Volatility and Risk
... an introduction to the concepts
Topical: Natural Gas and Propane prices soar ...

Source: Energy Information Administration, data are from reports released Jan 23, 2014.
... massive inventory drawdown due to record-setting cold weather

The “polar vortex” in early January pushed temperatures in the East to record lows ...

... which in turn caused a huge drawdown in seasonal natural gas inventories.
Nat gas is in backwardation ...

How to play it?? A14S to A15L spread? Long on May (will this be resolved by then)? We must watch the spread at least.

### Henry Hub Natural Gas Futures

**View Product List**

**Quotes | Contract Specifications | Performance Bonds / Margin**

<table>
<thead>
<tr>
<th>Month</th>
<th>Chars</th>
<th>Last</th>
<th>Change</th>
<th>Prior Settle</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feb 2014</td>
<td>OPT</td>
<td>4.910</td>
<td>+0.083</td>
<td>4.847</td>
</tr>
<tr>
<td>Mar 2014</td>
<td>OPT</td>
<td>4.730</td>
<td>+0.056</td>
<td>4.674</td>
</tr>
<tr>
<td>Apr 2014</td>
<td>OPT</td>
<td>4.297</td>
<td>+0.011</td>
<td>4.286</td>
</tr>
<tr>
<td>May 2014</td>
<td>OPT</td>
<td>4.225</td>
<td>-0.009</td>
<td>4.234</td>
</tr>
<tr>
<td>Jun 2014</td>
<td>OPT</td>
<td>4.271</td>
<td>+0.020</td>
<td>4.251</td>
</tr>
<tr>
<td>Jul 2014</td>
<td>OPT</td>
<td>4.273</td>
<td>-</td>
<td>4.273</td>
</tr>
<tr>
<td>Aug 2014</td>
<td>OPT</td>
<td>4.272</td>
<td>-</td>
<td>4.272</td>
</tr>
<tr>
<td>Sep 2014</td>
<td>OPT</td>
<td>-</td>
<td>-</td>
<td>4.251</td>
</tr>
<tr>
<td>Oct 2014</td>
<td>OPT</td>
<td>4.262</td>
<td>-0.003</td>
<td>4.265</td>
</tr>
<tr>
<td>Nov 2014</td>
<td>OPT</td>
<td>-</td>
<td>-</td>
<td>4.310</td>
</tr>
<tr>
<td>Dec 2014</td>
<td>OPT</td>
<td>-</td>
<td>-</td>
<td>4.420</td>
</tr>
<tr>
<td>Jan 2015</td>
<td>OPT</td>
<td>-</td>
<td>-</td>
<td>4.512</td>
</tr>
<tr>
<td>Feb 2015</td>
<td>OPT</td>
<td>-</td>
<td>-</td>
<td>4.466</td>
</tr>
<tr>
<td>Mar 2015</td>
<td>OPT</td>
<td>-</td>
<td>-</td>
<td>4.395</td>
</tr>
<tr>
<td>Apr 2015</td>
<td>OPT</td>
<td>-</td>
<td>-</td>
<td>4.397</td>
</tr>
</tbody>
</table>

**NG G4-1 [10] - NATURAL GAS ELECTRONIC**

**Last:** 4.910  **Change:** +0.003  **High:** 4.910  **Low:** 3.997  **4/20/2014**
The role played by uncertainty

The finance markets exist, and financial assets exist, and their prices exist, and their prices are variable and even sometimes volatile, because of uncertainty about the future value of financial assets, such as a company's profits, the value of a home, the future income stream from a bond, or in the case of esoteric derivatives, the future value of a price. Precise predictability is impossible. Even general fuzzy predictability (with a wide range and a wide time horizon) can be difficult and unreliable.

Solid financial theory is less about becoming a better predictor and more about making decisions that do not rely upon predictions. To "win" in a zero- or positive-sum game, you only need a slight edge.
The relevance of this ...

Asset price volatility, such as volatility in the stock market, is a major source of risk. We must try to find reliable measures of risk if we want to minimize risk while targeting yields or other measures of performance from financial asset portfolios.

We often have access to time-series data (in other words, the history of the data series) and we can use common statistical techniques to find useful patterns of information in that data. When we have historical series like the graph shown next, whether of stock indexes, individual stocks, yields on bonds, or futures and options values, we can assess risk up to a point.

Common sense tells us that we are likely to make some use of measures of dispersion like variance or standard deviation as a point and then gradually refine it.
Common assumptions made about the price performance over time of primary assets and their derivatives

- (For purposes of mathematical ease and because historical data conform to this assumption *within limits*), since the time of Black, Scholes, and Merton, prices paths and their growth rates are assumed to be continuous.

- The price behavior of a financial asset (FA) is independent of its past price behavior (Markov Chain)
  - also referred to as a "random number walk"
  - this is very debatable, but is the basis for a lot of modern modeling
  - this denies the possibility of so-called "technical analysis" and "charting."
  - was this done to make the math models, like Black-Scholes, work, or because it is true?
  - is not meant to imply that the price of a share of stock is unrelated to its price the previous day – this is not pure Brownian Motion

- The past price behavior of a FA may be filtered in a way that gives some reliable indicator of the risk associated with the FA.
... assumptions (continued)

- The previous assumption implies that we can use historical time series data for FAs to *partly* estimate measures of their risk (caveat: the past doesn't always repeat itself).
- We typically assume that the *rates of return* for FAs can be represented as random variables that conform to a Gaussian (normal) probability distribution
  - ... which further typically implies that the raw data from which the rates of return were calculated conform to an asymmetric distribution like lognormal.
  - ... and this assumption requires that when working with raw time series data that it be converted to continuous growth rates before risk estimates are made.
  - ... and at some point, this assumption must be to a test, like Anderson-Darling or Kolmogorov-Smirnov normalcy test.
Example of a high risk, high gain / high loss stock

Dendreon (DNDN), a biotechnology research company who makes Provenge, an experimental and controversial prostate cancer vaccine.

DNDN Yahoo Beta was 5.29 compared to 0.62 for XOM (Exxon-Mobile) on 1/27/2012.

High gain (Stage III FDA clinical trial success)  
High loss (vaccine costs $93,000 per treatment and may not work).

Source: finance.yahoo
Question: When using time-series data ...

Consider a financial asset with a variable price, like a stock. Common sense tells us that volatile, and hence risky, stocks will have a higher standard deviation than quiet stocks.

Is the standard deviation of the historical time series for stock continuous growth rates, especially when compared to each other, a useful measure of risk?

Riskier??

Ummm.. maybe. It's a good starting point.
The assumed probability distribution of FA continuous growth rates (or similar): Gaussian (normal)

We will use the **standard deviations** of these distributions as our first proxy for risk.
The Standard Normal Distribution
which we will use a lot

Dividing a normal distribution with mean 0 by its standard deviation produces the standard normal distribution, where we can describe the probability of a number being X standard deviations away from its mean. Shown is the probability of a value being less than +1 SD.
Memo slide: The Normal Distribution

Gaussian (Normal) Probability Density Function

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

Probability Cumulative Distribution Function

\[ F(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{x} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt \]

or

\[ F(x; \mu, \sigma^2) = \frac{1 + \text{errorfunction} \left( \frac{x - \mu}{\sigma \sqrt{2}} \right)}{2} \]
The Taylor Series expansion of the error function is:

$$
ef(z) = \frac{2}{\sqrt{\pi}} \sum_{i=0}^{\infty} \frac{(-1)^n z^{2n+1}}{n!(2n+1)} = \frac{2}{\sqrt{\pi}} \left( z - \frac{z^3}{3} + \frac{z^5}{10} - \frac{z^7}{42} + \cdots \right)
$$

You don’t have to go very far in the series before convergence and this is trivial to code. My little Excel equiv of NORMDIST:

Also useful:

If $X$ is $N(\mu, \sigma^2)$

Then $aX+b$ is $N(a\mu + b, a^2 \sigma^2)$
How do we represent relative risk?

The log-normal distributions on the right are transformed from normal daily continuous growth distributions, with the same mean (zero), but the top has a standard deviation of 0.05 and the bottom has a standard deviation of 0.08.

The bottom has a wider dispersion, so although there is a higher probability of a great gain, there is also a higher probability of a great loss. That is seen as “riskier.”
Other distributions for risk

Lognormal distributions and other skewed distributions are also used to represent the nature of risk (and transformation from normal to lognormal and vice-versa is trivial) ...

This distribution is skewed ... it is a transformed log-normal plot of two slides ago.
Memo: Power-law models for catastrophic risk, also called “tail risk,” a hot subject these days.

I found this is an article exploring models used to predict the probability of a rare event like 9/11:

Reminder from previous lecture.

**Typical first steps in estimating measures of dispersion (risk proxies)**

Given some *sample* from a *population* set of prices, such as the daily closing price of CSCO from Jan 2009 to Dec 2010, a typical first step is to convert the data to continuous growth rates:

$$CGR_{px} = R_p = \ln\left(\frac{P_t}{P_{t-1}}\right)$$

for each paired observation.

In Excel:

<table>
<thead>
<tr>
<th>Sample</th>
<th>Price</th>
<th>CGRP</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>22.34</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>22.86</td>
<td>0.0230</td>
</tr>
<tr>
<td>3</td>
<td>23.01</td>
<td>0.0065</td>
</tr>
<tr>
<td>4</td>
<td>22.79</td>
<td>-0.0096</td>
</tr>
<tr>
<td>5</td>
<td>23.41</td>
<td>0.0268</td>
</tr>
<tr>
<td>6</td>
<td>23.56</td>
<td>0.0064</td>
</tr>
</tbody>
</table>

Original sample (in part)  
Converted data
Elementary estimates of historical yield (alpha) and risk (beta)

Using the transformed continuous growth rates from the previous slide ($R_i$) we calculate the mean growth rate, which is our *alpha* estimate …

\[
R_p = \frac{\sum_{i=1}^{n} R_i}{n} = \mu_p
\]

… then we calculate the variance of the same …

\[
V_p = \frac{\sum_{i=1}^{n} (R_i - \bar{R})^2}{n} = \sigma_p^2
\]

… and finally, the square root of variance, the standard deviation, is our *beta*, or risk proxy.

\[
SD_p = \sqrt{V_p} = \sigma_p
\]
The Sharpe Ratio

Historical Sharpe Ratio: The ratio of the stock's (or other FA) historical rate of return over its volatility, over the same period:

$$SR = \frac{\mu}{\sigma}$$

Some versions make calculate this as an opportunity-cost return by subtracting the risk-free interest rate from yield:

$$SR = \frac{(\mu - r)}{\sigma}$$

Investment strategists may replace mu with their alpha and historical standard deviation with implied volatility.
The elementary \textit{Beta} risk estimator from 104

\[
B_x = \frac{COV_{xs}}{V_s}
\]

and

\[
COV_{xs} = CC_{xs} \times SD_x \times SD_s
\]

which is the covariance of the financial asset and the index (or portfolio) divided by the variance of the index (or the portfolio). But variance and covariance of what? As we will now understand, never unconverted data, but either log growth rates or yields (the latter used when evaluating the volatility of yield-bearing financial assets).
... and the Beta Rules of Thumb

When comparing the growth rate of a stock price to that of an index like the S&P 500, then (subject to a qualification),

- If Beta > 1, then the asset is more volatile than the index. If the asset were added to the index, the index would become more volatile.
- If Beta > 0 and <1, the asset is less volatile than the index. If the asset were added to the index, the index would become less volatile.
- If Beta is negative, the financial asset in question tends to move in the opposite direction of the index (the covariance is negative). Bond yields relative to the S&P 500 might be an example.

Can be more volatile not reflected in Beta if CC is very low but positive.
DWBH: We have a dedicated lecture about the VIX.

Market Volatility Proxy – The VIX

The VIX is a popular and very useful generic volatility indicator for the S&P 500. It is calculated using an options pricing model (which makes it comparable to calculating volatility for individual stocks). We review the technique later. If you trade options, you know where the VIX is.

Great time to trade options (this hit 80 in late 2008).

A panic:

Lousy time to trade options.

Source: finance.yahoo
Using a histogram …

In homework 1 we take our daily continuous growth rate data and, among other things, arrange it into histograms like the one shown on the right.

A histogram divides the full range of your data into an arbitrary range of odd-numbered intervals of equal size, like 11 in our example.

Then the program counts the frequency of observations for each interval (such as 60 for the center interval in the diagram shown) and maps each interval as a bar. In our model we overlay the histogram with a mapping of the continuous density distribution taken from our estimate of the mean and standard deviation of the same data, assuming a normal distribution. The comparison shows how “close” we are to normal, and also identifies any anomalies like outliers. The Gaussian fit should also be checked with a normalcy test, like Kolmogorov-Smirnov, which is done in our homework.
Is anything actually going to fit? ...

Sure, maybe with some bias, but not bad. here is DIA July 2010-2011, daily data.
How “normal” can we expect our converted data to be?
Evaluation by histogram: SPY

Example:

- SPY
- Daily data
- 2/4/2011 start
- 2/3/2012 end
- n = 252
- DCGR = 0.00018
- \( \sigma = 0.01456 \)

(This is the HW)
How “normal” can we expect our converted data to be?

Evaluation by histogram

Example:

- GLD
- Daily data
- 11/24/2010 start
- 11/22/2011 end
- n = 252
- DCGR = 0.00083
- \( \sigma = 0.01239 \)

... again, Anderson-Darling or Kolmogorov-Smirnov et.al. (easy in MatLab) would actually be used to evaluate this.
How “normal” can we expect our converted data to be?
Evaluation by histogram

Example:

- DNDN
- Weekly data
- 8/15/05 start
- 8/6/07 end
- n = 104
- MWGR = 0.0024
- σ = 0.17315

Getting rid of the outliers means getting rid of this.
Time-adjusting volatility
(converting daily to annual and so-forth)

1. \( V_t = tV_d \)

2. \( \sigma_t = \sqrt{t} \sigma_d \)

3. \( \sigma_a = \sqrt{365} \sigma_d \approx 19.1 \sigma_d \)

Note: When you convert one to the other, though, you may be introducing bias (error). This is why I like to use and stay with daily volatility measures.

Because the variance of any interval that is \( t \) times longer than a shorter interval is simply equal to \( t \) times the variance of the shorter interval, the proportion of the standard deviations is equal to the square root of \( t \). For example, annual volatility is equal to about 19 times daily volatility.
Introduction to “tail risk”

This segment from an earlier DIA histogram shows a typical fat tail bias, where the historical frequency of observations outside of 2 or even 3 standard deviations is a huge multiple above the expected statistical frequency for that range of observations.

What is the operational implications of this?? The implied “safety” of a calculated improbable event is not really there. The probability of a “six-sigma event” is much higher than six sigma.

[Discussion of limitation on Gaussian approaches].
Important to remember

- Why are we using continuous growth rates (converted) when using raw price data?
  - because the raw data will never have a normal distribution or anything close to it (normally it just trends with a fractal pattern) so we can't use any statistical procedures that requires the normal distribution assumption, nor can we do any clever math work that requires the same assumption
- But sometimes the data that we are using are already in rates (such as bond and note yields) so may not need to be converted.
- At some point in advanced empirical work, you have to use some kind of normalcy test (such as Anderson-Darling) to see if your data are normally distributed.
About drift and volatility

We are going to regard the path of securities (and their derivatives) prices as a Markov Process with actual price behavior over time reflecting drift and volatility, where the latter is represented by a Gaussian distribution. The resulting pattern will reflect randomness with a trend. We are searching for mathematical processes that have this feature.

Drift rate (alpha)  volatility (beta)

Hull refers to the “variance rate.”
Conclusion

We began this by searching for useful measures of volatility and risk when using historical data for financial variables.

We know that there are complications of bias when using arithmetic estimators of means and dispersion when using discrete growth rates.

We concluded that standard deviations of the continuous growth rates (the natural logs) of time-series variables provide reasonable starting estimates of volatility and risk.

With these starting points, we know that we can tweak these to possibly better measures of risk (like the volatility formula), but even more than that.
Conversion to log growth and calculating average log growth rates and SDs, evaluating distributions (and setting the foundation for more complex problems).