

# Об эллиптической модели для сети рынка.

Петр Колданов

НИУ Высшая Школа Экономики

10 марта 2017 г.

## 1 Введение

Модели многомерного распределения доходностей акций используются при решении многих экономических задач. В частности, модели многомерного распределения необходимы для построения портфеля инвестиций и управления риском. Одной из наиболее популярных многомерных моделей является эллиптическая модель распределения [1]. Недавно в работе [2] изучена согласованность реальных данных с эллиптической моделью распределения и показано, что совместное распределение реальных доходностей акций фондового рынка не соответствует эллиптической модели. Этот результат получен с помощью изучения мер зависимостей между парами акций. Однако, как отмечается авторами этой работы, их методика проверки отличается от статистической методологии проверки гипотез. В настоящей работе мы дополняем проведенный анализ методологией множественной проверки гипотез. В случае большого фондового рынка число пар акций рынка велико и необходимо принимать во внимание так называемый эффект множественности [3]. Такой эффект заключается в том, что при увеличении числа проверяемых гипотез вероятность хотя бы одного ложного отвержения стремится к 1 даже при истинности всех проверяемых гипотез. Процедуры множественного сравнения корректируют статистические выводы в соответствии с эффектом множественности и приводят к получению более корректных решений. В настоящей

работе мы анализируем соответствие реальных данных эллиптической модели с точки зрения множественной проверки гипотез [4]. Наш подход основан на проверке свойства симметрии знаков эллиптической модели для двумерных распределений каждой пары доходностей акций. Основная цель заключается в обнаружении пар акций, для которых гипотеза симметрии знаков отвергается и изучении соответствующего графа отвержений. Для обнаружения таких пар акций мы предлагаем использовать процедуру множественной проверки гипотез, предложенную Холмом [5]. В нашем подходе уровень значимости множественного теста задается вероятностью хотя бы одной ошибки 1-го рода тестами проверки индивидуальных гипотез, известного как FWER. Известно, что процедура Холма гарантирует строгий контроль FWER. Таким образом, в настоящей работе мы развиваем подход к проверке одного из свойств многомерного эллиптического распределения и иллюстрируем его практическим примером.

Статья организована следующим образом: в разделе 2 приведена постановка задачи и кратко описана процедура множественной проверки, в разделе 3 приведены практические результаты проверки гипотезы симметрии на реальных данных рынка США и Великобритании.

## 2 Статистическая процедура проверки симметрии вероятностной модели

Рассмотрим вектор случайных величин

$$X = \begin{pmatrix} X_1 \\ X_2 \\ \dots \\ X_N \end{pmatrix}$$

с произвольным многомерным распределением  $F_{X_1, \dots, X_N}(x_1, \dots, x_N)$ .

Задача проверки гипотезы симметрии распределения случайного вектора  $X$  формулируется следующим образом: требуется построить правило выбора одной

из гипотез

$$\begin{aligned}
H_{1,\dots,N} &: p_{1,1}^{i,j} = p_{0,0}^{i,j}; \quad p_{1,0}^{i,j} = p_{0,1}^{i,j}; \quad \forall i, j = 1, \dots, N \\
H_{1,\dots,N-1} &: p_{1,1}^{i,j} = p_{0,0}^{i,j}; \quad p_{1,0}^{i,j} = p_{0,1}^{i,j}; \quad \forall i, j = 1, \dots, N-1 \\
&\dots \\
H_{i_1,\dots,i_K} &: p_{1,1}^{i,j} = p_{0,0}^{i,j}; \quad p_{1,0}^{i,j} = p_{0,1}^{i,j}; \quad \forall i, j \in \{i_1, \dots, i_K\} \\
&\dots
\end{aligned} \tag{1}$$

где

$$\begin{aligned}
p_{1,1}^{i,j} &= P[(X_i - E(X_i)) > 0, (X_j - E(X_j)) > 0]; \\
p_{0,0}^{i,j} &= P[(X_i - E(X_i)) < 0, (X_j - E(X_j)) < 0]; \\
p_{1,0}^{i,j} &= P[(X_i - E(X_i)) > 0, (X_j - E(X_j)) < 0]; \\
p_{0,1}^{i,j} &= P[(X_i - E(X_i)) < 0, (X_j - E(X_j)) > 0]
\end{aligned} \tag{2}$$

При этом гипотеза  $H_{i_1,\dots,i_K}$  означает, что распределение случайных величин  $X_i, X_j : i, j \in \{i_1, \dots, i_K\}$  является симметричным.

Известно, что функция плотности эллиптического распределения обладает следующим свойством симметрии:

$$f(-(x - \mu); \mu, \Lambda) = f((x - \mu); \mu, \Lambda)$$

т.е. для эллиптических распределений гипотеза  $H_{1,\dots,N}$  из (1) истинна.

Поэтому отвержение гипотез  $H_I : \{i_1, i_2, \dots, i_K\} \subset I$  и не отвержение гипотезы  $H_{i_1,\dots,i_K}$  из (1) означает, что только случайные величины  $X_i : i \in \{i_1, \dots, i_K\}$  могут иметь эллиптическое распределение.

Для изучения симметрии распределения случайного вектора  $X$  будут использоваться индивидуальные гипотезы о симметрии знаков соответствующей пары случайных величин  $X_i, X_j$ .

$$\begin{aligned}
h_1^{i,j} : p_{1,1}^{i,j} = p_{0,0}^{i,j} \text{ против } k_1^{i,j} : p_{1,1}^{i,j} \neq p_{0,0}^{i,j}; \quad i, j = 1, \dots, N; \quad i \neq j \\
h_2^{i,j} : p_{1,0}^{i,j} = p_{0,1}^{i,j} \text{ против } k_2^{i,j} : p_{1,0}^{i,j} \neq p_{0,1}^{i,j}; \quad i, j = 1, \dots, N; \quad i \neq j
\end{aligned} \tag{3}$$

**Статистические тесты проверки индивидуальных гипотез** Для построения статистического теста проверки индивидуальных гипотез (3) рассмотрим ста-

ТИСТИКИ:

$$T_{1,1}^{i,j} = \sum_{t=0}^n T_{1,1}^{i,j}(t), \quad T_{0,0}^{i,j} = \sum_{t=0}^n T_{0,0}^{i,j}(t), \quad (4)$$

$$T_{1,0}^{i,j} = \sum_{t=0}^n T_{1,0}^{i,j}(t), \quad T_{0,1}^{i,j} = \sum_{t=0}^n T_{0,1}^{i,j}(t)$$

где

$$T_{1,1}^{i,j}(t) = \begin{cases} 1, & X_i(t) \geq 0, X_i(t) \geq 0 \\ 0, & \text{иначе} \end{cases} \quad T_{0,0}^{i,j}(t) = \begin{cases} 1, & X_i(t) < 0, X_i(t) < 0 \\ 0, & \text{иначе} \end{cases}$$

$$T_{1,0}^{i,j}(t) = \begin{cases} 1, & X_i(t) \geq 0, X_i(t) < 0 \\ 0, & \text{иначе} \end{cases} \quad T_{0,1}^{i,j}(t) = \begin{cases} 1, & X_i(t) < 0, X_i(t) \geq 0 \\ 0, & \text{иначе} \end{cases}$$

В последующем изложении мы опустим индексы  $i, j$  в (4) для упрощения обозначений. Совместное распределение статистик (4) можно записать в виде:

$$\begin{aligned} & P(T_{1,1} = k_1, T_{0,0} = k_2, T_{1,0} = k_3, T_{0,1} = k_4) = \\ & = C \exp\left\{k_1 \ln \frac{p_{1,1}}{p_{0,0}} + (k_1 + k_2) \ln \frac{p_{0,0}}{p_{0,1}} + k_3 \ln \frac{p_{1,0}}{p_{0,1}}\right\} \end{aligned}$$

где

$$C = \frac{n!}{k_1!k_2!k_3!k_4!} (1 - p_{1,1} - p_{0,0} - p_{1,0})^n$$

Тогда равномерно наиболее мощный несмещенный тест проверки гипотез  $h_{i,j}^1$  является тестом структуры Неймана и имеет вид [4]:

$$\varphi_{i,j}^1 = \begin{cases} 0, & C_1(k, k_3) < k_1 < C_2(k, k_3) \\ 1, & \text{else} \end{cases} \quad (5)$$

где  $k_1, k_2, k_3, k_4$  - наблюдаемые значения статистик  $T_{1,1}^{i,j}, T_{0,0}^{i,j}, T_{1,0}^{i,j}, T_{0,1}^{i,j}$ ,  $k = k_1 + k_2$ . Константы  $C_1, C_2$  определяются из условного распределения статистики  $T_{1,1}$  при условиях  $T_{1,1} + T_{0,0} = k$ ,  $T_{1,0} = k_3$  и предположении, что гипотеза  $h_{i,j}^1$  истинна.

Условное распределение статистики  $T_{1,1}$  является биномиальным распределением, не зависящим от  $T_{1,0}$ . При условии, что гипотеза  $h_{i,j}^1$  истинна, выполняется  $p_{1,1} = p_{0,0}$ . Поэтому условное распределение статистики  $T_{1,1}$  при условиях  $T_{1,1} + T_{0,0} = k$ ,  $T_{1,0} = k_3$  и предположении, что гипотеза  $h_{i,j}^1$  истинна, является

биномиальным с параметрами  $k, \frac{1}{2}$ . Поэтому равномерно наиболее мощный несмещенный тест (5) имеет вид:

$$\varphi_{i,j}^1 = \begin{cases} 0, & C_1(k) < k_1 < C_2(k) \\ 1, & \text{else} \end{cases} \quad (6)$$

где  $k = k_1 + k_2$ ,  $C_1(k)$  и  $C_2(k)$  определяются из

$$C_1(k) = \max\left\{C : \left(\frac{1}{2}\right)^k \sum_{i=0}^C C_k^i \leq \frac{\alpha}{2}\right\}$$

$$C_2(k) = \min\left\{C : \left(\frac{1}{2}\right)^k \sum_{i=C}^k C_k^i \leq \frac{\alpha}{2}\right\}$$

р-значение теста (6) вычисляется из

$$p_{i,j}^1 = 2 \min\left\{\left(\frac{1}{2}\right)^k \sum_{i=k_1}^k C_k^i, \left(\frac{1}{2}\right)^k \sum_{i=0}^{k_1} C_k^i\right\} \quad (7)$$

Таковыми же рассуждениями можно построить РНМ тест проверки гипотезы  $h_{i,j}^2$ . Этот тест имеет вид:

$$\varphi_{i,j}^2 = \begin{cases} 0, & C_1(m) < k_3 < C_2(m) \\ 1, & \text{else} \end{cases} \quad (8)$$

где  $m = k_3 + k_4$ ,  $C_1(m)$  и  $C_2(m)$  определяется из (6) р-значение теста (8) вычисляется из

$$p_{i,j}^2 = 2 \min\left\{\left(\frac{1}{2}\right)^m \sum_{i=k_3}^m C_m^i, \left(\frac{1}{2}\right)^m \sum_{i=0}^{k_3} C_m^i\right\} \quad (9)$$

Заметим, что по построению все индивидуальные тесты являются свободными от распределения равномерно наиболее мощными тестами структуры Неймана.

### 3 Проверка гипотезы симметрии совместного распределения доходностей акций по реальным данным

Результаты применения процедуры Холма с индивидуальными тестами (6), (8) удобно представить с помощью графа отвержений: вершинами графа являются случайные величины  $X_i : i = 1, \dots, L$ , ребро  $(i, j)$  включается в граф отвержений тогда и только тогда, когда гипотеза  $h_{i,j}$  отвергается процедурой Холма.

Ниже приведены экспериментальные результаты применения построенной в разделе 2 процедуры проверки гипотезы эллиптичности распределения доходностей акций по данным рынков США и Великобритании. Для исследования выбраны 100 акций с каждого рынка с наибольшим объемом торгов за период с 01.01.2006 по 31.12.2013. Проведено сравнение результатов для различных периодов наблюдения: 8 периодов по 1 году, 4 периода по 2 года, 2 периода по 4 года и 1 период в 8 лет. Уровни значимости процедуры одновременной проверки многих гипотез равны  $\alpha = 0,1$  и  $\alpha = 0,5$ . Для описания результатов процедуры одновременной проверки многих гипотез использовано понятие *графа отвержений*. Интересно определить множество вершин  $S$  минимального размера, такое что после удаления этих вершин из рассмотрения граф отвержений становится пустым. В теории графов такое множество называется минимальным вершинным покрытием. Будем называть *хабами (hubs)* вершины в  $S$ . Очевидно, что размер  $S$  имеет практический смысл. Результаты приведены в таблицах и рисунках ниже. Каждая таблица содержит следующую информацию: первый столбец описывает год наблюдения, второй столбец содержит число хабов/число ребер в графе отвержений за различные периоды наблюдения. Например столбец 3 содержит число хабов/число ребер в графе отвержений за двухгодовые периоды наблюдения (2006-2007, 2008-2009, 2010-2011, 2012-2013), последний столбец содержит число хабов/число ребер в графе отвержений за 8-летний период наблюдения 2006-2013.

Таблица 1. Рынок США.  $\alpha = 0.1; p_{1,1} = p_{0,0}$

	1 год	2 года	4 года	8 лет
2006	1/1	0/0	0/0	0/0
2007	0/0			
2008	0/0	0/0		
2009	0/0			
2010	0/0	0/0	0/0	
2011	0/0			
2012	0/0	0/0		
2013	0/0			

Таблица 2. Рынок США.  $\alpha = 0.5; p_{1,1} = p_{0,0}$

	1 год	2 года	4 года	8 лет
2006	2/5	1/2	2/3	1/2
2007	0/0			
2008	0/0	0/0		
2009	0/0			
2010	0/0	0/0	0/0	
2011	0/0			
2012	0/0	0/0		
2013	0/0			

В таблицах 1-2 приведены результаты применения процедуры Холма для проверки индивидуальных гипотез  $h_{i,j}^1$  по данным рынка США с уровнями значимости  $\alpha = 0.1$ ,  $\alpha = 0.5$ . Заметим, что при  $\alpha = 0.1$  есть только один непустой граф отвержений, соответствующий 2006 году. Соответствующими хабами являются Dell Inc.(DELL) или Compañia Cervcerias Unidas S.A. (CCU). Для  $\alpha = 0.5$  хабами для различных периодов являются Dell Inc. (DELL) и 3M Company (MMM) в 2006, Dell Inc. (DELL) в 2006-2007, McDonalds Corp. (MCD) и Altria Group Inc. (MO) в 2006-2009, и McDonalds Corp. (MCD) в 2006-2013.

Таблица 3. Рынок США.  $\alpha = 0.1; p_{0,1} = p_{1,0}$

	1 год	2 года	4 года	8 лет
2006	2/24	1/3	0/0	2/5
2007	0/0			
2008	0/0	1/2		
2009	0/0			
2010	0/0	0/0	0/0	
2011	0/0			
2012	0/0	0/0		
2013	0/0			

Таблица 4. Рынок США.  $\alpha = 0.5; p_{0,1} = p_{1,0}$

	1 год	2 года	4 года	8 лет
2006	2/44	1/14	2/4	3/12
2007	1/3			
2008	0/0	1/3		
2009	1/1			
2010	0/0	0/0	2/4	
2011	0/0			
2012	3/10	0/0		
2013	0/0			

В таблицах 3-4 приведены результаты применения процедуры Холма для проверки гипотез  $h_{i,j}^2$  по данным рынка США, с уровнями значимости  $\alpha = 0.1$ ,  $\alpha = 0.5$ . Заметим, что при  $\alpha = 0.1$  есть только 4 непустых графа отвержений, соответствующих периодам наблюдений 2006, 2006-2007, 2008-2009, 2006-2013. Соответствующими хабами являются Dell Inc. (DELL) и 3M Company (MMM) в 2006, Dell Inc. (DELL) в 2006-2007, Altria Group Inc. (MO) в 2008-2009, и McDonalds Corp. (MCD) и Sprint Corporation (S) в 2006-2013. Для  $\alpha = 0.5$  хабами являются Dell Inc. (DELL) и 3M Company (MMM) в 2006, Dell Inc. (DELL) в 2006-2007, McDonalds Group (MCD) и Altria Group Inc. (MO) в 2006-2009, и McDonalds Corp. (MCD), Altria Group Inc. (MO) и Chevron Corporation (CVX) в 2006-2013. На рисунке 1 приведен граф отвержений за 2006 год.

Таблица 5. Рынок Великобритании.  $\alpha = 0.1; p_{1,1} = p_{0,0}$



	1 год	2 года	4 года	8 лет
2006	5/149	4/130	5/181	48/2121
2007	1/47			
2008	0/0	1/1		
2009	0/0			
2010	2/3	1/1	24/819	
2011	0/0			
2012	2/7	16/400		
2013	5/49			

Таблица 6. Рынок Великобритании.  $\alpha = 0.1; p_{0,1} = p_{1,0}$

	1 год	2 года	4 года	8 лет
2006	1/4	1/44	1/43	1/28
2007	2/9			
2008	1/1	0/0		
2009	0/0			
2010	0/0	0/0	0/0	
2011	0/0			
2012	0/0	0/0		
2013	1/1			

В таблицах 5-6 приведены результаты применения процедуры Холма для проверки гипотез  $h_{i,j}^1, h_{i,j}^2$  для рынка Великобритании, с уровнем значимости  $\alpha = 0.1$ . Заметим, что для всех периодов число хабов мало по сравнению с числом ребер графов отвержений. Одна акция, Bankers Ord (BNKR), находится в множестве хабов графов отвержений для обоих типов гипотез в периоды 2006, 2007, 2006-2007, 2006-2009, 2006-2013. Можно заметить, что рынок Великобритании более диверсифицирован, чем рынок США при характеристике диверсификации рынка графом отвержений.

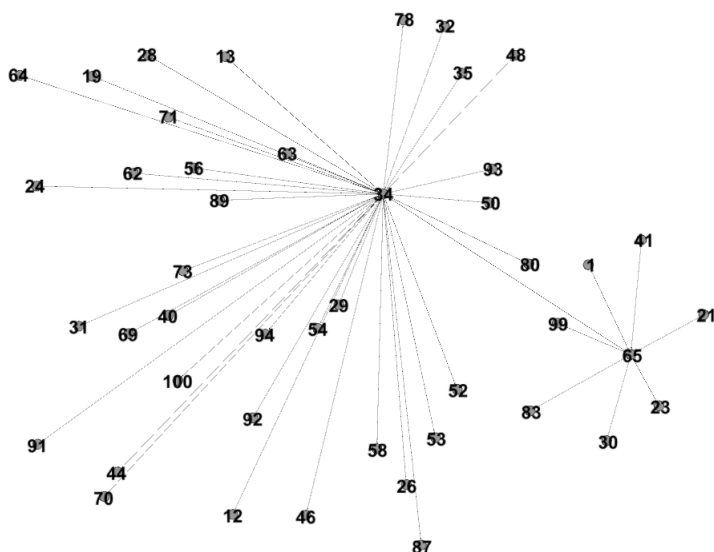


Рис. 1: Граф отвержений для гипотезы  $h_{i,j}^2$ ,  $\alpha = 0.5$ , Фондовый рынок США. Год наблюдения - 2006.

На рисунке показан вид графа отвержений. Общий вывод который можно сформулировать из анализа этого графа, заключается в следующем: при удалении из рассмотрения акций двух компаний (Dell Inc. (DELL) и 3M Company (MMM) с номерами 34, 65) гипотеза симметрии совместного распределения доходностей акций не отвергается процедурой Холма при уровне значимости  $\alpha = 0.5$ .

**Acknowledgments:** Работа поддержана грантом РФНФ № 15-32-01052.

## Список литературы

- [1] F K Gupta, T Varga, and T Bodnar. *Elliptically Contoured Models in Statistics and Portfolio Theory*. Springer, 2013.
- [2] R. Chicheportiche and Bouchaud J-P. The joint distribution of stock returns is not elliptical. *International Journal of Theoretical and Applied Finance*, 15(3), 2012.
- [3] F. Bretz, T. Hothorn, and P. Westfall. *Multiple Comparisons Using R*. Taylor & Francis Group, 2011.

- [4] E L Lehmann and J P Romano. *Testing statistical hypotheses*. Springer, New York., 2005.
- [5] S Holm. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2):65–70, 1979.