

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Запорізький національний технічний університет

**ПРОГРАМА, МЕТОДИЧНІ ВКАЗІВКИ
ТА КОНТРОЛЬНІ ЗАВДАННЯ**

з дисципліни

**“КОМП’ЮТЕРНА ТЕХНІКА В НАУКОВИХ
ДОСЛІДЖЕННЯХ”**

для студентів напрямку 123 “Комп’ютерна інженерія”
спеціалізацій (освітніх програм)
”Комп’ютерні системи та мережі”,
”Спеціалізовані комп’ютерні системи”
заочної форми навчання

2017 р.

Програма, методичні вказівки та контрольні завдання з дисципліни "Комп'ютерна техніка в наукових дослідженнях" для студентів напряму 123 "Комп'ютерна інженерія" спеціалізацій (освітніх програм) "Комп'ютерні системи та мережі", "Спеціалізовані комп'ютерні системи" заочної форми навчання/ Укл.: К.М. Касьян, М.М. Касьян. - Запоріжжя: ЗНТУ, 2017. - 46 с.

Укладачі: К.М. Касьян, доцент, к.т.н.
М.М. Касьян, доцент, к.т.н.

Рецензент М.Ю. Тягунова, доцент, к.т.н.

Відповідальний
за випуск: К.М. Касьян, доцент, к.т.н.

Затверджено
на засіданні кафедри КСМ
Протокол № 4 від 14.11.2017

Затверджено
на засіданні НМК КНТ
Протокол № 4 від 28.11.2017

ЗАГАЛЬНІ МЕТОДИЧНІ ВКАЗІВКИ

Вивчення дисципліни полягає в засвоєнні студентами основних методів аналізу різного роду експериментальних даних при наявності випадкових та непередбачених впливів, умінні виявляти закономірності на тлі випадків, аналізувати їх та будувати достовірні прогнози. Студенти повинні ознайомитись з основними стадіями генерації експериментальних даних у вигляді часових рядів, їх обробки та аналізу за допомогою статистичних пакетів, а також документування отриманих результатів за допомогою спеціалізованого редактора.

Дисципліна "Комп'ютерна техніка в наукових дослідженнях" вивчається студентами заочної форми навчання в одинадцятому семестрі. Основна форма вивчення дисципліни - самостійна проробка матеріалу за літературою, вказаною по розділах цих методичних вказівок.

Увесь матеріал дисципліни складається з розділів. На початку кожного розділу дається відповідна частина програми дисципліни, потім коротко викладаються основні положення, на які повинна бути звернена увага при вивченні теоретичних питань даного розділу по рекомендованій літературі. Для самоконтролю ступеня засвоєння проробленого матеріалу студент повинен користуватися запитаннями для самоперевірки, що подаються в кожному розділі вказівок.

У процесі роботи над дисципліною студент зобов'язаний виконати одну контрольну роботу та ряд лабораторних робіт. По кожній виконаній лабораторній роботі студент повинен скласти звіт та захистити її перед викладачем.

Успішне засвоєння дисципліни "Комп'ютерна техніка в наукових дослідженнях" можливе тільки при систематичній роботі над підручниками. Матеріал, що вивчається рекомендується конспектувати.

1 РОБОЧА ПРОГРАМА І МЕТОДИЧНІ ВКАЗІВКИ ЩОДО ВИВЧЕННЯ ТЕМ ДИСЦИПЛІНИ

1.1 Вступ. Поняття часового ряду

Поняття часового ряду. Цілі та етапи аналізу часових рядів. Структура часового ряду: детермінована і випадкова складові часового ряду. [1(с. 330-341)].

1.1.1 Методичні вказівки

У повсякденному житті і професійній діяльності, а також у наукових дослідженнях ми постійно зіштовхуємося з подіями і явищами з невизначеним результатом. При цьому постійно приходиться приймати в подібних невизначених, зв'язаних з багатьма випадками ситуаціях свої рішення, іноді дуже важливі. У побуті чи в нескладному бізнесі ми можемо приймати такі рішення на основі здорового глузду, інтуїції, попереднього досвіду. Однак у більш серйозному бізнесі, в умовах жорсткої конкуренції, рішення повинні прийматися на основі ретельного аналізу наявної інформації, бути обґрунтованими і доказовими.

Для рішення задач, зв'язаних з аналізом даних при наявності випадкових і непередбачених впливів, математиками та іншими дослідниками (біологами, психологами, економістами і т.д.) за останні двісті років був вироблений могутній і гнучкий арсенал методів, названих у сукупності математичною статистикою (а також прикладною статистикою чи аналізом даних). Ці методи дозволяють виявляти закономірності на тлі випадків, робити обґрунтовані висновки і прогнози, давати оцінки імовірностей їхнього виконання або невиконання.

Широкому впровадженню методів аналізу даних у 60-х і 70-х роках ХХ століття чимало сприяла поява комп'ютерів, а починаючи з 80-х років ХХ століття - персональних комп'ютерів. Статистичні програмні пакети зробили методи аналізу даних більш доступними і наочними: тепер уже не було потрібно вручну виконувати трудомісткі розрахунки по складних формулах, будувати таблиці і графіки - усю цю чорнову роботу взяв на себе комп'ютер, а людині залишилася головним чином творча робота: постановка задач, вибір методів їхнього рішення та інтерпретація результатів.

Результатом появи могутніх і зручних пакетів для аналізу даних на персональних комп'ютерах стало різке розширення і зміна кола

споживачів методів аналізу даних. Якщо раніш ці методи розглядалися головним чином як інструмент наукових досліджень то, починаючи із середини 1980-х років основними покупцями статистичних пакетів стали вже не наукові, а комерційні організації, а також урядові і медичні установи. Таким чином, методи аналізу даних і статистичні пакети для комп'ютерів стали типовим і загальноновживаним інструментом планових, аналітичних, маркетингових відділів виробничих і торгових корпорацій, банків і страхових компаній, урядових і медичних установ.

Важлива задача — прогнозування майбутнього поведіння деякого часового ряду: зміни курсу долара, цін і попиту на продукцію чи сировину і т.д. Для такого часового ряду за допомогою статистичного пакета програм підбирають деяке аналітичне рівняння - будують регресійну модель. Якщо ми припускаємо, що на цікавлячий нас показник впливають деякі інші фактори, їх теж можна включити в модель, попередньо (за допомогою того ж статистичного пакета) перевіривши істотність (значимість) цього впливу. Потім на основі побудованої моделі можна зробити прогноз і вказати його точність.

У багатьох технологічних процесах необхідно систематично контролювати стан процесу, щоб вчасно втрутитися при відхиленнях його від нормального режиму і запобігти тим самим втрати від випуску неякісної продукції. Для цього використовуються статистичні методи контролю якості, повсюдне і неухильне застосування яких багато в чому визначило разючі успіхи японської промисловості.

Методи статистичного аналізу є універсальними і можуть застосовуватися в самих різних галузях людської діяльності. Скажемо, про рокування курсу долара і прогноз попиту на автомобілі робляться за допомогою тих самих процедур. У переважній більшості випадків усі потрібні користувачам задачі по аналізу даних можуть бути вирішені за допомогою універсальних статистичних пакетів. Зрозуміло, є і такі області людської діяльності, у яких виникають специфічні, що не зустрічаються в інших областях, задачі по аналізу даних, і тому вимагаються специфічні статистичні засоби. Однак таких областей дуже мало.

Часовим рядом називається набір значень $x_1, x_2, \dots, x_n$ досліджуваної величини, зареєстрованих у послідовні моменти часу: $t_1, t_2, \dots, t_n$.

При практичному вивченні часових рядів дослідник на підставі спостереженого відрізка часового ряду (кінцевої довжини) повинен зробити висновки про властивості цього ряду і про вірогідносний механізм, що породжує цей ряд. Частіше за все при вивченні часових рядів ставляться наступні цілі:

- стислий опис характерних особливостей ряду;
- підбір статистичної моделі (моделей), що описує часовий ряд;
- прогноз майбутніх значень на основі минулих спостережень;
- управління процесом, що породжує часовий ряд.

На практиці ці і подібні цілі досяжні далеко не завжди і далеко не повною мірою. Часто цьому перешкоджає недостатній об'єм спостережень (недостатня тривалість); ще частіше — статистична структура часового ряду, що змінюється з часом. Через ці зміни значення минулих спостережень знецінюється, і вони вже не допомагають передбачати майбутнє.

Звичайно при практичному аналізі часових рядів послідовно проходять наступні етапи:

- графічне зображення і опис поведінки часового ряду;
- виділення і видалення закономірних складових часового ряду, залежних від часу: тренда, сезонних і циклічних складових;
- виділення і видалення низько або височастотних складових процесу (фільтрація);
- дослідження випадкової складової часового ряду, що залишилася після видалення наведених вище складових;
- побудова (підбір) математичної моделі для опису випадкової складової і перевірка її адекватності;
- прогнозування майбутнього розвитку процесу, що представлений часовим рядом;
- дослідження взаємодій між різними часовими рядами.

Для вирішення вказаних вище (а також багатьох інших) задач дослідниками запропонована велика кількість різних методів. Найпоширеніші з них:

- *кореляційний аналіз* дозволяє виявити істотні періодичні залежності та їх *лаги* (затримки) усередині одного процесу (автокореляція) або між декількома процесами (кроскореляція);

- *спектральний аналіз* дозволяє знаходити періодичні і квазі-періодичні складові часового ряду;
- *згладжування і фільтрацію* призначено для перетворення часових рядів з метою видалення з них високочастотних або сезонних коливань;
- моделі *авторегресії і ковзного середнього* є особливо корисними для опису і прогнозування процесів, що проявляють однорідні коливання навколо середнього значення;
- *прогнозування* дозволяє на основі підібраної моделі поведінки часового ряду передбачати його значення в майбутньому.

Часовий ряд прийнято описувати за допомогою закономірної не випадкової і випадкової складових.

Невипадкова (детермінована) складова часового ряду являє собою функцію від часу, що обчислюється в дискретні моменти часу.

Випадкова складова – набір випадкових величин, розподіл яких невідомий.

Форми розкладання (декомпозиції) часового ряду на детерміновану і випадкову компоненти можуть розрізнятися.

1.1.2 Запитання для самоперевірки

1.1.2.1 Що таке часовий ряд?

1.1.2.2 Які цілі ставляться при вивченні часових рядів?

1.1.2.3 Що перешкоджає досягненню цілей, що ставляться при вивченні часових рядів на практиці?

1.1.2.4 Які етапи послідовно проходять при практичному аналізі часових рядів?

1.1.2.5 Які методи запропоновані дослідниками для вирішення вказаних вище задач аналізу часових рядів?

1.1.2.6 Що називається детермінованою складовою часового ряду?

1.1.2.7 Що називається випадковою складовою часового ряду?

1.2 Моделі часового ряду, тренда та випадкової складової часового ряду

Адитивна і мультиплікативна моделі часового ряду. Компоненти детермінованої складової часового ряду: тренд, сезонна і циклічна

компоненти. Моделі тренда: лінійна, поліноміальна, логарифмічна, логістична, полігармонійна. Моделі випадкової складової часового ряду: білий шум, процес ковзного середнього, процес авторегресії. [1(с. 341-348), 6(с.4-13)].

1.2.1 Методичні вказівки

Адитивною моделлю часового ряду називається представлення ряду у вигляді суми детермінованої і випадкової компонент, а саме:

$$x_t = d_t + \varepsilon_t \quad \text{при} \quad t = 1, \dots, n.$$

Мультиплікативною моделлю часового ряду називається представлення ряду у вигляді перемноження детермінованої і випадкової компонент, а саме:

$$x_t = d_t \times \varepsilon_t \quad \text{при} \quad t = 1, \dots, n.$$

Способи опису детермінованих компонент часового ряду сильно залежать від області застосування. В економічних (і багатьох інших) застосуваннях у детермінованої компоненти часового ряду d_t звичайно виділяють три складові частини: тренд tr_t , сезонну компоненту s_t і циклічну компоненту c_t . Для адитивної моделі часового ряду можна записати:

$$d_t = tr_t + s_t + c_t, \quad \text{при} \quad t = 1, \dots, n.$$

Трендом часового ряду tr_t при $t = 1, \dots, n$ називають плавно змінну, не циклічну компоненту, що описує чистий вплив довгострокових факторів, ефект яких позначається поступово.

Сезонна компонента s_t часового ряду при $t = 1, \dots, n$ описує поведінку, що змінюється регулярно протягом заданого періоду (року, місяця, тижня, дня і т.і.). Вона складається з послідовності циклів, що майже повторюються.

Циклічна компонента c_t часового ряду описує тривалі періоди відносного підйому і спаду. Вона складається з циклів, що змінюються по амплітуді і довжині.

Приведемо моделі трендів, що найбільш часто використовуються при аналізі економічних часових рядів, а також в багатьох інших областях. По-перше, це проста лінійна модель

$$tr_t = b_0 + b_1 \cdot t,$$

яка, не дивлячись на свою простоту, виявляється корисною в багатьох реальних задачах. Якщо нелінійний характер тренда очевидний, то може підійти одна з наступних моделей:

– поліноміальна: $tr_t = b_0 + b_1 t + b_2 t^2 + \dots + b_n t^n$, де значення ступені полінома n в практичних задачах рідко перевищує 5;

– логарифмічна: $tr_t = \exp(b_0 + b_1 t)$. Ця модель частіше за все застосовується для даних, що мають тенденцію зберігати постійні темпи приросту;

– логістична: $tr_t = \frac{a}{1 + b \cdot e^{-ct}}$,

– Гомперца: $\log(tr_t) = a - b \cdot r^t$, де $0 < r < 1$.

Дві останні моделі задають криві тренда S-образної форми. Вони відповідають процесам з поступово зростаючими темпами зростання в початковій стадії і поступово затухаючими темпами зростання в кінці. Необхідність подібних моделей обумовлена неможливістю багатьох економічних процесів тривалий час розвиватися з постійними темпами зростання або за поліноміальними моделями, у зв'язку з їх досить швидким зростанням (або зменшенням).

Перше уявлення про можливий характер тренда дає графічне представлення часового ряду. При прогнозуванні тренд використовують в першу чергу для довготривалих прогнозів. Точність короткострокових прогнозів, заснованих тільки на підбраній кривій тренда, як правило, недостатня.

В технічних додатках ми часто знаємо фізичні закони або технічні характеристики механізмів, що генерують досліджувані часові ряди. Це істотно полегшує дослідження. Один з типів моделей часових рядів, що часто використовується в технічних додатках - полігармонійна модель. Найпростіший варіант полігармонійної моделі часового ряду - це косінусоїдальна модель:

$$x_t = a \cos(\omega t + \theta) + \varepsilon_t.$$

Тут детермінованою компонентою є косінусоїдальна функція з амплітудою a , частотою ω , періодом $2\pi/\omega$ и фазою θ . Величини a , ω та θ є константами.

Коло даних, що описуються чисто косінусоїдальною моделлю невелике. По-перше, часто зустрічаються періодичні залежності, які

описуються не косінусоїдальною, а складнішою функцією. По-друге, звичайно в процесі, що вивчається, можна виділити не одну, а декілька періодичні компонент з різними періодами.

Як відомо з математичного аналізу, будь-яку гладку періодичну функцію $G(t)$ з періодом p (тобто функцію, для якої $G(t + kp) = G(t)$ для будь-якого цілого k) можна представити у вигляді ряду Фур'є:

$$G(t) = \sum_{j=1}^p a_j \cos(j\omega t + \theta_j),$$

де $\omega = 2\pi/p$ називається основною (Найквістовою) частотою, a_j , θ_j — деякі параметри. Частоти j називаються гармоніками основної частоти.

Функцію, що є сумою декількох періодичних функцій з різними періодами, можна задати у вигляді $G(t) = \sum_k G_k(t) = \sum_{j,k} a_{jk} \cos(j\omega_k t + \theta_{jk})$.

Говорять, що часовий ряд описується полігармонійною моделлю, якщо він представлений у вигляді:

$$x_t = \sum_{j,k} a_{jk} \cos(j\omega_k t + \theta_{jk}) + \varepsilon_t,$$

де $\omega_k = 2\pi/p_k$, а ε_t є білим шумом.

Математично найпростішою моделлю випадкової компоненти часового ряду є послідовність незалежних випадкових величин. *Білим шумом* називають часовий ряд (випадковий процес) з нульовим середнім, якщо його складові випадкові величини незалежні і розподілені однаково при всіх t .

Послідовності незалежних випадкових величин далеко не завжди адекватно описують випадкові компоненти часових рядів. Теорією і практикою для опису випадкових послідовностей вироблені і більш складні моделі.

Нехай $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n, \dots$ - незалежні однаково розподілені випадкові величини (білий шум).

Процесом ковзного середнього (першого порядку) із середнім μ називають процес $X(t)$:

$$X(t) = \varepsilon_t + \theta \varepsilon_{t-1} + \mu,$$

де θ - деякий числовий коефіцієнт, а μ - константа.

У процесі ковзного середнього статистично залежні тільки сусідні величини $X(t-1)$ і $X(t)$. Значення процесу, що розділені проміжком часу 2 і більше, статистично незалежні, тому що в їхньому формуванні беруть участь різні складові ε_t . З цієї причини процеси ковзного середнього є безпосереднім і найпростішим узагальненням процесів білого шуму.

Процесом авторегресії (першого порядку) із середнім значенням μ називають випадковий процес $X(t)$, що задовольняє співвідношенню:

$$X(t) - \mu = \phi \cdot (X(t-1) - \mu) + \varepsilon_t,$$

де ϕ і μ - деякі числа.

Члени процесу авторегресії, що розділені проміжком часу $h > 0$, не стають незалежними, яким би великим ні було h . Однак залежність між ними швидко убуває з ростом h , якщо $|\phi| < 1$. Саме такі процеси авторегресії звичайно зустрічаються в прикладних задачах.

1.2.2 Запитання для самоперевірки

1.2.2.1 Що таке адитивна і мультиплікативна моделі часового ряду?

1.2.2.2 Що є компонентами детермінованої складової часового ряду?

1.2.2.3 Що таке тренд часового ряду?

1.2.2.4 Для якого прогнозування (довго чи короткострокового) точність прогнозу на основі моделі тренда вище?

1.2.2.5 Що таке сезонна складова часового ряду?

1.2.2.6 Що таке циклічна складова часового ряду?

1.2.2.7 Які моделі тренда найбільш часто використовуються при аналізі економічних часових рядів?

1.2.2.8 Які моделі тренда, що використовуються при аналізі технічних часових рядів ви знаєте?

1.2.2.9 Що таке білий шум?

1.2.2.10 Який процес називають процесом ковзного середнього першого порядку?

1.2.2.11 Який процес називають процесом авторегресії першого порядку?

1.3 Числові характеристики часових рядів

Числові характеристики часових рядів: математичне очікування, коваріаційна функція, дисперсія, стандартне відхилення, кореляційна функція. Поняття стаціонарного процесу (сталість математичного чекання і дисперсії, властивість коваріаційної функції. Автокореляційна функція стаціонарного процесу. [1(с. 348-352)].

1.3.1 Методичні вказівки

Числові характеристики часових рядів вводяться в повній аналогії з числовими характеристиками випадкових величин.

Математичне очікування (перший момент) випадкового процесу $X(t)$ - це функція $m(t)$, така, що для кожного t значення функції $m(t)$ є математичним очікуванням випадкової величини $X(t)$:

$$m(t) = M X(t) .$$

Функцію $m(t)$ часто називають середнім значенням процесу $X(t)$. Вона використовується для опису систематичної зміни процесу. Наприклад, для випадкового процесу, що допускає запис у вигляді адитивної моделі, середнє значення дорівнює $tr_t + s_t + c_t$. Під словом «усереднювання» тут розуміється усереднювання випадкової величини $X(t)$ при незмінному t , а не усереднювання за часом.

Коваріаційна функція випадкового процесу $X(t)$ (коротко $\text{cov}(X(t), X(s))$) — це величина

$$B(s, t) = \text{cov}(X(t), X(s)) = M[(X(t) - m(t)) (X(s) - m(s))] .$$

Вона є функцією пари змінних (t, s) . Іноді її називають функцією других центральних моментів.

Значення коваріаційної функції при $t = s$ задає *дисперсію* випадкового процесу $DX(t) = \text{cov}(X(t), X(t))$. Квадратний корінь з $\text{cov}(X(t), X(t))$ називають *стандартним відхиленням* $\sigma(t)$ випадкового процесу $X(t)$:

$$\sigma(t) = \sqrt{\text{cov}(X(t), X(t))} .$$

Кореляційна функція випадкового процесу $X(t)$ — це величина:

$$\text{corr}(X(t), X(s)) = \frac{\text{cov}(X(t), X(s))}{\sigma(t)\sigma(s)}.$$

Як і коваріаційна функція, кореляційна функція також залежить від пари змінних (t, s) .

При фіксованих t і s $\text{corr}(X(t), X(s))$ за визначенням є коефіцієнтом кореляції випадкових величин $X(t)$ і $X(s)$. З визначення $\text{cov}(X(t), X(s))$ і $\text{corr}(X(t), X(s))$ виходить їх симетрія щодо t і s :

$$\text{cov}(X(t), X(s)) = \text{cov}(X(s), X(t)),$$

$$\text{corr}(X(t), X(s)) = \text{corr}(X(s), X(t)).$$

З визначення стаціонарності виходить, що для будь-якого s, t і будь-якого τ :

$$m(t + \tau) = m(t), \quad B(s + \tau, t + \tau) = B(s, t).$$

Положив $\tau = -t$, отримаємо, що

$$m(t) = m(0), \quad B(s, t) = B(s - t, 0).$$

Звідси витікає, що у стаціонарного процесу функції $m(t)$ і $\sigma(t)$ постійні, а коваріаційна функція $B(s, t)$ реально залежить не від пари (s, t) , як в загальному випадку, а від $|s - t|$. Так само можна переконатися, що і кореляційна функція стаціонарного процесу є функцією $|s - t|$. Розглянемо $t = s + k, k > 0$. За визначенням

$$r(k) = \text{corr}(X(t), X(s)) = \text{corr}(X(t), X(t + k)).$$

Автокореляційною функцією стаціонарного процесу $X(t)$ називають функцію $r(k) = \text{corr}(X(t), X(t + k))$, де $k > 0$ - ціле число.

Величину k часто називають затримкою, або лагом. Вона указує відстань між членами часового ряду, для яких обчислюється коефіцієнт кореляції.

Найпоширенішим випадком порушення стаціонарності на практиці є зміна середнього значення $m(t)$ із зміною часу t . В тих випадках, коли $m(t)$ вдається тим або іншим способом оцінити, перетворення $Y(t) = X(t) - m(t)$ перетворює процес на стаціонарний. Далі $Y(t)$ вивчають як стаціонарний, використовуючи для цього його специфічні властивості.

1.3.2 Запитання для самоперевірки

1.3.2.1 Що таке математичне очікування випадкового процесу?

1.3.2.2 Що таке коваріаційна функція випадкового процесу?

1.3.2.3 Що таке дисперсія випадкового процесу?

1.3.2.4 Що таке стандартне відхилення випадкового процесу?

1.3.2.5 Що таке кореляційна функція випадкового процесу?

1.3.2.6 Що таке стаціонарний випадковий процес?

1.3.2.7 Що таке автокореляційна функція випадкового процесу?

1.3.2.8 Наведіть приклади порушення стаціонарності випадкового процесу.

1.4 Вибіркові оцінки числових характеристик часового ряду

Вибіркові оцінки числових характеристик часового ряду (вибіркове середнє, вибірка дисперсія, вибірка автокореляційна функція). Графік вибіркової автокореляційної функції (корелограма). [1(с. 352-358), 6(с.4-13)].

1.4.1 Методичні вказівки

У кожний фіксований момент часу t випадковий процес $X(t)$ є випадковою величиною. Отже, для побудови оцінок його моментів $m(t)$ и $B(s, t)$ теоретично можна використовувати ті ж методи, що і для звичайних випадкових величин. При цьому потрібна деяка сукупність незалежних реалізацій цієї випадкової величини X , одержаних при повторенні досліду в незмінних умовах. Іншими словами, потрібна вибірка $x_1, \dots, x_k$. Застосування цієї методики до випадкового процесу вимагає від нас набору реалізацій (траєкторій) цього процесу $x_1(t), \dots, x_k(t)$.

У технічних додатках можливості для незалежних повторень досліду іноді є. Скажімо, вивчаючи коливання напруги в електричних мережах протягом доби, ми можемо вважати часові ряди, одержані в різні доби, незалежними реалізаціями одного випадкового процесу. Проте в економічних, соціальних, демографічних і подібних процесах ми звичайно маємо справу з єдиною траєкторією розвитку, повторити яку неможливо. Тому при вивченні статистичних властивостей таких процесів доводиться обходитися цією самою єдиною реалізацією. Проте довжина її може зростати.

По єдиній реалізації процесу $X(t)$ ми не можемо скласти оцінки для його середнього, дисперсії, коваріації і т.д., як ми зробили б це,

маючи в своєму розпорядженні вибірку. Але деякі схожі середні величини скласти можна.

Маючи ряд $x(t_1), \dots, x(t_n)$ послідовних спостережень випадкового процесу $X(t)$, можна скласти «середнє по реалізації»:

$$\bar{m} = \frac{1}{n} \sum_{i=1}^n x(t_i). \quad (1.1)$$

Виявляється, за деяких умов це середнє може служити оцінкою математичного очікування процесу $X(t)$. Першою з таких умов є стаціонарність випадкового процесу $X(t)$. Оскільки для стаціонарного процесу всі моменти часу рівноправні і його числові характеристики незмінні в часі, як оцінку $m(t) = m$ природно розглянути саме $\bar{m}$.

Розглянемо питання про точність цієї оцінки. Природно, хотілося б, щоб ця оцінка $\bar{m}$ наближалася до невідомого істинного значення із зростанням числа спостережень n , тобто була б спроможною. Оскільки відхилення оцінки від істинного значення можна описати за допомогою її дисперсії, то для спроможності достатньо, щоб

$$D\bar{m} \rightarrow 0 \quad \text{при} \quad n \rightarrow \infty. \quad (1.2)$$

На жаль, одна лише стаціонарність випадкового процесу не забезпечує виконання (1.2).

Теорема Слуцкого: для стаціонарного в широкому значенні випадкового процесу $X(t)$ оцінка його середнього значення (1.1) спроможна тоді і тільки тоді, коли:

$$\frac{1}{n} \sum_{t=0}^{n-1} r_t \rightarrow 0 \quad \text{при} \quad n \rightarrow \infty, \quad (1.3)$$

де r_t — автокореляційна функція процесу.

Звернемо увагу на те, що для виконання (1.3) достатнє, щоб $r_t \rightarrow 0$ при $t \rightarrow \infty$. Останнє зауваження дозволяє на практиці судити про те, чи можна використовувати усереднювання по одній реалізації для отримання спроможних оцінок його характеристик. Таким чином, теорема Слуцкого підкреслює важливість аналізу поведінки автокореляційної функції випадкового процесу.

Методика отримання оцінок значень автокореляційної функції $r(k)$ багато в чому нагадує випадок двох вибірок. Розберемо її устрій на оцінці $r(1)$ — кореляції між сусідніми членами часового ряду X_t і X_{t+1} .

Утворюємо з часового ряду $x_1, x_2, \dots, x_n$ сукупність з $n - 1$ пар: $(x_1, x_2), (x_2, x_3), \dots, (x_{n-1}, x_n)$. Перший елемент кожної пари, через стаціонарність, ми можемо розглядати як реалізацію випадкової величини X_t , а другий - як реалізацію випадкової величини X_{t+1} . Тоді оцінка коефіцієнта кореляції між X_t і X_{t+1} може бути записана у вигляді:

$$\bar{r}_1 = \frac{\sum_{t=1}^{n-1} (x_t - \bar{x}_{(1)}) (x_{t+1} - \bar{x}_{(2)})}{\sqrt{\left[\sum_{t=1}^{n-1} (x_t - \bar{x}_{(1)})^2 \sum_{t=1}^{n-1} (x_{t+1} - \bar{x}_{(2)})^2 \right]}}, \quad (1.4)$$

де

$$\bar{x}_{(1)} = \sum_{t=1}^{n-1} x_t / (n-1), \quad \bar{x}_{(2)} = \sum_{t=2}^n x_t / (n-1)$$

відповідно оцінки середніх значень величин X_t і X_{t+1} .

При великих значеннях n , враховуючи що $\bar{x}_{(1)} \approx \bar{x}_{(2)} \approx \bar{x}$ і $n/(n-1) \approx 1$, вираз (1.4) часто замінюють набагато простішим:

$$\bar{r}_1 = \frac{\sum_{t=1}^{n-1} (x_t - \bar{x})(x_{t+1} - \bar{x})}{\sum_{t=1}^n (x_t - \bar{x})^2} \quad (1.5)$$

Аналогічним чином може бути визначена оцінка кореляції між X_t і X_{t+k} або k -го члена автокореляційної функції r_k :

$$\bar{r}_k = \frac{\sum_{t=1}^{n-k} (x_t - \bar{x})(x_{t+k} - \bar{x})}{\sum_{t=1}^n (x_t - \bar{x})^2}. \quad (1.6)$$

Звернемо увагу, що точність наближення (1.6) помітно знижується із зростанням лага k , як через погіршення точності використаних вище замін, так і через зменшення числа спостережень, що використо-

вуються для обчислення оцінки $\bar{r}_k$. Тому на практиці звичайно обмежуються вивченням невеликого числа перших членів автокореляційної функції.

Функцію $\bar{r}_k$ аргументу k при $k = 1, 2, \dots$ називають *вибірковою автокореляційною функцією* або просто автокореляційною функцією. (При $k=0$ $\bar{r}_k$ за визначенням рівно 1 і це значення звичайно виключають з розгляду як не несуче ніякої інформації.) Графік вибіркової автокореляційної функції називають *корелограмою*. На цьому графіку (см. рис. 1.1), окрім значень самої функції, звичайно указують довірчі інтервали цієї функції в припущенні, що значення автокореляційної функції рівні 0 для всіх $k \neq 0$.

Вивчення властивостей вибірових оцінок автокореляційної функції часового ряду - в загальному випадку досить складна і до кінця не вирішена задача.

М.Бартлетом в 1946 р. для випадку нескінченного дискретного часового ряду ($-\infty < t < \infty$) був вказаний вираз дисперсії оцінки $\bar{r}_k$ для гауссовского процесу:

$$D\bar{r}_k = \frac{1}{n} \sum_{t=-\infty}^{t=\infty} [r_t^2 + r_{t-k}r_{t+k} - 4r_k r_t r_{t+k} + 2r_t^2 r_k^2]. \quad (1.7)$$

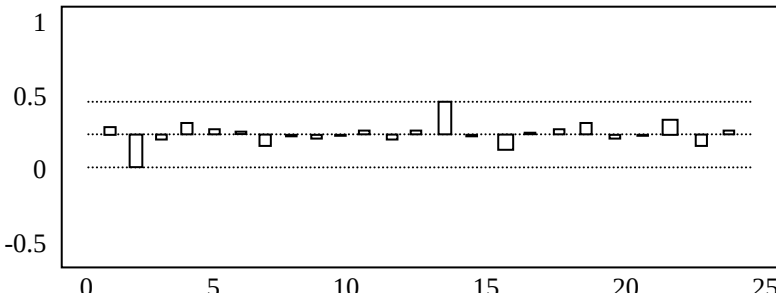


Рисунок 1.1– Коррелограма з довірчими інтервалами (при рівності нулю автокореляційної функції для всіх $k \neq 0$)

Цей результат показує, що ми не можемо оцінити по кінцевому відрізку часового ряду дисперсію оцінки $\bar{r}_k$, оскільки вона залежить

від нескінченного невідомого числа автокореляцій r_t . Тому на практиці доводиться задовольнятися лише наближеннями для виразу (1.7).

Вкажемо властивості оцінок $\overline{r_k}$ для найпростішого і практично важливого випадку. А саме, розглянемо властивості оцінок автокореляцій для часового ряду, що є стаціонарною послідовністю незалежних нормально розподілених випадкових величин або, іншими словами, гауссовским білим шумом. В цьому випадку для будь-яких k , не рівних нулю, за визначенням $r_k = 0$. Таким чином, всі доданки, що стоять під знаком суми у виразі (1.7), рівні нулю, окрім $r_0^2 = 1$. Звідси дисперсія $\overline{r_k}$ дорівнює:

$$D\overline{r_k} = \frac{1}{n}.$$

Звернемо увагу на те, що оцінка $\overline{r_k}$ у формі (1.6) є зміщеною. Можна показати, що, $M\overline{r_k} \approx -\frac{1}{n}$, проте величина цього зміщення прагне нуля із зростанням об'єму ряду і не така суттєва в прикладному аналізі.

Іншою важливою властивістю оцінки $\overline{r_k}$ є її асимптотична нормальність при $n \rightarrow \infty$.

Таким чином, для кожного окремого значення $\overline{r_k}$ ми можемо вказати приблизний 95% довірчий інтервал у вигляді: $-1/n \pm 2/\sqrt{n}$. Межі цього довірчого інтервалу звичайно наносяться на графік корелограми і називаються довірчою трубкою. Вони певною мірою дозволяють судити про те, наскільки процес, що вивчається, нагадує білий шум. Вказівка 95% довірчих меж для кожного коефіцієнта автокореляційної функції окремо не означає, що з 95% вірогідністю всі дані оцінки $\overline{r_k}$ одночасно потрапляють в довірчу трубку. Так, розглядаючи 20 перших оцінок $\overline{r_k}$ для гауссовського білого шуму, досить часто можна спостерігати, що одна або дві з оцінок виходять за межі довірчої трубки. Ця обставина також ускладнює інтерпретацію корелограми

1.4.2 Запитання для самоперевірки

1.4.2.1 В чому полягає складність визначення числових характеристик часового ряду?

1.4.2.2 Що таке вибіркове середнє часового ряду?

1.4.2.3 Що таке вибіркова дисперсія часового ряду?

1.4.2.4 Що таке вибіркова автокореляційна функція часового ряду?

1.4.2.5 Що таке затримка або лаг?

1.4.2.6 В якому діапазоні і чому саме так беруть значення лагу для розрахунку автокореляційної функції?

1.4.2.7 Що таке довірча трубка або довірчий інтервал?

1.4.2.8 Чому на корелограмі білого шуму коефіцієнти кореляції можуть виходити за довірчу трубку?

1.5 Методи зведення часового ряду до стаціонарного

Виділення тренда за допомогою методу найменших квадратів. Виділення тренда за допомогою простих різницевих операторів. Виділення сезонної компоненти часового ряду. Метод ковзних середніх. Сезонні різницеві оператори. Перетворення шкали вимірів значень часового ряду: логарифмічне перетворення, перетворення Боксу-Коксу. [1(с. 363-386), 6(с.5-17)].

1.5.1 Методичні вказівки

Після вивчення графіка часового ряду звичайно пробують виділити в часовому ряду тренд, сезонні і циклічні компоненти. Після їх виключення часовий ряд повинен стати стаціонарним. Крім того, для полегшення подальшого аналізу іноді використовуються перетворення значень часового ряду (точніше, тієї шкали, в якій вони зміряні) - це дозволяє наблизити розподіл значень часового ряду до нормального або зробити дисперсію цих значень більш постійною (інакше кажучи, стабілізувати дисперсію).

Для оцінки і видалення трендів з часових рядів частіше за все використовується метод найменших квадратів. Кажучи мовою регресійного аналізу, значення часового ряду x_t розглядають як відгук (залежну змінну), а час t — як чинник, що впливає на відгук (незалежну змінну):

$$X_{ti} = f(t_i, \theta) + \varepsilon_i, \quad i = 1, \dots, n,$$

де f — функція тренда (вона звичайно передбачається гладкою), θ — невідомі нам параметри (параметри моделі часового ряду), а ε_i — незалежні і однаково розподілені випадкові величини, розподіл яких ми припускаємо нормальним. Метод найменших квадратів полягає в тому, що ми вибираємо функцію тренда так, щоб

$$\sum_{i=1}^n [x_{ti} - f(t_i, \theta)]^2 \rightarrow \min_{\theta}.$$

Для часових рядів типово, що статистичні передумови регресійного аналізу виконуються не повністю. Це особливо стосується припущення про незалежність випадкових відхилень. Для часових рядів характерна саме взаємна залежність його членів (принаймні, не далеко віддалених за часом). Проте, оцінки тренда і в цих умовах звичайно виявляються розумними, якщо вибрана адекватна модель тренда і якщо серед спостережень немає великих викидів. Згадані порушення передумов регресійного аналізу позначаються не стільки на значеннях оцінок, скільки на їх статистичних властивостях. Так, за наявності помітної залежності між членами часового ряду оцінки дисперсії, засновані на залишковій сумі квадратів, дають неправильні результати. Неправильними виявляються і довірчі інтервали для коефіцієнтів моделі, і т.д. В кращому разі їх можна розглядати як дуже наближені.

Це положення може бути частково виправлено, якщо застосовувати модифіковані алгоритми методу найменших квадратів, такі як зважений метод найменших квадратів або метод найменших квадратів для корельованих спостережень. Проте для цих методів потрібна додаткова інформація про те, як змінюється дисперсія спостережень або їх кореляція. Якщо ж така інформація недоступна, дослідникам доводиться застосовувати класичний метод найменших квадратів, не дивлячись на вказані недоліки.

Разом з методом якнайменших квадратів, для видалення тренда можна використовувати і ряд інших методів. Одним з них є метод переходу від початкового ряду до ряду різниць сусідніх значень ряду. В більш загальному вигляді ця ідея описується за допомогою застосування до ряду різницевих операторів різних порядків.

Процедура переходу від ряду x_t при $t = 1, \dots, n$ до ряду $y_t = x_t - x_{t-1} = \nabla x_t$ при $t = 2, \dots, n$ називається взяттям перших різниць, а оператор ∇ називається простим різницевим оператором першого порядку.

Відмітимо, що довжина ряду перших різниць y_t на одиницю менше, ніж довжина початкового ряду x_t . Покажемо, як діє різницевий оператор на часовому ряді x_t , що містить простий лінійний тренд:

$$y_t = \nabla x_t = x_t - x_{t-1} = b_0 + b_1 \cdot t + \varepsilon_t - b_0 - b_1(t-1) - \varepsilon_{t-1} = b_1 + \varepsilon_t - \varepsilon_{t-1}$$

Видно, що на відміну від ряду x_t , перетворений ряд y_t вже не містить тренда, проте структура випадкової компоненти в ньому вже інша. Так, якщо ε_t була послідовністю незалежних випадкових величин, то послідовність $\varepsilon_t - \varepsilon_{t-1}$, $t = 2, \dots, n$, цією властивістю вже не володіє. Кореляція між сусідніми членами цієї послідовності дорівнює -0.5.

Аналогічним чином можна ввести різницевий оператор другого і більш високих порядків. Так, простий різницевий оператор другого порядку перетворює ряд x_t до ряду y_t , де

$$y_t = \nabla^2 x_t = \nabla (\nabla x_t) = \nabla (x_t - x_{t-1}) = \nabla x_t - \nabla x_{t-1} = x_t - 2x_{t-1} + x_{t-2}.$$

Часто для запису різницевих операторів використовують оператор B «зсуву назад»: $Bx_t = x_{t-1}$. При цьому

$$\nabla x_t = (1 - B) x_t, \quad \nabla^2 x_t = (1 - B)^2 x_t, \quad \nabla^k x_t = (1 - B)^k x_t.$$

Ясно, що довжина ряду $\nabla^k x_t$ на k одиниць менше довжини початкового ряду.

Прості різницеві оператори більш високих порядків дозволяють видаляти з ряду полиноміальні тренди відповідних порядків.

Різницеві оператори дійсно придатні для видалення трендів, особливо якщо не видна відповідна аналітична модель тренда. Недолік же методу різницевих операторів в тому, що не завжди ясно, як прикласти до початкового часового ряду результати статистичного аналізу його різниць. Зокрема, це відноситься до законів розподілу помилок. До того ж ці різниці можуть мати (і часто мають) набагато більш складну статистичну структуру, ніж початковий ряд.

Багато часових рядів, особливо економічні, містять сезонні компоненти. Сезонні компоненти ряду можуть як представляти інтерес

самі по собі, так і виступати в ролі чинника, що заважає. В обох випадках задача дослідника - виділити і усунути їх з ряду.

Для цього є декілька способів. Їх вибір звичайно визначається моделлю підбраного часового ряду. Розглянемо дві найпоширеніші моделі опису економічних часових рядів. Перша з них включає тренд (tr_t), сезонну (s_t) і випадкову (ε_t) компоненти:

$$x_t = tr_t + s_t + \varepsilon_t \quad (1.8)$$

Інша модель, крім перерахованих вище компонент, включає ще і циклічну компоненту (c_t):

$$x_t = tr_t + s_t + c_t + \varepsilon_t \quad (1.9)$$

Припустимо, що часовий ряд $x_1, \dots, x_n$ може бути описаний адитивною моделлю (1.8). Нехай p - період послідовності s_t , так що $s_t = s_{t+p}$ для будь якого t . Наша задача - оцінити значення s_t по спостереженням x_t при тому, що величина p відома.

Для цього спочатку ми повинні оцінити тренд tr_t . Це можна зробити за допомогою методу найменших квадратів або його модифікацій. Позначимо через $\hat{tr}_t$ одержану оцінку тренда. Звичайно вона виражається у вигляді деякої достатньо гладкої функції залежної від часу t і одного або декількох невідомих параметрів. Оцінки цих параметрів і дає метод найменших квадратів.

Потім для кожного сезону i , $1 \leq i \leq p$, розглянемо всі різниці, що відносяться до нього

$$x_i - \hat{tr}_i, x_{i+p} - \hat{tr}_{i+p}, \dots, x_{i+mp} - \hat{tr}_{i+mp}, \quad (1.10)$$

(для простоти припускаємо, що в даному ряду міститься ціле число періодів, тобто $n = (m + 1)p$). Кожне з цих відхилень x_i від $\hat{tr}_i$ можна розглядати як результат впливу сезонних змін. Усереднювання цих різниць дає нам оцінку сезонної компоненти s_i . Як найпростішу оцінку можна взяти просте середнє, тобто покласти

$$\hat{s}_i = \frac{1}{m+1} \sum_{l=0}^m (x_{i+lp} - \hat{tr}_{i+lp}) \quad \text{для} \quad i = 1, \dots, p. \quad (1.11)$$

Як інші оцінки $\hat{S}_i$ можна взяти зважене середнє, цензоване середнє, медіану і т.д. Перераховані середні зменшують вплив спостережень, що різко виділяються.

Часто буває бажано, щоб сума сезонних ефектів дорівнювала нулю. Тоді переходять до скорегованих оцінок сезонних ефектів у вигляді

$$s_i^* = \hat{s}_i - \frac{1}{p} \sum_{i=1}^p \hat{s}_i.$$

В практичних задачах поширена ситуація, коли сезонні коливання пропорційні середньому значенню процесу в даний момент часу. Для опису подібних даних можна використовувати одну з наступних моделей:

$$\begin{aligned} x_t &= tr_t \cdot s_t + \varepsilon_t \\ x_t &= tr_t \cdot s_t \cdot \varepsilon_t. \end{aligned}$$

Перша з них є змішаною мультиплікативно-адитивною моделлю, друга - мультиплікативною моделлю часового ряду. Для моделі $x_t = tr_t \cdot s_t + \varepsilon_t$ при оцінці сезонних ефектів замість сукупності (1.10) розглядають сукупність (1.12) виражених у відсотках:

$$\frac{x_{i+lp}}{tr_{i+lp}} \cdot 100\% \quad \text{при} \quad l = 0, 1, 2, \dots, m. \quad (1.12)$$

В цьому випадку оцінкою сезонної компоненти або *сезонним індексом* називають величину:

$$\hat{s}_i = \frac{1}{m+1} \sum_{l=0}^m \left(\frac{x_{i+lp}}{tr_{i+lp}} \cdot 100\% \right), \quad \text{де} \quad 1 \leq i \leq p. \quad (1.13)$$

Так само, як і у разі адитивної моделі, замість середнього арифметичного в правій частині може фігурувати зважене або цензоване середнє, медіана або інші більш стійкі до грубих викидів оцінки. Сезонні індекси (1.13) особливо популярні при аналізі економічних часових рядів.

На практиці вважається, що оцінки сезонних ефектів недостатньо точні, якщо число періодів в досліджуваному сезонному часовому ряду менше п'ять-шести. Це означає, наприклад, що при розгляді місячних даних для достатньо точної оцінки сезонних ефектів необхідні, як мінімум, спостереження за п'ять-шість років.

Одержавши оцінки сезонних ефектів (1.11), в адитивній моделі легко провести видалення цих ефектів з даного ряду, віднімаючи їх з початкових значень ряду. Подібна процедура часто носить назву сезонного вирівнювання ряду або сезонної корекції ряду. Для мультиплікативно-адитивної моделі ця процедура зводиться до ділення значень початкового ряду на відповідні сезонні індекси і множенні на 100%.

За наявності у ряді циклічної компоненти розрахунок сезонних ефектів дещо відрізняється від описаного вище. В цьому випадку для з'ясування сезонних внесків необхідно оцінити не тільки тренд, але і циклічну компоненту. Простіше за все одночасно оцінити тренд і циклічну компоненту можна за допомогою ковзного середнього. Цей метод корисний і тоді, коли модель тренда не ясна.

Метод ковзних середніх - один з найстаріших і широко відомих способів згладжування часового ряду. Він заснований на переході від початкових значень ряду до їх середніх значень на інтервалі часу, довжина якого вибрана наперед. При цьому сам вибраний інтервал часу ковзає уздовж ряду.

Одержаний таким чином ряд ковзних середніх поводить себе набагато більш гладко, ніж початковий ряд, за рахунок усереднювання відхилень початкового ряду. Таким чином, ця процедура дає уявлення про загальну тенденцію поведінки ряду.

Застосовуючи метод ковзних середніх, можна використовувати різні види усереднювання значень ряду: середнє арифметичне, медіани та ін. До згладжування за допомогою медіани (медіанне згладжування) вдаються тоді, коли серед спостережень є викиди (дані, що різко виділяються). Величину інтервалу згладжування доцільно вибирати рівним або кратним періоду сезонності.

Дамо формальне визначення ковзного середнього спочатку для інтервалів згладжування, довжина яких задається непарними числами. Нехай $p = 2m + 1$. Позначимо через $\hat{x}_t$ результат усереднювання елементів ряду $x_{t-m}, \dots, x_{t-1}, x_t, x_{t+1}, \dots, x_{t+m}$.

Якщо обговорюване середнє є середнє арифметичне, то

$$\hat{x}_t = \frac{1}{2m+1} (x_{t-m} + \dots + x_{t-1} + x_t + x_{t+1} + \dots + x_{t+m}).$$

Для парних $p = 2m$ визначення дещо складніше. Причина в тому, що обчислене по аналогічних формулах (як середнє арифметичне, медіана і т.д.) усереднене значення не можна зіставити якому-небудь певному моменту часу t . Наприклад, середнє арифметичне

$$\frac{1}{2m} \sum_{t=1}^{t=2m} (x_t)$$

слід було б зіставити моменту часу $t = (2m + 1)/2$, але такого моменту в часовому ряду немає. А це сильно ускладнює подальше виділення сезонних ефектів.

Тому при парному інтервалі згладжування $2m$ в усереднюванні задіюють не $2m$, а $2m+1$ значень тимчасового ряду, але значення на краях інтервалу згладжування беруть з вагами $1/2$. Так, при використанні для усереднювання середнього арифметичного виходить наступна формула:

$$\hat{x}_l = \frac{1}{2m} \left(\frac{1}{2} x_{l-m} + x_{l-m+1} + \dots + x_{l+m-1} + \frac{1}{2} x_{l+m} \right).$$

Цей вираз задає величину простого ковзного середнього $\hat{x}_l$ для $l = m + 1, m + 2, \dots, n - m$ при парній величині інтервалу згладжування $p = 2m$.

Припустимо, що спостережуваний часовий ряд має структуру

$$x_t = tr_t + c_t + s_t + \varepsilon_t,$$

де $tr_t + c_t$ - тренд і циклічна складова, s_t - сезонна складова, а ε_t - випадкова складова ряду. Нехай p - період послідовності s_t , так що $s_t = s_{t+p}$ для будь-якого t . Нехай величина p нам відома. Ми хочемо оцінити значення s_t за спостереженнями x_t .

Порядок оцінки сезонних компонент в цьому випадку, в цілому, аналогічний розглянутому раніше, тільки замість оцінки тренда методом найменших квадратів використовують ковзне середнє як сумісну оцінку тренда і циклічної компоненти.

Для мультиплікативної моделі часового ряду, коли

$$x_t = tr_t \cdot c_t \cdot s_t \cdot \varepsilon_t$$

доцільно перейти до логарифмів $y_t = \log x_t$. Тоді

$$y_t = d_t + g_t + r_t + \delta_t,$$

де $d_t = \log tr_t$, $g_t = \log c_t$, $r_t = \log s_t$, $\delta_t = \log \varepsilon_t$. До ряду y_t можна застосувати викладену вище методику, починаючи з обчислення ковзних середніх і закінчуючи складанням оцінки $\hat{r}_i$ для r_i . Оцінкою для початкової величини $S_i = e^{r_i}$ буде служити $e^{\hat{r}_i}$, якщо $\log x$ — натуральний логарифм x , або $\hat{S}_i = 10^{\hat{r}_i}$, якщо наші логарифми десятинні.

Видалення сезонної компоненти проводиться так само, як і в розібраному вище випадку. Для адитивної моделі видалення сезонної компоненти зводиться до віднімання оціненої сезонної компоненти з початкового ряду. Для мультиплікативної моделі ця процедура полягає в діленні значень початкового ряду на відповідні сезонні індекси.

Ще один спосіб видалення сезонних компонент з ряду заснований на використанні спеціальних різницьових операторів, які називаються *сезонними*.

Нехай $x_1, \dots, x_n$ — реалізація часового ряду, а p — період його сезонності.

Процедура переходу від ряду x_t (при $t = 1, \dots, n$) до ряду $y_t = x_t - x_{t-p} = \nabla_p x_t$ (при $t = p+1, \dots, n$) називається взяттям першої сезонної різниці, а оператор ∇_p називається сезонним різницевим оператором з періодом p .

Перетворення $x_t - x_{t-p}$ може бути також записано за допомогою оператора зсуву назад B у вигляді:

$$y_t = x_t - x_{t-p} = (1 - B^p) x_t.$$

Як і у випадку з простими різницьовими операторами, іноді бувають корисні сезонні оператори більш високих порядків. Так, сезонний оператор другого порядку з періодом p задається співвідношенням:

$$\nabla_p^2 x_t = \nabla_p(\nabla_p x_t) = \nabla_p(x_t - x_{t-p}) = x_t - 2x_{t-p} + x_{t-2p}$$

або, за допомогою оператора зсуву назад B :

$$\nabla_p^2 x_t = (1 - B^p)^2 x_t = (1 - 2B^p + B^{2p}) x_t.$$

Прості і сезонні різницеві оператори можуть бути використані відповідно для видалення тренда і сезонної компоненти з часового ряду. Якщо часовий ряд одночасно містить обидві ці компоненти, то їх видалення можливе за допомогою послідовного застосування простих і сезонних операторів. Неважко перекоонатися, що порядок застосування цих операторів не суттєвий:

$$\nabla \nabla_p x_t = \nabla (x_t - x_{t-p}) = (x_t - x_{t-1}) - (x_{t-p} - x_{t-p-1}) = \nabla_p \nabla x_t.$$

До перетворень значень часового ряду (точніше - до перетворень тієї шкали, в якій зміряні значення часового ряду) вдаються звичайно з двох причин: або для того, щоб наблизити розподіл до нормального (наприклад, позбавитися від його скошеності), або для того, щоб зробити дисперсію часового ряду більш постійною (іншими словами, стабілізувати дисперсію часового ряду).

Нехай змінна x вживається для запису значень часового ряду. Розглянемо перетворення x в y за правилом $y = f(x)$, де f означає деяку певну функцію. (Звичайно f — монотонна функція; тоді від значень y можна однозначно повернутися до значень x .) Застосовуючи перетворення f до кожного члена ряду x_t , ми одержимо новий часовий ряд $y_t = f(x_t)$.

Частіше інших перетворення, що використовується, - логарифмічне, коли

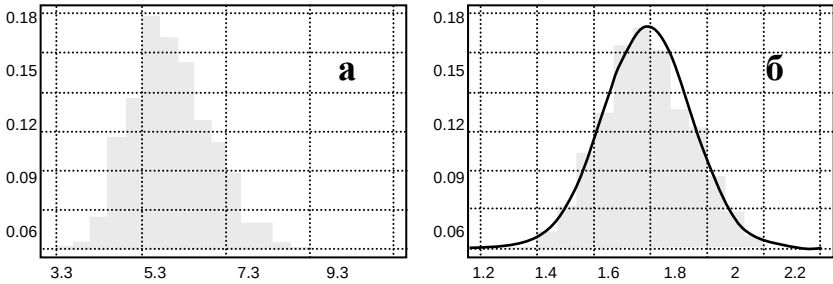
$$y = \log x, \quad \text{або} \quad y = \log (x + c),$$

де c — деяка постійна величина, вибір якої знаходиться у розпорядженні дослідника. При логарифмічному перетворенні

$$y_t = \log (x_t + c).$$

Логарифмічне перетворення можна застосовувати тільки до позитивних величин. В тих випадках, коли частина членів ряду x_t негативна, перед переходом до логарифмів до всіх членів **ряду** додають сталу c , добиваючись того, щоб $x_t + c > 0$ при всіх t .

Скошені (асиметричні) розподіли досить часто з'являються в економічній статистиці. Типовим прикладом є дані про душевий дохід: осіб з невеликими і середніми доходами набагато більше, ніж осіб з високими доходами. А останніх значно більше, ніж осіб з дуже високими доходами. Приблизна гістограма розподілу доходів приведена на рис. 1.2а. Прологарифмуємо дані про доходи і знов побудуємо гістограму. Вона приведена на рис. 1.2б. Видно, що в логарифмічній шкалі розподіл доходів близький до нормального (гаусовського).



а) вихідна шкала, б) логарифмічна шкала (для наочності на графік (б))

Рисунок 1.2 - Гістограма даних про доходи

Логарифмічне перетворення може виявитися корисним і при деяких порушеннях стаціонарності ряду що спостерігається. Припустимо, що ми спостерігаємо процес $x_t = b_t \cdot z_t$, де z_t — стаціонарний ряд, а b_t — деяка позитивна не випадкова послідовність. Позначивши $D z_t$ через σ^2 , одержимо, що $D x_t = \sigma^2 b_t^2$ змінюється в часі. Перехід до логарифмічної шкали $y_t = \log x_t$ дає

$$y_t = \log b_t + \log z_t.$$

При цьому ряд $\log z_t$ — стаціонарний, його дисперсія в часі не змінюється. Це дозволяє застосувати метод найменших квадратів для виділення тренду $\log b_t$ з ряду y_t .

Логарифмічне перетворення є окремим випадком деякого сімейства перетворень, яке ввели Дж.Бокс і Д.Кокс в 1964 р. Перетворення, що створюють це сімейство, залежать від параметра λ , $\lambda \geq 0$. Якщо повернутися до формули перетворень $y = f(x)$, то можна сказати, що тепер $y = f(x, \lambda)$, де значення $\lambda \geq 0$ дослідник може вибрати на свій розсуд. Бокс і Кокс запропонували наступну формулу

$$f(x, \lambda) = \begin{cases} (x_t^\lambda - 1) / \lambda & \text{при } \lambda > 0 \\ \log x_t & \text{при } \lambda = 0 \end{cases}.$$

Неважко переконатися, що при фіксованому λ функція $f(x, \lambda)$ монотонно зростає із зростанням x , і що $f(x, \lambda)$ безперервна не тільки по x , але і по λ , якщо $\lambda \geq 0$.

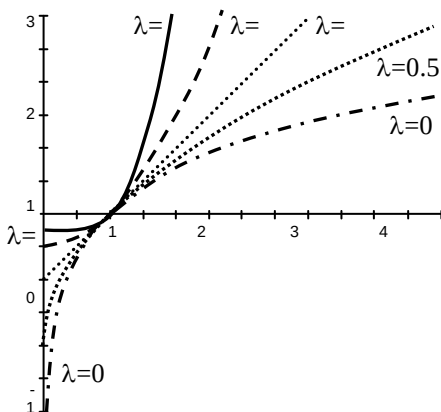


Рисунок 1.3 - Характер перетворення Боксу-Коксу при різних

Як видно з рисунку 1.3, перетворення Боксу-Коксу при $\lambda < 1$ розтягує відстані між малими значеннями і стискає його між великими за величиною значеннями даних. При $\lambda > 1$ спостерігається зворотна картина.

Слід зауважити, що застосування перетворення Боксу-Коксу до часових рядів може породжувати певні труднощі в їх подальшому аналізі. Річ у тому, що показник ступеня λ істотно впливає на кореляційну функцію процесу і здатний значно ускладнити подальший підбір моделі ряду.

Подібно логарифмічному, перетворення Боксу-Коксу можна застосовувати тільки до позитивних чисел. Якщо частина членів ряду x_t негативна, перш ніж застосувати до ряду перетворення Боксу-Коксу, до всіх членів ряду додають постійну c . Члени перетвореного ряду одержують за формулою

$$y_t = \frac{(x_t + c)^\lambda - 1}{\lambda},$$

якщо обране $\lambda > 0$. Для $\lambda = 0$ перетворення Боксу-Коксу діє як вже згадане логарифмічне: $y_t = \log(x_t + c)$.

1.5.2 Запитання для самоперевірки

1.5.2.1 Яка мета виділення з ряду детермінованих компонент?

1.5.2.2 В чому полягає виділення тренда за допомогою методу найменших квадратів?

1.5.2.3 Які є різновиди методу найменших квадратів, та для чого вони потрібні?

1.5.2.4 В чому полягає виділення тренда за допомогою простих різницевих операторів?

1.5.2.5 Як проводиться виділення сезонної компоненти часового ряду?

1.5.2.6 З якою метою застосовують метод ковзних середніх?

1.5.2.7 В чому полягає суть методу ковзних середніх?

1.5.2.8 Для чого застосовують сезонні різницеві оператори?

1.5.2.9 Яка мета перетворення шкали вимірів значень часового ряду?

1.5.2.10 В чому полягає логарифмічне перетворення часового ряду?

1.5.2.11 В чому полягає перетворення Боксу-Коксу?

1.6 Дослідження структури стаціонарної випадкової складової часового ряду

Корелограми часових рядів, що містять тренд або сезонну складову. Дослідження стаціонарної випадкової складової часового ряду. Корелограма білого шуму. Корелограми процесу ковзного середнього першого та вищих порядків. Корелограми процесу авторегресії першого та вищих порядків. Часткова автокореляційна функція процесів ковзного середнього. Часткова автокореляційна функція процесів авторегресії. [1(с. 387-394), 6(с.19-30)].

1.6.1 Методичні вказівки

Після видалення детермінованої компоненти часовий ряд повинний звестися до стаціонарного процесу. Так що наступним кроком після виділення детермінованої компоненти повинний бути аналіз залишків, тобто вивчення ряду, отриманого з вихідного часового ряду після виключення детермінованої компоненти.

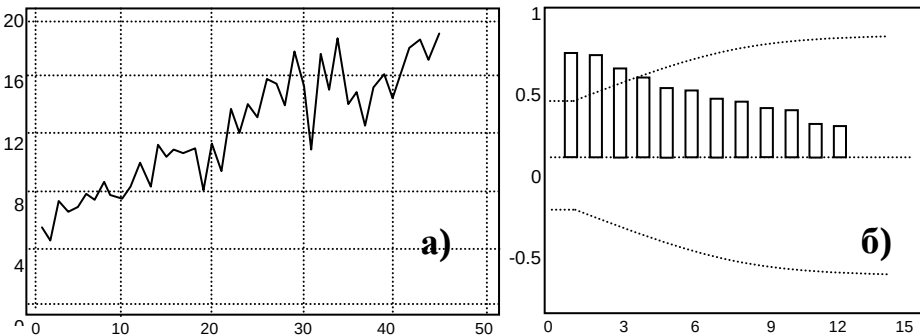
Як модель стаціонарних часових рядів найчастіше використовуються процеси авторегресії, ковзного середнього та їхні комбінації.

Для перевірки стаціонарності ряду залишків і оцінки його дисперсії на практиці найчастіше використовуються вибіркова автокореляційна (корелограма) і часткова автокореляційна функції.

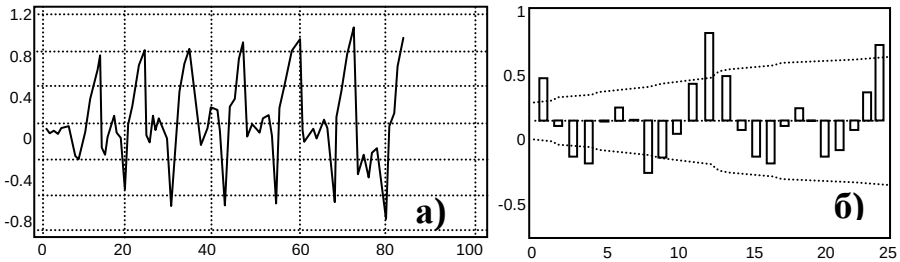
Аналіз корелограми - це порой досить непроста задача. Розглянемо поведінку корелограми для деяких нестационарних рядів. У цьому випадку варто пам'ятати, що корелограма практично не несе ніякої інформації про статистичну залежність чи незалежність членів часового ряду, однак вона може відбивати причини порушення стаціонарності.

Для часового ряду, що містить *тренд*, корелограма не прагне до нуля з ростом значення лага. Її характерне поведінку зображене на рис. 1.4.

Для ряду з *сезонними коливаннями* корелограма також буде містити періодичні сплески, що відповідають періоду сезонних коливань. Це дозволяє встановлювати передбачуваний період сезонності. Типове поведінку корелограми для ряду з сезонними коливаннями приведе- не на рис. 1.5.



а) вихідний часовий ряд; б) його корелограма
Рисунок 1.4 - Корелограма ряду, що містить тренд



а) вихідний часовий ряд; б) його корелограма
Рисунок 1.5 - Корелограма ряду з сезонними коливаннями

У випадку стаціонарних випадкових процесів корелограма показує корельованість значень часового ряду при різних відстанях між ними.

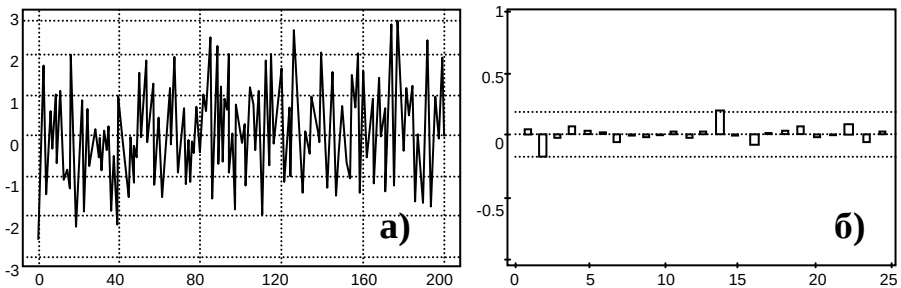
Автокореляційна функція r_k білого шуму дорівнює нулю для всіх $k \neq 0$. На рис. 1.6 зображена типова корелограма білого шуму. Для гаусовського білого шуму можна вказати 95% довірчий інтервал для кожного конкретного значення $\bar{r}_k$ у вигляді $-1/\sqrt{n} \pm 2/\sqrt{n}$. Він зображений на графіку корелограми пунктирними лініями. Якщо вибіркові оцінки кореляційної функції попадають у зазначені довірчі інтервали, то можна припустити, що значення процесу є білим шумом. Однак досить часто одне чи декілька значень вибіркової автокореляційної функції білого шуму можуть виходити з зазначених меж. Особливо часто цей ефект можна спостерігати при наявності невеликого числа спостережень.

Траєкторії багатьох стаціонарних випадкових процесів виглядають набагато більш гладко, чим траєкторії білого шуму. Це зв'язано з наявністю позитивної кореляції між двома чи декількома сусідніми членами подібних рядів. Якщо ж кореляція між сусідніми членами ряду негативна, то траєкторії подібних процесів будуть більш зламаними, чим траєкторії білого шуму. Найпростішим прикладом процесів, у яких залежні одне чи декілька сусідніх значень, є процеси *ковзного середнього*. На рис. 1.7 приведені графіки ста значень реалізації процесу ковзного середнього з коефіцієнтом $\theta = 0.75$ і його корелограми. На рис. 1.8 приведені аналогічні графіки при $\theta = -0.75$.

На графіках видно, що хоча отримані оцінки значень r_k при $k = 2, 3, \dots$ не дорівнюють нулю, вони значиме не відрізняються від нульових значень, тому що попадають у 95% довірчий інтервал, що побудований у припущенні рівності нулю відповідних значень автокореляційної функції.

Для процесів ковзного середнього другого порядку відрізняються від нуля тільки значення r_1 і r_2 , а всі наступні значення r_k при $k = 3, 4, \dots$ дорівнюють нулю. Нарешті, для процесів ковзного середнього порядку q відмінні від нуля тільки перші q значень автокореляційної функції. Строячи графіки корелограм для подібних процесів, можна на підставі зазначеної властивості зробити попередній висновок про можливий порядок процесу ковзного середнього, котрий може бути використаний для опису спостереженого ряду.

Зазначене правило гарне, якщо підібраний порядок моделі ковзного середнього невеликий, скажемо від одного до чотирьох-п'яти. Однак на практиці часто зустрічаються стаціонарні процеси з автокореляційною функцією помітно відмінною від нуля навіть при великих затримках. Згідно сформульованому правилу, їх можна намагатися описати процесами ковзного середнього високих порядків. Це приводить до великого числа коефіцієнтів процесу ковзного середнього, котрі підлягають подальшій оцінці. При цьому точність цих оцінок помітно знижується. Практична цінність таких багатопараметричних моделей ковзного середнього невелика. У цій ситуації краще спробувати описати часовий ряд за допомогою моделі авторегресії. Якщо і ці спроби не закінчаться успіхом - перейти до комбінованих моделей авторегресії - ковзного середнього.



а) вихідний ряд; б) його корелограма

Рисунок 1.6 – Корелограма білого шуму

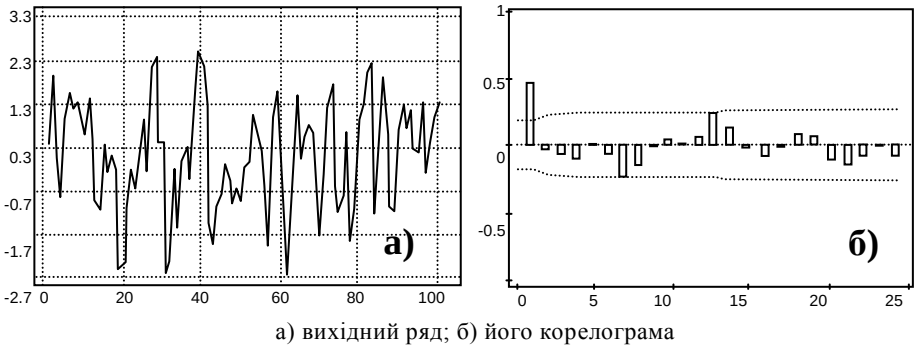


Рисунок 1.7 - Корелограма процесу ковзного середнього при $\theta = 0.75$

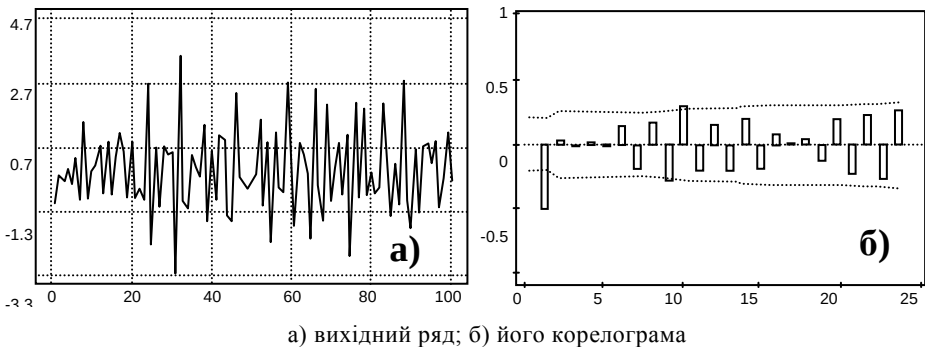


Рисунок 1.8 - Корелограма процесу ковзного середнього при $\theta = -0.75$

Нагадаємо, що найпростіший процес авторегресії першого порядку $X(t)$ з нульовим середнім задається співвідношенням:

$$X(t) = \phi X(t-1) + \varepsilon_t,$$

де ε_t не залежить від $X(t-1)$. Члени навіть цього найпростішого процесу не стають незалежними з ростом проміжку часу між ними. Однак за певних умов на коефіцієнти ця залежність швидко убиває.

Приведемо два типові графіки поведінки вибірових автокореляційних функцій цих процесів. Автокореляційні функції цих процесів з ростом лага або просто експоненційно загасають, або являють собою експоненційно загасаючі синусоїдальні хвилі.

На рис. 1.9 приведені графіки ста значень реалізації процесу авторегресії другого порядку AR(2):

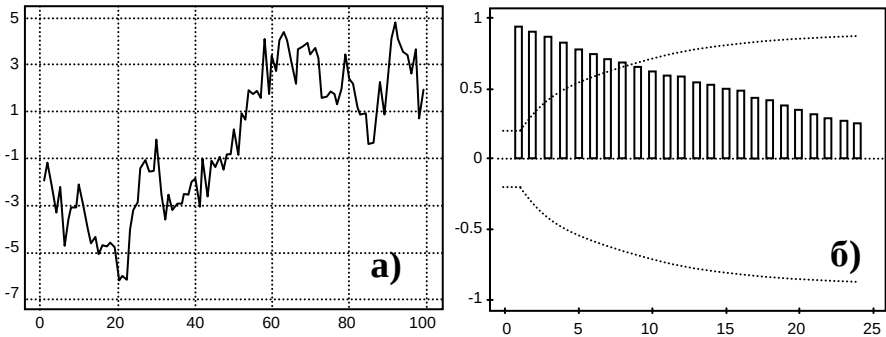
$$X(t) = \phi_1 X(t-1) + \phi_2 X(t-2) + \varepsilon_t,$$

при $\phi_1 = 0.7$ і $\phi_2 = 0.25$. Тут автокореляційна функція процесу і відповідно корелограма експоненційно загасають з ростом лага.

На рис. 1.10 приведені графіки ста значень реалізації AR(2) процесу при $\phi_1 = 0.7$ і $\phi_2 = -0.25$. Автокореляційна функція цього процесу і відповідно корелограма поведуться з ростом лага як експоненційно загасаюча синусоїда.

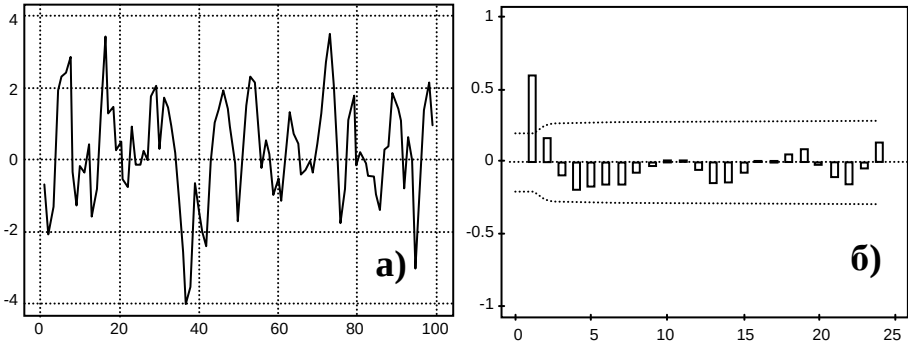
Для того, щоб по отриманій реалізації процесу підібрати модель авторегресії, необхідно попередньо вказати можливий порядок цієї моделі. Приведені приклади авторегресійних процесів показують, що безпосередньо з виду вибіркової автокореляційної функції цей висновок зробити досить важко. Цю задачу значно полегшує спеціально перетворена автокореляційна функція. Вона називається частковою автокореляційною функцією (ЧАКФ).

Для процесів авторегресії першого порядку всі значення вибіркової часткової автокореляційної функції, починаючи з другого, повинні значуще не відрізнятися від нуля, тобто потрапляти у відповідний довірчий інтервал.



а) вихідний ряд; б) його корелограма

Рисунок 1.9 - Корелограма AR(2) процесу при $\phi_1 = 0.7$, $\phi_2 = 0.25$



а) вихідний ряд; б) його корелограма

Рисунок 1.10 - Корелограма AR(2) процесу при $\phi_1 = 0.7$, $\phi_2 = -0.25$

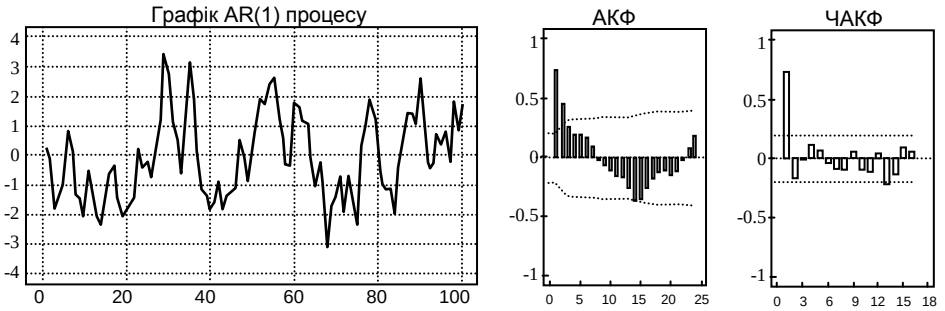


Рисунок 1.11 - Графік AR(1) процесу (значення $\phi = 0.75$), його вибіркова автокореляційна функція (АКФ) та вибіркова часткова автокореляційна функція (ЧАКФ)

Це і видно на рис. 1.11. Тут значення вибіркової ЧАКФ (на відміну від вибіркової АКФ), починаючи з другого, малі і значуще не відрізняються від нуля. Таким чином, за поведінкою вибіркової ЧАКФ легше з'ясувати вид моделі часового ряду.

Для процесу авторегресії другого порядку відмінні від нуля тільки перші два значення ЧАКФ. На рис. 1.12а і 1.12б зображені вибіркві ЧАКФ процесів, що представлені на рис. 1.9 і 1.10.

Для процесів ковзного середнього (МА- процесів), на відміну від процесів авторегресії, ЧАКФ при великих значеннях лага k не пе-

ретворюється в нуль, а експоненційно убуває. На рис. 1.13 приведені вибіркові ЧАКФ МА(1) процесів при $\theta = 0.75$ і $\theta = -0.75$.

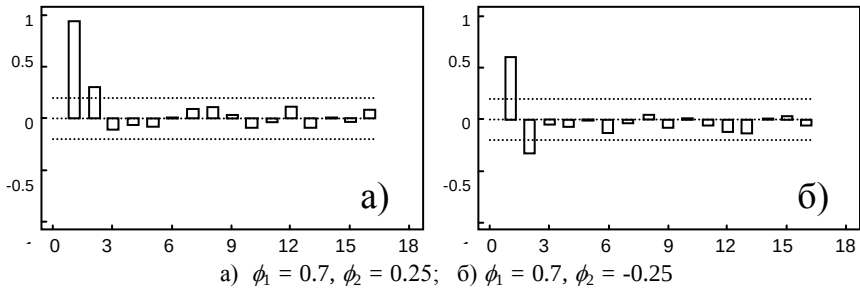


Рисунок 1.12 – Вибіркова ЧАКФ AR(2) процесів

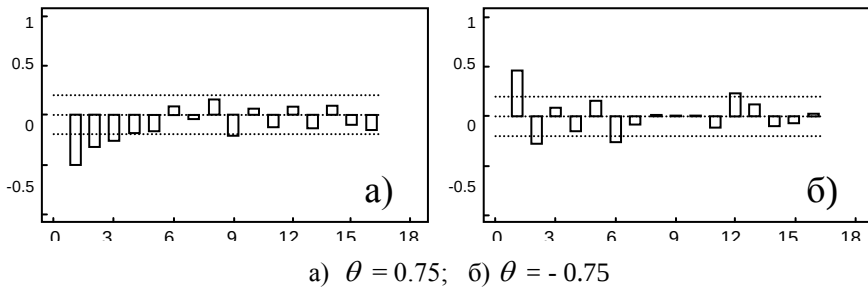


Рисунок 1.13 – Вибіркова ЧАКФ МА(1) процесів

1.6.2 Запитання для самоперевірки

1.6.2.1 Які цілі ставляться при вивченні ряду, отриманого з вихідного часового ряду після виключення детермінованої компоненти?

1.6.2.2 Які особливості корелограми часового ряду, що містить тренд?

1.6.2.3 Які особливості корелограми часового ряду, що містить сезонну компоненту?

1.6.2.4 Які особливості корелограми білого шуму?

1.6.2.5 Які особливості корелограми процесів ковзного середнього?

1.6.2.6 Які особливості корелограми процесів авторегресії?

1.6.2.7 В чому призначення часткової автокореляційної функції?

1.6.2.8 Які особливості часткової автокореляційної функції про-

цесів ковзного середнього?

1.6.2.9 Які особливості часткової автокореляційної функції процесів авторегресії?

2 ПЕРЕЛІК ЛАБОРАТОРНИХ РОБІТ

2.1 Генерація та аналіз часових рядів у математичному пакеті MathCAD.

2.2 Аналіз часових рядів у статистичному пакеті TableCurve.

2.3 Аналіз часових рядів у статистичному пакеті SPSS.

2.4 Підготовка наукових публікацій насичених математичними рівняннями за допомогою спеціалізованого редактора LaTeX.

3 КОНТРОЛЬНІ ЗАВДАННЯ

3.1 Загальні вказівки

По даній дисципліні виконується 1 контрольна робота.

Контрольна робота базується на вивченні розділів 1.1 - 1.6 робочої програми дисципліни, та виконанні лабораторних робіт №1-3.

3.2 Завдання на контрольну роботу

У контрольній роботі необхідно:

- згенерувати часові ряди для всіх приведених нижче моделей,
- побудувати їх графічне уявлення,
- розрахувати та проаналізувати числові характеристики 3 різних часових рядів (їх моделі трендів та випадкової складової обираються з таблиці 3.1 згідно варіанту).

Таблиця 3.1 –Варіанти завдання

Варіант №	Модель		Варіант №	Модель	
	тренд	вип.складова		тренд	вип.складова
1	1	1	11	1	1
	2	2		3	2
	3	3		5	3
2	4	1	12	1	1
	5	2		3	2
	6	3		5	3
3	1	2	13	2	2
	2	3		4	3
	3	1		6	1
4	4	2	14	2	2
	5	3		4	3
	6	1		6	1
5	1	3	15	1	3
	2	2		2	2
	3	1		4	1
6	4	3	16	1	3
	5	2		2	2
	6	1		4	1
7	1	2	17	2	2
	2	1		5	1
	3	3		6	3
8	4	2	18	2	2
	5	1		5	1
	6	3		6	3
9	1	3	19	1	3
	2	1		3	1
	3	2		4	2
10	4	3	20	1	3
	5	1		3	1
	6	2		6	2

Часовий ряд складається з двох компонент:

$$x_t = tr_t + \varepsilon_t,$$

де тренд задається наступними моделями:

- $tr_t = a + b \cdot t,$
- $tr_t = a + b \cdot t + c \cdot t^2 + d \cdot t^3 + \dots,$
- $tr_t = e^{a+bt},$
- $tr_t = \frac{a}{1+b \cdot e^{-ct}},$
- $\log tr_t = a - b \cdot r^t, r \in (0,1),$
- $tr_t = a \cdot \cos(\varpi \cdot t + \phi),$

а випадкова складова задається наступними моделями:

- $\varepsilon_t \in N(0, \sigma^2)$ - білий шум (нормальне розподілення з математичним очікуванням 0, дисперсією σ^2),
- $\varepsilon_t = \alpha_t + \theta \cdot \alpha_{t-1}, \quad \alpha_t \in N(0, \sigma^2)$ - ковзне середнє,
- $\varepsilon_t = \mu \cdot \varepsilon_{t-1} + \alpha_t, \quad \alpha_t \in N(0, \sigma^2)$ - процес авторегресії.

Побудувати графіки часового ряду x_t , тренду tr_t , випадкової складової ε_t , вибіркової автокореляційної функції $\bar{r}(k)$ (корелограму).

Розрахувати математичне очікування $\bar{\varepsilon}$ випадкової складової часового ряду.

Провести розрахунок дисперсії $\bar{\sigma}^2$ випадкової складової часового ряду.

розрахувати коефіцієнт Пірсона R и автокореляційну функцію $\bar{r}(k)$.

Математичне очікування $\bar{\varepsilon} = \frac{1}{n} \sum_{t=1}^n \varepsilon_t$,

дисперсія $\bar{\sigma}^2 = \frac{1}{n} \sum_{t=1}^n (\varepsilon_t - \bar{\varepsilon})^2$,

коефіцієнт Пірсона

$$R = \frac{\sum_{t=1}^n (x_t - \bar{x})(tr_t - \bar{tr})}{\sqrt{\sum_{t=1}^n (x_t - \bar{x})^2} * \sqrt{\sum_{t=1}^n (tr_t - \bar{tr})^2}},$$

$$\text{де } \bar{x} = \frac{1}{n} \sum_{t=1}^n x_t, \quad \bar{tr} = \frac{1}{n} \sum_{t=1}^n tr_t.$$

Коефіцієнт Пірсона показує ступень статистичної залежності часового ряду і тренду: чим ближче цей коефіцієнт до 1, тим краще підібраний тренд.

Автокореляційна функція

$$\bar{r}(k) = \frac{\sum_{t=1}^{n-k} (\varepsilon_t - \bar{\varepsilon}) \cdot (\varepsilon_{t+k} - \bar{\varepsilon})}{\sum_{t=1}^n (\varepsilon_t - \bar{\varepsilon})^2}; \quad k = 1, 2, 3, \dots, n/4.$$

Вона показує залежність між елементами випадкової складової. Якщо всі $\bar{r}(k)$ близькі до 0 ($|\bar{r}(k)| < 0.2$), то вважаємо всі елементи

випадкової складової незалежними, тобто ми маємо справу з білим шумом.

В контрольній роботі застосовується математичний пакет MathCad, що надає повний набір функціональних можливостей та забезпечує необхідну точність розрахунків.

Варіанти

3.3 Приклад виконання завдання

Генеруємо часовий ряд $x_t = tr_t + \varepsilon_t$ для лінійного тренда

$$tr_t = a + b \cdot t$$

і випадкової складової у вигляді білого шуму $\varepsilon_t \in N(0, \sigma^2)$.

Присвоїмо початкові значення і побудуємо графіки тренду та часового ряду:

$$a := 2$$

$$b := 5$$

$$\sigma := 10$$

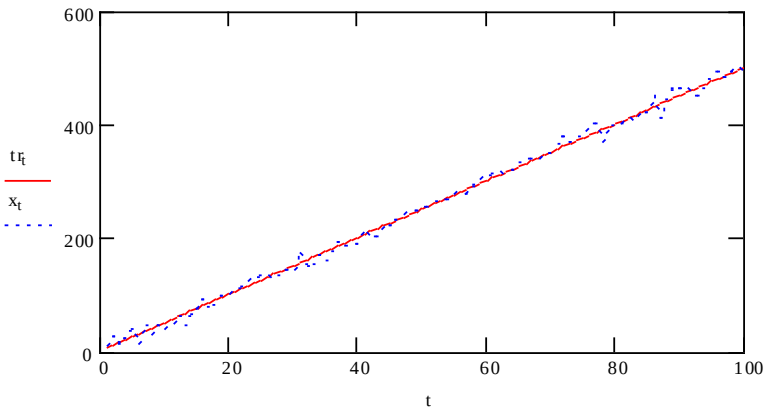
$$n := 100$$

$$t := 1..n$$

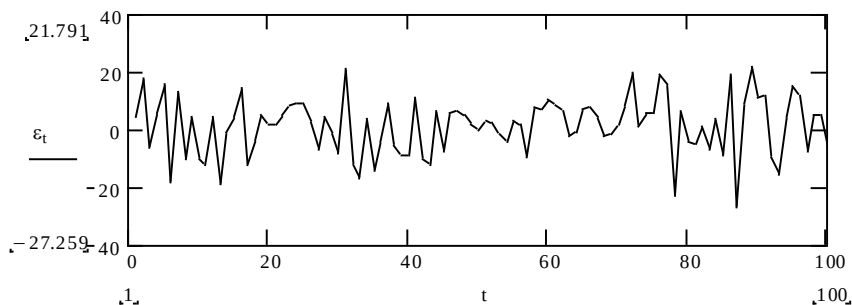
$$tr_t := a + b \cdot t$$

$$\varepsilon := \text{rnorm}(n, 0, \sigma)$$

$$x_t := tr_t + \varepsilon_t$$



Графік випадкової складової має вигляд



Оцінимо математичне очікування $\varepsilon_1 = M(\varepsilon)$ і дисперсію $D\varepsilon$ випадкової складової ε_t

$$\varepsilon_1 := \frac{1}{n} \cdot \sum_{t=1}^n \varepsilon_t$$

$$D\varepsilon := \frac{1}{n} \cdot \sum_{t=1}^n (\varepsilon_t - \varepsilon_1)^2$$

$$\varepsilon_1 = 1.162$$

$$D\varepsilon = 98.535$$

$$\sqrt{D\varepsilon} = 9.926$$

Знайдемо коефіцієнт Пірсона R

$$x_1 := \frac{1}{n} \cdot \sum_{t=1}^n x_t$$

$$tr_1 := \frac{1}{n} \cdot \sum_{t=1}^n tr_t$$

$$x_1 = 252.995$$

$$tr_1 = 254.5$$

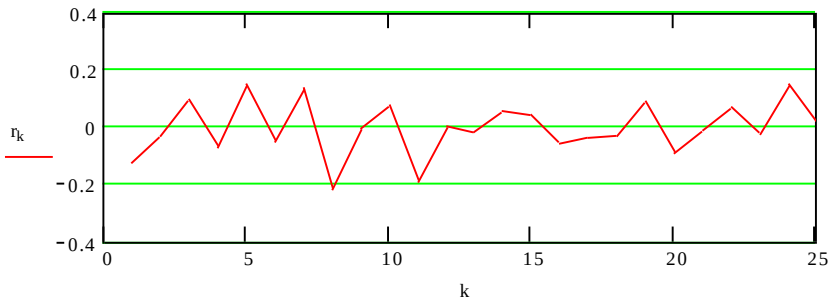
$$R := \frac{\sum_{t=1}^n (x_t - \bar{x}) \cdot (tr_t - \bar{tr})}{\sqrt{\sum_{t=1}^n (x_t - \bar{x})^2 \cdot \sum_{t=1}^n (tr_t - \bar{tr})^2}}$$

$$R = 0.998$$

Побудуємо графік автокореляційної функції r_k – корелограму

$$k := 1.. \frac{n}{4}$$

$$r_k := \frac{\sum_{t=1}^{n-k} (\varepsilon_t - \bar{\varepsilon}) \cdot (\varepsilon_{t+k} - \bar{\varepsilon})}{\sum_{t=1}^n (\varepsilon_t - \bar{\varepsilon})^2}$$



4 ПИТАННЯ ДО ІСПИТУ

- 4.1 Поняття часового ряду.
- 4.2 Аналіз часових рядів та його розділи.
- 4.3. Мета, етапи і методи аналізу часових рядів.
- 4.4 Детермінована і випадкова складові часового ряду.
- 4.5. Тренд, сезонна і циклічна компоненти.
- 4.6. Моделі тренда.
- 4.7. Моделі випадкової компоненти.
- 4.8. Числові характеристики часових рядів.
- 4.9. Процеси, стаціонарні в широкому змісті.
- 4.10. Оцінки числових характеристик часових рядів.
- 4.11. Порядок аналізу часових рядів.
- 4.12. Графічні методи аналізу часових рядів.
- 4.13. Виділення тренда методом найменших квадратів.
- 4.14. Виділення тренда за допомогою простих різницевих операторів.
- 4.15. Виділення сезонних ефектів.
- 4.16. Метод ковзних середніх.
- 4.17. Сезонні різницеві оператори.
- 4.18. Логарифмічне перетворення шкали.
- 4.19. Перетворення Боксу-Коксу.
- 4.20. Мета і методи аналізу структури стаціонарного часового ряду.
- 4.21. Графіки корелограми при наявності тренда і сезонних коливань.
- 4.22. Корелограма білого шуму.
- 4.23. Корелограма процесів ковзного середнього.
- 4.24. Корелограма процесів авторегресії.
- 4.25. Вибіркова часткова автокореляційна функція.
- 4.26. Графік часткової автокореляційної функції процесів авторегресії першого порядку.
- 4.27. Графік часткової автокореляційної функції процесів авторегресії другого порядку.
- 4.28. Графік часткової автокореляційної функції процесів ковзного середнього.

ЛІТЕРАТУРА

1. Тюрин, Ю.Н., Макаров, А.А. Анализ данных на компьютере: учебное пособие / Ю.Н. Тюрин, А.А. Макаров. – Под ред. В.Э. Фигурнова. - М.: ИНФА-М, 2003.- 544 с.
2. Горяинова, Е. Р., Панков, А. Р., Платонов, Е. Н. Прикладные методы анализа статистических данных: учеб. пособие / Е. Р. Горяинова, А. Р. Панков, Е. Н. Платонов ; Нац. исслед. ун-т «Высшая школа экономики». — М.: Изд. дом Высшей школы экономики, 2012. - 310 с.
3. Калинина, В. Н., Соловьев, В.И. Анализ данных. Компьютерный практикум: учебное пособие / В.Н. Калинина, В.И. Соловьев. - М.: КНОРУС, 2017. - 166 с.
4. Барсегян, А. А. Анализ данных и процессов: учеб. пособие / А. А. Барсегян, М. С. Куприянов, И. И. Холод, М. Д. Тесс, С. И. Елизаров. - 3-е изд., перераб. и доп. - СПб.: БХВ-Петербург, 2009. - 512 с.
5. Рыжиков, Ю.И. Решение научно-технических задач на персональном компьютере / Ю.И. Рыжиков. - СПб: КОРОНА принт, 2000.- 270 с.
6. Гайдышев, И. Анализ и обработка данных: Специальный справочник / И. Гайдышев. - СПб: Питер, 2001.- 752 с.
7. Бююль, А., Цёфель, П. SPSS: Искусство обработки информации. Анализ статистических данных и восстановление скрытых закономерностей: Пер. с нем. / А. Бююль, П. Цёфель. - СПб.: ДиаСофтЮП, 2005. - 608 с.
8. Кнут, Дональд, Э. Все про TEX: Перевод с англ. /Д. Кнут. – М.: Издательский дом «Вильямс», 2003. – 560 с.
9. Львовский, С.М. Набор и верстка в пакете LATEX /С.М. Львовский. - М.: МЦНМО, 2014. - 400 с.
10. Тутубалин, В.Н. Теория вероятностей и случайных процессов / В.Н. Тутубалин. - М.: Изд-во МГУ, 1992. - 400 с.
11. Кирьянов, Д.В. Самоучитель MathCad 13 / Д.В. Кирьянов. - СПб: БХВ-Петербург, 2006.- 528 с.
12. Методичні вказівки до виконання лабораторних робіт з дисципліни "Комп'ютерна техніка в наукових дослідженнях" для студентів напряму 123 "Комп'ютерна інженерія" спеціалізацій (освітніх програм) "Комп'ютерні системи та мережі", "Спеціалізовані комп'ютерні системи" усіх форм навчання/ Укл.: К.М. Касьян, М.М. Касьян.- Запоріжжя: ЗНТУ, 2017. - 38с.