

BPT 分布と時間予測モデルの接点

寺田吉彦

2025 年 1 月 31 日

1 逆ガウス分布の定義

逆ガウス分布は, $(0, \infty)$ 上の連続型確率分布である. ガウス分布に従う確率変数の逆数の従う分布とは異なるので注意が必要である. 逆ガウス分布は, 以下の確率密度関数によって定義され, 2 種類の parametrization がある.

$$f(t | \mu, \alpha) = \sqrt{\frac{\mu}{2\pi\alpha^2 t^3}} \exp \left\{ -\frac{(t - \mu)^2}{2\alpha^2 \mu t} \right\} \quad (t > 0, \mu > 0, \alpha > 0) \quad (1.1)$$

$$f(t | \theta, \beta) = \sqrt{\frac{\beta}{2\pi t^3}} \exp \left\{ -\frac{\beta t}{2} \left(\theta - \frac{1}{t} \right)^2 \right\} \quad (t > 0, \theta > 0, \beta > 0) \quad (1.2)$$

$$f(t | \mu, \lambda) = \sqrt{\frac{\lambda}{2\pi t^3}} \exp \left\{ -\frac{\lambda(t - \mu)^2}{2\mu^2 t} \right\} \quad (t > 0, \mu > 0, \lambda > 0) \quad (1.3)$$

式 (1.1) によって定義した逆ガウス分布を $\text{IG}(\mu, \alpha^2)$ または $\text{IG}_1(\mu, \alpha)$, 式 (1.2) によって定義した逆ガウス分布を $\text{IG}_2(\theta, \beta)$ と記す. この他にも様々な parametrization がある. 例えば, R package の `statmod` では式 (1.3) のパラメトライズを用いており, λ を shape parameter と呼んでいる. 式 (1.1) で定義した場合, 逆ガウス分布 $\text{IG}_1(\mu, \alpha)$ の平均と分散は, μ と $(\mu\alpha)^2 = \mu^3/\lambda$ となる. よって, μ は平均, α は変動係数 (coefficient of variation) を表している. 2 つの parametrization の関係は以下の通りである.

$$\mu = 1/\theta, \quad \alpha = \sqrt{\beta\theta} = \sqrt{\mu/\lambda}, \quad \lambda = \frac{\mu}{\alpha^2}$$

2 逆ガウス分布とブラウン運動

逆ガウス分布は, W_t を標準ブラウン運動として, 以下の確率過程の first passage time の分布として特徴付けられる.

$$X_0 = 0, \quad X_t = \eta t + \sigma W_t$$

ある閾値 τ に初めて到達するまでにかかる時間 (first passage time) は, 以下の逆ガウス分布に従う.

$$T_\tau := \inf\{t > 0 \mid X_t = \tau\} \sim \text{IG} \left(\frac{\tau}{\eta}, \frac{\sigma^2}{\eta\tau} \right) = \text{IG}_1 \left(\frac{\tau}{\eta}, \frac{\sigma}{\sqrt{\eta\tau}} \right) = \text{IG}_2 \left(\frac{\eta}{\tau}, \frac{\sigma^2}{\eta^2} \right)$$

3 時間予測モデルと BPT モデルの融合 (SSD-BPT model; Ogata, 2002)

隆起量（滑り量）の情報を組み込んだモデルを考える．逆ガウス分布とブラウン運動の関係において、閾値 τ が隆起量 u に依存しているモデルを考える．

時間予測モデルの定式化を思い出す．以下の記号を用いる．

- U_i : i 番目の地震発生時の隆起量
- O_i : i 番目の地震発生時刻 ($O_0 = 0$)
- $T_i = O_i - O_{i-1}$: 地震発生間隔

時間予測モデルでは, U_{i-1} が観測された下で, $i-1$ 番目の活動から i 番目の活動までの条件付き期待経過時間 $\mathbb{E}[T_i | U_{i-1}]$ に対して, 以下のようなモデルを考えている．

$$\mathbb{E}[T_i | U_{i-1}] = \frac{U_{i-1}}{\nu} = \beta U_{i-1}, \quad \beta := \frac{1}{\nu}$$

すなわち, 切片項が 0 の単回帰モデルとなっている．

時間予測モデルと BPT モデルの整合性を考えると, 閾値 τ と隆起量 u は比例関係にあると考えることが妥当である．

$$\beta u_{i-1} = \mathbb{E}[T_i | U_{i-1} = u_{i-1}] = \mathbb{E}[T_{\tau_i}] = \frac{\tau_i}{\eta}$$

そのため, 隆起量 U が閾値 τ と比例関係にあると仮定する．

$$u_i = a\tau_i$$

すなわち, η と σ は活動によらず一定で, 閾値だけが隆起量によって変化するというモデルを考える． T_τ の期待値が τ/η であることに注目することで, 以下の関係を考えることができる．

$$\beta u_{i-1} = \mathbb{E}[T_i | U_{i-1} = u_{i-1}] = \mathbb{E}[T_{\tau_i}] = \frac{\tau_i}{\eta} \Rightarrow \beta a = \frac{1}{\eta}$$

$\tau_i = \eta\beta u_{i-1}$ であるから, 以下の条件付きモデルを自然に導出できる．

$$\frac{\sigma^2}{\eta\tau_i} = \frac{\sigma^2}{\eta^2\beta u_{i-1}} = \beta \frac{(a\sigma)^2}{u_{i-1}} \Rightarrow T_i | U_{i-1} = u_{i-1} \sim \text{IG}\left(\beta u_{i-1}, \beta \frac{\gamma^2}{u_{i-1}}\right), \quad \gamma := a\sigma$$

パラメータの意味を考えると, 大きい隆起量（地震）ほど変動係数が小さくなるというモデリングになっている．

T_i の条件付き期待値と条件付き分散は, それぞれ以下で与えられる．

$$\mathbb{E}[T_i | U_{i-1} = u_{i-1}] = \beta u_{i-1}, \quad \text{Var}[T_i | U_{i-1} = u_{i-1}] = \beta^3 \gamma^2 u_{i-1}$$

このモデルは, Ogata (2002, JGR) では SSD-BPT モデルと呼ばれている．ここで, SSD は Slip-Size Dependent である．

4 周辺モデルも利用した推定（新たな提案手法, 不確実性の考慮なし）

記号の煩雑さを避けるため, 以下では T_i と U_{i-1} の index を i に統一する．隆起量が得られていないデータの利用を考える．データが少ないため, パラメトリックなモデルを考える．隆起量が互いに独立に区間 $[0, \infty)$ 上の切断正規分布 $\text{TN}(\mu_u, \sigma_u^2, 0, \infty)$ に従うと仮定する．そして, 以下のようなモデルを考える．

$$(T_i, U_i) \sim \text{IG}\left(\beta U_i, \frac{\beta \gamma^2}{U_i}\right) \times \text{TN}(\mu_u, \sigma_u^2, 0, \infty)$$

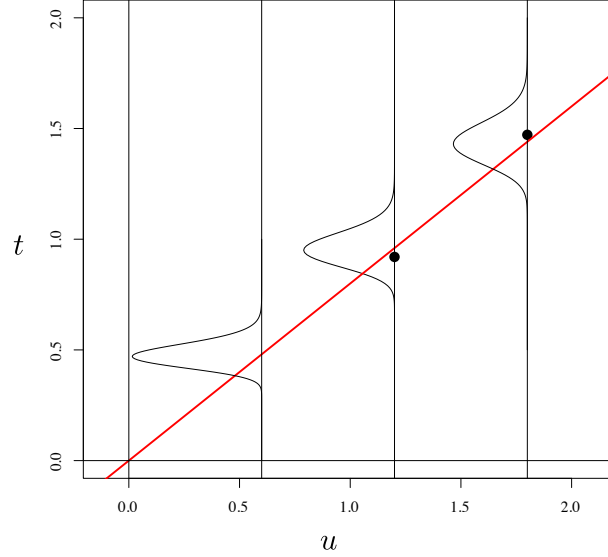


図 3.1: 時間予測 BPT モデルのイメージ図

他にもガンマ分布によるモデリングも考えられる（ガンマ-時間予測 BPT モデルと呼ぶ）。

$$(T_i, U_i) \sim \text{IG} \left(\beta U_i, \frac{\beta \gamma^2}{U_i} \right) \times \text{Gam}(\alpha_u, \lambda_u)$$

そして、周辺尤度の最大化を考えることで、隆起量が観測されていないデータも考慮した解析を考えることができる^{*1}。

$$f_T(t; \theta, \xi) = \int f_{T|U}(t | u; \theta) f_U(u; \xi) du, \quad \theta = (\beta, \gamma), \quad \xi = (\mu_u, \sigma_u^2) \text{ or } (\alpha_u, \lambda_u)$$

積分自体は、1次元なのでここでは一旦数値積分を用いて、周辺尤度を計算する。EM algorithm による最適化も考えられるが、計算が容易となるとは限らない。

従って、隆起量が観測されていないデータ $T_1, \dots, T_m$ と隆起量と時間間隔が両方とも観測されているデータ $(U_{m+1}, T_{m+1}), \dots, (U_{m+n}, T_{m+n})$ がそれぞれ周辺分布と同時分布から互いに独立に得られているとする。このとき、これらの観測データに対する尤度は以下で与えられる。

$$L(\theta, \xi | \mathcal{D}(m, n)) = \prod_{i=1}^m f_T(T_i; \theta, \xi) \times \prod_{j=1}^n f_U(U_{m+j}; \theta, \xi) \times \prod_{j=1}^n f_{T|U}(T_{m+j} | U_{m+j}; \theta)$$

1361 年 8 月 3 日の南海地震以降のデータを用いて、解析してみる。それ以前のデータでは、発生間隔が著しく異なるため、分布の同一性が崩れると判断した。上述の尤度関数に基づき隆起量の不確実性を考慮せずに最尤推定を適用すると以下のような結果が得られた。

- $\hat{\beta}_{\text{TN}} = 0.7986$, $\hat{\gamma}_{\text{TN}} = 0.0637$, $\hat{\mu}_u = 1.319$, $\hat{\sigma}_u^2 = 0.537^2$
- $\hat{\beta}_{\text{Gam}} = 0.7994$, $\hat{\gamma}_{\text{Gam}} = 0.0637$, $\hat{\alpha}_u = 33.51$, $\hat{\lambda}_u = 23.67$

例えば、ガンマ-SSD-BPT モデルを考えて、上述のパラメータを用いて実験し、従来の BPT モデルを当てはめてみる。具体的には、 $n = 5 \times 10^5$ の時間間隔を発生させ、そのデータを用いた最尤推定によって BPT モデルを当てはめた結果を以下に記す。発生間隔のみのデータを用いた場合、 $\alpha = 0.2$ が推定されることと整合している。

*1 最適化が難しいという問題点がある

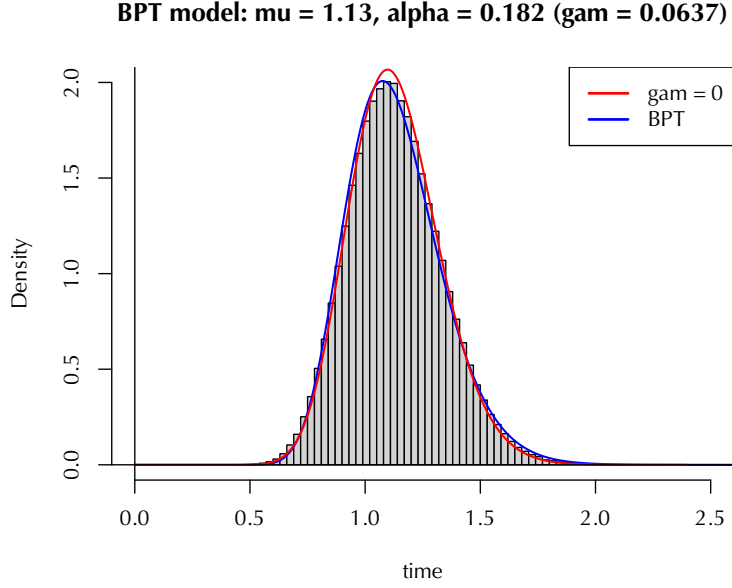


図 4.1: ガンマ-SSD-BPT モデルの周辺分布と BPT モデルによる近似

4.1 条件付きモデルの γ と周辺モデルの変動係数 α の関係

γ と α の関係について、混乱を招くため、それらの違いについて正確に述べる。まず、SSD-BPT モデルは隆起量を条件付けた下での条件付きモデルであるにも関わらず、隆起量のない場合の BPT 分布として解釈している点が理解の妨げとなっている。この点を考慮せず、以下の関係式だけに注目すると、 γ を小さく見積もることに抵抗感が生じてしまう。

$$\frac{\sqrt{\text{Var}(T | U = u_i)}}{\mathbb{E}[T | U = u_i]} = \alpha_i = \sqrt{\frac{\beta}{u_i}} \gamma, \quad \text{Var}(T | U = u) = \beta^3 \gamma^2 u, \quad \mathbb{E}[\text{Var}(T | U)] = \beta^3 \gamma^2 \mathbb{E}[U]$$

隆起量を固定した場合は状況を限定した場合、当然そのばらつきは、平均的に隆起量を考えない周辺分布のばらつきよりも小さくなるのが妥当であろう。つまり、上述の条件付きの変動係数 α_i は、極端な隆起量でない限り、周辺分布で推定された変動係数 $\alpha = 0.2$ よりも小さくなる方が自然である。この議論を踏まえて、 γ の事前分布の設計を行わなければ、不適切な解析結果を導くこととなる。特に、これまでの隆起量を考えない周辺分布に対する α に関する情報を上述の式から naive に変換して、 γ の事前分布を設計すると、大幅な過大評価となる可能性がある。例えば、以下の $\bar{\alpha}$ を周辺分布に対する α と考えて、変換すると過大評価となる可能性がある。

$$\bar{\alpha} = \frac{1}{n} \sum_{i=1}^n \alpha_i = \gamma \frac{1}{n} \sum_{i=1}^n \sqrt{\frac{\beta}{u_i}}$$

もう少し数学的にこの現象を説明する。上述は変動係数ではあるが、まず分散を考える。よく知られた事実として、条件付き分散の期待値は周辺分散よりも必ず小さくなる。

$$\begin{aligned} \text{Var}(T) &= \mathbb{E}[(T - \mathbb{E}[T])^2] = \mathbb{E}_U[\text{Var}(T | U)] + \mathbb{E}_U[(\mathbb{E}[T | U] - \mathbb{E}[T])^2] \\ &\geq \mathbb{E}_U[\text{Var}(T | U)] \end{aligned}$$

変動係数に対しては、このようなシンプルな計算では示すことができないが、本質的には同じ理由で過大評価の可能性はある。 $\bar{\alpha}$ は条件付き変動係数の期待値の推定量であり、周辺分布の変動係数とは別物であることが原因となっている。

例えば, ガンマ-SSD-BPT モデルの場合を考えてみよう. 周辺期待値と周辺分散は, $\mathbb{E}[T] = \beta\alpha_u/\lambda_u$ と

$$\begin{aligned}\text{Var}(T) &= \mathbb{E}_U [\text{Var}(T | U)] + \mathbb{E}_U [(\mathbb{E}[T | U] - \mathbb{E}[T])^2] \\ &= \beta^3\gamma^2\mathbb{E}[U] + \beta^2\text{Var}(U) = \frac{\beta^3\gamma^2\alpha_u}{\lambda_u} + \frac{\beta^2\alpha_u}{\lambda_u^2} = \frac{(\beta\gamma^2\lambda_u + 1)\beta^2\alpha_u}{\lambda_u^2}\end{aligned}$$

で与えられる. すなわち, 周辺分布の変動係数は,

$$\frac{\sqrt{\text{Var}(T)}}{\mathbb{E}[T]} = \frac{\beta\sqrt{(\beta\gamma^2\lambda_u + 1)\alpha_u}}{\lambda_u} \times \frac{\lambda_u}{\beta\alpha_u} = \sqrt{\frac{\beta\gamma^2\lambda_u + 1}{\alpha_u}}$$

となる.

次に, 条件付き変動係数の期待値を考えてみよう. Jensen の不等式より,

$$\mathbb{E} \left[\frac{\sqrt{\text{Var}(T | U)}}{\mathbb{E}[T | U]} \right] \leq \sqrt{\mathbb{E} \left[\frac{\text{Var}(T | U)}{\mathbb{E}[T | U]^2} \right]}$$

が成り立つことに注意する. 一方, 逆ガンマ分布の期待値より ($\alpha_u > 1$),

$$\mathbb{E} \left[\frac{\text{Var}(T | U)}{\mathbb{E}[T | U]^2} \right] = \beta\gamma^2\mathbb{E} \left[\frac{1}{U} \right] = \frac{\beta\gamma^2\lambda_u}{\alpha_u - 1}$$

を得る. したがって, 以下を得る.

$$\bar{\alpha} = \frac{1}{n} \sum_{i=1}^n \alpha_i \approx \mathbb{E} \left[\frac{\sqrt{\text{Var}(T | U)}}{\mathbb{E}[T | U]} \right] \leq \sqrt{\frac{\beta\gamma^2\lambda_u}{\alpha_u - 1}} \approx 0.049, \quad \frac{\sqrt{\text{Var}(T)}}{\mathbb{E}[T]} = \sqrt{\frac{\beta\gamma^2\lambda_u + 1}{\alpha_u}} \approx 0.179$$

数値は, 上述の推定結果 ($\hat{\beta}_{\text{Gam}} = 0.7994$, $\hat{\gamma}_{\text{Gam}} = 0.0637$, $\hat{\alpha}_u = 33.51$, $\hat{\lambda}_u = 23.67$) を当てはめた場合であり, 周辺分布の変動係数を見せると大幅な過大評価となることがわかる. また, 上述の計算では簡単のため, Jensen の不等式より上界の評価を行ったが, 条件付き変動係数の正確な期待値は

$$\mathbb{E} \left[\frac{\sqrt{\text{Var}(T | U)}}{\mathbb{E}[T | U]} \right] = \frac{\Gamma(\alpha_u - 1/2)}{\Gamma(\alpha_u)} \sqrt{\beta\gamma^2\lambda_u}$$

で与えられる.