# Сравнительный анализ оценок параметров процесса авторегрессии с остатками в виде скользящего среднего

Кузнецова Александра Андреевна, гр. 522

Санкт-Петербургский Государственный Университет
Математико-механический факультет
Кафедра статистического моделирования

Научный руководитель: к.ф.-м.н. Товстик Т.М.
Рецензент: к.ф.-м.н. Сизова А.Ф.

Санкт-Петербург
2008г.

## Основные понятия и определения

- Стохастическое уравнение стационарного процесса АРСС(p,q):

$$\sum_{j=0}^{p} \alpha_j x_{t-j} = \sum_{k=0}^{q} \beta_k \xi_{t-k}, \qquad (1)$$

$$\mathbf{E}\xi_j = 0, \quad \mathbf{E}\xi_k\xi_j = \delta_{kj}\sigma^2, \quad \alpha_0 = 1, \quad \beta_0 = 1.$$

- Если все корни полинома $\alpha(z) = \sum_{k=0}^{p} \alpha_k z^k$ по модулю больше единицы, то процесс обладает свойством стационарности.
- Корни полинома $\beta(z) = \sum_{k=0}^{q} \beta_k z^k$ должны быть по модулю больше единицы(понадобится для обратимости процесса).

# Постановка задачи

Мы предполагаем, что некоторый временной ряд порождается моделью АРСС(p,q), при этом возникает проблема оценок коэффициентов этой модели.

- **Задача:**
  **Провести сравнительный анализ методов, решающих данную проблему. Построить алгоритмы, реализующие эти методы :**

Задача будет решаться в несколько этапов:

- Оценка коэффициентов на основе метода Юла-Уокера(Метод 1).
- Оценка коэффициентов методом, предложенным J.Hannan и J.Rissanen(Метод 2).
- Оценка коэффициентов методом, основанном на методе наименьших квадратов(МНК)(Метод 3).
- Сравнительный анализ результатов, сделанных разными методами(а также методом, предложенным пакетом Statistica).

- Оценка коэффициентов $\alpha_k, \quad k = 1, \ldots, p$. Коэффициенты находятся из уравнения Юла-Уокера:

$$\begin{array}{l} \widehat{R}_{q+1} + \widehat{\alpha}_1 \widehat{R}_q + \ldots + \widehat{\alpha}_p \widehat{R}_{q+1-p} = 0, \\ \widehat{R}_{q+2} + \widehat{\alpha}_1 \widehat{R}_{q+1} + \ldots + \widehat{\alpha}_p \widehat{R}_{q+2-p} = 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \widehat{R}_{q+p} + \widehat{\alpha}_1 \widehat{R}_{q+p-1} + \ldots + \widehat{\alpha}_p \widehat{R}_q = 0, \end{array} \tag{2}$$

где

$$\widehat{R}_k = \frac{1}{N-k} \sum_{i=1}^{N-k} (x_{i+k} - \overline{x})(x_i - \overline{x}), \quad \overline{x} = \frac{1}{N} \sum_{i=1}^{N} x_i,$$

$$\max_{1 \le k \le p} |\alpha_k - \widehat{\alpha}_k| = O\Big(\sqrt{\frac{\log(\log N)}{N}}\Big)$$

## Оценка коэффициентов на основе метода Юла-Уокера(**Метод 1**)

- Оценка коэффициентов $\beta_k, \quad k = 1, \dots, q$

$$\frac{\alpha(z)}{\beta(z)} = \sum_{0}^{\infty} A_k z^k.$$

-разлагается при |z|<1 в сходящийся ряд, где
$\alpha(z) = \sum_{k=0}^{p} \alpha_k z^k, \quad \beta(z) = \sum_{k=0}^{q} \beta_k z^k$

$$\exists M : \sum_{M+1}^{\infty} |A_k|^2 < \varepsilon.$$

Уравнение (1) с точностью до $\varepsilon$ эквивалентно

$$\sum_{k=0}^{M} A_k x_{t-k} = \xi_t. \tag{3}$$

Оценки коэффициентов $\beta_k$ находятся из следующего приближенного равенства

$$\frac{\widehat{\alpha}(z)}{\sum_0^M \widehat{A}_k z^k} \approx \widehat{\beta}(z), \tag{4}$$

где $\widehat{A}_k$ получаются из АР(M) методом Юла-Уокера.

# Оценка коэффициентов методом, предложенным J.Hannan и J.Rissanen(Метод 2)

- Оценки коэффициентов $\beta_k, \quad k = 1, \ldots, q; \quad \alpha_k, \quad k = 1, \ldots, p$:

$$\widetilde{\sigma}^2_{p,q} = \min_{\beta_k, \alpha_j} \{ \frac{1}{N} \sum_{t=m}^{N} (\sum_{j=0}^{p} \alpha_j x_{t-j} - \sum_{k=1}^{q} \beta_k \widehat{\xi}_{t-k})^2 \}, \qquad (5)$$

где

$$\widehat{\xi}_t = \sum_{j=0}^{M} \widehat{A}_j x_{t-j},$$

$\widehat{A}_j$ находятся из АР(М) методом Юла-Уокера.
$m = max(M + p + 1, M + q + 1)$. Минимизируя оценку дисперсии $\widetilde{\sigma}^2_{p,q}$, находятся $\widehat{\beta}_k \;\; k = 1, \ldots, q, \quad \widehat{\alpha}_k, \quad k = 1, \ldots, p$.

## Оценка коэффициентов методом, основанном на методе наименьших квадратов(МНК)(**Метод 3**)

- Оценка коэффициентов $\alpha_k, \quad k = 1, \ldots, q$.
  В уравнении (2) Метода 1 по методу наименьших квадратов находятся оценки $\widehat{\alpha}_j$. Именно $\widehat{\alpha}_j$ обращает в минимум функцию:

$$\Phi(\widehat{\alpha}_j) = \sum_{k=q+1}^{\nu} (\sum_{j=0}^{p} \widehat{\alpha}_j \widehat{R}_{k-j})^2, \tag{6}$$

где $\nu > p + q$. Вычисление минимума приводит к системе:

$$\sum_{j=0}^{p} \widehat{\alpha}_j \sum_{k=q+1}^{\nu} \widehat{R}_{k-j} \widehat{R}_{k-t} = 0, \;\; t = 1, \ldots, p.$$

# Оценка коэффициентов методом, основанным на методе наименьших квадратов(МНК)(Метод 3)

- Оценка коэффициентов $\beta_k, \quad k = 1, \ldots, q$. Из Метода 1 получаем систему уравнений:

$$\begin{cases} \widehat{\alpha}_1 = \widehat{\beta}_1 + \widehat{A}_1, \\ \widehat{\alpha}_2 = \widehat{\beta}_1 \widehat{A}_1 + \widehat{\beta}_2 \widehat{A}_1 + \widehat{A}_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \widehat{\alpha}_q = \widehat{\beta}_1 \widehat{A}_{q-1} + \widehat{\beta}_2 \widehat{A}_{q-2} + \ldots + \widehat{\beta}_q + \widehat{A}_q, \\ \widehat{\alpha}_{q+1} = \widehat{\beta}_1 \widehat{A}_q + \widehat{\beta}_2 \widehat{A}_{q-1} + \ldots + \widehat{\beta}_q \widehat{A}_1 + \widehat{A}_{q+1}, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \widehat{\alpha}_M = \widehat{\beta}_1 \widehat{A}_{M-1} + \widehat{\beta}_2 \widehat{A}_{M-2} + \ldots + \widehat{\beta}_q \widehat{A}_{M-q} + \widehat{A}_M, \end{cases} \tag{7}$$

где $\widehat{\alpha}_{p+1} = 0, \ldots, \widehat{\alpha}_M = 0, \quad \widehat{A}_{-k} = 0 \quad k = 1, \ldots, q$.
Методом наименьших квадратов получим:

$$\sum_{k=1}^{M} (\widehat{\alpha}_k - \widehat{A}_k - \sum_{j=1}^{q} \widehat{A}_{k-j} \widehat{\beta}_j)^2 \to \min_{\widehat{\beta}_j},$$

## Сравнительный анализ результатов, сделанных разными методами(а также методом предложенным пакетом Statistica)

Критерии методов следующие:

- Первый критерий (предполагает известность истиных автокорреляций модели):

$$\sigma_1 = \Big(\frac{1}{\nu+1}\sum_{j=0}^{\nu}(\rho_j - \widetilde{\rho}_j)^2\Big)\Big)^{\frac{1}{2}}, \tag{8}$$

  где $\nu \geq p+q$ ,$\rho_j = \frac{R(j)}{R(0)}$ - истиные коэффициенты автокорреляции, а $\widetilde{\rho}_j = \frac{\widetilde{R}(j)}{\widetilde{R}(0)}$ - автокорреляции модели, полученной одним из методов.
- Второй критерий (на основе выборочных автокорреляций):

$$\sigma_2 = \Big(\frac{1}{\nu+1}\sum_{j=0}^{\nu}(\widehat{\rho}_j - \widetilde{\rho}_j)^2\Big)\Big)^{\frac{1}{2}}, \tag{9}$$

  где $\widehat{\rho}_j = \frac{\widehat{R}(j)}{\widehat{R}(0)}$-выборочные коэффициенты автокорреляций
- Третий критерий ( из статьи J.Hannan и J.Rissanen):

$$\sigma_3 = \frac{1}{N}\sum_{t=m}^{N}\{\sum_{j=0}^{p}\widehat{\alpha}_j x_{t-j} - \sum_{j=1}^{q}\widehat{\beta}_j\widehat{\xi}_{t-j}\}^2. \tag{10}$$

Истиная модель при **p=3, q=2**, объем выборки $N = 180$

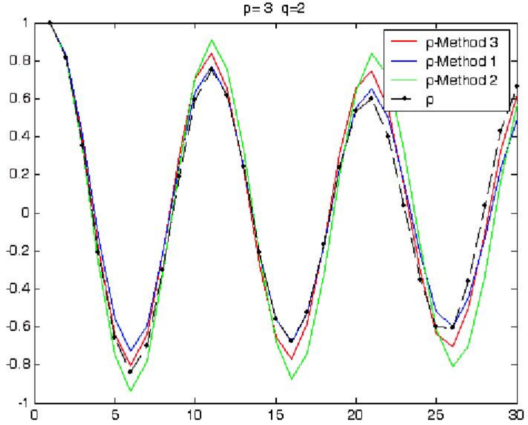



Рис.: График автокорреляций модели АРСС(3,2)

## Сравнительный анализ полученных результатов, сделанных разными методами(а также методом, предложенным пакетом Statistica)

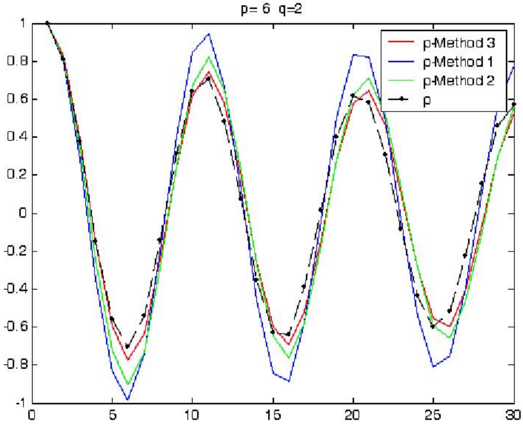



Рис.: График автокорреляций модели АРСС(6,2)

# Сравнительный анализ полученных результатов сделанных разными методами(а также методом предложенным пакетом Statistica)

Выделенные данные соответствуют выбранной модели.
По первому критерию $\sigma_1$:

| **p=3, q=2** | | | |
|---|---|---|---|
| | $p = 3, \ q = 2$ | $p = 6, \ q = 2$ | $p = 2, \ q = 2$ |
| Метод1(Юла-Уокера) | 0.1083 | 0.0787 | 0.0891 |
| Метод2(J.Hannan и J.Rissanen) | 0.0904 | 0.0699 | 0.0717 |
| Метод3(МНК) | 0.0383 | 0.0765 | 0.0940 |

По второму критерию $\sigma_2$:

| **p=3, q=2** | | | |
|---|---|---|---|
| | $p = 3, \ q = 2$ | $p = 6, \ q = 2$ | $p = 2, \ q = 2$ |
| Метод1(Юла-Уокера) | 0.1258 | 0.1012 | 0.0957 |
| Метод2(J.Hannan и J.Rissanen) | 0.1131 | 0.1100 | 0.0932 |
| Метод3(МНК) | 0.0579 | 0.0710 | 0.0655 |

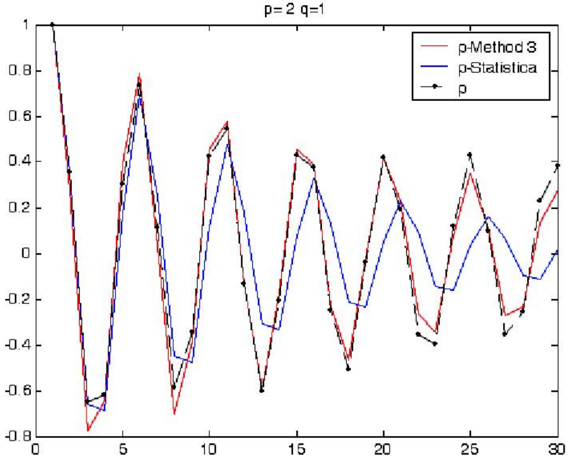
## Сравнительный анализ полученных результатов сделанных разными методами(а также методом предложенным пакетом Statistica)

По третьему критерию $\sigma_3$:

| **p=3, q=2** | | | |
|---|---|---|---|
| | $p = 3,\ q = 2$ | $p = 6,\ q = 2$ | $p = 2,\ q = 2$ |
| Метод1(Юла-Уокера) | 0.1256 | 0.1841 | 0.2298 |
| Метод2(J.Hannan и J.Rissanen) | 0.1180 | 0.1814 | 0.2207 |
| Метод3(МНК) | 0.1252 | 0.1836 | 0.2255 |

# Сравнительный анализ полученных результатов сделанных разными методами(а также методом предложенным пакетом Statistica)

Сравнение с результатами полученными пакетом Statistica



# Сравнительный анализ полученных результатов сделанных разными методами(а также методом предложенным пакетом Statistica)

Сравнение с результатами, полученными пакетом Statistica

Значения $\sigma_1$

| | p=2, q=1 | p=3, q=2 | p=4, q=2 |
|---|---|---|---|
| Метод Statistica | 0.1283 | 0.1848 | 0.3742 |
| Метод 3(МНК) | 0.0634 | 0.0879 | 0.0591 |
| Метод2(J.Hannan и J.Rissanen) | 0.1180 | 0.1014 | 0.1207 |
| Метод1(Юла-Уокера) | 0.1270 | 0.1214 | 0.1301 |

Значения $\sigma_3$

| | p=2, q=1 | p=3, q=2 | p=4, q=2 |
|---|---|---|---|
| Метод Statistica | 0.2381 | 0.2948 | 0.3870 |
| Метод 3(МНК) | 0.1476 | 0.1552 | 0.1235 |
| Метод2(J.Hannan и J.Rissanen) | 0.1321 | 0.1145 | 0.1200 |
| Метод1(Юла-Уокера) | 0.1481 | 0.1567 | 0.1301 |

# Заключение

- Были промоделированы стационарные процессы и оценены коэффициенты модели АРСС(p,q). Предложен новый метод оценок коэффициентов(Метод 3). В большинстве случаев он дает лучший результат по сравнении с другими методами по первому и второму критериям. По третьему критерию Метод 3 уступает Методу 2. Оценки коэффициентов, сделанные пакетом Statistica хуже, чем оценки, предложенные рассматриваемыми методами по трем критериям.
- Оценки коэффициентов $\alpha_k, \;\; k = 1, \ldots, p$, найденные методом наименьших квадратов (Метод 3) гораздо чаще подчиняются условиям стационарности процесса нежели, полученные другими методами, либо статистическим пакетом Statistica(за счет выбора $\nu$ в (6)).