令和7年度　お茶の水女子大学　共創工学部　文化情報工学科

学校推薦型選抜

小論文

※Ⅰ，Ⅱともに解答しなさい。

※答案用紙には、解答とともに計算過程も記入しなさい。

I　n 個の実数値データ $a_1, a_2, \ldots a_n$ に対して、関数

$$f(x) = \sum_{i=1}^{n} (x - a_i)^2 = (x - a_1)^2 + (x - a_2)^2 + \cdots + (x - a_n)^2$$

を考える。このとき、以下の各問いに答えよ。なお、答案用紙には、問いの番号を記して解答すること。

問 1. 関数 $f(x)$ が最小となる x の値と最小値を求めよ。

問 2. 関数 $f(x)$ の最小値を与える x の値と、その最小値は、データ $\{a_1, a_2, \ldots, a_n\}$ のどのような統計量と関連しているか述べよ。

II 次の文章を読み、後の問いに答えなさい。なお、答案用紙には、問いの番号を記して解答すること。

『キリストにならいて The Imitation of Christ (De imitatione Christi)』は、「第二の福音書」ともいわれる。カトリック信徒必読の書とされ、キリスト教を信じる人々の間では、『聖書』に次いで広く親しまれている書物である。しかし、修道士たちの信仰への営みについて書かれたこの書物の作者が誰であるかということについては、長い間論争が続いてきており、今日でも完全に決着がついたとは言えないようである。

『キリストにならいて』の作者としては、アウグスティヌス会の聖アグネス修道院副院長のトマス・ア・ケンピス（Thomas á Kempis, 1379(1380)-1471）、パリ大学総長のジャン・ジェルソン（Jean Gerson, 1363-1429）、ケンピスが属していた修道会の創立者ヘーラルト・ホロート（Gerard Groote, 1340-1384）の名が挙げられてきた。

文の長さの比較

ここで紹介するのは、同書の作者はケンピスとジェルソンのいずれかであるとされていた1930年代に、著名な統計学者のウドニー・ユール（George Udny Yule, 1871-1951）が行った研究である。

彼は、作者の文章の特徴は文の長さに現れると予想した。そこで、まず次の4人の作家、フランシス・ベーコン（Francis Bacon, 1561-1626, 英国の政治家）、コールリッジ（Samuel Taylor Coleridge, 1772-1834, 英国の詩人、批評家）、ラム（Charles Lamb, 1775-1834, 英国の随筆家、批評家）、マコーリイ（Thomas Babington Macaulay, 1800-1859, 歴史家、批評家、詩人）の文の長さを調べてみた。その結果、文の長さの平均値や文の長さのバラつき（四分位範囲と呼ばれる統計量で測定）は作家間で異なるが、同じ作家の作品では、ほぼ同じような数値となるという結果を得た。次にユールは、ケンピス、ジェルソンの著作と『キリストにならいて』から、いくつかの文をサンプルとして選び、それらの文の長さの平均値やバラつきを求めてみたところ、表4・3のような結果を得た。

表4・3 『キリストにならいて』と、ケンピス、ジェルソンの著作の文の長さとバラつき

	文の長さの平均値	バラつき（四分位範囲）
『キリストにならいて』	16.2語	9.2語
ケンピスの著作	17.9語	11.8語
ジェルソンの著作	23.4語	18.8語

これらの数値を見ると、ケンピスの文章はジェルソンの文章よりも短く、文の長さのバラつき（四分位範囲）も小さく、二人の文章には違いが見られる。また、『キリストにならいて』の文章は、ジェルソンの文章よりもケンピスの文章にずっと近い。このことから、ユールは『キリストにならいて』の作者は、ジェルソンよりもケンピスの可能性の方が高いと判断している。

語彙の豊富さの比較

さらにユールは、作家の有する語彙量という観点から、1944年に再度この問題に取り組んでいる。ただ、作家の有する語彙の量を正確に把握する方法については、著名な確率論や統計の専門家が試みてきたが現時点でも解決されていない難しい問題である。そこで彼は、語彙量そのものを把握するのではなく、語彙の豊富さの大小を比較できるようにするために、単語が出現する可能性を「ポアソン分布」と呼ばれる稀にしか生じない事象の分析に用いる確率分布でモデル化し、K特性値と呼ぶ指標を提案した。K特性値は、文章に出現する単語の種類が多いほど、つまり語彙が豊富なほど小さな値になり、反対に出現する単語の種類が少なく語彙が貧弱なほど大きな値となるようになっており、語彙量そのものは測定できなくても、作家の語彙の豊富さは比較できる。

さて、『キリストにならいて』、ケンピスの作品、およびジェルソンの作品にみる K 特性値を求めてみると、表4・4のような結果となった。

表4・4 『キリストにならいて』と、ケンピス、ジェルソンの著作の K特性値

作品	K特性値
『キリストにならいて』	84.2
ケンピスの著作1	113.7
ケンピスの著作2	110.9
ケンピスの著作3	100.1
ケンピスの著作4	66.9
ケンピスの著作5	59.7
ジェルソンの著作	35.9

さてケンピスとジェルソンを比較すると、ジェルソンの方が K特性値が小さく、したがって、ケンピスより語彙が豊富であることがわかる。そして問題となっている『キリストにならいて』の K特性値はケンピスの作品に近い値となっている。このような K特性値の比較から、ユールは文の長さを用いた分析と同様に、ケンピスの方がジェルソンより『キリストにならいて』の作者である可能性が高いと判断している。

しかし前述したように、ケンピス、ジェルソンの他に、ケンピスが属していた修道会の創立者ヘーラルト・ホロートが書いたものへ、ケンピスが修正加筆しながら編集したという説も検討する必要があり、いまだに完全には解決を見ていないといえよう。

（村上征勝『この本を書いたのは誰だ？統計で探る“文章の指紋”』より一部改変）

問 1. ユールの行った『キリストにならいて』の著者の推定方法は当時としては先進的であったが、現代では問題点もあるとされる。問題があるところを指摘しなさい。

問 2. 日本語の文章の書き手を識別するとき、文の長さを用いることは困難であることが多い。その理由を論じなさい。