

Pythonを使った測定データ解析の第一歩

○森 加奈恵、栗山 恵輔、新地 姉理華

佐賀大学 総合分析実験センター

1.はじめに

総合分析実験センターでは、職員のリスクリングとして月1回90分の時間を設け、Microsoft PowerAutomateの活用やPythonを使ったプログラミングの研修を行っている。職員は全員Pythonでのプログラミング経験はないため、ゼロからプログラムを組み立てることよりも、生成AIを活用してプログラムを作成することに重きを置いている。昨年度、利用料金請求用のデータ処理を自動化したことをきっかけに、利用料金の集計の自動化にも取り組んでいるところである。

研修を行っていく中で、Pythonで測定データの解析もできるのではないかと思うようになった。今回はプレートリーダーと質量分析のデータを用いて、データ解析のためのプログラムの作成に挑戦した。なお、本研究で用いた生成AIはMicrosoft Copilotである。

2.プレートリーダーの検量線4PL

プレートリーダーはセルの代わりにマイクロプレートに入ったサンプルを測定する分光光度計である。一般的には一度の測定で96サンプルのデータを取得する。この時、検量線は理論的には直線になるのだが、生物学的研究では通常測定値の上限と下限が存在するため、S字型のシグモイド曲線を描く。したがって、より良く近似できる4パラメーターロジスティック曲線(4 parameter logistic curve、以下4PL)が利用されている⁽¹⁾。4PLは4つのパラメーターを用いたシグモイド曲線であり、Excelの機能で描くことが難しいため、Pythonを使って描画するプログラムを検討した。一般的な4PLの式は以下の通り⁽²⁾：

$$y = \left(\frac{A - D}{1 + \left(\frac{x}{C}\right)^B} \right) + D$$

ここで、xは濃度、yは検出値、Aは検量線濃度ゼロでの検出値、Bは傾き、Cは変曲点、Dは検量線濃度最大での検出値である。

本研究では、まず、生成AIを用いて、検量点7点、未知サンプル24点、各点3回測定したとするテストデータを作成させた。次に、Pythonでプログラムを作成し、グラフを描画させた(図1)。また、求めた検量線から未知サンプルの濃度を定量させる。最後に、解析ソフトウェアでの解析結果と比較し、プログラムの正確性を検証する。

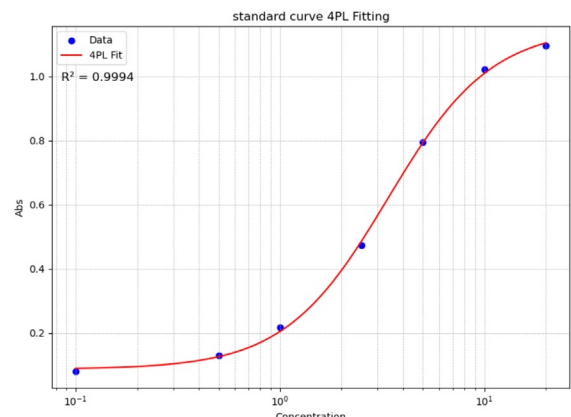


図1. Pythonで作成した4PLフィッティング検量線

3.質量分析結果の多変量解析

質量分析は近年、生体に含まれる全代謝物を対象に解析するメタボロミクス等の網羅的解析に利用されることも多い。非常に多くのパラメーターやサンプル数を有するため、単純な定量計算だけでなく、要素同士の関係を見出すために多変量解析を求められる。

本研究では、まず、生成AIに、サンプル数30、化合物数50のテストデータを作成させた。このデータは対数正規分布に従って生成されている。次に、プログラムを作成し、主成分分析(Principal Component Analysis、以下PCA)及びボルケーノプロットの描画を行った(図2・図3)。最後に、島津製作所から提供いただいたテストデータを用いて解析し、プログラムの結果を、本学に導入されている多変量解析ソフトウェアの結果と比較することで、プログラムが正しく動作したかどうかを検証する。



図 2. Python で作成した PCA

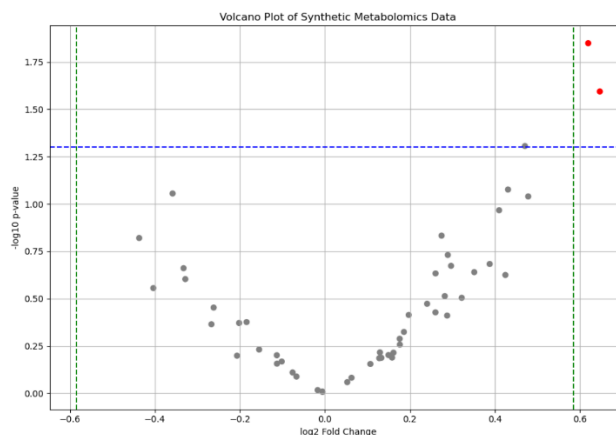


図 3. Python で作成したボルケーノプロット:
優位と判断したプロットが赤く表示されている

4.今後の展望

Pythonによるデータ解析を支援することができるようになれば、ユーザーの研究内容に合わせたオーダーメイドの解析を提供することもできる。また、有料の解析ソフトウェアを購入する必要がなくなれば、ソフトウェアの更新やPC更新等の維持費に悩まされることもなくなる可能性がある。

ただ、現時点では、ローカル環境で動作させているため、Python環境の構築が前提となっている。GoogleColab等のオンラインサービスもあるが、最低限プログラムコードを実行し、結果を理解できる必要がある。今後は、PyScript等を学び、WEBページ上でユーザーが簡単に利用できる環境を提供することを目指したい。

参考

- (1) トミーデジタルバイオロジー株式会社 第51回:今更聞けない、、、ELISAやLEGENDplexの検量曲線フィッティングモデルってどれを使ったら良いの? ?
https://www.digital-biology.co.jp/html/japanese/mail/TDBNews2023_11/BLI.pdf
- (2) MOLECULAR DEVICES SoftMax Pro ソフトウェアでベストcurve fitをセレクト
<https://www.moleculardevices.co.jp/app-note/selecting-best-curve-fit-in-softmax-pro-7-software.html>
- (3) 島津製作所 メタボロミクスをはじめよう
https://www.an.shimadzu.co.jp/sites/an.shimadzu.co.jp/files/pim/pim_document_file/an_jp/brochures/20405/u219-1053.pdf