全系列法データの最尤法による分析例

全系列法データの最尤法による分析例

岡　本　安　晴[1]

は　じ　め　に

100 g の肉と101 g の肉の重さの差は分からないが、100 g と200 g なら分かるというように、感覚の強さの差が弁別されるためにはその差がある程度以上必要である。この差が認められるのに必要な最小限の差の値は閾値（threshold）と呼ばれている。本行動科学科の心理学コースで開講されている初級者用実験において閾値をテーマとするものの1つとして「重さの弁別閾」（心理学実験（I）マニュアル、1990）がある。筆者は本学に着任以来この実験のインストラクターを拝命しているが、この場合の閾値は全系列法（method of complete series）データから算出される。加減乗除の算術演算のみで求まるので初級者には適した方法であるが、閾値の意味の定義が操作主義的であるため、感覚のメカニズムにおける意味付けが不明確である。感覚のメカニズムにおける意味付けを与えるためにはモデルを構成してその中で（弁別）閾の機能が設定されなければならない。

今、2つの刺激 S_1 、S_2 について考える。S_1 及び S_2 が惹起する感覚量を X_1 、X_2 とし X_1 、X_2 はある確率密度分布に従って変動するものとする。このとき閾値の古典的な定義は次のように与えられる（Laming, 1986, p. 18）。

X_1 と X_2 の差の弁別に必要な最小限の感覚量の差（閾値）を ΔX とおく（但し、Laming (1986) では Δ^+X と Δ^-X と正負で異なる値をとることが許されている）。

$$|X_1 - X_2| > \Delta X$$

であれば S_1 と S_2 の感覚の強さに差があると知覚されるが、

$$|X_1 - X_2| < \Delta X$$

であれば S_1 と S_2 により惹起される感覚量に違いがあるとは知覚されない。

しかし、この上のタイプのモデルでは、判断（judgment）の要因が無視されている。判断と感覚を区別する信号検出理論（Signal Detection Theory ; Green and Swets, 1966）による分析が今日広く行われているが、特に本稿で扱われるような「区別できません」等の評定を許す3カテゴリーのデータの分析においては判断の要因を無視することは不適当である。ここでは、信号検出理論の考え方に従ったモデルを設定して閾値を求めることにする。

1）　電子メール宛先：C00279@sinet. ad. jp

モ　デ　ル

全系列法（心理学実験（I）マニュアル、1990）においては、標準刺激S_Sと（2N＋1）個の比較刺激S_{-N}、…、S_0、…、S_Nが用意される。比較刺激のうちS_0はS_Sと物理的に同じものであり、隣り合うS_iとS_{i+1}との物理的な差は一定である。S_S及びS_iにより惹起される感覚量をX_S及びX_iとおき、X_S及びX_iはある確率分布に従う確率変数であると考える。いま、閾値を求めることが問題であるので、感覚連続体上において各S_iとS_Sの感覚量は狭い範囲内の隣接した位置に対応するようにとられていると考えられる。故に、X_SとX_iの分散の値により尺度の単位を適当に定めれば、S_S及びS_iの感覚量の平均値μ_S及びμ_iの数値をそれぞれの刺激の物理量に等しくとることができる。

刺激の比較の場合、順位誤差と呼ばれる効果を考慮する必要がある。この現象はLaming（1986, p. 23）に依ればFechner（1860）が最初に発見したものである。本稿で扱う全系列法による重りの比較では、先に挙錘されたものか後の方なのかという時間順位誤差（Time-order error）と、重りが左に置いてあったか右の方であったかという空間順位誤差（Space-order error）の両方が複合されているが、この効果を以下のようにモデル化する。

対の比較において、先に提示される方の感覚量の平均値は物理量の数値に等しくとり、後に提示される方の感覚量の平均値を物理量の数値よりτだけ大きくとる。つまり、先に提示されるより後に提示される方が感覚量はτだけ大きくなるとする。τは正負いづれの値でもよい定数である。

以上の仮定のもとで、標準刺激S_S、比較刺激S_iの順で重りが提示されたときにはS_iとS_Sの感覚量の差$X_i - X_S$は平均$\mu = (\mu_i + \tau) - \mu_S$の分布に従い、$S_i$、$S_S$の順のときには$S_i$と$S_S$の感覚量の差$X_i - X_S$は平均$\mu = \mu_i - (\mu_S + \tau)$の分布に従うことになる。差の分布の形は$\mu$をパラメータとする次の累積確率分布関数$\phi(x; \mu)$で与える。

$$\phi(x; \mu) = \phi_0(x - \mu)$$

ここで、$\phi_0(x)$は$x \to +\infty\ (-\infty)$のとき$\phi_0(x) \to 1\ (0)$なる単調増加関数である。

通常の一対比較法データの場合は、$\phi_0(x)$の具体的な形の細かい差異は実証的に区別し難い（岡本、1992）ので、本稿では計算速度を上げるため

$$\phi_0(x) = \frac{1}{1 + e^{-s \cdot x}}, \qquad s > 0$$

の形（logistic）を採用する。sが大きいほど感覚量の変動が小さくなる。

感覚量と反応との関係は信号検出理論に従って反応のカテゴーリーの基準値Cを設定し、感覚量の差がCより大きいとき「より重い」、$-C$より小さいとき「より軽い」、Cと$-C$の間にあるときには「同じ」（あるいは「区別できない」等）の反応がなされるも

のとする。$C(>0)$ の値は被験者内において選ばれるもので、性格や実験状況等の影響を受けると考えられる。

以上のことより、各反応の確率は以下のように与えられる。

1） 標準刺激 S_s → 比較刺激 S_i の順の場合。

P（S_i の方が重い） $= 1 - \phi(C;\ \mu_i - \mu_s + \tau)$

P（S_i と S_s は同じ重さ） $= \phi(C;\ \mu_i - \mu_s - \tau) - \phi(-C;\ \mu_i - \mu_s + \tau)$

P（S_i の方が軽い） $= \phi(-C;\ \mu_i - \mu_s + \tau)$

2） 比較刺激 S_i → 標準刺激 S_s の順の場合。

P（S_i の方が重い） $= 1 - \phi(C;\ \mu_i - \mu_s - \tau)$

P（S_i と S_s は同じ重さ） $= \phi(C;\ \mu_i - \mu_s - \tau) - \phi(-C;\ \mu_i - \mu_s - \tau)$

P（S_i の方が軽い） $= \phi(-C;\ \mu_i - \mu_s - \tau)$

標準刺激 S_s が第 j 順位（$j=1$（S_s が先）、2（S_s が後））のもとで与えられたときの反応が r であった頻度を $n(i,\ j,\ r)$ で表す。但し、

$$r = \begin{cases} h: S_i \text{の方が} S_s \text{より「重い」と評定} \\ e: S_i \text{と} S_s \text{は「同じ重さ」と評定} \\ l: S_i \text{の方が} S_s \text{「軽い」と評定} \end{cases}$$

と表記する。

このとき、尤度 L を次式で与える。

$$L = (\text{定数}) \cdot \prod_{\substack{i=-N,\ \cdots,\ N \\ j=1,\ 2 \\ r=h,\ e,\ l}} p(i,\ j,\ r)^{n(i,\ j,\ r)}$$

但し、

$p(i,\ j,\ r) = P$（S_i が与えられたときの反応が r である $|$ S_s が第 j 順位）

M 人の被験者のデータをまとめて扱うときには、M 人分のデータに対する尤度 L_{global} を考える。個々の被験者 k の尤度を L_k で表すとき、L_{global} を次式で与える。

$$L_{global} = \prod_{k=1}^{N} \mathrm{L}_k$$

AIC は定義より

$$\mathrm{AIC} = -2 \cdot max\ \{log L\} + 2 \cdot (\text{自由パラメータの数})$$

AIC を与える上式において M 人分のデータに対しては L を L_{global} でおきかえる。AIC の値を求めることにより種々の仮説の検討が可能となる。

さて、閾値については、信号検出理論に従ったモデルにおいては次のように与えられる。

反応のカテゴリー境界を与える値 C は個々の被験者内において実験条件等に応じて設

定されるもので感覚の弁別力とは別のものであり、判断の基準を与える認知的なものである。それに対して、感覚の弁別力はその感覚量の分布の分散に対応して考えられるべきものである。弁別の閾値δをS_sに対して75%の確率でより大きな感覚量を与える比較刺激S_δとの差

$$\delta = S_\delta - S_s$$

で定義する (cf. Falmagne, 1986)。

すなわち、S_δに対応する感覚量X_δの平均値をμ_δとおけば

$$\begin{aligned} 0.75 &= P(X_\delta - X_s > 0) \\ &= 1 - \phi(0\,;\,\mu_\delta - \mu_s) \\ &= 1 - \phi(0\,;\,\delta) \end{aligned}$$

である。尺度値が$\mu_s = S_s$、$\mu_\delta = S_\delta$となるように設定されていたことに注意。又、順序誤差τの影響は除いて考える。

$$\phi_0(x) = \frac{1}{1 + e^{-s \cdot x}}$$

のときには、δは次のように求まる。

$$0.75 = 1 - \frac{1}{1 + e^{-s(0-\delta)}}$$

$$\begin{aligned} \frac{1}{1 + e^{s\delta}} &= \frac{1}{4} \\ 4 &= 1 + e^{s\delta} \\ s\delta &= log\,3 \\ \delta &= \frac{log\,3}{s} \end{aligned}$$

このとき

$$\phi_0(x) = \frac{1}{1 + \mathrm{e}^{-(log 3/\delta)\,x}} = \frac{1}{1 + 3^{-(x/\delta)}}$$

Saito (1992) は一対比較法データの分析のための一般的なモデルについて論じている。しかし、本稿で扱うような全系列法データの場合は上に述べたようなモデルで十分である。本モデルの場合、被験者1人当たりの自由パラメータの数は最大s、τ、Cの3つである。

分　析　例

心理学コースの初級者向け実験「心理学実験（I）」の受講者のレポートから５人を適当に選び分析を試みた。「実験（I）」の「重さの弁別閾」のテーマにおける全系列法では、60 g の重りを標準刺激とし、$S_{-7}=39$ g から $S_7=81$ g までの 15 個の比較刺激が 3 g 間隔で用意される。それぞれの比較刺激と標準刺激との組み合わせは、標準刺激が先のものと後のものの２組用意し、計 15×２＝30 組構成される。この 30 組をランダムな順序で閉目した被験者に提示し、各対比較において後に提示されたものの方が「重い」か「等しい（又は、分からない）」か「軽い」かの判断を行わせる。この 30 組の提示が終わると、同じ 30 組が提示順序を改めてランダムに決め直した後、順に提示される。このランダムな順序による 30 組の提示は４回繰り返されるので、各比較刺激と標準刺激の組み合わせについて標準刺激が先に提示される場合と後に提示される場合のそれぞれの条件において４回の比較判断が求められることになる。対内の提示時間間隔は、マニュアルでは１秒とされているが、この時間間隔では履修者には難しいのでインストラクションを与えるときに「1.5 秒でよい」と説明してある。

データの例（被験者１）を表. 1 に示す。表中、第 j 順位とあるのは標準刺激が同じ対に組み合わされている比較刺激との関係において j 番目に提示される場合のことである。

比較刺激	重い		同じ		軽い	
	第１順位	第２順位	第１順位	第２順位	第１順位	第２順位
81 g	4	4	0	0	0	0
78 g	4	4	0	0	0	0
75 g	4	4	0	0	0	0
72 g	4	4	0	0	0	0
69 g	3	3	0	0	1	1
66 g	0	3	2	0	2	1
63 g	1	2	2	0	1	2
60 g	1	0	2	3	1	1
57 g	0	0	0	1	4	3
54 g	0	0	0	1	4	3
51 g	0	0	0	0	4	4
48 g	0	0	0	0	4	4
42 g	0	0	0	0	4	4
39 g	0	0	0	0	4	4

表1　データの例。被験者１のデータを示す。「重い」、「同じ」、「軽い」は比較刺激の方が標準刺激より「重い」「同じ（分からない）」「軽い」と判断されたことを表す。

5人のデータについて被験者別に分析した結果を表. 2に示す。パラメータの推定値は、対数尤度関数を目的関数としてRosenbrockの座標軸回転法（Rao, 1984）により最適化を行って求めた。表. 2における分析の場合はSimplex法でも本質的に（有効数字の範囲内で）同じ解が得られたが、後に示す表. 3における分析の場合にはSimplex法の収束値はRosenbrockの方法によるものに比べると不十分であった。

被験者	τ （自由パラメータ）				$\tau=0.0$ （固定）		
	AIC	s (75%閾値)	τ	C	AIC	s (75%閾値)	C
被験者1	103.5	0.308 (3.57)	−1.19	1.93	103.1	0.300 (3.66)	1.95
被験者2	140.8	0.183 (5.99)	−1.58	1.92	140.2	0.181 (6.07)	1.92
被験者3	160.0	0.158 (6.96)	1.38	2.66	158.9	0.158 (6.95)	2.62
被験者4	144.9	0.191 (5.77)	1.79	2.47	145.0	0.187 (5.88)	2.49
被験者5	122.2	0.246 (4.47)	0.399	2.20	120.3	0.245 (4.48)	2.20

表2　被験者別の分析結果。τを自由パラメータとする場合と、$\tau=0.0$に固定した（順位誤差を考えない）場合とに分けて分析。閾値δはsの値の下に（　）で囲んで示されている。

3つのパラメータ（s、τ、C）を自由パラメータとした場合と、$\tau=0$に固定した場合（順位誤差を考えない）とを比較すると、被験者4を除いて順位誤差を考慮しない場合の方がAICの値は小さい。被験者間におけるパラメータの推定値の差異を見ると、「等しい（分からない）」のカテゴリー境界の値Cの変化に比べて閾値δの被験者間の変動が大きい。

個人間の差異を検討するために5人分のデータをまとめて分析した結果が表. 3である。被験者kのパラメータを$s_k(\delta_k)$、τ_k、C_kと表せばパラメータの総数は全部で最大3×5＝15個である。15個の全てのパラメータが自由パラメータである場合を、$\{s_k\}$、$\{\tau_k\}$、$\{C_k\}$の3つのパラメータ群のうち1つの群に個人差が無い条件、及び$\tau_k=0$の条件と比較すると、$\{\tau_k\}$又は$\{C_k\}$に個人差が無い条件と$\tau_k=0$の条件の場合のAICの値が全てのパラメータが独立（自由パラメータ）である場合より小さくなっている。$\tau_k=0$の条件は$\{\tau_k\}$に個人差が無い条件よりAICの値が小さい。$\{C_k\}$に個人差が無く、かつ$\tau_k=0$である条件では、$\{C_k\}$に個人差が無い条件あるいは$\tau_k=0$の条件よりさらにAICの値は小さくなっている。しかし、さらに$\{C_k\}$に個人差が無く且つ$\tau_k=0$の条件

モ　デ　ル	AIC
各パラメータ独立	671.4
s_k に個人差無し	674.9
τ_k に個人差無し	669.6
$\tau_k=0.0$（順位誤差無し）	667.6
C_k に個人差無し	664.5
$\tau_k=0.0$ かつ C_k に個人差無し	660.6
s_k、τ_k、C_k に個人差無し	665.3
$\tau_k=0.0$ かつ s_k、C_k に個人差無し	663.3

表3　5人分のデータをまとめて扱った場合。条件別にAICの値を示す。

に $\{s_k\}$ に個人差が無い条件を加えるとAICの値は増加する。又，3つのパラメータ群 $\{s_k\}$、$\{\tau_k\}$、$\{C_k\}$ 全てに個人差が無いとする条件のAICも、$\{C_k\}$ に個人差が無くかつ $\tau_k=0$ の条件の場合より大きい値である。

以上のことから、AICの基準によれば、パラメータの個人差は閾値 δ_k $(=log 3/s_k)$ にのみ認められ、カテゴリー境界の個人差および順位誤差は認められないことが分かる。認知的要因の影響を受けやすい「等しい」という判断のカテゴリー境界に個人差が認められなかったことは受講者の実験時における心構えに大差が無かったことを示していると思われ、感覚過程における弁別閾の方は受講者の心構えに拘らず個々の被験者における差異が現れたと考えられる。つまり、感覚と判断の過程を区別する信号検出理論のモデルによる分析が成功したといえる。表.4に $\{C_k\}$ に個人差が無くかつ $\tau_k=0$ の条件におけるパラメータの推定値を示す。

被験者	s_k (δ_k)	τ_k	C_k
被験者1	0.294 (3.74)	0.0	2.22
被験者2	0.177 (6.20)	0.0	2.22
被験者3	0.162 (6.78)	0.0	2.22
被験者4	0.190 (5.79)	0.0	2.22
被験者5	0.245 (4.49)	0.0	2.22

表4　パラメータの推定値。表.3の「$\tau_k=0.0$ かつ C_k に個人差無し」の条件におけるパラメータの値を示す。

順位誤差が認められなかったことについては用いたデータが順位誤差の分析には不適切であった可能性だけを指摘しておく。

考　　察

刺激間の差が知覚されるのに必要な最小の差の値としての閾値を、より大きな感覚量を惹き起こす確率が75%である差の値とする定義が採用された。さらに、「同じ」という反応に対する基準値（境界）を設定して、反応の3カテゴリーの確率を与えるモデルを構成した。このモデルは、信号検出理論の考え方に従うものであり、感覚過程の特性に関わる閾値と判断過程の特性に関わる反応カテゴリーの基準値が区別されている。このモデルに基づいて全系列法データの分析が行われた。被験者間に反応カテゴリーの基準値に差は認められなかったが、閾値には差が認められた。これは、受講者の授業における実験に対する態度の差に比べて、感覚過程の個人差が大きかったことを意味している。

上のような感覚過程と判断過程とを区別する分析はそれに対応するモデルを構成することにより可能となるものであり、本稿における分析例からもその区別の必要性が分かる。また、一般に初級実験のような簡単な場合でもモデルに基づく分析が重要であることも示されていると思われる。

最後に、感覚量の差の分布について少しふれておく。

感覚量の差の確率分布の形としてlogistic分布が採用された。これは、分布の形における細かい差異は実証的には意味をもたないことがあること（岡本、1992）、logistic分布が計算上簡単な形をしていることが理由であった。しかし、logistic分布を用いることの理論的根拠もLuce（1959）により示されている。彼の選択モデルによれば、一対比較の確率はlogistic分布により与えられる（p. 40）のである。又、Hellström（1993）は、感覚量等の分布に正規分布が採用されることに対して批判を加えている。彼によれば、感覚過程のモデルの場合は裾（tail）の長い分布の方が適しており、彼の一対比較のデータの分析では正規分布よりlogistic分布、logistic分布よりt分布の方がモデルに適合度がよい。しかし、t分布を採用した場合は、計算速度の点ではlogistic分布に対して不利である。本稿で扱われたような分析例が低速のパーソナルコンピュータ（PC－9801vm 2）上で行われる場合にはlogistic分布の計算上の簡便さによるメリットは大きい。

引 用 文 献

Falmagne, J. C. (1986) Psychophysical measurement and theory. In : K. R. Boff, L. Kaufman and J. P. Thomas (eds.) Handbook of Perception and Human Performance, Vol. 1, John Wiley and Sons.

Gescheider, G. A. (1985) Psychophysics : Method, theory, and Application, 2nd ed. Lawrece Erlbaum Associates, Publishers.

Hellström, A. (1993) The normal distribution in scaling subjective stimulus differences : Less "normal" than we think ? *Perception & Psychophysics*, **54**, 82 - 92.

金沢大学文学部心理学研究室 (1990) 心理学実験(I)マニュアル(改訂版)

Laming, D. (1986) Sensory Analysis. Academic Press.

Luce, R. D. (1959) Individual Choice Behavior : A theoretical analysis. Greenwood Press, Publishers.

岡本安晴 (1992) 正規分布とロジステック分布のシュミレーションによる比較. 金沢大学文学部論集行動科学科篇、第12号、34 - 45.

Rao, S. S. (1984) Optimization : Theory and Applications, 2nd ed., John Wiley and sons.

Saito, T. (1992) Measurement of the Asymmetry Observed in Comparative Judgment. Hokkaido Behavioral Science Report, Series M, No.19, August.