

Up and down試験計画による感度試験データの最尤法解析

竹山象三*†

*岡山商科大学 〒700-8601 岡山市津島京町2丁目10-1

TEL: 086-252-0642 FAX: 086-255-6947

†Corresponding address: takeyama@po.osu.ac.jp

2010年8月5日受付 2011年10月17日受理

要旨

Up and down試験計画によって得られた感度試験データを最尤法によって解析する方法を提案した。この解析方法を、Microsoft Excelに備わったソルバー機能を利用し、Excelシート上に実現した。試料の臨界刺激量の分布の母平均 μ 、母標準偏差 σ が最尤法によって効果的に推定されることを数値シミュレーションによって確認した。また、 μ 、 σ の信頼領域を評価する方法を開発した。その結果、種々の信頼率の μ 、 σ の信頼領域は、点推定値を中心とした同心円にはならず、 μ 、 σ の信頼限界を別々に評価することの危険性が示された。点推定値とデータセット中の各試験結果との整合性を評価する方法をも提案した。この方法を用い得れば、点推定結果に影響の強い試験結果を絞り込むことができる。

1. はじめに

火薬類の感度試験において、事前の試験結果を「爆」と判定するか「不爆」と判定するかに基づいて、つぎの試験で負荷する刺激量（たとえば、落ついの高さ、静電気エネルギーなど）を調整するUp and down型の試験計画の代表的なものとして、Dixon法¹⁾がある。そして、Dixon法では、試料が「爆」となるのに必要な刺激量（以下、臨界刺激量）の分布の母平均 μ および母標準偏差 σ の推定式が提案されている。

Dixon法の推定式は、臨界刺激量の対数値が正規分布すると仮定して、得られた試験データセットの尤度（ゆうど）を求め、これを最大とする μ 、 σ を近似的に求めるものである。しかし、最近では、パソコンソフトMicrosoft Excelのソルバー機能を用いると、最大尤度を与える μ 、 σ を直接的に求めることができる。

本報告では、Up and down試験計画で得られたデータについて、ソルバー機能を用いて尤度方程式を近似を用いずに解くことにより、臨界刺激量分布の母平均 μ および母標準偏差 σ の精密解を求める方法および得られたデータセット中の各試験の妥当性を評価する方法について提案する。

2. Up and down試験計画と最尤法による母数推定

Up and down試験計画から得られる試験データセットの尤度を求めるに際して設定した刺激-反応モデルをFig. 1に示す。

感度試験に供されるサンプルごとの臨界刺激量 x はばらついており、母平均 μ 、母標準偏差 σ の正規分布に従っていると仮定する（Dixonは、臨界刺激量の対数値が正規

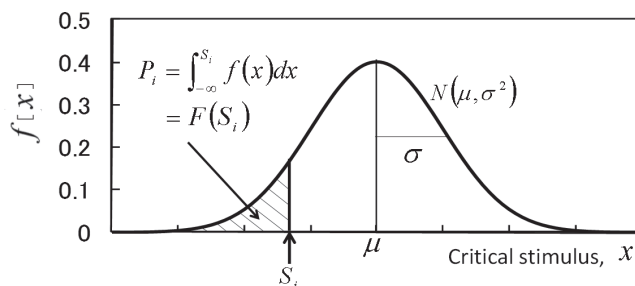


Fig. 1 Stimulus vs. response probability model as basis of sensitivity test results analysis.

分布に従っているとしているが、説明を簡単にするために、このように仮定する)。i回目の試験で負荷される刺激量 S_i に対して、サンプルが「爆」となる確率 P_i は(1)式で求められ、Fig. 1の斜線部の面積で表される。

$$P_i = Pr(x < S_i) = \int_{-\infty}^{S_i} f(x) dx = \int_{-\infty}^{S_i} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx \equiv F(S_i) \quad (1)$$

$f(x)$: 確率密度関数 $F(x)$: 分布関数

一方、i回目の試験結果が、「不爆」となる確率は、(2)式で表される。

$$1 - P_i = 1 - F(S_i) \quad (2)$$

つぎのような3回の試験結果が得られたとする。

- 1 回目: 刺激量 S_1 を負荷して試験を行い、結果が「爆」
- 2 回目: 刺激量 S_2 を負荷して試験を行い、結果が「不爆」
- 3 回目: 刺激量 S_3 を負荷して試験を行い、結果が「爆」

このような3回の試験のデータセットが得られる確率、す

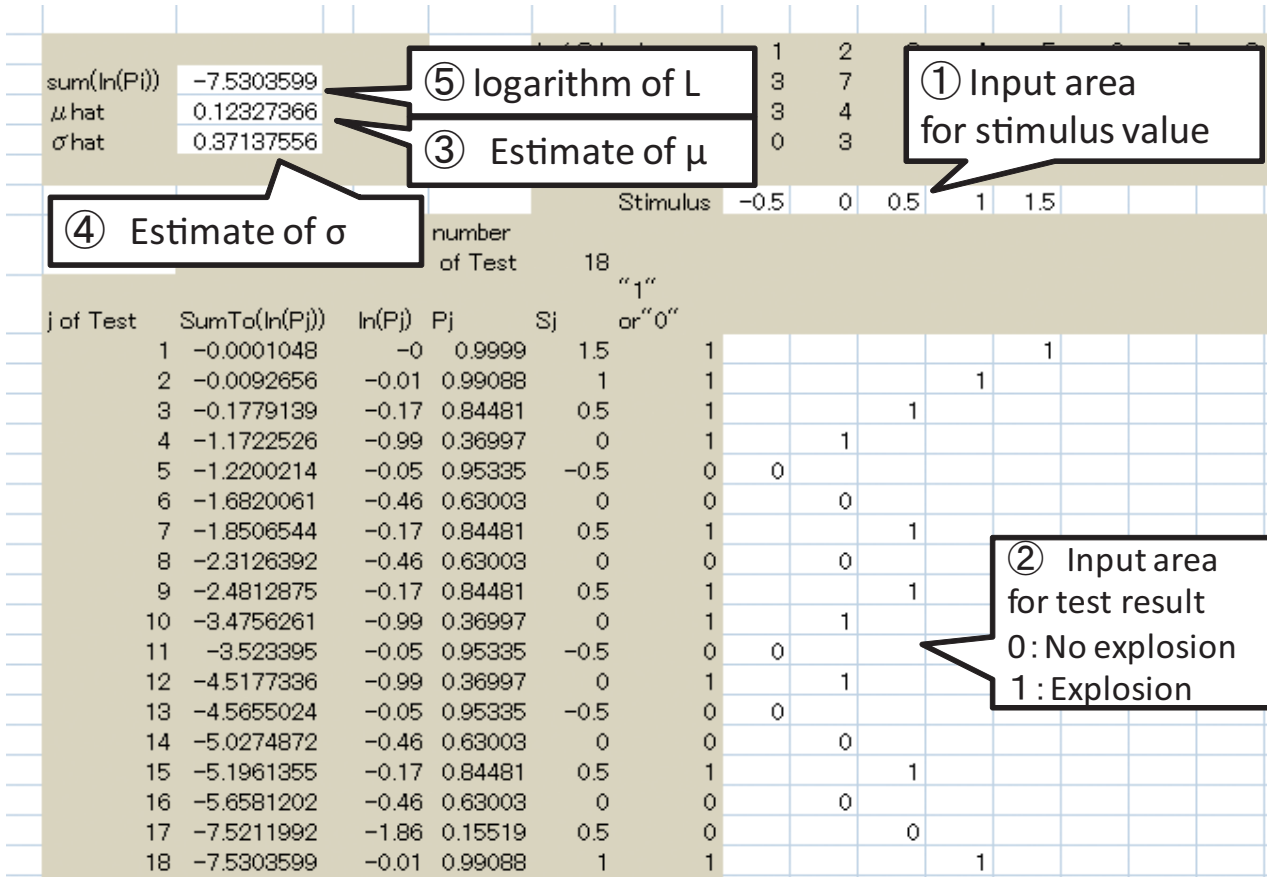


Fig. 2 Brief overview of Excel sheet for maximum likelihood analysis.

なわち、尤度 L は、(3)式で得られる。

$$L = P_1 \cdot (1 - P_2) \cdot P_3 = F(S_1) \cdot (1 - F(S_2)) \cdot F(S_3)$$

$$= \left\{ \int_{-\infty}^{S_1} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx \right\}$$

$$\left\{ 1 - \int_{-\infty}^{S_2} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx \right\}$$

$$\left\{ \int_{-\infty}^{S_3} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx \right\} \quad (3)$$

最大の尤度 L が得られる(3)式の μ , σ を求め、 μ , σ の推定値 $\hat{\mu}$, $\hat{\sigma}$ とするのが、最尤推定法(最尤法)である。Excelのソルバー機能を用いると尤度 L を最大にする μ , σ を簡単に求めることができる。本報告では、この最尤法による推定をExcelのソルバー機能を使用してExcelシート上に実現した。Fig. 2にその概要を示し、その利用の手順を以下に示す。

- ①の領域に試験で負荷した刺激量の値を入力する。
- ③④のセルに μ , σ の推定の初期値となる適当な値を入力する。
- ②の領域に上から下へを試験順序として試験結果を入力する。このとき、③④のセルの値が μ , σ である場合の各試験結果が得られる確率が P_i の列に、また、入力された試験結果全体の尤度 L の対数値が⑤のセルに表示される。
- 「目的セル」を⑤のセル、「目標値」を最大値、「変化させるセル」を③、④のセル、「制約条件」を「④のセルの値 > 0 」として、ソルバーを実行する。

- ⑤のセル(尤度の対数値)が最大となる μ の値, σ の値が③, ④のセルに表示される。

3. 最尤推定値の有効性の確認

最尤推定値の有効性を確認するために、つぎのようなUp and down試験計画のシミュレーションを行って試験データセットを得、これに最尤法を適用した。

- ① 臨界刺激量 x のデータとして、 $\mu = 0.0$, $\sigma = 1.0$ の正規分布に従う乱数を10000個用意する。
- ② 負荷する刺激量をTable 1のように設定する。
- ③ 1回目 ($i = 1$) の試験で負荷する刺激量 S_1 に $E_j (= E_{12})$ を選択する。
- ④ 臨界刺激量 x_i をランダムに1個取り出し、 i 回目の試験で選択されている刺激量 S_i と比較し、
 - ・ $x_i < S_i$ のとき、「爆」と判定し、つぎの試験で負荷する刺激量 S_{i+1} に1段低い刺激量 E_{j-1} を選択する。
 - ・ $x_i > S_i$ のとき、「不爆」と判定し、つぎの試験で負荷する刺激量 S_{i+1} に1段高い刺激量 E_{j+1} を選択する。
- ⑤ ④を60回繰り返し、1セットのUp and down試験データセットとする。その一例をFig. 3に示す。
- ⑥ この結果にソルバー機能を利用した最尤法を適用し、 μ , σ を推定する。
- ⑦ ③~⑥を1000セット繰り返す。

各セットのデータから推定された $\hat{\mu}$, $\hat{\sigma}$ の分布をFig. 4に示す。

Table 1 Loaded stimulus in Up and down test simulation.

Stimulus E_i	E_0	E_1	E_2	E_3	E_4	E_5	E_6	E_7	E_8	E_9	E_{10}	E_{11}	E_{12}
Preset value	-3.0	-2.5	-2.0	-1.5	-1.0	-0.5	0.0	0.5	1.0	1.5	2.0	2.5	3.0

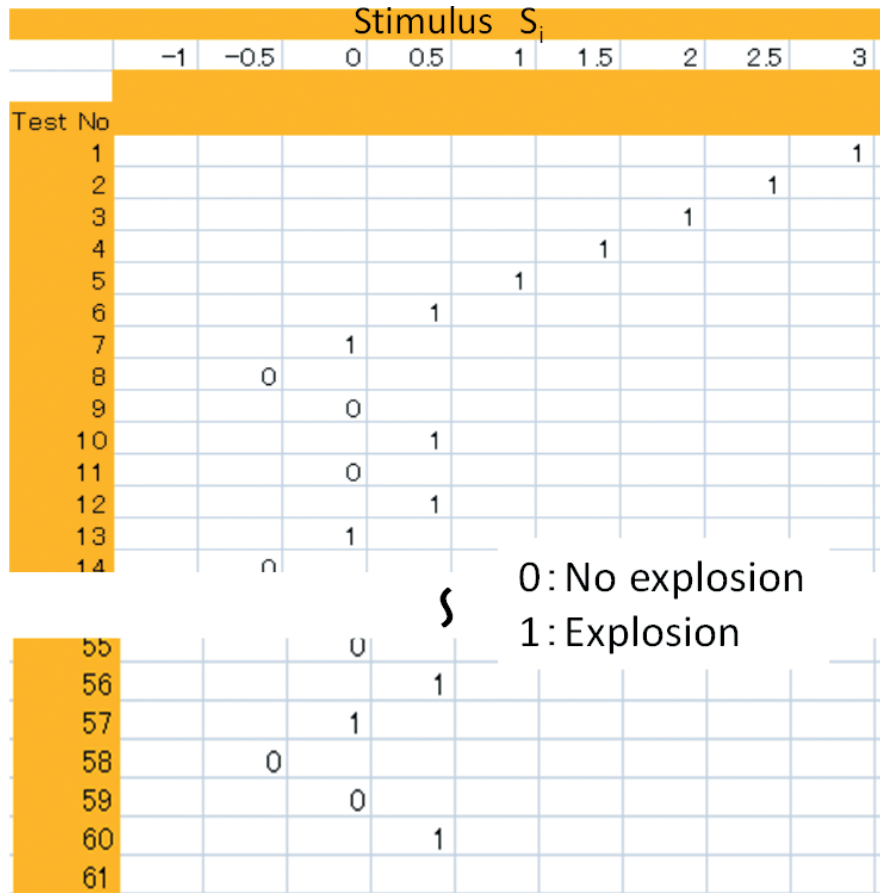


Fig. 3 An example of Up and down test simulation.

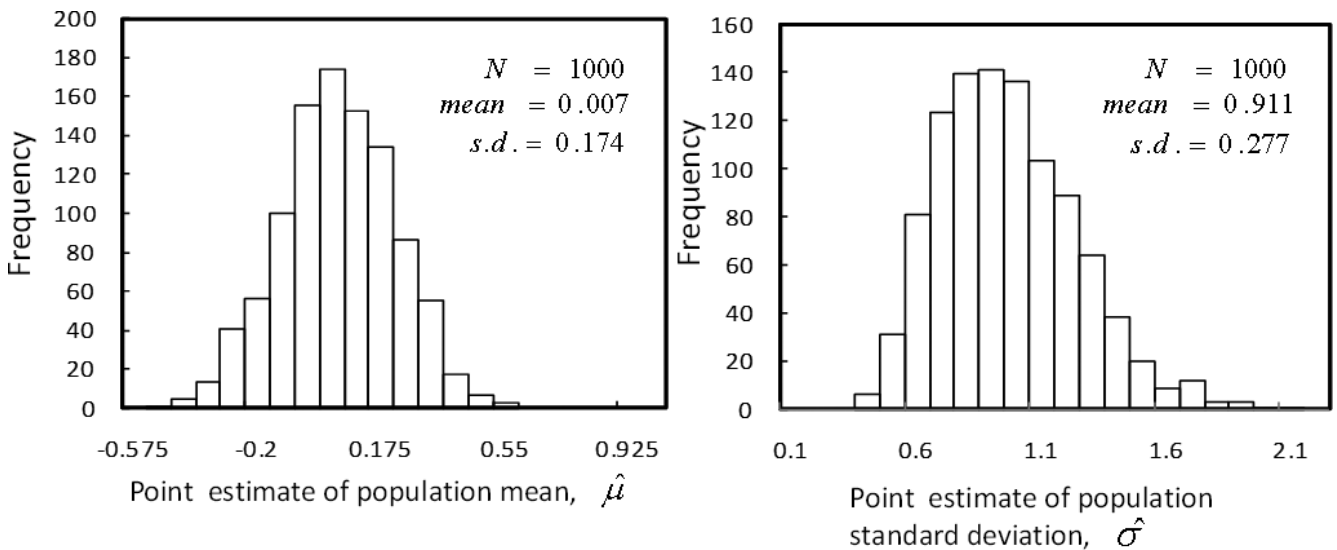


Fig. 4 Distributions of maximum likelihood point estimate.

μ の推定値 $\hat{\mu}$ は、シミュレーションで設定されている $\mu = 0.0$ を中心として、ほぼ左右対称な分布となっている。また、 σ の推定値 $\hat{\sigma}$ は、左右対称の分布ではなく、右に裾引いた分布になっているが、その平均値は 0.911 であり、シミュレーションで設定した $\sigma = 1.0$ とほぼ等しい値が得

られている。これらのことより、臨界刺激量の母平均 μ 、母標準偏差 σ の推定に最尤法は有効であると判断できる。

4. 最尤推定値の信頼領域

最尤推定値について、つぎのことが知られている²⁾³⁾。

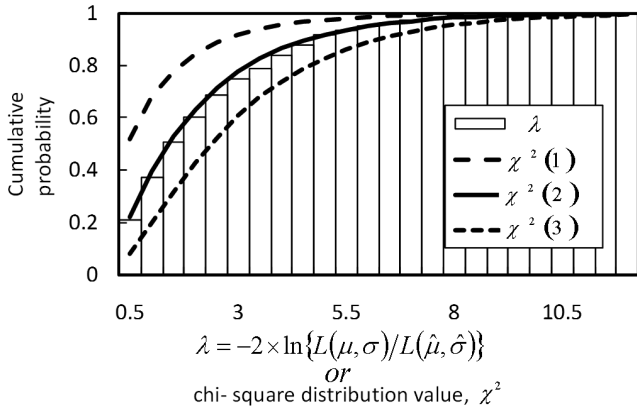


Fig. 5 Comparison of cumulative probability between λ and χ^2 .

- ① 臨界刺激量の母集団の母数（本報告では、 μ および σ ）を用いて求めた試験データセットの尤度を $L(\mu, \sigma)$ とする。
- ② その母集団から得られる試験データセットに最尤法を適用して得られる推定値（本報告では、 $\hat{\mu}$ および $\hat{\sigma}$ ）を得た最大尤度を $L(\hat{\mu}, \hat{\sigma})$ とする。
- ③ $\lambda = -2 \times \ln\{L(\mu, \sigma) / L(\hat{\mu}, \hat{\sigma})\}$
 $= -2 \times [\ln\{L(\mu, \sigma)\} - \ln\{L(\hat{\mu}, \hat{\sigma})\}]$
 としたとき、 λ は χ^2 分布に従う。（尤度の対数値を対数尤度と呼ぶ）

このことが、本研究で対象としているUp and down試験計画から得られる最尤推定値にも適用できることを確認するために、3.のシミュレーションで得た1000組のデータセットごとに求めた λ を自由度が1, 2, 3の χ^2 分布と累積確率で比較し、Fig. 5に示した。

自由度が1および3の χ^2 分布の累積確率の曲線は、 λ の累積比率へのあてはまりが悪いが、自由度が2のとき、非常にあてはまりの良いことがわかる。したがって、①～③のことは、本研究で対象としているUp and down試験計

画から得られる最尤推定値にも適用できると考え、本研究では、自由度2の χ^2 分布を用いて、本節で述べる最尤推定値の信頼限界を求めることにする。

λ が自由度2の χ^2 分布に従うとき、 $1-\alpha$ の確率で得られる λ の値の範囲は、 χ^2 分布の上側100 α %点である $\chi^2(2, \alpha)$ より小さい範囲となり、(4)式が得られる。なお、 λ が小さい側については、推定値である $\hat{\mu}$, $\hat{\sigma}$ が母数 μ , σ に近づくのであるから、 λ の小さい側には制限を設定せず、 α のすべてを χ^2 分布の上側にとった。

$$Pr(\lambda = -2 \times [\ln\{L(\mu, \sigma)\} - \ln\{L(\hat{\mu}, \hat{\sigma})\}] < \chi^2(2, \alpha)) = 1 - \alpha \quad (4)$$

(4)式を利用すると、各種の母数 μ , σ の値の組み合わせにおける推定値 $\hat{\mu}$, $\hat{\sigma}$ が得られる確率 $1-\alpha$ が計算できる。たとえば、母数を $\mu = 0.3$, $\sigma = 1.2$ と仮定した場合に（ただし、このデータセットを得るときに仮定した μ , σ は0.0, 1.0である。）、Fig. 3で示したデータセットが得られる確率を求めることができる。Fig. 3のデータセットでは、 $\hat{\mu} = -0.162$, $\hat{\sigma} = 1.015$ と推定されたとき、最大対数尤度 $\ln\{L(\hat{\mu}, \hat{\sigma})\} = -33.4167$ が得られた。一方、このデータセットに $\mu = 0.3$, $\sigma = 1.2$ を仮定すると、対数尤度は、 $\ln\{L(\mu, \sigma)\} = -36.1280$ となる。よって、 λ は、つぎのように求められる。

$$\lambda = -2 \times [\ln\{L(\mu, \sigma)\} - \ln\{L(\hat{\mu}, \hat{\sigma})\}] = -2 \times \{(-36.1280) - (-33.4167)\} \approx 5.4226$$

母数が、 $\mu = 0.3$, $\sigma = 1.2$ であった場合に、5.4226以下の λ の値が得られる確率は、 χ^2 分布を利用してつぎのように得られる。

$$Pr(\lambda \leq \chi^2(2, \alpha)) = Pr(5.4226 \leq \chi^2(2, 0.066)) = 1 - 0.066 = 0.934$$

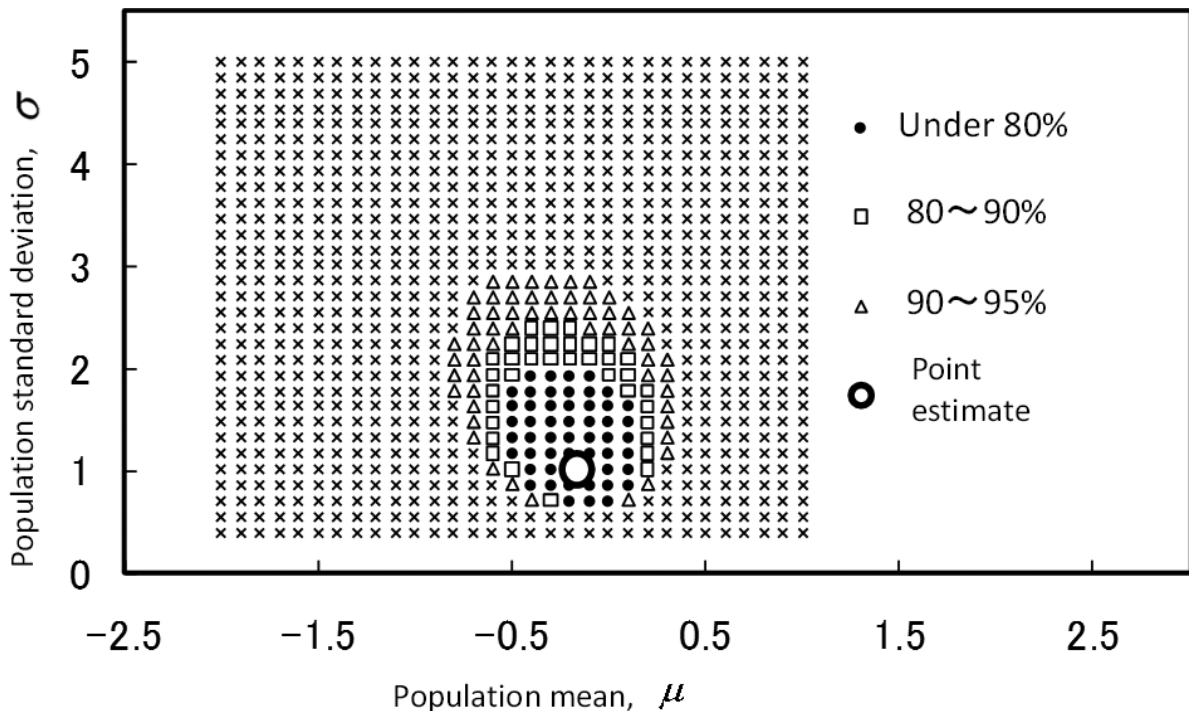


Fig. 6 Confidence area of population mean μ and population standard deviation σ .

すなわち、母数が $\mu = 0.3$, $\sigma = 1.2$ のときでも、 $\hat{\mu} = -0.162$, $\hat{\sigma} = 1.015$ と推定される確率が最大で 93.4% はあると計算される（確率 $1-\alpha$ は 1 つの母数の区間推定の信頼率と呼ばれるものに相当し、以下、信頼率と呼ぶ）。

Fig. 3 のデータセットについて、信頼率が 80%, 90%, 95% のそれぞれの値以下となる μ , σ の値の組み合わせを Fig. 6 に示す。各信頼率の領域は、点推定値 $\hat{\mu} = -0.162$, $\hat{\sigma} = 1.015$ を中心に広がっている。

そして、信頼率を高く設定すると排除される μ , σ の値の組み合わせが減少するので、その領域は広がっている。しかし、その広がり形状は、点推定値からみると、 σ の大きい方向、 μ のやや小さい方向へ幅広くなっており、点推定値を中心とした同心円ではない。これは、 μ , σ の区間推定を行う場合、 μ , σ の信頼限界を単独に求めるのではなく、Fig. 6 のように、 μ , σ の組み合わせ信頼限界として推定する必要性を示している。

Fig. 6 で・印と□印の境界で示される信頼率 80% の信頼限界に着目してみる。信頼率が 80% の信頼限界は、 μ については $-0.50 \sim 0.10$, σ については $0.71 \sim 1.93$ の範囲で歪んだ円形となっている。これは、この信頼限界の領域内に 80% の確率で、真の $\mu (= 0.0)$, $\sigma (= 1.0)$ の値の組み合わせが存在しているということを意味する。このことは、この組み合わせ信頼限界の領域内のどの μ , σ の値の組み合わせであっても、最大 80% の信頼率で、真の μ , σ の値の組み合わせの候補となることを意味している。したがって、信頼限界の領域の広さは、点推定値が真値からどれだけ離れた値を示す可能性があるのかを表す指標となる。信頼限界の領域が広いと、得られた点推定値の精度が低いといえ、逆に、狭いと点推定値の精度が高いといえる。化学物質の感度を比較する場合、点推定値の比較だけでなく、信頼限界の領域の広さを比較することにより点推定値の精度についても考慮することも重要である。

本研究で作成した Excel シートでは、Fig. 6 で示したような図出力を備えており、信頼領域をも考慮して、種々の化学物質の感度を比較することを可能としている。

5. データセットの診断

Dixon 型の Up and down 試験計画における推定式では、データセットおよび各データと、そこから推定された $\hat{\mu}$, $\hat{\sigma}$ の値との整合性を検討する手段が用意されていない。そのためか、これまでに発表された論文では、推定結果のみが提示されている場合が多い。Up and down 試験計画から得られる試験結果は、①刺激量負荷、②レスポンスの確認、③レスポンスに応じた次回試験での負荷刺激量の選択、を①～③を繰り返した結果として得られるデータセットが解析対象となる。したがって、この一連のデータの中に、「爆」または「不爆」が異常に長く続くなど異常なデータが含まれれば後続する試験の流れが変化し、それはデータセットの解析結果へ影響する。したがって、データセットのなかに異常なデータが含まれていないかを確認することは、データセットから推定される $\hat{\mu}$, $\hat{\sigma}$ の値を吟味する上で重要である。

以上の考えから、データセット中の各試験結果とそこから推定された $\hat{\mu}$, $\hat{\sigma}$ の値との整合性を検討する図として Fig. 7 を示す。

Fig. 7 は、最尤法によって推定された $\hat{\mu}$, $\hat{\sigma}$ のもとで、データセット内の i 回目の試験結果が得られる確率 P_i と i 回目の試験結果までの対数尤度の推移図である。 P_i を確認することにより、その試験結果が、 $\hat{\mu}$, $\hat{\sigma}$ のもとに得られる可能性を確率としてチェックすることができる。また、尤度の計算では、試験を繰り返すごとに、その試験結果が得られる確率（1 以下）が乗じられるので、試験が繰り返されるたびに、尤度および対数尤度は小さい値になってゆく。したがって、起こりそうにもない結果、起こる確率の低い結果が得られると、 i 回目の試験結果までの対数尤度は、激しく低下する推移を示す。

たとえば、矢印でマークされた 38 回目の試験に着目してみる。試験結果は「爆」であるが、「爆」となる確率は、0.067 であり、比較的起こりにくい試験結果であると判定される（一般には、起こらないと判断する確率として 5% や 1% が用いられる。したがって、特に異常な結果というわけではないが、例示のため、この試験結果に着目する。）。また、対数尤度推移の観点から見ても、対数尤度の低下が、この

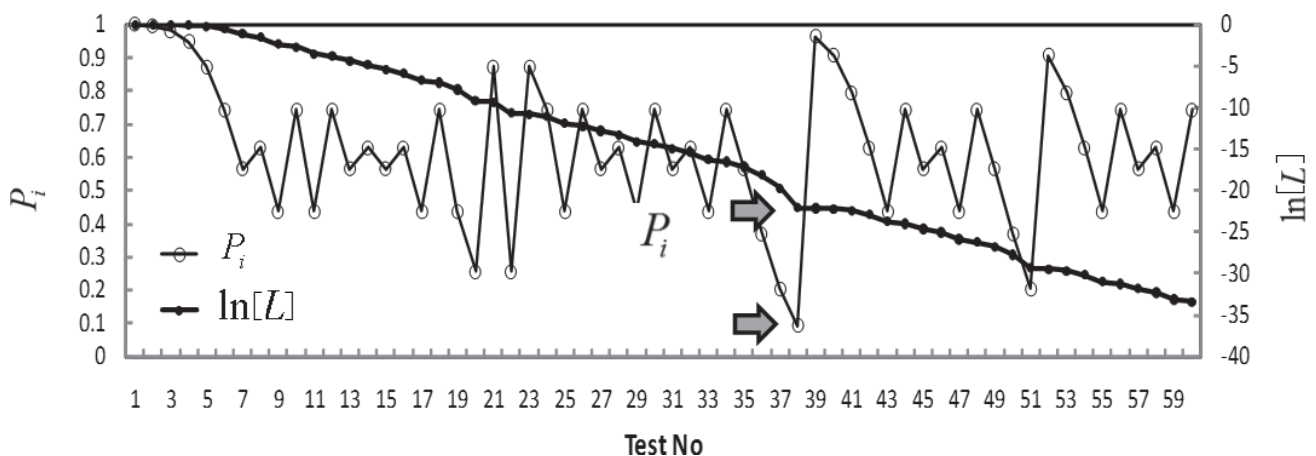


Fig. 7 Profile of P_i and $\ln(L)$.
 P_i : Probability of No.i test results observed under the population mean μ and population standard deviation σ .
 $\ln(L)$: Logarithm natural of the likelihood from No.1 test to No.i test.

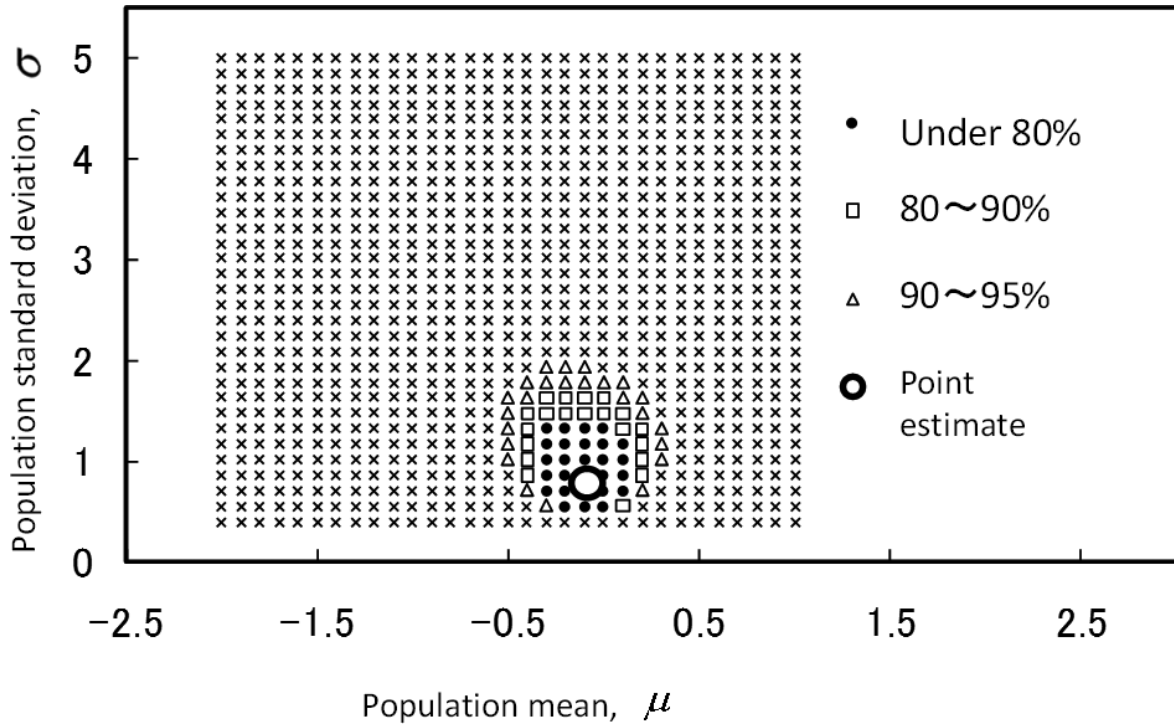


Fig. 8 Confidence area of population mean μ and population standard deviation σ in the absence of No.38 test result.

試験結果により顕著になっている。このことから、他の回の試験結果に比べて、この38回目の試験結果は、解析結果に影響の大きい試験結果と推測される。

そこで、38回目の試験結果を除外した計59回のデータセットを解析することを考えてみる。Dixon法では、データセットの中に、正規分布から異常に逸脱する試験結果、「爆」や「不爆」が異常に長く続くなど異常な試験結果が発見されても、その試験結果を削除することはできない。なぜなら、Dixonの推定式の導出の前提であるUp and down試験手順の制限を逸脱することになるからである。その点、本報告の解析では、尤度が最大となる μ 、 σ を何の制約も設けずに求めているので、任意の試験結果を削除しても解析は可能である。

38回目の試験結果を除外した計59回のデータセットからの推定結果をFig. 8に示す。

Fig. 6にくらべて信頼領域が狭くなったことから、38回目の試験結果が、推定の精度を非常に下げていることがわかる。

以上のように、 P_i および対数尤度推移を吟味することにより、推定された $\hat{\mu}$ 、 $\hat{\sigma}$ に対するデータセット中の試験結果の妥当性を評価でき、さらには、推定結果へ影響の大きい試験結果を抽出できるといえる。

6. おわり

Up and down試験計画により得られるデータセットから、最尤法により臨界刺激量の母平均 μ 、母標準偏差 σ を推定することの有効性が確認された。また、 μ 、 σ の値の組み合わせの信頼領域をも求めることができた。これにより、信頼率をも考慮した化学物質の感度比較が可能となる。そして、これらのことの実行を、Excelシート上にExcelのソルバー機能を使用することにより実現した。

なお、本研究の解析では、尤度方程式をそのまま解いているので、Dixonの推定式を使用するために設けられたUp and down試験を計画する上での制約（負荷する刺激量の間隔を等しくすること、3.④の規則により次回の試験で負荷する刺激量を選択することなど）を設ける必要がない。したがって、試験計画の実施後、途中の試験結果を削除しても解析は可能であることは、注目される点である。

文献

- 1) W. J. Dixon, A. M. Mood, J. Am. Stat. Assoc. 43, 109 (1948)
- 2) 芳賀敏郎, 「最小2乗法, 最尤法」, <http://stat.sci.kagoshima-u.ac.jp/~cse/work/2004/haga/ml/ml.pdf> (accessed 3-October-2011)(online)
- 3) 稲垣宣生, 「数理統計学」 p. 272 (2003), 裳華房

Maximum likelihood analysis of sensitivity test data according to Up and down test plan

Shozo Takeyama^{*†}

Maximum likelihood analysis method of sensitivity test data according to Up and down test plan is proposed. That method is designed in Microsoft Excel sheet utilizing the Solver function. Maximum likelihood point estimates of population mean μ and population standard deviation σ of the critical stimulus distribution of samples are made certain of good estimates by numerical simulation. And confidence area of μ, σ are also derived. As results, the shapes of confidence area of μ, σ with various confidence levels are not concentric circles. Therefore it is not recommended to evaluate confidence limit of μ, σ independently. Consistency check method between the estimates $\hat{\mu}, \hat{\sigma}$ and the each test results in Up and down data set is also proposed. By this method, the impact test result on estimation of parameters is can be found.

Keywords : sensitivity test data analysis, maximum likelihood analysis, up and down test plan

*Okayama Shoka University,
2-10-1, Tsushimakyomachi, Okayama-shi, 700-8601, JAPAN
TEL : +081-86-252-0642

† Corresponding address : takeyama@po.osu.ac.jp