4.1

Where we are

- Last week: Interest Rate Risk and an Introduction to Value at Risk (VaR) (Chapter 8-9)
- This week
  - Finish-up a few items for VaR: Including, Marginal VaR and back-testing; then move to Volatility (Chapter 10)
- Next week: Correlation and the Copula model (Chapter 11)

4.2

Assignment

- For September 30th (This Week)
  - Read: Hull Chapters 10 - 11 (Volatility, Correlation and the Copula model)
  - Problems (Due September 30th)
    - Chapter 8: 6, 7, 10, 11; 16
    - Chapter 9: 1, 3, 4, 5 (a) – (d); 12 (a) – (d)
  - Problems (Due Oct 7th)
    - Chapter 9: 5 (e); 12 (e)
    - Chapter 10 (Oct 7): 1, 2, 5, 7, 8, 9, 11, 15, 16; 19, 22

3.3

Assignment

- For Oct 7th (Next Week)
  - Read: Hull Chapter 11 (Correlation and Copula models)
  - Problems (Due Oct 7th)
    - Chapter 9: 5 (e); 12 (e)
    - Chapter 10 (Oct 7): 1, 2, 5, 7, 8, 9, 11, 15, 16; 19, 22
  - Problems (Due Oct 15th)
    - Chapter 10 (Oct 15th): 21
    - Chapter 11 (Oct 15th): 1, 2, 5, 6, 8, 10, 14, 19
Assignment

- Midterm: October 30, 2013
- Final Exam
  - Wednesday, December 18th; 9am - Noon
  - Shaffer 202

Coherent Risk Measure

- A number of properties have been proposed as desirable for a risk measure if it is to be used to set capital requirements
- A risk measure that satisfies all these properties is called a coherent risk measure
- These properties follow …

Coherent Risk Measures

- Properties of coherent risk measures
  - If one portfolio always produces a worse outcome than another its risk measure should be greater (monotonicity)
  - If we add an amount of cash $K$ to a portfolio its risk measure should go down by $K$ (translation invariance) – buffer against loss
  - Changing the size of a portfolio by $\lambda$ should result in the risk measure being multiplied by $\lambda$ (homogeneity) – 2x the PF; 2x loss
  - The risk measures for two portfolios after they have been merged should be no greater than the sum of their risk measures before they were merged (subadditivity) – allows for diversification to reduce risk

VaR vs Expected Shortfall

- VaR satisfies the first three conditions but not always does it satisfy the fourth one
- Expected shortfall (C-VaR) satisfies all four conditions.
VaR vs Expected Shortfall

Example 9.6: Two $10 million one-year loans each of which has a 1.25% chance of defaulting. All recoveries between 0 and 100% are equally likely. If there is no default the loan leads to a profit of $0.2 million. If one loan defaults it is certain that the other one will not default.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neither loan defaults</td>
<td>97.50%</td>
</tr>
<tr>
<td>Loan 1 defaults; Loan 2 does not default</td>
<td>1.25%</td>
</tr>
<tr>
<td>Loan 2 defaults; Loan 1 does not default</td>
<td>1.25%</td>
</tr>
<tr>
<td>Both loans default</td>
<td>0.00%</td>
</tr>
</tbody>
</table>

4.9

VaR vs Expected Shortfall

Example 9.6 (Continued):
- Single Loan 1-yr 99% VaR is $2m
  - 1.25% chance of loss; if loss, then 80% chance it is > $2m (uniform)
  - Unconditional prob of loss > $2m is 80% of 1.25% = 1% (99% VaR)
- Two Loan PF 1-yr 99% VaR is $5.8m
  - Default occurs 1.25% of time, but never together
  - Pr a default occurs is 2.5% (1.0 - .975 = probability of a default)
  - If there is a loss, then 40% chance it is > $6m
  - Unconditional prob of loss > $6m is 40% of 2.5% = 1% (99% VaR)
  - A profit of $.2m on other loan plus loss of $6m => $5.8
- Two Loans Separately = 2 + 2 = $4m
- Two Loans Together = $5.8m > Separately => No Subadd.

4.10

VaR vs Expected Shortfall

Example 9.8: Consider same situation again, but for C-VaR
- C-VaR (Expected Shortfall) from 1 loan for 1-year and 99% confidence level is
  - Expected Loss conditional on loss > $2m = 99% 1-yr VaR
  - W/ uniform loss [0,$10m]; expected loss, conditioned on > $2m, is halfway along the interval [$2m,$10m] = $6m
- VaR for PF of 2 loans = $5.8m
  - C-VaR is expected loss on PF conditional on loss > $5.8m
  - When 1 loan defaults, other doesn’t; outcomes = uniform [+$.2,-$9.8m]
  - Expected loss given we are in the part of the distribution [$5.8m,$9.8m] is $7.8m
- Note: $6m + $6m > $7.8m, so C-VaR is subadditive

4.11

Spectral Risk Measures

- Risk measures are characterized by the weights assigned to quantiles of the loss distribution
- VaR assigns all weight to λth quantile
- Expected shortfall assigns equal weight to all quantiles greater than the λth quantile and zero to all below the λth quantile
- We now consider alternatives …

4.12
Spectral Risk Measures

- A “spectral risk measure” may be defined by making other assumptions about weights assigned to quantiles of the loss distribution
- A spectral risk measure is coherent, that is,
  - It satisfies the Subadditivity condition, IF
  - The weights are a non-decreasing function of the quantiles – such as with Conditional VaR
- Another possibility for a risk measure is to make the weight assigned to q-th quantile proportional to $e^{-\gamma q^2}$ where $\gamma$ is a constant
  - This is called the exponential spectral risk measure …

Normal Distribution Assumption

- The simplest assumption is that daily gains/losses are normally distributed and independent
- It is then easy to calculate 99%, 1-Day VaR from the standard deviation of the 1-day loss distribution
  - (1-day 99% VaR = 2.33$\sigma$)
  - Since $N(-2.33) = 0.01$ or $N(2.33) = 0.99$
- The $N$-day VaR equals $\sqrt{N}$ times the one-day VaR
  - Assuming the daily changes are i.i.d. normal, $N(0, \sigma^2)$
  - Regulators allow banks to calculate the 10 day VaR as $\sqrt{10}$ times the one-day VaR

Independence Assumption in VaR Calculations (Equation 9.3, page 193)

- When daily changes in a portfolio are identically distributed and independent the variance over $N$ days is $N$ times the variance over one day
  $$N \text{-Day VaR} = \sqrt{N} \times 1\text{-Day VaR}$$
- When there is first-order autocorrelation, correlation in the changes equal to $\rho$, the multiplier of the variance, $\sigma^2$, is increased from $N$ to $N + 2(N-1)\rho + 2(N-2)\rho^2 + 2(N-3)\rho^3 + \ldots 2\rho^{N-1}$
  - Where the correlation between $\Delta P_i$ and $\Delta P_j$ is $\rho^j$ and $\Delta P_i$ is the change in the PF value on day $i$
Impact of Autocorrelation: Ratio of N-day VaR to 1-day VaR (Table 9.1, page 193)

<table>
<thead>
<tr>
<th>N</th>
<th>1</th>
<th>2</th>
<th>5</th>
<th>10</th>
<th>50</th>
<th>250</th>
</tr>
</thead>
<tbody>
<tr>
<td>ρ=0</td>
<td>1.0</td>
<td>1.41</td>
<td>2.24</td>
<td>3.16</td>
<td>7.07</td>
<td>15.81</td>
</tr>
<tr>
<td>ρ=0.05</td>
<td>1.0</td>
<td>1.45</td>
<td>2.33</td>
<td>3.31</td>
<td>7.43</td>
<td>16.62</td>
</tr>
<tr>
<td>ρ=0.1</td>
<td>1.0</td>
<td>1.48</td>
<td>2.42</td>
<td>3.46</td>
<td>7.80</td>
<td>17.47</td>
</tr>
<tr>
<td>ρ=0.2</td>
<td>1.0</td>
<td>1.55</td>
<td>2.62</td>
<td>3.79</td>
<td>8.62</td>
<td>19.35</td>
</tr>
</tbody>
</table>

As correlation is present & increases from zero, the approximation of multiplying the variance by N can more seriously understate VaR

Choice of VaR Parameters

- Time horizon should depend on how quickly portfolio can be unwound. Regulators in effect use 1-day for bank market risk and 1-year for credit/operational risk. Fund managers often use one month.
- Confidence level depends on objectives. Regulators use 99% for market risk and 99.9% for credit/operational risk. A bank wanting to maintain a AA credit rating will often use 99.97% for internal calculations.
  - VaR for high confidence levels cannot be observed directly from data and must be inferred in some way.
  - One approach: Since $\text{VaR}(X) = \sigma N^{-1}(X)$ then for the same loss distribution:
    \[
    \text{VaR}(X) = \text{VaR}(X^*) \Rightarrow \text{VaR}(X^*) = \frac{N^{-1}(X^*)}{N^{-1}(X)} \text{VaR}(X)
    \]
    - Determine one confidence level VaR from another.
    - Better approach is to use EVT.

VaR PF Measures: An Amount $x_i$ is Invested in the $i^{th}$ Sub-PF

- Marginal VaR: Sensitivity of VaR to the size of the $i^{th}$ subportfolio, $x_i$: $\frac{\partial \text{VaR}}{\partial x_i}$
  - Marginal VaR is related to the CAPM and if a sub-PF’s beta is high, the marginal VaR will be large.
- Incremental VaR: Incremental effect of $i^{th}$ sub-PF on VaR
  - The difference in VaR with/without the $i^{th}$ sub-PF
  - Can calculate w/brute force: Find VaR w/o the $i^{th}$ sub-PF

VaR PF Measures: An Amount $x_i$ is Invested in the $i^{th}$ Sub-PF

- Component VaR: Marginal effect of $i^{th}$ sub-PF on VaR leads to the formula:
  \[
  \text{VaR} = \sum_{i} C_i
  \]
  - Where the partitioned component’s sub-PF Component VaR:
    \[
    C_i = x_i \frac{\partial \text{VaR}}{\partial x_i}
    \]
  - Add-up to the total VaR
  - Alternatively, consistent VaR (same time horizon and confidence) may be aggregated
    \[
    \text{VaR}_{\text{cons}} = \sqrt{\sum_i \sum_j \text{VaR}_i \text{VaR}_j \rho_{ij}}
    \]
Example of Component VaR

- Consider the $100 million PF of assets:
  
  $\text{VaR} = 1.645 \times 10.3\% \times $100 \times 100\% = $16.9$

  - As $\sigma^2 = \sum_{i=1}^{3} w_i^2 \sigma_i^2 + \sum_{i<j}^{3} w_i w_j \rho_{ij} \sigma_i \sigma_j$
  
  - Demonstrates the benefit of diversification

<table>
<thead>
<tr>
<th>Asset</th>
<th>Exp Ret %</th>
<th>Vol %</th>
<th>Corr</th>
<th>Alloc %</th>
<th>95% VR</th>
<th>VAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>US Stocks</td>
<td>1</td>
<td>13.8</td>
<td>1.0</td>
<td>60.0</td>
<td>$15.3</td>
<td></td>
</tr>
<tr>
<td>US Bonds</td>
<td>2</td>
<td>8.40</td>
<td>0.2</td>
<td>1.0</td>
<td>$0.9</td>
<td></td>
</tr>
<tr>
<td>Intl Stocks</td>
<td>3</td>
<td>9.60</td>
<td>0.04</td>
<td>1.0</td>
<td>$5.9</td>
<td></td>
</tr>
<tr>
<td>Portfolio</td>
<td>12.00</td>
<td>10.3</td>
<td></td>
<td>100.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Back-testing

- Back-testing a VaR calculation methodology involves looking at how often exceptions (loss > VaR) occur.

  - If exceptions occur more or less frequently than the model VaR predicts is there a problem?
  
  - We compare what we see with what we expect and test if the exceptions are consistent with the population to some level of confidence

Example of Component VaR

- With our definition of marginal VaR:
  
  Increasing the size of the Allocation to US Stocks from 60% to 61% increases PF VaR from $10.3 to $10.315, an increase of .144% (similarly for other assets)

<table>
<thead>
<tr>
<th>Asset</th>
<th>% Alloc</th>
<th>Marginal Comp Risk, C,</th>
</tr>
</thead>
<tbody>
<tr>
<td>US Stocks</td>
<td>60.0</td>
<td>0.144 8.63</td>
</tr>
<tr>
<td>US Bonds</td>
<td>7.7</td>
<td>0.028 0.21</td>
</tr>
<tr>
<td>Intl Stocks</td>
<td>32.3</td>
<td>0.0451 1.46</td>
</tr>
<tr>
<td>Portfolio</td>
<td>100.</td>
<td>10.3</td>
</tr>
</tbody>
</table>

Back-testing

- Suppose that the theoretical probability of an exception is $p (=1-X/100)$. The probability of $m$ or more exceptions in $n$ days is:

  $\sum_{k=m}^{n} \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}$

  - An often used confidence level in statistics is 5%
    
    - If the probability of VaR level being exceeded $m$ or more days is less than 5%, we reject the hypothesis that the probability of an exception is $p$
    
    - If the probability of the VaR level being exceeded on $m$ or more days is greater than 5%, the hypothesis is not rejected and that the probability of an exception is $p$ – the model for VaR is a good one
Example of Backtesting

- Suppose we back-test VaR w/600 days of data. VaR confidence is 99% and we observe 9 exceptions (the expected number of exceptions is 6)
  - Should we reject the model?
  - Probability of 9 or more exceptions in EXCEL is
    \[1 - \text{BINOMDIST}(8,600,0.01,\text{TRUE}) = 0.152 > .05\]
  - At 5% confidence level we should not reject the model
- Probability of 11 or more exceptions in EXCEL is
  \[1 - \text{BINOMDIST}(10,600,0.01,\text{TRUE}) = 0.019 < .05\]
- Probability of 10 or more exceptions in EXCEL is
  \[1 - \text{BINOMDIST}(9,600,0.01,\text{TRUE}) > .05\]
- The model is still a good one up thru 10 observations w/95% confidence – not too low an estimate for risk

Back-testing

- In the example, the model was for a 99% VaR
  - In 600 observations that equated to 6 times
  - With 9 exceptions, we were able to conclude that the model was still good – not too risky
  - On the other hand, if we observed 2 exceptions should we reject the model as being too conservative
- Here, if the probability of an exception is \( p = 1 - X/100 \), the probability of \( m \) or fewer exceptions is
  \[\sum_{k=0}^{m} \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}\]
  - This is compared to 5% as before

Example of Backtesting

- Suppose again we back-test VaR w/600 days of data. VaR confidence is 99% and we observe 1 exception)
  - Should we reject the model as too conservative?
  - Probability of 0 or 1 exceptions in EXCEL is
    \[\text{BINOMDIST}(1,600,0.01,\text{TRUE}) = 0.017 < .05\]
  - At 5% confidence level we should reject the model
  - However, if the number of exceptions had been 2 or more (up to 6) we would not reject the model with 95% confidence

Back-testing

- Alternatively, there is a relatively powerful 2-sided test.
  - If the probability of an exception under VaR is \( p \) and \( m \) exceptions in \( n \) trials are observed, then
    \[-2\ln \left( (1-p)^m p^0 \right) + 2\ln \left( (1-m/n)^m (m/n)^0 \right)\]
    should be chi-squared with 1-degree of freedom
  - The value of the statistic is very high for either low or high occurrences of exceptions
  - There is a probability of 5%that the chi-squared variable with 1-degree of freedom is greater than 3.84
  - If the above is greater than 3.84 we should reject the model – either at the low end or the high end
Example of Backtesting

- Suppose we again have a situation like the previous two examples where we back-test 99% VaR with 600 days of data.
  - The value of the statistic is greater than 3.84 when the number of exceptions is 1 and less or 12 and more.
  - Therefore we accept the VaR model when $2 \leq m \leq 11$ and reject otherwise.

The End for VaR

- For now ... we shall return and address more about VaR later.

Background on Volatility

- The volatility of a variable is the standard deviation of its return with the return being expressed with continuous compounding.
- The variance rate is the square of volatility.
- Implied volatilities are the volatilities implied from option prices.
- Normally days when markets are closed are ignored in volatility calculations (252 days per year; see Business Snapshot 10.1, page 207).

Implied Volatilities

- Of the variables needed to price an option the one that cannot be observed directly is volatility.
- We can therefore imply a volatilities from market prices and vice versa.
Price as the Market Variable

- For the non-dividend paying stock, we assumed a model for stock price movement where the expected percentage return is independent of stock price.
- If $S$ is the stock price at time $t$, the expected drift rate in $S$ should be assumed to be $\mu \times S$ for some constant parameter $\mu$.
- This means that in a short interval of time, $\Delta t$, the expected increase in $S$ is $\mu S \Delta t$.
- The parameter $\mu$ is the expected annual rate of return on the stock.
- If volatility of the stock price is always zero, then this model implies $\Delta S = \mu S \Delta t$ and as $dS = \mu S dt$ or $\frac{dS}{S} = \mu dt$.
- When the variance rate is zero, the stock price grows at a continuously compounded rate of $\mu$ per unit time.

Price as the Market Variable

- In practice a stock price does exhibit variance from expectation.
- A reasonable assumption is that the variability of the percentage return in a short period, $\Delta t$, is the same regardless of stock price.
- This suggests that the standard deviation in a short period of time, $\Delta t$, should be proportional, $\sigma$, to the stock price and leads to the model $dS = \mu S dt + \sigma S \, dz$ or $\frac{dS}{S} = \mu dt + \sigma dz$.
- The parameter $\sigma$ is the volatility of the market variable per year.
- The discrete time version is $\Delta S = \mu \Delta t + \sigma \sqrt{\Delta t}$ or $\Delta S = \mu S \Delta t + \sigma S \epsilon \sqrt{\Delta t}$.
- Where $\Delta S$ is the change in stock price in a small time interval $\Delta t$ and $\epsilon$ is standard normal $(0,1)$.

Price as the Market Variable

- In summary, $\Delta S/S$, is normally distributed with mean $\mu \Delta t$ and standard deviation $\sigma \sqrt{\Delta t}$.
- Consider the process followed by $\ln S$, where $S$ follows geometric Brownian motion.
- Define $G = \ln S$, then since $\frac{dG}{S} = \frac{1}{S} \frac{dS}{S} = -\frac{1}{S} + \frac{dG}{S}$ and $\frac{dG}{S} = 0$ it follows from Ito’s lemma that the process followed by $G$ is $dG = \left( \mu - \frac{\sigma^2}{2} \right) dt + \sigma dz$.
- Since $\mu$ and $\sigma$ are constant, $G = \ln S$ follows a generalized Wiener process with constant drift rate $\mu - \frac{\sigma^2}{2}$ and constant variance rate $\sigma^2$.

Price as the Market Variable

- This means that the change in $\ln S$ between 0 and a future time $T$ is normally distributed with mean $(\mu - \sigma^2/2)T$ and variance $\sigma^2 T$.
- $\ln S_T = \ln S_0 \sim \phi \left( \left( \mu - \frac{\sigma^2}{2} \right) T, \sigma^2 T \right)$ or $\ln S_T \sim \phi \left( \ln S_0 + \left( \mu - \frac{\sigma^2}{2} \right) T, \sigma^2 T \right)$.
- Where $\phi(m,v)$ denotes a normal distribution with mean $m$ and variance $v$.
- Remember that a basic property of the Wiener process $dz$ is that $[dz]^2$ is normally distributed with mean $0$ and standard deviation $\sigma^2 T$.
- The equation above shows that $\ln S_T$ is normally distributed.
- Furthermore, a variable is said to have a lognormal distribution if the natural logarithm of the variable is normally distributed.
Price as the Market Variable

- For the non-dividend paying stock, we assume a model for stock price movement where the expected percentage return is independent of stock price
  - For this process, as we have developed,
    \[ \frac{\Delta S}{S} = \phi(\mu \Delta t, \sigma^2 \Delta t) \]
  - And we have shown that
    \[ \ln \frac{S_t}{S_0} = \ln S_0 - \phi \left( \mu - \frac{\sigma^2}{2} \right) T, \sigma^2 T \]
  - In either case, same variance rate (volatility)
  - We use both models interchangeably in quantifying volatility

\[ \text{4.37} \]

Price as the Market Variable

- If we consider the process followed by \( \ln S \), where \( S \) follows geometric Brownian motion and
  - Define \( G = \ln S \), then since \( \frac{\partial G}{\partial S} = \frac{1}{S} \), \( \frac{\partial^2 G}{\partial S^2} = -\frac{1}{S^2} \), and \( \frac{\partial G}{\partial t} = 0 \) it follows from Ito’s lemma that the process followed by \( G \) is
    \[ dG = \left( \mu - \frac{\sigma^2}{2} \right) dt + \sigma dz \]
  - We get
    \[ \ln \frac{S_t}{S_0} = \ln S_0 - \phi \left( \mu - \frac{\sigma^2}{2} \right) T, \sigma^2 T \]
  - In either case, same variance rate (volatility)
  - We use both models interchangeably in quantifying volatility

\[ \text{4.39} \]

Price as the Market Variable

- Determining volatility from historical price data
  - We usually observe prices at fixed intervals (daily, weekly, or monthly)
  - Define \( n+1 \): Number of observations
  - \( \pi_i \): Stock price at the end of the \( i \)th interval, \( i = 0,1,...,n \)
  - \( \tau \): Length of the time interval in years
  - Let \( u_i = \ln \left( \frac{S_i}{S_{i-1}} \right) \) for \( i = 1, 2, ..., n \)
  - So the usual estimate, \( s \), of the standard deviation of the \( u_i \) is given by
    \[ s = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (u_i - \bar{u})^2} \] or
    \[ s = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} u_i^2 - \frac{1}{n(n-1)} \left( \sum_{i=1}^{n} u_i \right)^2} \]
  - Where \( \bar{u} \) is the mean of the \( u_i \): \( \bar{u} = \frac{1}{n} \sum_{i=1}^{n} u_i \)

\[ \text{4.40} \]
Price as the Market Variable

- From \( \ln \frac{S_T}{S_t} \sim \phi \left( \frac{(\mu - \sigma^2/2)T, \sigma\sqrt{T}}{\sqrt{2\pi}} \right) \) and \( \frac{S_t}{S_{t-1}} \sim \phi \left( \frac{\mu - \sigma^2/2}{\sigma\sqrt{\Delta t}} \right) \)

we see that the variance of the \( u_i \) is \( \sigma^2/\tau/n \) (where \( \tau/n \) is the time interval associated with each \( u_i \)).

- The variable \( s \) is therefore an estimate of \( \sigma \sqrt{\tau} \) and it follows that \( \sigma \) itself can be estimated as \( \hat{\sigma} = \frac{s}{\sqrt{2n}} \).

- The standard error of this estimate is approximately \( \sigma/\sqrt{2n} \).

- More data adds to the accuracy of the estimate, but the assumption that volatility is a constant is only an expedience; it does change over time and old data may not be relevant for predicting future volatility.

- Most analyses for volatility derived VaR depends on the Ito process with Gaussian randomness in the time series.

- Is that a good/reasonable assumption?

Are Daily Changes in Exchange Rates Normally Distributed? Table 10.1, page 209

<table>
<thead>
<tr>
<th></th>
<th>Real World (%)</th>
<th>Normal Model (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;1 SD</td>
<td>25.04</td>
<td>31.73</td>
</tr>
<tr>
<td>&gt;2SD</td>
<td>5.27</td>
<td>4.55</td>
</tr>
<tr>
<td>&gt;3SD</td>
<td>1.34</td>
<td>0.27</td>
</tr>
<tr>
<td>&gt;4SD</td>
<td>0.29</td>
<td>0.01</td>
</tr>
<tr>
<td>&gt;5SD</td>
<td>0.08</td>
<td>0.00</td>
</tr>
<tr>
<td>&gt;6SD</td>
<td>0.03</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Alternatives to Normal Distributions: The Power Law:

\[
Prob(v > x) = Kx^{-\alpha}
\]

- Seems to fit the tail behavior of the returns for many market variables better than the normal distribution.

- where \( x \) is the number of standard deviations.

- and two parameters define the exponential distribution.

- So we have

\[
\ln[Prob(v > x)] = \ln K - \alpha \ln x
\]
Alternatives to Normal Distributions: The Power Law

Let $x$ be the number of standard deviations from Table 10.2 for the Real World data.

<table>
<thead>
<tr>
<th>$x$</th>
<th>$\ln x$</th>
<th>$\text{Prob}(\nu &gt; x)$</th>
<th>$\ln[\text{Prob}(\nu &gt; x)]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.000</td>
<td>0.12520</td>
<td>−2.078</td>
</tr>
<tr>
<td>2</td>
<td>0.693</td>
<td>0.02635</td>
<td>−3.636</td>
</tr>
<tr>
<td>3</td>
<td>1.099</td>
<td>0.00670</td>
<td>−5.006</td>
</tr>
<tr>
<td>4</td>
<td>1.386</td>
<td>0.00145</td>
<td>−6.536</td>
</tr>
<tr>
<td>5</td>
<td>1.609</td>
<td>0.00040</td>
<td>−7.824</td>
</tr>
<tr>
<td>6</td>
<td>1.792</td>
<td>0.00015</td>
<td>−8.805</td>
</tr>
</tbody>
</table>

Estimating Volatility – Standard Approach

- The alternative to the normal distribution – the power law – shows promise in characterizing the tail behavior (we will come back to this idea later)
- In the mean time most risk managers still work with, and prefer, volatility-based VaR
- We look at standard techniques and refinements that are popular
Estimating Volatility – Standard Approach

- Define \( \sigma_n \) as the volatility per day between day \( n-1 \) and day \( n \), as estimated at end of day \( n-1 \)
- Define \( S_i \) as the value of market variable at end of day \( i \)
- Define \( u_i = \ln(S_i/S_{i-1}) \), the continuously compounded return during day \( i \), and where we use only the last \( m \) days of data

\[
\sigma_n^2 = \frac{1}{m-1} \sum_{i=m}^{n}(u_{n-1} - \pi)^2
\]

\[
\pi = \frac{1}{m} \sum_{i=m}^{n} u_{n-1}
\]

for an unbiased estimate of the variance rate per day, \( \sigma_n^2 \)

Estimating Volatility – Refining the Standard Approach

- For purpose of monitoring daily volatility, the previous estimate is changed in the following ways
  - Define \( u_i \) as \( (S_i - S_{i-1})/S_{i-1} \) the percent day-change
  - Assume that the mean value of \( u_i \) is zero
  - Replace \( m-1 \) by \( m \) (invokes a maximum likelihood estimate vs. an unbiased estimate – more later)

This gives

\[
\sigma_n^2 = \frac{1}{m} \sum_{i=m}^{n} u_{n-1}^2
\]

Estimating Volatility – Variations to the Standard Approach

- Instead of assigning equal weights to the observations \( u_i \), we can alternatively set

\[
\sigma_n^2 = \sum_{i=1}^{n} \alpha_i u_i^2, \quad \text{where} \quad \sum_{i=1}^{n} \alpha_i = 1
\]

- Where, if we choose \( \alpha_i < \alpha_j \), for \( i > j \), less weight is given to older observations

- An extension to this idea is to assume there is a long run average variance rate, \( V_L \), and that this should be given some weight also,

\[
\sigma_n^2 = \gamma V_L + \sum_{i=1}^{n} \alpha_i u_i^2, \quad \text{where} \quad \gamma + \sum_{i=1}^{n} \alpha_i = 1
\]

- The ARCH(m) model (AutoRegressive Conditional Heteroscedacity)
  - We can alternatively denote \( \gamma V_L = \omega \)

Estimating Volatility – Refining the Standard Approach

- In an exponentially weighted moving average model (EWMA), the weights assigned to the \( u_i^2 \) decline exponentially moving back through time
  - In \( \sigma_n^2 = \sum_{i=1}^{n} \alpha_i u_i^2 \) where \( \sum_{i=1}^{n} \alpha_i = 1 \) we let \( \alpha_i = \lambda \alpha_{i-1} \)
  - This leads to a recursive relationship

\[
\sigma_n^2 = \lambda \sigma_{n-1}^2 + (1-\lambda)u_{n-1}^2
\]

- Whereby little data needs to be saved; only the last estimate and the new update (estimate for day \( n \))

- A small \( \lambda \) leads to heavy weight being put on most recent observation; a large \( \lambda \) provides an estimate that changes more slowly, with less volatility in the estimate

- RiskMetrics has found the value \( \lambda = 0.94 \) to be satisfactory
**Example**

Suppose that $\lambda = 0.90$, the volatility estimated for a market variable for day $n-1$ is 1% per day, and during day $n-1$ the market variable increased by 2%. This means that $\sigma^2_{L-1} = (0.01)^2 = 0.0001$ and $\sigma^2_{L-1} = 0.02^2 = 0.0004$. Equation (21.7) gives

$$\sigma^2 = 0.9 \times 0.0001 + 0.1 \times 0.0004 = 0.00013$$

The estimate of the volatility, $\sigma_n$, for day $n$ is therefore $\sqrt{0.00013}$, or 1.14%, per day. Note that the expected value of $\sigma^2_{L-1}$ is $\sigma^2_{L-1}$, or 0.0001. In this example, the realized value of $\sigma^2_{L-1}$ is greater than the expected value, and as a result our volatility estimate increases. If the realized value of $\sigma^2_{L-1}$ had been less than its expected value, our estimate of the volatility would have decreased.

---

**Estimating Volatility – Variations to the Standard Approach**

- GARCH “(1,1)” indicates the estimate is calculated from the single most recent $u^2$ and the single most recent estimate of the variance rate
- The more general GARCH “(p,q)” estimates from the p most recent $u^2$ and the q most recent estimates of the variance rate
- Setting $\omega = \gamma V_L$, the GARCH (1,1) model can also be written as $\sigma^2 = \omega + \alpha u^2_{L-1} + \beta \sigma^2_{L-1}$ where $V_L = \frac{\omega}{1-\alpha-\beta}$
  - For stable GARCH, $\alpha + \beta < 1$ so weight on $V_L$ is > 0
  - For GARCH (p,q) $\sigma^2 = \omega + \sum_{i=1}^{p} \alpha_i u^2_{L-i} + \sum_{j=1}^{q} \beta_j \sigma^2_{L-j}$

---

**Example**

Suppose that a GARCH(1,1) model is estimated from daily data as

$$\sigma^2 = 0.000002 + 0.13 \sigma^2_{L-1} + 0.86 \sigma^2_{L-1}$$

This corresponds to $\alpha = 0.13$, $\beta = 0.86$, and $\omega = 0.000002$. Because $\gamma = 1 - \alpha - \beta$, it follows that $\gamma = 0.01$. Because $\omega = \gamma V_L$, it follows that $V_L = 0.0002$. In other words, the long-run average variance per day implied by the model is 0.0002. This corresponds to a volatility of $\sqrt{0.0002} = 0.014$, or 1.4%.

Suppose that the estimate of the volatility on day $n-1$ is 1.6% per day, so that $\sigma^2_{L-1} = 0.016^2 = 0.000256$, and that on day $n$ the market variable decreased by 1%, so that $\sigma^2_{L-1} = 0.01^2 = 0.0001$. Then

$$\sigma^2 = 0.000002 + 0.13 \times 0.0001 + 0.86 \times 0.000256 = 0.00023516$$

The new estimate of the volatility is therefore $\sqrt{0.00023516} = 0.153$, or 1.53%, per day.
Estimating Volatility – Parameterizing the Models

- Key to making the Volatility Estimation approaches effective is to have a way to parameterize the models that gives good results in reproducing the data
- In maximum likelihood methods we choose parameters that maximize the likelihood of the model to predict the data observed
- Remember maximum likelihood techniques?

A simple example of the maximum likelihood (ML) technique in statistics

- We observe that a certain event happens one time in ten trials. What is our estimate of the proportion of the time, \( p \), that it happens?
- The probability of the event happening on one particular trial and not on the others is \( p(1-p)^9 \)
- We maximize to obtain a ML estimate
  \[
  L(p) = p(1-p)^9, \text{ then } \max_p L(p) \text{ is found by solving } L'(p)=0
  \]
- Result: \( p=0.1 \)

A more pertinent example

- Estimate the variance of \( m \) observations, \( u_1, u_2, \ldots, u_m \) from a normal distribution with mean zero
- The likelihood of the \( m \) observations appearing as they were observed is
  \[
  \prod_{i=1}^{m} \frac{1}{\sqrt{2\pi v}} \exp\left(\frac{-u_i^2}{2v}\right)
  \]
- The ML estimate of \( \nu \)
  \[
  \max_{\nu} \prod_{i=1}^{m} \frac{1}{\sqrt{2\pi \nu}} \exp\left(\frac{-u_i^2}{2\nu}\right)
  \]
- Is equivalent to taking logarithms and maximizing
  \[
  \sum_{i=1}^{m} -\ln(\nu) - \frac{u_i^2}{2\nu} \text{ or } -m\ln(\nu) - \sum_{i=1}^{m} \frac{u_i^2}{2\nu}
  \]
- Which results in:
  \[
  \nu = \frac{1}{m} \sum_{i=1}^{m} u_i^2
  \]

For application of ML estimation to parameters of GARCH\((1,1)\), the problem is to find the best parameters \((\omega, \alpha, \beta)\) in the expression for the estimator \( \sigma_i^2 = \omega + \alpha u_{i-1}^2 + \beta \sigma_{i-1}^2 = \nu_i \)

- This is
  \[
  \max_{\omega, \alpha, \beta} \prod_{i=1}^{m} \frac{1}{\sqrt{2\pi \nu_i}} \exp\left(\frac{-u_i^2}{2\nu_i}\right)
  \]
- Or equivalently
  \[
  \max_{\omega, \alpha, \beta} \sum_{i=1}^{m} \left[ -\ln(\nu_i) - \frac{u_i^2}{2\nu_i} \right]
  \]
Estimating Volatility – Parameterizing the Models (Example: Table 10.4)

- An example of ML estimation for the GARCH(1,1) model
- Data for Yen/US$ exchange rate
- See the Table 10.4
- Start with trial values of \( \omega, \alpha, \text{ and } \beta \)
- Update variances: \( v_i = \omega + \alpha u_{i-1}^2 + \beta v_{i-1} \)
- Calculate: \( \sum_{i=1}^{n} \left[ -\ln(v_i) - \frac{u_i^2}{v_i} \right] \)
- Use solver to search for values of \( \omega, \alpha, \text{ and } \beta \) that maximize this objective function
- Important note: set up spreadsheet so that you are searching for three numbers that are the same order of magnitude (?!)

\[
\sum_{i} \left[ -\ln(v_i) - \frac{u_i^2}{v_i} \right]
\]

Estimating Volatility – Parameterizing the Models (Table 10.4)

One way of implementing GARCH(1,1) that increases stability is by using Variance Targeting

- We set the long-run average volatility equal to the sample variance
- Calculated from the data or another reasonable value
- Only two other parameters need be estimated
- Indeed, since \( V_L = \frac{\omega}{1-\alpha - \beta} \) \( \omega = V_L (1-\alpha - \beta) \)
- When EWMA is used, \( \omega = 0, \alpha = 1 - \lambda, \text{ and } \beta = \lambda \)
- And only \( \lambda \) need be estimated

Estimating Volatility – Parameterizing the Models

- How different/good are the various estimators
- For comparison in the example:
  - 3-parameter G(1,1): Objective function = 22,063.5763
  - 2-parameter w/variance targeting: = 22,063.5274
  - 1-parameter EWMA: = 21,995.8377

\[
\begin{array}{cccccc}
\text{Date} & \text{Day} & S_i & u_i & v_i \alpha \beta = \omega_i - \gamma_i^2 & -\ln(v_i) - u_i^2/v_i \\
06-Jan-88 & 1 & 0.007728 & & & 9.6283 \\
07-Jan-88 & 2 & 0.007759 & 0.06599 & & \\
08-Jan-88 & 3 & 0.007746 & -0.004242 & 0.00004355 & 8.1329 \\
11-Jan-88 & 4 & 0.007816 & 0.00907 & 0.00004198 & 9.8568 \\
12-Jan-88 & 5 & 0.007837 & 0.002677 & 0.00004475 & 7.1529 \\
13-Jan-88 & 6 & 0.007924 & 0.011101 & 0.00004200 & \\
13-Aug-97 & 2421 & 0.008643 & 0.003374 & 0.00007626 & 9.3321 \\
14-Aug-97 & 2422 & 0.008493 & -0.017309 & 0.00007092 & 5.3204 \\
15-Aug-97 & 2423 & 0.008495 & 0.001444 & 0.00008417 & 9.3824 \\
\end{array}
\]

\( \omega = 0.00000176 \quad \alpha = 0.0626 \quad \beta = 0.8976 \)
Estimating Volatility – Parameterizing the Models

- Eliminating autocorrelation: If the estimate for $\frac{u_i^2}{\sigma_i^2}$ is free of any of the autocorrelation present in the $u_i^2$, the model for $\sigma_i^2$ has succeeded in explaining that autocorrelation and can be judged to be a “good model”
- See the data in Table 10.5 on the next page for yen-dollar from before
  - The table shows that the autocorrelation are positive for $u_i^2$ for all lags between 1 and 15
  - On the other hand, for $u_i^2/\sigma_i^2$, some are positive and some are negative – and always of smaller magnitude – suggesting a “good result”

Estimating Volatility – Parameterizing the Models

<table>
<thead>
<tr>
<th>Time lag</th>
<th>Autocorrelation for $u_i^2$</th>
<th>Autocorrelation for $u_i^2/\sigma_i^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.072</td>
<td>0.000</td>
</tr>
<tr>
<td>2</td>
<td>0.041</td>
<td>-0.010</td>
</tr>
<tr>
<td>3</td>
<td>0.057</td>
<td>0.005</td>
</tr>
<tr>
<td>4</td>
<td>0.107</td>
<td>0.000</td>
</tr>
<tr>
<td>5</td>
<td>0.075</td>
<td>0.011</td>
</tr>
<tr>
<td>6</td>
<td>0.066</td>
<td>0.009</td>
</tr>
<tr>
<td>7</td>
<td>0.019</td>
<td>-0.034</td>
</tr>
<tr>
<td>8</td>
<td>0.085</td>
<td>0.015</td>
</tr>
<tr>
<td>9</td>
<td>0.054</td>
<td>0.009</td>
</tr>
<tr>
<td>10</td>
<td>0.030</td>
<td>-0.022</td>
</tr>
<tr>
<td>11</td>
<td>0.038</td>
<td>-0.004</td>
</tr>
<tr>
<td>12</td>
<td>0.038</td>
<td>-0.021</td>
</tr>
<tr>
<td>13</td>
<td>0.057</td>
<td>-0.001</td>
</tr>
<tr>
<td>14</td>
<td>0.040</td>
<td>0.004</td>
</tr>
<tr>
<td>15</td>
<td>0.007</td>
<td>-0.026</td>
</tr>
</tbody>
</table>

But, we can be more scientific using the Ljung-Box statistic to confirm this observation …

- If a series has $m$ observations the Ljung-Box statistic is
  $$m \sum_{k=4}^{K} W_k \eta_k^2$$
  where: $\eta_k$ is the autocorrelation of lag $k$
  $K$ is the number of lags, and
  $$w_j = \frac{m + 2}{m - k}$$
- For $K = 15$, zero autocorrelation can be rejected with 95% confidence when L-B statistic is greater than 25
- For Table 10.5 where $K$ is 15 :
  - For the $u_i^2$, the L-B = $123 > 25$ => autocorrelation is present
  - For the $u_i^2/\sigma_i^2$, the L-B = 8.2 < 25 => autocorrelation removed