4章 確率変数の変数変換

目的　確率分布のクラスの拡張,

4.1 線形変換

$$\begin{cases} Y = aX + b & \text{一変数} \\ \vec{Y} = \mathbb{A}\vec{X} + \vec{b} & \text{多変数} \end{cases}$$

前提条件　X, $F_X(x)$, $f_X(x)$; 連続型が既知
r.v.　cdf　pdf

$Y = g(X)$, $X = g^{-1}(Y)$
$y = g(x)$, $x = g^{-1}(y)$

$Y = aX + b \quad (a > 0)$ の分布関数

$F_Y(y)$ を $F_X(g(y))$ と表している.

$$\begin{aligned} F_Y(y) &= P(Y \le y) & Y, y \\ &= P(aX + b \le y) & X, y \\ &= P\left(X \le \underbrace{\frac{y-b}{a}}_{g^{-1}(y)}\right) & X, y \\ &= F_X\left(\underbrace{\frac{y-b}{a}}_{g^{-1}(y)}\right) & X, y \end{aligned}$$

$x = \frac{y-b}{a} = g(y)$

$f_Y(y) = f_X(g(y))\, g'(y)$

$$\begin{aligned} f_Y(y) &= dF_Y(y)/dy & \text{y で微分} \\ &= dF_X\left(\frac{y-b}{a}\right)/dy \\ &= f\left(\frac{y-b}{a}\right)\cdot\left(\frac{y-b}{a}\right)' \\ &= f\left(\frac{y-b}{a}\right)\cdot\frac{1}{|a|} \end{aligned}$$

$Y=aX+b$, $a<0$ の場合

$$F_Y(y) = P(Y \leq y)$$
$$= P(aX+b \leq y)$$
$$= P(X \geq (y-b)/a)$$
$$= 1 - F_X(\underbrace{(y-b)/a}_{g(y)})$$

$$f_Y(y) = d(F_Y(y))/dy$$
$$= d(1-F_X((y-b)/a))/dy$$
$$= -f_X((y-b)/a)\cdot a$$

$a \neq 0$
$$f_Y(y) = f_X\left(\frac{y-b}{a}\right)\cdot\frac{1}{|a|}$$

線形変換とアフィン変換

＋平行移動

(回転, 拡大, 縮小, 剪断)

$F_X(x) = P(X \leq x)$ 累積分布関数と確率関数

$\frac{d}{dx} F_X(x) = f_X(x)$ —① 〃 と確率密度 〃

$x = g(y)$の時, $f_Y(y)$は? $f_X(x)$をもとにどう計算できる?

$$\left\{\begin{array}{l} F_Y(y) = \boxed{P(Y \leq y)} \\ \frac{d}{dy} F_Y(y) = f_Y(y) \end{array}\right.$$

結びつけるための確率,
とにかく求めたい事, 性質

$x = \frac{y-b}{a}$

$Y = g(X), \Leftrightarrow X = g^{-1}(Y)$

$$\boxed{\begin{aligned} P(Y \leq y) &= P(g(X) \leq y) \\ &= P(X \leq \underline{g^{-1}(y)}) \\ &= F_X(g^{-1}(y)) \end{aligned}}$$

符号が変化することがある?

$Y = aX + b$

$a<0$の時,

$\therefore F_Y(y) = F_X(\underbrace{g^{-1}(y)}_{x=})$

$$\frac{d}{dy} F_Y(y) = \frac{dF_X}{dx} \cdot \frac{dx}{dy} = F_X'(x) \cdot (g^{-1}(y))'$$

$x = g^{-1}(y)$

$g^{-1}(y)$

正規分布の 線型変換.

$X \sim N(\mu, \sigma^2)$ の時, $Y = aX + b$ の pdf は

$$f_Y(y) = f_X\left(\frac{y-b}{a}\right)\frac{1}{|a|}$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(y-(a\mu+b))^2}{2a^2\sigma^2}\right\}\frac{1}{|a|}$$

$$= \frac{1}{\sqrt{2\pi a^2\sigma^2}} \exp\left\{-\frac{\{y-(a\mu+b)\}^2}{2a^2\sigma^2}\right\}$$

$$Y = aX + b \sim N(a\mu+b, (a\sigma)^2)$$

つまり, 正規分布は 線型変換しても 正規分布

$$Z = \frac{X-\mu}{\sigma}, \Rightarrow a = \frac{1}{\sigma}, b = -\frac{\mu}{\sigma}$$

$$\sim N(0, 1)$$

$Z \sim N(0,1)$ の時 $X = \mu + \sigma Z$ の分布は

$$X \sim N(\mu, \sigma^2)$$

stat4
4.1-4

多次元の線形変換

n次元

$\vec{X} \sim f_X(\vec{x})$, 連続型

正規 正方行列 A, $n\times n$ による $\vec{X}$ の線形変換
n次元ベクトル $\vec{b}$

$$\vec{Y} = A\vec{X} + \vec{b}$$

導出は 4.4.1

$$\sim f_Y(\vec{y}) = f_X(A^{-1}(\vec{y}-\vec{b}))\frac{1}{|A|}$$

n次元 正規分布 の線形変換

$$f_X(\vec{x}) = \frac{1}{\sqrt{(2\pi)^n}|\Sigma|^{1/2}}\exp\left\{-\frac{1}{2}(\vec{x}-\vec{\mu})^t\Sigma^{-1}(\vec{x}-\vec{\mu})\right\}$$

$$f_Y(\vec{y}) = \frac{1}{\sqrt{(2\pi)^n}|\Sigma|^{1/2}}\frac{1}{|A|}$$

$$\times\exp\left\{-\frac{1}{2}\left(A^{-1}(\vec{y}-b)-\vec{\mu}\right)^t\Sigma^{-1} \quad ''\right\}$$

$E\cdot\vec{\mu} = A^{-1}A$

$$A^{-1}\left[(\vec{y}-b)-A\vec{\mu}\right]$$

$$\left(\mathbb{A}^{-1}(\vec{y}-\vec{b}) - \mathbb{A}^{-1}\cdot\mathbb{A}\,\vec{\mu}\right)^t$$

$$\left[\boxed{\mathbb{A}^{-1}}\boxed{\{(\vec{y}-\vec{b}) - \mathbb{A}\vec{\mu}\}}\right]^t$$

$$\{(\vec{y}-\vec{b}) - \mathbb{A}^{-1}\mu\}^t\,[\mathbb{A}^{-1}]^t$$

$$= \frac{1}{\sqrt{(2\pi)^n}\,|\mathbb{A}\Sigma\mathbb{A}^t|^{1/2}}$$

$$\exp\left\{-\frac{1}{2}\{(\vec{y}-\vec{b})-\mathbb{A}\vec{\mu}\}(\mathbb{A}\Sigma\mathbb{A}^t)^{-1}\{(\vec{y}-\vec{b})-\mathbb{A}\vec{\mu}\}\right\}$$

$$\therefore\quad \vec{Y} \sim N_n(\mathbb{A}\vec{\mu}+\vec{b},\ \mathbb{A}\Sigma\mathbb{A}')$$

変数の変換

$y = f(x)$, $\left\{ \begin{array}{l} f'(x)\text{が連続} \\ \text{単調関数} \end{array} \right.$

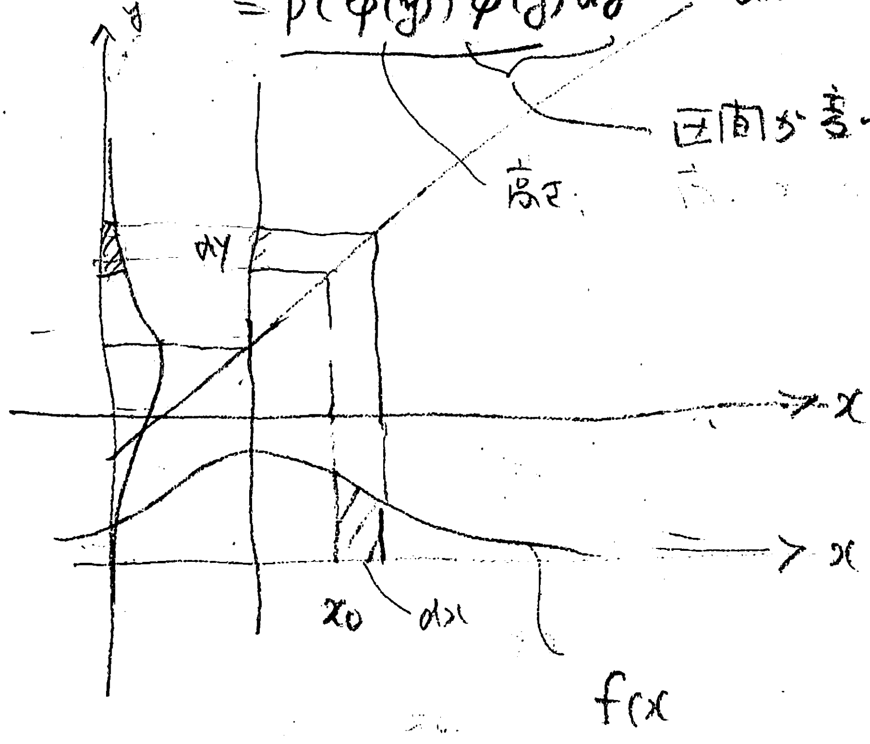
x軸上の分布 $p(x)$,
$y=f(x)$によって y軸上に移される分布 $g(x)$

⇓

$x = \varphi(y)$ とすると $(\varphi = f^{-1})$

$$\underline{g(y)}\,dy = p(x)\,dx$$
$$= \underline{p(\varphi(y))\,\varphi'(y)\,dy}$$

高さ
区間が変わる



分布 $p(x)$で、事象が (x_0, x_0+dx)の間になる
確率は、 $p(x_0)\,dx$

xとyの対応が $f(x)$, $x=\varphi(y)$

同じ事象をyで変換する

$$g(y_0)\,dy = p(x_0)\,dx, \quad y_0 = f(x_0)$$
$$x_0 = \varphi(y_0)$$

Stat
小針 p49-2

命題 2.7

$y=f(x)$の逆関数を $x_i=\varphi_i(y)$とし，

$\varphi_i(y)$は有限個の点を除いて

連続な導関数をもっているとする。

x軸上の分布 $p(x)$を $y=f(x)$によって

y 〃 $g(y)$に移す。

$$g(y)=\sum_i p(\varphi_i(y))\,\varphi_i'(y)$$

証明

y軸上の区間$(y, y+dy)$を考える。

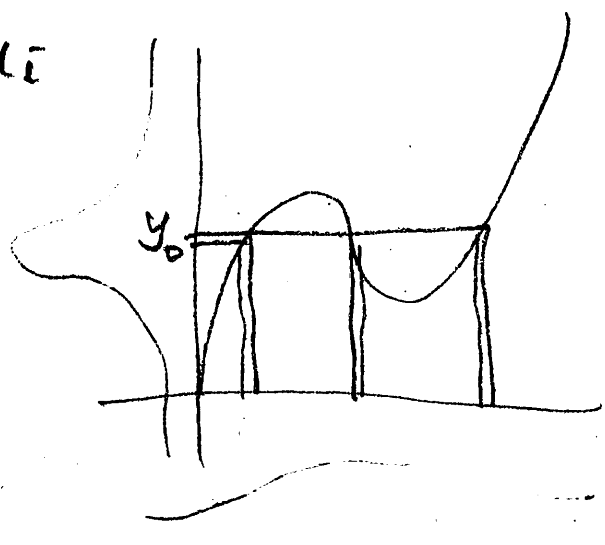
ここになる確率は$g(y)dy$

x軸上の区間に対応させると

$$(\varphi_i(y),\ \varphi_i(y)+\underbrace{\varphi_i'(y)dy}_{= dx_i})$$

$$\therefore \sum_i p(\varphi_i(y))\,\varphi_i'(y)dy$$

$$=\sum_i p(x_i)\,dx_i$$



4.2 独立な確率変数の和の分布

$X = X_1 + \cdots + X_n$ が従う分布

4.2.1 pdf に基づいた和の確率分布の導出

例　コイン投げ　　生起確率 p のベルヌーイ分布
n回の 〃　　$X_1, \cdots, X_n$ の確率変数
X_i 表なら $X_i = 1$
　　裏なら $X_i = 0$

$X = X_1 + \cdots + X_n \sim ?$

離散型　$X = x$ となる事象の定義

$$\mathcal{X} = \{\vec{x} = (x_1, \cdots, x_n) ; x_i = 0, 1 ; \sum x_i = x\}$$

$$\therefore \quad f(x) = P(X = x)$$

$$= \sum_{x \in \mathcal{X}} f(\vec{x})$$

$$= \sum_{x \in \mathcal{X}} f_1(x_1) \cdots f_n(x_n)$$

$$= \sum_{x \in \mathcal{X}} p^{x_1}(1-p)^{1-x_1} \cdots p^{x_n}(1-p)^{1-x_n}$$

$$= \sum_{x \in \mathcal{X}} p^{x}(1-p)^{n-x} = {}_nC_x\, p^{x}(1-p)^{n-x}$$

連続型

$X=\sum_{1}^{n} X_i$ の分布関数は

$$F(x) = P(X \leq x) = \int_{\sum x_i < x} f(\vec{x})\, d\vec{x}$$

$$= \int_{-\infty}^{\infty} \cdots \boxed{\int_{-\infty}^{x - \sum_{2}^{n} x_i} f(\vec{x})\, dx_1} \cdot d\vec{x}^{(1)}$$

xで微分して

$$f(x) = \vec{\int}_{-\infty}^{\infty} f(x - \sum_{2}^{n} x_i) \cdot d\vec{x}^{(1)}$$

独立だと

たたみこみ

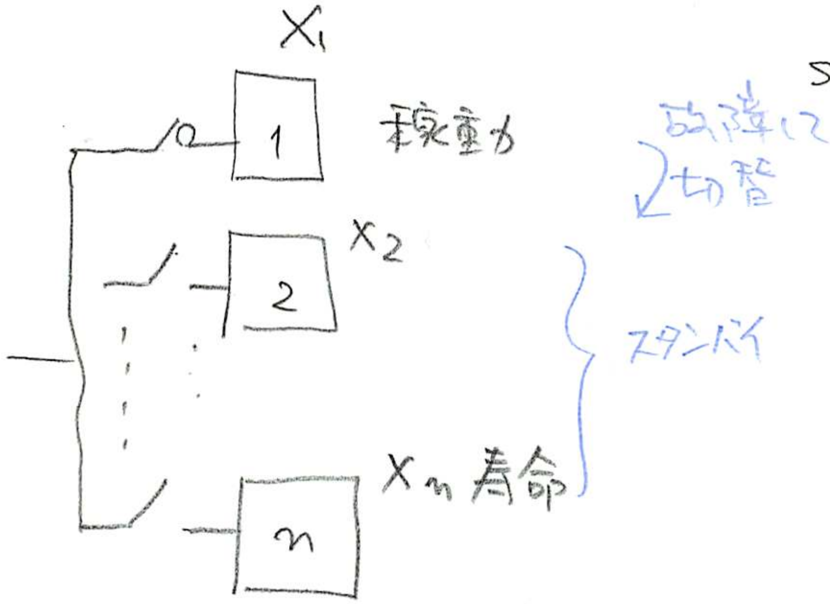
$$f(x) = \vec{\int}_{\infty}^{\infty} f_1(x - \sum x_i)\, f(x_2) \cdots f(x_n)\, d\vec{x}^{(1)}$$

$$\frac{d}{dx} \int_{-\infty}^{x-\alpha} f(x_1, \cdots, x_n)\, dx_1$$

$$= f(x - \alpha, \cdots x_n)$$

stat4.4.2-3

並列システムの寿命



$\Gamma(n,\lambda)$

スイッチ

システム全体の寿命 $X = \sum_{1}^{n} x_i \sim \Gamma(n,\lambda)$

$x_i \sim \exp_x(\lambda)$

$\lambda e^{-\lambda x}$

n=2の場合の確認

$$f(x) = \int_{-\infty}^{\infty} f_1(x - x_2) f_2(x_2)\, dx_2$$

$x_1 + x_2 = x$ となる　　$x - x_2 > 0$, $x_2 > 0$ なので

$$= \int_0^x f_1(x - x_2) \cdot f(x_2)\, dx_2$$

$$= \int_0^x (\lambda e^{-\lambda(x - x_2)})(\lambda e^{-\lambda x_2})\, dx_2$$

$$= \int_0^x \lambda^2 e^{-\lambda x}\, dx_2$$

$$= \lambda^2 e^{-\lambda x}$$

Γ-節後で

$$= \Gamma(2,\lambda)$$

4.2.2 モーメント母関数に基づいた和の確率分布の導出

原理

準備

$$X_1, \cdots, X_n \overset{\text{独立に}}{\sim} f_1(x), \cdots, f_n(x) \quad \text{をもつ}$$

$$\sim \psi_1(t), \cdots, \psi_n(t)$$

$$\Downarrow$$

$X = \sum X_i$ は X_i の独立性から.

和のモーメント母関数

$$\psi_X(t) = E[e^{tX}]$$
$$= E[e^{t(X_1 + \cdots X_n)}]$$
$$= E[e^{tX_1} e^{tX_2} \cdots e^{tX_n}]$$

独立

$$= E[e^{tX_1}] \cdot E[e^{tX_2}] \cdots E[e^{tX_n}]$$
$$= \psi_1(t) \cdots \psi_n(t)$$

各変数のモーメント母関数の積

二項分布の例

$$X_i \;\; i=1:n \sim B(m_i ; p) \;\; i=1:n$$

$B(m;\theta)$ の $\psi(t) = \{(1-p)+pe^t\}^m$

$X = \sum X_i$ の $\psi(t)$ は

$$\psi_X(t) = \psi_1(t) \cdots \psi_n(t)$$
$$= \{(1-p)+pe^t\}^{\sum m_i}$$

$$\psi_X(t) = \{(1-p) + pe^t\}^{\Sigma m_i}$$

$$\parallel$$

$B(\Sigma m_i ; p)$ 二項分布の母関数

$\downarrow$

二項分布の和の分布は二項分布になる

ガンマ分布の場合

$$X_i \sim \Gamma(m_i, \lambda)$$

$$\Downarrow$$

$$\sum X_i = X \sim \Gamma(\Sigma m_i, \lambda)$$

χ^2 分布の場合

$$Y = \sum_i^n X_i^2, \quad X_{i=1,n} \sim N(0,1)$$

$$X_i^2 \sim \Gamma(1/2, 1/2)$$

基礎事項より

[A.4.4]の問題

$$Y \sim \Gamma(n/2, 1/2)$$

正規分布の和の分布

前提　　独立に
$X_{i\ 1\sim n} \sim N(\mu_i, \sigma_i^2)$

$N(\mu, \sigma^2)$のモーメント母関数は
$$\psi(t) = \exp\left\{ \mu t + \left(\frac{\sigma^2}{2}\right) t^2 \right\}$$

線形結合について考える
$$X = \sum a_i X_i$$

$$\begin{aligned}\psi_X(t) &= E[e^{tX}] \\ &= E[e^{t\sum a_i X_i}] \\ &= \prod E[e^{t a_i X_i}] \\ &= \prod \exp\left\{ \mu_i (a_i t) + \left(\frac{\sigma_i^2}{2}\right)(t a_i)^2 \right\} \\ &= \exp\left\{ \mu_+ t + \left(\frac{\sigma_+^2}{2}\right) t^2 \right\}\end{aligned}$$

定ぎ

$$\left\{ \begin{aligned} \mu_+ &= \sum a_i \mu_i \\ \sigma_+^2 &= \sum a_i^2 \sigma_i^2 \end{aligned} \right\} \text{の母関数}$$

$$X \sim N(\mu_+, \sigma_+^2)$$

再生性

確率分布の再生性
(reproductivity)

同一の確率分布に基づく,
独立な
確率変数の和　ΣX_i, $\Sigma u(X_i)$
の分布が,
同じ確率分布に従がうこと

$$
\begin{aligned}
\psi_{\Sigma X}(t) &= E[e^{t\Sigma X}] \\
&= \Pi E[e^{tX_i}] \\
&= \Pi \psi_{X_i}(t) \\
&= \psi_X(\ \text{—}\ ;\Sigma\Theta)
\end{aligned}
$$

再生性をもつ分布の性質

畳み込み演算について閉じている。

再生性をもつ分布

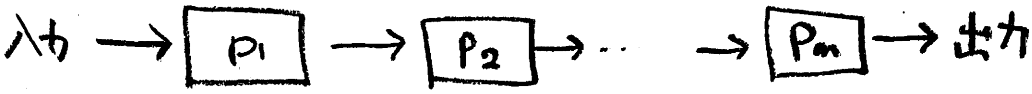
べき乗?

正規分布	$N(\Sigma\mu_i, \Sigma\sigma_i^2)$
ガンマ分布	$\Gamma(\Sigma n, \lambda)$
二項分布	$B(\Sigma n, p)$
ポアソン分布	$P_o(\Sigma\lambda)$
χ^2 分布	$\chi^2_{\underbrace{\Sigma n}_{\text{自由度}}}$

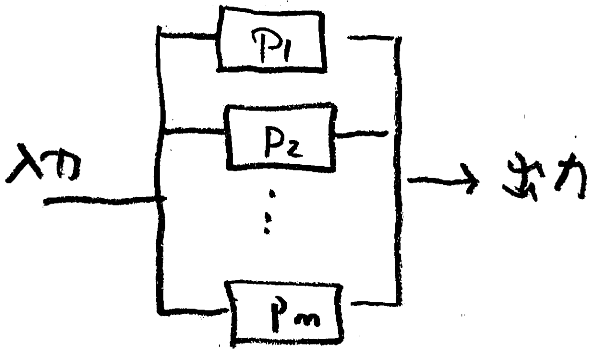
4.3 確率変数の最大値と最小値の分布

例. 多重システムの寿命.(直列と並列)

series 直列 cascade series



parallel // 並列 parallel



Seriesの $X_s = X_{\min x_i} = \min\{X_i : i=1:n\}$

$$
\begin{aligned}
P(X_{min} \le x) &= 1 - P(X_{min} > x) \\
&= 1 - P(\min\{X_i\} > x) \\
&= 1 - P(X_1 > x, \cdots, X_n > x) \\
&= 1 - \prod P(X_i > x) \\
&= 1 - \prod (1 - F_i(x))
\end{aligned}
$$

並列システムの

一台でも動いていれば、
システム全体として稼動（入力に対し出力がでる）

$$X_{//} = X_{max} = \max\{X_i : i=1:n\}$$

システム全体の寿命の分布関数

x時までは動く、x以上で動かない

$$\begin{aligned}
&P(X_{max} \leq x) \\
&= P(\max\{X_i : i=1:n\} \leq x) \\
&= P(X_1 \leq x, \ldots, X_n \leq x) && \text{※独立性} \\
&= P(X_1 \leq x) \cdot P(X_2 \leq x) \cdots P(X_n \leq x) \\
&= F_1(x) \cdot F_2(x) \cdots F_n(x)
\end{aligned}$$

変数変換された確率変数の確率分布

変数変換する ⇒ 合成関数を作る

⇒ 新しい確率分布を得る.

要請　$X \sim f_X(x)$ をもつ

唯一
逆関数

変数変換 $g(x)$：狭義単調増加,
連続微分可能 ではならない

表わすこと　$Y = g(X)$ の確率分布を表現すること.

やり方
(一般的)

$$P(Y \leq y) = P(g(X) \leq y)$$
$$= P(X \leq g^{-1}(y))$$
$$= \int_{-\infty}^{g^{-1}(y)} f_X(x)\,dx$$

$$F_Y(y) = F_X(g^{-1}(y))$$

$$F_Y(y) = F_X(g^{-1}(y))$$

dx/dy

$$f_Y(y) = f_X(g^{-1}(y))\,(g^{-1}(y))'$$

$f_X($

線形変換を例に

$$g(x) = ax+b \quad (a \neq 0)$$

$$x = g^{-1}(y) = (y-b)/a$$

$$dg^{-1}(y)/dy = 1/a$$

$$\therefore \quad f_Y(y) = f_X\left(\frac{y-b}{a}\right)\frac{1}{|a|}$$

(変数変換による分布関数)

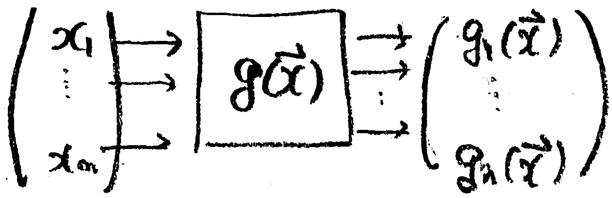
多次元での表現

前提

$\vec{X} = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$ は連続型.

確率分布関数は $f_{\vec{X}}(\vec{x})$ をもつ.

変数変換関数 $\vec{g}(\vec{x}) : \mathbb{R}^m \to \mathbb{R}^n$



後で考えよう

2019.9.10

$\vec{g}(\vec{x})$ の行列表現

$$\begin{bmatrix} \lambda x . g_{11}(x) \\ \\ \\ \\ \end{bmatrix}$$

$\vec{Y} = \vec{g}(\vec{X})$ の分布関数

$$P(\vec{Y} \le \vec{y}) = P(\vec{g}(\vec{X}) \le \vec{y})$$

$$= \int_{\vec{g}(\vec{x}) \le \vec{y}} \vec{f}_{\vec{X}}(\vec{x}) d\vec{x}$$

$$= \int_{\vec{z} \le \vec{y}} \vec{f}_{\vec{X}}(\vec{g}^{-1}(\vec{z})) \left| J(\vec{g}^{-1}(\vec{z})) \right| d\vec{z}$$

$$\vec{z} = \vec{g}(\vec{x})$$

$$P(\vec{Y} \le \vec{y}) = F_{\vec{Y}}(\vec{y})$$

$$F_{\vec{Y}}(\vec{y}) = \int_{\vec{z} \le \vec{y}} \vec{f}_{\vec{X}}(\vec{g}^{-1}(\vec{z})) \left| J(\vec{g}^{-1}(\vec{z})) \right| d\vec{z}$$

両辺を微分して

$$f_{\vec{Y}}(\vec{y}) = \vec{f}_{\vec{X}}(\vec{g}^{-1}(\vec{y})) \left| J(\vec{g}^{-1}(y)) \right|$$

$J(\vec{g}(\vec{x}))$

多変数ベクトル値関数の勾配

$$\vec{J_{\vec{f}}} = D_{\vec{x}} \cdot \vec{f}, \qquad \vec{f} = \begin{pmatrix} f_1(\vec{x}) \\ \vdots \\ f_n(\vec{x}) \end{pmatrix}$$

$$\begin{array}{c} \partial\vec{x} \\ \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{pmatrix} \end{array} (\vec{f})^T$$

$$= \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & & & \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix}$$

$$\left(\frac{\partial \vec{f}}{\partial x_1}, \frac{\partial \vec{f}}{\partial x_2}, \cdots, \frac{\partial \vec{f}}{\partial x_n} \right)$$

$\vec{x}$のx_n方向微分

Wikipediaの表記
僕は $\vec{x}$ にしてない.

$$\frac{\partial \vec{f}}{\partial \vec{x}}(\vec{p}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\vec{p}) & \cdots & \frac{\partial f_1}{\partial x_n}(\vec{p}) \\ \vdots & \ddots & \\ \frac{\partial f_n}{\partial x_1}(\vec{p}) & \cdots & \frac{\partial f_n}{\partial x_n}(\vec{p}) \end{bmatrix}$$

$\vec{p}$ はある点.

t分布の密度fの導出

前提

$X_1 \sim N(0,1)$

$X_2 \sim \chi^2_n$

X_1とX_2は独立

変数変換 Y_1

$Y_1 = X_1 / \sqrt{X_2/n}$ の密度fは?

$$\vec{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \vec{g}(\vec{x}) = \begin{pmatrix} x_1/\sqrt{x_2/n} \\ x_2 \end{pmatrix}$$

逆変換

$$\vec{x} = \vec{g}^{-1}(\vec{y}) = \begin{pmatrix} y_1\sqrt{y_2/n} \\ y_2 \end{pmatrix}$$

$$J(\vec{g}^{-1}(\vec{y})) = \det\left(\frac{\partial \vec{g}^{-1}(\vec{y})}{\partial \vec{y}^t} \right)$$

$$= \det$$