●参考資料２　経年分析の方法等に関する補足説明

・対数線形回帰モデル

環境中に残留している化学物質の濃度減少は、1次反応（図1（左）：濃度の高低によらず、ある一定の期間において一定の割合で減少する反応）を仮定するとき、図1（右）に示すように濃度の対数と時間との関係は線形で回帰できるため、対数線形回帰モデルを利用することとした。

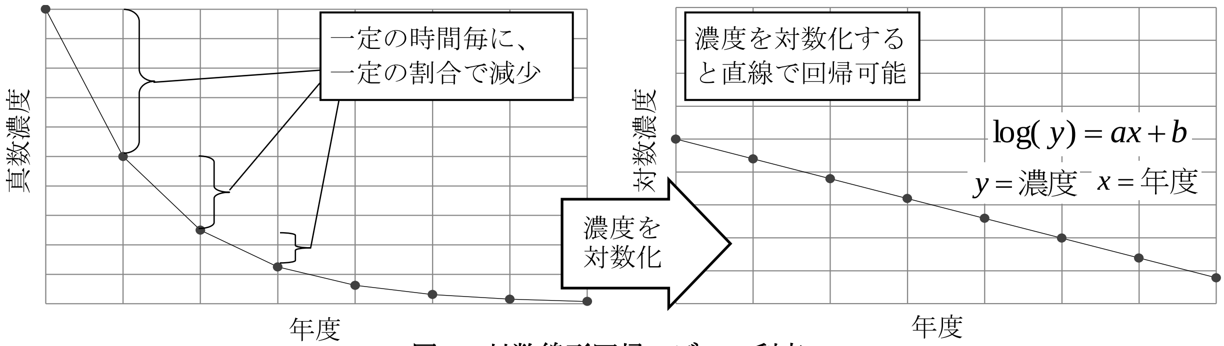


図１　対数線形回帰モデルの利点

環境中に残留している化学物質の濃度分布が、図２（左）に示すような傾向であるとき、図２（右）のように対数正規分布で近似することが出来る。

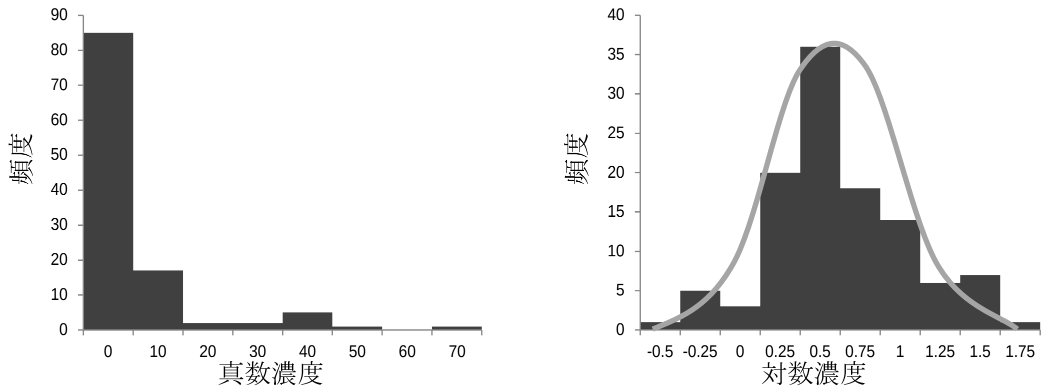


図２　真数及び対数による濃度分布

更に、図3に示すとおり、真数において作成する線形回帰の傾きは時間に対して一定の割合で濃度が減少する場合、低濃度に比べ、より高濃度のデータ変動の影響を受けやすい。しかし、対数濃度では高濃度と低濃度でデータの変動の影響は等価となるため、全体の傾向を一つの傾きで評価できる。

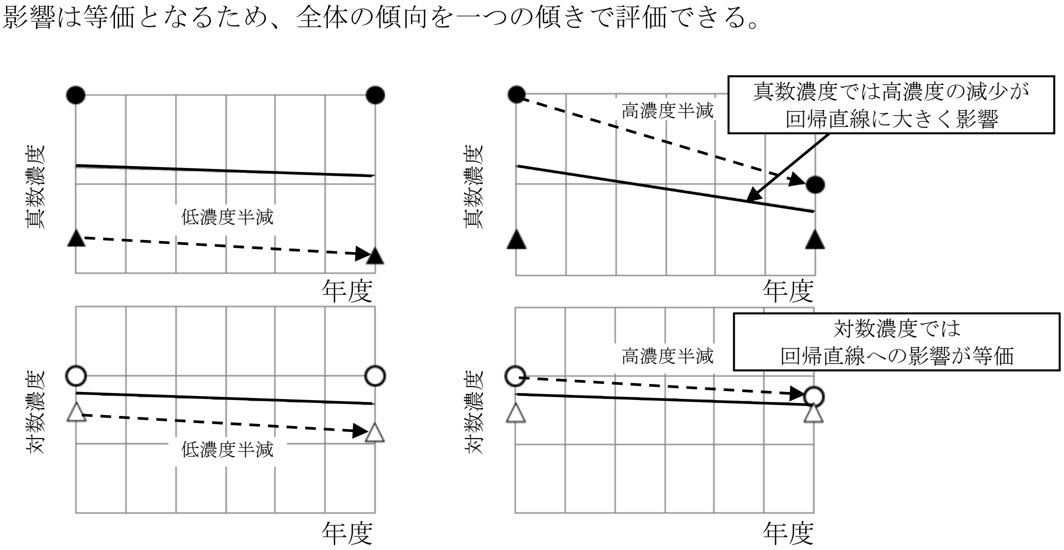


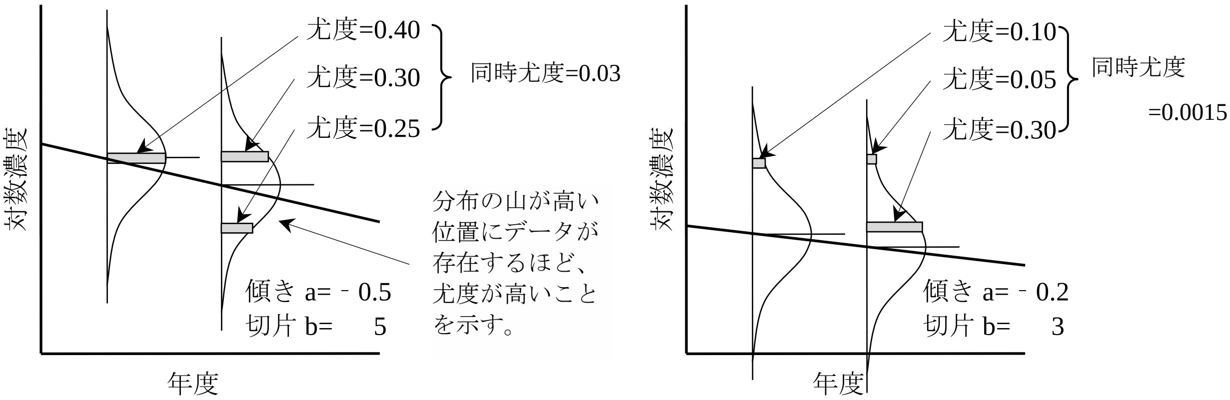
図３　対数線形回帰モデルにおける濃度変動の影響

・実データに基づいた残差分布による最尤推定法

経年変化解析を行うために直線回帰を行う場合には、最小二乗法による回帰直線がよく利用される手法であるが、前提条件として残差分布が正規分布である必要がある。しかし、本手法では、回帰直線を算出する際にパラメトリックな残差分布を仮定せず、正規分布以外のデータについても直線回帰を行うことができる。

最尤推定法とは「最も尤もらしい」パラメータを探索する方法である。回帰直線を算出する場合には傾きa及び切片bの２つのパラメータに対して様々な値を代入し、その結果として算出された回帰直線が「最も尤もらしいとき」のパラメータを最も適した回帰直線とすることとした。

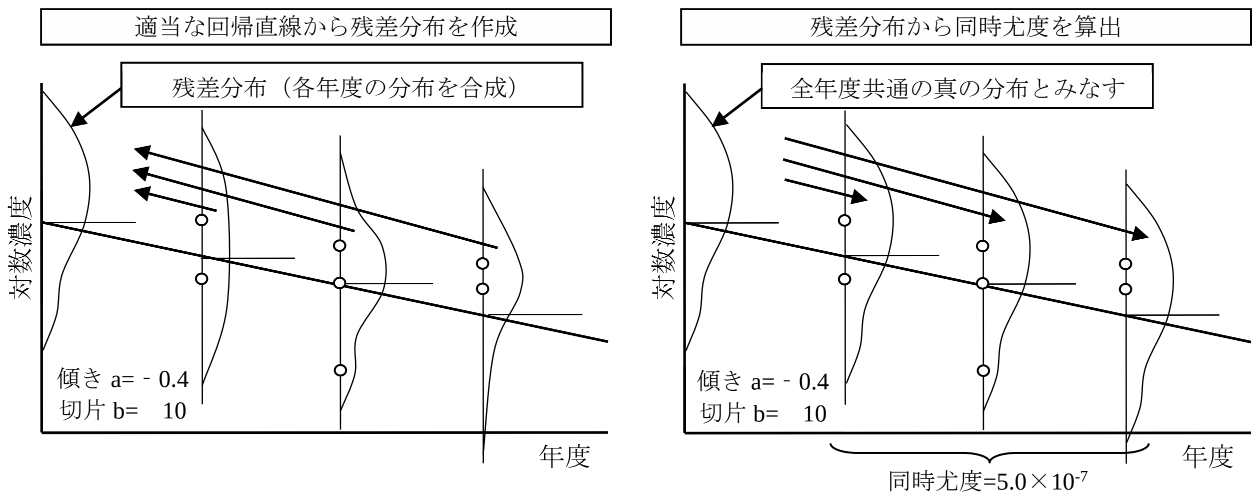
この「最も尤もらしいとき」とは、図4に示すように、回帰直線を算出した際に各データの尤度が最も高くなる事とし、データが複数ある場合には各データの尤度を全て掛け算した値（同時尤度）が最も高くなることとした。また、各データの尤度は、母集団の確率密度分布において、その分布のどの位置にデータが存在するかによって求めることができる。



より同時尤度が高い左図の回帰直線がより適しており、最も尤もらしい回帰直線は、a= - 0.5, b= 5 であると判断する。

図 4　最尤法による最適な回帰直線の決定方法

各解析データはそれぞれで特徴的な分布を持っている場合が多いため、経年変化解析には図5に示すように、回帰直線からの残差で表した各年度の残差分布を作成し、その後足し合わせた各年度共通の残差分布を用いて解析することとした。



例において、適当な回帰直線 a= - 0.4, b= 10 による同時尤度は 5.0×10^{-7} である。
同様に様々な回帰直線で同時尤度を算出し、最も平均尤度の大きい回帰直線を最適な回帰直線とした。

図5　最尤法に用いる残差分布の算出と最適な回帰直線の決定方法

・AIC（赤池情報量規準）

AIC（赤池情報量規準）とは、有効なモデルの選択基準の代表的な指標である。

回帰モデルではパラメータを増やすほどデータに対する誤差は小さくなる傾向にあるが、モデルが複雑となるため、必ずしも良いモデルになるとは限らない。AICは、パラメータ数が増えることを不適として評価を修正する性質を持つことから、パラメータ数を考慮してより良いモデルを把握するための指標となる。また、モデルの母集団の分布に制限もない。これらの理由からAICを用いて最適なパラメータ数のモデルを選択することとした。以下にAICの算出式を示す。

$$AIC = -2 \times \text{最大対数尤度} + 2 \times \text{モデルのパラメータ数}$$

最尤法を用いて求めた回帰直線は、図6に示すように年度をパラメータとする１次式である。この対数線形回帰モデルから計算されるAIC_1と、回帰直線の傾きが偶然の変動によるもので全体を代表する一定値から変動しないと考える0次式（傾き0における対数線形回帰直線モデル）から計算されるAIC_0を比較し、どちらがより適切なモデルであるかを判断した。通常、AICの値の小さいモデルが適切と判断する。更に、AICの差が少ない場合にも安全性を見込んで適切に判断できるよう、ベイズの定理を利用して事後確率の考え方を導入した。

$$p_1 = \exp\{-0.5AIC_1\} / (\exp\{-0.5AIC_0\} + \exp\{-0.5AIC_1\})$$

（p_1:１次のモデルの事後確率、AIC_1:１次式における AIC、AIC_0:０次式における AIC）

１次モデルのAIC事後確率 p_1は０から１の値をとり、１に近い値ほど１次式に近い事を示す。１次式のAIC事後確率 p_1が0.950以上の場合には安全性を見込んだ上で、経年変化において傾きを持つことが適切と判断した。また、0.950のしきい値は危険率5%の考え方を参考に設定することとした。

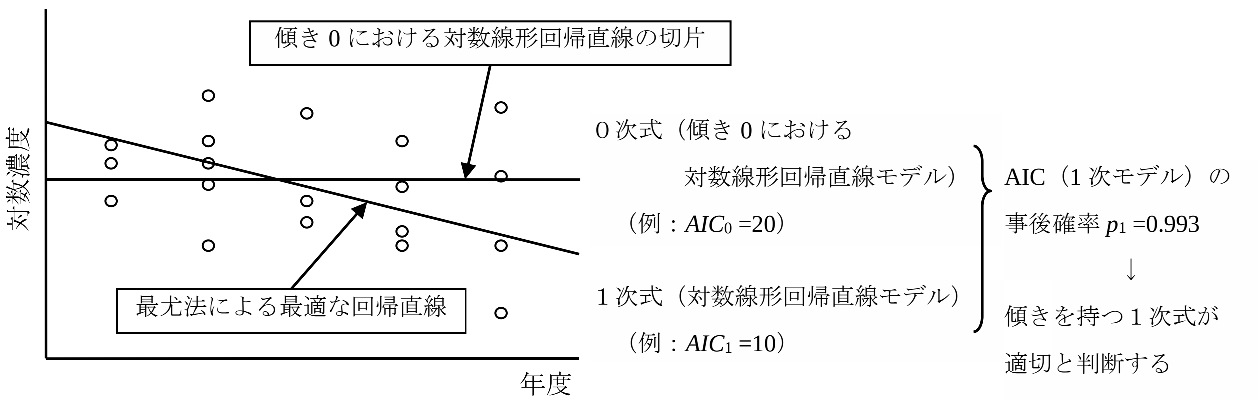


図 6　AIC を利用した傾きの有無の判断方法

・ブートストラップ法による平均値の差の検定

一般的に用いられる t 検定による平均値の差の検定は、前提として正規性が得られている２つのデータ群間を比較する場合に用いる手法である。しかし、ブートストラップ法による平均値の差の検定では、ランダムサンプリングによる繰り返し抽出によって正規性のない母集団に漸近正規性を持たせることが可能なため、平均値を比較する２群の各データがどのような分布であっても平均値の差の検定を行うことが可能となる。

具体的には図7に示すように、前期6か年（A群）と後期6か年（B群）において有意に濃度差があるか確認するために、平均値の差の検定を実施した。2つの標本に対し、それぞれ無作為に抽出した際の平均値を求め、それを繰り返すことにより得られる平均値の分布は t 分布であるが、自由度が極めて大きいことからそれぞれ正規分布とみなすことが出来ることを利用し、標本間で差があるか検定する方法である。

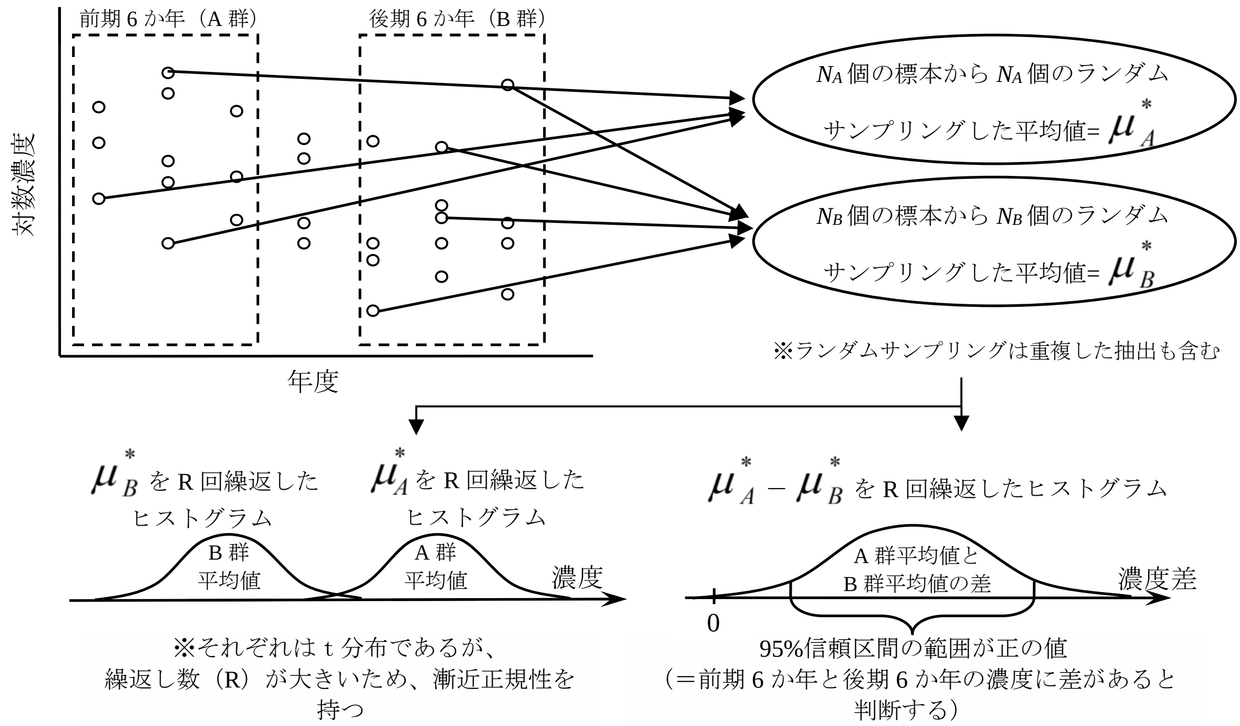


図 7　ブートストラップ法による平均値の差の検定手法

繰り返し抽出して算出した平均値の差の分布において、95%信頼区間が正（負）の範囲にある場合、前期6か年と比較して後期6か年は有意に低（高）値であると判断した。

・検出した検体数の割合による経年変化傾向の解析

解析の対象とした期間における最も高い検出下限値に検出した検体数の割合に着目し、その検出下限値を下回る地点を「低濃度地点」と定義し、その低濃度地点数の増加（または減少）傾向を確認することとした。各物質の調査結果は、検出・不検出のデータとして考え、二項分布を想定した最尤法による回帰直線を算出することとし、同時尤度の値が最も高い際のパラメータa,bを尤もらしい回帰直線とした。なお、安全を見込んだ上で事後確率の考え方を利用し、１次式のAIC事後確率P_1が0.950以上の場合に傾きを持つと判断した。

また、各年度で検出下限値が異なることによる影響を除外するため、調査年度で最も高い検出下限値を統一検出下限値とし、統一した検出下限値における各調査年度の低濃度地点数の割合を算出して解析した。