

## 市販飲料の香気分析データに対するデータサイエンスの適用例

(2022年5月20日受付)

(2022年10月11日受理)

芹野 武<sup>a, b)</sup>、平原嘉親<sup>a, c)</sup>

- a) 東京農業大学 総合研究所研究会 食・農データサイエンス部会 幹事  
b) アジレント・テクノロジー株式会社  
c) 摂南大学農学部食品栄養学科

### Application of data science to the aroma analytical results of beverages commercially available

(Received May 20, 2022)

(Accepted October 11, 2022)

Takeshi Serino<sup>a, b)</sup>, Yoshichika Hirahara<sup>a, c)</sup>

- a) Secretary of Data Science Committee on Food and Agricultural Research, Research Institute for Agricultural and Life Sciences  
b) Agilent Technologies Japan, Ltd.  
c) Department of Food Science and Human Nutrition, Faculty of Agriculture, Setsunan University

#### Abstract

Applying the statistical methods on the food analysis data plays important role in the field of food safety, food development, research, and quality control, etc. In this paper, an application example by data science in food analysis using the aroma component data of commercial canned coffee obtained by analysis was introduced. The aroma components of 6 types of fragrance-free products and 6 types of fragrance-added products of canned coffee were analyzed using headspace GC/MS. Analytical data of aroma components contained in coffee was statistically analyzed by R program. Four hundred ninety aroma compounds were extracted by the estimation of GC/MS library search (NIST MS library). Principal component analysis, cluster analysis, False Discovery Rate (FDR) test, etc. were applied to the estimated compounds, and they were classified into two groups, fragrance-added and fragrance-free coffee. Statistical methods are one of the important tools in food analysis as demonstrated by this article, and analytical ability is especially useful for food research and food development.

**Keywords** : 食品データ、統計解析、主成分分析、クラスター分析、FDR

food data, statistical analysis, principal components analysis, cluster analysis, FDR

## I 緒言

食品成分や食品中の有害化学物質の分析は、食品、添加物等の規格基準<sup>1)</sup>などへの適合性の判断をはじめ食品の開発、研究、品質管理などさまざまな分野で用いられている。近年では、食品分析に精度管理が導入され<sup>2)</sup>データの信頼性確保が益々重要になってきている。これら分析データは客観的な評価が必要であり、統計的な手法を用いたデータの取り扱いが重要になる。統計手法では、サンプリングでは均一性や代表性、定量では検量線の相関性や再現性、解析では差異解析や多変量解析、妥当性では検定などが用いられており、食品分析に必要と

される統計的手法を理解し、目的に応じて正しく適用することが重要である。しかし、統計で扱う数式が複雑であり、様々なデータをどのように取り扱えばよいのか判断に迷うことが多く、統計的処理を苦手とする分析者も少なくない。そこで、本稿では、分析データによる統計解析を例としたデータサイエンスの事例を紹介する。

コーヒーは世界三大嗜好品のひとつと言われ、多くの国で飲まれている。コーヒー豆には、アラビカ種とロブスタ種などがある。アラビカ種の豆は酸味が強く、花のような甘い香りが特徴的である。日本では、カフェだけでなく自動販売機やコンビニエンスストアなどにおけるコーヒーマシンなどコーヒーを楽しめる文化が広がっ