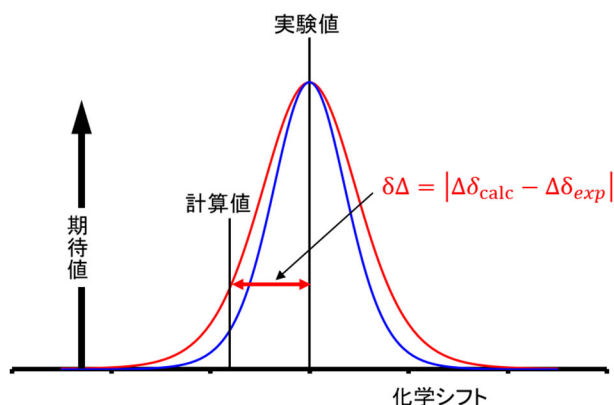


DP4 の考え方(エクセルを活用する)

DP4 は Spartan での Spreadsheet でも計算できますが、例えば分子内エーテルか、ジオールなのかといったような分子式が違う場合や水酸基やケトンの位置異性体といった場合には、現在の Spartan では対応できません。また、DP4 解析に使用されるパラメータが公開されていないため、論文に報告するときに困ります。しかし、DP4 解析はエクセルの機能で十分に対応可能です。

まずは、DP4 の原理から考えてみましょう。正しい構造では、真の値(実測値)と計算値の化学シフトが完全に一致すればよいのですが、どんな計算手法を使用しても必ず差異($\Delta\delta$)があります。正しい構造の場合、分子全体を通して $\Delta\delta$ が小さい計算手法が精度の高いといえます。[精度を示す指標は複数ありますが、多くの場合、標準偏差(STD=偏差二乗平均平方根)を用います。]

次に、真偽不明の異性体の化学シフトを考えます。計算精度にかかわらず $\Delta\delta$ が小さい場合と大きい場合では、小さい場合のほうが「もっともらしい」、すなわち、正しいという確率が高いことになります。正誤二つの異性体を比較する場合、一つの炭素核だけでは、 $\Delta\delta$ によほどの開きがない限り判定が難しいですが、複数の炭素核で比較すると優位なものになります。数学的にはベイズ推定というそうです。以上の考え方を数値化して異性体間の確率を比較する方法が DP4 です。



ある炭素核 (C_i)について、実験値と計算値の $\Delta\delta$ と期待値との関係は以下ようになります。すなわち計算値が標準偏差と比較して、ずれた割合を期待値に変換します。図では、赤色の分布曲線で示しています。計算精度が高い場合は、青い分布曲線になります。すなわち、差異と期待値との関係がより鋭敏になります。差異 $\Delta\delta$ 、標準偏差 σ とすると期待値 P_i は

$$P_i = \left[1 - T^v \left(\frac{|\Delta\delta|}{\sigma} \right) \right]$$

で表されます。 T^v は自由度 v とするスチューデントの T 分布関数です。分布関数としては正規分布関数が一般的ですが、数値予想などでは T 分布関数を用いるのだそうです。正規分布関数を用いても大差ないと思います。Goodman の論文でも正規分布関数でもよいと書かれていますが、計算式がほんの少し変わるだけですので T 分布関数を使用したほうが無難です。

必要なパラメータとしては標準偏差 (σ) と自由度 (v) です。多くの場合、これらを調べることは大変です。Goodman らの DP4 のオリジナル論文にある以下のパラメータを用いることにします。

	^{13}C	^1H
標準偏差 (σ)	2.306 (ppm)	0.185 (ppm)
自由度 (v)	11.38	14.18

Spartan'18 の計算手法は Goodman らの方法より計算精度が高く、Wavefunction 社によると $\sigma = 1.77$ ppm を用いることができるそうです。確かにこの値を用いると DP4 のスコアは鋭敏に変化することになりますが、実質的には影響がないことと、論文などで説明することが面倒なので Goodman らのパラメータを用い、本文、実験の部で「Goodman らのパラメータを用いた」と記述した方が便利だと思います。ちなみに DP4+は基底関数のレベルを高くして、計算結果を経験的な補正を加えることで s の値が小さくなり、スコアが構造に鋭敏に反応するようになったというのみで、考え方は DP4 と全く同じです。

なお、Spartan を用いてデフォルト条件で計算した場合、計算後、汎関数の性質に由来する誤差を経験的パラメータで補正して化学シフトが表示されています。したがって、Goodman の論文で記載されている、実験値と計算値との相関から切片や傾きを求めて補正するという面倒な作業は必要ありません。むしろ補正してはいけないと考えるべきです。補正前のデータは Output ファイルに記載されています。

Spartan で DP4 を計算する具体的方法です。

- ① i 番目炭素核で、 $|\delta\Delta_i|$ を求めます。エクセルで i 番目の炭素核を求める場合、

$$\Delta\delta_i = \text{ABS}(\delta_{\text{calc}} - \delta_{\text{exp}})$$

です。二乗して平方根を求めても同じです。

- ② T 分布関数を用いて各炭素の期待値を求めます。エクセルで Goodman のパラメータを使用する場合、

$$P_i = (1 - \text{T.DIST}(|\delta\Delta_i|/2.308, 11.38, \text{true}))$$

です。Spartan での計算精度を反映させる場合、2.308 の代わりに、1.77 を代入します。自由度はそのまま構わないそうです。プロトンの場合は

$$P_i = (1 - \text{T.DIST}(|\delta\Delta_i|/0.185, 14.18, \text{true}))$$

です。Spartan'18 ではプロトンの化学シフトの徹底的な最適化は不明ですが、Spartan'20 では経験補正が加わったと聞いています。従って、実験値と計算値との相関から切片や傾きを求めて補正するという作業は必要ありません。

- ③ 各炭素の期待値の積を求めます。エクセルでは PRODUCT 関数を用います。

とても小さい数字になりますが、その異性体の期待値となります。

- ④ ①～③の操作で比較するすべての異性体における期待値を求め、異性体間での期待値の比を百分率で求めたものが DP4 です。