Financial Time Series

Topic 9: Modelling Return Distributions

Hung Chen
Department of Mathematics
National Taiwan University
6/5/2002
OUTLINE

1. Empirical finding on return distributions
2. Models for Return Distributions
3. Determining the Tail Shape of a Return Distribution
4. Estimation of $\mu$ and $\sigma$ in Ito’s process
5. Revisit the Pricing of European Option
Empirical finding on return distributions

- Characteristics:
  - Fat tails
  - High peakedness (excess kurtosis)
  - Skew

- Three return series: Figure 5.1
  - daily returns of the London FT30 from 1935 to 1994
  - daily returns of the S&P500
  - daily dollar/sterling exchange rate

- Descriptive statistics:
  Splus command summary: It gives min, 1st Qu, median, mean, 3rd Qu, max.
  mean: 0.022, 0.020, −0.008
  sd: 1.004, 1.154, 0.647
  median: 0.000, 0.047, 0.000
  range: 23.1, 33.1, 13.2
  kurt: 14.53, 25.04, 6.51
• Graphical representations:
  – smoothed function of the histogram
  – Q-Q plots: Plot empirical cumulative distribution against normal distribution.

  Histogram

• pdf:

  \[ P(-h/2 \leq X < h/2) = \int_{-h/2}^{h/2} f(x) dx \]

• law of large numbers:

  \[ P(-h/2 \leq X < h/2) \approx n^{-1} \# \{ X_i \in [-h/2, h/2) \} \]

• approximation:

  \[ P(-h/2 \leq X < h/2) \approx f(\xi)h \]

• density estimate:

  \[ \hat{f}_h(x) = (nh)^{-1} \sum_{i=1}^{n} \sum_{j} I(X_i \in B_j)I(x \in B_j) \]

Consider the kernel density estimate

\[ \hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^{n} K \left( \frac{x - X_i}{h} \right), \]

where \( K \) is a kernel function with the following properties
• Kernel function is symmetric around 0 and integrate to 1.

• Kernel is a density function and the kernel estimate is a density too.

• Kernel estimates do not depend on any choice of origin.

Quantile-Quantile plots

• If $X$ is a continuous random variable with a strictly increasing distribution function, $F$, the $p$th quantile of the distribution is the value of $x$ such that $F(x) = p$ or $x_p = F^{-1}(p)$.

• In a Q-Q plot, the quantiles of one distribution are plotted against those of another.

• Suppose $G(y) = F(y - h)$. Then

$$y_p = x_p + h$$

and a Q-Q plot would be a straight line with slope 1 and intercept $h$.

• Let $r_1, \ldots, r_n$ be the returns of a portfolio in the sample period. The order statistics of the sample are these
values arranged in increasing order. 
$r_{(1)}$ is the sample minimum and $r_{(n)}$ the sample maximum.

- For $\ell = np$,
  \[
  \sqrt{n}(r_{\ell} - x_p) \sim N(0, p(1 - p)/f^2(x_p)),
  \]
  if $f(x_p) \neq 0$. 


Models for Return Distributions

- fat-tailed and highly peaked
- Stable distributions
  It is a generalization of normal in that they are stable under addition.
  But non-normal stable distributions do not have a finite variance.
- Mixture distributions
  Hierarchical model approach: regression

Facts:

- \( EX = E(E(X|Y)), X|Y \sim Bin(Y, p), Y \sim Poisson(\lambda) \)
- \( VarX = E(Var(X|Y)) + Var(E(X|Y)) \)
- Mixed discrete-continuous distribution
  Slice a data set into different, supposedly more homogeneous, subsets.
- Scale-Mixture Normals
  As an example, consider

\[
    r_t \sim (1 - \alpha)N(\mu, \sigma_1^2) + \alpha N(\mu, \sigma_2^2),
\]

where \( 0 \leq \alpha \leq 1 \), \( \sigma_1^2 \) is small, and \( \sigma_2^2 \) is large.
The large value of $\sigma^2$ enables the mixture to put more mass at the tails of its distribution.
Note that
\[
E(r_t) = EE(r_t|I) = E\mu = \mu
\]
\[
V(r_t) = EV(r_t|I) + V(E(r_t|I))
\]
\[
= E\sigma_1^2 + V(\mu)
\]
\[
= (1 - \alpha)\sigma_1^2 + \alpha\sigma_2^2
\]
- Normal and Stable
- Student or double Weibull distributions

- Stable distributions:
  - Characteristic function:
    \[
    \varphi(t) = \int_{-\infty}^{\infty} \exp(itx)dF(x)
    \]
  - The symmetric (about zero) stable characteristic function:
    \[
    \varphi(t) = \exp(-\sigma^\alpha |t|^\alpha)
    \]
    where $0 < \alpha \leq 2$ is the characteristic exponent and $\sigma$ is a scale parameter.
  - Probability distribution
    \[
    F(X) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-\sigma^\alpha |t|^\alpha) \exp(-itX)dt
    \]
\[ = \frac{1}{\pi} \int_0^\infty \exp(-\sigma^\alpha |t|^\alpha) \cos(tX) dt. \] (1)

− Normal distribution: \( \alpha = 2 \)
− For \( \alpha < 2 \), all moments greater than \( \alpha \) are infinite.
− Fat tail: relative to the normal
− regular variation at infinity

\[ \lim_{s \to \infty} \frac{1 - F(sX)}{1 - F(s)} = X^{-\alpha} \]

The stable distribution displays a power declining tail, \( X^{-\alpha} \), rather than an exponential decline as is the case with the normal.

− \( \alpha \): the tail index
− Why the stable distribution should be an appropriate generating process for financial data?

Mandelbrot (1963): The limiting distribution of an appropriately scaled sum of independent and identically distributed random variables exists then it must be a member of the stable class, even if these random variables have infinite variance.
— If daily returns follow a stable distribution, then weakly, monthly and quarterly returns can be viewed as the sum of daily returns, they too will follow stable distributions having identical characteristic exponents.

• Volatility clustering

— GARCH class of models: serial correlation of conditional variances

— Consider $ARCH(1)$ with normal innovations’ process for $X_t$.

\[ X_t = U_t \sigma_t \]  \hspace{1cm} (2)

where $U_t \sim NID(0, 1)$ and

\[ \sigma_t^2 = w + \beta X_{t-1}^2. \] \hspace{1cm} (3)

From the above two equations, we have

\[ X_t^2 = wU_t^2 + \beta U_t^2 X_{t-1}^2 = B_t + A_t X_{t-1}^2. \] \hspace{1cm} (4)

— $X_t$ is serially uncorrelated but is not independent.

— $ARCH(1)$ process may also exhibit fat tails.
de Haan et al. (1989) show that the $X_t$ of (4) regularly varies at infinity and has a tail index $\zeta$ defined explicitly by the equation

$$\Gamma\left(\frac{\zeta + 1}{2}\right) = \pi^{1/2}(2\beta)^{-\zeta/2}.$$
Determining the tail shape of a return distribution
Estimation of $\mu$ and $\sigma$ in Ito’s process

- Treating price of an asset as a random variable that evolves over time, the prices series forms a stochastic process.

- The observed price series is a realization of the underlying stochastic process.

- Consider two types of stochastic processes.
  
  - Discrete-time stochastic process: Consider the daily closing price of IBM stock on the NYSE. Here the price change occurs only at the closing of a trading day.

  - Continuous-time stochastic processes: Assume the price changes continuously even when the stock is not traded.

  - A continuous-time continuous stochastic process can be written as $\{x(\eta, t)\}$, where $t$ denotes time and is continuous in $[0, \infty)$. For a given $t$, $x(\eta, t)$ is a continuous random variable defined on a probability space and $\eta$ is an element of the space.

  - For a given $\eta$, $\{x(\eta, t)\}$ is a time series with value depending on time.
The counterpart of white noise process to a discrete-time econometric model in a continuous-time model is the Wiener Process, which is also known as Brownian motion.

A continuous-time stochastic process \( \{w_t\} \) is a Wiener Process if it satisfies

- \( \Delta w_t = w_{t+\Delta t} - w_t = \epsilon \sqrt{\Delta t} \), where \( \epsilon \sim N(0, 1) \).
  - \( \Delta w_t \sim N(0, \Delta t) \).
  - Define \( N = t/\Delta t \). Then
    \[
    w_t - w_0 = \sum_{i=1}^{N} \epsilon_i \sqrt{\Delta t} \sim N(0, t).
    \]

- \( \Delta w_t \) is independent of \( w_j \) for all \( j < t \).
  - This is a Markov property.
  - \( w_{t_1+\Delta_1 t} - w_{t_1} \) and \( w_{t_2+\Delta_2 t} - w_{t_2} \) are independent for any two non-overlapping time intervals \( \Delta_1 \) and \( \Delta_2 \).

This suggests that we can simulate Wiener process on the unit time interval \([0, 1]\) using the following statistical property.

**Property:**
• Assume that \( \{z_t\}_{t=1}^n \) is a sequence of independent standard normal random variables.

• For any \( t \in [0, 1] \), let \( \lfloor nt \rfloor \) be the integer part of \( nt \).

• Define \( w_{n,t} = n^{-1/2} \sum_{i=1}^{\lfloor nt \rfloor} z_i \).

• \( w_{n,t} \) converges in distribution to the Wiener process \( w_t \) as \( n \) goes to infinity.

Let \( P_t \) be the price of a security at time \( t \), which is continuous in \( [0, \infty) \). In the literature, it is common to assume that \( P_t \) follows the following Ito’s process

\[
dP_t = \mu(P_t, t)dt + \sigma(P_t, t)dW_t. \tag{5}
\]

Here

• \( \mu \) and \( \sigma \) are referred to as the drift and volatility parameters of the process \( P_t \).

• \( W_t \) is a Wiener process.

• Use the notation \( dy \) for a small change in the variable \( y \).

• When \( \mu(P_t, t) = \mu P_t \) and \( \sigma(P_t, t) = \sigma P_t \) where \( \mu \) and \( \sigma \) are constants, apply the Ito’s
lemma to obtain
\[ d \ln(P_t) = \left( \mu - \frac{\sigma^2}{2} \right) dt + \sigma dW_t. \]

Hence, the change in logarithm of price (log return) between current time \( t \) and some future time \( T \) is normally distributed with mean \((\mu - \sigma^2/2)(T - t)\) and variance \(\sigma^2(T - t)\).

- To simulate \( P_t \), we can use the following recursive form
  \[ \ln P_{t+\Delta} = \ln P_t + (\mu - \sigma^2/2)\Delta + \sigma \Delta^{1/2} Z, \]
  where \( Z \sim N(0, 1) \).

  How do we estimate these two unknown parameters \( \mu \) and \( \sigma \)?

- Assume that we have \( n + 1 \) observations of stock price \( P_i \) before time \( t \) at equally spaced time interval \( \Delta \).

- Denote the observed prices as \( \{P_0, P_1, \ldots, P_n\} \) and let \( r_i = \ln(P_i) - \ln(P_{i-1}) \) for \( i = 1, \ldots, n \).

- \( r_i \) is normally distributed with mean \((\mu - \sigma^2/2)\Delta \) and variance \(\sigma^2\Delta \).
  \( r_i \)'s are not serially correlated.
• Let \( \bar{r} \) and \( s_r \) be the sample mean and standard deviation of the data.

• Estimate \( \sigma \) by \( s_r / \sqrt{\Delta} \) and \( \mu \) by
  \[
  \frac{\bar{r}}{\Delta} + \frac{\hat{\sigma}^2}{2n}.
  \]

• Consider the daily log returns of IBM in 1998.
  
  – It has 252 observations.
  
  – The sample ACF of the data indicates that the log returns are indeed serially uncorrelated. The Ljung-Box statistic gives \( Q(10) = 4.9 \), which is highly insignificant compared with a chi-square distribution with 10 degrees of freedom.
  
  – Assume that the price of IBM in 1998 follows the Ito’s process.
  
  – \( \bar{r} = 0.002276, \ s_r = 0.01915, \) and \( \Delta = 1/252 \) year.
  
  – \( \hat{\sigma} = s_r / \sqrt{\Delta} = 0.3040. \)
  
  – \( \hat{\mu} = 0.6198 \)
  
  – The estimated expected return is 61.98% and the standard deviation is 30.4% per annum for IBM in 1998.
Distributions of stock prices and log returns

- Conditional on the price $P_t$ at time $t$, the log price at time $T > t$ is

$$\ln(P_T) = \ln(P_t) + X_{t,T}$$

$$X_{t,T} \sim N((\mu - \sigma^2/2)(T - t), \sigma^2(T - t)).$$

This gives the information on the future price of $P_T$.

- Lognormal distribution with parameters $\mu$ and $\sigma^2$:

$$E(Y) = \exp(\mu + \sigma^2/2)$$

$$V(Y) = \exp(2\mu + \sigma^2)[\exp(\sigma^2) - 1].$$

- The conditional mean and variance of $P_T$ given $P_t$ are

$$E(P_T) = P_t \exp[\mu(T - t)],$$

$$V(P_T) = P_t^2 \exp[2\mu(T - t)]\{\exp[\sigma^2(T - t)] - 1\}.$$

- Suppose the current price of stock A is $50, the expected return is 15% per annum, and the volatility is 40% per annum. Then the expected price and variance of stock A in 6-month are

$$E(P_T) = 50 \exp(0.15 \times 0.5) = 53.89$$
and
\[ V(P_T) = 2500 \exp(0.3 \times 0.5)[\exp(0.16 \times 0.5) - 1] \]
\[ = 241.92. \]

The standard deviation is around 15.55.

- Let \( r \) be the continuously compounded rate of return per annum from time \( t \) to \( T \).

Note that
\[ P_T = P_t \exp[r(T - t)] \]
and
\[ r = \frac{1}{T - t} \ln \frac{P_T}{P_t}. \]

- The distribution of \( r \) is
\[ N(\mu - \sigma^2 / 2, \sigma^2 / (T - t)). \]

- Consider a stock with an expected rate of return of 15% per annum and a volatility of 10% per annum. The distribution of \( r \) is \( N(.15 - .01/2, (.1/\sqrt{2})^2) \).

A 95% CI for \( r \) is \((0.145 \pm 1.96 \times 0.071) = (0.6\%, 28.4\%). \)

What is the effect on using estimated \( \mu \) and \( \sigma^2 \)?

Refer to next topic for the European call option.
Revisit the Pricing of European Option

Black-Scholes Pricing Formulas

Risk-Neutral World

• The expected return on all securities is the risk-free interest rate $r$.

• The present value of any cash flow can be obtained by discounting its expected value at the risk-free rate.

• Under the no arbitrage assumption, the portfolio $V_t$ must be riskless during the small time interval. Here $V_t$ is the value of the portfolio. The portfolio must instantaneously earn the same rate of return as other short-term risk-free securities.

• The expected value of a European call option at maturity in a risk-neutral world is

$$E_*[\max(P_T - K, 0)]$$

where $E_*$ denotes expected value in a risk-neutral world.
• The price of the call option at time $t$ is

$$c_t = \exp[-r(T - t)]E_x[\max(P_T - K, 0)].$$  

(7)

We need to specify the distribution of $P_T$.

• In a risk-neutral world, we have $\mu = r$ and by (6),

$$\ln(P_T) \sim \ln(P_t)$$

$$+ N((r - \sigma^2/2)(T - t), \sigma^2(T - t)).$$

• Let $g(P_T)$ be the probability density function of $P_T$. Then the price of the call option in (7) is

$$c_t = \exp[-r(T - t)] \int_{K}^{\infty} (P_T - K)g(P_T)dP_T.$$  

The above formular holds for general price process.

Under Black-Scholes model, the distribution is log-normal with mean $\ln(P_t) + (r - \sigma^2/2)(T - t)$ and variance $\sigma^2(T - t)$.

• By changing variable in the integration and some algebraic calculations, we have

$$c_t = P_t\Phi(h_1) - K \exp[-r(T-t)]\Phi(h_2)$$  

(8)
where $\Phi(x)$ is the cumulative distribution function of the standard normal random variable evaluated at $x$,

$$
\begin{align*}
    h_1 &= \frac{\ln(P_t/K) + (r + \sigma^2/2)(T - t)}{\sigma \sqrt{T - t}} \\
    h_2 &= \frac{\ln(P_t/K) + (r - \sigma^2/2)(T - t)}{\sigma \sqrt{T - t}} \\
    &= h_1 - \sigma \sqrt{T - t}.
\end{align*}
$$

- Similarly, the price of a European put option is

$$
    p_t = K \exp[-r(T - t)]\Phi(h_2) - P_t\Phi(h_1).
$$

(9)

- How the price $c_t$ depends on the estimated $\sigma^2$?
  
  differentiate $c_t$ with respect to $\sigma$

- Can we solve $c_t$ without using analytic technique?
Example 2

• Suppose the current price of Intel stock is $80 per share with volatility $\sigma = 20\%$ per annum.

• The risk-free interest is 8\% per annum.

• Find the price of a European call option on Intel with as striking price of $90$ that will expire in 3 months?

  $- P_t = 80, \ K = 90, \ T - t = 0.25, \ \sigma = 0.2$ and $r = 0.08$.

  $- We \ have$

  $h_1 = \frac{\ln(80/90) + (0.08 + 0.2^2/2)0.25}{0.2\sqrt{0.25}}$

  $= -0.9278$

  $h_2 = h_1 - 0.2\sqrt{0.25} = -1.0278.$

  $- \ \Phi(-0.9278) = 0.1768, \ \Phi(-1.0278) = 0.1520.$

  $- The \ price \ of \ a \ European \ call \ option \ is$

  $c_t = 80 \times \Phi(-0.9278)$

  $-90 \times \Phi(-1.0278) \exp(-0.02) = 0.46.$

  $- The \ stock \ price \ has \ to \ rise \ by \$1.05 \ for$ the purchaser of the put option to break even
\* $90$ may be too high.
Now we find the price of a European call option on Intel with as striking price of $85$ that will expire in 3 months?

\[ P_t = 80, \; K = 85, \; T - t = 0.25, \; \sigma = 0.2 \]
and \( r = 0.08. \)

\[ h_1 = \frac{\ln(80/85) + (0.08 + .2^2/2)0.25}{0.2\sqrt{0.25}} = -0.356246 \]
\[ h_2 = h_1 - 0.2\sqrt{0.25} = -0.456246. \]

\[ \Phi(-0.356246) = 0.3608, \; \Phi(-0.456246) = 0.3241. \]

\[ \text{The price of a European call option is} \]
\[ c_t = 80 \times \Phi(-0.356246) \]
\[ -85 \times \Phi(-0.456246) \exp(-0.02) = 1.86. \]

\[ \text{The stock price has to rise by $6.86 for} \]
\[ \text{the purchaser of the put option to break even} \]

\* Under the above assumptions, the price of a European put option is
\[ p_t = 85 \exp(-0.02) \Phi(0.456246) \]
\[-80 \times \Phi(-0.356246) = 5.18.\]

The stock has to fall 0.18 for the purchaser of the put option to break even.
Call values when conditional variances change (product process)

- The Black-Scholes formula assumes the price logarithm followed a continuous-time Wiener process: daily returns then have independent and identical distributions.

- Suppose the conditional standard deviations $\sigma_t, \sigma_{t+1}, \cdots, \sigma_T$ are generated by the Gaussian process

  $$\ln(\sigma_{t+h}) - \alpha = \phi[\ln(\sigma_{t+h-1}) - \alpha] + \eta_t \quad (10)$$

  for $1 \leq h \leq T - t$ with $0 < \phi < 1$.

- The unconditional distribution of $\ln(\sigma_{t+h})$ is $N(\alpha, \beta^2)$, the unconditional variance of $\sigma_{t+h}$ is $\sigma_1^2 = \exp(2\alpha + 2\beta^2)$ and the $\eta_t$ are independently distributed as $N(0, \beta^2[1 - \phi^2])$.

- Suppose the return $X_t$ has distribution $N(\mu, \sigma_t^2)$ for some constant $\mu$.

- At time $t$, we know $P_t$ and $\sigma_t$. Then

  $$\ln(P_T) = \ln(P_t) + \sum_{i=1}^{T-t} r_{t+i}$$
where $r_{t+i} = \ln P_{t+i} - \ln P_{t+i-1}$.

To determine $c_t$, we need to find the distribution of $\sum_{i=1}^{T-t} r_{t+i}$. Under the above assumption, $\ln(P_T)$ is normally distributed with conditional mean $\ln(P_t) + (T - t)\mu$ and conditional variance

$$\sum_{h=1}^{T-t} E(\sigma_{t+h}^2 | \sigma_t)$$  \hspace{1cm} (11)

because the returns $r_{t+h}$ are uncorrelated.

- (11) can be evaluated via (10) to give

$$\ln(\sigma_{t+h} | \sigma_t) \sim N(\alpha + \phi^h [\ln(\sigma_t) - \alpha], \beta^2[1 - \phi^{2h}])$$

and hence

$$E(\sigma_{t+h}^2 | \sigma_t) = \exp\{2\alpha + 2\phi^h[\ln(\sigma_t) - \alpha] + 2\beta^2[1 - \phi^{2h}]\}.$$

- Note that we can still use $c_t$ derived under Black-Scholes model but we need to make corresponding change on different variance.

- Replace $\sigma\sqrt{T - t}$ by $\sigma_1$ and $\sigma^2$ by $\sigma_1^2/(T - t)$.