

$$t_i = \frac{Y + (-1)^i \sqrt{(u/v)^2(X^2 + Y^2) - X^2}}{v(1 - (u/v)^2)} \quad \sin \theta_i = \frac{1}{u} \left(v + \frac{Y}{t_i} \right), \quad \cos \theta_i = \frac{X}{ut_i}$$

$$f(x, y; t, \theta) = h(x, y) \exp \left\{ - \int_0^t h(x - su \cos \theta, y + (v - u \sin \theta)s) ds \right\}$$

である。これに対して異なる鯨がそこで発見される確率密度は

$$P_2 = D^2 \int_0^{2\pi} f(x_A, y_A; t_A, \theta) \frac{d\theta}{2\pi} \int_0^{2\pi} f(x_B, y_B; t_B, \theta') \frac{d\theta'}{2\pi}$$

である。これに観測誤差が加わると、密度が高くなるにつれ異なる個体である可能性が高くなる様子が表現されている。

DNA 配列データによるヒト上科の分類

長谷川 政 美

ヒト上科 Hominoidea とは、ヒトと類人猿を含む分類単位である。ヒトと類人猿とが系統進化の過程でいつ分かれたかという問題は、まだ十分に解明されていない。化石の証拠はもちろん重要な手掛りを与えるものではあるが、解釈の任意性や研究者の主観的な判断の入り込む余地があまりにも多く、客観的な評価基準に欠けていた。ところが近年、分子遺伝学の進歩により、DNA など分子レベルの証拠にこそ客観的な手掛りがあることが明らかになってきた¹⁾。われわれはミトコンドリア DNA の塩基配列データから、生物種が分岐した年代を推定するための統計的方法を開発し^{2,3,4)}、それをもとにヒト上科の進化における分岐年代を推定した⁴⁾。それによると、テナガザル、オランウータン、ゴリラ、チンパンジーがヒトに至る系統からそれぞれ分岐したのが、今から 1330 ± 150 , 1090 ± 120 , 370 ± 60 , 270 ± 60 万年前(±は1SD)ということになる。ただしミトコンドリア DNA は母性遺伝をするために、異種間で交雑が起った場合に、この DNA は種の壁を越えてそっくりそのまま別の種に転移することがある。そのようなことは種として分かれてから 100~150 万年くらいの期間は起こり得るので、ヒトとチンパンジーの祖先集団間でそのようなことが起きたとすると、この2種の分岐は 400~500 万年前までさかのぼるかもしれない。いずれにしてもこれまで常識的にはこの分岐は 1000 万年よりも古いと考えられていたのにくらべて、実際にははるかに新しいことのようにである。一般には、チンパンジー、ゴリラ、オランウータンはオランウータン科、ヒトだけがヒト科として分類されてきた。しかしこのような分類は明らかに系統関係を反映していない。ヒトとチンパンジーの分岐を実際よりもはるかに古く考えてきた原因の一つが、このような誤まった分類体系の使用にあったと思われる。われわれは系統関係を反映した新しい分類体系を提案した(図1)⁵⁾。

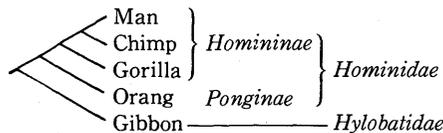


図1. ヒト上科の新しい分類体系(文献5)。

参 考 文 献

- 1) 長谷川政美(1984)「DNA からみた人類の起原と進化」海鳴社
- 2) Hasegawa, M., Yano, T. and Kishino, H. (1984) *Proc. Japan Acad.*, 60B, 95-98.
- 3) 長谷川政美, 岸野洋久, 矢野隆昭(1984) 統計研彙報 32, 59-70.
- 4) Hasegawa, M., Kishino, H. and Yano, T. (1985) *J. Mol. Evol.* (in press)

- 5) Hasegawa, M. and Yano, T. (1984) *Proc. Japan Acad.*, 60B, 389-392.

ノンパラメトリックなハザード関数の推定

鎌倉稔成

系列事象の統計的解析では、強度関数 (intensity function), あるいはハザード関数が重要な役割を演ずる。ハザード関数は信頼性理論の分野では故障率関数と呼ばれており、パラメトリック、ノンパラメトリック、数々の推定法がある (Singpurwalla & Wong, 1993)。よく用いられるのはワイブル分布、ガンマ分布、レーリー分布などのパラメトリックなモデルを仮定してそのパラメータを推定するといったかたちでハザード関数を求める方法である。ハザード関数をノンパラメトリックに求めるのは、密度関数をノンパラメトリックに求めるのと同程度のむずかしさがある。通常は、累積ハザード関数を推定することが多い。累積ハザード関数の推定量としては、Nelson の推定量、Perterson の推定量がよく使われている。ハザード関数そのものの推定方法としては、actuarial estimates, kernel estimates, などがあり、漸近理論の研究もある。ここでは、ノンパラメトリックに滑らかなハザード関数の推定を行う問題を考える。滑らかなハザード関数を推定する方法としては、Klotz (1982) が滑らかな生存時間分布関数の推定に用いた B スプラインの利用が考えられる。この他、Akaike (1980) の Bayes モデルの利用も 1 つの方法といえる。本小稿では B スプラインによるハザード関数の推定を議論する。

スプライン関数は区分的な多項式であり、プロットされた点を通る滑らかな曲線を得るために用いられており、種々の応用が考えられている (市田・吉本, 1979)。ここでは、基準化された B スプラインを用いる。Klotz (1982) では 2 階の B スプラインを用いており、滑らかな生存時間分布関数の推定を行うにはさほど問題はないが、ハザード関数の推定には 2 階の B スプラインでは十分とはいえない。尤度計算には B スプライン関数の積分が必要になるが、これは簡単には表現されない。筆者が数式処理プログラム REDUCE 3 を用いて $\int_0^x N_{k,r}(u)du$ の計算を行った結果を記述する。たとえば、2 階の B スプラインでは次のようになる。

- i) $x \leq \tau_k$
0
- ii) $\tau_k < x \leq \tau_{k+1}$
 $(\tau_k - x)^2 / 2(\tau_{k+1} - \tau_k)$
- iii) $\tau_{k+1} < x \leq \tau_{k+2}$
 $-(\tau_{k+2}\tau_{k+1} + \tau_{k+2}\tau_k - 2\tau_{k+2}x - \tau_{k+1}\tau_k + x^2) / 2(\tau_{k+2} - \tau_{k+1})$
- iv) $\tau_{k+2} < x$
 $(\tau_{k+2} - \tau_k) / 2$

第 3 研究部

医学データ解析における不適切問題の解法

田 辺 國 士

肺内ガス交換機能解析に登場するある不適切問題の統計的解法を与えた。問題は被験者の呼気に含まれるガスの分圧を測定し、この情報から肺内での換気・血流比 r の分布を推定することである。分布を換気血流比 r の対数 $R = \log r$ の密度函数 $p(R)$ で表現するとき、定常ガス交換における拡散平衡に関する物質保存の原理から、観測値 d と密度函数 $p(R)$ の間に

$$(1) \quad d = K(p) + \varepsilon$$