exemple d'erreur d'interprétation

L'inversion du conditionnel

→ Interprétation : « *il y a 1 chance sur 1 milliard que le suspect, s'il n'est pas la source de la trace, possède un profil génétique identique à celui trouvé sur la victime. Donc, il y a 1 chance sur 1 milliard qu'il ne soit pas la source de la trace et 99.9999999 % de chance qu'il le soit* »

« $P(\text{profils identiques} \mid \text{le suspect n'est pas la source de la trace}) = 1/1 \text{ milliard}$ »

égalité fallacieuse → $= P(\text{le suspect n'est pas la source de la trace} \mid \text{profils identiques})$

donc : $P(\text{le suspect est la source de la trace} \mid \text{profils identiques}) = 1 - 1/1 \text{ milliard} = 99.9999999 \% \gg$

→ cette interprétation fallacieuse s'appelle la *confusion des inverses*

→ dans cet exemple, elle revient à supposer implicitement que le suspect a *a priori* 1 chance sur 2 d'être la source de la trace (avant d'avoir considéré le résultat du test ADN)

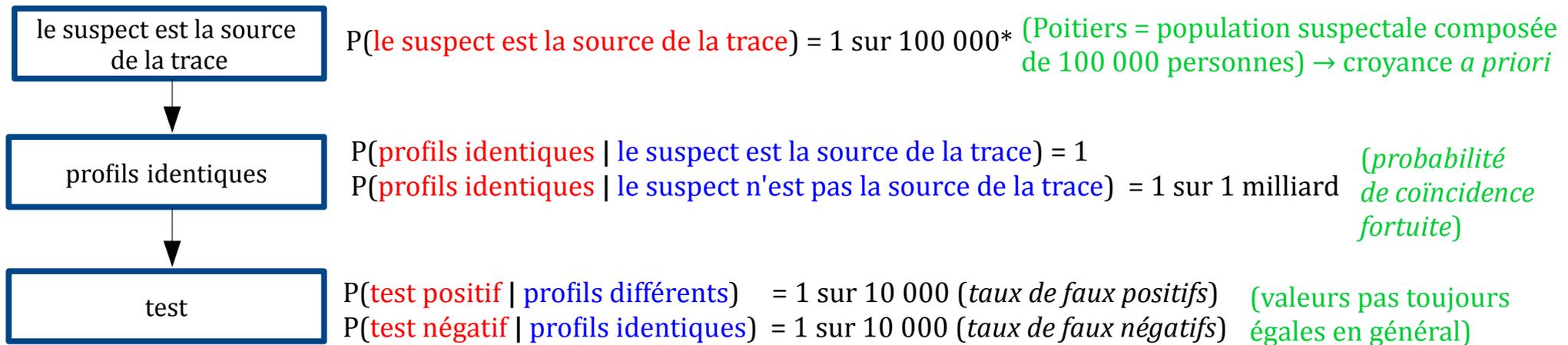
soit $P(\text{le suspect est la source de la trace}) = 0.5$ → cela n'est généralement pas le cas

Remarque « *a priori* » signifie ici « avant de prendre en compte le résultat du test ADN »



Un homicide a été commis à Poitiers et une trace de sang appartenant à l'agresseur a été retrouvée sur la victime. Selon les enquêteurs, l'agresseur est un habitant de Poitiers. Un suspect est arrêté et son profil génétique est identique à celui trouvé sur la victime

→ L'expert : « *il y a 1 chance sur 1 milliard pour qu'une personne choisie au hasard dans la population de Poitiers et n'étant pas la source de la trace présente un profil génétique identique à celui trouvé sur la victime* » → utilisation du théorème de Bayes → d'autres informations sont alors nécessaires



Les données sont représentées sous la forme d'un graphe appelé « *réseau bayésien* ». Une flèche représente une dépendance → représentation intuitive des données

Remarque on tient compte ici des faux positifs et négatifs des tests ADN pour plus de généralité

* si aucune autre preuve n'est acquise avant que le résultat du test ADN soit pris en compte, sinon il faut modifier cette probabilité en remplaçant $P(\text{le suspect est la source de la trace})$ par $P(\text{le suspect est la source de la trace} \mid \text{preuve } 1, \dots, \text{preuve } n)$



A partir du **théorème de Bayes**, on trouve :

→ $P(\text{le suspect n'est pas la source de la trace} \mid \text{test négatif}) = \underline{99.99\%}$

→ $P(\text{le suspect est la source de la trace} \mid \text{test positif}) = \underline{9.09\%}$ *

et $P(\text{le suspect n'est pas la source de la trace} \mid \text{test négatif}) \gg P(\text{le suspect est la source de la trace} \mid \text{test positif})$

Conclusions : dans cet exemple

- le test ADN positif ne permet pas ici à lui seul d'affirmer que le suspect est la source de la trace

→ d'autres preuves sont nécessaires → il faudrait alors remplacer $P(\text{le suspect est la source de la trace})$ par $P(\text{le suspect est la source de la trace} \mid \text{preuve 1, ..., preuve n})$ puis en déduire $P(\text{le suspect est la source de la trace} \mid \text{preuve 1, ..., preuve n, test positif})$

- le test ADN négatif permet à lui seul d'affirmer que le suspect n'est pas la source de la trace (probabilité = 99.99 %)

→ le fait de savoir que le test ADN est positif a malgré tout permis de gagner 4 ordres de grandeur sur la probabilité que le suspect soit la source de la trace (0.001 % → 9.09 %)

→ succès de la « preuve ADN » en criminalistique mais cet exemple montre qu'elle ne peut pas être considérée comme la « reine des preuves »

* *remarque* : la probabilité de coïncidence fortuite choisie ici est un ordre de grandeur que nous avons fixé pour illustrer notre propos. En pratique, cette valeur peut être de 1 sur plusieurs milliards. En outre, les taux de faux positifs et négatifs peuvent être aussi plus faibles. Pour une probabilité de coïncidence fortuite de 1/mille milliards et taux de faux positifs = taux de faux négatifs = 1/1million, on trouve aisément :

$P(\text{le suspect est la source de la trace} \mid \text{test positif}) \approx 90.9\%$ permettant alors de dire que le suspect est la source de la trace.



Résolution par le théorème de Bayes

a priori : avant prise en compte du résultat du test ADN
a posteriori : après prise en compte du résultat du test ADN

Tableau récapitulatif

Probabilité *a posteriori* en fonction de la probabilité *a priori* P(le suspect est la source de la trace | preuve 1,...,preuve n).

Notations. C : « le suspect est la source de la trace » ; I : « le suspect n'est pas la source de la trace », P₁,...,P_n : preuves obtenues avant le résultat du test ADN ; test P, N : test positif, négatif. Le taux de faux positifs et négatifs est fixé à 1 sur 10 000 et la probabilité de coïncidence fortuite à 1 sur 1 milliard. Les probabilités sont données en %

P(C P ₁ ,...,P _n)	P(C P ₁ ,...,P _n , test P)	P(I P ₁ ,...,P _n , test N)
0	0	100
→ 0.001	<u>9.09008265</u>	<u>99.9999999</u>
0.002	16.6654167	99.9999998
0.003	23.0755029	99.9999997
0.004	28.57	99.9999996
0.005	33.332	99.9999995
0.006	37.4988281	99.9999994

P(C P ₁ ,...,P _n)	P(C P ₁ ,...,P _n , test P)	P(I P ₁ ,...,P _n , test N)
0.980	98.99959	99.9999010
→ 0.981	<u>99.0006096</u>	99.9999009
0.982	99.0016271	99.9999008
0.983	99.0026426	99.9999007
0.984	99.0036561	99.9999006
0.985	99.0046675	99.9999005
0.986	99.0056768	99.9999004

→ la flèche bleue indique le résultat lorsque aucune autre preuve n'a été obtenue avant la prise en compte du résultat du test ADN : $P(C|P_1, \dots, P_n) = P(C) = 1/100\ 000 = 0.001\ %$ (population suspectale de 100 000 personnes)

→ en présence d'autres preuves, dès que $P(C|P_1, \dots, P_n) > 0.98\ %$, la probabilité que le suspect soit la source de la trace si le test ADN est positif dépasse 99 % (flèche rouge). La probabilité qu'il ne soit pas la source de la trace si le test ADN est négatif ne varie quasiment pas et reste supérieur à 99.99 %

→ nécessité ici d'accumuler d'autres preuves afin que le résultat positif du test ADN permette d'affirmer que le suspect est la source de la trace (à condition que la probabilité *a priori* augmente alors suffisamment) et/ou diminuer suffisamment le taux de faux positifs et/ou diminuer la probabilité de coïncidence fortuite (utiliser plus de marqueurs génétiques si possible, etc)



Problème précédent résolu simplement à l'aide du logiciel *Elvira*
→ création intuitive d'un réseau bayésien → résolution automatique du problème

The screenshot shows the Elvira software interface. The main window displays a Bayesian network diagram with three nodes: "Suspect est la source de la trace" (circled in black), "Profils identiques", and "Test". A dialog box titled "Node: Suspect est la source de la trace" is open, showing the configuration for this node. The dialog box has tabs for "Node", "Values", "Parents", and "Relation". The "Relation" tab is selected, showing "Relation Type" set to "General" and "Probabilistic" selected. The "Values" tab is also visible, showing a table of probabilities for the "yes" and "no" states. The "yes" state has a probability of 4.0E-5, and the "no" state has a probability of 0.99996. A red arrow points from the text "P(le suspect est la source de la trace | preuve 1) = 0.004 % = 0.00004*" to the value 4.0E-5 in the table.

State	Probability
yes	4.0E-5
no	0.99996

Exemple : preuve 1 = le suspect était sur les lieux du crime au moment des faits. Cette preuve a fait augmenter la probabilité que le suspect soit la source de la trace de sang retrouvée sur la victime. Sans cette preuve, cette probabilité valait 0.001 % (cf. slides précédents)

** la valeur choisie $P(\text{le suspect est la source de la trace} \mid \text{preuve 1}) = 0.004 \%$ est la probabilité que le suspect soit la source de la trace après avoir pris en compte la preuve n°1 mais avant d'avoir pris en compte le résultat du test ADN : cette probabilité peut être considérée comme une probabilité a priori → estimation difficile en pratique*



Utilisation du logiciel *Elvira*

Elvira

File Edit View Tasks Options Window Help

150% Edit Mode « Edit »

Nœud 2 : « Profils identiques »

C:\Users\Guillaume\Desktop\Cours Bayes\ADN.elv

Node: Profils identiques

Node Values Parents Relation

Relation Type General Probabilistic Deterministic

All Parameters Independent Parameters Values Probabilities CPT Canonical Parameters Net Compound

Suspect est ...	yes	no
yes	1.0	1.0E-9
no	0.0	0.999999999

$P(\text{profils identiques} \mid \text{le suspect n'est pas la source de la trace}) = 1/1 \text{ milliard}$ (probabilité de coïncidence fortuite)

$P(\text{profils identiques} \mid \text{le suspect est la source de la trace}) = 1$ (certitude)

OK Cancel Apply



Elvira

File Edit View Tasks Options Window Help

150% Edit

Mode « Edit »

Nœud 3 : « Test »

C:\Users\Guillaume\Desktop\Cours Bayes\ADN.elv

Node: Test

Relation Type: General

Probabilistic Deterministic

All Parameters Independent Parameters

Values Probabilities

CPT Canonical Parameters

Net Compound

Profils identi...	yes	no
positive	0.9999	1.0E-4
negative	1.0E-4	0.9999

$P(\text{test positif} \mid \text{profils différents}) = 1/10\ 000$ (faux positifs)

$P(\text{test négatif} \mid \text{profils identiques}) = 1/10\ 000$ (faux négatifs)

OK Cancel Apply



Utilisation du logiciel *Elvira*

Les tests ADN

Résolution automatique du problème → $P(\text{le suspect est la source de la trace} \mid \text{preuve 1, test positif}) = ?$

The screenshot shows the Elvira software interface in 'Inference' mode. The main window displays a Bayesian network diagram with three nodes: 'Suspect es...', 'Profils identiques', and 'Test'. Each node contains a table of probabilities for 'yes' and 'no' (or 'positive' and 'negative') outcomes. The 'Test' node is currently set to 'positive' with a probability of 1.00000. A red arrow points from the 'Test' node to the 'Suspect es...' node, indicating the direction of inference. A red arrow also points from the 'Suspect es...' node to the text $P(\text{le suspect est la source de la trace} \mid \text{preuve 1, test positif}) = 0.2857 = \underline{28.57 \%}$. A dashed arrow labeled 'calcul automatique' points from a gear icon to the text. A text box at the bottom right explains that double-clicking the 'positive' option in the 'Test' node sets its value to 1, representing a certain result.

Suspect es...

yes		0,28570
<input type="checkbox"/> no		0,71430

Profils identiques

yes		0,28571
<input type="checkbox"/> no		0,71429

Test

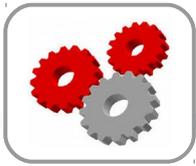
<input type="checkbox"/> positive		1,00000
negative		0,00000

$P(\text{le suspect est la source de la trace} \mid \text{preuve 1, test positif}) = 0.2857 = \underline{28.57 \%}$

calcul automatique

on double-clic ici pour mettre test positif à 1 (résultat du test ADN connu avec certitude : il est positif)

→ nécessité ici d'accumuler d'autres preuves afin que le résultat positif du test ADN permette d'affirmer, avec une probabilité suffisante, que le suspect est la source de la trace de sang (voir slide 11)



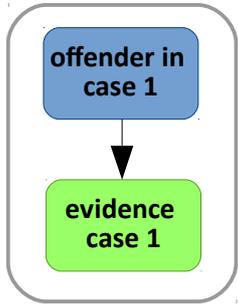
Plan

Introduction à l'approche bayésienne

L'approche bayésienne en criminalistique

Évaluation d'un indice selon deux propositions alternatives : le rapport de vraisemblance

Conclusion : avantages et inconvénients de l'approche bayésienne en criminalistique





Le rapport de vraisemblance : évaluation d'une trace

Afin d'évaluer une preuve de façon équilibrée (recommandation de l'ENFSI*), l'expert doit comparer :

- la probabilité d'observer la preuve si la proposition de l'accusation est vraie
→ **proposition (ou hypothèse) de l'accusation**

avec

- la probabilité d'observer la preuve si la proposition de la défense est vraie
→ **proposition (ou hypothèse) de la défense**

L'expert détermine le rapport entre ces deux probabilités appelé « rapport de vraisemblance » noté *RV*

▪ ce rapport permet de quantifier le poids apporté par une preuve au regard de deux propositions alternatives (accusation vs défense) :

$$RV = \frac{P(\text{observation de la preuve} \mid \text{proposition de l'accusation})}{P(\text{observation de la preuve} \mid \text{proposition de la défense})}$$

Remarque le rapport de vraisemblance porte sur les propositions des parties sur lesquelles la trace est susceptible d'apporter des éléments et non sur des probabilités de culpabilité ou d'innocence. La culpabilité ou l'innocence d'un individu est une décision de justice... celle-ci n'est donc pas du ressort de l'expert

* ENFSI Guideline For Evaluative Reporting in Forensic Science



Le rapport de vraisemblance : évaluation d'une trace

A partir du **théorème de Bayes**, on trouve :

$$\frac{P(\text{proposition de l'accusation} \mid \text{observation de la preuve})}{P(\text{proposition de la défense} \mid \text{observation de la preuve})} = \text{RV} \times \frac{P(\text{proposition de l'accusation})}{P(\text{proposition de la défense})}$$

↑
croyance actualisée (a posteriori) dans la véracité d'une proposition (domaine de compétence du juge qui a accès aux données de l'enquête)

↑
rapport de vraisemblance qui permet de quantifier le poids apporté par une preuve au regard de deux propositions alternatives (domaine de compétence de l'expert)

↑
croyance a priori dans la véracité d'une proposition : (domaine de compétence du juge qui a accès aux données de l'enquête)

Cette expression se compose de deux termes et permet :

- de rendre visible la répartition des rôles entre l'expert, qui se prononce sur la valeur probante de l'indice (RV) et le juge, qui a accès aux données de l'enquête (croyance *a priori*)
- de mettre en commun les résultats scientifiques de l'expert et ceux obtenus par les enquêteurs (croyance *a posteriori*)



Le rapport de vraisemblance : application

Application au cas d'une expertise génétique

Lors d'une soirée étudiante qui rassemble 100 hommes et 50 femmes, une femme déclare avoir été violée. Du sperme appartenant à l'agresseur présumé est retrouvé sur la plaignante*. D'après les enquêteurs, l'agresseur présumé est l'un des hommes de la soirée. Un suspect X est arrêté et les expertises génétiques ont montré que son profil génétique n'est pas différenciable de celui trouvé sur la plaignante**

→ L'expert : « *il y a 1 chance sur 1 milliard pour qu'une personne choisie au hasard dans la population de référence et n'étant pas la source de la trace présente un profil génétique identique à celui trouvé sur la plaignante* »

→ **il y a donc 1 chance sur 1 milliard que le suspect, s'il n'est pas la source de la trace, possède un profil génétique identique à celui trouvé sur la plaignante (probabilité de coïncidence fortuite)**

$P(\text{profils ADN identiques} \mid X \text{ n'est pas la source de la trace}) = 1/1 \text{ milliard}$

mais $P(X \text{ n'est pas la source de la trace} \mid \text{profils ADN identiques}) \neq 1/1 \text{ milliard}$

** en toute rigueur, nous préférons utiliser ici le terme « plaignante » plutôt que « victime » puisqu'à ce stade de l'enquête, on suppose que le viol n'est pas prouvé (faux témoignage possible...). On s'intéresse ici à la source de la trace de sperme et pas à l'activité ayant mené à cette trace.*

*** on suppose ici que les expertises ont été réalisées plusieurs fois par des laboratoires indépendants de sorte que les taux de faux positifs et négatifs sont proches de 0. Par soucis pédagogique, nous n'en tiendrons pas compte dans cet exemple. On suppose aussi que les prélèvements n'ont pas été pollués.*



Le rapport de vraisemblance : application

Travail de l'expert : calcul du rapport de vraisemblance RV

Deux propositions alternatives sont mises en compétition (résultat équilibré) :

- proposition de l'accusation : X est la source de la trace
- proposition de la défense : X n'est pas la source de la trace

$$RV = \frac{P(\text{profils ADN identiques} \mid X \text{ est la source de la trace})}{P(\text{profils ADN identiques} \mid X \text{ n'est pas la source de la trace})}$$

← accusation
← défense

$$RV = \frac{1}{1/1\,000\,000\,000}$$

← on est certain que si X est la source de la trace, les deux profils ADN (le sien et celui retrouvé sur la victime) vont être identiques

← probabilité de coïncidence fortuite dudit profil génétique dans la population de référence considérée (valeur issue de publications scientifiques)

RV = 1 000 000 000 (valeur donnée par l'expert)

→ l'expert : « les résultats observés soutiennent très fortement la proposition de l'accusation par rapport à la proposition de la défense »

↑ échelle de conclusions selon la valeur de RV

Mais pour aller plus loin, il faut calculer le rapport des probabilités *a posteriori*



Le rapport de vraisemblance : application

Travail du juge : calcul du rapport des probabilités *a priori*

Deux propositions alternatives sont mises en compétition :

- proposition de l'accusation : X est la source de la trace
- proposition de la défense : X n'est pas la source de la trace

$$\frac{\text{P(X est la source de la trace)}}{\text{P(X n'est pas la source de la trace)}} \leftarrow \begin{array}{l} \text{accusation} \\ \text{défense} \end{array} \quad (\text{rapport des probabilités } a \text{ priori})$$

$$= \frac{1/100}{99/100} \leftarrow \begin{array}{l} 1 \text{ chance sur } 100 \text{ que } X \text{ soit la source de la trace (la source de la trace est l'un des} \\ 100 \text{ hommes de la soirée)} \\ 99 \text{ chances sur } 100 \text{ que } X \text{ ne soit pas la source de la trace (ce peut être en effet un} \\ \text{des 99 autres hommes de la soirée)} \end{array}$$

$$= \underline{0.01} \quad (\text{valeur issue de l'enquête : nombre d'hommes présents à la soirée, etc.})$$

→ il faut maintenant multiplier cette valeur connue du juge par le rapport de vraisemblance donné par l'expert afin d'obtenir le rapport des probabilités *a posteriori* (voir slide 18)



Le rapport de vraisemblance : application

A partir du rapport de vraisemblance, calculé par l'expert, et du rapport des probabilités *a priori*, obtenu à partir des données de l'enquête, on en déduit le rapport des probabilités *a posteriori* (voir slide 18) :

$$\frac{P(\text{X est la source de la trace} \mid \text{profils ADN identiques})}{P(\text{X n'est pas la source de la trace} \mid \text{profils ADN identiques})}$$

(rapport des probabilités a posteriori)

$$= 1\,000\,000\,000 \times 0.01 = \underline{10\,000\,000} \text{ (10 millions)}$$

→ *degré de croyance*

(combinaison entre une valeur issue de publications scientifiques et une valeur issue de l'enquête)

Conclusion :

« puisque les profils ADN sont identiques, il est 10 millions de fois plus probable que X soit la source de la trace plutôt qu'un autre parmi les hommes présents à la soirée. »

→ le tribunal pourra alors conclure que X est la source de la trace

→ conformément aux recommandations de l'ENFSI, c'est au tribunal de statuer sur cette identification

Différentes activités, criminelles ou pas, peuvent expliquer la présence de sperme sur la plaignante...

- rapport sexuel consenti mais faux témoignage de la plaignante ?

- sperme de X placé intentionnellement par l'agresseur sur la victime ? (c.f. affaire J.L Cayez, 2005)

- etc.



Plan



Introduction à l'approche bayésienne

offender in
case 1

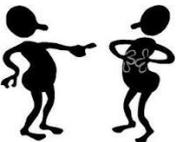


evidence
case 1

L'approche bayésienne en criminalistique



Évaluation d'un indice selon deux propositions alternatives : le rapport de vraisemblance



Conclusion : avantages et inconvénients de l'approche bayésienne en criminalistique



Conclusion

L'approche bayésienne en criminalistique : avantages

- **évaluation probabiliste d'une trace selon les propositions de l'accusation et de la défense et estimation du rapport de vraisemblance**
- **chaque nouveau fait observé fait évoluer ce rapport et donc le poids apporté par une preuve selon les propositions de l'accusation et de la défense**
- **le formalisme utilisé clarifie les domaines de compétence du juge et des experts qui les assistent**
- **Éviter autant que possible les raisonnements fallacieux et les erreurs judiciaires**
 - **rapport d'expertise équilibré conformément aux recommandations de l'ENFSI**
 - **approche logique, rigoureuse et universelle**



Conclusion

L'approche bayésienne en criminalistique : inconvénients

- données nécessaires pour évaluer une trace pas toujours disponibles (population de référence, probabilités *a priori*), choix des variables à prendre en compte ? indépendance de ces variables ?
- formalisme parfois difficile à utiliser notamment en balistique, comparaison d'écritures, biologie (mélange d'ADN...)
- pour une même probabilité, une personne peut concevoir un certain doute alors que d'autres perçoivent la même information avec davantage de confiance
- le formalisme bayésien peut rebuter des non-scientifiques (juges, avocats, jurés...)

Solutions proposées mais pas de consensus :

→ utilisation de graphes intuitifs ou « réseaux bayésiens », notamment aux assises

→ utilisation de logiciels intuitifs → calcul automatique des probabilités

→ réflexions nécessaires sur l'approche bayésienne dans un contexte judiciaire entre scientifiques et magistrats ?

→ formation continue des acteurs de la chaîne criminalistique ?

→ (in)formation préalable des jurés ?