

確率論 (Probability Theory) 第 13 週 正規分布

w.namamoto

1 確率変数の変換 (1 次元)

確率変数 X が確率密度関数 $f_x(x)$ を持つ確率分布に従うとする。

1.1 位置変換

確率分布の平均を μ だけ横にずらすことを、位置変換と呼ぶことがある。確率変数 Y の従う確率分布が、確率変数 X の従う確率分布から μ だけずれているとは、

$$Y \sim X + \mu \quad (1)$$

という関係にあることを言う。連続確率変数の場合には、 X と Y の標本空間は共通でなければならない。このとき Y の密度関数は

$$y = x + \mu \quad (2)$$

かつ $x \in \mathcal{X} \Rightarrow y \in \mathcal{Y}$ から、

$$f_y(y) = f_x(x - \mu) \quad (3)$$

となる。

これは実は、確率変数の関数の期待値を計算する際に、 $y = x + \mu$ と変数変換をすると、 $dx = dy$ と積分領域が変わらないことから

$$\int_{\mathcal{X}} g(y) f_y(y) dy = \int_{\mathcal{X}} g(y) f_x(y - \mu) dy \quad (4)$$

と右辺で積分ができることと同等である。

このことを以て、確率分布の μ による位置変換後の確率変数 Y の確率密度関数は $f(y - \mu)$ である、という。

1.2 尺度変換

確率変数の原点からの距離 (あるいは値) を $1/\sigma$ 倍することを、尺度変換と呼ぶことがある。 σ 倍することも、同じく尺度変換という。¹

確率論では、

1. 期待値が μ の非負の確率変数 X について、期待値を 1 にする変換: $Y \sim X/\mu$
2. 期待値が 0 で分散が σ^2 の確率変数 X について、分散を 1 にする変換: $Y \sim X/\sigma$

という方向の尺度変換と

1. 非負で期待値が 1 の確率変数 Y について、期待値を μ にする変換: $X \sim Y\mu$
2. 期待値が 0 で分散が 1 の確率変数 Y について、分散を σ^2 にする変換: $X \sim Y\sigma$

¹たとえば m 単位を km 単位に変換するときに $1/1000$ 倍することと、 km 単位を m 単位に変換するときに 1000 倍すること、どちらか単位の変換と言われるのと同じこと。

という上とは逆方向の尺度変換をよく見かける。いずれにせよ変換のヤコビアンは

$$dx = \mu dy \quad dy = \frac{1}{\mu} dx$$

あるいは

$$dx = \frac{1}{\sigma} dy \quad dy = \sigma dx$$

となる。

例えば指数分布の密度関数には

$$f(x) = \frac{1}{\mu} e^{-x/\mu} = \frac{\lambda^{-\lambda x}}{e} \quad (5)$$

のふたつの表現がある。平均が 1 の指数分布の密度関数が

$$f_1(z) = e^{-z} \quad (6)$$

と見比べると、式 (5) の真ん中の表現は Z を μ 倍すると期待値も μ になるという意味であり、右側の表現は Z を $1/\lambda$ 倍すると平均が $1/\lambda$ の指数分布になるという意味である。参考までに両者が同じ確率分布であるためには、 $\mu = 1/\lambda$ が成り立たなければならない。

確率変数 Y の尺度が、確率変数 X の尺度の σ 倍であるとは、

$$Y/\sigma \sim X \quad \text{or} \quad Y \sim X\sigma \quad (7)$$

という関係にあることを言う。このとき、連続確率変数の場合には、 X と Y の標本空間は共通であるなら、 Y の密度関数は

$$y = x\sigma \quad (8)$$

かつ $x \in X \Rightarrow y \in \mathcal{Y}$ から、

$$f_y(y) = \frac{1}{\sigma} f_x\left(\frac{y}{\sigma}\right) \quad (9)$$

となる。

これは実は、確率変数の関数の期待値を計算する際に、 $y = x\sigma$ と変数変換をすると、 $dx = dy/\sigma$ であり、ここでも積分領域が変わらないことから

$$\int_X g(y) f_y(y) dy = \int_X g(y) \frac{1}{\sigma} f_x\left(\frac{y}{\sigma}\right) dy \quad (10)$$

と右辺で積分ができることと同等である。

このことを以て、確率分布の μ による位置変換後の確率変数 Y の確率密度関数は $f_x(y/\sigma)$ である、という。

1.3 一般の変換

一般の変換の場合には、確率変数 X の密度関数を $f_x(x)$ とすると、一階微分可能かつ狭義単調な関数 $h(\cdot)$ を用いて $Y \sim h(X)$ という関係にある確率変数 Y の密度関数は、

$$f_y(y) = \frac{1}{\left| \frac{\partial h^{-1}(y)}{\partial x} \right|} f_x(h^{-1}(y)) \quad (11)$$

となる。

2 正規分布

正規分布は「誤差の分布」として知られている。ある量 X を計るとする。重量計に乗せてメモリを読む、ということは何回か繰り返して測定値を複数個、 X_1, X_2, \dots, X_n と得て、それらの平均 \bar{X}_n を X の測定結果、とする実験を考える。このとき \bar{X}_n は、 n が大きくなればなるほど、真値 X を中心とし、分散が σ^2/n と測定回数に反比例して小さくなる、正規分布に従うようになる。このことは中心極限定理によって正当化されるが、詳細は後に譲る。正規分布の密度関数は、次のように位置変換と尺度変換を用いて、得る。

平均が 0、分散が 1 の正規分布の確率密度関数は

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \quad (12)$$

で与えられる。 $\frac{1}{\sqrt{2\pi}}$ は、全積分を 1 にするための乗数で、基準化定数と呼ばれることがある。

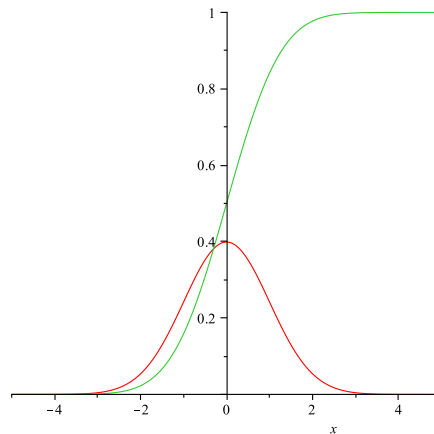


図 1: 標準正規分布の密度関数と累積分布関数

まず、平均が μ で分散が 1 の正規分布の密度関数は、位置変換を用いて、

$$f(x; \mu) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2}\right) \quad (13)$$

で与えられることが分かる。

次に、分散を σ^2 倍するということが尺度は σ 倍することに他ならない

$$V[aX] = a^2 V[X] \quad (14)$$

ことを思い出せば、平均が μ で分散が σ^2 の正規分布の密度関数は、

$$f(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (15)$$

で与えられることが分かる。指数の肩に入った $1/\sigma^2$ は $x-\mu$ を $(x-\mu)/\sigma$ に置き換えることによって出現し、指数の外の $1/\sigma$ は尺度変換のヤコビアンとして出現していることに注意する。これが、平均 μ 、分散 σ^2 の正規分布の密度関数である。

2.1 e^{-x^2} の積分

e^{-x^2} の曲線は美しいベル型を描くが、その不定積分は誤差関数と呼ばれ、解析的に陽な表現を持たない。また $(-\infty, \infty)$ 上の定積分は計算でき、 $\sqrt{\pi}$ となるが、計算には極座標に持ち込む意気込みが必要となる。

まず $I = \int_{-\infty}^{\infty} \exp(-x^2) dx$ と置く。指数の肩で円を描かせたいので、 $\int_{-\infty}^{\infty} \exp(-y^2) dx$ との積を考える。これも I

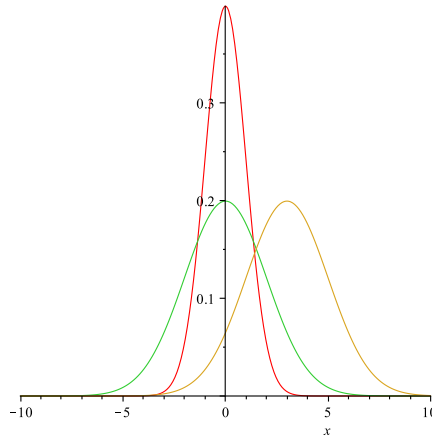


図 2: 標準正規分布の密度関数と累積分布関数

に等しいので、

$$\begin{aligned}
 I^2 &= \left\{ \int_{-\infty}^{\infty} \exp(-x^2) dx \right\} \left\{ \int_{-\infty}^{\infty} \exp(-y^2) dy \right\} \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp(-x^2) \exp(-y^2) dx dy \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp(-x^2 - y^2) dx dy \tag{16}
 \end{aligned}$$

2行目は積分順序を交換しただけ、3行目は指数の肩をまとめただけ。これを極座標表現 $(x, y) = (r \cos \theta, r \sin \theta)$ に書き換えると、

$$I^2 = \int_0^{2\pi} \int_0^{\infty} \exp(-r^2) r dr d\theta \tag{17}$$

となる。ここで被積分関数に現れた r は、以下に示すように、 $dx dy$ から $dr d\theta$ に変換するときの、ヤコビ行列の行列式である。

$$\begin{aligned}
 \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} &= \begin{vmatrix} \frac{\partial r \cos \theta}{\partial r} & \frac{\partial r \cos \theta}{\partial \theta} \\ \frac{\partial r \sin \theta}{\partial r} & \frac{\partial r \sin \theta}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{vmatrix} \\
 &= r(\cos^2 \theta + \sin^2 \theta) = r \tag{18}
 \end{aligned}$$

これより、

$$dx dy = \left| \frac{\partial(x, y)}{\partial(r, \theta)} \right| dr d\theta = r dr d\theta \tag{19}$$

と分かる。

さて、積分を続ける。被積分関数に θ が含まれないので、

$$\begin{aligned}
 I^2 &= \int_0^{2\pi} d\theta \int_0^{\infty} \exp(-r^2) r dr \\
 &= 2\pi \int_0^{\infty} \exp(-r^2) r dr \tag{20}
 \end{aligned}$$

ここで $u = r^2$ (つまり $r = \sqrt{u}$) と置換すると、

$$du = 2r dr \Leftrightarrow dr = \frac{1}{2r} du \tag{21}$$

かつ

$$u : 0 \rightarrow \infty \quad \text{when} \quad r : 0 \rightarrow \infty \tag{22}$$

よって、 I^2 は

$$I^2 = 2\pi \int_0^{\infty} \exp(-u) \frac{1}{2} du = \pi \int_0^{\infty} \exp(-u) du \tag{23}$$

となる。この積分は簡単で、

$$\int \exp(-u) du = -\exp(-u) + C \quad (24)$$

より、

$$I^2 = \pi [-\exp(-u)]_0^\infty = \pi [0 - (-1)] = 2\pi \quad (25)$$

を得る。

以上より、

$$I = \int_{-\infty}^{\infty} \exp(-x^2) dx = \sqrt{\pi} \quad (26)$$

となる。

2.2 誤差関数

誤差関数 $\operatorname{erf}(x)$ は $[0, x]$ 上の e^{-t^2} の定積分の $2/\sqrt{\pi}$ 倍として定義される。

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \quad (27)$$

数学では特殊関数の 1 つに数えられている。この関数は陽は表現を持たないが、被積分関数 e^{-t^2} を原点の回りでテイラー展開し、項別に定積分を求めれば、誤差関数自体のテイラー級数が得られる。

e^{-t^2} は偶関数であり、奇数次数の項はすべて消えることに注意しつつ、

$$\begin{aligned} \exp(-t^2) &= \sum_{k=0}^{\infty} \frac{1}{k!} (-1)^k t^{2k} \\ &= 1 - t^2 + \frac{1}{2}t^4 - \frac{1}{6}t^6 + \frac{1}{24}t^8 + O(t^{10}) \end{aligned} \quad (28)$$

となる。これを項別に積分すると、

$$\begin{aligned} \int_0^x \exp(-t^2) dt &= \sum_{k=0}^x \frac{1}{k!} (-1)^k \int_0^x t^{2k} dt \\ &= \sum_{k=0}^x \frac{1}{k!} (-1)^k \left[\frac{t^{2k+1}}{2k+1} \right]_0^x \\ &= \sum_{k=0}^x \frac{1}{k!} (-1)^k \frac{x^{2k+1}}{2k+1} \end{aligned} \quad (29)$$

となり、

$$\begin{aligned} \operatorname{erf}(x) &= \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \\ &= \frac{2}{\sqrt{\pi}} \sum_{k=0}^x \frac{1}{k!} (-1)^k \frac{x^{2k+1}}{2k+1} \\ &= \frac{2}{\sqrt{\pi}} \left\{ x - \frac{x^3}{3} + \frac{x^5}{10} - \frac{x^7}{42} + O(x^{11}) \right\} \end{aligned} \quad (30)$$

を得る。

2.3 期待値の計算

一番簡単なのは、関数

$$(x - \mu) \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) \quad (31)$$

が、点 $(\mu, 0)$ に対して奇関数となることを用いると、

$$\int_{-\infty}^{\mu} (x - \mu) \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) dx = - \int_{\mu}^{\infty} (x - \mu) \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) dx \quad (32)$$

が示せて、これより

$$\begin{aligned} \int_{-\infty}^{\infty} (x - \mu) \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) dx &= \int_{-\infty}^{\mu} (x - \mu) \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) dx + \int_{\mu}^{\infty} (x - \mu) \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) dx \\ &= \int_{-\infty}^{\mu} (x - \mu) \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) dx - \int_{-\infty}^{\mu} (x - \mu) \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) dx \\ &= 0 \end{aligned} \quad (33)$$

となることから、

$$\int_{-\infty}^{\infty} (x - \mu) \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) dx = E[x - \mu] = 0 \quad (34)$$

を得る。以上から

$$E[X] = \mu \quad (35)$$

を得る。

2.4 正規分布の分散

分散 σ^2 の計算は、部分積分を用いるものと、置換積分を用いるものと、あとモーメント母関数をテイラー展開するもの、があるらしい。以下は部分積分。

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

と

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) dx \\ &= \mu \end{aligned}$$

は所与とする。

$$\begin{aligned} V[X] &= E[(X - \mu)^2] \\ &= \int_{-\infty}^{\infty} (x - \mu)^2 \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) dx \\ &= \int_{-\infty}^{\infty} \{-(x - \mu)\} \left[\{-(x - \mu)\} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) \right] dx \end{aligned}$$

ここで

$$\int \{-(x - \mu)\} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) dx = \sigma^2 \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) + C$$

より、 $\{-(x - \mu)\}$ と $\{-(x - \mu)\} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right)$ について部分積分を用いれば

$$\begin{aligned} V[X] &= \left[\sigma^2 (x - \mu)^2 \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) \right]_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \sigma^2 \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) dx \\ &= \sigma^2 \end{aligned}$$

を得る。

2.5 モーメント母関数

分散やそれ以上のモーメントの計算も可能である。しかし、正規分布のモーメント母関数は

$$\begin{aligned}M_X(t) &= E[\exp(tX)] = \int_{-\infty}^{\infty} \exp(tx) \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx \\&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2} + tx\right) dx \\&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2 - 2\mu x + \mu^2 - 2\sigma^2 xt}{2\sigma^2}\right) dx \\&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2 - 2(\mu + \sigma^2 t)x + \mu^2}{2\sigma^2}\right) dx \\&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - \mu - \sigma^2 t)^2 - 2\mu\sigma^2 t - \sigma^4 t^2}{2\sigma^2}\right) dx \\&= \exp\left(\mu t + \frac{\sigma^2}{2} t^2\right) \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - \mu - \sigma^2 t)^2}{2\sigma^2}\right) dx \\&= \exp\left(\mu t + \frac{\sigma^2}{2} t^2\right)\end{aligned}\tag{36}$$

のように、積分を変形するだけで得られる。そのため実際には、モーメント母関数をテイラー展開する方が、速い。

2.6 正規分布の和の分布

モーメント母関数を用いて、次の関係が得られる。

$$N(\mu_1, \sigma_1^2) + N(\mu_2, \sigma_2^2) \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

3 レポート課題

#12-1 参考書の p.196 にあるのは、平均が 0、分散が 1 の (標準) 正規分布の $Pr[X \geq x] = \alpha$ となる x の値の表である。これを利用して、次の確率を求めてみよ。表の読み方のメモを残すこと。

- (1) $N(5, 2^2)$ の上側 5% 点 ($Pr[X \geq a] = 0.05$ となる a の値のこと)
- (1) $N(4, 3^2)$ の下側 5% 点 ($Pr[X \leq a] = 0.05$ となる a の値のこと)
- (1) $N(6, 4^2)$ で $X \geq 8$ となる確率 ($Pr[X \geq 8]$ の値のこと)

#12-2 正規分布 $N(\mu, \sigma^2)$ の分散を求めよ。(答えは σ^2 なので、計算過程を示せ、という問題)

#12-3 任意の 2 つの正規分布 $N(\mu_1, \sigma_1^2)$, $N(\mu_2, \sigma_2^2)$ の和の分布を求めよ。(これも答えは自明なので、計算過程を示せ、という問題)

#12-4 平均 0、分散 1 の正規分布に従う、互いに独立な確率変数 X_1, X_2 の同時分布を考える。

- (1) (X_1, X_2) の同時分布の確率密度関数を導け。
- (2) X_1 を σ_1 倍、 X_2 を σ_2 倍したときの、 $(\sigma_1 X_1, \sigma_2 X_2)$ の同時分布の確率密度関数を導け。
- (3) さらに原点回りに θ だけ反時計回りに回転したときの、回転後の同時分布の確率密度関数を導け。(線形代数における 2 次元ベクトルの原点回りの回転を思い出して)
- (4) さらに平均を (μ_1, μ_2) の移動させたときの、移動後の同時分布の確率密度関数を導け。

参考文献

- [1] 微分積分学の教科書.
- [2] 永田靖 (2005) 「統計学のための数学入門 30 講」朝倉書店.
- [3] 宮川雅巳 (1988) 「統計技法」共立出版.
- [4] 稲垣宣生 (2003) 「数理統計学」改訂版, 裳華房.