鈴木譲「ベイジアンネットワーク入門」(培風館)
1. 確率論の基礎
1.3 分布関数

鈴木譲

大阪大学

2010 年 4 月 22 日

あらまし

1. 分布関数
2. 確率密度関数
3. 多変量の確率密度関数

分布関数

$X : \Omega \to \mathbb{R}$: 確率変数

$$(X \leq x) \in \mathcal{F},\ x \in \mathbb{R}$$

$$\{\omega \in \Omega | X(\omega) \in D\} \in \mathcal{F},\ D \in \mathcal{B}$$

X の分布関数 F_X

$$F_X(x) := \mu(X \leq x),\ x \in \mathbb{R}$$

X の確率測度 μ_X

$$\mu_X(D) := \mu(X^{-1}(D)) = \mu(\{\omega \in \Omega | X(\omega) \in D\}),\quad D \in \mathcal{B}$$

X の測度 ν_X

$$\nu_X(D) := \nu(X^{-1}(D)) = \nu(\{\omega \in \Omega | X(\omega) \in D\}),\quad D \in \mathcal{B}$$

分布関数の性質

定理 1.2

1. 単調増加：$x < x' \Longrightarrow F_X(x) \leq F_X(x')$
2. $F_X(-\infty) = \lim_{x \to -\infty} F_X(x) = 0,\ F_X(+\infty) = \lim_{x \to +\infty} F_X(x) = 1$
3. 右連続：$\lim_{h \to +0} F_X(x+h) = F_X(x)$

証明：

1. $x_1 < x_2 \Longrightarrow \mu(X \leq x_1) \leq \mu(X \leq x_1) + \mu(x_1 < X \leq x_2) = \mu(X \leq x_2)$.
2. $A_n := (X \leq -n)$ とおくと、$\lim_{n \to \infty} A_n = \{\}$ と命題 1.2 より
$$\lim_{x \to -\infty} \mu(X \leq x) = \lim_{n \to \infty} \mu(A_n) = \mu(\lim_{n \to \infty} A_n) = \mu(\{\}) = 0.$$
$A_n := (X \leq n)$ とおくと、$\lim_{n \to \infty} A_n = \Omega$ と命題 1.2 より
$$\lim_{x \to \infty} \mu(X \leq x) = \lim_{n \to \infty} \mu(A_n) = \mu(\lim_{n \to \infty} A_n) = \mu(\Omega) = 1.$$
3. $B_n := (x < X \leq x + 1/n)$ に、$\lim_{n \to \infty} B_n = \{\}$ と命題 1.2 を適用

$\mathbb{R}$ における絶対連続性

X: 確率変数
λ_X: $\mathcal{B}$ 上の Lebesgue 測度

定理 1.3

以下の 2 条件は同値

1. 各 $D \in \mathcal{B}$ について
$$\lambda_X(D) = 0 \Longrightarrow \mu_X(D) = 0.$$
2. $$F_X(x) = \int_{-\infty}^{x} f_X(x)\,\lambda_X(dx)$$
を満たす $\mathcal{B}$ 上可測な f_X が存在

$\mathbb{R}$ における絶対連続性

F_X は絶対連続、μ_X は λ_X に関して絶対連続

定理 1.3 の同値な条件が成立する

任意の $\epsilon > 0$ と各 $k = 1, 2, \cdots$ について

$$\sum_{i=1}^{k} |b_i - a_i| < \delta(\epsilon) \Longrightarrow \sum_{i=1}^{k} |F_X(b_i) - F_X(a_i)| < \epsilon$$

なる $\delta(\epsilon) > 0$ が存在 $(a_i < b_i)$

一様連続

$k = 1$ の場合のみを要求

Cantor 関数

Cantor 集合の 3 進数表現に確率測度 μ_X を設定

$0, 2$ が確率 $1/2$、1 が確率 0

- $X \in [0, 1)$ の分布関数 F_X は連続。実際、各 $a \in [0, 1)$ に対して,

$$|x - a| < \delta := 3^{-n} \Longrightarrow |F_X(x) - F_X(a)| \leq \epsilon := 2^{-n}$$

- $\lambda(A) = 0, \mu_X(A) = 1$ であるので，F_X は絶対連続ではない。

定理 1.3 の証明

1. $\Longrightarrow$ 2. は、Radon-Nikodym の定理 (定理 1.2) から
2. を仮定
f_X は $\mathcal{B}$ 上可測ゆえ、$A_n := \{x \in \mathbb{R} | f_X(x) > n\}$ は $\mathcal{B}$ の要素 (事象)
$A_n \downarrow \{\}$ と命題 1.1 の 2. より，$\mu_X(A_n) \downarrow 0$

$$\mu_X(A_n) = \int_{A_n} f_X(x)\lambda_X(dx) < \frac{\epsilon_n}{2}$$

を満足する $\epsilon_n \to 0$, $\epsilon_n > 0$ なる列 $\{\epsilon_n\}$ が存在
$D \in \mathcal{B}$ に対して，$\lambda_X(D) < \delta := \epsilon_n/(2n)$ であれば

$$\begin{aligned}\mu_X(D) &= \int_D f_X(x)\lambda_X(dx) \\ &= \int_{D-A_n} f_X(x)\lambda_X(dx) + \int_{A_n} f_X(x)\lambda_X(dx) \le n\lambda_X(D) + \frac{\epsilon_n}{2} < \epsilon_n\end{aligned}$$

確率密度関数

確率変数 X が連続型

X が絶対連続な分布関数をもつ
(f_X を確率密度関数という)

f_X が連続のとき、$h \to 0$ で $(h > 0, h < 0$ とも$)$

$$\left| \frac{1}{h} \int_x^{x+h} f_X(y)\, dy - f_X(x) \right| \le \frac{1}{h} \int_x^{x+h} |f_X(y) - f_X(x)|\, dy$$
$$\le \quad \sup[|f_X(y) - f_X(x)|\,|\, x \le y \le x+h] \to 0$$

$$F'_X(x) = \frac{d}{dx} \int_{-\infty}^{x} f_X(y)\, dy = f_X(x)$$

- $f_X(x) \ge 0$
- $\int_{-\infty}^{\infty} f_X(x)\, dx = 1$

正規分布 $N(\mu, \sigma^2)$

$$F_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{x} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] dx$$

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$$

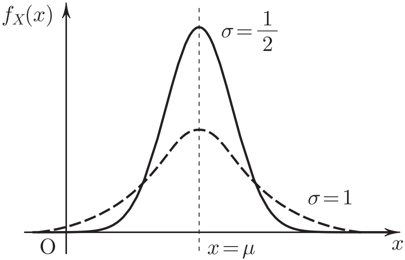

Figure: 正規分布

正規分布 $N(\mu, \sigma^2)$ (続)

$$I := \int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$$

$$I^2 = \iint_{\mathbb{R}^2} e^{-(x^2+y^2)}\, dxdy = \int_{0<\theta<2\pi} \int_{r=0}^{\infty} e^{-r^2} r\, drd\theta = \pi$$

標準正規分布

正規分布 $N(0,1)$

ガンマ分布 $\Gamma(\alpha, \lambda)$

$$f_X(x) = \begin{cases} \dfrac{1}{\Gamma(\alpha)} \lambda^{\alpha} x^{\alpha-1} e^{-\lambda x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

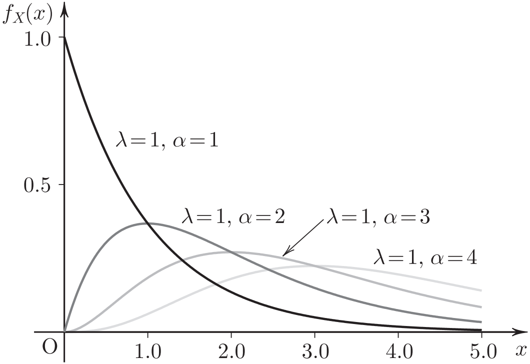

Figure: ガンマ分布

ガンマ関数 $\Gamma(\alpha)$

$$\Gamma(\alpha) := \int_0^\infty e^{-t} t^{\alpha-1} dt$$

$I = \sqrt{\pi}$ より、$t := x^2, \alpha := \frac{1}{2}$ とおくと、$\Gamma(\frac{1}{2}) = \sqrt{\pi}$

$$\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$$

自由度 n の χ^2 分布 χ^2_n

自由度 n の χ^2 分布

$\alpha = n/2, \lambda = 1/2$ のときのガンマ分布

$$f_X(x) = \frac{1}{2^{n/2}\Gamma(n/2)} x^{n/2-1} e^{-x/2}$$

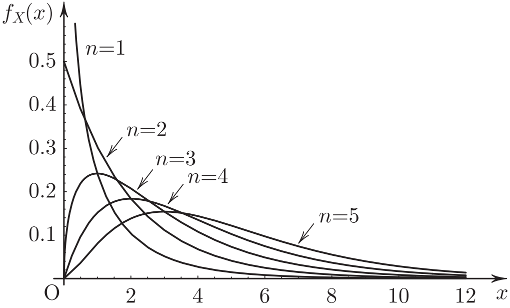

Figure: χ^2 分布

離散確率分布

$\mu_X(\{x\}) > 0$ なる $x \in \mathbb{R}$ が存在 $\Longrightarrow$ 確率密度関数は存在しない

$\mathrm{support}(\mu_X) := \{x \in \mathbb{R} \mid \mu_X(\{x\}) > 0\}$

$$\sum_{x \in \mathrm{support}(\mu_X)} \mu_X(\{x\}) = 1 \Longrightarrow \mathrm{support}(\mu_X)\ \text{は可算集合}$$

$\mu_X(\{x\}),\ x \in \mathrm{support}(\mu_X)$ (離散確率分布)

$F_X(x) = \sum_{y \le x} \mu_X(\{y\})$

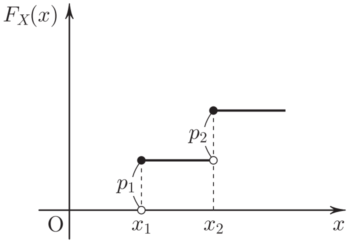

Figure: $F_X(x)$, $x \in \mathrm{support}(\mu) = \{x_1, x_2, \cdots\}$

離散確率分布 (続)

$A \in \mathcal{F}$

確率変数 $X = I_A$ の分布関数 $F_X(x) = \begin{cases} 0, & x < 0 \\ 1 - \mu(A), & 0 \le x < 1 \\ 1, & x \ge 1 \end{cases}$

確率変数の変換

X, Y: 確率変数
f_X: 確率変数 X の確率密度関数

定理 1.4

Y が単調な微分可能な関数 g を用いて $Y = g(X)$ とかけるとき、

$$f_Y(y) = f_X(g^{-1}(y))\left|\frac{dg^{-1}(y)}{dy}\right|$$

証明: テキスト (21 ページ)

(例) $X \sim N(0, 1), Y = X^2$ として、$y > 0$ では

$$f_Y(y) = 2f_X(\sqrt{y})\left|\frac{dx}{dy}\right| = 2\frac{1}{\sqrt{2\pi}}e^{-y/2}\left|-\frac{1}{2\sqrt{y}}\right| = \frac{1}{\sqrt{2\pi y}}e^{-y/2},$$

$y < 0$ では $f_Y(y) = 0$

n 個の確率変数 $X_1, \cdots, X_n$

$(X_i \in D_i) \in \mathcal{F},\ D_i \in \mathcal{B},\ i = 1, \cdots, n$ より、

$$(X_1 \in D_1, \cdots, X_n \in D_n) \in \mathcal{F},\ D_1, \cdots, D_n \in \mathcal{B}$$

$$(X_1 \leq x_1, \cdots, X_n \leq x_n) \in \mathcal{F}, x_1, \cdots, x_n \in \mathbb{R}$$

$$F_{X_1,\cdots,X_n}(x_1, \cdots, x_n) := \mu(X_1 \leq x_1, \cdots, X_n \leq x_n) \in \mathcal{F},\quad x_1, \cdots, x_n \in \mathbb{R},$$

$$\mu_{X_1,\cdots,X_n}(D_1, \cdots, D_n) := \mu(X_1 \in D_1, \cdots, X_n \in D_n)$$

定理 1.2，定理 1.3，定理 1.4 は，n 変数の場合に拡張
$f_{X_1,\cdots,X_n}$: $X_1, \cdots, X_n$ の同時密度関数
確率変数 $Y_1, \cdots, Y_n$ が微分可能な関数 $g_1, \cdots, g_n$ を用いて，
$Y_i := g_i(X_1, \cdots, X_n)$ とかけるとき、J をヤコビアンとして、

$$f_{Y_1,\cdots,Y_n}(y_1, \cdots, y_n) = f_{X_1,\cdots,X_n}(g_1^{-1}(y_1, \cdots, y_n), \cdots, g_n^{-1}(y_1, \cdots, y_n))|J|$$

$$x_j = g^{-1}(y_1, \cdots, y_n)$$

Direchlet 分布

$$w_{a,b}(\theta) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\theta^{a-1}(1-\theta)^{b-1}, \quad 0 \le \theta \le 1$$

$$w_{a,b}(\theta) \ge 0, \int_0^1 w_{a,b}(\theta)\, d\theta = 1$$

$u := x/(x+y), w := x+y$ とおくと，$x = uw$，$y = (1-u)w$ より、

$$\left| \begin{array}{cc} \partial x/\partial u & \partial x/\partial w \\ \partial w/\partial u & \partial w/\partial u \end{array} \right| = \left| \begin{array}{cc} w & u \\ -w & 1-u \end{array} \right| = w,$$

$$\begin{aligned} \Gamma(a)\Gamma(b) &= \int_0^\infty \int_0^\infty e^{-(x+y)} x^{a-1} y^{b-1}\, dxdy \\ &= \int_0^1 \int_0^\infty e^{-w}(uw)^{a-1}[(1-u)w]^{b-1}|J|\, dudw \\ &= \int_0^1 u^{a-1}(1-u)^{b-1} du \int_0^\infty e^{-w} w^{a+b-1}\, dw = \beta(a,b)\Gamma(a+b) \end{aligned}$$

Direchlet 分布 (続)

$$\beta(a,b) := \int_0^1 u^{a-1}(1-u)^{b-1}\,du$$

一般に，$m \geq 2$ では，

$$\theta := (\theta_1, \cdots, \theta_{m-1}),\ \theta_0 := 1 - \sum_{i=1}^{m-1} \theta_i,\ 0 \leq \theta_i \leq 1,\ i = 0, 1, \cdots, m-1$$

$a_0 > 0,\ a_1 > 0,\ \cdots,\ a_{m-1} > 0$ に対して，

$$w_{a_0,a_1,\cdots,a_{m-1}}(\theta) = \frac{\Gamma\left(\sum_{i=0}^{m-1} a_i\right)}{\prod_{i=0}^{m-1}\Gamma(a_i)} \prod_{i=0}^{m-1} \theta_i^{a_i}$$

$$\int w_{a_0,a_1,\cdots,a_{m-1}}(\theta)\,d\theta_1 \cdots d\theta_{m-1} = 1$$

同時確率密度関数と周辺確率密度関数

$f_{X_1,\cdots,X_n}$ が連続であれば,

$$f_{X_1,\cdots,X_n}(x_1,\cdots,x_n) = \frac{\partial^n}{\partial x_1 \cdots \partial x_n} F_{X_1,\cdots,X_n}(x_1,\cdots,x_n)$$

- $f_{X_1,\cdots,X_n}(x_1,\cdots,x_n) \geq 0$
- $\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{X_1,\cdots,X_n}(x_1,\cdots,x_n)\, dx_1 \cdots dx_n = 1$

各 $i = 1,\cdots,n$ について、

$$f_{X_i}(x_i) := \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{X_1,\cdots,X_n}(x_1,\cdots,x_n)\, dx_1 \cdots dx_{i-1} dx_{i+1} \cdots dx_n$$

確率変数によって生成される事象

確率変数 X によって生成された事象集合 $\mathcal{F}(X)$

$\{(X \in D), D \in \mathcal{B}\}$

確率変数 X のもとでの A の条件付確率

$\mathcal{G} = \mathcal{F}(X)$ のときの f_A

確率変数 $X_1, \cdots, X_n$ のもとでの A の条件付確率

$\mathcal{G} = \mathcal{F}(X_1, \cdots, X_n)$ のときの f_A

X, Y が独立

$\mathcal{F}(X), \mathcal{F}(Y) \subseteq \mathcal{F}$ が独立 $(\Longleftrightarrow F_{XY}(x,y) = F_X(x)F_Y(y), x, y \in \mathbb{R})$

$X_1, \cdots, X_n$ が独立

どの $X_i, X_j (i \neq j)$ も独立

確率変数によって生成される事象 (続)

例:
X: $X(\Omega)$ が可算集合、$\mu(X = x) > 0$, $x \in X(\Omega)$ となる確率変数
$\mathcal{G}$: $(X = x)$, $x \in X(\Omega)$ で生成される事象集合

各 $G \in \mathcal{G}$ で

$$\mu(A \cap G) = \sum_{x \in X(\Omega)} \frac{\mu(A \cap (X = x))}{\mu(X = x)} \mu(X = x)$$

$$X(\omega) = x \Longrightarrow f_A(\omega) = \frac{\mu(A \cap (X(\omega) = x))}{\mu(X(\omega) = x)}$$

確率変数によって生成される事象 (続)

例:

X, Y: 独立な確率変数の分布関数 F_X, F_Y が絶対連続

$$F_X(x) = \int_{-\infty}^{x} f_X(u)\,du,\ \ F_Y(y) = \int_{-\infty}^{y} f_Y(v)\,du$$

$$F_{XY}(x, y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{XY}(u, v)\,dudv$$

f_X, f_Y, f_{XY}(確率密度関数) が存在

$$F_{XY}(x, y) = F_X(x)F_Y(y), x, y \in \mathbb{R}$$

$$F_X(x)F_Y(y) - F_{XY}(x, y) = \int_{-\infty}^{x}\int_{-\infty}^{y} \{f_X(u)f_Y(v) - f_{XY}(u, v)\}\,dudv$$

$$f_X(x)f_Y(y) - f_{XY}(x, y) = 0$$

多変量正規分布

$U_1, \cdots, U_n$: 標準正規分布をもつ独立な確率変数
$\mu_1, \cdots, \mu_n \in \mathbb{R}$
$A \in \mathbb{R}^{n \times n}$: 正則な行列

$X_1, \cdots, X_n$ は n 変量正規分布にしたがう

$$\begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix} = A \begin{bmatrix} U_1 \\ \vdots \\ U_n \end{bmatrix} + \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_n \end{bmatrix}$$

$\mathbf{U} := {}^t(U_1, \cdots, U_n)$, $\mathbf{u} := {}^t(u_1, \cdots, u_n)$, $\boldsymbol{\mu} := {}^t(\mu_1, \cdots, \mu_n)$ とおくと,

$$F_{\mathbf{U}}(\mathbf{u}) = \prod_{i=1}^{n} F_{U_i}(u_i) = \int_{-\infty}^{u_1} \cdots \int_{-\infty}^{u_n} \left\{ \prod_{i=1}^{n} f_{U_i}(v_i) \right\} dv_1 \cdots dv_n$$

$$f_{\mathbf{U}}(\mathbf{u}) := \prod_{i=1}^{n} f_{U_i}(u_i) = \frac{1}{(2\pi)^{n/2}} \exp\left\{ -\frac{1}{2} {}^t\mathbf{u}\mathbf{u} \right\}$$

多変量正規分布 (続)

$\mathbf{x} := A\mathbf{u} + \boldsymbol{\mu},\ \mathbf{X} := A\mathbf{U} + \boldsymbol{\mu}$

$$\begin{aligned} f_{\mathbf{X}}(A^{-1}(\mathbf{x} - \boldsymbol{\mu})) &= f_{\mathbf{U}}(\boldsymbol{\mu}) \left| \frac{\partial \boldsymbol{\mu}}{\partial \mathbf{x}} \right| \\ &= \frac{1}{(2\pi)^{n/2}} \exp\left\{ -\frac{1}{2}\, {}^t(\mathbf{x} - \boldsymbol{\mu})\, {}^t(A^{-1}) A^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right\} |\det(A^{-1})| \end{aligned}$$

$C := A\,{}^tA$ とおくと、$\det(C^{-1}) = |\det(A^{-1})|^2$ より,

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{\sqrt{\det(C^{-1})}}{(2\pi)^{n/2}} \exp\left\{ -\frac{1}{2}\, {}^t(\mathbf{x} - \boldsymbol{\mu}) C^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right\}$$

多変量正規分布 (続)

逆に，正定値行列 C から、$C = A\,{}^t\!A$ を満足する正則行列 A が得られる。

$$X_i = a_{i1}U_1 + \cdots + a_{in}U_n,\ \ a_1, \cdots, a_n \in \mathbb{R}$$

独立な正規分布に従う確率変数の和は正規分布に従う

証明: テキスト (26-27 ページ)

独立でない正規分布に従う確率変数の和は正規分布に従う

正規分布と χ^2 分布

命題 1.5

$X_1, \cdots, X_n$ が標準正規分布に従う独立な確率変数であれば,
$\chi^2 := X_1^2 + \cdots + X_n^2$ は自由度 n の χ^2 分布に従う。

証明: テキスト (27-28 ページ)