

統計検定直前対策講座
1級 統計数理
－ 変数変換 －

茨城大学 工学部 情報工学科
新納浩幸

変数変換(1次元)

X : 確率変数
 $f(x)$: 密度関数
 $g : R \rightarrow R$



$g(X)$: 確率変数



分布はどうなる?
平均はどうなる?
分散はどうなる?

X 、 f 、 g からうまく求めたい

変数変換 (1次元)

$Y = g(X)$ の密度関数 $h(y)$

$$h(y) = f(x) \frac{dx}{dy}$$

ここも g の逆関数
が必要



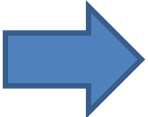
$y = g(x)$ なので、

$$h(y) = f(g^{-1}(y)) \frac{dx}{dy}$$

例)

$X : (0,1)$ の一様分布 $f(x) = 1$ 、

$Y = X^2$ の密度関数 $h(y)$ は？

 $y = x^2$ なので $x = y^{1/2}$

$$\frac{dx}{dy} = \frac{1}{2} y^{-1/2}$$

$$h(y) = 1 \cdot \frac{1}{2} y^{-1/2} = \frac{1}{2\sqrt{y}}$$

平均と分散

$$\begin{aligned} E(Y) &= E(g(X)) = \int yh(y)dy \\ &= \int g(x)f(x)\frac{dx}{dy}dy = \int g(x)f(x)dx \end{aligned}$$

$$\begin{aligned} V(Y) &= V(g(X)) = \int (y - E(Y))^2 h(y)dy \\ &= \int (g(x) - E(Y))^2 f(x)\frac{dx}{dy}dy = \int (g(x) - E(Y))^2 f(x)dx \end{aligned}$$

Y の密度関数 $h(y)$ を
必要としない

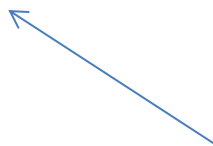
1次変換の平均と分散

$$E(aX + b) = aE(X) + b$$

$$V(aX + b) = a^2V(X)$$

注意！

aは2乗で外に出て、
bは消える




変数変換(2次元)

(X, Y) : 確率変数

$f(x, y)$: 同時密度関数

$$g_1 : (R, R) \rightarrow R$$

$$g_2 : (R, R) \rightarrow R$$

 $(U, V) = (g_1(X, Y), g_2(X, Y))$

: 確率変数

分布はどうなるか？

変数変換 (2次元)

(U, V) の同時密度関数 $h(u, v)$

$$h(u, v) = f(x, y) \| J \|$$

$$J = J \left(\frac{\partial(x, y)}{\partial(u, v)} \right) = \begin{pmatrix} \frac{\partial g_1^{-1}}{\partial u} & \frac{\partial g_1^{-1}}{\partial v} \\ \frac{\partial g_2^{-1}}{\partial u} & \frac{\partial g_2^{-1}}{\partial v} \end{pmatrix}$$

ヤコビアン
の行列式の
絶対値

ヤコビアンの注意

* 要素の順序は行列式の符号だけを変化させる、絶対値には影響しない
(u, v) でも (v, u) でもどちらでもよい

* 転置行列の行列式は不変

$$|J| = |J'|$$

(x, y), (u, v) どちらを行にするか列にするかは関係ない

* 逆関数を求めなくてもヤコビアンは求まる

$$\left| J \left(\frac{\partial(x, y)}{\partial(u, w)} \right) \right| = \frac{1}{\left| J \left(\frac{\partial(u, w)}{\partial(x, y)} \right) \right|}$$

どちらで微分しているか
大注意！！
また最終的には
逆関数は必要

代表的利用例

- * 確率変数のたたみ込みの公式
- * 正規分布の正規化定数の算出
- * ベータ関数とガンマ関数の関係

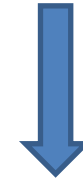


2012-問2, 2014-問2

たたみ込みの公式

関数 f を平行移動しながら関数 g を重ね足し合わせる二項演算

X, Y : 独立な確率変数



$Z = X + Y$ の分布を X と Y の **たたみ込み** から求める

$\left\{ \begin{array}{l} W = X \\ Z = X + Y \end{array} \right.$ により変数変換

$\frac{\partial(w, z)}{\partial(x, y)} = \begin{vmatrix} 1 & 0 \\ 1 & 1 \end{vmatrix} = 1$ の**逆数**がヤコビアン

$h(w, z) = f_X(w) f_Y(z - w)$ w で周辺化する

$$h(z) = \int f_X(w) f_Y(z - w) dw = \int f_X(x) f_Y(z - x) dx$$

たたみ込みによる正規分布の和

$$X \sim N(\mu_x, \sigma_x^2), Y \sim N(\mu_y, \sigma_y^2) : \text{独立}$$

$$Z = X + Y$$

$$h(z) = \int f_X(x) f_Y(z-x) dx$$

$$\begin{aligned} f_X(x) f_Y(z-x) &= \frac{1}{\sqrt{2\pi}\sigma_x} \exp\left(-\frac{(x-\mu_x)^2}{2\sigma_x^2}\right) \frac{1}{\sqrt{2\pi}\sigma_y} \exp\left(-\frac{(z-x-\mu_y)^2}{2\sigma_y^2}\right) \\ &= \frac{1}{2\pi\sigma_x\sigma_y} \exp\left(-\frac{Q}{2}\right) \end{aligned}$$

$$Q = \frac{\sigma_x^2 + \sigma_y^2}{\sigma_x^2 \sigma_y^2} \left\{ (x - \mu_x) - \frac{\sigma_y^2}{\sigma_x^2 + \sigma_y^2} (z - \mu_x - \mu_y) \right\} + \frac{1}{\sigma_x^2 + \sigma_y^2} (z - \mu_x - \mu_y)^2$$

$$h(z) = \int f_X(x) f_Y(z-x) dx$$

$$= \frac{1}{2\pi\sigma_x\sigma_y} \int \exp\left(-\frac{Q}{2}\right) dx$$

$$= \frac{1}{2\pi\sigma_x\sigma_y} \exp\left(-\frac{(z - \mu_x - \mu_y)^2}{2(\sigma_x^2 + \sigma_y^2)}\right)$$

$$\times \int \exp \frac{\sigma_x^2 + \sigma_y^2}{\sigma_x^2 \sigma_y^2} \left\{ (x - \mu_x) - \frac{\sigma_x^2}{\sigma_x^2 + \sigma_y^2} (z - \mu_x - \mu_y) \right\}^2 dx$$

$$= \frac{1}{\sqrt{2\pi(\sigma_x^2 + \sigma_y^2)}} \exp\left(-\frac{(z - \mu_x - \mu_y)^2}{2(\sigma_x^2 + \sigma_y^2)}\right)$$

$$\times \frac{\sqrt{\sigma_x^2 + \sigma_y^2}}{\sqrt{2\pi\sigma_x^2\sigma_y^2}} \int \exp\left[-\frac{\sigma_x^2 + \sigma_y^2}{2\sigma_x^2\sigma_y^2} \left\{ (x - \mu_x) - \frac{\sigma_x^2}{\sigma_x^2 + \sigma_y^2} (z - \mu_x - \mu_y) \right\}^2\right] dx$$

最後の積分の中は

$$N\left(\mu_x + \frac{\sigma_x^2}{\sigma_x^2 + \sigma_y^2} (z - \mu_x - \mu_y), \frac{\sigma_x^2 \sigma_y^2}{\sigma_x^2 + \sigma_y^2}\right)$$

の確率密度関数なので、積分値は 1

$$h(z) = \frac{1}{\sqrt{2\pi(\sigma_x^2 + \sigma_y^2)}} \exp\left(-\frac{(z - \mu_x - \mu_y)^2}{2(\sigma_x^2 + \sigma_y^2)}\right)$$



$$Z = X + Y \sim N(\mu_x + \mu_y, \sigma_x^2 + \sigma_y^2)$$

正規分布の正規化定数の算出

$$\int_{-\infty}^{\infty} e^{-x^2/2} dx = \frac{1}{c} \quad \text{と} \quad \text{おいて} \quad c = \frac{1}{\sqrt{2\pi}} \quad \text{となることの証明}$$

X, Y 独立、密度関数 $ce^{-x^2/2}$ を考える

$$\left\{ \begin{array}{l} X = r \cos \theta \\ Y = r \sin \theta \end{array} \right. \quad \text{により変数変換} \quad \leftarrow \text{逆関数で定義}$$

$$\frac{\partial(x, y)}{\partial(r, \theta)} = \begin{vmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{vmatrix} = r(\cos^2 \theta + \sin^2 \theta) = r \quad \text{がヤコビアン}$$

(X, Y) の同時密度関数は $c^2 e^{-(x^2+y^2)/2}$

(r, θ) の同時密度関数は $f(r, \theta) = c^2 r e^{-r^2/2}$

全体で積分する、極座標の変数変換なので範囲に注意、

$$\begin{aligned} 1 &= \iint f(r, \theta) d\theta dr = c^2 \int_0^\infty \int_0^{2\pi} r \exp(-r^2 / 2) d\theta dr \\ &= 2\pi c^2 \int_0^\infty r \exp(-r^2 / 2) dr = 2\pi c^2 \left[-\exp(-r^2 / 2) \right]_0^\infty = 2\pi c^2 \end{aligned}$$

➡ $c = \frac{1}{\sqrt{2\pi}}$

➡ $\int_{-\infty}^\infty e^{-x^2/2} dx = \sqrt{2\pi}$

偶関数なので

➡ $\int_0^\infty e^{-x^2/2} dx = \sqrt{\frac{\pi}{2}}$

$t = \frac{x}{\sqrt{2}}$

で変数変換して

➡ $\int_0^\infty e^{-x^2} dx = \sqrt{\pi}$



通常の公式

ガンマ関数とガンマ分布

$$\Gamma(m) = \int_0^{\infty} x^{m-1} e^{-x} dx \quad \text{ガンマ関数}$$

$$(1) \Gamma(1) = 1$$

$$(2) \Gamma(m+1) = m\Gamma(m)$$

$$(3) \Gamma(m) = (m-1)!$$

ただし m が整数のとき

ガンマ分布の密度関数

$$f(x) = x^{k-1} \frac{e^{-x/\theta}}{\Gamma(k)\theta^k}$$

k : 形状母数

θ : 尺度母数

$$X \sim Ga(k, \theta)$$

$$E(X) = k\theta$$

$$V(X) = k\theta^2$$

ガンマ分布と指数分布の関係

$$X \sim Ga(1, \theta) \quad \longrightarrow \quad f(x) = \frac{e^{-x/\theta}}{\theta} \quad \text{指数分布 } Ex(\theta)$$

ガンマ分布とカイ2乗分布の関係

$$X \sim Ga\left(\frac{n}{2}, 2\right) \quad \longrightarrow \quad f(x) = x^{n/2-1} \frac{e^{-x/2}}{\Gamma(n/2)2^{n/2}}$$

自由度 n のカイ2乗分布

ベータ関数とベータ分布

$$B(a, b) = \int_0^1 x^{a-1} (1-x)^{b-1} dx \quad \text{ベータ関数}$$

$$B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

これを示すのに
変数変換を利用

ベータ分布の密度関数

$$f(x) = \frac{x^{a-1} (1-x)^{b-1}}{B(a, b)}$$

a, b : 形状母数
 $X \sim Be(a, b)$

$$E(X) = \frac{a}{a+b} \quad V(X) = \frac{ab}{(a+b)^2 (a+b+1)}$$

ガンマ分布とベータ分布の関係

$X \sim Ga(a,1)$ 、 $Y \sim Ga(b,1)$ は独立とする

$$\left\{ \begin{array}{l} U = \frac{X}{X+Y} \quad \longrightarrow \quad U \sim Be(a,b) \\ V = X+Y \quad \longrightarrow \quad V \sim Ga(a+b,1) \end{array} \right.$$

最重要

上記は変数変換を用いて証明され、同時に以下のこともわかる

- * UとVは独立
- * ガンマ関数とベータ関数の関係
- * ガンマ分布の再生性 (和が同じ分布になる)

その他、正規分布、2項分布、ポアソン分布、
コーシー分布なども再生性をもつ

$$\begin{cases} U = \frac{X}{X+Y} \\ V = X+Y \end{cases}$$



逆関数を求めてから偏微分して
ヤコビアンを求める

$$\begin{cases} X = UV \\ Y = V(1-U) \end{cases}$$

$$\frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} v & u \\ -v & 1-u \end{vmatrix} = v(1-u) + uv = v$$

これがヤコビアン

(X, Y) の同時確率は X と Y が独立であることから

$$f_{XY}(x, y) = f_X(x) f_Y(y) = x^{a-1} \frac{e^{-x}}{\Gamma(a)} y^{b-1} \frac{e^{-y}}{\Gamma(b)}$$

(U, V) の同時確率は

$$x^{a-1} \frac{e^{-x}}{\Gamma(a)} y^{b-1} \frac{e^{-y}}{\Gamma(b)} \cdot v = \frac{1}{\Gamma(a)\Gamma(b)} (uv)^{a-1} (v(1-u))^{b-1} e^{-uv-v(1-u)} \cdot v$$

$$= \frac{1}{\Gamma(a)\Gamma(b)} u^{a-1} v^{a+b-1} (1-u)^{b-1} e^{-v}$$

$$= \frac{1}{\Gamma(a+b)} v^{a+b-1} e^{-v} \cdot \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} u^{a-1} (1-u)^{b-1}$$



v のみの関数
Ga(a+b) の密度関数

u のみの関数
Be(a,b) の密度関数

変数が分離されているので独立、
積分することで

$$B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} \quad \text{も示される}$$

過去問

どちらも非常に良問
変数変換がポイント

2012-問2

カイ2乗分布の再生性

(モーメント母関数を利用)

カイ2乗分布とベータ分布の関係

2014-問2

ガンマ分布の再生性

ガンマ分布とベータ分布の関係