

『カルストハンス』の匿名作者……

X^2 -検定の文献学的応用について

新 井 皓 士

2年ほど前のことである。毎年3月31日に催される経済学部清談会に出席しての帰路、私は多少の酔いも手伝って統計学の早川教授に向かい、自分が日頃接する機会の多い15、16世紀の時事的文書類には作者不詳の物が少なくないが、その匿名作者究明の一助として仮説検定の手法が利用できるのではないかと思い、ぼちぼち初歩的な統計学の本をのぞいている、とあやしげな気炎を吐いた事があった。にこにこ聞いておられた早川教授は、それならこんな本があります、とたちまち二、三の書名を挙げられ、私は早速それを、折り畳み傘と家人への手土産を入れて持っていた紙袋の上にメモしたのだったが、細心篤実な早川教授は翌日更に書信を寄せられ、前夜は少しアルコールが作用して人名を間違えたので、とあらためて老童の蒙を啓くの労をとられたのである。その一冊が池田貞男・西田英郎共著『社会のなかの統計学』¹⁾であったが、私はこの書のある種の感動と共に読み、且つその終章にある文体統計論の範例を足がかりとして、自己の研究領域への活用を試みる次第となった。いわゆる機械的ローテーション方式によって経済学部の一員になって17年余、専門を異にする種々の分野からの知的、人的刺戟にふんだんに恵まれてきた私であるが、右のきっかけは偶然性と意義においてとりわけ愉快かつ爽やかであるがゆえに、同じくこれまで幅広い知的刺戟を与えてきてくれた総合的学術誌『一橋論叢』誌上に、このことをあえて記し微意を表したいと思う。

I

『カルストハンス』が1521年初頭に版行をみた小冊子であり、数あるフルークシュリフテンの中でもまず有数の出来映えと人気を一時誇ったものであること、その内容が文書合戦におけるカトリック側の立て役者の一人トーマス・ムルナーを揶揄するものであること、その作者が題名の暗示するような農民層に属する者ではなく人文主義的教養人であること、などについてはすでに述べたところであるが²⁾、その際「筆者は……作者がヴァディアンではない可能性を想定し取り組んでみたのだが、結果的にはむしろヴァディアンの確率が高いという印象を強くした」と書いた。まことに重箱の隅をつつくような些細な問題といえはいえるのだが、「テキストのたった一語の解明の為でも一生を費やす覚悟のないものは学問の道をあきらめよ」などといううろ覚えの言葉をもっぱら頼りに、再びこの「著作者問題」にアプローチしてみたいと思う。今回その方法としてとりあげるのは、カイ自乗分布を利用した二つの分布の一様性の検定 (X^2 “goodness of fit” test) である。すなわち二つの標本分布、具体的には文書(群) A, B における単語長 (word-length) の分布を分類集計し、仮に両者が同一著作者によるものならばこの標本分布も同一型をなすと仮定してカイ自乗検定を行い、この仮説(いわゆる帰無仮説 null hypothesis) が棄却(reject) されるか採択(accept 保留) されるか、問うものだが、その手順はおおむね次のようになる。

(一) 作者不明の文献(群) A と、その幻の著作者である可能性が高い人物の書いた文献(群) B について、使用単語が何文字からなるかに基き 15 程度の階級にわけ頻度分布集計表を作成する。

(二) カイ自乗検定の為の計算表を作成し、 X^2 を計算する。その際、理論度数(期待度数) が 10 以下のクラスは前後のクラスと合併させ計算表を補正する。

(三) 自由度を考慮し、 X^2 分布表を参照して、仮説が棄却されるかどうか調べる。

単語長の頻度分布により作者問題を検討する基本的発想は、上掲の池田・西田に紹介されているブラインガー (Brinegar) 論文が指摘するように、T. C. メンデンホール (Mendenhall) に始まり、特定の著作者のテキストは少なくとも大規模な次元で見る限り無意識裏に使用されている単語の長短の比率が一定であることを示す、という前提に立っている。ただし別個の資料 (標本群) が同一の頻度分布の型を示す場合にはその作者が同一である可能性を否定しない留保的判断にとどまるが、まったく異なる型を示す場合は作者が同一である確率はきわめて低いと判断するものである。メンデンホールは頻度分布を千語当りに換算してグラフで表わし、例えばシェークスピアとベーコンの文体に相違がみられること、つまり両者が同一人物の著作群とはみなされえないことなどを示したが、その約 60 年後にブラインガーが『米国統計学会誌』(1963 年) に発表した論文「マーク・トゥエーンといわゆる QCS 文書」³⁾ は、基本的には同じ前提に立ちつつ、統計的仮説検定の方法を利用して、比較的小標本による確率論的な解答を求めたものである。本稿はこのブラインガーの研究を踏襲し、テキスト (= 標本群) の頻度分布 (= 実現度数) 集計方法において若干独自の工夫を加えたものであるが、このような方法による著作者問題の検討は、比喩的にいえば「指紋検査」より「血液型検査」⁴⁾ に近いとはいえ、今後なるべく多数ないし多種類の標本群について頻度分布を調査し整備ないしデータベース化すれば、少なくとも一定量以上の語数をもつ作者不詳文献について、ある程度の確率をもってその帰属を判断する資料とすることができるのではないかと思われる。

さて上の手順に従って実際に作業に従事してみると、(一) の過程が所用時間からいっても、また結果を左右する点からいっても、9 割 5 分以上の比重を占めることが明らかになる。(二) の過程は多少面倒とはいえ電卓やそろばんでも計算可能であるし、小型コンピュータでもあれば簡単なプログラミングによって、計算表を作らずともデータ入力だけで一挙に計算することができる。

(三) はその計算結果を統計表と対照させるだけであるから、少なくとも作業としては更に簡単である。問題は(一)だが、これを手作業で精確に行なうのは膨大な労力と注意力を必要とし、且つ検算もままならない。まだ電算機が開発されない1939年に『キリストにならひて』(De Imitatione Christi)の著作者問題を文体統計論的に論じ発表したG. U. ユール(Yule)が、単語の長さではなく文章の長さを標本の基準としたのは、少なくとも文章長(Sentence-length)の方が単語長よりも弁別・検算が容易であるということがひとつの理由ではないかと思われるし、ブラインガーも「(単語長による分別)数えあげは明らかに単調で間違えやすい作業であるから」⁵⁾云々と述べているから、あるいは手作業で1万以上の単語を分類集計したのかもしれない。私自身は当初ゲーテとシラーのテキストの一部を文章長について手作業で分類してみても……この時点ではまだ X^2 -testと本格的にとりくむ決心がつかず安直な気持ちで手作業にかかったのである……単純作業において人間はいかに誤りやすいか思い知らされた後、一定の形式でテキストをパーソナル・コンピュータに入力し、次にこのテキストファイルを読みだし単語長に従って分類し順次記録するプログラムと、このファイル化した分類データを集計・表示するプログラムとを、PC フォートランで作成して、作業の一応の効率化をはかった。

一応というのは、正書法その他が均整化する新しい時代はともかく、16世紀頃を対象とする場合には、テキスト・データの入力・作成自体が、機械的にはいかないからである。早い話が句読法や略語、文中の出典表示や引用などの処理、また誤記誤植であるか特殊な単語であるかなどの判断をまじえながら、試行錯誤で進まなければならない。入力作業時間そのものは今後スキャナーでとりこんだ画像データを文字データに読みかえるソフトを利用することによって劇的に(?)短縮できるかもしれないのだが、百パーセント完全に読みとることは期待できないであろうから、キーボードからの入力における単純なミスの排除と同様、結局全文を読みなおす校正的作業が不可欠であるし、上にふれたように時代その他の特殊条件は変わらない。

とはいえ、テキストがひとたび入力されればその単語長頻度分布の数え上げ

過程では、人間的ミスは排除されるわけであるから、依然として入力過程でのミスが絶無ではないとはいえ、数え上げ過程を念頭においたブラインガーの、(表中の) 数値が示すほどには勘定が精密でないかもしれないが、無作為抽出によるダブルチェックは「数え上げのエラーはごく小さく、重要な意味はもたない」ことを示している⁶⁾、という言葉は、少なくとも同じ程度に本研究の方法にもあてはまるといえよう。

以上おおざっぱにカイ自乗検定を行なう作業手順を示したが、計算の為の表の作り方や計算法については注にあげた書⁷⁾などを参照すればよいので、次に本題の『カルストハンス』著作者問題に入る前に、予備作業として行なった幾つかの検定結果を簡単に報告しておきたい。本来疑い深い私はこれらの予備作業を通じて次第に「非主観主義的」 X^2 -test に親しんでいったのである。

II

1. 『キリストにならひて』(De Imitatione Christi) の著作者問題

この高名な 15 世紀の信仰書は、一般にはトマス・ア・ケンピス (Thomas a Kempis) の著述と認められてきたが、彼に精神的影響を与えたヘールト・フロテ (Geert Groote), あるいはパリ大学のジェルソン (Jean Charlier de Gerson) を作者とする説も絶えなかった。統計学者 G. U. ユールはセンテンスの長さ(一文が何語の単語から成るか)を基準に、ケンピス (K) とジェルソン (G) の著述および『キリストにならひて』(I) の三者からランダムにとった標本群を記述統計学的に調べたのだが、その結果が度数分布表の形で池田・西田⁸⁾にも転載されているので、私はこれを X^2 テストにかけてみたのである。文章長をとっていることもあって、標本数が各群とも 1200 余と少ない割に階級数は 20 程度であるため、いわゆるスタージェス (Sturges) の階級数を参照して階級をそれぞれ 11 に補正した。その結果、自由度 (d. f. = degree of freedom) 10 で、I と K の適合度検定では $X^2=31.83$ 。I と G 1 では $X^2=239.07$ 。I と G 2 では $X^2=197.56$ 。G 1 と G 2 では $X^2=10.97$ という結果

をえた。自由度 10 における X^2 分布の、有意水準 (level of significance 危険率) 5%, 1%, 0.1% 点はそれぞれ 18.307, 23.209, 29.588 であるから、ケンピスにしてもジェルソンにしても計算された X^2 値はこれらのいずれよりも大きく、同一性の仮説は棄却される。但しケンピスの場合は有意水準 0.1% (上側確率 0.001) の値に接近しており、標本数を増やせばあるいは仮説が棄却されない余地を残すと考えられるのに対し、ジェルソンの場合はその可能性はまずないと思われる。そしてジェルソン自身の二つの標本群同志は小さな X^2 値を、すなわち両者の同一性の仮説が採択されることを示しており、この検定が著作者問題に関してかなり有効ではないかという予感を私に与えた。『キリストにならひて』については、ケンピスがフローテの草稿を下敷にしたとする説もあり、もしそうだとすれば上の値 (棄却はされるが、境界値に近い) も諾けるところである。いずれにしてもこの問題は今後ケンピス及びフローテの著作物から直接標本をとり、次に述べる単語長による検定をも含めてあらためて検討する必要がある。

2. 『若きヴェルテルの悩み』とそのパロディ『若きヴェルテルの喜び』

多感な青年が人妻に恋こがれたあげく思い詰めて自殺する、という、陳腐といえは陳腐な田舎町的话题を、ヨーロッパ的次元の近代小説に結晶化させた J. W. ゲーテのデヴュー作は、もはや現実的恋愛とは縁遠くなった年頃ないし時代にそぞろ読み返してみると、存外ほほえましくも清潔でせつない余韻を感じさせるものだが、旧体制なお揺るぎない 1774 年当時の逼塞した社会状況のなかではむしろ一種危険で不吉な感化力を感じさせるものだったらしく、分別ある人士の中には少なくとも青年が徒にヴェルテル的思想や行動に走らぬようにと、説教や文筆活動によって善導をこころみる動きも少なくなかった。就中ベルリンの啓蒙出版社主フリードリヒ・ニコライ (Nicolai) が、ヴェルテルの『なやみ』の大団円を改作した『喜び』は著名だが、この両書の初版を一巻に収めた写植本⁹⁾をテキストとして、私は次に単語長による X^2 検定を試みてみた。ニコライの作は文体まで模倣しようとしたパロディではないが、少なく

ともテーマが共通の地盤に根付くものであるから、ある程度の類似性は認められようし、逆にこの適合度の検定が予期に反し、両者を同一とする帰無仮説を棄却しないとすれば、この方法活用の前途は暗いことになる。

さてゲーテの作品からはその第二部より最終的に 6000 余語をとり、ニコライのそれは全編で 4800 余語を得た。これらは固有名詞その他を排除する作業を経た数であるが、標本の総量としては必ずしも十分とはいえないかもしれない。しかしニコライについては少なくともこのテーマに関してはこれ以上増やすことができないから、ゲーテの方を増やすしかないことになるが、この作業の過程で標本量、とくに対照群 (controle groupe 作者未詳のテキストに対する、作者が明かなテキスト。ここでは仮にニコライが匿名であったとすればゲーテの作品がそれにあたる) の標本量が、検定結果に微妙な影響を及ぼすことが如実になった。つまり入力の間隔に何度か経過時点的検定をしてみると、5000 語前後の段階までは両者同一性の仮説は棄却されなかったのである。具体的にはニコライの総語数 4810 に対し、ゲーテの総語数 5029 で検定をしてみると、 $X^2=34.352$ (自由度 13) となり、有意水準 0.1% (34.528) では微差ながら棄却されない。ゲーテの総語数を 6121 にあげた時点ではじめて、 $X^2=39.025$ (自由度 13) となり、有意水準ないし危険率 0.1% において両者同一性の仮説は棄却される、換言すれば、上の二作品が同一作者のもの (あるいは同一の文体) とする仮説は否定され、かつこの検定結果が誤っている (いわゆる第一種の過誤の) 確率は千分の一以下、という統計的結論に達する。また例えばニコライを三分割してその第三区分文書を取り、上の 6121 語のゲーテとの適合度を検定してみると自由度 12 における右側 5% 点 (21.026) を越え棄却されるが、3800 語、5000 語段階ではこの限界値内にとどまり保留される。この保留は直ちに両者が同一であるという結論には至らないとはいえ、対照群の標本量が微妙に検定結果に作用することを示している。なお一般に統計では有意水準として 5% ないし 1% をとるようであるが、文体統計論の適合度検定ではその対象とする資料の性質上、有意水準に 0.1% をとるのが適切とみなされている。

3. 『農民大衆の集会に与える』と T. ミュンツァー

15世紀末『ジギスムントの改革』その他の帝国・社会改造論や「哀れなコンラート」「ブントシュール」蜂起など種々の形をとって現われた社会変革運動と、贖宥状売買の是非を問うルターの『95箇条の提題』(1517)を起点として、忽ちカトリック教会の基盤を揺るがす運動に発展した宗教改革運動は、少なくとも大衆的次元ではわがちがたく結びついた制度的改革運動(Reformation)として理解され、いわゆるドイツ農民戦争に至るが、その渦中1525年に『農民大衆の集会に与える』(An die versammlung gemayner Pawerschafft)と題する一編の檄文が南部ドイツで発表され流布した。全11章から成り農民蜂起の神学的・歴史的弁護論と実践的・防衛的団結論を述べたこの文書は、起草地がおそらく上シュヴァーベンであり、1925年5月中旬にニュルンベルクのヘルツェル(Hieronymus Hölzel)の下で印刷されたことまではほぼ確実とされるが、匿名の著者についてはなお論議が尽きない。古くはツィンマーマンのミュンツァー説に始まり、ミュンツァーのみならずツヴィングリの思想的影響が指摘され、さらにはヴァルツフートの急進的説教師フープマイア説、『農民の十二箇条』に関係するメミンゲンの説教師シャペラー説まで様々な推測がなされている。このうちミュンツァー起草説は、彼が1524年9月下旬にミュールハウゼンから追放され翌1525年2月末に再び同市に戻るまでの期間をフランケン、シュヴァーベン、スイスに至る遍歴に過ごし就中ニュルンベルクで『暴露』『弁護論』¹⁰⁾を印刷せしめたという状況証拠に基づくものだが、現在では可能性が比較的薄いとされるこの説の当否を X^2 -testにより検証してみよう。対照(基準)群にはむしろシャペラーなどを取りたいのであるが、手元に適当な資料がなく、ミュンツァーに関しては私が初めてこの文書に接したカチェロフスキの『農民戦争文書集』¹¹⁾が、ミュンツァーの三通の書簡を冒頭に掲げたミュンヘン国立図書館蔵本を底本とするうえに、上掲の『暴露』と『弁護』をも併せ収めているので、校訂を含めて編集方針の一貫性という点で好都合だという事情もある。対照群としては『弁護論』全文と『暴露』の一部をとったが、これは後者は「ダニエル書第二章」講解のかたちで行なわれた一種の御前

説教なので、比較対照として前者の方がより適切であると判断した為である。

さて単語長の分布度数を調べるに当たって、原テキストから何を削除すべきかという問題が、上述の『ヴェルテル』の場合にくらべてこの資料ではいささか複雑である。細かな点は別として、第一に、当時の表記法が、句読法にしろ引用や出典表示にしろ、17世紀後半以後にくらべて遥かに錯綜していること、第二に、聖書その他の引用とイエス・キリスト始め諸聖人等の名称が重要な文成分を成していることが指摘されよう。句読法が安定しないことはこれらの文書を検定対象とするにあたって、文章長ではなく単語長をさしあたり選んだ理由でもあるが、試行錯誤の結果私は、聖書等の引用はそのままとりいれ、文中に出没する出典表示や固有名詞は原則として除去、略語略記は一定の方針で処理することとした。むろん標題も省いたが、直接間接を含め聖書からの引用文

図 1

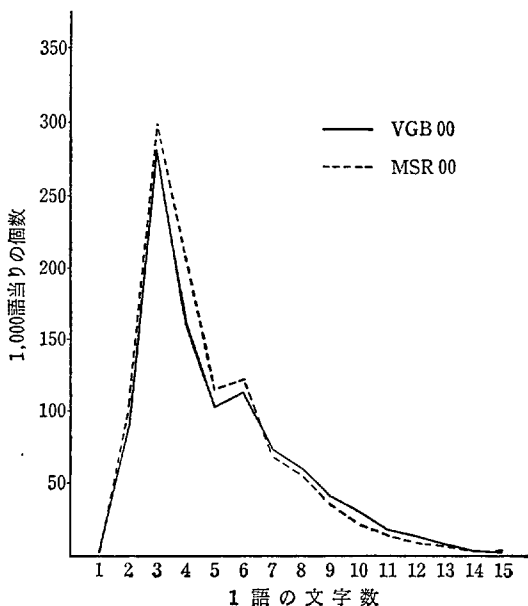


表 1

単語長	VGB00	MSR00	MSRFP
1	7	5	12
2	727	598	775
3	2381	1813	2556
4	1357	893	1211
5	876	700	981
6	959	719	1053
7	629	439	571
8	518	340	474
9	342	223	301
10	260	135	200
11	152	86	126
12	116	52	76
13	61	40	54
14	30	17	22
15以上	22	22	32
Sum	8437	6082	8444
Mn	5.096	4.905	4.925
StD	2.605	2.493	2.498
CoV	0.511	0.508	0.507

Mn : 平均値
StD : 標準偏差
CoV : 変異係数

はそのままとしたのは、これを除くと文を成さないケースがあり、また当時における聖書引用はルター訳とは限らないのみならず、引用文、典拠等の厳密でない場合も少なくない、つまり著作者固有の特徴を無意識裏に反映しているとみられるからである。

結果(図1表1参照)は、『農民大衆の集會に与える』(VGB 00)と『弁護論』(MSR 00)のみの適合度検定が、 $X^2=44.614$ (d. f.=13) となり、対象群に『暴露』の一部を加えた場合(MSRFP)が、 $X^2=56.145$ (d. f.=13)であった。これらはいずれも自由度13の0.1%点34.528より大きいから、統計理論上『農民大衆の集會に与える』文書がミュンツァーの起草したものである可能性は極めて小さいといえよう。なお念のため固有名詞等を削除しない『農民大衆の集會に与える』原テキスト(VGB)と処理済みテキスト(VGB 00)との適合度を検定してみると、 $X^2=1.223$ (d. f.=13)、同様に『弁護論』では、 $X^2=1.968$ (d.f.=13)であった。

III

『カルストハンス』(Karsthans)の著作者問題を再びとりあげたきっかけは、かつて『カルストハンス考』で類似性を指摘した『豺狼歌』(Das Wolfgesang)の、ユダス・ナザレイ(Judas Nazarei)という風変わりな筆名の主がやはり当時の推測通り「スイス、ザンクト・ガレンのヨアヒム・フォン・ヴァット、通称ヴァディアンまたはヴァディアース(1484—1551)」であることが判明したからである。といってもそれは何等新しい発見ではなく、NdL叢書中の一巻として1896年に刊行された『古きと新しき神、信仰、教えについて』(Vom alten vnd nuen Gott, Glauben, vnd Ler 以下『新旧信仰論』と略す)の解説において、校訂者キュック(Eduard Kück)がすでに論証しているのだが¹²⁾、私は最近ようやくこの書のコピーに目を通す機会を得、『豺狼歌』と『新旧信仰論』の著者名“Judas Nazarei”が、“J. Vadian arzet”のアナグラムであることを知ったのであった。ヴァディアンは往時の典型的な博学家(Polyhistor)で、桂冠詩人であるのみならず医学や法学をもウィーンで修めたが、

故郷ザンクト・ガレンへはまず医師として帰還したとされ、それがこのアナグラムにもあらわれていることになる。

もっともキュークは上の両書と『カルストハンス』の関連には触れていないし、一応通説とされる『カルストハンス』＝ヴァディアン作者説に対する疑問は依然として絶えないので、私は上に示した方法によって文体統計論的方向からこの問題を再検討してみようと思うのである。この場合、検証すべき作者不詳文献はいうまでもなく『カルストハンス』であり、作者が判明している対照群は『豺狼歌』¹³⁾と『新旧信仰論』ということになるが、のちに述べる理由によって始めに『カルストハンス』と『豺狼歌』のみの適合度検定を行い、次に『新旧信仰論』とこれらとの適合度を調べることにする。

1. 『カルストハンス』(KH) と 『豺狼歌』(WG)

結論から先に言えば、この両者の適合度は極めて良好、つまり両者が同一作者の手になるとする仮説は、固有名詞や出典その他の削除等の処理をほどこした標本(KH 00 と WG 00. 図2表2参照) 同士のみならず、『カルストハンス』におけるラテン語文や猫語(?擬声音など)を含めた原文通りの標本(KH 22 と WG 22)によっても、棄却されない。すなわち KH 00 と WG 00 とは、 $X^2=24.344$ (d. f.=13) であり、KH 22 と WG 22 では $X^2=18.128$ となって、共に有意水準 0.1% 点 (34.528) のみならず 1% 点 (27.688) をも下回っている。

2. 『豺狼歌』(WG) と 『新旧信仰論』(ANG)

『豺狼歌』はその表題から予測されるような韻文ないし歌謡形式ではなく、羊を欺く狼の比喻になぞらえて既存教会制度の弊害を福音書を論拠に批判する諷刺書だが、『新旧信仰論』はかかる教会制度を歴史的に通観し、教会権力の肥大化は「老獺な蛇＝サタン」に唆された現世的悪弊なることを主張する論証的性格の啓蒙文書である。ここでいう新旧は、単純に新教旧教という意味ではなく、既存教会こそ古来の由緒ある信仰に反しローマ首座制を固め神ならぬ

図 2

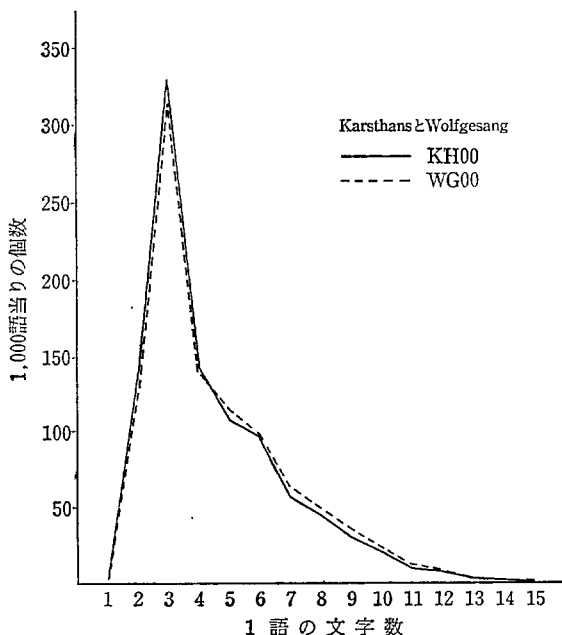


表 2

単語長	KH00	WG00
1	18	11
2	1034	1412
3	2546	3486
4	1111	1552
5	833	1271
6	754	1093
7	522	710
8	355	553
9	244	399
10	156	260
11	77	136
12	59	88
13	28	29
14	15	20
15以上	8	25
Sum	7760	11045
Mn	4.556	4.670
StD	2.345	2.410
CoV	0.514	0.516

「人間の」教えを広めてきた「新しい」偶像崇拜であり、福音書の教えに帰るのはむしろ「古い」正統的信仰だ、という意味で使われており、その当否は別として、新旧の価値観がなお古きを良しとする方向であることを窺わせるが、いずれにしても、テーマとしてもジャンルとしても、ムルナーの諷刺と反論を目的とする対話体の『カルストハンス』とは少々違いすぎる印象を与える。私がこの文書を直ちに『豺狼歌』と組み合わせて『カルストハンス』の対照群にしなかったのはこの点が主な理由だが¹⁴⁾、それでは同一作者によることが明らかかな『豺狼歌』と『新旧信仰論』の適合度はどうであろうか。後者からそのほぼ 4/5 をなす第一部をとって固有名詞や年号等を削除したもの (ANG 00) と、WG 00 との適合度をみると、 $X^2=31.108$ (d. f.=14) となり、自由度 14 の 0.1 % 点は 36.123 であるから、両者同一性の仮説は……当然ながら……採択 (保

留)される。但し無修正同士の『豺狼歌』(WG 22)と『新旧信仰論』(ANG 22)では、 $X^2=50.666$ (d. f.=14) となって同一性仮説が棄却されるのみか、『新旧信仰論』自体の無修正版標本 (ANG 22) と修正版標本 (ANG 00) ですら、 $X^2=49.902$ (d. f.=14) となって同一性仮説が棄却されてしまう。これは ANG 22 に固有名詞や年号等が約 1200 語含まれている結果と考えられ、ジャンルあるいは記述形態が異なるテキストを扱う場合は十分慎重な対応が必要なることを物語っているといえよう。

3. 『カルストハンス』と『新旧信仰論』、そして『新カルストハンス』

それでは対話体と論述体という異なる表現様式による『カルストハンス』と『新旧信仰論』の適合度はどうであろうか。KH 00 と ANG 00 とでは $X^2=43.693$ (d. f.=12), また ANG 00 と WG 00 を合計した対照群 (ANGWG) とでは $X^2=35.669$ (d. f.=13) となり、ともに自由度 12 及び 13 の 0.1% 点 (32.909/34.528) を越えるから、同一性の仮説は棄却されるのである。つまりこれまでの結果をまとめると、『カルストハンス』と『豺狼歌』は同一作者、『豺狼歌』と『新旧信仰論』も同一作者、しかし『新旧信仰論』と『カルストハンス』は同一作者とはいえない、という、一部が欠けたリング、あるいは『豺狼歌』を中にしたヤジロベエのような、いまひとつ落ち着かない状況になってしまった。無論これは表現様式の相違が大きく作用しているらしいが、そもそも考えてみると『カルストハンス』自体が全体としては会話体とはいえ、前半と後半では随分異なる印象を与えている。つまり前半は会話を通じてトーマス・ムルナーという人間を諷刺するおもむきであるとするれば、後半はムルナーのカトリック弁護論を論破しようとする論証的な傾向が内容的にも文体的にも強く、『新旧信仰論』と比較するには後半こそふさわしいと思われる。そこで私は『カルストハンス』をほぼ二分して、後半部 (KH 2 h) と WG 00, ANG 00, ANGWG との適合度を X^2 検定にかけてみた。その結果は、自由度 12 でそれぞれ、9.852, 24.693, 16.623 となり、いずれも同一性仮説は採択 (保留) されるのである。この場合 KH 2 h の総語数が 4520 とやや少ないのが気になる……経験的には

図3

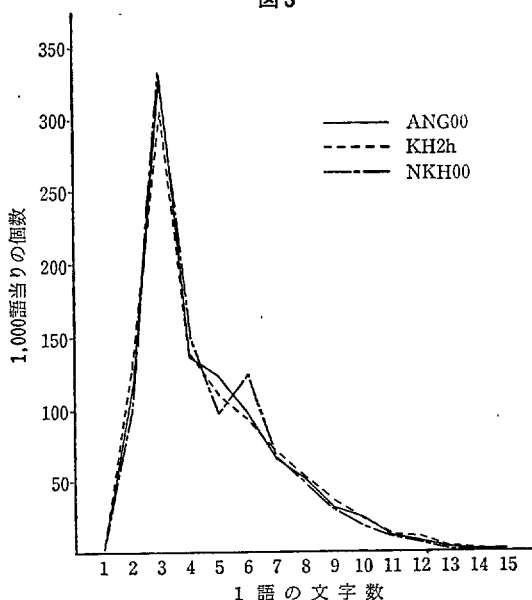


表3

単語長	ANG00	KH2h	NKH00
1	16	4	5
2	1891	577	1053
3	5503	1409	3327
4	2296	631	1528
5	2059	499	990
6	1637	425	1246
7	1096	320	677
8	869	239	501
9	530	167	311
10	416	112	185
11	193	54	106
12	127	48	73
13	60	19	33
14	24	10	5
15以上	23	6	20
Sum	16740	4520	10060
Mn	4.676	4.729	4.659
StD	2.367	2.464	2.300
CoV	0.506	0.521	0.494

語数が少ないと X^2 は小さ目になるようだ……が、だからといってこの結果が全く信頼できないものではないことは、次にふれる『新カルストハンス』と KH 2h の検定結果は、 $X^2=68.044$ (d. f.=12) となって同一性仮説がきちんと棄却されていることから認められよう。(図3表3参照)

『新カルストハンス』(Gesprechiehlin Neüw Karsthans) は、一時期 F. v. ジッキンゲンのもとに身を寄せ、U. v. フッテンのラテン語『対話篇』をドイツ語に訳したりしていた、後年のシュトラースブルク宗教改革の指導者 M. ブッツァー (Martin Bucer, Butzer, Bucerus) の作品であることはほぼ疑いないとされている。ジッキンゲンとカルストハンスの対話という形で、福音主義的立場から既成教会制度を批判するのだが、元祖『カルストハンス』ほど清新奇抜な趣はないかわりに論旨は平明である。この『新カルストハンス』の標本

(NKH 00) と上記の 4 標本およびそれら相互間の適合度検定結果を組合せ表のかたちで次に示そう。

	KH 00	KH 2 h	WG 00	ANG 00	NKH 00
KH 00	* *	$X^2=17.839$	$X^2=24.344$	$X^2=43.693$	$X^2=70.709$
KH 2 h	$X^2=13.839$	* *	$X^2=9.852$	$X^2=24.693$	$X^2=68.044$
WG 00	$X^2=24.344$	$X^2=9.852$	* *	$X^2=31.108$	$X^2=94.749$
ANG 00	$X^2=43.693$	$X^2=24.693$	$X^2=31.108$	* *	$X^2=100.797$
NKH 00	$X^2=70.709$	$X^2=68.044$	$X^2=94.749$	$X^2=100.797$	* *

(d. g. は WH 00/ANG 00 が 14, WG 00/KH 00 が 13, 他は 12)

4. 2 標本 t-検定

ブラインガーに倣って最後に私は 2 標本検定も試みてみた。 X^2 検定に使用した標本群 KH 00, WG 00, ANG 00 をそれぞれ語数 2000 程度のグループに分割し、それぞれのグループにおける 3 文字単語, 4 文字単語, 5 文字単語の比率を計算して表にした上で, t-test 用の計算プログラムを組んでみたのである。その結果は表 4 のようになるが, 得られた t の値と t-分布表によって判断すると, 結果はほぼ X^2 -検定による判断と一致すると言ってよいと思われる。ただ 2 標本 t-検定についてはまだ資料も経験も乏しい試技的段階なのでこれ以上のコメントは控えておきたい。

結 語

以上の過程をふまえて私は、『カルストハンス』の作者は通称ヴァディアン, すなわち Joachim von Watt (Vadianus; 1484—1551) である, と断定してもほぼ間違えないと考える。ヴァディアンについてもなお未見かつ問題視される文書が残っているが, 今後は更に種々の文献資料について philologisch-statistisch な接近を試み, 広義の歴史研究に資する一補助手段を確立したいと思う。

(1988. 12. 6)

表4

各〔分割〕文書における比率				
		3文字語	4文字語	5文字語
〔KH00〕	1	0.350	0.154	0.098
	2	0.352	0.143	0.108
	3	0.318	0.135	0.107
	4	0.305	0.145	0.115
平 均		0.331	0.144	0.107
〔WG00〕	1	0.312	0.118	0.120
	2	0.343	0.134	0.102
	3	0.298	0.151	0.118
	4	0.329	0.148	0.111
	5	0.303	0.153	0.122
平 均		0.317	0.141	0.115
〔ANG00〕	1	0.322	0.126	0.120
	2	0.328	0.126	0.114
	3	0.339	0.142	0.116
	4	0.323	0.142	0.140
	5	0.341	0.138	0.142
	6	0.351	0.128	0.113
	7	0.306	0.150	0.113
	8	0.325	0.142	0.122
平 均		0.329	0.137	0.123

2標本t-検定

KH00とWG00	
d.f.=7	
3文字語	t= 1.018
4文字語	t= 0.419
5文字語	t=-1.475

WG00とANG00	
d.f.=11	
3文字語	t=-1.370
4文字語	t= 0.621
5文字語	t=-1.297

ANG00とKH00	
d.f.=10	
3文字語	t=-0.177
4文字語	t=-1.414
5文字語	t= 2.377

本研究は1984～85年度、科学研究助成金のお蔭を被っている。おくれらせながら慎んで謝意を表わしたい。

- 1) 池田貞男・西田英郎：『社会のなかの統計学』, Tokyo ⁴1984
- 2) 新井：「カルストハンス考」(『一橋論叢』95巻3号, 1986)
- 3) Brinegar, Claude S.: Mark Twain and the Quintus Curtius Snodgrass letters: A statistical test of authorship. (in: Journal of the American Statistical Association, 58, 1963, p. 85—96)

- 4) ibid. p. 87
- 5) ibid. p. 92
- 6) ibid.
- 7) 石川栄助:『実務家のための新統計学』, Tokyo ¹⁸1983
ラオ (奥野・長田・篠崎・広崎・古河・矢島・鷲尾訳):『統計的推測とその応用』, Tokyo ³1980 など
- 8) 池田・西 p. 234
- 9) Die Leiden des jungen Werthers (Leipzig 1774) & Freuden des jungen Werthers (Berlin 1775), beide anonym: Dortmund 1978 (bibliophil. TB)
- 10) Müntzer, Thomas: Außlegung des andern vnterschryds Daniels deß propheten gepredigt auffm schlos zu Allstedt vor den tetigen thewren Herzcogen vnd vorstehern zu Sachssen durch Thomam Müntzer diener des wort gottes, Allstedt 1524 (邦訳『まやかしの信仰のあからさまな暴露』)
ders.: Hochverursachte Schutzrede vnd antwort, wider das Gaistloße Sanfft lebende fleysch zu Wittenberg, welches mit verklärter weyße, durch den Diepstal der heiligen schrift die erbermdliche Christenheit, also gantz jämerlichen hesudelt hat, Allstedt 1524 (邦訳『きわめてやむをえざる弁護論』)
- 11) Kaczerowsky, Klaus (Hg): Flugschriften des Bauernkrieges, Hamburg 1970
- 12) Judas Nazarei: Vom alten und neuen Gott, Glauben und Lehre (1521); Mit Abhandlung und Kommentar hrsg. v. Eduard Kück, Halle/S. 1896
- 13) Judas Nazarei: Das Wolffgesang. o.O., o.J., In: Flugschriftensammlung Gustav Freytag (Microfiche), Nr. 3940.
- 14) 『新旧信仰論』(注 12) は原典にきわめて忠実な校訂版であるが、それでも“Flugschriften aus der Reformationszeit” 叢書全体の編集方針として、例えば, mhd. iu を ü とする, など校訂者の手が加わっている。これが本書を直ちに対照群に加えられない第二の理由である。
- 15) Gespräch biechlin neüw Karsthans. o.O., o.J., In: Flugschriften des Frühen 16. Jahrhunderts (Microfiche), Nr. 475; cf. Böcking, Eduard (Hg): Ulrichs von Hutten Schriften, Bd. 4, p. 649—681
- 16) 参考の為に掲げたが, 表 1~3 の平均値, 標準偏差, 変異 (変動) 係数は, 単語長 15 以上のものを 15 とみなし計算しているので, 完全に厳密なものとはいえない。
* この拙ない論考を退官された加藤二郎教授に捧げる。

(一橋大学教授)