

A Estatística no Tribunal

Manuela Figueira Neves

Instituto Politécnico da Guarda

Escola Superior de Tecnologia e Gestão

Centro de Estatística e Aplicações da Universidade de Lisboa

mfigueira@ipg.pt

Resumo: Nas últimas décadas o uso de métodos estatísticos para fins legais teve um grande e rápido desenvolvimento e que continua nos dias de hoje. Os estatísticos são chamados a tribunal como peritos ou consultores. Eles avaliam e quantificam a incerteza.

Podem considerar-se diferentes abordagens sobre o uso da estatística em tribunal. Alguns defendem o uso da abordagem bayesiana, apesar de reconhecerem que o uso dos métodos bayesianos baseados em distribuições *a priori* subjectivas gera muita controvérsia em termos legais. A objecção a esta abordagem é que o decisor tem que impor informação *a priori* antes ou independentemente da evidência encontrada. Está-se perante a interpretação subjectivista de probabilidade.

Em alternativa é também usada a abordagem frequentista ou “clássica”.

Muita embora já se verifique algum acolhimento relativamente à participação de estatísticos no ambiente jurídico, há ainda muito a fazer por uma aproximação entre a estatística e o direito.

Palavras - chave: Teorema de Bayes, Estatística Forense.

1.Introdução

No último quarto do século passado a importância da estatística na ciência forense e na criminalística tornou-se evidente. Com o objectivo de melhorar a apresentação e uso da argumentação estatística no sistema legal, estatísticos e não só escreveram artigos e livros sobre probabilidade e estatística especialmente para advogados e juízes e também para cientistas forenses.

Devido aos avanços na ciência e na tecnologia, actualmente os cientistas forenses podem usar um grande número de métodos e técnicas, aplicáveis a vários tipos de evidência mesmo em quantidades muito reduzidas.

Sendo a ciência forense uma ciência aplicada que se baseia nas leis da física e da química ela tem conhecido um rápido desenvolvimento devido à recente e grande evolução de técnicas relacionadas com perfis de ADN. As técnicas usadas actualmente de forma rotineira em muitos países eram impensáveis há alguns anos atrás. Uma das consequências é que a força da evidência é apresentada em tribunal de forma numérica.

Os perfis de ADN têm actualmente um grande impacto na detecção de crimes. Pode dizer-se que o perfil de ADN é um dos maiores avanços em termos de ciência forense. Durante alguns anos assistiu-se a uma grande controvérsia nos *media* e nos tribunais sobre a validade da tecnologia e interpretação apropriada da evidência de ADN. Em alguns países existiram alguns e longos confrontos em tribunal. Esta controvérsia levou ao aparecimento de muitos trabalhos científicos sobre o tema. Foram publicados centenas de artigos sobre perfis de ADN, alguns sobre recolha de dados, outros sobre considerações teóricas sobre probabilidades, estatística e genética. A discussão foi proveitosa para o desenvolvimento do tema. Pode dizer-se que existe mais trabalho científico sobre estatística com perfis de ADN do que de qualquer outra técnica forense, incluindo impressões digitais. Dado o seu grande desenvolvimento, as técnicas de análise de perfis de ADN foram reconhecidas, por advogados, juizes e cientistas forenses, como método apropriado para analisar aquele tipo de prova. Nos últimos anos os desenvolvimentos permitiram também o aumento do número de tipos de traços e de técnicas de análise que podem ser presentes aos tribunais. Estes desenvolvimentos permitiram a análise de um maior número de traços e de traços de dimensão reduzida. Mas não devem ser só os resultados das análises que devem ser apresentados no tribunal. Os resultados apresentados devem ser avaliados e interpretados no contexto do caso em julgamento. Mas, apesar da grande quantidade de informação potencial há ainda no meio forense uma crescente tomada de consciência da situação.

2. A Evidência

Para os cientistas forenses e para os tribunais a evidência refere-se aos dados ou indícios relativos ao caso em análise que foram recolhidas pela polícia ou investigadores e apresentados em tribunal pela acusação e pela defesa.

Mais concretamente, o termo evidência é usado para designar os materiais cuja existência no local do crime ou no suspeito levam à avaliação da sua culpa em relação a determinado crime.

Podem considerar-se diversos tipos de evidência. Fibras têxteis¹ e a sua existência no suspeito ou no local do crime são objecto de vários estudos. As armas de fogo constituem outro tipo de evidência em que se procede à análise de balas encontradas na vítima ou no local do crime para determinar de que tipo de arma são provenientes.

As impressões digitais também são frequentemente usadas em análise forense. O objectivo é o de determinar se existe coincidência entre as impressões digitais encontradas no local do crime ou na vítima e as impressões digitais do suspeito ou as existentes em bases de dados.

¹ Por exemplo Evett *et al.* (1987) e Wakefield *et al.* (1991).

Relacionado com o crime de falsificação de documentos é habitual usar como evidência documentos ou manuscritos.

Os fragmentos de vidro são uma das evidências mais estudada e referida na bibliografia² sobre análise forense.

A utilização de perfis de ADN em casos de identificação forense generalizou-se nos últimos anos. Este tipo de evidência é desde a década de 80 a que tem merecido mais atenção e é objecto de estudo na literatura sobre análise forense. Este tipo de evidência pode ser usado na identificação do suspeito de determinado crime ou na quantificação da culpa do suspeito tendo ele sido identificado por outro tipo de evidência que não o ADN.

Em termos científicos, a análise da evidência envolve incerteza e não pode ignorar-se que cada caso tem as suas especificidades. A evidência encontrada pode servir para identificar o suspeito do crime e/ou para atestar da sua culpa ou provar a sua inocência. Há várias questões a ter em conta na análise da evidência. A variabilidade associada a qualquer observação científica pode ser estudada usando distribuições de probabilidade adequadas ao fenómeno em causa.

A evidência científica tal como ela é encontrada no mundo real é de alguma forma incompleta, de tal forma que a incerteza é um elemento importante e com o qual os cientistas forenses têm habitualmente de lidar.

A evidência por si só pouco diz, o seu significado tem de ser elucidado à luz das proposições em causa e do conhecimento anterior sobre o caso.

Há uma grande necessidade prática de os cientistas forenses informarem os seus clientes, advogados, júri, ou decisores do significado das suas descobertas sobre o caso.

Os cientistas forenses são chamados a qualificar e se possível quantificar o seu estado de conhecimento e a serem consultores na avaliação das incertezas associadas com as inferências que devem ser feitas a partir da evidência. Para esta tarefa deve a teoria das probabilidades ser considerada como fundamental no seu raciocínio.

3. Avaliação da evidência científica

Foi desenvolvido um conjunto de conceitos básicos e fundamentais aplicando-se especificamente na análise forense. Esses conceitos são essenciais para a prática da ciência forense e podem ser vistos como processos que se usam na tentativa de dar resposta às várias questões que surgem num processo de investigação: *quem, o quê, onde, porquê, quando e como*.

Tradicionalmente podem ser considerados cinco conceitos básicos: transferência, identificação, individualização, associação e reconstrução. Inman e Rudin (2002) sugerem um sexto conceito que considera a existência de divisão antes de ocorrer a transferência da evidência.

² Lindley (1977), Seheult (1978), Grove (1980) e Aitken e Lucy (2004), entre outros.

A identificação consiste na categorização da evidência. Este processo tenta responder à questão “*o que é isto?*”. Em muitas situações a análise forense termina no processo de identificação, ou seja, a identificação pode ser vista como um fim. Mas, a identificação pode ser uma etapa que leva à individualização. Para vários autores, a individualização é vista como a conclusão sobre a origem comum de duas amostras. Qualquer análise forense que leva à individualização baseia-se na comparação de pelo menos dois itens, com o fim de saber se eles partilham a mesma origem.

Para Inman e Rudin (2002), associação define-se como uma inferência sobre o contacto entre a origem do indício e um alvo ou objectivo. Tal inferência é baseada na detecção da evidência transferida. A origem e o alvo são definições operacionais dependentes da estrutura do caso. Em situações em que a transferência é detectada em ambas as direcções cada item é ao mesmo tempo origem e alvo. O processo de associação envolve a avaliação de toda a evidência a favor e contra uma origem comum. Pode ser visto como a comparação de duas hipóteses contrárias.

A probabilidade da evidência segundo as duas hipóteses contrárias é uma expressão da verosimilhança da evidência dado que o alvo e a origem estiveram em contacto físico em contraste com a verosimilhança da evidência dado que o alvo esteve em contacto com uma origem diferente e não relacionada. Este processo requer a combinação do peso da evidência definido na fase de individualização com informação adicional e também com suposições feitas pelo analista.

A reconstrução é definida como a ordenação de associações no tempo e no espaço. Tenta-se encontrar resposta para as questões *onde*, *como* e *quando*. A questão *quando* refere-se a uma ordenação relativa no tempo. Por exemplo, existiu contacto antes, durante ou depois de o crime ter ocorrido?

A questão do *porquê*, ou seja a motivação do crime, não pode ser respondida por considerações relativas a evidências físicas.

Sabe-se que os cientistas forenses devem usar probabilidades para expressar os seus graus de credibilidade e as leis das probabilidades para fazerem inferência, quando avaliam a força da evidência.

A avaliação da evidência científica no tribunal é a expressão do grau de convicção de cada indivíduo. Usualmente requer a combinação das frequências de referência com o conhecimento pessoal das circunstâncias do caso em particular.

As probabilidades são a expressão das frequências relativas e o uso de probabilidades como a linguagem quantitativa do conhecimento pessoal proporciona um único e potente mecanismo para tratar com a combinação de dados estatísticos “objectivos” e a avaliação “subjectiva” dos factos do caso.

Na investigação preliminar as provas são provas técnicas e cabem no âmbito das ciências forenses. As provas técnicas podem ser usadas para suportar a existência ou a natureza de um acto criminoso, ou para ajudar a demonstrar ligações entre elementos envolvidos num acto criminoso (por exemplo uma agressão entre o acusado e a vítima do acto de violação).

Os cientistas investigam a existência de ligações com a ajuda de uma análise comparativa de múltiplos traços (tais como ADN e fibras têxteis) que podem ter sido transferidos durante o acto.

O conceito de “odds” é importante na tarefa de avaliar a evidência. A evidência é avaliada a partir da probabilidade de certa proposição sobre um suspeito (antes de chegar a julgamento) ou do arguido (enquanto o julgamento decorre). Inicialmente, a discussão será em termos do efeito da evidência nas probabilidades de culpa e de inocência do suspeito. Podem ser considerados dois acontecimentos complementares. O rácio das probabilidades destes dois eventos representa a chance (odds) contra a inocência ou a chance a favor da culpa. De referir que os acontecimentos representam a culpa ou inocência do suspeito e não o facto de ele ser considerado culpado ou inocente no julgamento.

O teorema de Bayes é um elemento essencial na consideração das chances. De facto, ele permite a reavaliação, com base em informação adicional, da medida da incerteza de determinada situação.

A grande controvérsia relacionada com a evidência científica está relacionada com a forma como ela é apresentada em tribunal. No julgamento é uma tarefa difícil a de assegurar que o juiz e os membros do júri compreendam a evidência científica e as dificuldades aumentam quando o cientista forense apresenta resultados da avaliação da força da evidência. Esta avaliação está relacionada com probabilidades como forma de avaliar a incerteza.

O grau de credibilidade, num caso particular, baseado no conhecimento de dados estatísticos relevantes pode ser considerado mais “objectivo” que um grau de credibilidade não baseado no conhecimento de dados estatísticos, no sentido que pessoas razoáveis concordam que se existem dados estatísticos conhecidos, os graus de credibilidade das pessoas devem ser baseados ou ter em conta esses dados.

Isto não permite afirmar que a probabilidade estimada com base no conhecimento de dados estatísticos é “objectiva” no sentido em que existe uma ligação intrínseca entre frequências e probabilidades que é independente de toda a informação relacionada e de não-frequências.

Todo o julgamento probabilístico de qualquer caso, mesmo quando é baseado em frequências, tem uma componente baseada no conhecimento pessoal de quem faz o julgamento.

O melhor argumento para aplicação da teoria bayesiana na ciência forense é mostrar que a teoria concorda com a intuição pessoal, quando os problemas de inferência são simples e as intuições são plausíveis e se pode ir através deles, quando os problemas se tornam mais complicados e as intuições não são tão plausíveis.

É importante assegurar que a incerteza seja avaliada de forma correcta para evitar as chamadas “falácias” que podem ser apresentadas tanto pela defesa como pela acusação.

Nos últimos 20 anos os desenvolvimentos permitiram o aumento do número de tipos de evidência e de técnicas de análise que podem ser presentes aos tribunais. Estes desenvolvimentos permitiram a análise de um maior número de evidência e de evidência de dimensão reduzida.

Mas não devem ser só os resultados das análises que devem ser apresentados no tribunal. Os resultados apresentados devem ser avaliados e interpretados no contexto do caso em julgamento.

A evidência científica requer muito cuidado na sua interpretação. Deve pôr-se ênfase na importância da questão “o que é que os resultados nos dizem (significam) neste caso em particular?” Então, teoria, métodos e aplicações de probabilidade e estatística estão na base da avaliação da evidência científica.

O julgador deve usar a avaliação probabilística para entender a avaliação da evidência científica que é habitualmente apresentada numericamente. Os cientistas forenses devem dar ao tribunal uma avaliação que ilustra o ponto de vista dos seus resultados. Tal avaliação usa probabilidades para medir a incerteza. Existe no meio consenso de que a evidência científica deve ser tratada em termos probabilísticos.

A avaliação da prova científica pode ser vista como a comparação entre a evidência de materiais encontrados na cena de um crime e a evidência de materiais encontrados num suspeito, na roupa do suspeito ou no ambiente à sua volta.

O rácio de verosimilhança pode ser visto como o valor da evidência. A avaliação da evidência implica a determinação do valor do rácio de verosimilhança (LR), representado por V.

$$V = \frac{P(E|Hp, I)}{P(E|Hd, I)}$$

Esta expressão é o rácio de verosimilhança de “probable cause”, ou seja, o LR em que o suspeito foi detido por razões não relacionadas com o seu perfil de ADN.

Se o suspeito é detido com base no resultado de busca numa base de dados, então nesse caso o LR vem

$$LR = \frac{P(E, D|Hp, I)}{P(E, D|Hd, I)}$$

em que D representa a informação adicional de que existe coincidência entre o ADN do suspeito e o da amostra recolhida da cena do crime, não acontecendo tal com os restantes membros da base de dados.

Proposições relevantes

É princípio fundamental que não é possível para o cientista fazer considerações sobre a verdade de uma proposição sem considerar pelo menos uma proposição alternativa. No âmbito forense é conveniente reduzir o número de proposições a apenas uma para representar o ponto de vista da defesa e outra para expressar o ponto de vista da acusação.

É largamente aceite que para a avaliação da evidência científica, os cientistas forenses devem considerar diferentes proposições que representem alternativas propostas pela acusação e pela defesa para ilustrar os factos em análise. Estas alternativas são representações formais do esquema das circunstâncias. Os cientistas forenses analisam a evidência segundo estas proposições. A formulação das proposições é a base crucial para uma abordagem lógica à avaliação da evidência.

Na prática as proposições escolhidas dependem das circunstâncias do caso, das observações dos factos que foram feitas, dados anteriores que sejam do conhecimento do cientista e do seu próprio domínio de especialidade. Isso significa que muitas vezes as proposições definidas pelo cientista podem ser retiradas das impostas pelo júri. É óbvio que as proposições consideradas devem ser mutuamente exclusivas.

Este procedimento pode ser sintetizado da seguinte forma e usando três princípios.

- 1) A avaliação só faz sentido quando pelo menos uma proposição alternativa é usada. As duas proposições são habitualmente designadas por H_a e H_d (a de acusação e d de defesa).
- 2) A avaliação da evidência científica (E) considera a probabilidade da evidência dadas as proposições usadas, ou seja, $P(E|H_a)$ e $P(E|H_d)$.
- 3) A avaliação da evidência científica é levada a cabo através de um conjunto de circunstâncias, designado por I. A avaliação é condicionada não apenas às proposições em causa mas também pela estrutura e conteúdo do esquema.

Assim, as proposições desempenham um papel importante neste processo. Geralmente as proposições são consideradas aos pares (pode haver situações onde existirão 3 ou mais) e são mutuamente exclusivas.

A definição/identificação das proposições a considerar não é uma tarefa simples. Elas dependem das circunstâncias do caso, das observações que foram feitas e da informação anterior disponível.

Uma classificação³, ou hierarquia, habitualmente usada consiste em considerar 3 categorias: nível I – source level; nível II – activity level e nível III – crime level.

No nível I ou nível de fonte ou origem a avaliação depende da análise e medição dos materiais recolhidos no suspeito e na cena do crime. As proposições deste nível são definidas com base em observações, medições e análises. Apenas se consideram as informações recolhidas durante a análise das provas. É feita a comparação entre duas amostras assumindo que são provenientes da mesma população se o suspeito é culpado e de populações distintas se o suspeito não é culpado. A probabilidade da evidência dada a 1ª proposição do par será considerada após uma cuidada comparação de duas amostras. A probabilidade da evidência dada a 2ª proposição do par será considerada por comparação, em princípio pelo menos, entre uma das amostras e qualquer tipo de população de fontes alternativas.

No nível II ou nível de actividade as hipóteses estão relacionadas com uma actividade ou acção. A consequência dessa actividade é o contacto físico e a transferência de materiais da cena do crime ou vítima para o suspeito e vice-versa. A definição destas proposições também requer a existência de observações, medições e análises, mas há outras questões a ter em conta. Para avaliar a 1ª proposição de cada par, para além de comparar observações de amostras de materiais recolhidos com amostras de controlo, é necessário considerar a probabilidade de transferência e persistência da evidência.

³ Aitken e Taroni (2004)

No entanto, a avaliação da evidência neste nível não depende apenas dos materiais recolhidos dado que pode haver ausência de materiais. Por vezes a ausência de materiais é também evidência com interesse.

Uma distinção entre proposição de nível I e de nível II está relacionada com o facto de que a interpretação dos pares do 2º nível não depende necessariamente dos materiais transferidos que foram encontrados. Outra diferença consiste no facto de as proposições de nível I exigirem pouco em termos de informação circunstancial, mas as do nível II não podem ser definidas sem um conjunto de circunstâncias. Assim tem de existir um grande nível de interação entre o cientista, investigador e/ou advogados que provavelmente não é exigido no nível I.

O nível III, o do topo, relaciona-se com uma ofensa/crime. É o nível mais elevado na hierarquia e corresponde às proposições que o júri/juiz deve considerar no julgamento. Em termos jurídicos as hipóteses que têm realmente interesse são as desta categoria. Também são proposições de actividade mas referem-se a crimes. Tal como as de nível II não dependem de material de evidência. Nas proposições de nível III há geralmente considerações que estão fora do domínio da ciência. De grande importância, é claro, se um crime ocorreu ou não e em alguns casos isso pode estar relacionado com a decisão se a palavra de qualquer testemunha é ou não credível em relação a outra.

O cientista está geralmente na posição de auxiliar o júri na determinação da verdade em relação a proposições que são intermediárias em relação ao objectivo final.

Pré – avaliação do caso

O processo de avaliação deverá começar quando o cientista tem conhecimento do caso pela primeira vez. Este processo pode ser sumariado nos passos seguintes. Recolha da informação sobre o caso que o cientista pode necessitar, consideração das questões/proposições e conseqüentemente o nível que o cientista escolhe para analisar a evidência, identificação dos parâmetros relevantes que vão aparecer no rácio de verosimilhança, avaliação do valor esperado do rácio de verosimilhança dada a informação anterior disponível, determinação da estratégia de análise, realização de testes e análise de resultados e cálculo do rácio de verosimilhança e relato do seu valor.

Há uma tendência geral na comunidade forense de ver a fase de interpretação como estando próxima do fim da análise do caso/crime, na altura da preparação do relatório formal do caso. Mas existem muitas vantagens no alargamento desta perspectiva.

O cientista deve fazer uma apreciação adequada das circunstâncias do caso para poder formalizar de forma correcta o esquema que as diligências/observações/análises que devem ser feitas e o que se deve esperar como resultado.

Avaliação do caso

Seguidamente o cientista interpreta a evidência através do cálculo do rácio de verosimilhança (LR), que é central na formulação bayesiana.

$$LR = \frac{\text{probabilidade da evidência se a proposição da acusação é verdadeira}}{\text{probabilidade da evidência se a proposição da defesa é verdadeira}}$$

Na fase de pré-avaliação as expectativas podem transformar-se em distribuições de probabilidades do valor esperado da evidência dadas as circunstâncias.

Não é prática usual os cientistas forenses fazerem os seus relatórios em termos do LR (a não ser em situações de ADN). No entanto, o LR dá a ideia de uma escala verbal que incorpora a noção de suporte qualificado em termos equivalentes a uma escala numérica do LR.

Para interpretar a evidência, o cientista forense deve considerar diferentes proposições que representam as alegações da acusação e da defesa em tribunal. Houve sempre a tendência, no sistema de justiça, para o cientista ser chamado a tribunal pela acusação para expressar a sua opinião em relação à proposição da acusação e depois responder às proposições alternativas colocadas pela defesa. Em muitos casos esta situação, provavelmente injustamente, coloca o cientista como partidário/defensor da acusação, defendendo a posição da acusação dos ataques.

Reconhece-se que o cientista deve considerar as proposições da defesa mas a cultura criou um ambiente em que o cientista toma uma posição e na preparação para tribunal pensa em todas as possíveis proposições a serem colocadas pela defesa e mentalmente prepara as respostas para elas. A formação destes cientistas prepara-os para reagir e responder a questões do tipo “que diria de a defesa alegar que...”.

Um tema importante e muito discutível é o da individualização, ou seja, estabelecimento que determinado traço de evidência encontrado na cena do crime é proveniente de um único indivíduo ou fonte.

Diz-se que a criminalística é a ciência da individualização. Em todo o mundo e nos mais variados crimes procede-se à identificação forense da origem do traço encontrado na cena do crime: arma, mão, dedo, instrumento, etc.

Mas a capacidade de fazer tais tipos de identificação depende de uma série de premissas:

- 1) A existência de diversos tipos de identidades físicas e biológicas é única;
- 2) Deixam vestígios únicos;
- 3) As técnicas de observação, medição e inferência utilizadas na ciência forense são adequadas para ligar os traços ao objecto que os produziu.

Interessa também aqui o nível das proposições usadas. A fase inicial envolve a consideração por parte do cientista do que deve ser esperado obter na examinação. Este processo deve ser formalizado. Esta fase requer que o cientista defina claramente as proposições a serem testadas, à luz das circunstâncias.

Profiler

É um termo que já faz parte do vocabulário comum. Pode definir-se como um desenhador de perfis a partir de um conjunto de indícios, evidências, recolhidos na cena de um crime. É um técnico de investigação forense capaz de descrever e prever comportamentos ou factos futuros. É um especialista na definição de perfis psicológicos de pessoas, em especial de criminosos. Mediante os indícios relativos ao crime em análise, ele ajuda a esclarecer as questões sobre o tipo de personalidade do criminoso.

Ele faz a “autópsia psicológica” do crime, estuda todos os antecedentes do criminoso e especialmente aqueles relacionados com o crime.

A maior parte dos profilers são psicólogos e têm conhecimentos de criminologia. A formação base de um profiler deve ser ampla e abranger estudos de psicologia, avaliação de personalidade, criminologia, antropologia sociologia, ciências do comportamento.

Pode dizer-se que em Portugal não há profilers dado que não existe formação adequada. No entanto, algumas pessoas desempenham essas funções. Alguns bons técnicos de investigação exercem essa função. Tradicionalmente os investigadores baseavam-se no seu senso comum, intuição perícia e experiência para investigar. Mais recentemente, psicólogos aplicam teorias desenvolvidas em universidades e hospitais à investigação de crimes reais.

Specific case analysis e *offender profiling* são designações dadas a um conjunto de teorias e técnicas científicas e psicológicas que procuram relacionar o acusado, vítima e cena do crime numa investigação.

Pode colocar-se a questão se as actividades de um Profiler poderão ser consideradas arte ou ciência. Podemos considerar que é ciência com um pouco de arte.

4.O modelo bayesiano

É frequente argumentar-se que dados frequentistas são apenas uma parte, e não uma parte necessária, da abordagem bayesiana à avaliação da evidência. A abordagem bayesiana considera a probabilidade como uma medida de um grau de crença pessoal, específico de um indivíduo. Como tal pode variar de indivíduo para indivíduo porque a informação que cada um possui é geral e rigorosamente diferenciada (Paulino *et al.* (2003)).

Assim, é possível a combinação de probabilidades objectivas baseadas em dados (mas exigindo subjectividade na definição dos modelos subjacentes) e probabilidades subjectivas onde o conhecimento e experiência do cientista podem ser importantes para a sua determinação.

Também os juristas se interessam pela determinação de probabilidades usando o conceito subjectivo de probabilidade. O seu cálculo pode ser útil em questões associadas com a credibilidade de testemunhas e com as conclusões que podem ser tiradas dos seus testemunhos.

A definição frequentista de probabilidade envolve uma grande sequência de repetições de determinado fenómeno em condições idênticas. Segundo esta abordagem a probabilidade de determinado acontecimento é definido como o limite da frequência relativa de ocorrência do acontecimento quando se repete um grande número de vezes.

As probabilidades frequentistas são probabilidades objectivas. São objectivas no sentido em que existe um conjunto de circunstâncias bem definidas de repetição de experiências tal que as correspondentes probabilidades são definidas de modo que os pontos de vista pessoais não vão alterar os seus valores. Qualquer pessoa que considere tais circunstâncias definirá o mesmo valor para as probabilidades.

O modelo frequencista assume que as frequências relativas obtidas numa longa sequência de repetições, obtidas de forma idêntica e fisicamente independentes umas das outras. No entanto, tais circunstâncias enfrentam muitas dificuldades. Em muitas situações não podemos assumir resultados igualmente prováveis e não podemos observar as ocorrências para determinar as frequências relativas. Por vezes o acontecimento de interesse já ocorreu ou só ocorre uma vez.

A ideia subjacente às probabilidades subjectivas é que a probabilidade que um acontecimento ocorra reflecte uma medida da crença pessoal na ocorrência desse acontecimento. Probabilidade subjectiva, por vezes designada por probabilidade bayesiana, é definida como o grau de crença detida por uma pessoa e baseada em todo o seu conhecimento, experiência e informação, relativamente a determinado acontecimento.

Acontecimentos ou parâmetros de interesse nas mais diversas áreas, por exemplo na análise forense, não são habitualmente o resultado de processos repetitivos ou replicáveis. Tais acontecimentos são singulares, únicos ou de tipo único. Nestas situações calculam-se probabilidades recorrendo ao conceito subjectivo de probabilidade.

A avaliação subjectiva de probabilidades pode ser vista como parte do esquema bayesiano. O teorema de Bayes indica como uma avaliação inicial da probabilidade deve ser alterada como resposta à aquisição de nova informação.

As probabilidades *a priori* são determinadas através das crenças pessoais sobre o assunto. A escolha dessas probabilidades é o resultado de procedimentos complexos e muito dependentes do contexto, envolvendo elementos psicológicos. Estes elementos têm um papel essencial na avaliação das probabilidades. No entanto, a formalização de probabilidades em termos subjectivos ou personalistas exige que o investigador satisfaça o princípio de coerência ou de consistência. Consequentemente, diferentes investigadores possuem, em regra, diferentes distribuições *a priori* para o mesmo parâmetro sem deixarem, necessariamente de ser coerentes (Paulino *et al.* (2003)).

A característica essencial da inferência bayesiana é a passagem da probabilidade *a priori* para a probabilidade *a posteriori*. Isto é visto como um pré-requisito para qualquer indução válida. Pode ser visto como uma simples mudança de opinião com resposta ao aparecimento de nova informação.

O paradigma bayesiano modela o processo científico. De facto, a ciência progride com os cientistas a alterarem as suas opiniões à medida que a informação se acumula e com os cientistas a tentarem persuadir outros cientistas da validade das suas opiniões.

Elicitação de opiniões, actualização depois da observação de dados e quantificação da incerteza usando distribuições de probabilidades são partes da estatística bayesiana.

Durante décadas a evidência relativa a fluidos corporais foi apresentada em tribunal sem recorrer ao modelo bayesiano e é natural sugerir-se que a apresentação da evidência através do rácio de verosimilhança é complicada e desnecessária. Porque não dar ao tribunal uma estimativa da frequência do tipo de ADN observado? Em casos simples pode assumir-se que a abordagem frequencista é tão eficaz como o rácio de verosimilhança, mas quando o caso se complica a abordagem frequencista falha.

Quando o caso é complicado a abordagem bayesiana é inevitável. Mas esta abordagem não é universalmente aceite. O modelo bayesiano não é resposta a todos os problemas do processo legal, mas pode dizer-se que é o melhor método para perceber a interpretação da evidência científica.

Podemos considerar que no âmbito da avaliação da evidência científica no contexto legal a abordagem bayesiana é dominante. Mas não se pode esquecer que á apenas um modelo matemático e como tal tem limitações, de entre as quais as mais importantes dizem respeito a dificuldades na definição das probabilidades *a priori* e ao facto do conceito de dúvida razoável não ser quantificável. Efectivamente, a determinação e interpretação da distribuição *a priori* contam-se entre os pontos mais melindrosos e controversos da teoria bayesiana e constituem um dos principais obstáculos à respectiva implementação (Paulino *et al.* (2003)).

Para avaliar a incerteza relacionada com uma dada proposição é necessário considerar pelo menos uma proposição alternativa. Calcula-se o odds rácio a partir do teorema de Bayes

$$\frac{P(H_a|E,I)}{P(H_d|E,I)} = \frac{P(E|H_a,I)}{P(E|H_d,I)} \times \frac{P(H_a|I)}{P(H_d|I)}$$

Esta equação é de importância capital na interpretação forense e permite uma clara distinção entre o papel do júri/juiz e do cientista.

O juiz (ou o tribunal) precisa de resposta a questões tais como:

- i) Qual a probabilidade de que a proposição da acusação seja verdadeira dada a evidência? Ou seja, qual o valor de $P(H_a|E,I)$.
- ii) Qual a probabilidade de que a proposição da defesa seja verdadeira dada a evidência? Ou seja, qual o valor de $P(H_d|E,I)$.

Por outro lado o cientista forense quer ver respondidas questões de tipo completamente diferente:

- i) Qual a probabilidade da evidência se a proposição da acusação é verdadeira? Ou seja, qual o valor de $P(E|H_p,I)$.
- ii) Qual a probabilidade da evidência se a proposição da defesa é verdadeira? Ou seja, qual o valor de $P(E|H_d,I)$.

O cientista deve sempre colocar e responder a questões do tipo “qual a probabilidade da evidência dada a proposição” e não “qual a probabilidade da proposição dada a evidência”. O cientista preocupa-se em analisar o rácio de verosimilhança e o tribunal as probabilidades a posterior. O cientista deve avaliar a evidência não apenas sob a condicionante H_a e H_d mas também segundo a condicionante I , outra informação relevante.

Bibliografia

Aitken, C.G.G. e Taroni, F.(2004) “*Statistics and the Evaluation of Evidence for Forensic Scientists*”, 2ª edição, John Wiley & Sons, Ltd.

- Aitkin, M. (1991) "Posterior Bayes Factors". *Journal of the American Statistical Society*, Serie B, **53**, 111-142.
- Balding, D. J. (2005) "Weight – of – evidence for Forensic DNA Profiles". John Wiley & Sons, Ltd.
- Balding, D.J. (1995) "Estimating Products in Forensic Identification using DNA Profiles". *Journal of the American Statistical Association*, **90**, 839-844.
- Balding, D.J. (1999) "When can a DNA profile be regarded as unique". *Science & Justice*, **39**, 257-260.
- Bayarri, M.J. e Berger, J.O. (2004) "The Interplay of Bayesian and Frequentist Analysis". *Statistical Science*, **19**, 58-80.
- Biedermann, A., Taroni, F. e Garbolino, P. (2007) "Equal prior probabilities: can one do any better?". *Forensic Science International*, **172**, 85-93.
- Curran, J. M., Walsh, S. J. e Buckleton, J. (2007) "Empirical testing of estimated DNA frequencies". *Forensic Science International: Genetics*, **1**, 267-272.
- Dawid, A.P. (2004) "Probability, Causality and the Empirical World: A Bayes - de Finetti – Popper - Borel Synthesis". *Statistical Science*, **19**, 44-57.
- Dawid, A.P. e Mortera, J. (1996) "Coherent use of Identification Evidence". *Journal of the Royal Statistical Society*, serie B, **58**, 425-443.
- Dawid, A.P. e Mortera, J. (1998) "Forensic Identification with Imperfect Evidence". *Biometrika*, **85**, 4, 835-849.
- Evetts, I. W. e Weir, B. S. (1998) "Interpreting DNA Evidence – Statistical Genetics for Forensic Scientists". Sinauer Associates, Inc.
- Evetts, I. W. (1996) "Expert evidence and forensic misconceptions of the nature of exact science". *Science & Justice*, **36**, 118-122.
- Evetts, I. W., Cagne, P. E. e Aitken, C. G. G. (1987) "Evaluation of the likelihood ratio for fibre transfer evidence in criminal cases". *Applied Statistics*, **36**, 174-180.
- Fienberg, S.E. e Straf, M.L. (1991) "Statistical Evidence in the US Courts: an Appraisal". *Journal of the Royal Statistical Society*, serie A, **154**, 49-59.
- Foreman, L.A., Smith, A.F.M. e Evetts, I.W. (1997) "Bayesian Analysis of DNA Profiling Data in Forensic Identification Applications". *Journal of the Royal Statistical Society*, serie A, **160**, 429-469.
- Garbolino, P. e Taroni, F. (2002) "Evaluation of scientific evidence using Bayesian networks". *Forensic Science International*, **125**, 149-155.
- Good, I. J. (1992) "The Bayes/non Bayes compromise: a brief review". *Journal of the American Statistical Association*, **87**, 597-606.
- Grove, D.M. (1980) "The interpretation of forensic evidence using a likelihood ratio". *Biometrika*, **67**, 243-246.
- Inman, K. e Rudin, N. (2002) "The origin of evidence". *Forensic Science International*, **126**, 11-16.
- Lindley, D.V. (1991) "Subjective Probability, Decision Analysis and their Legal Consequences". *Journal of the Royal Statistical Society*, serie A, **154**, 83-92.
- Neves, Manuela Figueira e Rosado, Fernando (2007) "A Importância da Estatística na Ciência Forense", XV Congresso Anual da Sociedade Portuguesa de Estatística, Lisboa.

- Neves, Manuela Figueira e Rosado, Fernando (2007) “On the Importance of Statistics in Forensic Science”, 56th Session do International Statistical Institute, Lisboa.
- Neves, Manuela Figueira e Rosado, Fernando (2008) “Sobre a Crescente Importância da Estatística na Ciência Forense”, Technical Report CEAUL 1/2008.
- Neves, Manuela Figueira e Rosado, Fernando (2008) “Forensic Statistics and Outlier Identification”, *7th International Conference on Forensic Inference and Statistics*, Lausanne, Suíça.
- Neves, Manuela Figueira e Rosado, Fernando (2009) “Bayes no Tribunal”, XVII Congresso Anual da Sociedade Portuguesa de Estatística, Sesimbra.
- Neves, Manuela Figueira (2009) “A Estatística no Tribunal”, XVII Congresso Anual da Sociedade Portuguesa de Estatística, Sesimbra.
- Paulino, C. D.; Turkman, M.A.A. e Murteira, B.J.F (2003) “Estatística Bayesiana”. Ed. Fundação Calouste Gulbenkian.
- Taroni, F., Aitken, C.G.G., Garbolino, P. e Biedermann, A. (2006) “*Bayesian Networks and Probabilistic Inference in Forensic Science*”, John Wiley & Sons, Ltd.