Fundamental Analysis of Securities Trading
(IV) Pairs Trading B

Kuan-Lun WANG\textsuperscript{1,2}

National Taiwan University

Preliminary Draft: February 13, 2020

\textsuperscript{1}E-mail: polyphonicared@gmail.com
\textsuperscript{2}Home Page: https://www.csie.ntu.edu.tw/~d06922002/

---

Short Biodata

- Research interests:
  - time series models.
  - simulation modeling.
  - portfolio choice.

- Central themes of my application:
  - multivariate pairs trading in real time.
  - assets searching with a long-run equilibrium.
  - riskless portfolio building.

- Current work:
  - cointegration test.
  - structural change analysis.
  - the probability estimation of mean reversion.
How can we measure the principal level of the factor?

\[
\begin{align*}
\text{price}_A(t) &= a_A + b_A \text{factor}(t) + \epsilon_A(t); \\
\text{price}_B(t) &= a_B + b_B \text{factor}(t) + \epsilon_B(t).
\end{align*}
\]

Clearly, we have

\[
\text{factor}(t) = -\frac{a_B - \text{price}_B(t) + \epsilon_B(t)}{b_B}
\]

So, we have

\[
\begin{align*}
\text{price}_A(t) &= a_A + b_A \text{factor}(t) + \epsilon_A(t) \\
&= a_A + b_A \left( -\frac{a_B - \text{price}_B(t) + \epsilon_B(t)}{b_B} \right) + \epsilon_A(t) \\
&= \frac{a_A b_B - a_B b_A}{b_B} + \frac{b_A}{b_B} \text{price}_B(t) + \frac{b_B \epsilon_A(t) - b_A \epsilon_B(t)}{b_B}
\end{align*}
\]
The following form without factor(t)

\[ \text{price}_A(t) = \frac{a_A b_B - a_B b_A}{b_B} + \frac{b_A}{b_B} \text{price}_B(t) + \frac{b_B \epsilon_A(t) - b_A \epsilon_B(t)}{b_B}. \]

Moreover, we can estimate \( b_A/b_B \) without factor(t).

**Remark: Our Portfolio**

We buy (or short) \( b_B A \) and short (or buy) \( b_A B \).

- Function relation: \( y = f(x) \). i.e., \( x \implies y \)
- Statistical relation: \( (y, x) \). i.e., \( x \implies y \)
  - \( x \) maybe imply \( y \).
  - \( y = f(x) + \epsilon \), where \( \epsilon \) is a small error.
  - if we known more information of \( x \) and \( y \), then we guess there is a function relation.
Pearson Correlation Coefficient [1]

If the standard deviations $\sigma_1$ and $\sigma_2$ are positive, then

$$\rho = \frac{\text{Cov}(X_1, X_2)}{\sigma_1 \sigma_2}$$

is called the Pearson correlation coefficient (PCC) of $X_1$ and $X_2$.

PCC measures the variance of $\epsilon$ in the following regression:

$$y = a + bx + \epsilon,$$

where $\epsilon \sim (0, \sigma^2_\epsilon)$. (Why?)

Figure: Examples of scatter diagrams with different values of correlation coefficient $\rho$
**Measurement (4/10)**

**Definition: Expectation [1]**

Let $X$ be a random variable. If $X$ is a continuous random variable with p.d.f. $f(x)$ and

$$\int_{-\infty}^{\infty} |x| f(x) \, dx < \infty,$$

then the expectation of $X$ is

$$E(X) = \int_{-\infty}^{\infty} x f(x) \, dx.$$

**Definition: Co-Variance [1]**

$$\text{Cov}(X, Y) = E((X - E(X))(Y - E(Y))).$$

---

**Measurement (5/10)**

**Definition: Expectation [1] –Continued**

If $X$ is a discrete random variable with p.m.f. $p(x)$ and

$$\sum_{x} |x| p(x) < \infty,$$

then the expectation of $X$ is

$$E(x) = \sum_{x} xp(x).$$
Assume the following variable pairs are linearly independent

- \( \text{factor}(t) \) and \( \epsilon_A(t) \),
- \( \text{factor}(t) \) and \( \epsilon_B(t) \),
- \( \epsilon_A(t) \) and \( \epsilon_B(t) \),

We analyze all terms in PCC.

\[
\rho = \frac{\text{Cov}(X_1, X_2)}{\sigma_1 \sigma_2} = \frac{\text{Cov}(X_1, X_2)}{\sqrt{\text{Var}[X_1] \text{Var}[X_2]}}.
\]

Remark: PCC

The form of PCC is

\[
\rho = \frac{\text{Cov}(X_1, X_2)}{\sigma_1 \sigma_2} = \frac{\text{Cov}(X_1, X_2)}{\sqrt{\text{Var}[X_1] \text{Var}[X_2]}}.
\]

Rule: Variance

Let \( X, Y \) and \( Z \) be variables, and let \( a, b \) and \( c \) be two constants. Then

\[
\text{Var}[aX + bY + c] = a^2 \text{Var}[X] + b^2 \text{Var}[Y];
\]

By the definition of \( \text{price}_A(t) \), we have

\[
\text{Var}[\text{price}_A(t)] = \text{Var}[aA + b(A\text{factor}(t)) + \epsilon_A] \\
= b^2 \text{Var}[\text{factor}(t)] + \text{Var}[\epsilon_A(t)].
\]

Similarly, we have

\[
\text{Var}[\text{price}_B(t)] = b^2 \text{Var}[\text{factor}(t)] + \text{Var}[\epsilon_B(t)].
\]
**Measurement (8/10)**

**Rule: Co-Variance**

Let $X$, $Y$ and $Z$ be variables, and let $a$, $b$ and $c$ be two constants. Then

$$\text{Cov}[aX + bY + c, Z] = a \text{Cov}[X, Z] + b \text{Cov}[Y, Z].$$

By the definitions of $\text{price}_A(t)$ and $\text{price}_B(t)$ and the linearly independent assumption, we have

\[
\text{Cov}[\text{price}_A(t), \text{price}_B(t)] = \text{Cov}[a_A + b_A \text{factor}(t) + \epsilon_A(t), \text{price}_B(t)]
= b_A \text{Cov}[\text{factor}(t), \text{price}_B(t)] + \text{Cov}[\epsilon_A(t), \text{price}_B(t)]
= b_A \text{Cov}[\text{factor}(t), \text{price}_B(t)]
= b_A \text{Cov}[\text{factor}(t), a_B + b_B \text{factor}(t) + \epsilon_B(t)]
= b_A (b_B \text{Cov}[\text{factor}(t), \text{factor}(t)] + \text{Cov}[\text{factor}(t), \epsilon_B(t)])
= b_A b_B \text{Var}[\text{factor}(t)].
\]

**Measurement (9/10)**

Then we have

\[
|\rho| = \frac{\text{Cov}[\text{price}_A, \text{price}_B]}{\sqrt{\text{Var}[\text{price}_A] \text{Var}[\text{price}_B]}}
= \sqrt{\frac{b_A^2 \text{Var}[\text{factor}(t)]}{b_A^2 \text{Var}[\text{factor}(t)] + \text{Var}[\epsilon_A]}} \sqrt{\frac{b_B^2 \text{Var}[\text{factor}(t)]}{b_B^2 \text{Var}[\text{factor}(t)] + \text{Var}[\epsilon_B]}}
\]
Clearly, we have the following inequality
\[
0 \leq \frac{b_A^2 \text{Var}[\text{factor}(t)]}{b_A^2 \text{Var}[\text{factor}(t)] + \text{Var}[\epsilon_A]} \cdot \sqrt{\frac{b_B^2 \text{Var}[\text{factor}(t)]}{b_B^2 \text{Var}[\text{factor}(t)] + \text{Var}[\epsilon_B]}} \leq 1
\]

Therefore, factor(t) is a high principal level factor if \(|\rho| \to 1\).

- Number of security: 937 (at 09.18.18)
- Number of combination: \(\binom{937}{2} = 438,516\)
  - It’s so large, we only need top-rank. (e.g. 0.01)
  - We can use any KPI!!
Searching (2/10)

Listing 1: example

<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
</tr>
</tbody>
</table>

Bernoulli Trial

In the theory of probability and statistics, a Bernoulli trial (or binomial trial) is a random experiment with exactly two possible outcomes, “success” and “failure”, in which the probability of success is the same every time the experiment is conducted. It is named after Jacob Bernoulli, a 17th-century Swiss mathematician, who analyzed them in his Ars Conjectandi (1713).

https://en.wikipedia.org/wiki/Bernoulli_trial
Example: Bernoulli Trial [2]

Out of millions of instant lottery tickets, suppose that 20% are winners. If five such tickets are purchased, then \((0, 0, 0, 1, 0)\) is a possible observed sequence in which the fourth ticket is a winner and the other four are losers. Assuming independence among winning and losing tickets, we observe that the probability of this outcome is

\[
(0.8)(0.8)(0.8)(0.2)(0.8) = (0.2)(0.8)^4.
\]

Listing 2: example

<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>All Combination 438,516</td>
</tr>
<tr>
<td>-----------------------------------</td>
</tr>
<tr>
<td>Top-Rank</td>
</tr>
<tr>
<td>0.01=4385.16=&gt;4385</td>
</tr>
<tr>
<td>Some in Top-Rank</td>
</tr>
<tr>
<td>0.05=219.25=219</td>
</tr>
<tr>
<td>Our Goal</td>
</tr>
<tr>
<td>How about 10?</td>
</tr>
<tr>
<td>-----------------------------------</td>
</tr>
<tr>
<td>How many samples do we need?</td>
</tr>
<tr>
<td>-----------------------------------</td>
</tr>
</tbody>
</table>
Negative Binomial Distribution

In probability theory and statistics, the negative binomial distribution is a discrete probability distribution of the number of successes in a sequence of independent and identically distributed Bernoulli trials before a specified (non-random) number of failures (denoted \( r \)) occurs. For example, if we define a 1 as failure, all non-1s as successes, and we throw a dice repeatedly until 1 appears the third time (\( r = \) three failures), then the probability distribution of the number of non-1s that appeared will be a negative binomial distribution.


Negative Binomial Distribution [2]

We observe a sequence of Bernoulli trials until exactly \( r \) successes occur, where \( r \) is a fixed positive integer. Let the random variable \( X \) denote the number of trials needed to observe the \( r \)th success. The p.m.f. of \( X \) is

\[
\binom{x-1}{r-1} p^r (1-p)^{x-r}.
\]

We say that \( X \) has a negative binomial distribution.
Example: Bernoulli Trial–Continued

Let $X$ denote the number of tickets needed to win one game. The some c.d.f value of $X$ is

\[
F(1) = 0.2^1 = 0.2 \\
F(2) = F(1) + (0.8)(0.2) = 0.36 \\
F(3) = F(2) + (0.8)^2(0.2) = 0.488 \\
F(4) = F(3) + (0.8)^3(0.2) = 0.5904 \\
\vdots \quad \vdots \\
F(13) \approx 0.9450 \\
F(14) \approx 0.9560
\]

Example: Bernoulli Trial–Continued

Let $X$ denote the number of tickets needed to win two games. The same c.d.f value of $X$ is

\[
F(2) = 0.2^2 = 0.04 \\
F(3) = F(2) + \binom{2}{1}(0.8)(0.2)^2 = 0.104 \\
F(4) = F(3) + \binom{3}{1}(0.8)^2(0.2)^2 \approx 0.1808 \\
F(5) = F(4) + \binom{4}{1}(0.8)^3(0.2)^2 \approx 0.2627 \\
\vdots \quad \vdots \\
F(21) \approx 0.9424 \\
F(22) \approx 0.9520
\]
Warning

For example, we need to classify the stock set by the Industry Classification Benchmark (ICB) and then choose pairs discretionarily from each classification of industry to confirm correlation.

http://dx.doi.org/10.6842%2fNCTU.2015.00722

Program

1. Search some pairs with high PCC in Top rank;
   - $|\rho|$ vs. $\rho$
2. Estimate the pairs share rate;
   - How many samples do we need?
3. Use a technical indicator.
Programming Type

<table>
<thead>
<tr>
<th></th>
<th>Top-Down</th>
<th>Bottom-Up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>Scale</td>
<td>Large</td>
<td>Small</td>
</tr>
<tr>
<td>Small-Scale</td>
<td>Slow</td>
<td>Fast</td>
</tr>
<tr>
<td>Large-Scale</td>
<td>Fast</td>
<td>Slow</td>
</tr>
<tr>
<td>Dependent</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>Testing</td>
<td>Complication</td>
<td>Simple</td>
</tr>
</tbody>
</table>

Table: Programming Type

Writing Process:

1. Estimate the pairs share rate;
2. Use a technical indicator;
3. Search some pairs has high PCC in Top rank;

We use a bottom-up process to realize the top-down programming type.