正規分布の和を直接計算

1　正規分布の密度関数

独立な２つの正規分布に対して、和がやはり正規分布になることを示す。証明の方法はいくつか知られているが、直接計算しても導かれることをつぎで示そう。

確率変数 X, Y は独立であり、$X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$ ならば、$X+Y \sim N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$

（証明）まず和の公式：

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(z-y) f_Y(y) dy$$

をもちいる。積分される関数は

$$\begin{aligned} f_X(z-y)f_Y(y) &= \frac{1}{2\pi\sigma_1\sigma_2}\exp\left(-\frac{(z-y-\mu_1)^2}{2\sigma_1^2} - \frac{(y-\mu_2)^2}{2\sigma_2^2}\right) \\ &= \frac{1}{2\pi\sigma_1\sigma_2}\exp\left(-\frac{1}{2}\{\frac{(z-y-\mu_1)^2}{\sigma_1^2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\}\right) \end{aligned}$$

この指数部を計算していくと、

$$\begin{aligned} \frac{(z-y-\mu_1)^2}{\sigma_1^2} + \frac{(y-\mu_2)^2}{\sigma_2^2} &= \frac{1}{\sigma_1^2}\left\{y^2 - 2(z-\mu_1)y + (z-\mu_1)^2\right\} + \frac{1}{\sigma_2^2}\left\{y^2 - 2\mu_2 y + \mu_2^2\right\} \\ &= \left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right)y^2 - 2\left(\frac{z-\mu_1}{\sigma_1^2} + \frac{\mu_2}{\sigma_2^2}\right)y + \frac{(z-\mu_1)^2}{\sigma_1^2} + \frac{\mu_2^2}{\sigma_2^2} \\ &= Ay^2 - 2By + C = A\Big(y - \frac{B}{A}\Big)^2 - \frac{B^2}{A} + C \end{aligned}$$

ここで　$A = \dfrac{1}{\sigma_1^2} + \dfrac{1}{\sigma_2^2} = \dfrac{\sigma_1^2+\sigma_2^2}{\sigma_1^2\sigma_2^2}$,　$B = \dfrac{z-\mu_1}{\sigma_1^2} + \dfrac{\mu_2}{\sigma_2^2}$,　$C = \dfrac{(z-\mu_1)^2}{\sigma_1^2} + \dfrac{\mu_2^2}{\sigma_2^2}$ とおく。さらにこれを計算する。

$$\begin{aligned} \frac{B^2}{A} - C &= \frac{\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2}\left(\frac{z-\mu_1}{\sigma_1^2} + \frac{\mu_2}{\sigma_2^2}\right)^2 - \left(\frac{(z-\mu_1)^2}{\sigma_1^2} + \frac{\mu_2^2}{\sigma_2^2}\right) \\ &= \frac{\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2}\left(\frac{(z-\mu_1)^2}{\sigma_1^4} + 2\frac{(z-\mu_1)\mu_2}{\sigma_1^2\sigma_2^2} + \frac{\mu_2^2}{\sigma_2^4}\right)^2 - \frac{(z-\mu_1)^2}{\sigma_1^2} - \frac{\mu_2^2}{\sigma_2^2} \\ &= \frac{1}{\sigma_1^2+\sigma_2^2}\frac{\sigma_2^2}{\sigma_1^2}(z-\mu_1)^2 + \frac{2}{\sigma_1^2+\sigma_2^2}(z-\mu_1)\mu_2 + \frac{1}{\sigma_1^2+\sigma_2^2}\frac{\sigma_1^2}{\sigma_2^2}\mu_2^2 - \frac{(z-\mu_1)^2}{\sigma_1^2} - \frac{\mu_2^2}{\sigma_2^2} \\ &= \left\{\frac{1}{\sigma_1^2+\sigma_2^2}\frac{\sigma_2^2}{\sigma_1^2} - \frac{1}{\sigma_1^2}\right\}(z-\mu_1)^2 + 2\frac{z-\mu_1}{\sigma_1^2+\sigma_2^2}\mu_2 + \left\{\frac{1}{\sigma_1^2+\sigma_2^2}\frac{\sigma_1^2}{\sigma_2^2} - \frac{1}{\sigma_2^2}\right\}\mu_2^2 \\ &= \frac{-1}{\sigma_1^2+\sigma_2^2} + 2\frac{z-\mu_1}{\sigma_1^2+\sigma_2^2}\mu_2 - \frac{-1}{\sigma_1^2+\sigma_2^2}\mu_2^2 \\ &= \frac{-1}{\sigma_1^2+\sigma_2^2}(z-\mu_1-\mu_2)^2 \end{aligned}$$

したがって指数部は

$$A\Big(y-\frac{B}{A}\Big)^2-\frac{B^2}{A}+C=\frac{\sigma_1^2+\sigma_2^2}{\sigma_1^2\sigma_2^2}\left(y-\frac{B}{A}\right)^2+\frac{-1}{\sigma_1^2+\sigma_2^2}(z-\mu_1-\mu_2)^2$$

正規分布の密度関数は積分をすると1になることから、

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

右辺は μ に依存しないことに注意する。つまり

$$\int_{-\infty}^{\infty}\exp\left(-\frac{1}{2}\frac{\sigma_1^2+\sigma_2^2}{\sigma_1^2\sigma_2^2}(y-\frac{B}{A})^2\right)dy=\sqrt{2\pi}\frac{\sigma_1\sigma_2}{\sqrt{\sigma_1^2+\sigma_2^2}}$$

したがって

$$\begin{aligned}
f_{X+Y}(z) &= \int_{-\infty}^{\infty}f_X(z-y)f_Y(y)dy\\
&= \frac{1}{2\pi\sigma_1\sigma_2}\int_{-\infty}^{\infty}\exp\left(-\frac{1}{2}\frac{\sigma_1^2+\sigma_2^2}{\sigma_1^2\sigma_2^2}(y-\frac{B}{A})^2\right)dy\times\exp\left(-\frac{(z-\mu_1-\mu_2)^2}{2(\sigma_1^2+\sigma_2^2)}\right)\\
&= \frac{1}{\sqrt{2\pi}\sqrt{\sigma_1^2+\sigma_2^2}}\exp\left(-\frac{(z-\mu_1-\mu_2)^2}{2(\sigma_1^2+\sigma_2^2)}\right)
\end{aligned}$$

最後の式は $X+Y\sim N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2)$ に対する密度関数であることを示している。(証明終わり)

問 1.1
定数 a,c に対し、確率変数 X が $X\sim N(\mu,\sigma^2)$ ならば、

$$aX+c\sim N(a\mu+c,a^2\sigma^2)$$

であることを示せ。

問 1.2
定数 a,b に対し、確率変数 X,Y は独立であり、$X\sim N(\mu_1,\sigma_1^2)$, $Y\sim N(\mu_2,\sigma_2^2)$ ならば、

$$aX+bY\sim N(a\mu_1+b\mu_2,a^2\sigma_1^2+b^2\sigma_2^2)$$

であることを示せ。

2 密度関数の計算

正規分布の密度関数が積分して、1に等しくなることの証明は初等的にはたいへん難しい。多く用いられるのは、ガンマ関数 $\Gamma(\alpha)=\int_0^{\infty}x^{\alpha-1}e^{-x}dx$, $\alpha>0$ において、$\Gamma(1/2)=\sqrt{\pi}$ となることがよく知られている。またこの積分値を計算のために $\phi(x)\phi(y)$ を考えて、この2変数関数 (x,y) の2重積分を極座標変換することも積分の練習問題としてよく使われる。ここではある関数を微分して定数になるから、恒等式の関係を満たすことがわかり、この極限を計算することで、正規分布の密度関数積分、あるいはパラメータ 1/2 のガンマ関数計算値として得られることを述べる。この関係式は S.B.Daniel(Cambridge Univ.) による。

密度関数 $\phi(x)=\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$ に対して $\int_{-\infty}^{\infty}\phi(x)dx=1$ を満たす。

（証明）すべての $x>0$ に対して

$$\left(\int_0^\infty e^{-t^2}dt\right)^2+\int_0^\infty \frac{e^{-x^2(1+t^2)}}{1+t^2}dt=\frac{\pi}{4}$$

を満たす。なぜならば、変数 x で微分をすると、左辺の第 1 項は $2e^{-x^2}\int_0^x e^{-t^2}dt$ であり、これは第 2 項を微分して変数変換したものに符号を変えたものに等しい。したがってその和はゼロとなる。すなわち元の第 1 項と第 2 項の積分の和は定数であることがわかる。このとき $x=0$ とすれば、$\int_0^1 \dfrac{dt}{1+t^2}=\tan^{-1}1=\dfrac{\pi}{4}$ さらに $x\to\infty$ とおけば、$\displaystyle\int_0^\infty \frac{e^{-x^2(1+t^2)}}{1+t^2}dt\to 0$ だから、$\displaystyle\left(\int_0^\infty e^{-t^2}dt\right)^2=\frac{\pi}{4}$ となり、証明される。

問 2.1

一般 $N(\mu,\sigma^2)$ の場合には変数変換を行う。これを示せ。

問 2.2

ガンマ関数において $\displaystyle\Gamma(1/2)=\int_0^\infty \sqrt{x}e^{-x}dx=\sqrt{\pi}$ の関係式は正規分布の密度関数を積分することと同じであることを確かめよ。

問 2.3

２項分布においても独立ならば、和を保存する性質（再生性）をもつ。これを示せ。$X\sim Binom(n,p), Y\sim Binom(m,p)$ ならば、$X+Y\sim Binom(n+m,p)$ が成り立つ。

問 2.4

幾何分布（負の 2 項分布）、指数分布（ガンマ分布）などについてもこのような再生性をもつ。分布の密度関数を調べて直接計算してみよ。