数理解析研究所講究録
第 1916 巻 2014 年 186-195

不均衡データと変形指数型分布族

慶應義塾大学・理工学部　清 智也
Tomonari Sei
Faculty of Science and Technology,
Keio University

概要

Sei (2014) では，不均衡データに対して二項回帰モデルを適用すると，極限として，変形指数型分布族（q-指数型分布族）の形の強度測度を持つポアソン点過程が得られることを示した．本稿では，この結果の簡単な復習をした後，ポアソン点過程モデルの情報幾何構造，特にダイバージェンスについて考察する．

キーワード 二項回帰，ダイバージェンス，不均衡データ，ポアソン点過程，q-指数型分布族．

1 はじめに

$\{(X_i, Y_i)\}_{i=1}^m$ を，$\mathbb{R}^p \times \{0,1\}$ 上の独立同一分布に従う m 個の確率変数とする．また，X_i の周辺分布，および X_i を条件づけた下での Y_i の条件付き分布をそれぞれ

$$\begin{aligned} &P(X_i \in dx) = F(dx), \\ &P(Y_i = 1 \mid X_i, a, b) = G(a + b^T X_i), \quad a \in \mathbb{R}, \quad b \in \mathbb{R}^p, \end{aligned} \tag{1}$$

と仮定する．ここで，$G(\cdot)$ は 1 次元の累積分布関数であり，その逆関数 $G^{-1}(p) = \sup\{z : G(z) \le p\}$ は一般化線形モデルにおけるリンク関数である．以下，式 (1) の条件付き分布族を二項回帰モデルと呼ぶ．また，a, b を回帰係数といい，G を逆リンク関数という．

逆リンク関数 G としてよく使われるのはロジスティック分布

$$G(x) = \frac{e^x}{e^x + 1}$$

である．この場合，式 (1) は $(Y_i, Y_i X_i)$ を十分統計量とする指数型分布族となる．

本稿では，$Y_i = 1$ となるデータ（正例）がほとんどない場合を考察する．モデルの言葉で言い直すと，式 (1) の条件付き確率がほとんどゼロの場合を考える．そのような状況が生ずる例として，不正検出，医療診断，政策分析などがある (Bolton and Hand, 2002; Chawla et al., 2004; Jin et al., 2005; King and Zeng, 2001).

以下では,「確率がほとんどゼロ」という極限と,「データが増える」という極限を同時に扱うため，三角列を考える．すなわち，$\{(X_i, Y_i)\}_{i=1}^m$ は独立同一分布に従うが，その分布自体は m に応じて変化してもよいと考える．このような設定として，ポアソンの少数の法則がよく知られている：もしある実数 $\lambda > 0$ が存在して

$$P(Y_i = 1) = \lambda/m + o(m^{-1}), \quad i = 1, \ldots, m, \tag{2}$$

ならば，$\sum_{i=1}^{m} Y_i$ は $m \to \infty$ のもとで平均 λ のポアソン分布に分布収束する．以後，式 (2) が満たされるような極限の取り方を不均衡極限と呼ぶことにする．

さて，不均衡極限を考える上では，回帰係数 (a,b) は m に依存してよいことになる．これを (a_m, b_m) と書く．一方で，やや不自然な仮定ではあるが，X_i の周辺分布 F と逆リンク関数 G は m に依存しないと仮定する．この仮定を外すことは今後の課題である．

Warton and Shepherd (2010) は，ロジスティック回帰モデルが，不均衡極限のもとで，ポアソン点過程モデルに分布収束することを示した．Owen (2007) や Baddeley et al. (2010) も関連した結果を得ている．これらの結果は大まかには以下のように示すことができる．

モデル (1) でロジスティック分布 $G(z) = e^z/(1+e^z)$ を用いた場合を考える．固定した $\alpha \in \mathbb{R}$ と $\beta \in \mathbb{R}^p$ に対し，回帰係数を $a_m(\alpha) = -\log m + \alpha$，$b_m(\beta) = \beta$ とおく．すると，$m \to \infty$ の下で

$$P(Y_i = 1 \mid X_i, a_m(\alpha), b_m(\beta)) = \frac{e^{-\log m + \alpha + \beta^T X_i}}{1 + e^{-\log m + \alpha + \beta^T X_i}} = \frac{e^{\alpha + \beta^T X_i}}{m} + o(m^{-1}) \qquad (3)$$

が成り立つ．さらにベイズの定理から，$Y_i = 1$ の下での X_i の（F に対する）条件付き密度は，形式的に，

$$\frac{e^{\beta^T X_i}}{\int e^{\beta^T x} F(dx)} + o(1) \qquad (4)$$

と計算される．これは十分統計量を X_i とする指数型分布族である (Owen, 2007)．

また，式 (3) より，$\mathbb{R}^p$ の任意のコンパクト集合 A に対して，確率 $P(Y_i = 1, X_i \in A)$ は近似的に $m^{-1} \int_A e^{\alpha + \beta^T x} F(dx)$ と表される．よって，ポアソンの少数の法則より，$Y_i = 1$ かつ $X_i \in A$ となるような観測値の個数は，平均 $\int_A e^{\alpha + \beta^T x} F(dx)$ のポアソン分布に近似的に従うことが分かる．これは強度測度（intensity measure）$e^{\alpha + \beta^T x} F(dx)$ のポアソン点過程である．ポアソン点過程については例えば間瀬・武田 (2001) を参照されたい．

Sei (2014) は，ロジスティック回帰モデル以外の二項回帰モデルの不均衡極限を考えた．ロジスティック回帰に対する結果と同様，その極限はポアソン点過程になる．注目すべき点は，収束先の点過程の強度測度が，一般に q-指数型分布族と呼ばれるクラスになることである．q-指数型分布族とは，変形指数型分布族，あるいは α-分布族とも呼ばれ，実数 q を使って特徴づけられる確率分布族であり，統計物理や情報幾何において近年脚光を浴びている（Amari (1985); Amari and Nagaoka (2000); Amari and Ohara (2011); 松添 (2013); Naudts (2002, 2010); Tsallis (1988)）．この結果は 2 節でレビューされる．

また 3 節では，Sei (2014) で詳しく触れなかった話題として，ポアソン点過程モデルにおけるダイバージェンスについて考察する．

2 二項回帰の不均衡極限

実数 q に対し，q-指数関数を

$$\exp_q(z) = \begin{cases} e^z, & \text{if } q = 1, \\ [1 + (1-q)z]_+^{1/(1-q)}, & \text{if } q \neq 1, \end{cases} \tag{5}$$

により定義する．ここで，$[z]_+ = \max(z, 0)$, $[0]_+^{-1} = \infty$ と約束する．この変換はパラメータ $\lambda = 1 - q$ の Box-Cox 変換の逆変換に他ならない．特に，$q < 1$ かつ $z \leq -1/(1-q)$ のとき $\exp_q(z) = 0$ であり，$q > 1$ かつ $z \geq -1/(1-q)$ のとき $\exp_q(z) = \infty$ となる．関数 $\exp_q(z)$ は $q \geq 0$ のとき，またそのときに限り凸関数である．

さて，二項回帰モデル (1) において，G に関する次の仮定を設ける．

仮定 1. ある $q > 0$, $c_m \in \mathbb{R}$ および $d_m > 0$ が存在し，各 $z \in \mathbb{R}$ に対して

$$G(c_m + d_m z) = \frac{1}{m}\exp_q(z) + o(m^{-1}), \quad m \to \infty \tag{6}$$

が成り立つ．

極値理論によれば，式 (6) 以外の漸近形は存在しない（例えば de Haan and Ferreira (2006, Theorem 1.1.2 and 1.1.3)）．実数 q は G の左裾の構造を決定している．例えばロジスティック分布は，仮定 1 を満たし，$q = 1$, $c_m = -\log m$, $d_m = 1$ である．その他の例については表 1 にまとめておく．

仮定 1 の c_m, d_m を用いて，$(\alpha, \beta) \in \mathbb{R} \times \mathbb{R}^p$ に対して

$$a_m(\alpha) = c_m + d_m\alpha \quad \text{and} \quad b_m(\beta) = d_m\beta \tag{7}$$

と定義する．また，真の回帰係数が $(a_m(\alpha), b_m(\beta))$ のときの $\{(X_i, Y_i)\}_{i=1}^m$ の確率法則を $P_{m,\alpha,\beta}$ と記す．

さて，式 (3) の類推が仮定 1 から得られる．実際，

$$\begin{aligned} P_{m,\alpha,\beta}(Y_i = 1 \mid X_i) &= G(a_m(\alpha) + b_m(\beta)^T X_i) \\ &= G(c_m + d_m(\alpha + \beta^T X_i)) \\ &= \frac{1}{m}\exp_q(\alpha + \beta^T X_i) + o(m^{-1}) \end{aligned}$$

となる．よって，ロジスティック回帰のときと同様，二項回帰モデルはポアソン点過程に収束することが期待される．これを以下示す．

主結果を述べる前に次の補題を用意する．

補題 1. $(\alpha, \beta) \in \mathbb{R} \times \mathbb{R}^p$ とする．また，A を $\mathbb{R}^p$ のコンパクト集合とし，$\forall x \in A$ に対し $\exp_q(\alpha + \beta^T x) < \infty$ と仮定する．このとき次の式が成り立つ：

$$P_{m,\alpha,\beta}(Y_i = 1, X_i \in A) = \frac{\lambda(A)}{m} + o(m^{-1}). \tag{8}$$

ただし $\lambda(A) = \int_A \exp_q(\alpha + \beta^T x) F(dx)$ とする．

表 1: 仮定 1 を満たす分布の代表例と，対応する q，および c_m, d_m の例を示す．定数 $C_{\kappa\lambda}$, C_κ および C^*_κ は分布の正規化定数である．l_m は $\log m$ の略である．GPD は Generalized Pareto distribution の略である．また，関数 $\gamma_q(z)$ は方程式 $\exp_q(z-\gamma_q(z))+\exp_q(-\gamma_q(z))=1$ の一意解として定義される (Ding et al., 2011).

Name	Distribution $G(z)$	q	c_m	d_m
logistic	$e^z/(1+e^z)$	1	$-\log m$	1
normal	$\int_{-\infty}^z (2\pi)^{-1/2}e^{-u^2/2}du$	1	$-(2l_m)^{1/2}+\frac{\log(l_m)+\log(4\pi)}{2(2l_m)^{1/2}}$	$(2l_m)^{-1/2}$
Gumbel	$1-\exp(-e^z)$	1	$-\log m$	1
beta	$C_{\kappa\lambda}\int_0^z u^{\kappa-1}(1-u)^{\lambda-1}du$	$1-\kappa^{-1}$	$\kappa^{1/\kappa}(mC_{\kappa\lambda})^{-1/\kappa}$	c_m/κ
gamma	$C_\kappa\int_0^z u^{\kappa-1}e^{-u}du$	$1-\kappa^{-1}$	$\kappa^{1/\kappa}(mC_\kappa)^{-1/\kappa}$	c_m/κ
Weibull	$1-\exp(-z^\kappa)$	$1-\kappa^{-1}$	$\kappa^{1/\kappa}m^{-1/\kappa}$	c_m/κ
Student	$C^*_\kappa\int_{-\infty}^z(1+u^2/\kappa)^{-(\kappa+1)/2}du$	$1+\kappa^{-1}$	$-\kappa^{-1/\kappa}(mC^*_\kappa)^{1/\kappa}$	$-c_m/\kappa$
Fréchet	$1-\exp(-(-z)^{-\kappa})$	$1+\kappa^{-1}$	$-\kappa^{-1/\kappa}m^{1/\kappa}$	$-c_m/\kappa$
GPD	$\exp_q(z)$	q	$(-1+m^{q-1})(1-q)^{-1}$	m^{q-1}
t-logistic	$\exp_q(z-\gamma_q(z))$, $q>1$	q	$-m^{q-1}(q-1)^{-1}$	m^{q-1}
	$\exp_q(z-\gamma_q(z))$, $0<q<1$	q	$(-1+m^{q-1})(1-q)^{-1}$	m^{q-1}
	$\exp_q(z-\gamma_q(z))$, $q<0$	0	$-(1-q)^{-1}+m^{-1}$	m^{-1}

証明．$t:=\alpha+\beta^TX_i$ の確率分布を $F^*(dt)$ とおく．また，$A^*=\{\alpha+\beta^Tx\mid x\in A\}$ と定義する．仮定より，A^* はコンパクトである．このとき，

$$P_{m,\alpha,\beta}(Y_i=1,X_i\in A)=\int_A G(a_m(\alpha)+b_m(\beta)^Tx)F(dx)$$
$$=\int_A G(c_m+d_m(\alpha+\beta^Tx))F(dx)$$
$$=\int_{A^*} G(c_m+d_mt)F^*(dt)$$

と書ける．式 (8) を示すには，

$$\int_{A^*} G(c_m+d_mt)F^*(dt)=\frac{1}{m}\int_{A^*}\exp_q(t)F^*(dt)+o(m^{-1})$$

を言えば十分である．仮定 1 より，各 $t\in A^*$ に対して $mG(c_m+d_mt)=\exp_q(t)+o(1)$ となる．よって，$mG(c_m+d_mt)$ が $t\in A^*$ について一様に $\exp_q(t)$ に収束することを示せばよい．ところが，$mG(c_m+d_mt)$ は t について単調であり，かつ $\exp_q(t)$ は $t\in A^*$ について連続であるから，この一様収束性は一般論から導かれる（例えば Galambos (1987, Lemma 2.10.1)）． □

データ $\{(X_i,Y_i)\}_{i=1}^m$ に対して，

$$N_m(A)=\sharp\{i\mid X_i\in A,\ Y_i=1\},\quad A\subset\mathbb{R}^p,$$

によって点過程 N_m を定義する．この式は，集合 A に属すような X_i のうち，$Y_i=1$ となるようなものの個数を表している．

定理 1. $P_{m,\alpha,\beta}$ の下で，点過程 N_m は次の強度測度を持つポアソン点過程 N に法則収束する：

$$\lambda(dx) = \exp_q(\alpha + \beta^T x)F(dx). \tag{9}$$

正確には，等式

$$\begin{aligned}&\lim_{m\to\infty} P_{m,\alpha,\beta}\left(N_m(A_j) = \nu_j,\ j = 1,\ldots,J\right)\\ &= P\left(N(A_j) = \nu_j,\ j = 1,\ldots,J\right) = \prod_{j=1}^{J} \frac{\lambda(A_j)^{\nu_j} e^{-\lambda(A_j)}}{\nu_j!}\end{aligned} \tag{10}$$

が，任意の正整数 J，非負整数 ν_j，互いに排反なコンパクト集合 $A_j \subset \mathbb{R}^p$ （で $\exp_q(\alpha + \beta^T x) < \infty, x \in A_j$, を満たすもの）に対して成り立つ．

式 (10) は，Embrechts et al. (1997) にある点過程の法則収束の定義と整合的である．

定理 1 の証明. $\{A_j\}_{j=1}^J$ を互いに排反なコンパクト集合とする．$\{(X_i, Y_i)\}_{i=1}^m$ は独立同一分布に従うので，確率変数 $\{N_m(A_j)\}_{j=1}^J$ の同時分布は

$$P_{m,\alpha,\beta}\left(N_m(A_j) = \nu_j,\ 1 \le j \le J\right) = \prod_{j=1}^{J} (p_{m,j})^{\nu_j} \left(1 - \sum_j p_{m,j}\right)^{m - \sum_j \nu_j}$$

という多項分布に従う．ただし，

$$p_{m,j} = P_{m,\alpha,\beta}(X_i \in A_j,\ Y_i = 1), \quad 1 \le j \le J,$$

とおいた．したがって補題 1 より，$(N_m(A_1), \ldots, N_m(A_J))$ は独立なポアソン確率変数に法則収束し，その平均パラメータは $(\lambda(A_1), \ldots, \lambda(A_J))$ である． □

定理 1 より，特にロジスティック回帰モデルは強度測度 $\exp(\alpha + \beta^T x)F(dx)$ のポアソン点過程モデルに収束する．これは Warton and Shepherd (2010) が示した事実と整合的である．

定義 1. 実数 $q \in \mathbb{R}$ に対し，式 (9) を強度測度の q-指数型分布族と呼ぶ．対応する点過程の確率法則を $P^{(q)}_{\alpha,\beta}$ と記す．

強度測度の q-指数型分布族は，確率測度の q-指数型分布族に密接に関係している．強度測度 (9) の全強度を

$$\Lambda_q(\alpha, \beta) = \int_{\mathbb{R}^p} \exp_q(\alpha + \beta^T x)F(dx) \tag{11}$$

と記すことにし，$\Lambda_q(\alpha, \beta) < \infty$ と仮定しよう．すると，$P^{(q)}_{\alpha,\beta}$ の尤度は

$$\frac{e^{-\Lambda_q(\alpha,\beta)}}{n!} \prod_{i=1}^{n} \exp_q(\alpha + \beta^T x_i) \tag{12}$$

と書ける．ここで，n の基準測度は非負整数上の計数測度であり，各 i に対する x_i の基準測度は $F(dx_i)$ とする．式 (12) において，数 n は観測値の個数であり，その周辺分布は平均 $\Lambda_q(\alpha,\beta)$ のポアソン分布である．n を条件づけたとき，各 x_i は独立に

$$\frac{\exp_q(\alpha+\beta^T x_i)}{\Lambda_q(\alpha,\beta)} \tag{13}$$

という密度をもつ確率分布に従う．式 (13) は q-指数型分布族，変形指数型分布族，あるいは α-分布族（$\alpha = 2q-1$）と呼ばれる．密度 (13) は，適切な θ と $\psi_q(\theta)$ を選んで $\exp_q(\theta^T x_i - \psi_q(\theta))$ という形に書くこともできる．(例えば Amari and Ohara (2011))．しかし，本稿ではこの表現を用いない．その理由は，この表現を用いても，ポアソン点過程の尤度関数 (12) において $\Lambda_q(\alpha,\beta)$ が残ってしまうためである．

図 1 に q-指数型分布族への収束のイメージを示す．

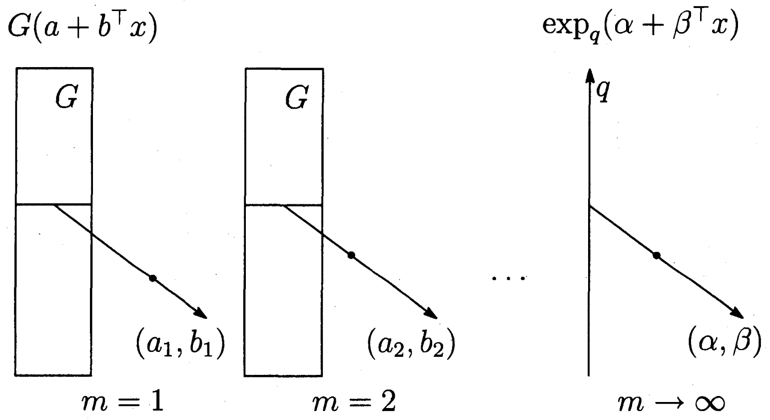

図 1: q-指数型分布族への収束のイメージを示す．モデル空間の確率分布は，G と a, b の選び方によって定まる．ただし $F(dX_i)$ は固定して考える．一方，不均衡極限においては G がいずれかの実数 q に対応する（または極限を持たない）．

3　ポアソン点過程のダイバージェンス

ポアソン点過程の q-指数型分布族に対する統計的推測法を構築するための前段階として，この節ではポアソン点過程間のダイバージェンスを計算する．なお，以下の議論はユークリッド空間 $\mathbb{R}^p$ を適当な可測空間に置き換えても同様に成り立つ．

$\mathbb{R}^p$ 上の確率測度 $F(dx)$ を固定する．F に対して絶対連続な 2 つの有限測度 μ, λ を考え ($\mu(\mathbb{R}^p), \nu(\mathbb{R}^p) < \infty$)，これらの（確率とは限らない）密度関数を r_μ, r_λ とおく．すなわち，

$$\mu(dx) = r_\mu(x)F(dx), \quad \lambda(dx) = r_\lambda(x)F(dx)$$

とする．以下，μ は真の強度測度，λ はモデルの強度測度と考える．特に，q-指数型分布族（定義 1）は，$r_\lambda = \exp_q(\alpha+\beta^T x)$ と表される．

強度測度 μ のポアソン点過程は，次で定義される確率変数 n と確率変数列 $(x_i)_{i=1}^n$ の組と同一視できる（たとえば間瀬・武田 (2001) の命題2.7）：

$$n \sim \frac{\mu(\mathbb{R}^p)^n}{n!} e^{-\mu(\mathbb{R}^p)} \quad \text{（ポアソン分布）} \tag{14}$$

$$x_i | n \sim \frac{\mu(dx_i)}{\mu(\mathbb{R}^p)} = \frac{r_\mu(x_i) F(dx_i)}{\mu(\mathbb{R}^p)}, \quad i = 1, \ldots, n, \quad \text{(i.i.d.)}. \tag{15}$$

これによって集合 $\Omega = \cup_{n=0}^{\infty} (\mathbb{R}^p)^n$ (disjoint union) の上の確率測度 P_μ が定義される：

$$P_\mu = e^{-\mu(\mathbb{R}^p)} \sum_{n=0}^{\infty} \frac{\mu^n}{n!}.$$

ただし μ^n は，μ の n 重の直積測度である．すると，P_F に対する P_μ の密度関数（尤度関数）は

$$p_\mu(x) = e^{-\mu(\mathbb{R}^p)+1} \prod_{i=1}^{n} r_\mu(x_i), \quad x = (x_i)_{i=1}^n \in \Omega, \tag{16}$$

となる．

密度関数が求まったので，これをダイバージェンスの定義に代入してその式を吟味することができる．以下では P_μ に関する期待値を E_μ と書く．

ポアソン点過程 P_μ と P_λ の間の Kullback–Leibler (KL) ダイバージェンスは以下のようになる：

$$\begin{aligned}
\mathrm{KL}(P_\mu \| P_\lambda) &= E_\mu \left[\log \frac{p_\mu}{p_\lambda} \right] \\
&= E_\mu \left[-\mu(\mathbb{R}^p) + \lambda(\mathbb{R}^p) + \sum_{j=1}^{n} \log \frac{r_\mu(x_j)}{r_\lambda(x_j)} \right] \\
&= -\mu(\mathbb{R}^p) + \lambda(\mathbb{R}^p) + E_\mu \left[n E_\mu \left[\log \frac{r_\mu(x_1)}{r_\lambda(x_1)} | n \right] \right] \\
&= -\mu(\mathbb{R}^p) + \lambda(\mathbb{R}^p) + \mu(\mathbb{R}^p) \int \frac{\mu(dx_1)}{\mu(\mathbb{R}^p)} \log \frac{r_\mu(x_1)}{r_\nu(x_1)} \\
&= \int_{\mathbb{R}^p} \left(-r_\mu + r_\lambda + r_\mu \log \frac{r_\mu}{r_\lambda} \right) dF \\
&= \mathrm{KL}(\mu \| \lambda).
\end{aligned}$$

すなわち，ポアソン点過程間の KL ダイバージェンスは，それらの強度測度の間の（拡張された）KL ダイバージェンスに等しい．

次に，α-ダイバージェンスは以下のようになる：

$$
\begin{aligned}
D_\alpha(P_\mu \| P_\lambda) &= \frac{1}{1-\alpha^2} E_\mu\left[1-\left(\frac{p_\mu}{p_\lambda}\right)^{\frac{-1+\alpha}{2}}\right] \\
&= \frac{1}{1-\alpha^2} E_\mu\left[1-\left\{\left(\prod_{i=1}^{n}\frac{r_\mu(x_i)}{r_\lambda(x_i)}\right)e^{-\mu(\mathbb{R}^p)+\lambda(\mathbb{R}^p)}\right\}^{\frac{-1+\alpha}{2}}\right] \\
&= \frac{1}{1-\alpha^2} E_\mu\left[1-\left(E_\mu\left[\left(\frac{r_\mu(x_1)}{r_\lambda(x_1)}\right)^{\frac{-1+\alpha}{2}}\middle| n\right]\right)^{n} e^{\frac{-1+\alpha}{2}(-\mu(\mathbb{R}^p)+\lambda(\mathbb{R}^p))}\right] \\
&= \frac{1}{1-\alpha^2}\left\{1-\exp\left(\int r_\mu^{\frac{1+\alpha}{2}} r_\lambda^{\frac{1-\alpha}{2}} dF - \frac{1+\alpha}{2}\mu(\mathbb{R}^p) - \frac{1-\alpha}{2}\lambda(\mathbb{R}^p)\right)\right\} \\
&= \frac{1}{1-\alpha^2}\left\{1-\exp\left(-(1-\alpha^2)D_\alpha(\mu\|\nu)\right)\right\}
\end{aligned}
$$

すなわち，ポアソン点過程間の α-ダイバージェンスは，それらの強度測度の間の α-ダイバージェンスの単調増加関数となる．

なお，上の式を $D_\alpha(\mu\|\nu)$ について解くと，

$$
\begin{aligned}
D_\alpha(\mu\|\lambda) &= \frac{1}{-1+\alpha^2}\log\left\{1-(1-\alpha^2)D_\alpha(P_\mu\|P_\lambda)\right\} \\
&= \frac{1}{-1+\alpha^2}\log\int p_\mu^{\frac{1+\alpha}{2}} p_\lambda^{\frac{1-\alpha}{2}} dP_F
\end{aligned}
$$

となり，これは Rényi ダイバージェンスを $(1+\alpha)$ で割った量となる．

次に，γ-ダイバージェンス (Fujisawa and Eguchi, 2008) を求める．γ-ダイバージェンスは，基準測度を P_F とするとき，

$$
d_\gamma(p_\mu\|p_\lambda) = -\frac{1}{\gamma}\log\int p_\mu p_\lambda^\gamma dP_F + \frac{1}{1+\gamma}\log\int p_\lambda^{1+\gamma} dP_F + \frac{1}{\gamma(1+\gamma)}\log\int p_\mu^{1+\gamma} dP_F
$$

と定義される．第一項の積分を計算すると

$$
\begin{aligned}
\int p_\mu p_\lambda^\gamma dP_F &= \int p_\lambda^\gamma dP_\mu \\
&= E_\mu\left[e^{-\gamma\lambda(\mathbb{R}^p)+\gamma}\prod_{i=1}^{n} r_\lambda(x_i)^\gamma\right] \\
&= e^{-\gamma\lambda(\mathbb{R}^p)+\gamma} E_\mu\left[\left(\frac{\int r_\mu r_\lambda^\gamma dF}{\mu(\mathbb{R}^p)}\right)^n\right] \\
&= \exp\left\{-\gamma\lambda(\mathbb{R}^p)+\gamma+\int r_\mu r_\lambda^\gamma dF - \mu(\mathbb{R}^p)\right\}
\end{aligned}
$$

となる．よって，

$$
\begin{aligned}
d_\gamma(p_\mu \| p_\lambda) = &-\frac{1}{\gamma}\left(-\gamma\lambda(\mathbb{R}^p)+\gamma+\int r_\mu r_\lambda^\gamma dF-\mu(\mathbb{R}^p)\right)\\
&+\frac{1}{1+\gamma}\left(-\gamma\lambda(\mathbb{R}^p)+\gamma+\int r_\lambda^{1+\gamma} dF-\lambda(\mathbb{R}^p)\right)\\
&+\frac{1}{\gamma(1+\gamma)}\left(-\gamma\mu(\mathbb{R}^p)+\gamma+\int r_\mu^{1+\gamma} dF-\mu(\mathbb{R}^p)\right)\\
=&-\frac{1}{\gamma}\int r_\mu r_\lambda^\gamma dF+\frac{1}{1+\gamma}\int r_\lambda^{1+\gamma}dF+\frac{1}{\gamma(1+\gamma)}\int r_\mu^{1+\gamma}dF
\end{aligned}
$$

となる．最後の式は，有限密度 r_μ, r_λ の間の β-ダイバージェンス (density power divergence, Basu et al. (1998)) で $\beta=\gamma$ とおいたものとなっている．

以上の対応関係をまとめると表2のようになる．結果として，ポアソン点過程間に定義される著名なダイバージェンスは，強度測度間に定義される自然なダイバージェンスに対応する．ところで，確率測度の空間における q-指数型分布族については，その推定関数の幾何学的特徴付けが議論されている（たとえば松添 (2013)）．この類推が，強度測度（有限測度）の空間における q-指数型分布族（定義1）に対しても成り立つものと考えられる．その解明については今後の課題としたい．

表 2: ダイバージェンスの対応.

点過程	強度測度
KL	KL
α (Rényi)	α
γ	β

参考文献

Amari, S., 1985. Differential-geometrical methods in statistics. Berlin: Springer.

Amari, S., Nagaoka, H., 2000. Methods of information geometry (Translations of Mathematical Monographs). Oxford University Press.

Amari, S., Ohara, A., 2011. Geometry of q-exponential family of probability distributions. Entropy 13, 1170–1185.

Baddeley, A., Berman, M., Fisher, N.I., Hardegen, A., Milne, R.K., Schuhmacher, D., Shah, R., Turner, R., 2010. Spatial logistic regression and change-of-support in Poisson point processes. Electron. J. Statist. 4, 1151–1201.

Basu, A., Harris, I.R., Hjort, N.L., Jones, M.C., 1998. Robust and efficient estimation by minimising a density power divergence, Biometrika, 85, 549–559.

Bolton, R.J., Hand, D.J., 2002. Statistical fraud detection: a review. Statist. Sci. 17, 235–249.

Chawla, N.V., Japkowicz, N., Koltz, A., 2004. Editorial: special issue on learning from imbalanced data sets. ACM SIGKDD Explorations Newsletter 6, 1–6.

Ding, N., Vishwanathan, S.V.N., Warmuth, M., Denchev, V., 2011. t-logistic regression. J. Mach. Learn. Res. 12, 1–55.

Embrechts, P., Klüppelberg, C., Mikosch, T., 1997. Modelling extremal events. Berlin: Springer.

Fujisawa, H., Eguchi, S., 2008. Robust parameter estimation with a small bias against heavy contamination, J. Multivariate Anal., 99, 2053–2081.

Galambos, J., 1987. The asymptotic theory of extreme order statistics. Malarbar: Robert E. Krieger Publishing Company.

de Haan, L., Ferreira, A., 2006. Extreme value theory, an introduction. New York: Springer.

Jin, Y., Rejesus, R.M., Little, B.B., 2005. Binary choice models for rare events data: a crop insurance fraud application. Applied Economics 37, 841–848.

King, G., Zeng, L., 2001. Logistic regression in rare events data. Political Analisis 9, 137–163.

間瀬 茂, 武田 純, 2001. 空間データモデリング – 空間統計学の応用, データサイエンス・シリーズ 7. 共立出版.

松添 博, 2013. 統計多様体と推定関数の幾何学, 2013 年度日本数学会年会予稿.

Naudts, J., 2002. Deformed exponentials and logarithms in generalized thermostatistics. Physica A 316, 323–334.

Naudts, J., 2010. The q-exponential family in statistical physics. J. Phys.: Conf. Ser. 201, 012003.

Owen, A.B., 2007. Infinitely imbalanced logistic regression. J. Mach. Learn. Res. 8, 761–773.

Sei, T., 2014. Infinitely imbalanced binomial regression and deformed exponential families. J. Statist. Plan. Infer. in Press.

Tsallis, C., 1988. Possible generalization of Boltzmann-Gibbs statistics. J. Statist. Phys. 52, 479–487.

Warton, D.I., Shepherd, L.C., 2010. Poisson point process models solve the "pseudo-absence problem" for presence only data in ecology. Ann. Applied Statist. 4, 1383–1402.