Scale Mixture Distributions

2014.03.21

1 Preliminaries

1.1 Normal(Gaussian) Distribution $\mathbf{N}(x \mid \mu, \sigma^2)$

The probability density function is given by:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp(-\frac{(x-\mu)^2}{2\sigma^2})$$

1.2 Generalized Inverse Gaussian Distributions

1.2.1 Gamma Distribution: $\mathbf{G}(x \mid \tau, \theta)$

The probability density function is given by:

$$f(x) = \frac{\theta^\tau}{\Gamma(\tau)} x^{\tau-1} \exp(-\theta x)$$

1.2.2 Inverse Gamma Distribution: $\mathbf{IG}(x \mid \tau, \theta)$

The probability density function is given by:

$$f(x) = \frac{\theta^\tau}{\Gamma(\tau)} x^{-(1+\tau)} \exp(-\theta x^{-1})$$

A special case is levy distribution. Let $x = x - \mu, \tau = \frac{1}{2}, \theta = \frac{c}{2}$, we will get the probability density function of levy distribution:

$$f(x) = \sqrt{\frac{c}{2\pi}} \frac{e^{-\frac{c}{2(x-\mu)}}}{(x-\mu)^{\frac{3}{2}}}$$

1.2.3 Generalized Inverse Gaussian Distribution: $\mathbf{GIG}(x \mid \gamma, \beta, \alpha)$

The probablility density function is given by:

$$f(x) = \frac{(\alpha/\beta)^{\gamma/2}}{2K_\gamma(\sqrt{\alpha\beta})} x^{\gamma-1} \exp(-(\alpha x + \beta x^{-1})/2)$$

1.3 Student's t-distribution

The probability density function is given by:

$$f(x) = \frac{\Gamma(\frac{v+1}{2})}{\Gamma(\frac{v}{2})} (1 + \frac{(x-\mu)^2}{v\sigma^2})^{-\frac{v+1}{2}} \frac{1}{\sqrt{v\pi}\sigma}$$

2 Gaussian GIG Mixture

Let $X \sim \mathrm{N}(x \mid \mu, \Sigma)$, $\Sigma \sim \mathrm{GIG}(x \mid \gamma, \beta, \alpha)$. The marginal density of x is defined by

$$\begin{aligned} f(x) =& \int_0^\infty N(x \mid \mu, \Sigma) GIG(\Sigma \mid \gamma, \beta, \alpha) \mathrm{d}\Sigma \\ =& \int_0^\infty \frac{1}{\sqrt{2\pi\Sigma}} \exp(-\frac{(x-\mu)^2}{2\Sigma}) \frac{(\alpha/\beta)^{\gamma/2}}{2K_\gamma(\sqrt{\alpha\beta})} \Sigma^{\gamma-1} \exp(-(\alpha\Sigma + \beta\Sigma^{-1})/2) \mathrm{d}\Sigma \\ =& C \int_0^\infty \Sigma^{\gamma-\frac{1}{2}-1} \exp(-\frac{\alpha}{2}\Sigma - (\beta + (x-\mu)^2)\frac{1}{2}\Sigma^{-1}) \mathrm{d}\Sigma, \end{aligned}$$

where $C = \frac{1}{\sqrt{2\pi}} \frac{(\alpha/\beta)^{\gamma/2}}{2K_\gamma(\sqrt{\alpha\beta})}$

Since

$$\int_0^\infty \frac{(\alpha/\beta)^{\gamma/2}}{2K_\gamma(\sqrt{\alpha\beta})} x^{\gamma-1} \exp(-(\alpha x + \beta x^{-1})/2) \mathrm{d}x = 1,$$

and let $\alpha = \alpha,\ \beta = \beta + (x-\mu)^2,\ \gamma = \gamma - \frac{1}{2}$,
so we will have

$$\int_0^\infty \Sigma^{\gamma-\frac{1}{2}-1} \exp(-\frac{\alpha}{2}\Sigma - (\beta + (x-\mu)^2)\frac{1}{2}\Sigma^{-1}) \mathrm{d}\Sigma = \frac{2K_{\gamma-\frac{1}{2}}(\sqrt{\alpha(\beta + (x-\mu)^2)})}{\alpha^{\frac{\gamma}{2}-\frac{1}{4}}} (\beta + (x-\mu)^2)^{\frac{\gamma}{2}-\frac{1}{4}}$$

.

Eventually,

$$\begin{aligned} f(x) =& \frac{1}{\sqrt{2\pi}} \frac{(\alpha/\beta)^{\gamma/2}}{2K_\gamma(\sqrt{\alpha\beta})} \frac{2K_{\gamma-\frac{1}{2}}\sqrt{\alpha(\beta + (x-\mu)^2)}}{\alpha^{\frac{\gamma}{2}-\frac{1}{4}}} (\beta + (x-\mu)^2)^{\frac{\gamma}{2}-\frac{1}{4}} \\ =& \frac{\alpha^{\frac{1}{4}} K_{\gamma-\frac{1}{2}} \sqrt{\alpha(\beta + (x-\mu)^2)} (\beta + (x-\mu)^2)^{\frac{\gamma}{2}-\frac{1}{4}}}{\sqrt{2\pi} K_\gamma \sqrt{\alpha\beta} \beta^{\frac{\gamma}{2}}} \end{aligned}$$

2.1 Student's t-distribution

2.1.1 Gaussian and Inverse Gamma Mixture

Let $X \sim \mathrm{N}(x \mid \mu, \Sigma)$, $\Sigma \sim \mathrm{IG}(x \mid \tau/2, \tau/2\lambda)$. The marginal density of x is defined by

$$\begin{aligned} f(x) =& \int_0^\infty N(x \mid \mu, \Sigma) IG(\Sigma \mid \tau/2, \tau/2\lambda) \mathrm{d}\Sigma \\ =& \int_0^\infty \frac{1}{\sqrt{2\pi\Sigma}} \exp\left(-\frac{(x-\mu)^2}{2\Sigma}\right) \frac{(\tau/2\lambda)^{\tau/2}}{\Gamma(\tau/2)} \Sigma^{-(1+\frac{\tau}{2})} \exp\left(-\frac{\tau}{2\lambda}\Sigma^{-1}\right) \mathrm{d}\Sigma \\ =& C \int_0^\infty \Sigma^{-(1+\frac{\tau+1}{2})} \exp\left(-\frac{\tau + \lambda(x-\mu)^2}{2\lambda}\Sigma^{-1}\right) \mathrm{d}\Sigma, \end{aligned}$$

where $C = \frac{1}{\sqrt{2\pi}} \frac{(\tau/2\lambda)^{\frac{\tau}{2}}}{\Gamma(\tau/2)}$

Since $\int_0^{+\infty} \frac{\theta^\tau}{\Gamma(\tau)} \Sigma^{-(1+\tau)} \exp(-\theta\Sigma^{-1}) \mathrm{d}\Sigma = 1$,

So we will have $\int_0^{+\infty} \Sigma^{-(1+\tau)} \exp(-\theta\Sigma^{-1}) \mathrm{d}\Sigma = \frac{\Gamma(\tau)}{\theta^\tau}$.
And let $\tau = \frac{\tau+1}{2}, \theta = \frac{\tau+\lambda(x-\mu)^2}{2\lambda}$.
Eventually,

$$\begin{aligned} f(x) &= \frac{1}{\sqrt{2\pi}} \frac{(\tau/2\lambda)^{\frac{\tau}{2}}}{\Gamma(\tau/2)} \frac{\Gamma(\frac{\tau+1}{2})(2\lambda)^{\frac{\tau+1}{2}}}{(\tau+\lambda(x-\mu)^2)^{\frac{\tau+1}{2}}} \\ &= \frac{1}{\sqrt{\pi}} \frac{\Gamma(\frac{\tau+1}{2})}{\Gamma(\frac{\tau}{2})} \left(\frac{\lambda}{\tau}\right)^{\frac{1}{2}} \left(1 + \frac{\lambda}{\tau}(x-\mu)^2\right)^{-\frac{\tau+1}{2}} \end{aligned}$$

2.1.2 Gaussian and Gamma Mixture

Let $X \sim \mathrm{N}(x \mid \mu, \frac{\sigma^2}{\gamma})$, $\gamma \sim \mathrm{G}(x \mid \frac{v}{2}, \frac{v}{2})$. The marginal density of x is defined by

$$\begin{aligned} f(x) &= \int_0^\infty N(x \mid \mu, \frac{\sigma^2}{\gamma}) G(\gamma \mid \frac{v}{2}, \frac{v}{2}) \mathrm{d}\gamma \\ &= \int_0^\infty \frac{\gamma^{\frac{1}{2}}}{\sqrt{2\pi}\sigma} \exp(-\frac{(x-\mu)^2}{2\sigma^2}\gamma) \frac{(\frac{v}{2})^{\frac{v}{2}}}{\Gamma(\frac{v}{2})} \gamma^{\frac{v}{2}-1} \exp(-\frac{v}{2}\gamma) \mathrm{d}\gamma \\ &= C \int_0^\infty \gamma^{\frac{v+1}{2}-1} \exp(-(\frac{v}{2} + \frac{(x-\mu)^2}{2\sigma^2})\gamma) \mathrm{d}\gamma, \end{aligned}$$

$$\textit{where } C = \frac{1}{\sqrt{2\pi}\sigma} \frac{(\frac{v}{2}^{\frac{v}{2}})}{\Gamma(\frac{v}{2})}$$

Since

$$\int_0^\infty \frac{\theta^\tau}{\Gamma(\tau)} x^{\tau-1} \exp(-\theta x) \mathrm{d}x = 1$$

so we will have

$$\int_0^\infty x^{\tau-1} \exp(-\theta x) \mathrm{d}x = \frac{\Gamma(\tau)}{\theta^\tau}$$

And let $\tau = \frac{v+1}{2},\ \theta = \frac{v}{2} + \frac{(x-\mu)^2}{2\sigma^2}$
Eventually,

$$f(x) = \frac{1}{\sqrt{v\pi}\sigma} \frac{\Gamma(\frac{v+1}{2})}{\Gamma(\frac{v}{2})} (1 + \frac{(x-\mu)^2}{v\sigma^2})^{\frac{v+1}{2}}$$

2.2 Gaussian Gamma Mixture

Let $X \sim \mathrm{N}(x \mid \mu, \Sigma)$, $\Sigma \sim \mathrm{G}(x \mid \gamma, \alpha/2)$. The marginal density of x is defined by

$$\begin{aligned} f(x) &= \int_0^\infty N(x \mid \mu, \Sigma) \mathrm{G}(\Sigma \mid \gamma, \alpha/2) \mathrm{d}\Sigma \\ &= \int_0^\infty \frac{1}{\sqrt{2\pi\Sigma}} \exp(-\frac{(x-\mu)^2}{2\Sigma}) \frac{(\frac{\alpha}{2})^\gamma}{\Gamma(\gamma)} \Sigma^{\gamma-1} \exp(-\frac{\alpha}{2}\Sigma) \mathrm{d}\Sigma \\ &= C \int_0^\infty \Sigma^{\gamma-\frac{1}{2}-1} \exp(-\frac{\alpha\Sigma}{2} - \frac{(x-\mu)^2}{2\Sigma}) \mathrm{d}\Sigma, \end{aligned}$$

$$\textit{where } C = \frac{1}{\sqrt{2\pi}} \frac{(\frac{\alpha}{2})^\gamma}{\Gamma(\gamma)}$$

Since

$$\int_0^\infty \frac{(\alpha/\beta)^{\gamma/2}}{2K_\gamma(\sqrt{\alpha\beta})} x^{\gamma-1} \exp(-(\alpha x + \beta x^{-1})/2)dx = 1,$$

so we will have

$$\int_0^\infty \Sigma^{\gamma-1} \exp(-(\alpha\Sigma + \beta\Sigma^{-1})/2)\mathrm{d}\Sigma = \frac{2K_\gamma(\sqrt{\alpha\beta})}{\alpha^{\frac{\gamma}{2}}}\beta^{\frac{\gamma}{2}}$$

Let $\alpha = \alpha,\ \beta = (x-\mu)^2,\ \gamma = \gamma - \frac{1}{2}$,
Eventually,

$$\begin{aligned} f(x) =& \frac{1}{\sqrt{2\pi}} \frac{(\frac{\alpha}{2})^\gamma}{\Gamma(\gamma)} \frac{(x-\mu)^{\gamma-\frac{1}{2}} 2K_{\gamma-\frac{1}{2}} \sqrt{\alpha(x-\mu)^2}}{\alpha^{\frac{\gamma}{2}-\frac{1}{4}}} \\ =& \frac{\alpha^{\frac{2\gamma+1}{4}}}{2^{\gamma-\frac{1}{2}}} \frac{1}{\Gamma(\gamma)} (x-\mu)^{\gamma-\frac{1}{2}} K_{\gamma-\frac{1}{2}} \sqrt{\alpha(x-\mu)^2} \end{aligned}$$

2.3 Scale Mixtures of Normal Distribution

X = Z/V, Z has a standard normal distribution and V is positive and is independent of Z.

Let the CDF V be G_v, what will the probability density of X be?

From definition, we know:

$$\begin{aligned} F_x(x) =& P\{X \leq x\} \\ =& P\{\frac{Z}{V} \leq x\} \\ =& P\{Z \leq Vx\} \\ =& P\{Z \leq vx\}P\{V = v\} \\ =& \int_0^{+\infty} \int_{-\infty}^{vx} \frac{1}{\sqrt{2\pi}} \exp\{-\frac{z^2}{2}\}\mathrm{d}z\mathrm{d}G_v(v) \\ f_x(x) =& \int_0^\infty \frac{v}{\sqrt{2\pi}} \exp -\frac{v^2x^2}{2} dG_v(v) \end{aligned}$$

3 Theorem

Theorem 3.1. *(Bernstein) Let $g : (0,\infty) \to \mathbb{R}$ be a completely monotone function. Then it is the Laplace transform of unique measure μ on $[0,\infty]$, i.e. for all $\lambda > 0$,*

$$g(\lambda) = \mathcal{L}(\mu;\lambda) = \int_{[0,\infty)} e^{-\lambda t}\mu(dt)$$

. Conversely, whenever $\mathcal{L}(\mu;\lambda) < \infty$ for every $\lambda > 0$, $\lambda \mapsto \mathcal{L}(\mu;\lambda)$ is a completely monotone function.

Theorem 3.2. *A function $f(x)$ can be represented as a Gaussian scale mixture iff $f(\sqrt{x})$ is completely monotone on $(0,\infty)$.*

Proof.

Let $g(x) = f(\sqrt{x})$.

$$
\begin{aligned}
& f(\sqrt{x}) \textit{ is completely monotone,} \\
\Longleftrightarrow & g(x) \textit{ is completely monotone.}
\end{aligned}
$$

By Bernstein :

$$
\begin{aligned}
\Longleftrightarrow & g(x) = \int_0^\infty e^{-xt}\mu(\mathrm{dt}) \\
\Longleftrightarrow & f(\sqrt{x}) = \int_0^\infty e^{-xt}\mu(\mathrm{dt}) \\
\Longleftrightarrow & f(x) = \int_0^\infty e^{-x^2 t}\mu(\mathrm{dt}) = C\int_0^\infty N(x \mid 0, \frac{1}{2t})\mu(\mathrm{dt}), \quad and \int_0^\infty \mu(dt) = 1 \\
\Longleftrightarrow & f(x) \textit{ can be represented as a Gaussian scale mixture.}
\end{aligned}
$$

□

Theorem 3.3. *If $f(x) > 0$, then $e^{-uf(x)}$ is completely monotone for every $u > 0$ iff $f'(x)$ is completely monotone.*

Proof. If $e^{-uf(x)}$ is completely monotone for every $u > 0$:

$$
e^{-\mu f(x)} = \sum_{j=0}^{\infty} \frac{(-1)^j \mu^j}{j!}[f(x)]^j
$$

and all of its formal derivatives converge uniformly, so we can calculate $\frac{d^n}{dx^n}e^{-\mu f(x)}$ by termwise differentiation. Since $e^{-\mu f}$ is completely monotone, we have:

$$
0 \leq (-1)^n \frac{d^n}{dx^n} e^{-\mu f(x)} = \sum_{j=1}^{\infty} \frac{\mu^j}{j!}(-1)^{n+j}\frac{d^n}{dx^n}[f(x)]^j
$$

As $\mu > 0$, dividing μ, there is:

$$
0 \leq (-1)^{n+1}\frac{d^n}{dx^n} f(x) + \sum_{j=2}^{\infty} \frac{\mu^{j-1}}{j!}(-1)^{n+j}\frac{d^n}{dx^n}[f(x)]^j
$$

Then let $\mu \to 0$:

$$
0 \leq (-1)^{n-1}\frac{d^{n-1}}{dx^{n-1}} f'(x)
$$

Eventually, $f'(x)$ is completely monotone.

If $f'(x)$ is completely monotone:

$$
(-1)^{n-1}\frac{d^n}{dx^n} f(x) \geq 0
$$

Let $g(\lambda) = e^{-\lambda}, \ \lambda = f(x)$:

$$
h(x) = e^{-f(x)} = g(\lambda) \circ f(x)
$$

And there is a formula for the n-th derivative of the composition $h = g \circ f$:

$$h^{(n)}(\lambda) = \sum_{(m,i_1,\dots,i_l)} \frac{n!}{i_1!\dots i_l!} g^{(m)}(f(\lambda)) \prod_{j=1}^{l} (\frac{f^{(j)}(\lambda)}{j!})^{i_j},$$

where $\sum_{j=1}^{l} j \cdot i_j = n$ and $\sum_{j=1}^{l} i_j = m$.

We can see that $n = m + \sum_{j=1}^{l} (j-1) \cdot i_j$.

We have $(-1)^m g^{(m)}(f(x)) \geq 0$ and $(-1)^{j-1} f^{(j)} \lambda \geq 0$.

So $(-1)^n h^{(n)}(x) \geq 0$ which means $e^{-f(x)}$ is completely monotone.

And $e^{-\mu f(x)}$ is completely monotone.

$\square$