

## 変数変換

### 変数変換による確率密度関数の変化

連続確率変数 $X$ の確率密度関数を $f(x)$ とする.ここで $X$ と1対1対応する新たな確率変数 $Y = g(X)$ について考える.ただし $g(X)$ は微分可能で単調増加する関数とする.

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

左辺は $Y$ の累積分布関数,右辺は $X$ の累積分布関数なので

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

両辺を微分して

$$\begin{aligned} f_Y(y) &= \frac{d}{dy} F_X(g^{-1}(y)) \\ &= f_X(g^{-1}(y)) \frac{d}{dy} g^{-1}(y) \\ &= \frac{f_X(g^{-1}(y))}{g'(x)} \\ &= \frac{f_X(g^{-1}(y))}{|g'(g^{-1}(y))|} \quad (\text{because } x = g^{-1}(y)) \end{aligned}$$

が得られた.単調減少の場合 $g'(x) < 0$ となるため、一般化して絶対値で表現されている.

2変数 $X, Y$ の確率密度関数を $f(x, y)$ とし,変数変換 $(Z, W) = (u(X, Y), v(X, Y))$ について考える.ただし逆変換 $(X, Y) = (s(Z, W), t(Z, W))$ が存在するものとする.この変換のヤコビアンは

$$J(z, w) = \frac{\partial(x, y)}{\partial(z, w)} = \det \begin{pmatrix} \frac{\partial s}{\partial z} & \frac{\partial s}{\partial w} \\ \frac{\partial t}{\partial z} & \frac{\partial t}{\partial w} \end{pmatrix}$$

と計算される. $(Z, W)$ の確率密度関数は

$$f_{Z, W}(z, w) = f_{X, Y}(s(z, w), t(z, w)) |J(z, w)|$$

として与えられる.

### 確率変数の線形結合の分布

2つの独立な確率変数 $X, Y$ についてその線形結合 $aX + bY$ の分布について考える. $aX + bY$ のモーメント母関数 $E[e^{\theta(aX+bY)}] = E[e^{\theta aX}]E[e^{\theta bY}]$ を計算し,既存の分布のモーメント母関数

と一致するか調べる方法がある。しかしこの方法では導出されたモーメント母関数が未知の場合求めたい分布 $aX + bY$ を求める事ができない。そこで先程の変数変換を用いて導出する事ができる。

$Z = aX + bY, W = Y$ とし $(Z, W)$ の分布を考える。逆変換は $(X, Y) = (\frac{Z-bW}{a}, W)$ でありヤコビアンは

$$J(X, Y) = \left| \begin{pmatrix} \frac{\partial Z}{\partial X} & \frac{\partial Z}{\partial Y} \\ \frac{\partial W}{\partial X} & \frac{\partial W}{\partial Y} \end{pmatrix} \right| = \left| \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} \right| = a$$

となる。

$(Z, W)$ の確率密度関数は $(X, Y)$ が独立なので

$$f_{Z,W}(z, w) = \frac{f_X(\frac{z-bw}{a}) f_Y(w)}{|a|}$$

になる。 $W$ について積分することで $Z$ の確率密度関数が得られる

$$f_Z = \frac{1}{|a|} \int_{-\infty}^{\infty} f_X\left(\frac{z-bw}{a}\right) f_Y(w) dw$$


---

積が積み重なることで得られるデータは対数を取ることで正規分布になることがある。この際は対数変換を行うとよい。人口、株価、所得などのデータには対数変換が行われる事が多い。目的変数が $x > 0$ の場合にのみ使える。

裾野が長い極端な分布を正規分布に従うようにする方法としてはべき乗変換もある。 $x^a$ という変換をおこなう。この際 $a$ の値をどのような値にするかも重要な問題となる。

- ex) 簡単なテストの点数の分布を取りたい場合に $a = 2$ の変換を行うことで僅かな点数の違いを大きくして正規分布に近づける。

べき乗変換と対数変換をひとまとめにした変換として **Box-Cox 変換**がある。パラメータ $\lambda$ に 対して

$$y(\lambda) = \begin{cases} \frac{x^\lambda - 1}{\lambda} (\lambda \neq 0) \\ \log x (\lambda = 0) \end{cases}$$

という変換である。 $\lambda \neq 0$ の場合 $x = 1$ 周辺を0付近にして $x = 1$ 付近で微分が1になり, $\lambda = 0$ の場合対数変換と滑らかにつながり微分可能なような変換となっている。最も正規分布に当てはまりのよい $\lambda$ を探索することでデータにフィットする変換を探索することができる。

確率 $p$ のような0から1の値しかとらないものを $-\infty$ から $\infty$ を取る値に変換したいときはロジット変換 $\log \frac{p}{1-p}$ を行う。ロジット変換を行った分布に対して線形回帰を行い予測した値を逆変換することで確率に対して回帰を行うという手法がロジスティック回帰である。これは $p$ を $a + bx$ のロジスティック変換で表すことと同値である。

$$\frac{1}{1 + e^{-(a+bx)}}$$

0から1までの範囲しか取らないものを $-\infty$ から $\infty$ を取る値に変換する方法として**プロビット変換**がある。プロビット変換は標準正規分布の累積分布関数 $\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt$ の逆関数 $\Phi^{-1}(x)$ によって変換する方法である。