Introduction to mathematical statistics ゼミ (第 3 回)

担当: 伊藤真道

未定

3 Chap.3. Some Special Distribution

3.7 Mixture Distribution / 混合分布

本節では，一般の場合の分布の混合について述べる．k 個の分布が，それぞれ，平均 μ_i,，分散 σ_i^2 と台 (support)$\mathcal{S}_i$ である確率密度関数 $f_i(x)$ をを持つとする ($i = 1, \ldots, k$)．また，$p_1 + p_2 + \cdots + p_k = 1$ であるような，正数 p_i を仮定する．今，$\mathcal{S} = \cup_{i=1}^{k} \mathcal{S}_i$ とし，以下のような関数を考える．

$$f(x) = p_1 f_1(x) + p_2 f_2(x) + \cdots + p_k f_k(x) = \sum_{i=1}^{k} p_i f_i(x),\ x \in \mathcal{S} \tag{1}$$

この関数は，確率密度関数である．

証明 3.0. まず，$f_i(x), p_i$ の非負性から $\forall x \in \mathcal{S}$ について $f(x) \geq 0$．さらに，

$$\begin{aligned}
\int_{-\infty}^{\infty} f(x)dx &= \int_{-\infty}^{\infty} \sum_{i=1}^{k} p_i f_i(x)dx \\
&= \sum_{i=1}^{k} p_i \int_{-\infty}^{\infty} f_i(x)dx \cdots (*) \\
&= \sum_{i=1}^{k} p_i \int_{-\infty}^{\infty} f_i(x)dx \\
&= \sum_{i=1}^{k} p_i \\
&= 1
\end{aligned}$$

以上から，$f(x)$ は確率密度関数である．

なお，(∗) で積分と和を入れ替えているが，必ずしもこれが成り立つとは限らない．本節では，これが成り立つことを仮定する．以上から，$f(x)$ は連続型確率変数 X の密度関数であることがわかった．X の平均は

$$E[X] = \sum_{i=1}^{k} p_i \int_{-\infty}^{\infty} x f_i(x) dx = \sum_{i=1}^{k} p_i \mu_i = \bar{\mu}$$

のように，μ_i の重み付き平均で与えられ，分散は，

$$\begin{aligned}
Var(X) &= \sum_{i=1}^{k} \int_{-\infty}^{\infty} (x-\bar{\mu})^2 f_{(}x) dx \\
&= \sum_{i=1}^{k} p_i \int_{-\infty}^{\infty} ((x-\mu_i)+(\mu_i-\bar{\mu}))^2 f_i(x) dx \\
&= \sum_{i=1}^{k} p_i \int_{-\infty}^{\infty} [(x-\mu_i)^2 f_i(x) + (\mu_i-\bar{\mu})^2 f_i(x) - 2(x-\mu_i)(\mu_i-\bar{\mu}) f_i(x)] dx \\
&= \sum_{i=1}^{k} p_i (Var(X_i) + \mu_i - \bar{\mu})^2) \ (\because \sum_{i=1}^{k} p_i(\mu_i-\bar{\mu}) = 0) \\
&= \sum_{i=1}^{k} p_i \sigma_i^2 + \sum_{i=1}^{k} p_i (\mu - \bar{\mu}^2)
\end{aligned}$$

のように，σ_i^2 の重み付き平均と，μ_i の重み付き分散の和になる．これらの性質は，混合分布について成り立つものであって，確率変数の線型結合とは関係ないということに注意されたい．
¡note¿

分布の混合は，時には混ぜ合わせ?(compounding) と呼ばれる．さらに，混ぜ合わせる分布の数は有限でなくても良い．

例 3.7.1 では，対数ガンマ関数

$$f_1(x) = \begin{cases} \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{-(1+\beta)/\beta} (\log x)^{\alpha-1} & x > 1 \\ 0 & elsewhere \end{cases}$$

を用いる $(\alpha, \beta > 0)$．ここでは名前と密度関数の紹介だけする．

例 **3.1.** $X_\theta \sim Po(\theta)$ とする．それぞれのパラメータ θ が異なる無限個のポアソン分布の混ぜ合わせを考える．重み関数を θ の分布，ガンマ関数とする．$x = 0, 1, 2, \ldots$ に対する，混合された分布の確率質量関数 *(probability mass function, pmf)* は，

$$
\begin{aligned}
p(x) &= \int_0^\infty \left[\frac{1}{\beta^\alpha \Gamma(\alpha)} \theta^{\alpha-1} e^{-\theta/\beta}\right]\left[e^{-\theta}\frac{\theta^x}{x!}\right] d\theta \\
&= \frac{1}{\Gamma(\alpha)\beta^\alpha x!} \int_0^\infty \theta^{x+\alpha-1} e^{-\theta(1+\beta)/\beta} d\theta \\
&= \frac{1}{\Gamma(\alpha)\beta^\alpha x!} \int_0^\infty \big[\frac{\beta}{1+\beta} t\big]^{x+\alpha-1} e^{-t} \left|\frac{\beta}{1+\beta}\right| dt \\
&\quad (t = \theta(1+\beta)/\beta) \text{ とおいて変数変換} \\
&= \frac{1}{\Gamma(\alpha)\beta^\alpha x!} \left[\frac{\beta}{1+beta}\right]^{x+\alpha} \int_0^\infty t^{x+\alpha-1} e^{-t} dt \\
&= \frac{\beta^x}{\Gamma(\alpha)(1+\beta)^{x+\alpha} x!} \times \frac{\Gamma(x+\alpha)}{1^{x+\alpha}} \\
&= \frac{\Gamma(x+\alpha)\beta^x}{\Gamma(\alpha)(1+\beta)^{x+\alpha} x!}
\end{aligned}
$$

と変形できる．$\alpha = r, \beta = (1-p)/p, 0 < p < 1$ の時，この確率質量関数は，

$$
p(x) = \frac{(r+x-1)!}{(r-1)!} \frac{p^r (1-p)^x}{x!}, x = 0, 1, 2, \ldots
$$

となる．つまり．この分布の混ぜ合わせは，連続する独立な試行において，成功確率が p の事象の r 回の成功を得るために必要な試行の数の分布と等しい．これは負の二項分布と呼ばれる．

¡note¿

混合分布を構成する際，元となる X の確率密度関数を，何らかのパラメータ θ が与えられたもとでの条件付き密度関数 $f(x|\theta)$ とみなし，重み関数を θ の密度関数 $g(\theta)$ で表すとする．この時，X, θ の同時密度関数は，$h(x, \theta) = f(x|\theta)g(\theta)$ であり，混合分布は，X の周辺密度関数

$$
h(x) = \int_\theta g(\theta) f(x|\theta) d\theta
$$

で与えられる．

例 3.2. $X|\theta \sim N(0, 1/\theta), \theta \sim Gam(\alpha, \beta)$ とする．この時，X, θ の同時密度関数は，

$$f(x|\theta)g(\theta) = \Bigg[\sqrt{\frac{\theta}{2\pi}}\exp(-\frac{\theta x^2}{2})\Bigg]\Bigg[\frac{1}{\Gamma(\alpha)\beta^\alpha}\theta^{\alpha-1}\exp(-\theta/\beta)\Bigg],\ -\infty < x < \infty, 0 < \theta < \infty$$

である．X の周辺密度関数を求めるには θ を積分消去すれば良いので，

$$\begin{aligned}
h(x) &= \frac{1}{\sqrt{2\pi}\Gamma(\alpha)\beta^\alpha}\int_0^\infty \theta^{1/2+\alpha-1}\exp(-\theta(\frac{x^2}{2}+\frac{1}{\beta}))d\theta \\
&= \frac{1}{\sqrt{2\pi}\Gamma(\alpha)\beta^\alpha}\int_0^\infty \theta^{\alpha+1/2-1}\exp(-\theta(\frac{\beta x^2+2}{2\beta}))d\theta \\
&= \frac{1}{\sqrt{2\pi}\Gamma(\alpha)\beta^\alpha}\times\frac{\Gamma(\alpha+1/2)}{\big[\frac{\beta x^2+2}{2\beta}\big]^{\alpha+1/2}} \\
&= \frac{\Gamma(\alpha+1/2)}{\sqrt{2\pi}\Gamma(\alpha)\beta^\alpha}\big[\frac{2\beta}{\beta x^2+2}\big]^{\alpha+1/2}
\end{aligned}$$

$\alpha = r/2, \beta = 2/r$ の時，

$$\begin{aligned}
h(x) &= \frac{\Gamma((r+1)/2)}{\sqrt{2\pi}\Gamma(r/2)(2/r)^{(r/2)}}\big[\frac{4/r}{(2/r)x^2+2}\big]^{(r+1)/2} \\
&= \frac{\Gamma((r+1)/2)}{\sqrt{2\pi}\Gamma(r/2)(2/r)^{(r/2)}}\big[\frac{4/r}{2\frac{x^2+r}{r}}\big]^{(r+1)/2} \\
&= \frac{\Gamma((r+1)/2)}{\sqrt{2\pi}\Gamma(r/2)(2/r)^{(r/2)}}\big[\frac{1}{x^2+r}\big]^{(r+1)/2}2^{(r+1)/2} \\
&= \frac{\Gamma((r+1)/2)}{\sqrt{\pi}\Gamma(r/2)}r^{(r/2)}\big[\frac{1}{x^2+r}\big]^{(r+1)/2} \\
&= \frac{\Gamma((r+1)/2)}{\sqrt{r\pi}\Gamma(r/2)}\big[\frac{r}{x^2+r}\big]^{(r+1)/2} \\
&= \frac{\Gamma((r+1)/2)}{\sqrt{\pi}\Gamma(r/2)}r^{(r/2)}\big[1+\frac{x^2}{r}\big]^{-(r+1)/2}\ (\sim t_r)
\end{aligned}$$

のように，X 自由度 r の t 分布 *(Student's t-distribution)* を持つ．この例では，分布の混合を用いて，一般化された t 分布 *(generalized t-distribution)* の導出を行なった．

¡note¿

例 3.3. この例では，分布の裾の長い *(heavy tailed)*，歪んだ *(skew)* 分布の混合による導出について考える．$X|\theta \sim Gam(k, 1/\theta), \theta \sim Gam(\alpha, \beta)$ とする．X の周辺分布は，

$$\begin{aligned}
h(x) &= \int_0^\infty f(x|\theta)g(\theta)d\theta \\
&= \int_0^\infty \left[\frac{x^{k-1}e^{-\theta x}}{(1/\theta)^k\Gamma(k)}\right]\left[\frac{\theta^{\alpha-1}e^{-\theta/\beta}}{\beta^\alpha\Gamma(\alpha)}\right]d\theta \\
&= \frac{x^{k-1}}{\Gamma(k)\beta^\alpha\Gamma(\alpha)}\int_0^\infty \theta^{k+\alpha-1}e^{-\theta(\beta x+1)/\beta}d\theta \\
&= \frac{x^{k-1}}{\Gamma(k)\beta^\alpha\Gamma(\alpha)}\frac{\Gamma(\alpha+k)}{\left[\frac{\beta x+1}{\beta}\right]^{\alpha+k}} \\
&= \frac{\Gamma(\alpha+k)\beta^k x^{k-1}}{\Gamma(k)\Gamma(\alpha)(1+\beta x)^{\alpha+k}}
\end{aligned}$$

となる．これは，一般化パレート分布 (generalized Pareto distribution) および，F 分布の一般化の密度関数である．もし，$k=1$ ならば，

$$h(x) = \alpha\beta(1+\beta x)^{-(\alpha+1)}$$

となり，これはパレート分布の密度関数である．これらの混合分布のいずれも，元となったガンマ分布よりも，裾が分厚くなっている．

パレート分布の分布関数は，

$$\begin{aligned}
H(x) &= \int_0^x \alpha\beta(1+\beta t)^{-(\alpha+1)}dt \\
&= \left[-(1+\beta t)^{-\alpha}\right]_0^x \\
&= 1-(1+\beta x)^{-\alpha},\ 0 \le x < \infty
\end{aligned}$$

となる．これと $X = Y^\tau, \tau > 0$ を用いて，別の有用な裾の長い分布を構成できる．Y の分布関数は，

$$G(y) = P[Y \le y] = P[X^{1/\tau} \le y] = P[X \le y^\tau]$$

となる．よって，Y の分布関数は，パレート分布の分布関数を用いて

$$G(y) = H(y^\tau) = 1-(1+\beta y^\tau)^{-\alpha},\ 0 < y < \infty$$

と表せる．よって，Y の確率密度関数は，

$$G'(y) = g(y) = \frac{\alpha\beta\tau y^{\tau-1}}{(1+\beta x^{\tau})^{\alpha+1}},\ 0 < y < \infty$$

となる．

これらに関係する分布は，変形パレート分布 *(transformed Pareto distribution)*，もしくは，*Burr* 分布 *(Burr distribution)* と呼ばれており，裾の重い分布のモデリングに有用だということが知られている．

¡note¿