1. Power law $f(x) = a x^{-b}$

① $f(cx) = a (cx)^{-b}$

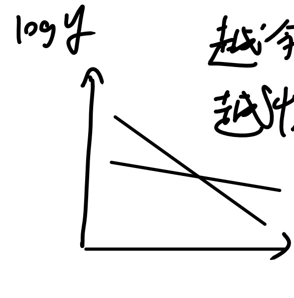
$= c^{-b} a x^{-b}$

$= c^{-b} f(x)$

$$\frac{f(x)}{f(cx)} = c^{b}$$

② $\log f(x) = \log a - b \log x$

$\log f(x) \propto \log x$



越斜，越快衰

③ b determines how fast to decay, how thin the tail is

E.g. $b = 2$ $\qquad b = 3$ $\qquad b = 4$

$\frac{f(x)}{f(2x)} = 2^2 = 4$ $\qquad 2^3 = 8$ $\qquad 2^4 = 16$

$\frac{f(x)}{f(3x)} = 3^2 = 9$ $\qquad 3^3 = 27$ $\qquad 3^4 = 81$

2. decay : exponential decay ae^{-bx}

power decay ax^{-b}

exponential squared decay ae^{-bx^2}

13.]	e^{-x}	e^{-x^2}	x^{-2}	x^{-3}	x^{-4}
$\frac{f(x)}{f(2x)}$	e^{x}	e^{3x^2}	4	8	16
$\frac{f(x)}{f(3x)}$	e^{2x}	e^{8x^2}	9	27	81

when $x=1$,

$e = 2.7$ $\quad e^3 = 20$

when $x=4$

$e^2 = 7.38$ $\quad e^{12} = 162754$

$e^2 = 7.38$ $\quad e^8 = 2980$

$e^4 = 54.6$ $\quad e^{32} = 7896296018$

3. Where to observe power law?

① upper tail of lognormal

详见5.

② Student-t

$$f(x) = \text{const} \cdot \left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}$$

$$\log f(x) = \log(\text{const}) - \frac{\nu+1}{2}\log\left(1+\frac{x^2}{\nu}\right)$$

当 x 足够大，"1" 可以忽略.

$$\log\left(1+\frac{x^2}{\nu}\right) \approx \log\left(\frac{x^2}{\nu}\right)$$

$$= 2\log x - \text{const}$$

$$\boxed{\log f(x) \propto \log x}$$

log-log plot

4. exponential squared decay ae^{-bx^2}

例：正态分布

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

$$\log f(x) = \text{const} - \frac{(x-\mu)^2}{2\sigma^2}$$

$$\boxed{\log f(x) \propto x^2}$$

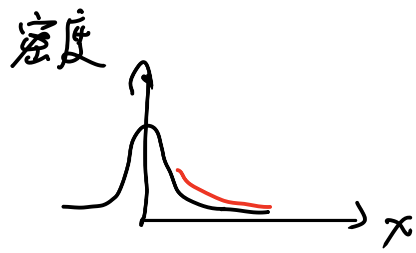
衰减得太快

导致出现极端事件概率太小

不符合现实

增加极端事件发生的概率

↳ 肥尾



* $E[e^{tX}]$

$$= \int e^{tx} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

$$= \frac{1}{\sqrt{2\pi}\sigma} \int e^{-\frac{(x-\mu)^2}{2\sigma^2}+tx} dx$$

$$I = -\frac{(x-\mu)^2}{2\sigma^2} + tx$$

$$= -\frac{(x-\mu)^2 - 2t\sigma^2 x}{2\sigma^2}$$

$$= -\frac{x^2 + \mu^2 - 2\mu x - 2t\sigma^2 x}{2\sigma^2}$$

$$= -\frac{x^2 - 2(\mu+t\sigma^2)x + (\mu+t\sigma^2)^2 - (\mu+t\sigma^2)^2 + \mu^2}{2\sigma^2}$$

$$= -\frac{(x-\mu-t\sigma^2)^2 - t^2\sigma^4 - 2\mu t\sigma^2}{2\sigma^2}$$

$$= -\frac{(x-(\mu+t\sigma^2))^2}{2\sigma^2} + \frac{t^2\sigma^2}{2} + \mu t$$

$$E[e^{tx}]$$

$$= \frac{1}{\sqrt{2\pi}\sigma}\int e^{I}\,dx$$

$$= e^{\mu t+\frac{\sigma^2 t^2}{2}} \underbrace{\frac{1}{\sqrt{2\pi}\sigma}\int e^{-\frac{(x-\mu^*)^2}{2\sigma^2}}\,dx}_{\text{所有概率之和为1}}$$

$$= e^{\mu t+\frac{\sigma^2 t^2}{2}}$$

5. lognormal distribution

$$Y = e^X \quad , \quad X \sim N(\mu, \sigma^2)$$

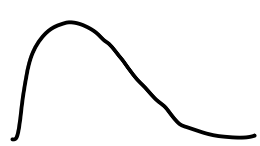
$$\ln Y = X \quad \sim N(\mu, \sigma^2)$$

X Normal
Y lognormal.

$$E[Y] = E[e^X]$$
$$= e^{\mu t + \frac{\sigma^2 t^2}{2}} \quad , \; t = 1$$
$$= e^{\mu + \frac{\sigma^2}{2}}$$

$$E[Y^2] = E[e^{2x}]$$
$$= e^{2\mu + 2\sigma^2}$$

$$Var(Y) = E[Y^2] - (EY)^2$$
$$= e^{2\mu + 2\sigma^2} - e^{2\mu + \sigma^2}$$
$$= e^{2\mu + \sigma^2}(e^{\sigma^2} - 1)$$



pdf ? $\text{pdf} = \frac{d}{dx}\,\text{cdf}$

$$f(x) = \frac{d}{dx} P(Y \le x)$$

$$= \frac{d}{dx} P(\ln Y \le \ln x)$$

$$= \frac{d}{dx} P(X \le \ln x), \quad X \sim N(\mu, \sigma^2)$$

$$= \frac{d}{dx} N(\ln x)$$

$$= N'(\ln x) \frac{d}{dx}(\ln x)$$

$$= \underbrace{N'(\ln x)}_{\text{Normal pdf}} \cdot \frac{1}{x}$$

$$= \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(\ln x-\mu)^2}{2\sigma^2}} \cdot \frac{1}{x}$$

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

x 足够大时 power decay

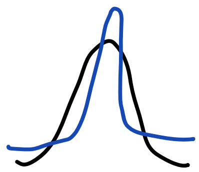
$\to$ log - log

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

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

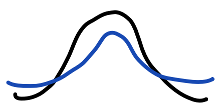
(?)

6. 目标：收益率不满足正态分布



kurtosis > 3.
(四阶矩)

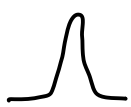
Student-t 不太行 只有肥尾



作为进阶，考虑 power law 相关

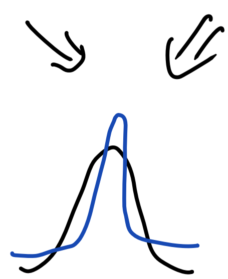
另外一个思路：

$50\% \cdot N(0, 0.5) + 50\% \cdot N(0, 5)$



peak

fat tail



Expectation
Maximization

$\uparrow$

Mixture Model $\rightarrow$ Gausian Mixture Model

$P(A)P(X_1 | A)$
$+P(B)P(X_2 | B)$

作为比较，
有一个完全不同的概念

multivariate normal

$Y = X_1 + X_2$

$X_1 \sim N(1, 2^2)$
$X_2 \sim N(2, 3^2)$

Y 是否也符合正态分布？

如果 Y 是正态分布，

$E[Y] = E[X_1] + E[X_2] = 3$

$$\begin{aligned}
var(Y) &= cov(Y, Y) \\
&= cov(X_1 + X_2, X_1 + X_2) \\
&= var(X_1) + var(X_2) \\
&\quad + 2cov(X_1, X_2) \\
&= 4 + 9 + 2 \cdot 2 \cdot 3 \cdot \rho
\end{aligned}$$