次回の南海地震の確率予測—その偏りと誤差について—

尾 形 良 彦*

Biases and Uncertainties when Estimating the Hazard of the Next Nankai Earthquake

Yosihiko OGATA *

Abstract

After a renewal process is estimated for data on repeated occurrences of large-scale earthquakes, we predict the hazard of the next rupture as a function of time using the estimated parameters. However, when the number of events in the data is small, the hazard function based on the maximum likelihood estimates (MLE) has not only a large uncertainty in the estimation but also sometimes has a serious bias. To see the uncertainty of estimated hazard functions, we use a Bayesian inference. In this manuscript I discuss the uncertainty of the estimated hazard functions of the forthcoming Nankaido earthquake based on historical data consisting of ten large events. Moreover, assuming the same distribution of the ratio of the time-interval length between the successive events to the slip size of the first event of the pair, an extended renewal process for the Shimazaki-Nakata's Time-predictable model can be considered. Thus we estimate not only the alternative hazard function of the next Nankaido event but also its uncertainty based on occurrence time data for the last three large events associated with records of coseismic uplift sizes. The same extended renewal model is applied to the same ten large events noted above associated with the estimated slip sizes that are calculated from Usami's magnitude estimates. The three datasets lead to significantly different predictions, even if the estimation uncertainty is taken into consideration.

Key words : hazard rate function, Lognormal distribution, Bayesian inference, posterior distribution, Time-predictable model, renewal processes, slip sizes

キーワード : 発生危険度関数, 対数正規分布, ベイズの推論, 事後分布, 発生時期予測可能モデル, 更新過程, 滑り量

I. はじめに

統計地震学では, 古地震や歴史地震の発生時刻データはイベント数が少ないため, ポアソン過程を除き最も簡単な点過程モデルとして更新過程が良く使われる。対数尤度を最大化すること (最尤

法) によって更新過程の発生間隔の確率分布のパラメタを推定できる (最尤推定値)。しかし, なるべくパラメタ数を節約することが推奨されている。例えば発生間隔に対数正規分布を考える場合, 分散を示すパラメタを固定したものを標準モデルとする。そして, 2つのパラメタを使ったモデルと

* 統計数理研究所

* The Institute of Statistical Mathematics, Tokyo

AIC（赤池情報量規準）によって比較し、良い方のモデルを使って将来の発生確率を予測するのである（地震調査推進本部地震調査委員会長期評価部会，1998；Ogata，1999 など）。

しかし特に地震数が少ない場合、また地震間隔がほぼ揃った場合、最尤法を適用すると直感に合わない結果が得られる場合がある。本稿では、限られたデータ数（情報）のもと、最尤法の限界を超えるために尤度関数に立ち返り、ベイズ的推論によってパラメタの推定値や予測確率などの誤差について議論する。具体的な解析例として、以下のデータに基き南海地震の確率予測やその誤差の分布を議論する。

II. データ

本稿では南海地震に関する2種類のデータを考える。第一に地震調査推進本部地震調査委員会長期評価部会（1998）の表3.1にあるdata set I（または、宇津，1999，表10.5）の9個のイベントに、1200年代前半のイベント（寒川，1996）を仮に1233年として、加えた以下の10個のデータに基いて考察する（データ1）。

二つめはShimazaki and Nakata（1980）にある、地震に伴った室戸岬の隆起量（室津港の水位変化）を含む最近の3個のイベント（データ2）を解析する。この隆起量のデータも使って、発生時期予測可能モデル（Time-predictable model：Shimazaki and Nakata，1980）の誤差を調べるために、更新過程を拡張した統計的モデルをⅢ章紹介する。

三つめは、発生年は最初のデータと同じものであるが、宇佐美（1987）に基く各推定最大および最小マグニチュードの平均値から地震モーメントの推定値を $\log M_0 = 1.5M + 16.1$ によって求め、さらにこれが滑り量に比例しているとし、1946年の地震の滑り量が1.0であるとした場合の他の地震の相対な滑り量を使って、データ2の場合と同様に発生時期予測可能モデルの更新過程拡張版をあてはめる。1233年のマグニチュードについてはKumagai（1996）に従った（データ3）。名古屋大学理学部（1996）とKumagai（1996）はこれと同様なデータによってモンテカルロ手法で解析し、データによっては近い将来を中心（2006 ± 10 年）に相対的に高い発生確率の可能性を予測している。

以上のデータに加えて、最後の地震（1946年の南海地震）以来今日（2000年12月末）まで南海地震が起きなかったという事実を踏んで以下の章において解析する。

III. モデルの推定と確率予測

地震の発生年の系列 $t_1 < t_2 < \dots < t_i < \dots$ について時間間隔 $X_{i+1} = t_{i+1} - t_i$ ($i = 1, 2, \dots$) を考える。これらが互いに独立で同じ確率分布に従っているとき、この系列は更新過程と呼ばれる。特にこれが指数分布である場合は定常ポアソン過程である。以下、間隔分布の密度関数を $f(x|\theta)$ ，分布関数を $F(x|\theta)$ で記すが、本稿での間隔分布は対数正規分布

データ1 (Dataset 1; year in A.D.)

発生年	684	887	1099	1233	1361	1498	1605	1707	1854	1946
-----	-----	-----	------	------	------	------	------	------	------	------

データ2 (Dataset 2; year in A.D. with uplift size)

発生年（隆起量）	1707 (1.8 m)	1854 (1.2 m)	1946 (1.15 m)
----------	--------------	--------------	---------------

データ3 (Dataset 3; year in A.D. with the estimated magnitudes (Usami, 1987))

発生年 (A.D.)	684	887	1099	1233	1361	1498	1605	1707	1854	1946
推定最大マグニチュード	8 $\frac{1}{4}$	8.0	8.0	8.1	8 $\frac{1}{4}$	8.2	7.9	8.4	8.4	8.0
推定最小マグニチュード	8 $\frac{1}{4}$	8.5	8.3	8.3	8.5	8.4	7.9	8.4	8.4	8.0
推定滑り量の比 (M_0 ratio)	1.83	1.83	1.45	1.62	2.43	2.05	0.82	2.57	2.57	1.0

$$f(\mathbf{x}|\mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma\mathbf{x}} \exp\left\{-\frac{(\ln \mathbf{x} - \mu)^2}{2\sigma^2}\right\} \quad (1)$$

に限ることにする。ここで $\mathbf{x}$ は地震の発生間隔で、 μ と σ はそれぞれ、間隔の対数 ($\ln \mathbf{x}$: $\ln$ は自然対数 $\log_e$) が正規分布に従うと仮定したときの平均と標準偏差である。以下、パラメタを $\theta = (\mu, \sigma)$ としてベクトル表記する。

現在の時刻を $\mathbf{T}$ として、現在まで n 個のイベントのデータ $\{t_i: i=1, 2, \dots, n\}$ が得られたときの対数尤度関数は

$$\begin{aligned} \ln L(\theta|t_1, \dots, t_n, \mathbf{T}) &= \sum_{i=2}^n \ln f(t_i - t_{i-1}|\theta) \\ &+ \ln \mathcal{F}(\mathbf{T} - t_n|\theta) \end{aligned} \quad (2)$$

と書ける (例えば, Ogata, 1978; Brillinger, 1982; Daley and Vere-Jones, 1988; Davis *et al.*, 1989)。ここで、最後の項の

$$\begin{aligned} \mathcal{F}(\mathbf{x}|\theta) &= \text{Prob}(\text{Time} > \mathbf{x}) = \int_{\mathbf{x}}^{\infty} f(t|\theta) dt \\ &= 1 - F(\mathbf{x}|\theta) \end{aligned} \quad (3)$$

は生存関数と呼ばれ、時間が $\mathbf{x}$ 年経ってもイベントが起きない確率を示し、今回のデータの場合、(2) 式の最後の項は 1946 年の南海地震から現在に至るまで南海地震が起きていないという情報を示す。

いま仮にパラメタの推定値 $\hat{\theta}$ が得られたものとする。例えば最尤推定値 (MLE) は対数尤度 (2) を最大化するようなパラメタ値であり、もし (2) 式の最後の項がなければ $\{\ln(t_i - t_{i-1}): i=2, 3, \dots, n\}$ の標本平均と標本標準偏差に一致するが、実際は $\mathbf{T} - t_n$ の値によって少しずつ違ってくるので計算機で数値的に求めなければならない。次章では最尤推定値以外の推定値なども議論する。

最後のイベント発生から $\mathbf{x}$ 年の経過時刻に於いて次のイベントが起きる危険度関数は

$$\begin{aligned} \lambda(\mathbf{x}|\theta) &= \lim_{\Delta \downarrow 0} \frac{\text{Prob}\{\mathbf{x} < \text{Time} < \mathbf{x} + \Delta | \text{Time} > \mathbf{x}\}}{\Delta} \\ &= \frac{f(\mathbf{x}|\theta)}{\mathcal{F}(\mathbf{x}|\theta)} \end{aligned} \quad (4)$$

で定義され、最後の等式は条件付き確率の計算で導かれる。

ここで最後のイベント発生から現在までの経過年を $\mathbf{x}_0$ とすれば、現在から将来 $\mathbf{y}$ 年間の内に次のイベントが起きる確率は次の様に計算できる。

$$\begin{aligned} P(\mathbf{x}_0; \mathbf{y}|\theta) &= 1 - \exp\left\{-\int_{\mathbf{x}_0}^{\mathbf{x}_0+\mathbf{y}} \lambda(\mathbf{x}|\theta) d\mathbf{x}\right\} \\ &= 1 - \frac{\mathcal{F}(\mathbf{x}_0 + \mathbf{y}|\theta)}{\mathcal{F}(\mathbf{x}_0|\theta)} \end{aligned} \quad (5)$$

さらに、これを微分することによって、現在から将来 $\mathbf{y}$ 年間までに次のイベントが起きる確率の密度関数

$$\begin{aligned} P(\mathbf{x}_0; \mathbf{y}|\theta) &= \lambda(\mathbf{x}_0 + \mathbf{y}|\theta) \exp\left\{-\int_{\mathbf{x}_0}^{\mathbf{x}_0+\mathbf{y}} \lambda(\mathbf{x}|\theta) d\mathbf{x}\right\} \\ &= \frac{\mathcal{F}(\mathbf{x}_0 + \mathbf{y}|\theta)}{\mathcal{F}(\mathbf{x}_0|\theta)} \cdot f(\mathbf{x}_0 + \mathbf{y}|\theta) \end{aligned} \quad (6)$$

が得られる。

他方、データ 2 や 3 のように、 n 個のイベントのデータ $\{t_i: i=1, 2, \dots, n\}$ に対応する滑り量データ $\{L_i: i=1, 2, \dots, n\}$ が使えるならば、発生時期予測可能モデル (Shimazaki and Nakata, 1980) を支持する更新過程の拡張版を考えることができる。すなわち地震発生時間間隔を滑り量で割り算 (規格化) したものが全て同じ対数正規分布に従うものと仮定して次のように定義する。

$$\begin{aligned} \ln L(\theta|t_1, \dots, t_n, \mathbf{T}; L_1, \dots, L_n) &= \\ \sum_{i=2}^n \ln f\left(\frac{t_i - t_{i-1}}{L_{i-1}}|\theta\right) &+ \ln \mathcal{F}\left(\frac{\mathbf{T} - t_n}{L_n}|\theta\right) \end{aligned} \quad (7)$$

このモデルの場合最後のイベントの滑り量を L_n としたとき、最後のイベントから時間経過 $\mathbf{y}$ 年で、次のイベントが起きる危険度関数は

$$\lambda(x|\theta) = \lim_{\Delta \downarrow 0} \frac{\text{Prob}\{x < \text{Time} < x + \Delta | \text{Time} > x\}}{\Delta} \\ = \frac{f(x/L_n|\theta)}{\mathcal{F}(x/L_n|\theta)} \cdot L_n^{-1} \quad (8)$$

であり、同様にして (5) 式の累積確率および (6) 式の密度関数に対応するものを計算できる。

IV. ベイズ的推論と予測危険度関数の誤差

ベイズ的推論の著書については、例えば Box and Tiao (1973) や O'Hagan (1994) を参照されたい。本章ではデータもベクトル X で略記することにする。尤度関数 $L(\theta|X)$ はデータと更新過程モデルの係わりについての全ての統計的情報を持っている。ベイズ的モデルを定義するために、パラメタ θ に関する適当な事前分布 $\pi(\theta)$ を与え、尤度関数を掛け合わせ規格化した事後分布

$$f(\theta|X) = \frac{L(\theta|X) \pi(\theta)}{\int_{\Theta} L(\theta|X) \pi(\theta) d\theta} \quad (9)$$

はデータを与えたときのモデルのパラメタ値の確からしさの分布を示す。 Θ はパラメタが定義されている空間であり、後ほど具体的に述べる。最尤法が典型的に優れているのは尤度関数が対称で周辺部の裾が軽い場合である。データ数が十分でないと、非対称で裾が重くなっている場合が多く、最尤推定値 $\hat{\theta}_M$ は多かれ少なかれ偏ったものとなり、事後分布の代表値としては適当でなくなる。代わりにの推定値としては事後分布の平均値

$$\hat{\theta}_B = \int_{\Theta} \theta f(\theta|X) d\theta = \frac{\int_{\Theta} \theta L(\theta|X) \pi(\theta) d\theta}{\int_{\Theta} L(\theta|X) \pi(\theta) d\theta} \quad (10)$$

が考えられる。これはベイズ推定値とも呼ばれる。

これらの事後分布の代表値である MLE $\hat{\theta}_M$ やベイズ推定値 $\hat{\theta}_B$ を (4) 式に代入して予測のための危険度関数を

$$\hat{\lambda}_M(x) = \lambda(x|\hat{\theta}_M) \quad (11)$$

$$\hat{\lambda}_B(x) = \lambda(x|\hat{\theta}_B) \quad (12)$$

のように定義し、将来の発生確率 (5) を計算する

ことができる。しかし、データが少ない場合には、推定値の誤差が大きく、それを調べるには事後分布全体の様子が重要であり、これをもとに以下のような統計的推論を考えることができる。すなわち、危険度関数の事後分布にかんする平均

$$\hat{\lambda}_P(x) = \hat{\lambda}_P(x|X) = \int_{\Theta} \lambda(x|\theta) f(\theta|X) d\theta \quad (13)$$

(例えば, Rhoades *et al.*, 1994; Ogata, 1999) は、(11) や (12) のような推定値を代入した予測関数より平均的な意味で優れた予測危険度関数の推定になっていることが Akaike (1978) の場合と同様に示される。これをベイズ的予測危険度関数 (Bayesian predictive hazard function) と呼ぶことにする。

次に、予測危険度関数や確率予測の誤差分布を調べるためにパラメタについて以下のような規格化をして議論を行う。対数正規分布のパラメタ $\theta = (\mu, \sigma)$ は領域 $\Theta = (-\infty, \infty) \times (0, \infty)$ で定義されているが、シミュレーションおよび以下の図中の事後分布の二次元表示の便宜上、これらを変換して $\Theta_0 = (0, 1) \times (0, 1)$ 上の標準化された $\theta_0 = (\mu_0, \sigma_0)$ で考えることにする。本稿では例えば

$$\mu - \frac{T-t_1}{n} = \ln \frac{\mu_0}{1-\mu_0} \text{ および } \sigma = -\frac{\ln(1-\sigma_0)}{2} \quad (14)$$

で変換することにする。ここで $(T-t_1)/n$ はイベント間の平均時間間隔の仮の推定値である。以後本稿では、事前分布 $\pi(\theta)$ は変換した空間 Θ_0 で一様分布 ($\pi(\theta_0)=1$) になるようなものを考えることにする。すなわち

$$\pi(\theta) = \pi(\mu) \cdot \pi(\sigma) = \frac{d\mu_0}{d\mu} \cdot \frac{d\sigma_0}{d\sigma} \\ = \frac{e^{-\mu}}{(1+e^{-\mu})^2} \cdot 2e^{-2\sigma} \quad (15)$$

である。さらにパラメタ空間 Θ_0 での事後分布は、 Θ での対応する領域上の事後分布の積分値が互いに等しくなるように、変数変換のヤコビ行列を施したものでなければならない。従って変換したパ

ラメタの事後分布は

$$f_0(\mu_0, \sigma_0) = L \left(\frac{T-t_1}{n} + \ln \frac{\mu_0}{1-\mu_0}, -\frac{\ln(1-\sigma_0)}{2} \right) / (\text{Normalizing const.}) \quad (16)$$

となる。ここで規格化定数は分子の尤度関数をパラメタ空間 Θ_0 上で積分したものである。以後この事後分布でパラメタをサンプリングして、その推定値などの誤差分布のシミュレーションをすることにする。

V. 南海地震の確率予測について (データ 1 の場合)

本章では、II 章の 10 個の歴史地震データに対数正規分布の更新過程をあてはめたとときの危険度とその誤差について調べる。図 1 は (16) 式の事後分布の等高線図である。

これらに基く最尤推定値とベイズ推定値は各々 $\hat{\mu}_M = 4.90$ ($\hat{\sigma}_M = 0.268$) そして $\hat{\mu}_B = 4.91$ ($\hat{\sigma}_B = 0.314$) である。 μ に関して最尤推定値とベイズ推定値がほとんど同じ値であるのに、 σ (パラッキ) の推定値に違いがでている。これは事後分布が μ に関して対称であるのに対して、 σ に関して非対称で裾が重い分布形になっているためである (式 (1) から σ に関しては逆ベキで減衰している)。

(16) 式の事後確率分布 $f_0(\mu_0, \sigma_0)$ から独立に 100 組の $\theta_0 = (\mu_0, \sigma_0)$ 値の標本をサンプリングして、それらの値をパラメタに代入した 100 組の危険度関数 $\lambda(x|\theta_0)$ を示したものが図 2A に示してある。これらがデータ 1 の地震と対数正規分布の更新過程モデルで推定される危険度関数の誤差の有り様を表わしている。これに対して図 2B は事後分布の代表値としての最尤推定値とベイズ推定値を代入した危険度関数 (10 式および 11 式参照) とベイズ的予測危険度関数 (12 式参照) を図示している。ここでベイズ的予測危険度関数は図 1A の全ての関数を重ね合わせ平均をとったものと殆ど同じものであることに注意する。

図 2B から、10 個の地震ぐらいだと、それぞれの代表的な推定値による予測危険度関数は大きく変わらないことがわかる。

図 1 データ 1 に対する対数正規分布更新過程モデルによって計算された、パラメタ μ_0 (横軸) と σ_0 (縦軸) に関する、事後分布の密度関数 (16 式) の等高線。
ただし対数スケールで -10 から 0 までの等高線で、 -11 以下は省略した。上の $+$ はベイズ推定値 (事後分布の平均値)、下の $+$ は最尤推定値の位置を示す。

Fig. 1 Heights of the posterior density distribution in (16) of μ_0 (the horizontal axis) and σ_0 (the vertical axis) for the Dataset 1 are contoured at unit intervals on a logarithmic scale for the range from -10 to 0 , but no contours are shown below the value -10 . The lower and upper plus signs show the MLE and Bayes estimate (posterior mean), respectively.

図 2A の危険度標本関数に対応する 100 組の累積確率予測関数 (5) が図 3A に示してある。各関数が或る時刻 t を横断する 100 個の点をヒストグラムで表わすと、これらが時刻 t までの確率予測の誤差分布を表わす。すなわちデータ 1 の地震に基いて対数正規分布の更新過程モデルで推定された時刻 t までの予測確率の誤差分布を表わしている。同様に図 2B に対応する累積確率予測関数が図 3B に示してある。10 個の地震ぐらいであれば、それぞれの代表的な推定値による予測危険確率が大きく変わらないことがわかる。図 3C と 3D はそれぞれ図 3A と 3B に対応する予測累積確率の密度関数である。このように、データ 1 と対数正規分布の更新過程モデルに基くと、2070 ~

図 2 A: 図 1 で示した事後分布 (16) からサンプリングされた 100 組のパラメタ値を代入して得られた 100 組の危険度関数. 推定危険度関数の誤差の有り様を表わしている. 水平の点線は平均発生率 (100 年あたり 0.76 イベント) の水準, 垂直の点線は予測開始時点 (2001 年 1 月) を示す.
B: 3 つの代表的な推定危険度関数. M は最尤推定値によるもの (11 式), B はベイズ推定値 (事後分布平均, 9 式) によるもの (11 式), および P はベイズの予測危険度関数 (13 式).

Fig. 2 A: One hundred samples of hazard functions of the elapsed time after the last event, representing the variability of the estimation. Their parameter values are simulated from the posterior probability distribution in (16) of the renewal model with a lognormal distribution applied to the Dataset 1. The horizontal dashed line indicates the level of the mean hazard rate throughout the whole time span (i.e., 0.0076 event/year), and the vertical dotted line indicates the current time (January 2001). B: Estimated hazard rate functions with the MLE in (11), with the Bayes estimate in (12), and the Bayesian predictive hazard function in (13).

2080 年頃を中心に次の南海地震の発生が考えられそうである。

VI. 南海地震の確率予測について (データ 2 の場合)

Ⅲ章の (7) 式に示すように, 地震発生時間間隔をスリップ量で除算 (規格化) することによって, 発生時期予測可能モデル (Time-predictable model: Shimazaki and Nakata, 1980) を支持するような更新過程の拡張版を考えることができる。ただし, データ 2 で示すように隆起量データがあるのは最近の 3 個のイベントだけである。最後のイベントから現在に至るまで南海地震が発生していないという情報を尤度に加えたとしても, まことにイベント数が少ないデータである。しかし, この場合でも事後分布は定義されており, 前章と同様に危険度とそのばらつきの有り様について議論できる。図 4 は (16) 式の事後分布の等高線図である。

これらに基く最尤推定値とベイズ推定値は各々 $\hat{\mu}_M = 4.38$ ($\hat{\sigma}_M = 0.034$) そして $\hat{\mu}_B = 4.39$ ($\hat{\sigma}_B = 0.163$) である。 μ に関して最尤推定値とベイズ推定値がほとんど同じ値であるのに, バラツキの推定値 $\hat{\sigma}$ に関して極めて大きな違いがでている。これは事後分布が μ に関して対称であるのに対して, σ に関して非対称で裾が極めて重い分布形になっているためであるが, その上, 規格化されたイベント間隔 $\{(t_i - t_{i-1}) / L_{i-1} : i = 2, 3\}$ が殆ど互いに等しいので, σ の最尤推定値は極めて小さい。図 4 を見る限り, σ の最尤推定値は事後確率分布の代表値としては極めて不適である。これに対してベイズ推定値の方がわれわれの統計的な常識に近い。

前章と同様にして (16) 式の事後確率分布 $f_0(\mu_0, \sigma_0)$ から独立に 100 組の (μ_0, σ_0) をサンプリングして, それらの値を危険度関数 (4) のパラメタに代入した 100 組の標本関数が図 5A に示されている。これらがデータ 1 の 3 個の地震と対数

図 3 A：現在（縦点線）を起点とした累積確率関数群（図 2A に対応，数式 5 参照）。
 B：図 2B に対応する累積確率関数。
 C，D：それぞれ図 3A と図 3B に対応する確率密度関数群（6 式参照）。

Fig. 3 A: Cumulative probability functions of time starting from the present (the vertical dashed line), corresponding to the hazard rate functions in Fig. 2A using the equation in (5). These represent the variability of the probability estimate. B: Probability functions based on MLE, Bayes estimate, and Bayesian predictive hazard function in (13). C and D: The probability density functions in (6) corresponding to Figs. 3A and 3B, respectively.

正規分布の更新過程モデルで推定される危険度関数のばらつきの有り様を示す。まことに大きな不確定性を表わしている。これも 3 個あまりしかないデータのせいであろう。いずれにしても図 5A の示すところは，第一に，危険度関数が 2040 年前で鋭い立ち上がりを見せている様に，典型的に発生時期予測可能モデルに従っている可能性と，第二に他の危険度関数のパターンでも，規格化された 2 つのイベント間隔がたまたま等しくなった可能性（宇津，1994）が混在していることを示している。従って，図 3B と異なり，図 5B では最尤推

定値とベイズ推定値を代入した危険度関数（10 式および 11 式）とベイズ的予測危険度関数（12 式）それぞれの予測危険度関数は大きく違っていることがわかるが，II 章で述べたように平均的な意味ではベイズ的予測危険度関数に基く確率予測が最も優れていると思われる。

図 5A の危険度標本関数に対応する 100 組の累積確率予測関数（5）が図 6A に示してある。これらがデータ 2 から推定された累積確率予測関数の誤差分布を表わしている。同様に図 5B に対応する代表的な累積確率予測関数が図 6B に示して

ある。図 6C と 6D はそれぞれ図 5A と 5B に対応する予測確率の密度関数である。それぞれの予測確率が大きく違っていることがわかる。データ 2

図 4 データ 2 に関する対数正規拡張版更新過程モデルの事後分布の等高線（尤度関数は数式 7 参照）。その他は図 1 の説明を参照。

Fig. 4 Heights of the posterior density distribution in (16) of the extended renewal process with the log-normal distribution for Dataset 2. See the caption of Fig. 1 for details.

図 5 A: 図 4 で示した事後分布からサンプリングした 100 組のパラメータを代入して得られた危険度関数群。横の点線は平均発生率（100 年あたり 1.02 イベント）の水準。
B: 図 2 の説明文参照。

Fig. 5 A: One hundred samples of hazard rate functions of time, starting from the occurrence of the last event, simulated from the posterior probability distribution in (16) for the Dataset 2. The horizontal dashed line indicates the mean hazard rate (0.0102 event/year) and the vertical dashed lines indicate the current time (January 2001). B: The estimated hazard rate functions are similarly defined as those in Figs. 2A and 2B.

と対数正規分布の更新過程拡張版モデルに基づく 2040 年頃を中心に次の南海地震が考えられ、2000～2040 年の発生確率は 6 割にも上ることがわかる。

VII. 南海地震の確率予測について (データ 3 の場合)

前章と同様に、地震発生時間間隔を滑り量で除算し規格化して拡張版対数正規更新過程をあてはめた。まず、発生年はデータ 1 と同じなので、データ 3 の推定滑り量を考慮した拡張版更新過程モデルを使った方が適合度が良いのか否かを見るためのモデル比較ができる。結果はデータ 1 の対数正規モデルと比べて AIC が 2.8 だけ小さくなり、データ 3 に与えられている推定滑り量を考慮して確率予測をした方が多少良いと考えられる。

図 7 は (16) 式の事後分布の等高線図である。

これらに基く最尤推定値とベイズ推定値は各々 $\hat{\mu}_M = 4.33$ ($\hat{\sigma}_M = 0.402$) そして $\hat{\mu}_B = 4.34$ ($\hat{\sigma}_B = 0.453$) である。 μ に関して最尤推定値とベイズ推定値がほとんど同じ値であるのに、バラツキ σ の推定値に関して多少の違いがでている。前章の場

図 6 A, B: それぞれ図 5A と図 5B に対応する累積確率関数群。
C, D: それぞれ図 5A と図 5B に対応する確率密度関数群。

Fig. 6 A: Cumulative probability functions of time, starting from the current time (January 2001) derived from the hazard rate functions in Fig. 5A using the first part of the equation in (5). B: The cumulative probability functions based on MLE, Bayes estimate, and Bayesian predictive hazard function in (13). C and D: The probability density functions in (6) corresponding to the functions in Figs. 5A and 5B, respectively.

合と比べてパラツキの推定値 $\hat{\sigma}_M$ と $\hat{\sigma}_B$ が大きいのはデータに発生時期予測可能モデルに合わない部分があることを示している。

(16) 式の事後確率分布 $f_0(\mu_0, \sigma_0)$ から独立に 100 組の (μ_0, σ_0) をサンプリングして、それらの値を危険度関数 (4) のパラメタに代入した 100 組の標本関数が図 8A に示されている。これらが推定滑り量で規格化された 10 個の地震発生間隔を対数正規分布ではめた拡張版更新過程モデルで推定される危険度関数のパラツキの有り様を示

す。その上で図 8A の示すところは、既に現在でも平均危険度の 2 倍前後の危険度であることである。これはデータ 1 や 2 のモデルの危険度関数には見られなかった特徴である。図 8B は事後分布の代表値としての最尤推定値とベイズ推定値を代入した危険度関数 (11 式および 12 式参照) とベイズ的予測危険度関数 (13 式参照) を図示している。これらはほぼ似かよっていることがわかる。

図 8A の危険度標本関数に対応する 100 組の累積確率予測関数 (5) が図 9A に示してある。これ

らがデータ 3 から推定された累積確率予測関数の誤差分布を表わしている。同様にして図 8B に対応する代表的な累積確率予測関数が図 9B に示してある。図 9C と 9D はそれぞれ図 8A と 8B に対

応する予測確率の密度関数である。データ 3 と対数正規分布の拡張版更新過程モデルに基づく 2020 年頃を最大に次の南海地震が考えられるが、2000 ~ 2040 年の発生確率は 5 割程度であることがわかる。

VIII. 考 察

有限個のデータによって推定される危険度関数は多かれ少なかれ誤差を伴う。最尤推定値を使う場合は対数尤度関数の最大値付近の二階微分行列 (Hessian matrix) を使う近似的方法 (例えば、井元, 1999) があるが、データ数が少ない場合は V 章や VI 章で示したように事後分布をつかってシミュレーションする方法が有効である。

V 章, VI 章, および VII 章の解析結果は、誤差を考慮したとしても、互いの予測確率の分布パターンに大きな差があることを示している。地震調査推進本部地震調査委員会長期評価部会 (1998, 3.1.1.2 節) および地震予知総合研究振興会 (1999, 6.4 節) はまずデータ 1 と類似のデータと近似尤度で 2070 ~ 2080 年を中心とした確率予測も出しているが、Working Group on California Earthquake Probabilities (1988, 1995) に基いて発生時期予

図 7 対数正規拡張版更新過程モデルのデータ 3 に関する事後分布の等高線。その他は図 1 の説明と同じ。

Fig. 7 Heights of the posterior density distribution in (16) of the extended renewal process with the log-normal distribution for Dataset 3. See the caption of Fig. 1 for details.

図 8 A: 図 7 で示した事後分布からサンプリングした 100 組のパラメタを代入して得られた危険度関数群。B: 図 2 の説明文参照。

Fig. 8 A: One hundred samples of hazard rate functions of time, starting from the occurrence of the last event, simulated from the posterior probability distribution in (16) for Dataset 3. The horizontal and vertical dashed lines are the same as those in Fig. 2A. B: The estimated hazard rate functions are similarly defined as those in Fig. 2B.

図 9 A, B: それぞれ図 8A と図 8B に対応する累積確率関数群。
C, D: それぞれ図 8A と図 8B に対応する確率密度関数群。

Fig. 9 A: Cumulative probability functions of time, starting from the current time (January 2001) derived from the hazard rate functions in Fig. 8A using the first part of the equation in (5). B: The cumulative probability functions based on MLE, Bayes estimate, and Bayesian predictive hazard function in (13). C and D: The probability density functions in (6) corresponding to each sample path of the function in Figs. 8A and 8B, respectively.

測可能モデルからの偏差と対数正規分布による地震発生率の偏差を合わせた偏差として、事前に0.2または0.3と与えることによって昭和の南海地震を起点に確率を計算している。これによると2040年頃を中心に発生が予測されている。また Hori and Oike (1996) は西南日本の内陸地震の活動のトレンドをもとに次の南海地震が2040年前後を中心に発生が予測されるとある(地震予知総合研究振興会, 1998, 6.4.1 節も参照)。これらは凡そ本稿Ⅵ章の結果と調和的である。Ⅶ章の解析結果は相

対的には2020年前後に最も起こり易いかたちになっており、これは名古屋大学理学部(1996)の結果(2006 ± 10 年)ほど差し迫ったものではない。しかも、ここでの確率密度は比較的なだらかで2040年まででも5割程度の発生確率でしかない。またⅤ章とⅦ章のモデルの適合度の優劣のAICによる比較結果やパラメタ $\hat{\theta}$ の大きさは古地震や歴史地震のスリップサイズやマグニチュードの推定をなるべく精度を上げ、モデルの不確定性を小さくする必要があることも示している。

図4Aと4Bのように σ が小さくなる場合は、図5Aと5Bに見られるように、極端に異なる危険度関数の推定が得られる。通常の更新過程をあてはめる場合にも地震の発生間隔が等間隔に近くなるとこの傾向は著しい。小田原付近の歴史地震やSan Andreas断層のParkfield地区の地震などのように、時間間隔が良く揃ったデータが日本や世界で散見されるが、高々4~5の間隔が揃っている場合が殆どである。このようなデータが示す固有地震は、実際に地震間隔が揃うような発生のメカニズムを持っている可能性もあろうが、普通のパラメタ値の更新過程で発生したイベントがたまたま揃った可能性も少なくない(宇津, 1994)。後者の可能性を無視できないのに最尤推定値は常に前者の場合に偏ったパラメタ値になるので要注意である。

IX. ま と め

固有地震と思われるデータに対して更新過程の分布のパラメタが推定されると、それを危険度関数に代入して次の地震発生の確率予測を出す。しかしデータ数が少ないとき最尤推定値を採用すると、予測危険度関数の誤差が大きだけでなく危険度関数や確率予測について偏りが生ずる場合がある。予測危険度関数の誤差や偏りを調べるには尤度関数(事後分布)全体を見る必要がある。本稿ではベイズの推論を採用し、まず南海地震の10個の発生年データに対数正規分布型更新過程をあてはめ、予測危険度関数の誤差分布について議論した。さらに地震に伴った隆起量が与えられている最新の3つの南海地震の発生データにあてはめるために、地震発生間隔を隆起量で規格化したものが同じ対数正規分布に従うと仮定する発生時期予測可能モデルの更新過程の拡張版にあてはめ、その結果の予測危険度関数の誤差(バラツキ)について議論した。さらに最初に解析した10個のイベント(発生年)に対して宇佐美による推定マグニチュードを使って「滑り量」を推定し発生時期予測可能モデルの更新過程拡張版にあてはめた。以上の3種類のデータにあてはめた2種類のモデルは、誤差を考慮しても、それぞれの確

率予測にも互いに無視できない違いが見られた。従って滑り量データの妥当性の議論が重要にならざるを得ない。

文 献

- Akaike, H. (1978): A new look at the Bayes procedure. *Biometrika*, **65**, 53-59.
- Box, G.P.B. and Tiao, G.C. (1973): *Bayesian inference in statistical analysis*. Addison-Wesley, Reading, Massachusetts, 588p.
- Brillinger, D.R. (1982): Seismic risk assessment: Some statistical aspect. *Earthquake Predict. Res.*, **1**, 183-185.
- Daley, D.J. and Vere-Jones, D. (1988): *An introduction to the theory of point processes*. Springer-Verlag, New York.
- Davis, P.M., Jackson, D.D. and Kagan, Y.Y. (1989): The longer it has been since the last earthquake, the longer the expected time till the next? *Bull. Seismol. Soc. Amer.*, **79**, 1439-1456.
- Hori, T. and Oike, K. (1996): A statistical model of temporal variation of seismicity in the Inner Zone of Southwest Japan related to the interplate earthquakes along the Nankai trough. *J. Phys. Earth.*, **44**, 349-356.
- 井元政二郎 (1999): 更新過程における予測確率の誤差. 地震II, **52**, 337-340.
- Kumagai, H. (1996): Time sequence and the recurrence models for large earthquakes along the Nankai trough revisited. *Geophys. Res. Lett.*, **23**, 1139-1142.
- 名古屋大学理学部 (1996): 南海道地震の発生時系列と再来モデルの再考. 地震予知連絡会会報, **56**, 573-578.
- Ogata, Y. (1978): The asymptotic behaviour of maximum likelihood estimators for stationary point processes. *Ann. Inst. Statist. Math.*, **30**, 243-261.
- Ogata, Y. (1999): Estimating the hazard of rupture using uncertain occurrence times of paleoearthquake. *J. Geophys. Res.*, **104**, 17995-18014.
- O'Hagan, A. (1994): *Kendall's advanced theory of statistics. 2B Bayesian inference*. Arnold, London, 332p.
- Rhoades, D.A., Van Dissen, R.J. and Dowrik, D.J. (1994): On the handling of uncertainties in estimating the hazard of rupture on a fault segment. *J. Geophys. Res.*, **99**, 13701-13712.
- 寒川 旭 (1996): 地震考古学 (中公新書1096). 中央公論社, 251p.
- Shimazaki, K. and Nakata, T. (1980): Time-predictable recurrence model for large earthquakes. *Geophys. Res. Lett.*, **7**, 279-282.
- 宇佐美龍夫 (1987): 新編日本被害地震総覧. 東京大学出版会, 435p.
- 宇津徳治 (1994): 時間間隔が揃った数個の地震. 地震II, **47**, 93-95.

宇津徳治(1999): 地震活動総説. 東京大学出版会, 876p.
Working Group on California Earthquake Probabilities (1988): Probability of large earthquakes occurring in California on the San Andreas Fault. *U.S. Geological Survey Open-File Report*, #88-398.
Working Group on California Earthquake Probabilities (1995): Seismic hazards in Southern California: Probable earthquakes, 1994 to 2024. *Bull. Seismol. Soc. Amer.*, **85**, 379-439.
地震調査推進本部地震調査委員会長期評価部会 (1998): (試案)長期的な地震発生確率の評価手法およびその適

用例について. 74p.
地震予知総合研究振興会 (1998): 「南海トラフにおける海溝型巨大地震災害軽減のための地震発生機構のモデル化・観測システムの高度化に関する総合研究」平成9年度科学技術振興調整費成果報告書, 493p.
地震予知総合研究振興会 (1999): 「南海トラフにおける海溝型巨大地震災害軽減のための地震発生機構のモデル化・観測システムの高度化に関する総合研究」平成10年度科学技術振興調整費成果報告書, 287p.

(2001年5月7日受付, 2001年7月6日受理)