

# 米国における統計的証拠の意義と技法 —もう一つの法情報学—

田 中 規久雄

はじめに

1. 科学的証拠に対する米連邦最高裁の態度
2. 科学的証拠に対する統計的証拠の位置
3. 統計的証拠の基礎技法
4. 統計的証拠の機能
5. 統計的証拠解釈における誤謬
  - 5-1. 処理の誤謬
  - 5-2. 解釈の誤謬
    - 5-2-1. 冤罪率の誤謬
    - 5-2-2. 無差別原理の誤謬
    - 5-2-3. 訴追者の誤謬
    - 5-2-4. 弁護者の誤謬

おわりに

はじめに

法情報学(legal informatics)の対象とする範囲は、基礎法学的な色彩を有する「法の情報科学(information science of law)」から、実際的なデータベース、法律人工知能(LAI, legal artificial intelligence)、法律実務や法学教育に有用な法情報システムの構築といった「法の情報工学(information technology of law)」までに至る<sup>1)</sup>。

---

1) 筆者は、これらをコンピュータに載せられるようにモデル化する研究分野を「計算法理学(computational jurisprudence)」と呼んでいる。

この中で、基礎法学としての法情報学が寄与する分野に法統計学(legal statistics)がある。法統計学は広い意味では社会統計学として、法社会学<sup>2)</sup>や刑事学<sup>3)</sup>の研究方法の一つとなるのみならず、科学的証拠(scientific evidence)の作成、解釈といった法律家の実務において一定の理解が必要とされる分野である。

O. W. Holmes は、「現在は太文字 (black letter) が重視されるが、将来は統計学の素養を持ち経済学を知る者が合理的な法律研究をする者となるだろう。」<sup>4)</sup> という有名な言葉を残しているが、それから百十余年、科学の発達と共に、今日の我が国でも、公害裁判や DNA 鑑定を初めとして、科学的証拠は専門家証人 (expert witness) に任せておけばよいものではなく、法律家自身が誤った解釈をしない為にも<sup>5)</sup> 一定の理解が求められるものとなっている<sup>6)</sup>。

## 1. 科学的証拠に対する米連邦最高裁の態度

本稿は、科学的証拠の許容性(admissibility)を論ずるものではないが、最初に米連邦最高裁の態度を簡単に見ておこう<sup>7)</sup>。

米国で科学的証拠に対する基準を初めて明確にしたのは、1923 年、D.C.(ワシントン)の連邦控訴裁が下した Frye 判決<sup>8)</sup> だといわれる。Frye 判決は、専門家証言 (expert testimony) の「根拠は、その専門分野での一般的承認 (general acceptance) を得た事が十分立証されねばならない。」

---

2) 川島武宜編『法社会学講座 3・法社会学の基礎 1』「第 III 章 法社会学研究の技法」1972 年等。

3) 浜井浩一『犯罪統計入門』2006 年等。

4) Oliver W. Holmes, THE PATH OF THE LAW, 10 HARV. L. REV. 457, 469 (1897). 'black letter' を「(教科書等の)太文字」としたが、「条文」や「訓詁学」等とも訳されている。

5) 統計が現象に対する誤導や偏見をもたらす危険性については、古典的名著とされる、Darrell Huff, HOW TO LIE WITH STATISTICS, 1954. (高木秀玄訳『統計でウソをつく法』1968 年)以来、枚挙に暇がない。

6) 「不確実な科学的状況での法的意思決定プロジェクト・法と科学」(<http://www.law-science.org/top.html>)、デジタル・フォレンジック研究会 (<http://www.digitalforensic.jp/>)等、参照。

7) 詳しくは、井上正仁「科学的証拠の証拠能力(1)、(2)」研修(法務総合研究所)、1995 年、560 号 3-16 頁、562 号 6-14 頁、徳永光「DNA 証拠の許容性：Daubert 判決の解釈とその適用」一橋法学 1 巻 3 号、2002 年、807-860 頁等、参照。

8) Frye v. United States, 54 App. D. C. 46, 293 F. 1013, No.3968 (D.C. Cir. 1923).

とするものだった。しかし、その後も裁判所は揺れ動き、1975年「連邦証拠規則(Federal Rules of Evidence)」<sup>9)</sup>も下記のように定めるのみで、一般的承認というFrye基準(Frye test)を内包するものなのか、排除するものなのかは不明で混乱は解決されなかった。

#### 連邦証拠規則(1975年当時版、要訳)<sup>10)</sup>

##### 第4章 関連性(relevancy)とその限界

401条 「関連性ある証拠(relevant evidence)」の定義：その証拠がない場合よりも、事実が存在する蓋然性をより高くあるいは低くする力(tendency)のある証拠である。

402条 関連性ある証拠には一般的に証拠能力が認められる(admissible)が、関連性のない証拠には認められない。

403条 偏見、混乱、無用な引き伸ばしといった根拠による関連性ある証拠の排除：そうした危険が証明力(probative value)を上回るときには排除する。

##### 第7章 意見と専門家証言

702条 専門家による証言：事実認定者(trier)の証拠の評価や事実の判定を助ける(assist)場合には専門家証言が許される。

703条 専門家証言による意見表明の根拠：専門家証人が意見や推論形成に際して合理的に依拠した(reasonably relied upon)特定の分野における事実やデータ自体(facts or data)に証拠能力は不要である。

9) PUBLIC LAW 93-595-JAN. 2, 1975, 88 STAT., pp.1931-32, p.1937. 民刑事共通に適用される。裁判基準としては以下の通りである。民事訴訟の証明度(standard of proof):相手方の証拠よりも説得的な証拠を提出して主張事実の可能性が「ないよりはる (more-likely-than-not)、50%超原則」、「証拠の優越(preponderance of evidence)」、「単なる優勢(mere preponderance)」の証明で足りる。名誉毀損(libel and slander)、子供の監護権(child custody)、詐欺(fraud)、不当威圧(undue influence)、父子関係確定等の特定の訴訟：高めの証明度である「明白かつ説得的な証明(clear and convincing proof)」もしくは「相当高度の優勢(heavy preponderance)」が必要である。刑事訴訟：犯罪事実について「合理的な疑いを越える証明(proof beyond a reasonable doubt, preponderance so great as to eliminate all reasonable doubt)」が必要である。田村陽子「アメリカ民事訴訟における証明論：『法と経済学』的分析説を中心に」立命館法学 339・340号、2011年、202頁注11、219-220頁。

10) なお、現連邦訴訟規則は、後述するような連邦最高裁判決に沿った改正がなされている。

こうした中、1993年のDaubert判決<sup>11)</sup>で連邦最高裁が態度を明らかにする。Blackmun判事の法廷意見は、Frye基準を破棄し、上記連邦証拠規則に従って、事実審(trial court)は「信頼性(reliable)」と「関連性(relevant)」の基準に従い、危険な証拠は排除するといった「柔軟」な判断が求められるとのルール(Daubert standard)を示した。ここで信頼性とは、科学として正確かという「科学性(scientific)」、関連性とはそのケースの判断にとって有用かという「有用性(helpful)」を指していると考えるとわかりやすい。そしてDaubert判決はこの科学性と有用性について、事実審の判事は『「門番(gatekeeper)』として行動せねばならず、専門家の証言前に、提出された科学的専門知識の妥当性を評価せねばならない。』としたのである。この事はRehnquist首席判事の一部反対意見にも表れているように、裁判官に「素人科学者」となれという事になるとの危惧を抱かせたが、今日ではそうではなく、「洗練された科学の消費者(sophisticated consumers of science)」である事が期待されているというように理解されている<sup>12)</sup>。更にその後のKumho判決<sup>13)</sup>では、一般人の証言(lay witness)の意見証言とは異なり<sup>14)</sup>、専門家証言においては個人的見解もDaubert基準において許容性を認めるとされた<sup>15)</sup>。

---

11) Daubert v. Merrell Dow Pharmaceuticals, Inc. 509 U.S. 113 S.Ct. 2786, 125 L.Ed.2d 469 (1993).

12) 前掲注7、徳永論文、832頁。

13) Kumho Tire Co. v. Carmichael, 526 U.S. 137, 152 (1999). なお現連邦訴訟規則702条「専門家証言」は、(a)「認定者の証拠の理解や事実の決定に有用」という従来の規定の他、(b)十分な事実またはデータに基づき、(c)信頼性のある原理および方法の結果であり、(d)その原理および方法を事件事実信頼性をもって適用したならば、専門家証人は、意見またはその他の形式による証言をしてよい旨定めている。

14) 一般人の証言については、知覚(perception)に基づいた合理性あるもの、あるいは有用性あるものといった制約が課されている。現連邦訴訟規則701条「素人証人による意見証言」、参照。

15) この事は、「主観確率」すなわちBayes確率を認める事になるものと思われる、母数(試行の繰り返し、事象空間)のない「一回限りの事象」についても専門家の「確信の程度」(主観確率)に証明力を与える事になる点で要検討である(後述)。飯田耕司『意思決定分析の理論：不確実性への挑戦』2006年、35-36頁、参照。但し、裁判は常に一回限りの試行であって、我が国のような自由心証主義下における裁判官の心証については、これが主観確率である事はまず疑いなく思われる。太田勝造『社会科学の理論とモデル7：法律』2000年、59-113頁、参照。

## 2. 科学的証拠に対する統計的証拠の位置

科学的証拠は米国でも我が国と同様、文字通り証拠法(evidence)の一環として講学されているが、我が国と異なり、かなり科学に立ち入った内容となっているようである。例えば、Daubert 判決でも引用されている代表的な文献の最新版である、Edward J. Imwinkelried, Paul C. Giannelli, SCIENTIFIC EVIDENCE, 5th Edition, 2012 では、総論的な法的解説の他、嘘発見器(polygraph and deception test)、心理社会科学的証拠(psychological and social science evidence)、音声鑑定(voice identification)、中性子放射化分析(NAA, neutron activation analysis)、催眠検査(hypnosis)、咬合痕鑑定(bite mark and dental identification)、線条痕等鑑定(firearms, toolmark and related examination)、統計的証拠(statistical evidence)、指紋鑑定(fingerprint)、DNA 以外の遺伝子マーカー(genetic markers other than DNA)、DNA 鑑定(DNA genetic marker)、病理鑑定(pathology)、毒物鑑定(toxicology)、筆跡等の書面鑑定(question document examination)、飲酒検査(alcohol testing)、薬物検査(drug identification)、証跡証拠(trace evidence)、機器分析(instrumental analysis)、放火証拠(arson)、事故再現(accident reconstruction)、速度計測(measurement of speed)、爆発物検査(explosive)といった実際の科学的証拠に即した技術的内容が法的側面と共に記されている。

以上のように科学的証拠の一つとして取り扱われている統計的証拠<sup>16)</sup>であるが実際に利用される事件分野にはどのようなものがあるだろうか。幾つかの当該分野の論文集などを見ると、米国では概ね以下のような分野の事件に対し統計的証拠が用いられている。

例えば、Carnegie-Mellon 大学統計学科のメンバーが編集した、Morris H. DeGroot, Stephen E. Fienberg, Joseph B. Kadane, ed., STATISTICS AND THE LAW, 1994 によると、雇用差別(employment discrimination)<sup>17)</sup>、独占禁止訴訟(antitrust litigation)、価格協定(price-fixing)、設備損害(equipment

16) 米国においては純粋な確率論的証拠も統計的証拠として扱われているので、以下区別しないで統計的証拠と呼ぶ事とする。

17) 勝田卓也「雇用差別訴訟における統計的証拠の利用について：アメリカにおける『系統的な異なる取扱い』訴訟を中心に」早稲田法学 74 巻 2 号、1999 年、27-82 頁、参照。

〈352〉 米国における統計的証拠の意義と技法 (田中)

damage)、窃盗(theft)、学校財務(school finance)、統計家証人の指名(court-appointed statistical expert)、ビデオポーカー(electric draw poker)、親子確認(disputed paternity)、問題ある選挙(contested election)、環境規制(environmental regulation)等があがっている。さらに、その編者の一人が14年後に纏めた Joseph B. Kadane ed., STATISTICS IN THE LAW, 2008 では、基本的な適用領域として、雇用差別(discrimination in employment)、黒人差別(driving while black)、民族差別(racial steering)、差別的陪審員構成(jury discrimination)をあげ、周辺の適用領域として、ビデオポーカー、試験での不正行為(cheating on examination)、税務監査(tax audit)、投票改竄(vote tampering)、性犯罪者の予防拘禁(incarceration)、特許侵害(patent misconduct)等があげられている<sup>18)</sup>。

### 3. 統計的証拠の基礎技法

そこで上記のような問題対象に適用される「技法」に特化して、米国の著名な統計的証拠の体系書を幾つか瞥見してみよう。

まず、Michael O. Finkelstein, Bruce Levin, STATISTICS FOR LAWYER, 2nd ed., 2001 を見る。本書は Springer の Statistics for Social Science and Public Policy 叢書の一冊である。

1章「記述統計(descriptive statistics)」: 代表値(central location)、散布度(dispersion)、相関(correlation)、2群の分散(variance)等。2章「裁判への適用」。3章「確率」: 基本計算、Bayesの定理(Bayes' theorem)、モンテカルロ法(Monte Carlo methods)等。4章「確率分布(probability distributions)」: 二項分布(binomial distribution)、正規分布(normal distribution)と中心極限定理(central limit theorem)、統計的仮説検定、超幾何分布(hypergeometric distribution)、正規性検定(test of normality)、Poisson分布(distribution)、幾何並びに指数分布(exponential distribution)等。5章「2群の統計的推計」: Fisherの正確性検定(exact test)、 $\chi^2$ 乗(chi-squared)並びにzスコア検定(z-score test)、信頼区間(confidence intervals)、仮説検定の検定力(power)、最尤推定(maximum likelihood

---

18) 英国の例であるが、貴族院(最高裁)の判決についての分析等もある。David Robertson, JUDICIAL DISCRETION IN THE HOUSE OF LORDS, 1998, pp.34-71.

estimation)。6章「3群以上の比較」：適合性検定(goodness of fit test)、Bonferroniの不等式、 $\phi^2$ 乗(phi-squared)と $\tau B$ (tau B)等。7章「比較技法」：Studentのt検定と信頼区間等。8章「層別分析」：Mantel-HaenzelとFisherの方法等。9章「標本抽出(sampling)問題」：無作為抽出(random sampling)等。10章「疫学(epidemiology)」。11章「生存分析(survival analysis)」。12章「ノンパラメトリック技法(nonparametric method)」。13章「回帰分析(regression model)」：重回帰(multiple regression)、回帰式(regression equation)、回帰係数(regression coefficient)、信頼区間と予測区間(confidence and prediction interval)等。14章「より複雑な回帰分析」。

この書においては、法律的な意味での証拠をマテリアルとして扱うものの、それを扱う統計的技法としては上記のように主な統計技法のほとんどが扱われており、法律家がすべてを理解する必要はないように思われ、本書は時宜に応じて利用されるリファレンスマニュアルというべき業績である<sup>19)</sup>。

なお、上記のSTATISTICS FOR LAWYERのサブセット版ともいえる、Finkelstein単独の執筆となるBASIC CONCEPTS OF PROBABILITY AND STATISTICS IN THE LAW, 2009では下記のような項目があがっており、基礎技法としてはこの辺りの範囲が標準的に思える。

1章「確率」：Bayesの定理を含む。2章「記述統計」：代表値、散布度、標準偏差(standard deviation, SD)、相関係数(correlation coefficient)。3章「複合事象(compound event)」。4章「有意性(significance)」。5章「確率変数(random variables)とその分布」：期待値、分散、相関係数、二項分布、超幾何分布、正規分布、Poisson分布、Studentのt分布、幾何分布と指数分布(exponential distribution)等。6章「信頼区間」。7章「検出力」。8章「サンプリング」：無作為抽出、偏向(bias)問題等。9章「疫学」：コホート研究(cohort studies)、対照研究(case-control studies)、偏向と交絡(confounding)、関連と因果。10章「証拠の結合」。11章、12章「回帰分析」。

次に、米国のLaw School教科書水準のものとして、West GroupのAmerican Casebook Series®の一冊である、David L. Faigman, David H.

19) 飯田高「分析ツールの説明書」アメリカ法2003-2、417-422頁。

Kaye, Michael J. Saks, Joseph Sanders, SCIENCE IN THE LAW: STANDARDS, STATISTICS AND RESEARCH ISSUES, 2002 を、あげられている例も含め少し詳しく見ていく。統計的証拠については第5章、調査方法については第6章、疫学（epidemiology）については第8章で扱われているが<sup>20)</sup>、ここでは第5章の一般的統計技法について取り扱う。（5章1節は法的議論。）

### 5章2節「データ収集」

調査計画 (design of a study)、証言証拠 (anecdotal evidence)、観察調査 (observational studies)、対照実験 (controlled experiments)、無作為抽出 (random sampling)、[例] 死刑の効果等、関連 (association) と因果 (causation)、[例] 死刑肯定派の陪審員と死刑評決、全数調査 (census)、母集団 (population)、標本 (sample)、選択偏向 (selection bias)、確率的抽出法 (probability sampling method)、[例] 陪審員の選出、計測の正確性 (accuracy of measurements)、信頼性 (reliability)、妥当性 (validity)。

### 5章3節「データ表現」

要約統計量 (summaries)、割合やパーセンテージ (rates or percentage)、記述統計 (descriptive statistics)、平均 (mean)、中央値 (median)、標準偏差、時系列データ (time series data)、第1種の誤謬 (type I error)、偽陽性 (false positive)、第2種の誤謬 (type II error)、偽陰性 (false negative)、[例] 煙草の好みのアンケート広告、妊娠検査、頻度 (actual numbers)、母数 (base)、[例] 老年犯罪率の変化、表 (table)、二元配置表 (cross-tab, contingency table)、[例] 性別入試合格率<sup>21)</sup>、差別的陪審構成、オッズ比 (odds ratio)、グラフ表現 (graph portray data)、傾向 (trends)、軸の尺度問題 (scales on the axes matter)、[例] Reagan-Bush 政権下の連邦の負債、度数分布図 (histogram)、区間 (bins, class interval)、回分 (batch)、幹葉図 (stem-and-leaf plot)、代表値 (center of a distribution)、相加平均 (mean, arithmetic mean)、中央値 (median)、最頻値 (mode)、外れ値 (outlier)、[例] 1975-1985 の平均過誤 (malpractice) 賠償額、変動 (variability)、範囲 (range)、四分位範囲 (interquartile range)、標準偏差、変動係数 (coefficient of variation)、少

20) 本稿では疫学については触れない。詳しくは、門口正人他編『民事訴訟法大系第2巻総論II』2004年、35-43頁、参照。

21) 前掲注19、飯田論文、417-419頁、参照。



数データ (small data set)、箱髭図 (box-and-whisker plot)。

#### 5章4節「データからの推論」

偶然誤差 (random error, sampling error, chance error, statistical error)、推定 (estimation)、標準誤差<sup>22)</sup> (standard error, SE)、信頼区間 (confidence intervals)、信頼水準 (confidence level)、[例] 信頼区間の解釈、有意確率 (significance probabilities)、仮説検定 (hypothesis tests, significance testing)、有意水準 (significance level)、帰無仮説 (null hypothesis)、対立仮説 (alternative hypothesis)、p 値 (p-value)、[例] p 値の解釈、検出率 (power)、事後確率 (posterior probability)、[例] 就職性差別、差別的効果 (disparate impact)、主観主義 / Bayes 基準 (subjectivist / Bayes' rule)、客観主義 / 頻度主義 (objectivist / frequentist)、系統誤差 (systematic bias)、[例] ドラッグと出産異常、標本量 (sample size)、[例] 黒人差別、片側あるいは両側検定 (one- or two-tailed test)、区間推定 (interval estimate)、[例] 従業員の横領、事後確率 (posterior probability)、事前確率 (prior probability)、操作特性 (operating characteristics)、逆確率 (converse or inverse probability)。

#### 5章5節「相関と回帰」

回帰分析、散布図 (scatter diagram, scatterplot, scattergram)、相関係数 (r)、[例] 学歴と収入、重相関係数 (multiple correlation coefficient)、[例] 学部成績及び LSAT 成績と law school 初年度成績、交絡変数あるいは第3変数 (third variable)、独立変数 (independent variable, 説明変数)、従属変数 (dependent variable, 目的変数)、決定係数 (proportion of variance explained, 寄与率)、回帰直線 (regression line)、傾き (slope)、切片 (intercept)、標準化 (standardized)、集約データ相関 (ecological correlation)、[例] 人種的少数者の選挙権、集約化 (compactness)、投票行動回帰 (ecological regression)、統計モデル (statistical model)、[例] Hooke の法則、最小二乗法 (method of least squares)、予測誤差 (prediction error)、残差 (residual)、ガウス分布 (Gauss' model)、[例] 給与の性差別、条件付き確率 (conditional probability)。

5章「補遺」下記の事項について、簡単な数学的解説を行っている。

主観確率 (信念の程度)、客観確率 (頻度)、独立事象 (independent

22) 原著は、裁判所は標準誤差と標準偏差を同じものとしているとする。See, p.184, note6. cf., p.190, note31.

event)、乗法定理(multiplication or product rule)、従属事象(dependent event)、条件付き確率、 $H_0$ (H-sub-zero)、 $H_1$ (H-sub-one)、Bayes 統計、[例] 犯行現場の血液型、期待値(expected value)、標準偏差、尤度(likelihood)、標準誤差、正規分布、信頼水準、p 値、有意水準、中心極限定理。

以上のように本書は、法律家が Daubert 基準に沿った「洗練された科学の消費者」となるに必要な事柄について、数式をほとんど用いず、また個々の検定技法等に深入りする事なく、豊富な判例や事例を引用しつつ統計的証拠解釈の留意点について解説する好著である。

#### 4. 統計的証拠の機能

しかし、統計を証拠とする事には、根強い反論もある。ここでは、著名な「不正入場者のパラドックス(Paradox of the Gatecrasher)」<sup>23)</sup>を例にあげよう。

設例は、以下のようなものである。

ロデオを見るのに、499 人が切符を買って入った。ところが、1000 人が座席に座っていた。誰が不正入場者かは特定できない。1000 人中 501 人が不正入場者なのだから、この 1000 人中のある 1 人が不正入場者である確率は 0.501 であり、それは 1000 人全員にいえる。それ故(50% 超原則により)、興行主は 1000 人全員に入場料を払わせる事ができる。つまり、最初に正規に切符を買った人 499 人と合わせて 1499 人から料金を取れる事になる<sup>24)</sup>。これはおかしいという訳である。

---

23) Laurence Jonathan Cohen, THE PROBABLE AND THE PROVABLE, 1977, pp.77-81. 田尾桃二、加藤新太郎共編『民事事実認定』1999 年、209 頁、参照。

24) I. Ayres は、まず平均と SD を想定して、95%(2SD)ルールを適用する事を概算法として提唱している。(但し、多くの分布は歪んでおり、例えば非正規乱数では 2SD の範囲に落ちるのは 75% だとしている。) Ian Ayres, SUPER CRUNCHERS: WHY THINKING-BY-NUMBERS IS THE NEW WAY TO BE SMART, 2007. (山形浩生訳『その数学が戦略を決める』2007 年、326-348 頁、446-447 頁。) Fisher が有意水準(危険率)を一応 5% においたのも、± 2SD から離れるのは 4.56%、これを丸めて 5% としたという事であるので(増山元三郎『でたらめの世界』1969 年、50 頁。)、一定の経験的信頼度はあるといえよう。この概算法を Cohen の設例に適用してみよう。不正入場者数の発生率を正規分布で近似する。入場者 100 人の場合、不正入場者は 5% で 5 名となる。不正入場 0 名未満はないのでそこまでを 2SD とすると、SD は概ね  $5/2=2.5$  となり、 $5+2SD$  は 10 名で、これを越えるのは 5% 以下という事になる。実際計算してみると不正入場者 9 名では 5.5% で、10 名をこえて初めて 5% を切り 2.28% と

この設例は、統計的証拠が論理的判断と矛盾する事を示している<sup>25)</sup>。しかし逆にいえば、事実認定は論理だけでは決まらないという事も示している<sup>26)</sup>。すなわち統計的証拠の証明力は事実認定にとって決定的な重要性を持つのである。

上述したように統計的証拠は他の科学的証拠と平行に扱われているとはいえ、実はすべての証拠が結局は統計的な証明力で量られる点で、統計的証拠への理解は他の科学的証拠に先立って重要なのである。

例えば目撃証言といった非科学的証拠ですら結局は統計的証拠である事を、プロスペクト理論で著名な D. Kahneman と A. Tversky の設例<sup>27)</sup>によって示す。若干簡単にするが、設例は以下のようなものである。

ある夜、ひき逃げ事件があった。今この街には青の車と緑の車しか走っていないとする。目撃者が青の車だったと証言したので、その目撃者の夜半における青の車と緑の車の識別率を鑑定すると 80% であった。

これだけでは、認知バイアス(cognitive bias)により、犯行車が青である確率は 80% と誤解する人が多いだろう。しかしここで、街を走る車

なるので、概算法としては有用だと思われる。なお、不正入場者が 17 名以上である事態が起るのは 100 万回に 1 回であり、不正入場平均 5 名としたときの Cohen の想定には現実性はない。逆に Cohen の想定が正規分布の近似で現実的なものである為には、不正入場者平均 25 名以上を想定する必要があるが、入場者の 1/4 が不正入場というのは経験的に首肯できない。(とはいえポアソン分布で見ると、平均 5 人のとき不正入場者がいないという事態が起こるのは 0.674% なので、人数はともかく、ほぼ 99.3% の確率で不正入場者は存在する事となる。)

25) David Salsburg, THE LADY TASTING TEA: HOW STATISTICS REVOLUTIONIZED SCIENCE IN THE TWENTIETH CENTURY 2001. (竹内恵行、熊谷悦生訳『統計学を拓いた異才たち』2010年、429-432頁)。なお、Cohen の問題提起に対する Glanville Williams の反論については、庭山英雄、田中嘉之「リチャード・エグレストン『証拠・証明および確率』(1978年)」中京法学 16 卷 3 号、1982 年、15-18 頁、参照。

26) こうした問題の解決の為には、例えば、論理法学の立場(吉野一『法律人工知能:法的知識の解明と法的推論の実現』2000年、16-33頁、224-232頁、参照。)の中に統計的証拠の理論が取り込まれる事が期待される。なお、夏井高人『裁判実務とコンピュータ』1993年、292-307頁、参照。

27) Daniel Kahneman, Amos Tversky, Subjective probability: A judgment under representativeness, JUDGEMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES, 1982, pp.32-47. See also, A. Tversky, D. Kahneman, Judgment under Uncertainty: Heuristics and Biases, SCIENCE, New Series, Vol.185, No.4157, Sep.27, 1974, pp.1124-1131. And see, Deborah J. Bennett, RANDOMNESS, 1998, pp.2-3. (江原摩美訳『確率とデタラメの世界:偶然の数学はどのように進化したか』2001年。)

〈358〉 米国における統計的証拠の意義と技法（田中）

が100台しかないとして、その内青の車が15台、緑の車が85台だとすると、目撃者が100台全部見たとき、青の車を見たという確率は以下のようになる。

青の車を青と証言する台数(証言は正しい)： $0.8 \times 15=12$ 台

緑の車を青と証言する台数(証言は誤り)： $(1-0.8) \times 85=17$ 台<sup>28)</sup>

つまり、目撃者は100台中、 $(12+17)=29$ 台の車を青の車と証言し、内12台について正しく判断している訳であるから、目撃証言の証明力は、 $12/29=41.4\%$ に過ぎないという結論になる。付言すると、この場合緑の車が犯行車である確率が $17/29=58.6\%$ であり、実は緑の車が犯行車だという確率の方が高い事にも留意すべきであろう。

## 5. 統計的証拠解釈における誤謬

Daubert判決は、許容性基準を示した上に、事実審における指針として、「検証可能性」、「同僚審査」、「誤謬率、標準」、「一般的承認」の4項目をあげていることから<sup>29)</sup>、「誤謬率、標準」が重要なポイントの一つである事は示されているとあってよいだろう。

統計的証拠の解釈について、上述した体系書類や専門家の間で問題視されるトピックは多岐にわたるが、上記の理由から、以下、統計論の立場から証拠の証明力や冤罪率を決定するという意味で最も重要な誤謬率(error rate)を主題とする。

### 5-1. 処理の誤謬

どんな証拠であろうが、資料の採取保存状況や計測機器の状態、並びに人間による誤認識や取り扱いのミス<sup>30)</sup>、場合によっては証拠の捏造等、理想的な状態での鑑定精度に比べて、実際の鑑定結果の精度が落ちる

---

28) 実は緑の車を緑と認定する率が同時に0.8となる訳でなく、青と誤認する率は必ずしも $(1-0.8)$ ではないのだが、そうした問題点については後述。なお、実際には「わからない(判別不能)」の割合もありうるだろう。

29) Daubert, at 594.

30) 徳永光「刑事裁判における科学的証拠の利用：DNA鑑定に関する日本の状況をアメリカにおける議論と比較して」一橋研究25巻2号、2002年、15頁、27頁注71。

というのは容易に想定できるであろう。

ある事件で<sup>31)</sup>、FBIがA潜在指紋とB潜在指紋が容疑者甲の指紋と一致するかどうかについて、35の州の鑑定機関から指紋鑑定結果を得た。AB共に一致するとした機関が21、Aには一致するがBには一致しないとした機関が8、Aには一致しないがBには一致するとした機関が6あった。

この場合、Aに一致が $(21+8)/35=0.829$ 、Bに一致が $(21+6)/35=0.771$ となる。理想状態での指紋鑑定の一般的な本人一致率(真陽性率)がほぼ1とされる事から見れば、真実が一致不一致の何れであるとしても、理想状態での鑑定誤差以外の処理の誤謬による誤差が如何に大きいかかわかるだろう。

それ故、例えばDNA鑑定についても、できれば捜査側からだけでなく弁護側からも、複数の異なる機関による鑑定を突き合わせて証拠とする事が望ましいように思われる<sup>32)</sup>。

## 5-2. 解釈の誤謬

以下、Kahneman等のいう認知バイアスに起因する誤謬問題を幾つか取りあげる。

### 5-2-1. 冤罪率の誤謬

鑑定が理想的に行われたとしても、解釈上の誤謬は発生する。その内最も問題となるのは冤罪率であろう。

冤罪率を考える前提として、まず「偽陽性問題」と「偽陰性問題」を理解する必要がある。簡単にいえば、偽陽性問題は非罹患者に検査結果陽性を与えてしまう事であり、偽陰性問題とは罹患者に検査結果陰性を与えてしまう事である。これは今日の確率統計学がいうところの、検定に対する「第1種の誤謬」と「第2種の誤謬」の観念を拡張したものと捉えてよいだろう。通常裁判では、被告(人)が有責(有罪)である事を立

31) United States v. Mitchell, 145 F.3d 572, 576-79 (3d Cir. 1998).

32) 前掲注30、徳永論文、2頁。こうした事情から、本稿で例とする鑑定精度は、説明の為の数値という事もあり、理論値よりかなり低く見積もっている。

証しようとするので、以下簡単な為、偽陽性(第1種の誤謬)を「無実なのに有罪としてしまう」つまり冤罪の事とし、偽陰性(第2種の誤謬)を「犯人なのに無罪としてしまう」事とする<sup>33)</sup>。(以下、民事でも刑事になぞらえてこのような表現をする。) 検定でいえば、偽陽性率とは「p 値」の事であり、1- 偽陽性率、つまり無実の人を無罪とする割合(真陰性率)とは「検出力」の事である。(なお、1- 偽陰性率が真陽性率である。)

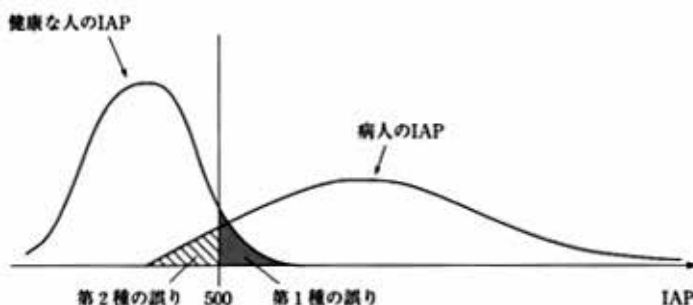


図1 IAPによる診断の誤り<sup>34)</sup>

IAP(免疫抑制酸性蛋白)検査を例にすると、検査値が500以上でも偽陽性は存在するし、500未満でも偽陰性は存在する事がわかるだろう(図1)。更に、偽陽性率と偽陰性率は一般に全く独立である事もよくわかる。

33) 医学や公衆衛生学分野では、陽性を真陽性と判断する精度の事を「感度(sensitivity)」、陰性を真陰性と判断できる精度の事を「特異度(specificity)」という等、分野によって様々な呼び方がされる事には留意。判例には、偽陽性を積極的誤判定(false identification)、偽陰性を消極的誤判定(false elimination)と呼ぶものもある。Cf., *United States v. Smith*, 869F. 2d348 (7th Cir. 1989). なお、医療検査では、真陽性率-偽陽性率が最大になるように閾値(threshold level)を決定するという方法がとられる事があるが、閾値によって真陽性率、偽陽性率が変わるので、これをどのように決定するかが問題となる(図1参照)。しかし、裁判においては真陽性よりも偽陽性(冤罪)の減少が重視されるので、医療検査と同じような基準で閾値の設定をする事はできない。因みにこうした感度・特異度のトレードオフの関係はROC(Receiver Operating Characteristic)曲線(受信者操作特性曲線、感度・特異度曲線)で表される。

34) 宮川公男『統計学でリスクと向き合う [新版]』2007年、211頁。

先述の Kahneman と Tversky の例を取りあげ、冤罪率について考えてみよう。

元の設例に加えて目撃者の識別力が以下の通りと仮定する。この二つが独立である事は重要である。

青の車を見た場合：青とする (0.8)、緑とする (0.2)。

緑の車を見た場合：緑とする (0.6)、青とする (0.4)。

証言の分布は以下ようになる。(図 1 参照、500 以上が青に見えた事象とすればわかり易い。)

以下、「青の車を見た」との目撃証言に沿って、青の車の冤罪率を考える。

真陽性数 (True Positive, 青に見えた | 青の車) :  $0.8 \times 15 = 12$  台

偽陽性数 (False Positive, 青に見えた | 緑の車) :  $0.4 \times 85 = 34$  台

真陰性数 (True Negative, 緑に見えた | 緑の車) :  $0.6 \times 85 = 51$  台

偽陰性数 (False Negative, 緑に見えた | 青の車) :  $0.2 \times 15 = 3$  台

青の車にとっての冤罪率(証言青だが、犯行車は緑) :  $34 / (12 + 34) = 0.74$

以上のように、「対象となる台数」といった事象空間(母集団)を前提とせずに統計的証拠を用いるのは誤解を招く原因となるという事は夙に指摘されている<sup>35)</sup>。この事は本節において述べる、認知バイアスに起因する解釈の誤謬に共通する問題である。

更に問題となるのは鑑定精度がいわれるとき、偽陽性率しか公表されない事である。例えば、指紋鑑定の精度は 1/10 万等としかいわれない。これは偽陽性率の事であって、真陽性率は明示されないのである。おそらく暗黙裡に真陽性率は 1、すなわち、本人が一致するのは間違いないとされているのであろう。しかし、もしそれが本当なら、容疑者が本人の場合、先述の FBI の指紋鑑定結果は全 35 機関すべて一致したであろう。今日、DNA 鑑定の偽陽性率(出現頻度<sup>36)</sup>)は 1/4 兆 7000 億といわれ

35) See, Stephen E. Fienberg ed., THE EVOLVING ROLE OF STATISTICAL ASSESSMENTS AS EVIDENCE IN THE COURTS, National Research Council 1988, published 1989, p.282. 小林秀之『新証拠法 [第 2 版]』2003 年、75 頁。

36) 我が国の裁判所が使用する語法。出現頻度 10 万分の 1 というのは、100 万人でいえば 10 人は同じ指紋と鑑定されるという事である。その中に本人がいるかどうかは無関係である。統計学では偶然一致率(random match rate)といっているので、本人を除くという誤解を招く可能性があるが、出現頻度と同じ意味で

ているが、上述の如く、実際の真陽性率ほどの程度なのだろうか。

FBIの指紋鑑定为例からも、それを1と仮定するのは現実的ではないだろう。処理の誤謬も含めて考えれば、指紋鑑定であろうとDNA鑑定であろうと、その偽陽性率に関わりなく、真陽性率はある程度1を下回っているだろう。それ故、鑑定精度については、偽陽性率だけでなく処理の誤謬も含めた実際的な真陽性率が幾ら位なのかという事も公表すべきである。少なくとも、本人のものとして扱われた指紋やDNAが実際の程度の割合で本当に本人に一致していたのか程度の実験調査は必要であろう。

Kahneman等の例に戻ろう。真陽性率を1とすると、上記の例は次のようになる。(以下、真陰性率と偽陰性率は不要なので省略する。)

真陽性数(青に見えた | 青の車) :  $1 \times 15 = 15$  台

偽陽性数(青に見えた | 緑の車) :  $0.4 \times 85 = 34$  台

青の車にとっての冤罪率(証言青だが、犯行車は緑) :  $34 / (15 + 34) = 0.69$

-----  
以上のように、通常、真陽性率は1より低いのだが、真陽性率を暗黙裡に1としてしまうと、冤罪率が低く見えてしまうのである。

今、犯人が1人、それ以外の容疑者(以下、他人とも呼ぶ。)がn人いたとする。つまり母集団を想定する訳である。

この母集団に対して何らかの鑑定をしたとする。識別力と実数(母数)は以下の通りである。他人に対する冤罪率を考える。

真陽性率(一致 | 犯人) : a、実数 : a

偽陽性率(一致 | 他人) : b、実数 : nb

他人にとっての冤罪率(鑑定と一致したが、犯人ではない) :  $(nb) / (a + nb)$

-----  
容疑者n人中最大でも冤罪者を1人までしか出さないとすると、冤罪率  $= nb / (a + nb) < 1/n$ 。故に、 $b/a < 1 / (n(n-1))$ 。つまり容疑者数が大きい程、(偽陽性率 / 真陽性率)が小さくなくてはならないことがわかる。大抵の鑑定でaは暗黙裡にはほぼ1とされているようであるので1としたとして、例えば指紋鑑定レベルで  $b = 1/10$ 万とすると、容疑者が316人をこえると、

---

ある。但し本稿では、簡単な為、本人は出現頻度を含めていない。



冤罪率  $<1/n$  にはならない。DNA 鑑定レベルで  $b=1/100$  万としても、容疑者が 1000 人をこえると、冤罪率  $<1/n$  にはならない。また逆に冤罪率を固定して、冤罪率  $<1/1$  万にするとした場合、 $nb/(a+nb) < 1/1$  万、仮に  $a=1$  としても、 $b < 1/(9999n)$  となり、指紋鑑定レベルの  $1/10$  万の偽陽性率では、容疑者  $n < 10$  人程度、DNA 鑑定レベルの  $1/100$  万の偽陽性率では、 $n < 100$  人程度の容疑者しか扱えない事となる。即ち、他の証拠等でこの程度まで容疑者が絞られた上でないと、鑑定結果が予定しない冤罪を生み出す可能性があるということである。

### 5-2-2. 無差別原理の誤謬

若干余談となるが、かの Laplace (1749-1827) は、冤罪率  $1/1000$  の仮定を立て、まず陪審員の誤判確率を最大  $0.5$  とした。そこで被告人が犯人である確率の初期設定を  $0.5$  とし、最大冤罪率を  $1/1000$  とする為には、陪審員が何人必要かを計算した。陪審員を  $n$  人とすると、 $1/1000 > 0.5 (0.5)^n = 0.5^{(n+1)}$ 、 $n=9$  のとき成立するので、陪審員は 9 名、それも有罪判決は全員一致が必要とした<sup>37)</sup>。計算結果はともかく、問題は被告人であるというだけで何の証拠も見ずに、犯人である初期確率を  $0.5$  とする無差別原理 (principle of indifference) の適用である。「無罪の推定」を統計的に解釈すれば、何の証拠も出されていない状態では、母集団を考え、我が国の場合なら少なくとも日本人の 1 人程度 (約 1 億分の 1) で犯人と考えるべきであろう。1 回限りの試行の確率を無差別原理で決めてしまう事は、サイコロを振って 1 の目が出る確率を、事象を「1 が出る」、「1 が出ない」に分け、無差別原理を適用して出るか出ないかは  $1/2$  だとする事につながる。あくまで正 6 面体という事象空間に基づき、主観確率を  $1/6$  と設定するのが通常であり、例えば 1 の目の彫が深い等の証拠から 1 が出るのは  $1/6$  より大きくなりそうだといった推論をすべきであろう。もし被告人が犯人である初期確率を  $0.5$  とするなら、筆者にはその

37) Frederick W. Truscott, Frederick L. Emory, trans., A PHILOSOPHICAL ESSAY ON PROBABILITIES. by Pierre S. Laplace from 6th french ed., pp.132-139. 天野徹『社会統計学へのアプローチ：思想と方法』2006 年、4-5 頁。なお、我が国の場合、裁判官 3 名と裁判員 6 名、計 9 名で審理するのは、この Laplace の計算と関係するのかどうかは不案内だが、米国のように陪審員が 12 名いるときには、全員一致なら Laplace の計算では誤判率は 1 万分の 1.22 以下となる。

根拠は、「裁判にかけられている事自体が証拠で、その証明力が0.5ある。」あるいは、「裁判は有罪か無罪かの2事象を決定するものだから0.5である。」といった理由しか考えられない。あくまで母集団(サイコロなら6、日本人なら1億等)を前提として考えるべきである。

再び Kahneman 等の例で考えよう。Kahneman 等のいうように、青の車を見た目の目撃者の、青の車に対する識別力が0.8と鑑定された場合、青の車が犯行車である確率が80%だとする予断を持つ者は多いであろう。これは、証言がない段階での心証(事前確率)が、無差別原理により、青も緑も犯行車である確率はタイで1/2だとしているからである。無意識の無差別原理の適用により、青の車が犯行車であるとの心証は、 $(0.5 \times 0.8) / ((0.5 \times 0.8) + 0.5(1-0.8)) = 0.8$ となる訳である。この場合、無差別原理の適用という予断(認知バイアス)を持っていたが為に、真実の目撃証言の証明力0.414に対して誤った心証0.8を持ってしまったのである。

本当に青の車が犯行車である証明力 $x$ が80%である為には、Kahneman 等の例で青の車15台、緑の車85台のとき、目撃者の識別力を $w$ とすると、 $x=0.8=(15w)/(15w+85(1-w))$ 、 $w=0.958$ でなければならないのである。

当然の事ではあるが、以上のように母集団や証拠の解釈に誤った予断を持つことは、判断を誤らせる。

### 5-2-3. 訴追者の誤謬

更に、誤謬問題でよく指摘されるのが、「訴追者の誤謬(prosecutor's fallacy)」と「弁護者の誤謬(defendant's fallacy)」であり<sup>38)</sup>、これらも Kahneman 等のいう、認知バイアスの一種である。

「訴追者の誤謬」から見よう。この問題に関してよく引き合いに出されるのが、所謂 Collins 事件である<sup>39)</sup>。

---

38) Hans Zeisel, David H. Kaye, PROVE IT WITH FIGURES: EMPIRICAL METHODS IN LAW AND LITIGATION 1997. (細谷雄三訳『数字で立証する：裁判と統計』2012年、244-245頁。)伊勢田哲治「DNA鑑定の倫理的考察」『ヒトゲノム計画に関する倫理的問題』(京都大学大学院文学研究科倫理学研究室)、1998年(<http://www.ethics.bun.kyoto-u.ac.jp/genome/genome95/43iseda.html>)、William C. Thompson, Edward L. Schumann, Interpretation of Statistical Evidence Criminal Trials, LAW AND HUMAN BEHAVIOR, Vol. 11, No. 3, 1987, pp.167-187, 参照。

39) People v. Collins 68 Cal.2d 319, 66 Cal.Rptr. 497, 438 P.2d (1968).

検察側は下記の証拠を提出し(表 1)、これらすべてが同時に成り立つ確率は 1200 万分の 1 だとした。

表 1 Collins 事件における検察側の証拠

証拠	個別確率
金髪の若い女性	1/3
ポニーテールの若い女性	1/10
一部黄色の車	1/10
口髭の男性	1/4
あご髭の黒人男性	1/10
同じ車に乗っている異人種カップル	1/1000
上記全てが同時に成り立つ確率	1/1200 万

この偽陽性率 1/1200 万をもって、一致者が犯人でない確率(無実率)を 1/1200 万とするのが「訴追者の誤謬」である。しかし今、母数を 1200 万とすると、特徴に合致する対象は 1 である。今、容疑を受けている一致者(被告人)が母数に含まれてないとすれば、犯人でない確率は 1/2 となる<sup>40)</sup>。

この事は、以下のように示される。

真陽性率(一致 | 犯人) : 1、実数 : 1

偽陽性率(一致 | 他人) : 1/1200 万、実数 : 1200 万 (1/1200 万) = 1

真陰性率(不一致 | 他人) : 1 - (1/1200 万)、実数 : 1200 万 - 1

偽陰性率(不一致 | 犯人) : 0

犯人 / 一致者 = 1 / (1 + 1)

1200 万といわずとも、一致者の偽陽性率に比べ、一致者の無実率 = 偽陽性数 / (真陽性数 + 偽陽性数) は通常かなり大きい。

例えば、大統領は 1 名であるが(真陽性数 1)、大統領に人相が一致す

40) なお、本件では、証拠が大雑把な推計に過ぎない事はもとより、各証拠の独立性も大きな問題となった。何故なら確率の乗法定理が成り立つ為には事象がそれぞれ独立していなければならないからである。例えば、「口髭の男性」と「あご髭の黒人男性」が独立事象かどうかは極めて疑わしい。前掲注 35、小林書、11-14 頁、77-78 頁。

る他人の存在確率（偽陽性率）が 1/100 万としよう。訴追者の誤謬は、人相が一致した人が大統領でない確率は 1/100 万であるとする。実は対象となる人口が 1 億とすれば、人相が一致する他人（1 億（1/100 万）=100 人）が、大統領でない確率は  $1 - (1/(100+1)) = 99\%$  である。大統領であるのは、 $\text{真陽性数} / (\text{真陽性数} + \text{偽陽性数}) = 1/(1+100) = 1\%$  となる。

Collins 事件と同様の例として、我が国の冤罪事件として有名な所謂「足利事件」を取りあげよう。第一審は、当時の MCT118DNA 型鑑定に基づき、以下の証拠を認定して、「以上の血液型及び DNA 型を持った者の日本人における出現頻度は、1000 人中 1.2 人程度であると算出された。」<sup>41)</sup> とした(表 2)。

表 2 足利事件における検察側の証拠

証拠	個別確率
DNA16-26 型の出現頻度	0.83/100
血液型 B 型の出現頻度	22.1/100
Le(a-b+)型：分泌型	67.8/100
上記全てが同時に成り立つ確率	1.2437/1000

首都圏の人口を 3500 万とすれば、上記の特性に合致する人はなんと 43530 人もいる。勿論この科学的証拠だけで有罪とした訳ではないだろうが(後述)、この証拠だけで見る限り、この容疑者が犯人である確率は  $1/(43530+1) = 0.0023\%$  である。それでも、犯人であれば 100% 一致する（真陽性率=1）という前提である。

先述した FBI の指紋鑑定の精度のように、保存状態、人的ミス等の処理の誤謬により真陽性率が 0.8 程度になっているとすればどうなるのであろうか。今、容疑者の母集団を 1000 名と現実的な数字にする。

真陽性率（一致 | 犯人）：0.8、実数：0.8

偽陽性率（一致 | 他人）：1.2/1000、実数：1000(1.2/1000)=1.2

真陰性率（不一致 | 他人）：1-(1.2/1000)、実数：1000(998.8/1000)=998.8

偽陰性率（不一致 | 犯人）：0.2、実数：0.2

---

41) 宇都宮地判平 5.7.7 刑集 54 卷 6 号 670 頁。再審請求の抗告審も鑑定結果は認めている。東京高決平 21.6.23 東高刑時報 60 卷 91 頁。

訴追者の誤謬は、偽陽性率が  $1.2/1000$  なので、一致者が無実である確率も  $1.2/1000$  とするが、実際には一致者が無実である確率は、 $1.2/(0.8+1.2)=0.6$  となる。またもし、証明力 = 犯人 / (犯人 + 他人) なら、例えば証明力が  $0.8$  より大きくなる為には、検査の真陽性率を  $a$ 、容疑者が  $b$  人とする、証明力  $= a/(a+b(1.2/1000)) > 0.8$ 、故に  $a > 0.0048b$ 、 $a \leq 1$  なので、この検査では、 $b$  すなわち容疑者が  $208$  人をこえると、この証拠が証明力  $80\%$  をこえる事ができないという事になる。つまり出現頻度が  $1000$  人に  $1.2$  人の証拠であっても、証明力が  $0.8$  以上でありうる射程は、容疑者  $b=208$  人までなのである。真陽性率  $a=0.9$  となれば、容疑者  $b$  は  $187$  人以下でなければならない。逆にいえば、冤罪率の誤謬と同様、他の証拠等でそこまで絞られている事が要求される訳である。(この事は次の「弁護者の誤謬」に関連する。)

以上のように、母集団を考える事は統計的証拠解釈にとって決定的に重要なのである。

#### 5-2-4. 弁護者の誤謬

次に「弁護者の誤謬」について見よう。

上述足利事件の例でいえば、母集団を首都圏の住人全員  $3500$  万人とすると、被告人が犯人である確率は  $1/3500$  万に過ぎず、それは証拠にならないとするのが「弁護者の誤謬」である。

しかし、実際は他の証拠等により、もっと少数に絞りこまれているのが通常である。例えばこれを  $100$  人と考えると、一致者が犯人である確率は、 $1/(1+0.12)=0.893$  となるのが真実である<sup>42)</sup>。

足利事件では、他の証拠による容疑者がどれだけいたのかわからないので、ここでは著名な例である O. J. Simpson 事件<sup>43)</sup>を取りあげよう。

この事件では DNA 鑑定を含み、様々な統計的証拠からの争点がある

42) See, Tim Thompson, Sue Black, ed., FORENSIC HUMAN IDENTIFICATION: AN INTRODUCTION 2007, p.19. And also, Norman Fenton, Martin Neil, The "Jury Observation Fallacy" and the use of Bayesian Networks to present Probabilistic Legal Arguments, BULLETIN OF THE IMA 2000, 36-6, p.2.

43) The People of the State of California v. Orenthal James Simpson (1995).

が<sup>44)</sup>、ここでは、DV(domestic violence)問題を取りあげる<sup>45)</sup>。

著名なアメリカンフットボールプレイヤーである O. J. Simpson が元妻を殺したとして起訴され、最終的に無罪となった。検察側は O. J. Simpson がその元妻に DV をふるっていた事を証拠に「平手打ちは殺人の前奏曲」だとした。それに対して弁護団は、著名な Harvard Law School 教授 Alan Dershowitz のアイデアに基づき、「1992 年には約 400 万の DV があるが、その内殺害までに至るのは 1432 件なので、パートナーに DV をふるう男性の内、殺害までに至るのは概ね 1/2500 未満となる。」と反論した。

ここに「弁護者の誤謬」がある。（この場合、むしろ「弁護者の詭弁」といった方が正鵠を射ているかもしれない。）つまり、被告人が犯人である確率は 1/2500 であり、これは被告人を犯人とする証拠にはならないという訳である。しかし、実は DV を受けている女性を殺した犯人である可能性のある容疑者は 400 万人でなく、もっと絞られている。DV を受けている女性を殺した犯人はパートナーかそれ以外しかないのである。

この事件では、DV を受けていて、そのパートナー以外から殺害されている女性のデータがない為、以下のような推論を試みる。

1993 年の統計では一般的な米国女性は年間 10 万人に 5 人が殺害されているそうである。米国女性人口を概ね 1 億とすると、年間 5000 人が殺害されている事となる。その内、DV を受けた挙句パートナーに殺害されるのが、1432 件なので、殺害された女性の内、DV を受けたパートナーから殺害されている者の割合は、 $1432/5000 = \text{約 } 28.6\%$  である。これは弁護団のいう割合より約 700 倍も高い。

---

44) 例えば本件の DNA 鑑定問題については、Jeffrey S. Rosenthal, STRUCK BY LIGHTNING: THE CURIOUS WORLD OF PROBABILITIES, 2005. (中村義作監修、柴田裕之訳『運は数学にまかせなさい: 確率・統計に学ぶ処世術』2010 年、216-220 頁。) Jonathan J. Koehler, One in Millions, Billions and Trillions: Lessons from People v. Collins (1968) for People v. Simpson (1995), JOURNAL OF LEGAL EDUCATION, Vol 47, 1997, pp.214-223, 関西医科大学法医学講座「O. J. シンプソン事件の DNA 鑑定」(<http://www3.kmu.ac.jp/legalmed/DNA/simpson.html>) 等、参照。

45) 以下、Gerd Gigerenzer, CALCULATED RISKS: HOW TO KNOW WHEN NUMBERS DECEIVE YOU, 2002. (吉田利子訳『リスク・リテラシーが身につく統計的思考法: 初歩からベイズ推定まで』2010 年、222-229 頁、参照。)

もう少し細かく見てみよう。DVを受けて、他人に殺された人を  $x$  人とする。パートナー以外(他人)に殺害されている女性の比率は、DVを受けるか受けないかには関係がないので同じとする(仮定1)。次にDVを受けなかった女性が9600万で、DVを受けた女性の24倍となるので、DVを受けた事がなく他人に殺害された女性は  $24x$  人となる。簡単な為、パートナーにDVを受けて、そのパートナーに殺されている人を、弁護団の数値を借り  $400$  万  $(1/2500)=1600$  とする。パートナーにDVを受けて、他人に殺された人の割合  $=x/400$  万、パートナーにDVを受けずに、そのパートナーに殺されている人の割合  $= (5000-1600-25x)/9600$  万、パートナーにDVを受けずに、他人に殺された人の割合  $= 24x/9600$  万となる。更に通常、「パートナーにDVを受けて、そのパートナーに殺されている人の割合」の方が、「パートナーにDVを受けずに、そのパートナーに殺されている人の割合」より大きいと考えられるので(仮定2)、 $1600/400$  万  $> (3400-25x)/9600$  万、整理すると  $x > (3400-1600 \times 24)/24$  ( $x \geq 0$  なら常に成立)。また、 $(5000-1600) > 25x$  から、 $0 \leq x \leq 136$  となる(図2)。

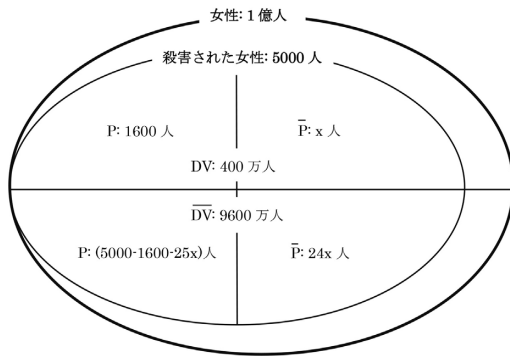


図2 米国における殺害された女性の分布<sup>46)</sup>

弁護団に一番有利な  $x=136$  をとる。DVを受けて殺害されている女性は  $1600+136=1736$  人、その内パートナーに殺されている女性は  $1600$  人なので、DVを受けて殺害されている女性中、パートナーが犯人である

46) P はパートナーに殺害された女性。DV は DV を受けていた女性。

〈370〉 米国における統計的証拠の意義と技法（田中）

割合は、実に 92.2% となるのである<sup>47)</sup>。

この事件の DV という争点に限っては、検察側が「DV を受けていて、そのパートナー以外に殺害された女性の数 x」を統計的証拠として出せなかった事がターニングポイントとなった。

弁護団の誤謬(詭弁)は、「DV を受けていて、パートナーもしくは他人に殺害されている女性(つまり、DV を受けていて誰かに殺害された女性)」の母集団(最大でも 1736 人)を無視して、母集団を DV を受けている女性 400 万と過大評価している点にある。

先述の大統領の例でいう。大統領と同じ人相の人を D.C. で見たという目撃証言があるとすると、人相が一致した人が大統領である確率は、 $1/100=1\%$  に過ぎないので証明力がないとするのが弁護者の誤謬である。しかし、D.C. の人口を考慮すると人相が同じ他人は 0.6 人程度しか存在しないので、目撃者が見たとき大統領が D.C. にいたとすれば、目撃者が見た人が大統領である確率は  $1/(1+0.6)=62.5\%$  なのである。

## おわりに

本稿の主張を一言でいえば、「心証と証拠の証明力を峻別し、証明力に関しては統計的証拠解釈の誤謬を避ける為に、無差別原理の濫用を避けた適切な母数を想定して、条件付き確率の計算を行うべきである。」という事となる。

我が国最高裁も、「DNA 型鑑定を例にとれば、同一の DNA 型となる出現頻度が 1 万人に 1 人の場合、人口 10 万人の街であれば、その中に 10 人いる事になり、決して出現頻度が低い数値とはいえない。しかし、これを対象資料の DNA 型と被告人の DNA 型が一致する割合としてパーセンテージで表した場合、99.99% となり、この数字を常識的に考えると、ほぼ被告人が犯人であるように見えてしまう。このように DNA 型の出現頻度をパーセンテージで表してしまうと、パーセンテージが日常生活において果たす役割やイメージから、科学的な意味合いで理解されてし

---

47) とはいえ、被告人に対する個人的確率(原因)と統計的確率(相関)を混同してはならない。混同例として、Marder v. G.D. Searle & Co. 630 F.Supp. 1087(U.S.D. Maryland 1986), aff'd 814 F.2d 655 (4th Cir. 1987)。



まう危険性がある。」としている<sup>48)</sup>。本稿に引きつけていえば、適切な母集団の想定を行い、上述したような解釈の誤謬を避けようという事になる。

この事は、現在の裁判員制度の下、素人の裁判員が統計的証拠を正しく理解する為にも重要である。その為の市民的教養としての法統計教育の必要性も主張されており<sup>49)</sup>、こうした問題への理解は法曹、市民に共通な課題となっている。その意味で、法学部、法科大学院等の教育において、もう少し取りあげられてもよいのではないだろうか。

### 〔謝辞〕

大阪大学大学院法学研究科の養老真一教授には、常日頃から数理的な理解についての御教示を受け、本稿にも貴重な御示唆を頂き、特に5節において筆者の直観が形となったのは養老教授の御指導、御助言の賜物である。誠に感謝の念に堪えない。

最後になったが、我が国法情報学の開祖の御一人である松浦好治先生は筆者の大学院時代からの恩師であり、今日まで多くの学恩を受けてきた。今後もそれに報いる事を目標とすると共に、先生の益々の御活躍を祈念する次第である。

---

48) 最高裁事務総局『科学的証拠とこれを用いた裁判の在り方』司法研究報告 64 輯 2 号、2013 年、59 頁。

49) 松原克志「法化社会における『科学』教育」日本科学教育学会年会論文集 29、2005 年、11-12 頁。

