研究会報告

見かけの複雑さの中の真実：あるDNA鑑定をめぐって

柳川　堯 *

1　序

DNAパラダイムでは、ヒトのDNAは個人によって完全に相異するとする。DNA鑑定は、このパラダイムを前提として血液や精液等の体液や毛根の細胞片等からDNAを抽出して個人識別を行う。鑑定法の開発者 Jeffereys, A.(1985) は、その論文の中でDNA鑑定の個人識別力を 3×10^{-10} と評価している。彼の方法は、日本でもいち早く注目され、新しい技術の開発や改良が重ねらて各地の県警レベルで実際の事件に供され始めている。新聞やＴＶ等のマスコミによればその個人識別力は「血液型鑑定と併用することで百万分の１程度」であるという。

個人識別力が「百万分の１」であるとは、「犯行現場から採集されたDNAが被告のものでない確率は百万分の１」を意味すると理解されている。この理解が新聞やＴＶ等のマスメディアによって流布され、また裁判官、検事、弁護士等法曹界の人々もそのように理解しており、裁判に大きな影響を与えている。この理解は正しいのであろうか。本小論では、個人識別力「百万分の１」が何を意味するのかについて考察する。さらに具体例として興掛訴訟での三澤鑑定書を考察することによって、DNA鑑定は実際に信頼できないものもあることを指摘する。最後に、より合理的な個人識別の方法について言及する。

2　Match-Binning 法

DNA鑑定は二つの仮説

H_o: the crime sample is from the suspect

H_1: the crime sample is from the other person

のどちらかをデータに基づいて決定する。ここではデータとして、電気泳動法を用いて測定される父方と母方から由来する二つのDNAの fragment length を考える、つまりデータは

(x_1, x_2): fragment length of the suspect

(y_1, y_2): fragment length of the crime

である。一般性を失うことなく $x_1 \le x_2 \quad y_1 \le y_2$ としておく(1).

現行の鑑定では、仮説 H_o と H_1 の決定は、Match-Binning 法とよばれるつぎのような考え方でおこなわれている。

*九大数理

(1) $x_1 \neq y_1$, または $x_2 \neq y_2 \Longrightarrow H_1$と決定

(2) $x_1 = y_1$, かつ $x_2 = y_2 \Longrightarrow H_o$と決定する。しかし、偶然 the other person の fragment length が (y_1, y_2) と一致する可能性がある。そこでこの可能性を確率的に評価する目的で、まず

(i) 第 3 者、即ち日本人集団の fragment length のデータベースを作成し、fragment length の分布を調べる。つぎに

(ii) このデータベースからランダムに一人を抽出したとき、抽出された人の fragment length (Z_1, Z_2) が fragment length of the crime (y_1, y_2) と一致する確率を計算する、即ち関係式

第 3 者の fragment length が偶然 (y_1, y_2) と一致する確率

$$= 2P[Z_1 = y_1, \quad Z_2 = y_2|H_1]^{(2)}$$
$$= 2P[Z_1 = y_1|H_1]P[Z_2 = y_2|H_1]^{(3)}$$

から求める。DNA 鑑定法の「個人識別力」とはこの確率のことである。

注 1) 短いのと長いののどちらが父方に由来するかは分からない.

注 2) 条件 $y_1 \leq y_2$ より確率は 2 倍されなければならない.

注 3) Hardy-Wienberg を仮定する。この仮定より Z_1 と Z_2の独立性が導かれる。確率 $P[Z_1 = y_1|H_1]$ 及び $P[Z_2 = y_2|H_1]$ の値はデータベースから求める。例えば三澤鑑定書では y_1=262, y_2=301 で $P[Z_1 = 262|H_1] = 0.0385, \quad P[Z_2 = 301|H_1] = 0.0231$ と推定されている。なお三澤鑑定書では注 2 で述べた 2 倍が脱落しており

$$個人識別力 = 0.0385 \times 0.0231 = 8.8 \times 10^{-4}$$

とされている。

3 Match-Binning 法に対する批判

3.1 個人識別力は意味をなさない

前節で見たように DNA 鑑定の個人識別力とされているのは、データベースからランダムに抽出した一人の人の fragment length が crime sample の fragment length と一致する確率のことである。三澤鑑定ではこの確率を上記のように 8.8×10^{-4}と見積もっている。かくして三澤鑑定では「crime sample が第三者のものである確率は極めて微小である。よって、この crime sample は suspect のもの以外ではありえない」と推論しているが、この確率は、正しい推論の根拠になりえない。なぜなら、

(i) 抽出した一人の人の fragment length が crime sample 以外の fragment length と一致する確率も 10^{-4}のオーダーでありえるからである。つまり話しは 10^{-4}のオーダーの世界の中の話しであって、この世界の中で 8.8×10^{-4}が果して微小であるかどうかというのは別問題である。このことを理解するためのポイントは「測定誤差」である。具体的に例に基づいて考えよう。

(ii)「身長 172cm」という情報に対して「身長 172cm、体重 65kg、男性、左利き、眼鏡着用」という情報は一見して個人識別力が高いように思われる。上記のような評価式を適用して確率的に評価すればこの情報の個人識別力は 10^{-4}のオーダーくらいでありえる。しかしながら、その中の「左利き」が決定的に間違った情報ならば、個人識別は決定的に間違ってしまう。このとき「身長 172cm」という情報の方が「身長 172cm、体重 65kg、男性、左利き、眼鏡着用」という情報よりもむしろ高い個人識別力をもつことは自明である。これは極端な例であるが、「左利き」が約 7 割の信頼度しかおけな

いような場合はよくある。このとき、「身長172cm、体重65kg、男性、左利き、眼鏡着用」という情報に基づく個人識別は決定的に間違うことはないが、約3割程度の間違いをもたらす。にもかかわらず上記の評価式で個人識別力を求めるとこの情報の個人識別力は10^{-4}のオーダーでありえる。(4)

(iii) つまりこの鑑定には「百万分の一」の個人識別力があるといってもそれだけのことで、この確率を「crime sample が suspect のそれと一致する蓋然性が高い」という推論の根拠とするには難がある。「左利きの信頼度約7割」というような測定誤差を考慮にいれた蓋然性推論をその根拠に置き換えるべきである。

注 4) 表がでる確率が 1/2(つまり単独では識別力をもたない)であるコイン投げを18回ほど組み合わせる(つまり18回続けて投げる)と、ここでいう個人識別力 $=0.5\times0.5\times\ldots\times0.5$ と計算され、この値はほぼ百万分の一のオーダーとなる。しかしながら、このような鑑定はまったく識別力をもたないことは自明である。

3.2 測定誤差の無視

電気泳動法による fragment length の測定は、fragment length(bp) が既知のマーカーを試料に加えて電圧をかけ、100bp(或いは200bp)からのマーカーの移動距離 (mm) を測定し、移動距離と fragment length(bp) の関係を示す基準線を求め、この基準線を利用して試料の移動距離 (mm) からその fragment length(bp) を逆算する。移動距離の測定は、相当大きな誤差を伴う。米国では、このことがよく認識され fragment length の一致は適当な幅をつけて評価されている、即ちあらかじめ与えた定数 δ_1とδ_2に対して $|x_1-y_1|>\delta_1$ または $|x_2-y_2|>\delta_2 \Longrightarrow \mathrm{H}_1$ と決定される。日本では、誤差が実際にあるにもかかわらず、これを無視している。誤差は以上の測定誤差のほか、一回の電気誘導では、8乃至10本の fragment しか同一ゲル上でとりあつかえないという制約があって、多数の fragment を測定してデータベースを作成するには、相異するいくつものゲル上で実験や測定を行わなければならないが、このとき微妙な測定環境の差からも誤差が生じる。言い替えれば求められる基準線にかなり大きなバラツキがみられる。三澤鑑定に基づいてこれらの誤差を試算すると、fragment length (bp) は 2bp や 4bp 程度でまとめても意味をなさない、即ち三澤鑑定では、二つの fragment length が相違してもそれが 10bp 程度以下なら真の差異のためではなく誤差のための差異にすぎない可能性があることが推察された。

これらの測定誤差は、crime sample および suspect sample の測定に対しても当然生じていると考えなければならない。実際、三澤鑑定に使用された原資料では、次のようなデータが与えられている。

1. (被告の新鮮血)

短バンド	長バンド
$119mm : 259.621bp$	$136mm : 297.523bp$
$120mm : 261.658bp$	$137mm : 300.000bp$
$121mm : 263.716bp$	$138mm : 302.509bp$

2. （16の1,10,1の毛髪資料）

短バンド	長バンド
$119mm : 259.621bp$	$136.5mm : 298.757bp$
$119.5mm : 260.637bp$	$137mm : 300.000bp$
$120mm : 261.658bp$	$137.5mm : 301.251bp$

測定にゆらぎがあるため幅をつけて吟味しているが、どうしたことか鑑定書では、これを乱暴にもエイヤッ！とまとめて $(x_1, x_2) = (y_1 . y_2) = (262, 301)$ としている。誤差を考慮すると、$(x_1, x_2) = (260, 303) : (y_1 . y_2) = (262, 299)$ である場合も完全には否定できない。前項でも指摘したように、測定誤差を確率的につかまえて、このような可能性をも考慮に入れて蓋然性の推論を行わない限り、科学的な個人識別は達成できない。

3.3 データベースの信頼性

「日本人の fragment length のデータベース」を作成するとして三澤鑑定では血縁関係を持たない65人の人が抽出され fragment length が測定された。抽出の仕方の詳細は不明であるが、そのデータベースの fragment length の分布は、同一の方法（ＡＣＴＰ２）で大分医科大グループが作成した fragment length の分布と有意に相違する（カイ二乗検定の p 値＜0.001)。このことから明らかなように分布の地域差、測定の系統誤差などに十分な配慮をおこなっておかなければ信頼度の高いデータベースはえられない。

4 妥当な一つの方法

前節で指摘したように、本来本質的な測定誤差があるにもかかわらず、これを無視するのは蓋然性の推論を行う上で致命的な誤りである。本節では、妥当な方法として、誤差を確率分布としてとらえ、尤度の概念を利用して蓋然性の推論を行う方法を提示する。前節までと同様に
H_o: the crime sample is from the suspect
H_1: the crime sample is from the other person
としておく。

データは、fragment length of the suspect (x), fragment length of the crime(y)、およびデータベースとして収集された測定値（$\underline{z} = (z_1, z_2, ..., z_m)$）である。このデータに基づいて H_oか H_1を判定するには、このデータが得られたという条件の下で H_oである確率と、H_1である確率を比較すればよい、つまり比 $P(H_o|x, y, \underline{z})/P(H_1|x, y, \underline{z})$ の大きさを評価すればよい。この比が大きければおおきいほど"the crime sample is from the suspect"である蓋然性が大きい。ところで次の等式が成立する：

$$\frac{P(H_o|x, y, \underline{z})}{P(H_1|x, y, \underline{z})} = L \times \frac{P(H_o|\underline{z})}{P(H_1|\underline{z})}$$

ここに

$$L = \frac{P(x, y|H_o, \underline{z})}{P(x, y|H_1, \underline{z})}$$

は数理統計学で尤度比とよばれる量である。上の等式から明らかなように尤度比Lの値が大きければおおきいほど'the crime sample is from the suspect'である蓋然性が大きい。

以下では尤度比Lを求めるための数学モデルについて考察する。いま、測定誤差をモデル化するため観測値x,yに対応する確率変数を

$$X = (X_1, X_2) \qquad \text{(fragment length of the suspect)}$$
$$Y = (Y_1, Y_2) \qquad \text{(fragment length of the crime)}$$

として

$$X_1 = \mu_1 + \epsilon_1 \qquad X_2 = \mu_1 + \epsilon_2$$
$$Y_1 = \mu_2 + \epsilon'_1 \qquad Y_2 = \mu_2 + \epsilon'_2$$

とおく。ここに $\epsilon_1, \epsilon_2, \epsilon'_1, \epsilon'_2$ は観測誤差を表す互いに独立な確率変数で同一分布 $F(x/\sigma_o)$ に従うと仮定する。(5)

このモデルでは H_o, H_1 はそれぞれ

$$H_o : \mu_1 = \mu_2 \qquad H_1 : \mu_1 \neq \mu_2$$

と表される。ところで上の解説で登場した確率 $P(H_o|x, y, \underline{z})$ が妥当性をもつためには H_o (および H_1) は確率事象でなければいけない。そのため μ_1, μ_2 は、先験分布 $G((\mu-\theta)/\sigma)$ に従う互いに独立な確率変数であると仮定する。この先験分布に関する推定は経験的ベイズ法を適用して $\underline{z}$ に基づいて行うことにする。問題は分布関数Fの尺度母数 σ_o の推定である。これは同一資料を繰り返し測定しなければ推定できない (6)。ここでは、仮に既知としておく。

さて、上述の枠組みでは H_o の下では $\mu = \mu_1 = \mu_2$ は分布 $G((\mu-\theta)/\sigma)$ に従うので、データxとyが得られたときの尤度は

$$L_1 = \frac{1}{\sigma_o^4} \int f(\frac{x_1-\mu}{\sigma_o}) f(\frac{x_2-\mu}{\sigma_o}) f(\frac{y_1-\mu}{\sigma_o}) f(\frac{y_2-\mu}{\sigma_o}) dG(\frac{\mu-\theta}{\sigma})$$

また、H_1 の下では $\mu_1, \mu_2 (\mu_1 \neq \mu_2)$ は互いに独立に分布 $G((\mu-\theta)/\sigma)$ に従うので、データxとyが得られたときの尤度は

$$L_2 = \frac{1}{\sigma_o^4} \int f(\frac{x_1-\mu_1}{\sigma_o}) f(\frac{x_2-\mu_1}{\sigma_o}) dG(\frac{\mu_1-\theta}{\sigma_o}) \int f(\frac{y_1-\mu_2}{\sigma_o}) f(\frac{y_2-\mu_2}{\sigma_o}) dG(\frac{\mu_2-\theta}{\sigma})$$

で与えられる。これらから尤度比 $L = L_1/L_2$ が求まる。この式から実際に尤度比Lの値を計算するには測定を繰り返して得られるデータに基づいて分布関数Fに関する知見を得ておく必要がある。尤も簡単な方法はデータを適当に変換して近似的に正規分布に従うとして正規分布の理論を適用すればよい。

注5) 観測誤差 ϵ_1 と ϵ_2 は同一ゲル等を通じて混入する誤差であるから実際には相関があると推察されるが、相関を調べる計量的データが入手できないので作業仮説として独立性を仮定した。(ϵ_1, ϵ_2) と $(\epsilon'_1, \epsilon'_2)$ の独立性はさほど問題にならないと思う。

注6) 三澤鑑定では、前節で与えたデータが、測定値として求められた被告人の唯一のデータである。せめて被告人の毛髪を10本ほど繰り返し測定して測定値のゆらぎに関する情報を得ておくべきであった。