情報論的学習理論テクニカルレポート 2009
Technical Report on Information-Based Induction Sciences 2009 (IBIS2009)

点過程を特徴付ける統計量の時間変化を推定する
Estimation of the time-varying parameters characterizing point events

下川 文明* Takeaki Shimokawa 篠本 滋† Shigeru Shinomoto

Abstract:
We selected a set of inter-event interval(IEI) metrics that may efficiently characterize patterns of event occurrences and determined the function that may extract these characteristics. We found that the set of efficient metrics is the mean IEI and the mean log IEI, which represent the rate and the irregularity respectively, and the most suitable function is the gamma distribution function. We constructed Bayes method equipped with the gamma distribution function for estimating the instantaneous rate and irregularity of occurrence for a given event sequence. We confirmed that the Bayes method can capture the instantaneous rate and irregularity reasonably well even when a event sequence is generated from the log-normal and inverse-Gaussian distributions.

Keywords: Bayesian estimation, point process, rate, irregularity

1 まえがき

本研究ではイベントの発生を確率過程，つまり点過程 (point process) としてとらえる．そしてその確率過程の性質をイベントの発生時刻の列 $\{t_i\}_{i=1}^n = \{t_1, t_2, \cdots, t_n\}$ のデータからうまく引き出すために，点過程をよりよく特徴づけするような統計量を考え，さらにそれらの統計量の時間変化を統計推定により捉えることを目的とする．

イベント時系列の例としては，地震，Web ページへのアクセス，経済現象や事件・故障の発生，などが挙げられるが，我々はこれまで神経細胞の出力するスパイク時系列について特に解析を行ってきた．一般に，イベント時系列を特徴付ける量として一番に考えられるのは発生率である．神経科学でもスパイク時系列の発生率に情報がのっているという考えは広く受け入れられてきた (firing-rate code)．しかし，発生率以外の出現パターンにも情報が埋め込まれているということが言われており (spike-timing code)，それを読み取ることも重要であるだろう．発生率以外の特徴量としては不規則性，つまりイベント間間隔 (inter-event interval; IEI) のバラつき，があげられる．図 1 に発生率と不規則性によってイベント時系列を特徴付け，二次元平面で表したものを示す．

*京都大学大学院理学研究科, 606-8502 京都市左京区北白川追分町, tel. 075-753-3763, e-mail shimokawa@ton.scphys.kyoto-u.ac.jp, Department of Physics, Kyoto University, Sakyo-ku, Kyoto 606-8502
†同上, e-mail shinomoto@scphys.kyoto-u.ac.jp,

我々は近年その不規則性に注目して研究を行ってきた．スパイク時系列の不規則性はそれを発する個々の神経細胞毎に異なり，それが発生率変動や時間に対してある程度一定 (図 1, 横方向のみの変化) であることを明らかにした [1]．さらに，運動野では規則的なスパイク時系列が出やすいのに対し，視覚野・前頭前野では不規則なスパイク時系列が出やすいといった，脳の領野毎に傾向の違いがあることを発見した [2]．しかし，同一の神経細胞から出力されるスパイク時系列であっても，実験タスクの内容によって不規則性が大きく変化する (図 1, 縦方向にも変化) という研究報告もなされている [3, 4]．

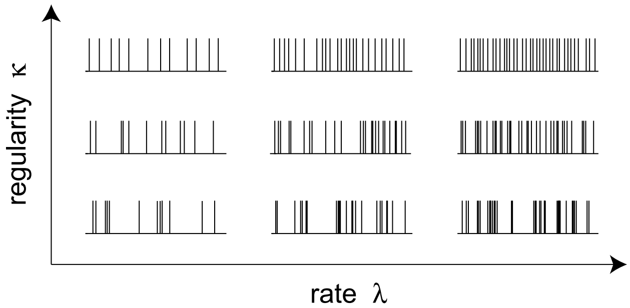

図 1: イベント時系列をラスター図で描き，その特性を発生率と不規則性の二次元平面で表した．横軸が発生率，縦軸が不規則性を表している．

しかし，不規則性についてはイベント間間隔 (IEI) の

バラつきを表す量であるという共通認識はあるものの，そのままIEIの分散として定義するのがよいのかどうか，不規則性を含むIEIの確率分布モデルとしてはどの分布関数を使うべきか，そもそも不規則性よりもうまく点過程を特徴づける量はないのかどうか，といった点についてはあまり議論されてこなかった．さらに，このような不規則性の推定においては，その変動を捉えるためのシステマティックな推定手法は考えられておらず，厳密に不規則性の変動について議論することは難しかった．

そこで本研究では，まず，不規則性がIEI分布を記述する統計量としてどのような位置づけになるのかについて情報理論的に捉え直した．その結果，不規則性は発生率の次に自然にでてくる統計量であること，発生率と不規則性を記述するIEI分布のモデルとしてはガンマ分布を用いるのが一番自然であることを導き出した．また，ベイズ推定法により発火率と不規則性の時間変化を同時に推定できる手法を開発し [5]，シミュレーションによりその性能を確かめた．

2　発生率と不規則性

本章ではまず発生率と不規則性が，情報幾何的に互いに直交するというよい条件を持つ統計量の組として導き出されることを示す．

2.1　本研究の前提

本研究では前提として，イベントの発生が renewal process，つまりイベント間間隔 $T_i = t_{i+1} - t_i$ が確率分布から独立に得られることを仮定する．すると，その確率過程の特徴はIEI分布 $p(T)$ を記述することによって完全に決定される．しかし，IEI分布が時間変化するとし，それをデータのみから捉えるような場合を考えると，IEI分布を完全に記述することは不可能であり，少数のパラメータでうまく記述する必要があるだろう．パラメータは情報幾何的に互いに直交していることが望ましい．直交していれば推定の際にパラメータ間の相関が漸近的になくなるからである [6, 7, 8]．そこで，本研究では発生率 $\lambda = 1/E(T) \equiv 1/\int Tp(T)dt$ に加えて，もう一つの統計量を考えたときに，二つのパラメータが情報幾何的に直交になるような統計量を探すことにする．

また，もう一つの前提として本研究では，IEI分布が時間スケール変換について不変であると仮定する，

$$p(T)dT = g(\Lambda)d\Lambda, \tag{1}$$

ここで，$\Lambda \equiv \lambda T$ である．つまり任意のIEI分布が発生率 λ で時間をスケールし直すと g という分布になるとする．

また，本研究においてIEI分布については最大エントロピー原理 [9] を適用することにする．つまり統計量の条件を満たすものの内，エントロピーが最大の分布を考えることにする．これは，分布を仮定する際，知っている情報以外にどの情報も用いないことを意味し，（エントロピーが最大でない分布を用いることは，何らかの情報により不確実性を減らしていることになる.）無情報事前分布として後に用いるベイズ推定とも関係する．

2.2　直交する統計量の組

上記の条件の下ではIEI分布 $p(T)$ の不確実性を表す量である differential entropy は以下の様に書ける，

$$\begin{aligned} h &= -\int_0^\infty p(T)\log p(T)dT & (2)\\ &= -\log\lambda - \int_0^\infty g(\Lambda)\log g(\Lambda)d\Lambda. & (3)\end{aligned}$$

この式より，確率分布の規格化条件 $\int p(T)dt = 1$，発生率の条件 $E[\Lambda] = 1$ の下で，h を最大化させるIEI分布は，指数分布 $g(\Lambda) = \exp(-\Lambda)$ であり，対応する確率過程はポアソン過程であることがわかる．

さて，では発生率 λ に加えて，ある統計量の値 $E[A(\Lambda)] = \eta$ が得られたものとする．すると，それらの束縛条件の下でエントロピー最大となる分布は，ラグランジュの未定乗数法から導き出せて以下の様に書ける [10]，

$$g(\Lambda) = \exp\left[-\{1 + a + b\Lambda + cA(\Lambda)\}\right]. \tag{4}$$

ここで，λ と η が情報幾何的に直交，つまり Fisher 情報行列が対角化されているという条件を課す，

$$E\left[\frac{\partial^2}{\partial\lambda\partial\eta}\log p(T)\right] = 0. \tag{5}$$

すると，以下の関係式が出てくる，

$$\frac{1}{\lambda}\frac{\partial b}{\partial\eta} + \frac{\partial c}{\partial\eta}E[TA'(\lambda T)] = 0. \tag{6}$$

また，確率分布においては一般に次の等式が成り立つ，

$$E\left[\frac{\partial}{\partial\lambda}\log p(T)\right] = \frac{\partial}{\partial\lambda}\int_0^\infty dTp(T) = 0. \tag{7}$$

この等式を η で微分することで以下の関係式が出てくる，

$$\frac{1}{\lambda}\frac{\partial b}{\partial\eta} + \frac{\partial c}{\partial\eta}E[TA'(\lambda T)] + c\frac{\partial}{\partial\eta}E[TA'(\lambda T)] = 0. \tag{8}$$

式 (6),(8) により，

$$\frac{\partial}{\partial\eta}E[TA'(\lambda T)] = 0. \tag{9}$$

つまり，条件を満たす統計量 $A(\lambda T)$ は

$$A(\lambda T) \propto \log\lambda T, \tag{10}$$

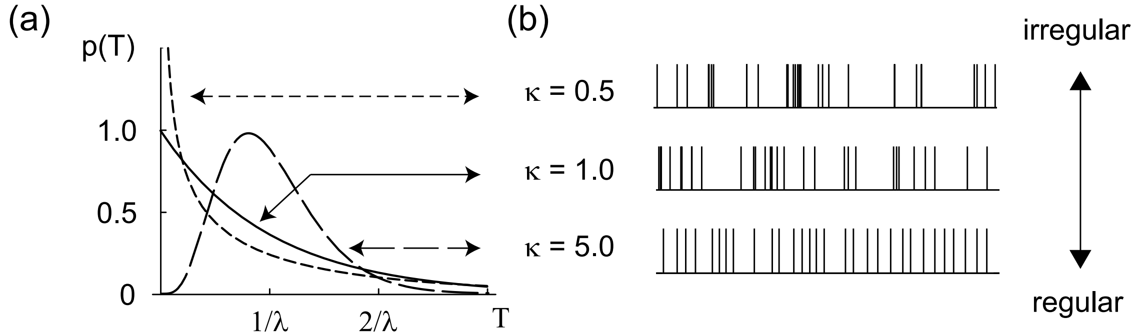

図 2: (a) 不規則性を表すシェイプパラメータを $\kappa = 0.5, 1.0, 5.0$ と変えたときのガンマ分布．(b) それら三種類のイベント間間隔分布から生成したイベント時系列の例をラスター図で表したもの．

に限られることが導き出せた．この統計量は

$$E[\log \lambda T] = \frac{E[(T-\bar{T})^2]}{2\bar{T}^2} - \frac{E[(T-\bar{T})^3]}{3\bar{T}^3} + \cdots, \quad (11)$$

と書き換えるとわかるように，バラつき具合を発生率で規格化したような，IEI 分布の不規則性を表す統計量であることが見て取れる．

これらの統計量の組,$1/E[T]$ と $E[\log \lambda T]$, のもとでエントロピー最大となる分布はガンマ分布である，

$$p(T \mid \lambda, \kappa) = \frac{\lambda^\kappa \kappa^\kappa}{\Gamma(\kappa)} T^{\kappa-1} \exp\left(-\lambda \kappa T\right). \quad (12)$$

ここで，簡単のため $\eta = -\log \kappa + \psi(\kappa)$ と書き直した．図 2 にガンマ分布と，それにより生成したイベント時系列の例を示す．

以上により，発生率（平均値）と情報幾何的に直交である統計量は不規則性（対数平均値）のみであること，それら二つのパラメータのみで表される（エントロピー最大である）分布はガンマ分布であることが示された．

3　時間変化する統計量の推定

本章では発生率と不規則性の時間変動を捉えるためのベイズ推定法 [5] について紹介し，シミュレーションによりその推定精度を確かめる．

3.1　ベイズ推定

本研究では発生率と不規則性の時間変動を捉えるために，イベント時系列は時間変動する 2 つのパラメータ，発生率 $\{\lambda(t)\}$ と不規則性 $\{\kappa(t)\}$(図 3(a))，を持つ確率過程により生じる (図 3(b)) と仮定し，それをベイズ法により推定する (図 3(c)) という方法を取る．

生成の確率モデルとしては，発生率と不規則性が与えられた元での無情報事前分布であるガンマ分布，式 (12), を用いる．ただし今回，それらのパラメータ $\{\lambda(t)\}$, $\{\kappa(t)\}$ は変動するので，イベント時系列 $\{t_i\}_{i=1}^n$ の発生確率は以下のように与えられるものとする，

$$p(\{t_i\}_{i=1}^n \mid \{\lambda(t)\}, \{\kappa(t)\}) = \prod_{i=1}^{n-1} p(T_i \mid \lambda(t_i), \kappa(t_i)). \quad (13)$$

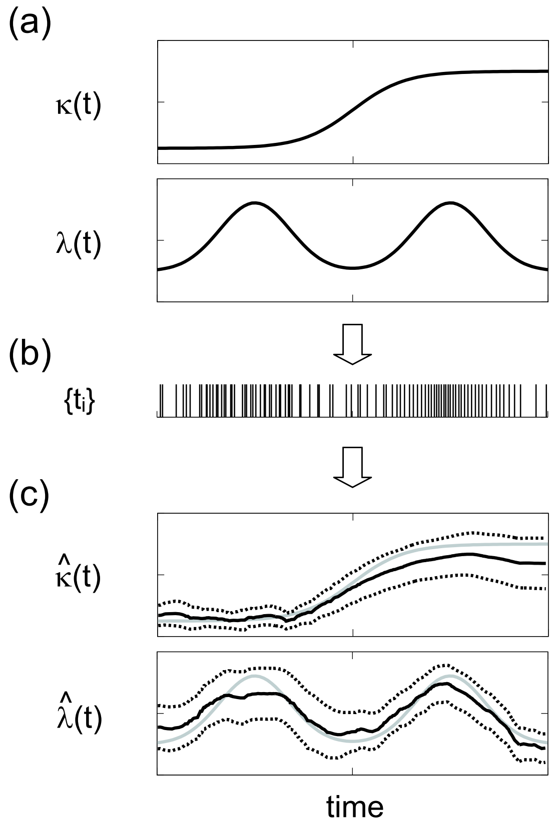

図 3: (a) 発火率 $\{\lambda(t)\}$ と不規則性 $\{\kappa(t)\}$ をパラメータとして持つガンマ過程により (b) イベント時系列を生成した．(c) それをベイズ法により推定した．破線は推定値の 95 %信頼領域を表す．

推定にはベイズ法を用いる．条件付確率 Eq.(13) を逆転させ，イベント時系列 $\{t_i\}_{i=1}^n$ が与えられたときの，

パラメータ $\{\lambda(t)\}$,$\{\kappa(t)\}$ の確率を求める，

$$p(\{\lambda(t)\},\{\kappa(t)\}\,|\,\{t_i\}_{i=1}^n) = \frac{p(\{t_i\}_{i=1}^n\,|\,\{\lambda(t)\},\{\kappa(t)\})\,p(\{\lambda(t)\})\,p(\{\kappa(t)\})}{p(\{t_i\}_{i=1}^n)}. \quad (14)$$

事前分布は Gaussian process prior を導入し，$\lambda(t)$ と $\kappa(t)$ の大きな変動に対してペナルティを課し，オーバーフィッティングを防ぐ，

$$p(\{\lambda(t)\};\gamma_\lambda) = \frac{1}{Z(\gamma_\lambda)}\exp\left[-\frac{1}{2\gamma_\lambda^2}\int_0^T\left(\frac{d\lambda(t)}{dt}\right)^2 dt\right], \quad (15)$$

$$p(\{\kappa(t)\};\gamma_\kappa) = \frac{1}{Z(\gamma_\kappa)}\exp\left[-\frac{1}{2\gamma_\kappa^2}\int_0^T\left(\frac{d\kappa(t)}{dt}\right)^2 dt\right], \quad (16)$$

ここで，γ_λ と γ_κ はそれぞれ変動の程度を表すハイパーパラメータである．これらの値は周辺尤度 $p(\{t_i\}_{i=1}^n;\gamma_\lambda,\gamma_\kappa)$ 最大化により適切な値を求める [11]．周辺尤度最大化のための数値計算には EM 法を用いる [12]．

ハイパーパラメータ決定後，事後分布, 式 (14), を最大化する $\lambda(t)$ と $\kappa(t)$ を推定値 (MAP 推定値) とする．具体的な数値計算の方法に関しては [5] を参照してもらいたい．

3.2　シミュレーション

推定の際には以上の様に，イベント時系列はガンマ分布から生成されたものである，と仮定する．これは 2 章で述べたように，ガンマ分布が発生率と不規則性のみを与えられたときの無情報事前分布だからである [9]．しかし，実際のイベント時系列はガンマ分布とは異なる分布によって生じることがほとんどであるだろう．そのため，イベントがもし異なる分布により発生した場合であっても，ガンマ分布を仮定した推定でロバストに発生率と不規則性を推定できるかどうかを確かめる必要がある．（もちろんイベントを生成する確率過程の関数形がわかっていれば，その関数形を生成モデルとして仮定してベイズ推定を行うのが一番よい．しかし現実的には真の関数形は分からないことがほとんどであり，そのときは無情報事前分布を用いるのが妥当となる．）

そこで，対数正規分布や逆ガウス分布といった，ガンマ分布とは異なる分布においても，以下で定義する発生率 λ と不規則性 κ で分布をパラメトライズし，

$$\lambda = 1/E[T], \quad (17)$$

$$\log\kappa - \psi(\kappa) = \log E[T] - E[\log T]. \quad (18)$$

それらを以下のように時間変動させ，

$$\lambda(t) = 1 + 0.5\sin(t/\tau_1), \quad (19)$$

$$\kappa(t) = 1 + 0.5\sin(t/\tau_2 + \pi/2), \quad (20)$$

イベント時系列を発生させたときに，今回の推定法でどの程度の誤差で推定ができるのかを調べた．

図 4 にその結果を示す．横軸は変動時間 τ_1 or τ_2 の逆数を表し（それぞれもう一つの τ_2 or τ_1 は 100 に固定してある），縦軸は積分二乗誤差 (Integrated squared error; ISE) を表す．図 4 から分かるようにある程度推定がうまくいくことが確かめられた．

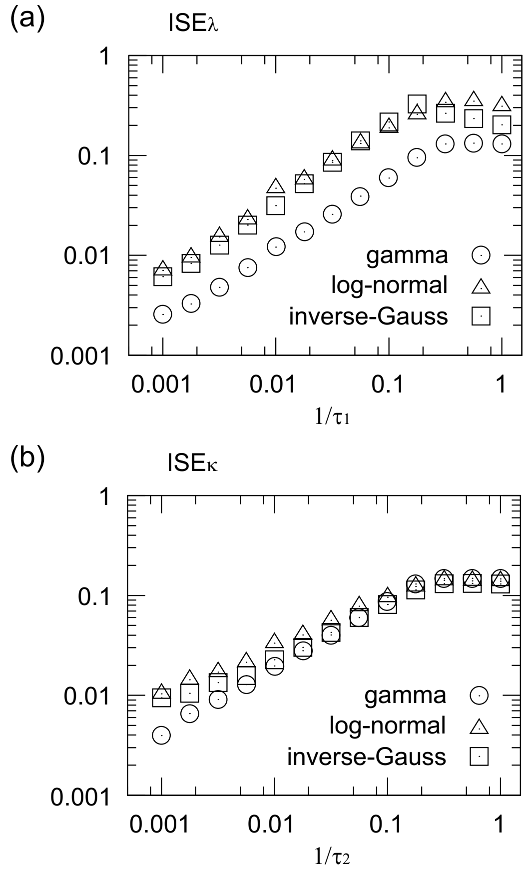

図 4: “○”，“△”，“□”はイベント時系列を生成したモデルがそれぞれ，ガンマ分布，対数正規分布，逆ガウス分布であることを表している．(a) 発生率の推定値の積分二乗誤差 $1/T\int_0^T(\hat{\lambda}(t)-\lambda(t))^2dt$．(b) 不規則性の推定値の積分二乗誤差 $1/T\int_0^T(\hat{\kappa}(t)-\kappa(t))^2dt$.

4　おわりに

イベント時系列を表す統計量として，まず一番に思い浮かぶのは発生率である．本研究では，その発火率に続く二番目の統計量として，互いが情報幾何的に直交するという良い条件を満たす統計量は不規則性であり，それに対応する確率分布はガンマ分布であることを導き出した．

さらに，それら発火率と不規則性の時間変化をイベント時系列のデータのみから推定する方法を紹介し，シ

ミュレーションによりその推定精度を確かめた.

参考文献

[1] S. Shinomoto, K. Shima, and J. Tanji (2003) "Differences in spiking patterns among cortical neurons." Neural Computation, **15**: 2823–2842.

[2] S. Shinomoto, H. Kim, T. Shimokawa et al. (2009) "Relating neuronal firing patterns to functional differentiation of cerebral cortex." PLoS Computational Biology, **5**: e1000433.

[3] R. M. Davies, G. L. Gerstein, and S. N. Baker, (2006) "Measurement of time-dependent changes in the irregularity of neural spiking." Journal of Neurophysiology, **96**: 906–918.

[4] J. F. Mitchell, K. A. Sundberg, and J. H. Reynolds, (2007) "Differential attention-dependent response modulation across cell classes in macaque visual area V4." Neuron, **55**: 131–141.

[5] T. Shimokawa and S. Shinomoto (2009) "Estimating instantaneous irregularity of neuronal firing." Neural Computation, **21**: 1931–1951.

[6] S. Amari and H. Nagaoka (2000) *Methods of information geometry.* Oxford: Oxford University Press.

[7] K. Ikeda (2005) "Information geometry of interspike intervals in spiking neurons." Neural Computation, **17**: 2719–2735.

[8] K. Miura, M. Okada, and S. Amari (2006) "Estimating spiking irregularities under changing environments." Neural Computation, **18**: 2359–2386.

[9] E. T. Jaynes (1957) "Information theory and statistical mechanics." Physical Review, **106**: 620–630.

[10] J. N. Kapur (1989) *Maximum-entropy models in science and engineering.* New York: Wiley.

[11] D. J. C. MacKay (1992) "Bayesian interpolation." Neural Computation, **4**: 415–447.

[12] A. P. Dempster, N. M. Laird, and D. B. Rubin (1977) "Maximum likelihood from incomplete data via the EM algorithm." Journal of the Royal Statistical Society, Series B, **39**: 1–38.