Благој Делипетрев; Наташа Стојковиќ; Зоран Утковски

МОДЕЛИРАЊЕ И СИМУЛАЦИИ - СКРИПТА

Штип, 2015

Благој Делипетрев; Наташа Стојковиќ; Зоран Утковски

МОДЕЛИРАЊЕ И СИМУЛАЦИИ - СКРИПТА

Автори: Доц. Д-р. Благој Делипетрев, М-р. Наташа Стојковиќ, Доц. Д-р. Зоран Утковски

МОДЕЛИРАЊЕ И СИМУЛАЦИИ - СКРИПТА

Рецензенти: Проф. Д-р. Владо Гичев, Вон. Проф. Д-р. Татјана Атанасова-Пачемска

Лектор: Слаѓан Спасовски

Уредник:

Техничко уредување:

Издавач: Факултет за Информатика
Универзитет „Гоце Делчев” - Штип

Објавено во е-библиотека:
https://e-lib.ugd.edu.mk

DOI:

CIP - Каталогизација во публикација

Национална и универзитетска библиотека

"Св. Климент Охридски", Скопје

519.85/.87:004.94(075.8)

УНИВЕРЗИТЕТ „ГОЦЕ ДЕЛЧЕВ” – ШТИП

ФАКУЛТЕТ ЗА ИНФОРМАТИКА

Доц. Д-р. Благој Делипетрев, М-р. Наташа Стојковиќ, Доц. Д-р. Зоран Утковски

МОДЕЛИРАЊЕ И СИМУЛАЦИИ

Скрипта

Штип, 2015

ПРЕДГОВОР

Скриптата „Моделирање и симулации" е наменета за студентите од прв циклус студии на Факултетот за информатика за предметот Моделирање и симулации кој се слуша во петтиот семестар, со фонд на часови 2+1+1, на студиските програми Компјутерски науки и Компјутерско инженерство и технологии.

Во ***првата*** глава е даден вовед во симулациите, се адресираат нивните предности и недостатоци, и се анализираат случаите кога тие се соодветна алатка за моделирање на системи. Покрај другото, објаснети се и сите чекори кои се потребни во процесот на симулација.

Во ***втората*** глава се разгледуваат техники за генерирање на случајни броеви. Дадени се средно-квадратниот метод, методот на линеарна конгруентност, адитивниот конгруентен метод и комбинираниот линеарно - конјугиран метод. Исто така, се разгледуваат и тестови за случајни броеви, тестови за фреквенција, тестови за монотоност и тестови за автокорелација.

Во ***третата*** глава се разгледуваат методите за генерирање на случајни променливи: техниката со инверзна трансформација, методот на конволуција и техниката за прифаќање и одбивање. Дадени се и постапките за генерирање на експоненцијална, рамномерна, триаголна, Вејбулова, геометриска, нормална, Ерлангова и Пуасонова распределба.

Во ***четвртата*** глава се анализира моделирањето на влезните податоци. Дадени се начини на собирање на влезни податоци и методи за идентификација на распределбите на податоците. Откако ќе бидат избрани распределбите на податоците, во следното поглавје се оценуваат нивните параметри.

Во ***петтата*** глава се разгледува една од најзначајните и најтешките задачи во развојот на моделот-верификацијата и валидацијата на моделот што се симулира. Дадени се различни критериуми за утврдување на валидноста на моделот.

Во ***шестата*** глава се воведуваат системи за масовно опслужување. Оваа глава е поделена на три подглавја:

1.1 Вовед во теоријата на масовно опслужување;

1.2 Прости системи за масовно опслужување и Маркови процеси;

1.3 Процеси на раѓање и умирање;

Во *првото* подглавје се објаснети поимите: популација од клиенти, влезен проток, редица на чекање, дисциплини на чекање и опслужување. Исто така, дадена е и Кендаловата класификација на системите за масовно опслужување, а се дефинира и прост поток на настани и време на опслужување.

Во *второто* поглавје од оваа глава се разгледуваат прости системи за масовно опслужување, Маркови процеси и Маркови вериги.

Во *третото* поглавје се објаснети процесите на раѓање и умирање. Објаснети се системите од облик: $M/M/n/\infty$ и $M/M/n/m$и затворените системи за масовно опслужување.

СОДРЖИНА

1. СИМУЛАЦИИ

1.1 Вовед во симулации

Симулацијата е имитација на операциите на процесите од реалниот свет или некој систем во тек на време. Однесувањето на системот во текот на времето се проучува со развој на таканаречен **симулационен модел.** Вообичаено е моделите да бидат претставени како множество од претпоставки кои се однесуваат за самиот систем. Овие претпоставки се изразени преку математички логички или симболички релации помеѓу ентитетите кои всушност се објекти од интерес во системот.

Симулациите може да бидат користени во стадиумот на дизајнирање пред системот да биде изграден, но и на веќе изградени системи со цел да се утврди дали потенцијалните промени ќе имаат влијание врз перформансите на системот. Значи, симулациите може да се користат или како алатка за да се предвиди како ќе влијаат промените во некој постоечки систем, или како алатка за предвидување на перформансите на некој нов систем под различно множество на услови.

Понекогаш моделот што се развива може да се реши математички. Тогаш решението може да биде добиено со користење на диференцијални равенки, теоријата на веројатност, алгебарски модели, или со други математички техники. Ова решение вообичаено се состои од еден или повеќе нумерички параметри кои се нарекуваат **мерка на перформансите** на системот. Меѓутоа, повеќето реални системи не можат да се решат математички бидејќи се премногу комплексни. Во овие случаи нумерички- базираните симулации може да бидат искористени да го имитираат однесувањето на системот во текот на времето. Вообичаено, податоците што се потребни за симулирање на некој систем се собираат преку набљудување на системот, додека податоците што се генерираат со симулациите се користат за оценување на перформансите на системот.

1.1.1 **Предности и недостатоци на симулациите**

Симулацијата е привлечна алатка за клиентите бидејќи со неа може да се предвиди како ќе се однесува реалниот систем уште додека е во стадиумот на дизајн. Истовремено, излезните резултати може директно да кореспондираат со излезите што ќе бидат добиени од реалниот систем. Од овие причини симулациите често се користат во техниката како соодветна алатка за развој на системот. Дополнително, при симулирањето може да се менуваат влезните податоци и да се набљудува однесувањето на системот при овие промени.

Секако, како и секоја друга метода, и симулацијата има свои предности и недостатоци.

Предности на симулациите се:

1. Нови политики, оперативни процедури, правила за решавање, информациони потоци може да бидат развиени без да се прекинуваат излезните операции на реалниот систем.
2. Нов хардверски дизајн, физички нацрти, транспортни системи може да бидат тестирани без потребни ресурси.
3. Може да се тестираат хипотези за тоа како и зошто некои феномени се случуваат.
4. Времето може да биде скратено или продолжено. Со скратување или продолжување на времето симулацијата овозможува забрзување или забавување на феноменот, за да може истиот повнимателно да се проучи.
5. Менаџерите често сакаат да знаат зошто даден феномен се појавува во реалниот систем. Со симулацијата може да се даде одговор на ова

прашање така што се врши реконструкција на сцената и се изведуваат микроскопски испитувања на системот за да се одреди зошто феноменот се појавил. Ова не може да се постигне во реалниот систем бидејќи тој не може целосно да се контролира.

6. Една од најголемите предности од користењето на симулациониот софтвер е тоа што кога еднаш ќе се развие валиден симулационен модел, може да се проучуваат нови оперативни процедури или методи без да се експериментира со реалниот систем. Модификациите се внесени во моделот и се набљудуваат влијанијата на овие промени на компјутер наместо во реалниот свет.
7. Модерните фабрики и организации се доста сложени, така што е невозможно да се знаат сите интеракции што постојат во даден момент. Симулациите овозможуваат подобро да се разберат интеракциите меѓу променливите кои ги сочинуваат таквите сложени системи.

Недостатоци на симулациите се:

1. Градењето на моделот бара специјален тренинг, тоа е уметност што се учи со тек на време и преку искуство. Ако два модели на систем се изградени од две различни индивидуи, тие може да имаат сличности, но многу е мала веројатноста да бидат исти.
2. Резултатите од симулацијата може да бидат тешки за интерпретација. Често, излезите од симулациите се случајни променливи (тие вообичаено се базираат на случајни влезови), и може да биде тешко да се забележи дали едно набљудување е резултат од внатрешните релации во системот или пак е случајно добиено.
3. Моделирањето на симулациите и нивната анализа може да одземе многу време и да биде многу скапо.
4. Немањето на доволно ресурси за анализа и моделирање може да резултира во симулационен модел што не ја задоволува целта за која е изграден.
5. Симулацијата да се користи во случаи кога аналитичкото решение е можно, па дури и препорачливо.

1.1.2 Систем и компоненти на системот

За моделирање на системот потребно е да се разберат концептите на системот и однесувањето на системот. Системот се дефинира како група од објекти кои се групирани на одреден начин се со цел да извршат некоја задача. Еден ваков пример е системот за производство на автомобили. Машинските елементи и работниците оперираат заедно со заедничка цел производство на високо квалитетни мотори. На системот често влијаат промените кои се случуваат надвор од системот. Овие промени се случуваат во околината на системот. Во моделирањето на системот потребно е да се одреди која е границата меѓу системот и неговата околина. Оваа одлука често зависи од целта на проучувањето. На пример ако разгледуваме како систем една фабрика, факторите кои го контролираат доаѓањето на нарачките може да се сметаат за надворешни фактори. За да се разбере и анализира системот, треба да бидат дефинирани следниве поими.

- **Ентитет** е објект кој е значаен за системот;
- **Атрибут** е својство на ентитетот**;**
- **Активност** претставува временски период со специфицирана должина;
- **Состојба** на системот се смета како множество од променливи со чија помош системот може да се опише во било кој момент;

- **Настан** се дефинира како моментално случување што може да ја промени состојбата на системот. За активностите и настаните кои се појавуваат внатре во системот се користи терминот **ендогени** или **внатрешни** активности односно настани, додека пак за активностите и настаните од околината на системот се користи терминот **егзогени** или **надворешни** настани, односно атрибути.

Понекогаш од интерес е со проучувањето на системот да се разбере релацијата помеѓу компонентите и да се предвиди како системот ќе се однесува под нови околности. Од големо значење е што новиот систем сè уште не мора да постои, тој може да биде во хипотетичка форма или во стадиум на дизајн. Затоа битно е да се нагласи дека проучувањата на системот се поврзани со моделот на системот.

Моделот е дефиниран како презентација на системот заради негово проучување. Во повеќето случаи потребно е да се земаат само оние аспекти од системот кои влијаат на проблемот кој се проучува. Овие аспекти се претставени во моделот на системот. Од друга страна моделот треба да биде детален за да се добие правилен заклучок за реалниот систем.

Моделите може да се поделат на **математички** и **физички**. Математичките модели користат симболичка нотација и математички равенства за претставување на системот, додека физичкиот модел е реално направен модел. Симулациониот модел е посебен вид на математички модел.

Симулационите модели според различни критериуми може да се поделат на: **статички** и **динамички**, **детерминистички** и **стохастички**, и **дискретни** и **непрекинати**. Статичкиот модел понекогаш наречен Метод на Монте Карло го претставува системот во даден момент на време. Динамичките симулациони модели ги претставуваат промените на системите во тек на време.

Симулационите модели кои не користат случајни променливи се наречени детерминистички. Детерминистиките модели имаат познато множество на влезови кои резултираат во единствено множество на излези. Стохастичките модели пак имаат една или повеќе случајни променливи како влез. Случајниот влез води кон случаен излез. Бидејќи излезите се случајни тие може да се сметаат само за проценка на вистинските карактеристики на системот. Поради тоа кај стохастичките системи главни карактеристики се: среден број на клиенти кои чекаат, средно време на чекање на клиент итн.

Дискретните модели се оние модели кај кои променливите кои ја одредуваат состојбата на системот се менуваат во дискретно множество точки од времето. Непрекинати системи се оние модели кај кои променливите кои ја одредуваат состојбата на системот се менуваат непрекинато во текот на времето.

Меѓутоа, значајно е да се нагласи дека дискретниот симулационен модел не се користи секогаш за моделирање на дискретен систем. Исто така и непрекинатиот симулационен модел не се користи секогаш за моделирање на непрекинат систем. Изборот на кој од овие два модели да се користи зависи од карактеристиките на систетот и од предметот на проучувањето.

1.1.3 Чекори во процесот на симулација

Формулирање на проблемот: Секое проучување започнува со дефинирање на проблемот. Дефиницијата на проблемот треба да се разгледа од страна на аналитичарот за да се осигури дека јасно го разбрал проблемот кој е опишан. Ако проблемот е дефиниран и разбран од страна на аналитичарот важно е извршителите на проектот да се согласат со таа формулација. Понекогаш може да се случи проблемот да мора да биде преформулиран во текот на проучувањето.

Собирање на податоци: Постои постојана врска меѓу конструкцијата на моделот и собирањето на потребните влезни податоци. Бидејќи комплексноста на моделот може да се промени потребните податоци исто така можат да се променат. Бидејќи собирањето на податоците одзема поголем дел од вкупното време потребно за да се изведе симулацијата, потребно е да се започне колку што е можно порано, вообичаено заедно со почетните фази на градењето на моделот. Целта на истражувањето го диктира и видот на податоците што треба да бидат собрани.

Развивање на моделот: Бидејќи најголем дел од системите во реалниот свет резултираат со модели кои бараат складирање на големо количество податоците и бараат многу пресметувања, моделот треба да бидат внесени во компјутерски препознатлив формат. Обично се употребува терминот „Програм“ иако во многу случаи возможно е да се постигне посакуваниот резултат со многу малку или без никакво кодирање.

Верификација и проверка на валидноста на моделот: Верификацијата се однесува на компјутерската програма која е подготвена за симулирачкиот модел. Дали компјутерскиот модел функционира правилно? Комплексните модели е потешко, ако не и невозможно, да се преведат успешно без при тоа да биде потребно дебагирање во поголема мера. Ако влезните параметри и логичката структура на моделот се точно претставени во компјутерот верификацијата е завршена. Во најголем дел логичкото расудување се употребува за да се заврши овој чекор. Проверка на валидноста претставува одредување дека моделот е точна претстава на реалниот систем. Тоа вообичаено се постигнува со калибрација на моделот, што е итеративен процес на споредување на моделот со моменталното однесување на системот, со цел да се подобри моделот. Процесот се повторува се додека точноста на моделот не се оцени како прифатлива.

Експериментирање со моделот и оптимизација: Алтернативите што треба да бидат симулирани треба да бидат детерминирани. Често пати, одлуката кои алтернативи да се симулираат може да биде резултат на повеќекратно извршување и анализирање на симулацијата. За секој дизајн на системот што е симулиран, треба да се донесат решенија во врска со должината за иницијализацијата, должината на повторувањето на симулацијата и бројот на повторувањето на симулацијата.

Имплементација: Успехот во фазата на имплементацијата зависи од тоа колку добро претходните чекори се изведени. Исто така зависи и од тоа колку темелно аналитичарот го вклучил и крајниот корисник на моделот за време на целиот процес на симулација. Ако корисникот на моделот бил темелно вклучен и ја разбира природата на моделот и неговите излезни податоци тогаш стремежот кон имплементацијата ќе се зголеми. Во спротивно, доколку моделот и претпоставките засновани врз него не се правилно поврзани, имплементацијата веројатно ќе страда, без оглед на валидноста на симулацискиот модел.

2. ГЕНЕРИРАЊЕ НА СЛУЧАЈНИ БРОЕВИ

Случајните броеви се неопходно основниот составен дел во симулирањето на речиси сите дискретни системи. Скоро сите програмски јазици имаат процедури, функции или објекти за генерирање на случајни броеви. Да напоменеме дека постојат и специјални јазици за симулирање кои генерираат случајни броеви што се користат за генерирање на времиња за појавување на настаните и за генерирање на други случајни променливи.

2.1 Својства на случајните броеви

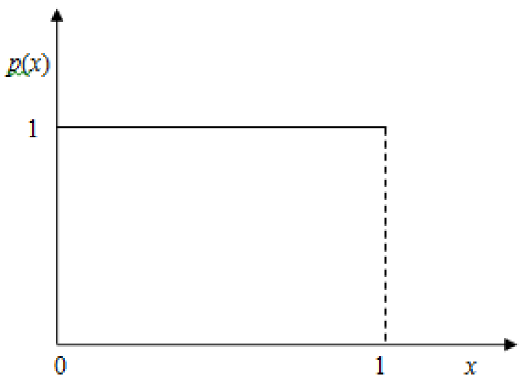
Низата од случајни броеви R_1, R_2, … мора да задоволува две својства:

- Броевите да бидат рамномерно распределени на интервалот[0,1],
- Броевите да бидат независни.

Секој случаен број R_i, i = 1,2 е независен примерок, поточно независна реализација на случајната променлива која има рамномерна распределбана [0,1].
Густината на распределба на рамномерната распределба на [0,1] е дадена со:

$$p(x) = \begin{cases} 1, & 0 \le x \le 1 \\ 0, & \text{инаку} \end{cases}$$

и се претставена на **Фиг. 2.1.**



Фиг. 2.1 Густина на распределба на случајни броеви

Очекуваната вредност (математичкото очекување) на секој R_i е дадена со равенството:

$$ER = \int_{-\infty}^{\infty} xp(x)dx = \int_0^1 xdx = \frac{x^2}{2}\Bigg|_0^1 = \frac{1}{2}$$

и дисперзијата е дадена со:

$$DR = \int_0^1 x^2 dx - (ER)^2 = \frac{x^3}{3}\Bigg|_0^1 - \left(\frac{1}{2}\right)^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}$$

Својства на низите кои ги задоволуваат условите за независност и рамномерност се следниве:

1. Ако интервалот (0,1) е поделен во *n* класи, или подинтервали со еднаква должина,очекуваниот број на набљудувања во секој интервал е *N*/*n* каде *N* е вкупниот број од набљудувања.

2. Веројатноста да се генерира вредност во било кој подинтервал не не зависи од претходно генерираните вредности.

Треба да се нагласи дека доколку се бараат вистински случајни броеви тоа не може да се постигне со користење на некој алгоритам. Остварување на вистински случајни броеви е можно со користење на некој процес кој сам за себе има случајно однесување. Примери за такви процеси се следниве:

- Термален шум (*thermal noise*);
- Атмосферски шум (*atmospheric noise*);
- Радиоактивно распаѓање (*Radioactive decay*).

Дадените примери бараат набавка на опрема или соодветен сервис. Затоа се оди кон користење на алгоритми, кои во кратко време можат да обезбедат голем број случајни броеви.

2.1.1 Генерирање на „псевдо"- случајни броеви

Случајни броеви добиени со некој алгоритам се нарекуваат **„псевдо" случајни броеви** (***pseudo-random numbers*)** и тие се всушност добиени со детерминистичка постапка.

Самиот збор „*pseudo*" значи лага (невистина),од каде што следува дека ќе генерираме „лажни" случајни броеви. Целта на било кој алгоритам за генерирање на псеудо-случајни броеви е да се добие низа од броеви помеѓу 0 и 1, кои ќе ги имитираат својствата на рамномерност и непрекинатост колку што е можно тоа. Општи барања кои треба да ги задоволуваат псеудо-случајни броеви се:

- Генерираните случајни броеви да се рамномерно распределени над интервалот [0,1].
- Секоја реализација да е независна од претходната, односно да нема корелација помеѓу нив.
- Можна е репликација (повторување на иста низа).
- Што подолг период на повторување на низа.
- Брзина, ефикасни и брзи генератори.
- Ефикасно користење на меморијата, т.е. да не е потребен голем мемориски простор.

Кога се генерираат псевдо-случајни броеви може да се појават следниве проблеми или грешки:

- Генерираните броеви да не бидат рамномерно распределени.
- Генерираните броеви може да примаат дискретни вредности наместо непрекинати.
- Аритметичката средина на генерираните броеви да биде многу голема или многу мала.

- Дисперзијата (варијансата) на генерираните броеви да биде многу голема или многу мала.
- Може да постои зависност помеѓу генерираните броеви, односно:
  - Автокорелација помеѓу броевите.
  - Броевите што следуваат да бидат многу поголеми или помали од соседните броеви.
  - Неколку броеви со вредност поголема од математичкото очекување може да бидат следени со броеви помали од математичкото очекување.

Особините на рамномерност и непрекинатост може да бидат тестирани и за оваа цел постојат генератори.

Вообичаено случајните броеви се симулираат на компјутер како дел од симулациите. За генерирање на случајните броеви може да бидат користени нумерички методи. Овие методи или рутини треба да ги задоволуваат следниве услови:

- Рутината треба да биде брза. Поединечните пресметки се евтини, но симулацијата може да побарува стотици или илјадници случајни броеви. Вкупните трошоци може да влијаат врз изборот на ефикасен метод за генерирање на случајни броеви.
- Рутините треба да бидат преносливи на различни компјутери, и идеални за различни програмски јазици. Со тоа се постигнува програмите за симулирање да дават резултати било каде да се извршуваат.
- Рутината треба да генерира броеви со многу долг период. Должината на периодот ја претставува должината на низата случајни броеви, пред да започне повторување на броевите по истиот редослед како претходниот редослед.
- Случајните броеви треба да може да се реплицираат, односно да има повторување на истата низа. Тргнувајќи од почетната точка треба да постои можност за генерирање на истото множество случајни броеви независно од системот.
- За крај, како најзначајно ќе го истакнеме својството кое вели дека случајните броеви треба да бидат добра апроксимација на низа од броеви кои ги задоволуваат својствата за независност и рамномерност.

2.1.2 Техники за генерирање на случајни броеви

2.1.2.1 Средно квадратен метод

Во овој метод се тргнува така што почетниот број X_0 е четирицифрен број. Овој број се квадрира и доколку се добие осум цифрен број е во ред, доколку тоа не е случај се додаваат нули од лево за да се добие осум цифрен број. Се вадат средните четири цифри за да се добие бројот X_1. Истата постапка потоа се повторува и за бројот X_1. За да се добијат случајни броеви во интервалот од (0,1), броевите X_0, X_1,… се делат со 1000.

Пример 2.1.

- X_0=1234 ; X_0^2= 1522756, бидејќи се добива седум цифрен број се додава 0 од лево и се добива бројот 01522756. Со вадење на средните четири цифри се добива бројот 5527. Истата постапка се повторува и за него. На крај броевите се делат со 1000. Подолу се дадени неколку случајни броеви генерирани на овој начин:

0.1234, 0.5227, 0.3215, 0.3362, 0.3030, 0.1809, 0. 2724, 0.4201, 0.6484, 0.0422, 0.1780, 0.1684, 0.8361, 0.8561, 0.2907,...

- Ако се почне со X_0 =2345 се добива следнава низа:

0.2345, 0.4990, 0.9001, 0.180, 0.324, 0.1049, 0.1004, 0.0080, 0.0040,...
нулите зад децималната запирка не се отстрануваат

- Ако се почне со X_0=2100се добива низата

0.2100, 04100, 08100, 0.6100, 04100 која се повторува.

2.1.2.2 Метод на линеарна конгруентност

Едноставен метод за генерирање на псевдо-случајни броеви, предложен од *Lehmer*[1951], зависи од изборот на параметри, и дава добри резултати кога се користи правилно. Со овој метод се генерира низата од цели броеви X_1, X_2,… помеѓу 0 и m-1 кои ја задоволуваат следната рекурзивна релација:

$$X_{i+1}=(aX_i + c) \bmod m \text{ за } i = 1, 2, \ldots \quad (2.1)$$

каде X_0 е почетна вредност или семе, a е константен множител, c прираснување, и m е модул. Можни се следниве два случаи:

1. Ако $c \neq 0$, тогаш имаме **мешан линеарен конгруентен метод.**
2. Ако $c = 0$, тогаш имаме **мултипликативен конгруентен метод**.

Изборот на параметрите a, c, m и X_0 драстично влијае на статистичките својства и должината на периодот.

Пример 2.2. Со овој метод да се генерира низа случајни броеви соX_0 = 27, a=17, c=43 и m=100. Случајните цели броеви треба да бидат рамномерно распределени на множеството {0, 1, 2, ..., 99}. За добивање на случајните броеви од 0 до 1 се користи следнава трансформација:

$$R_i = \frac{X_i}{m}. \quad (2.2)$$

X_i и R_i се добиваат на следниот начин:

X_0=27,

X_1=(17·27+43)mod 100 = 502 mod 100 = 2,

$R_1 = \frac{2}{100}$ =0.02,

X_2=(2·27+43)mod 100 = 77 mod 100 = 77,

$R_2 = \frac{77}{100}$ =0.77,

X_3=(77·27+43) mod 100 = 1352 mod 100 = 52,

$R_3 = \frac{52}{100}$ = 0.52.

На овој начин се добива низа од случајни броеви R_1, R_2,… кои се приближно рамномерно распределени и независни. Се разгледуваат и некои други значајни особини на случајните броеви:

- Густина на распределба (*densitydistribution*), колку различни броеви може да се генерираат во дадениот интервал.
- Периодот на генераторот (колку броеви се генерирани пред да се повтори истата низа од броеви со генераторот). Пожелни е периодите да се поголеми.
- Едноставни и брзи пресметки, без комплексни операции.

Исто така да се забележи дека броевите што се генерираат со равенството $R_i = \frac{X_i}{m}$ може да примаат вредности од множеството $I = \{0, \frac{1}{m}, \ldots, \frac{m-1}{m}\}$, значи секој X_i е цел број од множеството {0,1,…,m-1}. Исто така треба да се напомене дека добиената низа е реализација на дискретна случајна променлива со множество вредности I наместо на непрекината случајна променлива со рамномерна распределба на интервалот [0,1]. Обично модулот m е многу голем број. Во многу симулациони јазици вообичаено е да се избира $m=2^{31}-1$ или $m=2^{48}$. Како што е кажано поголемиот период овозможува поголема густина во интервалот. Максимален период може да биде добиен со соодветен избор на параметрите X_0, a, c и m.

- За m степен на 2, т.е. $m= 2^b$, и $c\neq0$, најдолгиот можен период е $P= m=2^b$, што се постигнува доколку c е взаемно прост со m (НЗД(c, m)=1), и a=1+4k, каде k е цел број. Во овој случај се добива, голем период и пресметките не се предолги. Овој пример се користи како генератор во Java (JDK) (*JDKRandom.java, TestRandom.java*)
- За m степен на 2, т.е. $m = 2^b$, и $c = 0$, најдолгиот можен период е $P = \frac{m}{4} = 2^{b-2}$, што се постигнува доколку X_0 е непарен и a=3+8k или a=5+8k, за некое k=0,1,2… Пресметката е поефикасна, но периодот е помал. Се бара добра почетна вредност што бара добро познавање на постапката.

- За m прост број и c=0, најдолгиот можен период е $P=m-1$, што се постигнува ако множителот a, го има својството најмалиот цел број k за кој $a^k - 1$ е делив со m, е $k = m - 1$.

Пример 2.3. Со користење на мултипликативниот конгруентен модел, да се најде периодот на генераторот за a = 13, $m = 2^6$ = 64, и X_0=1,2,3,4. Решението е дадено во Табела 2.1.

i	X_i	X_i	X_i	X_i
0	1	2	3	4
1	13	26	39	52
2	41	18	59	36
3	21	42	63	20
4	17	34	51	4
5	29	58	23	
6	57	50	43	

7	37	10	47	
8	33	2	35	
9	45		7	
10	9		27	
11	53		31	
12	49		19	
13	61		55	
14	25		11	
15	5		15	
16	1		3	

Табела 2.1

Може да се забележи дека за X_0=1 и X_0=3 низата има период 16, додека за X_0=2 има период 8, а за X_0=4 има период 4. Во овој пример $m=2^6=64$, и c=0, па најдолгиот можен период е $P=\frac{m}{4}=16$, Тој период се добива со случај кога X_0 е напарен, во нашиот случај се добива за X_0=1 и X_0=3. Додека за парно X_0, 2 или 4 периодот е помал од максималниот. Да се забележиме дека a=3+5k, што е едно од барањата за добивање на максимален период. „Пукнатините“–растојанието во низата од случајни броеви е доста големо (на пр. 5/64 - 1/64=0.0625.) Генераторот во примерот 2.3, не е погоден за практична примена, бидејќи неговиот период е доста мал и броевите не се доволно густо распределени. Меѓутоа со овој пример се согледува значајноста од соодветниот избор на параметрите, a, c, m и X_0.

Пример 2.4. Нека $m = 10^2 = 100$, $a = 19$, $c = 0$ и $X_0 = 63$, се генерира следнава низа од случајни броеви:

X_0=63,

X_1=19 · 63 mod 100 =1197 mod 100 = 97,

X_2=19 · 97 mod 100 = 1843 mod 100 = 43,

X_3=19 · 43 mod 100 = 817 mod 100 =17.

Пример 2.5. Последниот пример е за практично користење. Тој е тестиран од *Learmonth u Lewis, 1973 u Lewis*,1969. Вредностите на a,c и m се така избрани така што со голема веројатност се добиваат посакуваните особини. Нека $a = 7^5 = 16807$, $m = 2^{31} – 1$ = 2 147 483 647 (прост број) и c = 0. Со ова се овозможува должината на периодот да биде $P= m$-1.

Се избира X_0=123 457. Првите неколку генерирани броеви се:

$X_1 = 7^5 \cdot$ 123 457 mod (2^{31}-1) = 2 074 941 799 mod (2^{31}-1),

X_1 = 2 074 941 799,

$R_1 = X_1 / 2^{31}$=0.9662,

$X_2=7^5\cdot$ 2 074 941 799 mod (2^{31}-1) = 559 972 160,

$R_2= X_2/2^{31}$=0.2607,

$X_3=7^5*559\ 972\ 160 \bmod (2^{31}-1)= 1\ 645\ 535\ 613$,

$R_3 = X_3/2^{31} = 0.7662$.

2.1.2.3 Адитивен конгруентен метод

Се генерира низа на цели броеви во интервалот од 0 до m-1 согласно рекурзивната равенка:

$$X_i = (X_{i-1} + X_{i-k}) \bmod m, \quad i=1,2,3\ldots \text{ за } k> 2$$

За да се добие низата од случајни броеви кои се рамномерно распределени на [0,1] треба да се избере:

$$R_i = \frac{X_i}{m},$$

максималната должина на период е m^k.

Недостатоци на овој метод:

За k = 2, за три последователни случајни броја
R_{i-2} , R_{i-1} , R_i никогаш нема да важи
$R_{i-2}<R_i<R_{i-1}$или$R_{i-1}<R_i<R_{i-2}$

Забелешка: прирамномернараспределбаидватаредоследиимаатверојатност 1/6.

2.1.2.4 Комбиниран линеарно - конјугиран метод

Генераторите на случајни броеви со период 2^{31}-1, опишан во пример 2.4, не е соодветен за сите апликации. Примери за вакви апликации се симулациите на реалните системи во кои стотици или илјадници настани мора да бидат симулирани за да се забележи еден неуспешен настан, или симулирањето на комлексни компјутерски мрежи во кои стотици корисници извршуваат илјадници програми. За овие цели се потребни генератори кои обезбедуваат многу долги периоди. Овие генератори се комбинација од два или повеќе мултипликативни конгруентни генератори и на таков начин се добиваат генератори со добри статистички својства и подолг период. Овој генератор се добива со комбинација на два или повеќе мултипликативни конгруентни генератори, и на таков начин се постигнуваат добри статистички својства и подолг период.

Ако $W_{i,1}$, $W_{i,2}$,. . . $W_{i,k}$ се било кои независни дискретни случајни променливи (не е потребно да се идентично распределени), но едниот од нив на пример $W_{i,1}$ да е рамномерно распределен на интервалот [0, m_1-2], тогаш:

$$W_i = \left(\sum_{j=1}^{k} W_{i,j}\right) \bmod m_1 - 1,$$

е рамномерно распределен на [0, m_1-2].

За да се види како овој резултат може да се користи за комбинирани генератори, нека $X_{i,1}$, $X_{i,2}$,…, $X_{i,k}$се i-тите излези од k различни мултипликативни конгруентни генератори, каде j-тиот генератор има модул m_j, и множителот a_j е избран така што периодот е m_j-1. Тогаш j-тиот генератор го произведува целиот број $X_{i,j}$ кој има приближно рамномерна распределба на [0, m_j-1] и $W_{i,j}$ има приближно рамномерно распределба на [0, m_j-2]. *L'Ecuyer* (1988) предлага генераторите да се комбинираат на следниов начин:

$$X_i = \left(\sum_{j=1}^{k} (-1)^{j-1} X_{i,j} \right) \bmod m_1 - 1,$$

и

$$R_i = \begin{cases} \dfrac{X_i}{m_1}, X_i > 0 \\ \dfrac{m_1 - 1}{m_1}, X_i = 0 \end{cases}.$$

Максималниот можен период за овој генератор е :

$$P = \frac{(m_1 - 1)(m_2 - 1)\cdots(m_k - 1)}{2^{k-1}}.$$

Пример 2.6 За 32-битен компјутер *L'Ecuyer* (1988) ја сугерира комбинацијата од k=2 генератори со m_1=2147483563, a_1=40014,m_2=2147483399 и a_1=40692. Постапката е дадена со следниот алгоритам:

1. Се избира $X_{1,0}$ од интервалот [1, 2147483562] за првиот генератор и $X_{2,0}$ од интервалот [1, 2147483398] за вториот генератор. Се поставува j = 0,

2. За секој генератор посебно се пресметува:

   $X_{1,j+1}$ = 40014 $X_{1,j}$ mod 2147483563
   $X_{2,j+1}$ = 40692 $X_{2,j}$ mod 2147483399,

3. Се поставува:

   X_{j+1}=($X_{1,j+1}$ – $X_{2,j+1}$) mod 2147483562,

4. Се враќа

$$R_{j+1} = \begin{cases} \dfrac{X_{j+1}}{2147483563}, X_{j+1} > 0 \\ \dfrac{2147483562}{2147483562}, X_{j+1} = 0 \end{cases},$$

5. Се поставува j = j + 1 и се оди на чекор 2

Овој комбиниран генератор има период (m_1-1) (m_2-1)/2≈2x10^{18}.

2.2 Тестови за случајни броеви

За да обезбедиме случајните броеви да ги задоволуваат својствата на рамномерност и непрекинатост, се користат тестови за случајни броеви. Овие тестови се поделени во две класи во зависност од тоа доколку се испитува дали низата од случајни броеви го задоволува својствата за рамномерност, или пак својството непрекинатост. Ќе ги разгледаме следниве пет тестови за случајни броеви:

1. **Тестови за фреквенција**: Се користи **Колмогоров-Смирнов тестот** или **χ^2–(Хи - квадрат)** тестот за споредување на распределбата на множеството од случајни броеви со рамномерната распределба.

2. ***Run*** **тестови:** Тестови за отстапувањата на добиените вредности во споредба со очекуваните вредности. Статистика за споредба е χ^2 - тестот.

3. **Тестови за автокорелација:** Тестови за корелација меѓу случајните броеви.

4. **Тестови за празнина (*Gap tests*)**: Го даваат бројот на цифри меѓу две повторувања на една иста цифра, а потоа се користи Колмогоров-Смирнов тест за споредување со очекуваната големина на пукнатината (gap), т.е. растојанието.

5. **Покер тест**: Ги подредува броевите како што се подредуваат поделените карти во играта покер. Добиената распределба на броевите се споредува со таа што е очекувана со користење на χ^2 -тестот.

За тестирање на рамномерноста (при тестовите за фреквенција) се користат следниве хипотези:

$$H_0: R_i \sim U(0, 1),$$
$$H_1: R_i \nsim U(0, 1)$$

Доколку се прифати нултата хипотеза тоа значи дека броевите се рамномерно распредели на (0,1) во спротивно доколку се отфрли броевите не се рамномерно распределени на (0, 1).

За тестирање на независноста, хипотезите се следниве:

$$H_0: R_i \sim \text{ независни}$$
$$H_1: R_i \nsim \text{независни}$$

И во овој случај доколку се прифати, нултата хипотеза тоа значи дека броевите се независни. Во спротивно доколку се отфрли нултата хипотеза броевите се зависни.

За секој тест, постои **ниво на значајност**α, кое е всушност веројатноста нултата хипотеза, да биде отфрлена доколку таа е точна.

$$\alpha = P(\text{да се отфрли } H_0 \mid H_0 \text{ е вистинита}).$$

Обично за нивото на значајност се зема да биде од 0.01 до 0.05.

2.2.1 Тестови за фреквенција

Овие тестови се тестови за рамномерност на случајните броеви. Се користат два вида статистички тестови: Колмогоров - Смирнов и χ^2 -тестот. И двата теста го мерат степенот на согласност помеѓу распределбата на примерок од генерирани случајни броеви и теоретската рамномерна распределба. Во двата теста се тестира нултата хипотеза дека не постои значајна разлика помеѓу распределбата на генерираниот примерок и теоретската распределба.

2.3.1.1 Тест на Колмогоров-Смирнов

Овој тест ја споредува емпириска функција на распределба $S_N(x)$ на множество од N генерирани случајни броеви со функција на распределба $F(x)$ на случајната променлива $X \sim U$ (0, 1). За низата од случајни броеви R_1, R_2,…. R_N, емпириската функција на распределба е дефинира со:

$$S_N(x) = \frac{1}{N} \sum_{i:R_i<x} n_i .$$

За поголеми вредности на N треба да се очекува подобра апроксимација на функцијата $F(x)$. Колмогоров-Смирнов тестот се базира на најголемата отстапување меѓу $F(x)$ и $S_N(x)$ над рангот на случајната променлива т.е. тој е базиран на статистиката:

$$D = \max |F(x) - S_N(x)|. \quad (2.3)$$

Процедура за тестирање ги опфаќа следниве чекори:

Чекор 1. Случајните броеви се подредуваат во растечки редослед:

$$R_{(1)} \le R_{(2)} \le \ldots . \le R_N.$$

каде$R_{(i)}$ е i-от по големина случаен број.

Чекор 2. Се пресметува:

$$D^+ = \max_{1 \le i \le N} \{\frac{i}{N} - R_{(i)}\},$$

$$D^- = \max_{1 \le i \le N} \{R_{(i)} - \frac{i-1}{N}\}.$$

Чекор 3. Се пресметува D=max {D^+, D^-}.

Чекор 4. Се определува критична вредност D_α за дадено ниво на значајност α и дадена големина на примерокот N.

Чекор 5. Ако статистиката D е поголема од критичната вредност D_α тогаш нултата хипотеза дека случајните броеви се рамномерно распределени се отфрла. Во случај кога $D \le D_\alpha$, заклучуваме дека не е определено отстапување помеѓу распределбата на {R_1, R_2,…. R_N} и рамномерната распределба и нултата хипотеза се прифаќа.

Пример 2.7 (Колмогоров - Смирнов тест) Се претпоставува дека броевите 0.44, 0.81, 0.14,0.05 и 0.93 се генерирани и треба да се определи дали овие броеви се рамномерно распределени со помош на Колмогоров-Смирнов тестот со ниво на значајност α=0.05. Прво ги подредуваме броевите во растечи редослед, потоа ги пресметуваме D^+ и D^-. Резултатите се дадени во табелата подолу:

$R_{(i)}$	0.05	0.14	0.44	0.81	0.93
$\frac{i}{N}$	0.20	0.40	0.60	0.80	1.00
$\frac{i}{N} - R_{(i)}$	0.15	0.26	0.16	0.01	0.07
$R_{(i)} - \frac{i-1}{N}$	-	0.06	0.04	0.21	0.13

Табела 2.2 Пресметки за Колмогоров-Смирнов тестот

Од табелата се гледа дека $D^+ = 0.26$ и $D^-=0.21$, бидејќи за D се зема поголемата вредност $D= 0.26$. Критичната вредност D_α за $\alpha=0.05$ и $N=5$ е 0.0565. Бидејќи $D \le D_\alpha$ хипотезата дека случајните броеви се рамномерно распределени се прифаќа.

2.3.1.2 χ^2 – (Хи-квадрат) тест

За користење на χ^2 - тестот потребно е да се подели интервалот (0,1) во n подинтервали со еднаква должина. Овој тест ја користи следнава статистика:

$$\chi_0^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

каде O_iе бројот на елементи во примерокот во i-от подинтервал (i–тата класа), додека E_i е очекуваниот број на елементи во i-от подинтервал, т.е.

$$E_i = \frac{N}{n},$$

каде N е вкупниот број набљудувања, а n е вкупниот број на класи (подинтервали) Може да се покаже дека распределбата χ_0^2 е приближно χ^2 распределба со n-1 степен на слобода.

Пример 2.8(χ^2- тест) Со користење на χ^2 тестот, со ниво на значајностα да се види дали следните случајни броеви се рамномерно распределени:

0.34	0.90	0.25	0.89	0.87	0.44	0.12	0.21	0.46	0.67
0.83	0.76	0.79	0.64	0.70	0.81	0.94	0.74	0.22	0.74
0.96	0.99	0.77	0.67	0.56	0.41	0.52	0.73	0.99	0.02
0.47	0.30	0.17	0.82	0.56	0.05	0.45	0.31	0.78	0.05
0.19	0.71	0.23	0.19	0.82	0.93	0.65	0.37	0.39	0.42
0.99	0.17	0.99	0.46	0.05	0.66	0.10	0.42	0.18	0.49
0.37	0.51	0.54	0.01	0.81	0.28	0.69	0.34	0.75	0.49
0.72	0.43	0.56	0.97	0.30	0.94	0.96	0.58	0.73	0.05
0.06	0.39	0.84	0.24	0.40	0.64	0.40	0.19	0.79	0.62
0.18	0.26	0.97	0.88	0.64	0.47	0.60	0.11	0.29	0.78

Интервалот [0,1] е поделен на десет подинтервали [0,0.1), [0.1,0.2),..., [0.9,1.0). Вредноста на χ_0^2=3.4, како што е прикажано во долната табела. Критичната вредност $\chi_{0.05,9}^2$ =16.9 Бидејќи $\chi_0^2 \le \chi_{0.05,9}^2$, нултата хипотеза, дека низата од случајни броеви е рамномерно распределена се прифаќа.

интервал	O_i	E_i	$O_i- E_i$	$(O_i- E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
1	8	10	-2	4	0.4
2	8	10	-2	4	0.4
3	10	10	0	0	0.0
4	9	10	-1	1	0.1
5	12	10	2	4	0.4
6	8	10	-2	4	0.4

7	10	10	0	0	0
8	14	10	4	16	1.6
9	10	10	0	0	0
10	11	10	1	1	0.1
	100	100	0		3.4

Табела 2.3 Пресметки заχ^2 тестот

И Колмогоров-Смирновиот тест и χ^2 тестот се прифатливи за тестирање дали низата од случајни броеви е рамномерно распределена. Меѓутоа, Колмогоров-Смирновиот тест е помоќен тест и е попрепорачлив. Колмогоров-Смирновиот тест може да се користи и кај помали примероци од случајни броеви, додека χ^2- тестот е валиден за поголеми примероци кога $N \geq 50$.

2.3.2 Тестови со монотоност (*„Run“* тест)

Се разгледува следната низата од 40 случајни броеви генерирани од генератор на случајни броеви:

0.08	0.09	0.23	0.29	0.42	0.55	0.58	0.72	0.89	0.91
0.11	0.16	0.18	0.31	0.41	0.53	0.71	0.73	0.74	0.84
0.02	0.09	0.30	0.32	0.45	0.47	0.69	0.74	0.91	0.95
0.12	0.13	0.29	0.36	0.38	0.54	0.68	0.86	0.88	0.91

И Колмогоров-Смирнов тестот и χ^2 - тестот, покажуваат дека броевите се рамномено распределени.„***Run***“ тестот е тест за независност. Од горната низа се гледа дека броевите се подредени по големи во поднизи од по десет броеви. Ако броевите се прередат тогаш се поголеми сомневањата за нивната независност.

0.41	0.68	0.89	0.84	0.74	0.91	0.55	0.71	0.36	0.30
0.09	0.72	0.86	0.08	0.54	0.02	0.11	0.29	0.16	0.18
0.88	0.91	0.95	0.69	0.09	0.38	0.23	0.32	0.91	0.53
0.31	0.42	0.73	0.12	0.74	0.45	0.13	0.47	0.58	0.29

2.3.2.1 Монотоно растење или монотоно опаѓање (*Runup*или *Rundown*)

Кај „***run***“ тестот низата се поделува на поднизи, така што всушност „***run up***“ претствавува растечка подниза (низа од броеви каде секој број е следен со поголем број) и „***run down***“ претставува опаѓачка подниза (низа од броеви каде што секој број е следен со помал број). За илустрација се дава следнава секвенца од 15 броеви :

-0.87+0.15 +0.23 +.0.45 -0.69 -0.32 -0.30+0.19-0.24+0.18 +0.65 +0.82-0.93+0.22 0.81
Пред секој број се додава “+” или “-“ во зависност од тоа дали неговиот следбеник е поголем или помал број од нив. На последниот број не му се придружува знак затоа што тој нема следбеник. Во нашиот случај се добива следната низа од “+” и “-“:

\- + + + - - - + - + + + - +

Значи во овој пример има 8 монотони поднизи „*runs*".

Ако низата се состои од 15 броеви, тогаш максималниот број на поднизи е 14, а минималниот број е 1.

Нека A е вкупниот број на монотони поднизи („*runs*") во идеална рамномерно распределена низа од броеви. Тогаш математичкото очекување и дисперзијата на случајната променлива A се дадени со :

$$\mu_A = \frac{2N-1}{3}, \tag{2.4}$$

$$\sigma_A^2 = \frac{16N-29}{90}. \tag{2.5}$$

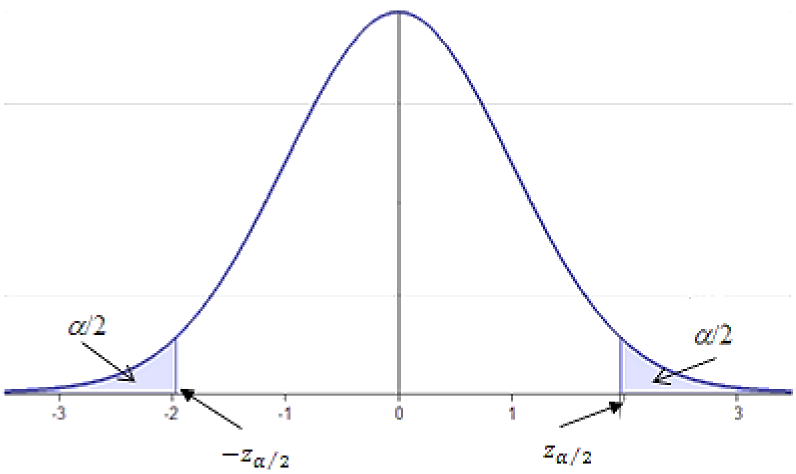
За N>20, распределбата на A е приближно нормална распределба $N(\mu_A, \sigma_A^2)$. Оваа апроксимација се користи како тест за независност на низата од случајни броеви. Се користи тестирање со нормирана нормална тест статистика:

$$Z_0 = \frac{A-\mu_A}{\sigma_A},$$

Со замена на μ_A и σ_A^2 се добива:

$$Z_0 = \frac{A-\frac{2N-1}{3}}{\sqrt{\frac{16N-29}{90}}}.$$

Каде Z_0 е нормирана нормална распределба, Z_0~$N(0,1)$. Хипотезата за независност не можеме да ја отфрлиме доколку $-z_{\alpha/2} \leq Z_0 \leq z_{\alpha/2}$, каде α е ниво на значајност. Критичната вредност и регионите на отфрлање се прикажани на **Фиг.2.2**:



Фиг. 2.2 Густина на распределба на нормирана нормална распределба

Пример 2.9 (Тест за монотоност) Со помош на тестот за монотоност, да се определи дали низата од 40 случајни броеви е независна низа од броеви, каде што нивото на значајност $\alpha/2$ =0.05

0.41	0.68	0.89	0.94	0.74	0.91	0.55	0.62	0.36	0.27
0.19	0.72	0.75	0.08	0.54	0.02	0.01	0.36	0.16	0.28

0.18	0.01	0.95	0.69	0.18	0.47	0.23	0.32	0.82	0.53
0.31	0.42	0.73	0.04	0.83	0.45	0.13	0.57	0.63	0.29

Низата од монотоно растечки и монотоно опаѓачки поднизи е следнава:

+ + +-+-+- - -+ +- +- - + - +- -+- -+-+ +- -+ +- + - - + + -

Имаме 26 монотоно растечки и моното опаѓачки поднизи во оваа низа. За *N*=40 и *A*=26, се добива:

$$\mu_A = \frac{2 \cdot 40 - 1}{3} = 26{,}33,$$

$$\sigma_A^2 = \frac{16 \cdot 40 - 29}{90} = 6{,}79,$$

$$z_0 = \frac{26 - 26{,}33}{\sqrt{6{,}79}} - 0{,}13.$$

Од таблицата за нормална распределба се чита$z_{0.025}$ = 1.96,па критичниот домен е C = (- ∞; -1,96) ∪ (1,96; ∞). Бидејќи $z_0 \notin C$можеме да заклучиме дека H_0 не може да биде отфрлена.

2.3.2.2. Поднизи над и под математичкото очекување (*Runs above u runs below*)

Тестот за монотони поднизи не е комплетно адекватен да ја процени независноста на низата од случајни броеви. Се разгледува следнава низа од случајни броеви:

063	0.72	0.79	0.81	0.52	0.94	0.83	0.93	0.87	0.67
0.54	0.83	0.89	0.55	0.88	0.77	0.74	0.95	0.82	0.86
0.43	0.32	0.36	0.18	0.08	0.19	0.18	0.27	0.36	0.34
0.31	0.45	0.49	0.43	0.46	0.35	0.28	0.39	0.47	0.41

Низата е иста како во претходниот пример:

+ + +-+-+- - -+ +- +- - + - +- -+- -+-+ +- -+ +- + - - + + -

Меѓутоа може да се забележи дека првите 20 броеви се поголеми од математичкото очекување (0.99+0.00)/2=0.495, додека последните 20 броеви се помали. Во овој случај може да се користи *runs* тестот само ако дефиницијата на *runs* е променета. Во овој случај ќе го додаваме знакот "+" пред бројот што е поголем од математичкото очекување, а "-" пред бројот што е помал од математичкото очекување.

Ја разгледуваме низата од 20 случајни броеви:

0.40	0.84	0.75	0.18	0.13	0.92	0.57	0.77	0.30	0.71
0.42	0.05	0.78	0.74	0.68	0.03	0.18	0.51	0.10	0.37

Низата од плусови и минуси е следнава:

– + + - - + + - + - - + + + - - + - -

Значи има 11 поднизи, 5 од нив се *runs above* (5 броја се поголеми од математичкото очекување), додека 6 се *runs below* (6 броја се помали од математичкото очекување). Нека n_1 е бројот на сите случајни броеви што се поголеми од математичкото очекување, додека n_2 бројот на сите случајни броеви што се помали математичкото очекување. Исто така да се забележува дека максималниот број на *runs* е $N = n_1 + n_2$, додека минималниот број на *runs* е 1. Нека случајната променлива B го означува бројот на поднизи под и над математичкото очекување. За дадени n_1 иn_2, математичкото очекување и дисперзијата на B за идеално рамномерно распределена низа броеви (со непрекината корекција, предложена од *Swed u Eisenhard*) се од облик:

$$\mu_B = \frac{2n_1n_2}{N} + \frac{1}{2}, \tag{2.6}$$

$$\sigma_B^2 = \frac{2n_1n_2(2n_1n_2 - 1)}{N^2(N-1)}. \tag{2.7}$$

За n_1 и n_2 поголеми од 20, B има приближно нормална распределба. Со нормирање на оваа распределба се добива нормирано нормално распределена случајна променлива :

$$Z_0 = \frac{b - \left(2n_1n_2/N\right) - 1/2}{\left[\dfrac{2n_1n_2(2n_1n_2 - N)}{N^2(N-1)}\right]^{1/2}}$$

Значи , $Z_0 \sim N(0,1)$. Хипотезата за независност не можеме да се отфрли доколку $-z_{\alpha/2} \leq Z_0 \leq z_{\alpha/2}$, каде α е ниво на значајност.

Пример 2.10 (*runs above u runs below*). Се разгледува истата низа од броеви како во пример 2,9. Во овој случај ја добиваме следната низа од плусови и минуси:

\- + + + + + + +- - - + + - +- - - - - - - + + - - - -++ - -+ - +- - + + -

n_1=18
n_2=22
$N= n_1+ n_2$=40
B=17

Се определува математичкото очекување и дисперзијата на случајната променлива B:

$$\mu_B = \frac{2 \cdot 18 \cdot 22}{40} + \frac{1}{2} = 20{,}3,$$

$$\sigma_B^2 = \frac{2 \cdot 18 \cdot 22(\cdot 18 \cdot 22 - 1)}{40^2 \cdot (40-1)} = 9{,}54.$$

Бидејќи n_2>20, нормалната апроксимација е прифатлива, за Z_0 се добива:

$$Z_0 = \frac{17 - 20.3}{\sqrt{9.54}} = -1.07$$

Критичниот домен е C = (- ∞; -1,96) ∪ (1,96; ∞). Бидејќи $z_0 \notin C$ можеме да заклучиме дека H_0 не може да биде отфрлена.

2.3.3 Тест за автокорелација

Со тестот за автокорелација се испитува зависноста помеѓу случајните броеви во низата од случајни броеви. Со овој тест се пресметува автокорелацијата меѓу секои m броеви започнувајќи со i-от број. Ја разгледуваме автокорелацијата ρ_{im} меѓу следниве броеви: R_i, R_{i+m}, R_{i+2m}, … , $R_{i+(M+1)m}$. Бројот M е најголемиот цел број кој го задоволува својството $i+(M+1)m \leq N$, каде N е вкупниот број на елементи во низата. Се разгледуваат следниве хипотези:

$$H_0 : \rho_{im} = 0;$$
$$H_1 : \rho_{im} \neq 0;$$

За големи вредности на M , распределбата на оценувачот ρ_{im} , означена со $\hat{\rho}_{im}$ е приближно нормално распределена, ако вредностите R_i, R_{i+m}, R_{i+2m}, … , $R_{i+(M+1)m}$ се независни за големи M. Се разгледува следната тест статистика:

$$Z_0 = \frac{\hat{\rho}_{im}}{\sigma_{\hat{\rho}_{im}}}.$$

Формулата за $\hat{\rho}_{im}$ и дисперзијата се дадени со:

$$\hat{\rho}_{im} = \frac{1}{M+1}\left[\sum_{k=0}^{M} R_{i+km} R_{i+(k+1)m}\right] - 0.25,$$

$$\sigma_{\hat{\rho}_{im}} = \frac{\sqrt{13M+7}}{12(M+1)}.$$

Каде Z_0 е нормирана нормално распределена случајна променлива, Z_0~N(0,1). Хипотезата за независност не можеме да ја отфрлиме доколку $-z_{\alpha/2} \leq Z_0 \leq z_{\alpha/2}$, каде α е ниво на значајност.

Пример 2.11 (Тест за автокорелација) : Се разгледува 3-от, 8-от, 13-от ... број од низата случајни броеви:

0.12	0.01	0.23	0.28	0.89	0.31	0.64	0.28	0.83	0.93
0.99	0.15	0.33	0.35	0.91	0.41	0.64	0.27	0.75	0.88
0.68	0.49	0.05	0.43	0.95	0.58	0.19	0.36	0.69	0.87

Нека i = 3, m = 5, N = 30, M = 4 (најголемиот цел број за кој $3+(M+1)5 \leq 30$)

$\hat{\rho}_{35}$= 1/(4+1)[0.23 ·0.28 +0.28·0.33 + 0.33·0.27 + 0.27· 0.05 + 0.05 ·0.36] – 0.25
= -0.1945
и

$$\sigma_{\hat{\rho}_{35}} = \frac{\sqrt{13\cdot 4+7}}{12(4+1)} = 0.1280,$$

Тест статистиката ја добива следната вредност:

$$Z_0 = -0.1945/0.1280 = -1.516$$

Оваа вредност не припаѓа на критичниот домен $C=(-\infty,-1,96)\cup(1,96,+\infty)$, па хипотезата за независност не може да се отфрли врз база на овој тест.

2.3.4 Тест на растојание (пукнатини - *Gap* тест)

Тестот на растојание се користи за да го определи значењето на интервалот меѓу повторувањето на една иста цифра. Пукнатина со должина *x* се нарекува растојанието меѓу две појавувања на една иста цифра. Во следниот пример е илустрирана должината на пукнатината за цифрата 3:

4, 1, 3, 5, 1, 7, 2, 8, 2, 0, 7, 9, 1 , 3, 5, 2, 7, 9, 4, 1, 6, 3
3, 9, 6, 3 , 4, 8, 2, 3, 9, 1, 4, 4, 6, 4, 8, 1, 3, 8, 9, 5, 5, 7
3, 9, 5, 9, 8, 5, 3,2, 2, 3, 7, 4, 7, 0, 3, 6, 3,5, 9, 9, 5, 5
5, 0, 4, 6, 8, 0, 4, 7, 0, 3, 3, 0, 9, 5, 7, 9, 5, 1, 6, 6, 3, 8
8, 8, 9, 2, 9, 1, 8, 5, 4, 4, 5, 0, 2, 3, 9, 7, 1, 2, 0, 3, 6, 3

Имаме 18 појавувања на цифрата 3, и 17 растојанија. Првото растојание е со должина 10.
Веројатноста за појавување првото растојание со должина 10 може да се определи на следниот начин:

P(растојание = 10)= *P* {после цифрата 3 следуваат точно 10 цифри различни од 3}= $(0.9)^{10}\cdot$ 0,1.

јасно е дека веројатноста некоја цифра да не е 3 е 0.9 а да биде 3 е 0.1. Во општ случај, веројатноста за појавување на пукнатина, може да се определи на следниот начин:

P(после цифрата *t* следуваат точно *x* цифри различни од *t*) =$(0.9)^x\cdot 0{,}1$, x=0,1,2…

Теоретски распределбата на растојанијата е геометриска распределба *Geo* (0,1). Теориската функција на оваа распределба е:

$$F(x) = P\{\text{растојание} < x\} = 0{,}1\sum_{n=0}^{x}(0{,}9)^n = 1 - 0{,}9^x \tag{2.8}$$

Нека $S_n(x)$ е емпириска функција на распределбана пукнатините од примерокот. Ако *N*е вкупен број на податоци, тогаш:

$$S_n(x) = \frac{\text{број на пукнатини} < x}{N}$$

Процедурата за овој тест е опишана во чекорите подолу:

Чекор 1.Се определува теориската распределба на фреквенцијата дадена со равенка (2.8) која се базира на избраните интервали.

Чекор 2.За набљудуваниот примерок од растојанија се определуваат кумулативните распределби за истите интервали.

Чекор 3.Се определува максималната разлика меѓу $F(x)$ и $S_N(x)$.

Чекор 4. Се определува критичната вредноста D_α, за дадена вредност на α и големина на примерокот *N*.

Чекор 5.Ако пресметаната вредност на *D*, е поголема од табличната вредност D_αтогаш нултата хипотеза за независност не може да биде отфрлена.

Пример 2.12 (Тест на растојание) . Се разгледува низа од 110 случајни цифри, со помош на овој тест да се определи дали цифрите се независни. Нивото на значајност

α = 0.05. Бројот на растојанија е даден со бројот на дадените податоци минус бројот на различните цифри, 110 - 10 = 100. Бројот на растојанија поврзани со различни цифри е даден во табелата подолу.

Цифра	0	1	2	3	4	5	6	7	8	9
Број на растојанија	7	8	8	17	10	13	7	8	9	13

Резултатите од тестот се дадени во табелата подолу. Критичната вредност на D е дадена со: $D_{0,05} = \frac{1.36}{\sqrt{100}} = 0.136$.

Значи $D = \max |F(x) - S_n(x)| = 0.0224$ е помала од $D_{0.05}$ нултата хипотеза за независност не може да биде отфрлена
.

Должина на пукнатина	Фреквенција	Релативна фреквенција	Кумулативна релативна фреквенција	F(x)	$\|F(x)-S_n(x)\|$
0-3	35	0.35	0.35	0.3439	0.0061
4-7	22	0.22	0.57	0.5695	0.0005
8-11	17	0.17	0.74	0.7176	0.0224
12-15	9	0.09	0.83	0.8147	0.0153
16-19	5	0.05	0.88	0.8784	0.0016
20-23	6	0.06	0.94	0.9202	0.0198
24-27	3	0.03	0.97	0.9497	0.0223
28-31	0	0	0.97	0.9657	0.0043
32-35	0	0	0.97	0.9775	0.0075
36-39	2	0.02	0.99	0.9852	0.0043
40-43	0	0	0.99	0.9903	0.0003
44-47	1	0.01	1.00	0.9936	0.0064

Табела 2.4 Пример за тест на растојанија

2.3.5 Покер тест

Покер тестот се базира на фреквенцијата со која некои цифри се повторуваат во дадена серија броеви. Следниот пример го покажува вообичаеното значење на повторувањето:0.255, 0.577, 0.331, 0.414, 0.828, 0.909, 0.303, 0.001, ..

Во секој случај се појавува пар од еднакви цифри во броевите што се генерирани. За трицифрените броеви ги разгледуваме следните веројатности:

1. Сите три цифри да бидат различни;
2. Сите цифри да бидат исти;
3. Да има две еднакви цифри.

P(трите цифри да се различни) = *P*(веројатноста втората цифра да е различна од првата)·*P*(третата цифра да е различна од првата и од втората) = 0,9 · 0,8=0,72.

P(трите цифри да се еднакви) = *P*(веројатноста втората цифра да е еднаква со првата) ·*P*(третата цифра да е еднаква со првата) = 0.1 · 0.1=0.01.

P(две цифри да се еднакви) = 1- 0.72- 0.01=0.27.

Истата веројатност може да се добие и на поинаков начин:

P(две цифри да се еднакви)= $\binom{3}{2}(0.1)(0.9) = 0.27$.

Пример 2.13 (Покер тест) Секвенцата од 1000 трицифрени броеви е генерирана, и анализата покажа дека 680 имаат три различни цифри, 289 содржат две еднакви цифри, и 31 содржат три еднакви цифри. Со помош на покер тест да се определи дали броевите се независни. Нека α=0.05. Резултати од тестот се дадени во табелата подолу. За анализа се користи Hi-квадрат тестот. Бројот на степени на слобода е за еден помал од бројот на класите (3 -1 = 2). Бидејќи $\chi_{0.05,2}$=5.99 > 47.65, хипотеза за независност во овој случај се отфрла.

Комбинации -*i*	O_i	E_i	$\frac{(O_i - E_i{}^2)}{E_i}$
Три различни цифри	680	720	2.22
Три еднакви цифри	31	10	44.10
Точно еден пар	289	270	1.33
Сума	1000	1000	47,65

Табела 2.5 Резултати од Покер тест

3. ГЕНЕРИРАЊЕ НА СЛУЧАЈНИ ПРОМЕНЛИВИ

Во овој дел се разгледани процедурите за генерирање на случајните променливи што се користат како влез во симулационите модели. Ги разгледуваме следниве три техники: Техника со инверзна трансформација, метод на конволуција и техника за прифаќање и одбивање. Сите техники претпоставуваат дека се познати случајните броеви $R_1, R_2, \ldots$. кои се рамномерно распределени на интервалот (0, 1). Сите случајни броеви R_i имаат густина на распределба

$$p_R(x) = \begin{cases} 1, & 0 \le x \le 1 \\ 0, & \text{инаку} \end{cases},$$

и функција на распределба:

$$F_R(x) = \begin{cases} 0, & x < 0 \\ x, & 0 \le x \le 1 \\ 1, & x \ge 1 \end{cases}$$

3.1 Техника со инверзна трансформација

Оваа техника се користи за генерирање на експоненцијална, рамномерна, *Weibull*, триаголна, емпириска и некои други распределби. Техниката ќе биде детално објаснета за експоненцијална распределба. Оваа техника е едноставна, но не е секогаш ефикасна.

3.1.1. Екпоненцијална распределба

Експоненцијалната распределба ја има следнава густина :

$$p(x) = \begin{cases} \lambda e^{-\lambda x}, & x \ge 0 \\ 0, & x < 0 \end{cases}.$$

и функција на распределба е дадена со :

$$F(x) = \int_{-\infty}^{x} f(t)dt = \begin{cases} 1 - e^{-\lambda x}, & x \ge 0 \\ 0, & x < 0 \end{cases}.$$

Параметарот λ се интерпретира како среден број на случувања во единица време. На пример ако интервалите меѓу две последователни доаѓања X_1, X_2,…. се екпоненцијално распределени со интензитет λ, тогаш λ се интерпретира како среден број од доаѓања во единица време, или интензитет на доаѓање. Да забележиме дека за секое i:

$$E(X_i) = \frac{1}{\lambda},$$

така да $1/\lambda$ е средното време меѓу две последователни пристигнувања. Целта е да се генерираат случајни променливи X_1, X_2, … кои ќе имаат експоненцијална распределба. Техниката со инверзна трансформација може да се користи за било која распределба, но таа е покорисна(поефикасна), кога функцијата на распределба $F(x)$ има поедноставна форма така што нејзината инверзна функција F^{-1}да може полесно да се пресмета. Процедурата за генерирање на експоненцијалната распределба ќе ја опишеме детално, чекор по чекор.

Чекор 1. Се пресметува функцијата на распределба за случајната променлива X. За експоненцијалната распределба, функцијата на распределба $F(x) = 1 - e^{-\lambda x}$, $x \geq 0$.

Чекор 2.Се става $F(X) = R$. За експоненцијалната распределба имаме $1- e^{-\lambda X} = R$. Значи Xе случајна променлива со експоненцијална распределба, од што следи дека и $1-e^{-\lambda x}$ е случајна променлива, означена со R. Покасно ќе покажеме дека ова случајна променлива е рамномерно распределена над интервалот (0,1).

Чекор 3.Ја решаваме равенката $F(X) = R$, така што го изразуваме X преку R. За експоненцијалната распределба решението е следното.

$$1- e^{-\lambda X} = R$$
$$e^{-\lambda X} = 1-R$$
$$-\lambda X = \ln(1-R)$$
$$X= -1/\lambda \ln(1-R) \tag{3.1}$$

Последната равенка е наречена генератор за експоненцијална распределба.

Чекор 4.Се генерира низа од рамномерно распределени случајни броеви R_1, R_2,…и се пресметуваат случајните вредности:

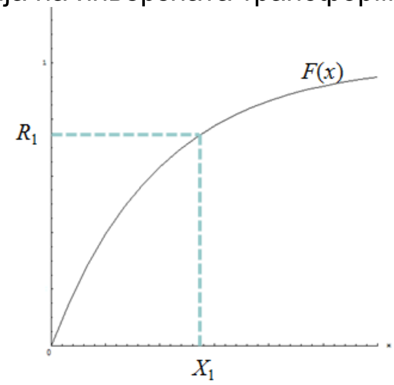
$$X_i= F^{-1}(R_i) \tag{3.2}$$

За експоненцијалната распределба се добива:

$$X_i = -1/\lambda \ln(1- R_i) \text{ , } i=1,3,.. \tag{3.3}$$

За поедноставување на пресметките обично $1-R_i$се заменува со R_i, $X_i= -1/\lambda \ln R_i$ ова е можно бидејќи и R_iи $1- R_i$ се рамномерно распределени на (0,1).

Графичката интерпретација на инверзната трансформација е дадена на **Фиг. 3.1**



Фиг. 3.1Графички приказ на техниката со инверзна трансформација

За добивање на вредноста на X_1со функција на распределба $F(x)$, прво случајниот број R_1е генериран, потоа се повлекува хоризонтална линија од R_1 до Фиг.от на функцијата на распределба, од оваа точка се повлекува вертикална линија, вредноста X_1се добива на пресекот на вертикалната линија со хоската.

Инверзната релација меѓу X_1и R_1за експоненцијалната распределба е дадена со:

$$R_1 = 1-e^{-X},$$

$$X_1=-\ln(1-R_1).$$

Во општ случај релација е:

$$R_1=F(X_1),$$

$$X_1=F^{-1}(R_1).$$

3.1.2 Рамномерна распределба

Нека случајната променлива X е рамномерно распределена на интервалот (a, b). Xќе ја има следнава густина на распределба:

$$p(x)=\begin{cases}\dfrac{1}{b-a}, & a \le x \le b\\ 0, & \text{инаку}\end{cases},$$

Чекор 1. Дадена е функцијата на распределба на рамномерната распределба :

$$F(x)=\begin{cases}0, & x<a\\ \dfrac{x-a}{b-a}, & a \le x \le b\\ 1, & x>b\end{cases},$$

Чекор 2. Се поставува $F(X)=(X-a)/(b-a) = R$,

Чекор 3. Xе изразено преку R, на следниот начин т.е.:

$$X = a+(b - a)R \tag{3.4}$$

3.1.3 Вејбулова распределба

Случајната променлива X има Вејбулова распределба со параметри αи β(α> 0, β> 0), со густина на распределба:

$$p(x)=\begin{cases}\dfrac{\beta}{\alpha^{\beta}}x^{\beta-1}e^{-\left(x/\alpha\right)^{\beta}}, & x \ge 0\\ 0, & \text{инаку}\end{cases},$$

Чекор 1. Функцијата на распределба е дадена со $F(x)=1-e^{-\left(x/\alpha\right)^{\beta}}$, $x \ge 0$.

Чекор 2. $F(x)=1-e^{-\left(x/\alpha\right)^{\beta}}=R$.

Чекор 3. Се изразува X преку R:

$$X = \alpha[-\ln(1-R)]^{1/\beta} \tag{3.5}$$

3.1.4 Триаголна распределба

Триаголната распределба е дадена со густината на распределба:

$$p(x)=\begin{cases} x, & 0\le x\le 1\\ 2-x, & 1\le x\le 2,\\ 0, & \text{инаку}\end{cases}$$

функцијата на распределба е:

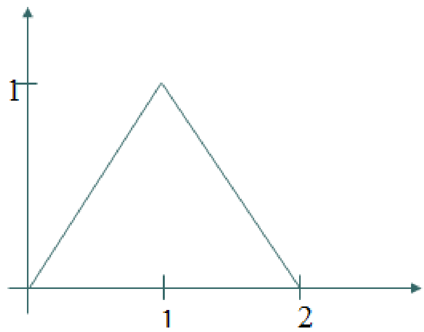
$$F(x)=\begin{cases} 0, & x\le 0\\ \dfrac{x^2}{2}, & 0<x\le 1\\ 1-\dfrac{(2-x)^2}{2}, & 1<x\le 2\\ 1, & x>2\end{cases},$$

за $0 \le X \le 1$

$$R = X^2/2 \tag{3.6}$$

за $1 \le X \le 2$

$$R=1-(2-X)^2/2 \tag{3.7}$$



Фиг. 3.2: Густина на веројатност на триаголна распределба

Од (3.6) za $0 \le X \le 1$ следи дека $0 \le R \le ½$, и во овој случај $X=\sqrt{2R}$, додека од (3.7) за $1\le X\le 2$, следи дека $1/2 \le R\le 1$ и во овој случај $X=2-\sqrt{2(1-R)}$ или :

$$X=\begin{cases} X=\sqrt{2R}, & 0\le R\le \dfrac{1}{2}\\ X=2-\sqrt{2(1-R)}, & \dfrac{1}{2}<R\le 1\end{cases}. \tag{3.8}$$

3.1.5 Емпириска непрекината распределба

Ако моделаторот не може да ја пронајде теориската распределба, што ќе доведе до добар модел за влезните податоци, тогаш е потребно да се користи емпириската распределба на податоците. Емпириската распределба покажува добри особини кога за влезниот процес е познато дека се зема на конечен број на вредности.

Пример 3.1 Пет набљудувања на времињата на група пожари од влезниот аларм се собираат. Овие податоци се:

2.76 1. 83 0.80 1.45 1.24

Пред да се соберат податоците, потребно е да се развие прелиминарен симулационен модел кој ќе ја користи распределбата на времињата базирана на овие пет набљудувања. Потребен е метод за генерирање на случајните променливи за распределбата на времињата. Претпоставуваме дека времето X има ранг, $0 \leq X \leq c$, каде c е непозната но ќе биде оценето со $\hat{c} = \max\{X_i : i = 1 \ldots n\} = 2.76$ каде $\{X_i | \ i=1, 2\ldots, n\}$ се набљудуваните податоци , а $n = 5$ е бројот на набљудувања. Податоците се подредуваат, од најмалите кон поголемите, значи $x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(n)}$. Се зема најмалата можна вредност да е нула, $x_{(0)} = 0$. Се зема веројатност од $1/n = 1/5$ за секој интервал $x_{(i-1)} \leq x \leq x_{(i)}$ како што се гледа во табелата:

i: $x_{(i-1)} < x \leq x_{(i)}$	Веројатност $1/n$	Кумулативна веројатност	a_i
1. $0 < x \leq 0.80$	0.2	0.2	4.00
2. $0.80 < x \leq 1.24$	0.2	0.4	2.20
3. $1.24 < x \leq 1.45$	0.2	0.6	1.05
4. $1.45 < x \leq 1.83$	0.2	0.8	1.90
5. $1.83 < x \leq 2.76$	0.2	1.0	4.65

Табела 3.1 Сумирање на податоците

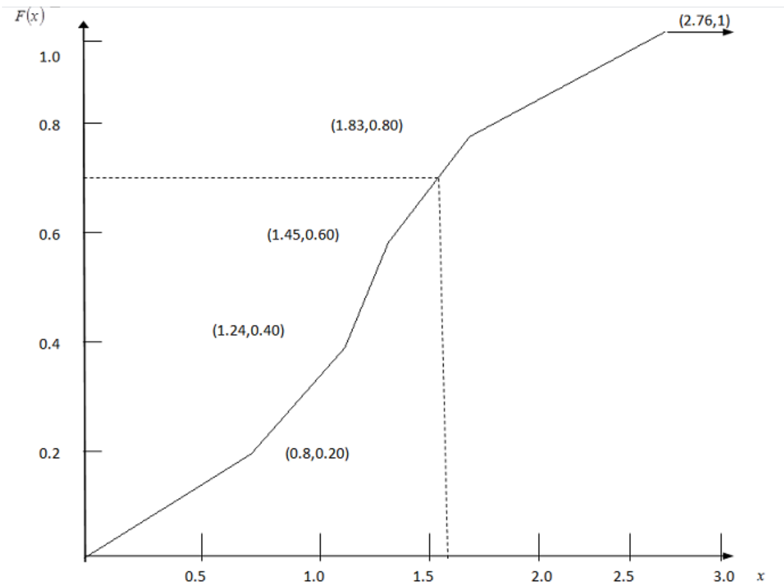
Емпириската функција на распределба $\hat{F}(x)$, е илустрирана на игура 3.3 .Стрмнините a_i се дадени со следново равенство:

$$a_i = \frac{x_{(i)} - x_{(i-1)}}{i/n - (i-1)/n} = \frac{x_{(i)} - x_{(i-1)}}{1/n} .$$

Додека инверзната функција на распределба се пресметува со:

$$X = \hat{F}^{-1}(R) = x_{(i-1)} + a_i\left(R - \frac{(i-1)}{n}\right). \tag{3.9}$$

каде што $(i-1)/n < R \leq i/n$.



Фиг. 3.3 Емпириска функција на распределба

За овој пример случајниот број R_1= 0.71 е генериран, овој број ќе припаѓа во третиот интервал, $0.6 < 0.71 \leq 0.8$, од каде се добива:

$$X_1 = x_{(4-1)} + a_4(R_1-(4-1)/5) = 1.45 + 1.9(0.71 - 0.6) = 1.66.$$

3.1.6 Генерирање на случајни променливи од дискретен тип

Сите дискретни распределби може да бидат генерирани со користење на техниката за инверзна трансформација, или нумерички преку *table-lookup* процедура, или во некои случаи алгебарски.

Пример 3.2 (Емпириска дискретна распределба) На крајот на денот, бројот на пратките донесени со брод може да биде 0, 1 или 2 , со релативни фреквенции 0.5, 0.3 и 0.2 соодветно, како што може да се види во табелата:

X	$p(x)$	$F(x)$
0	0.5	0.5
1	0.3	0.8
2	0.2	1.00

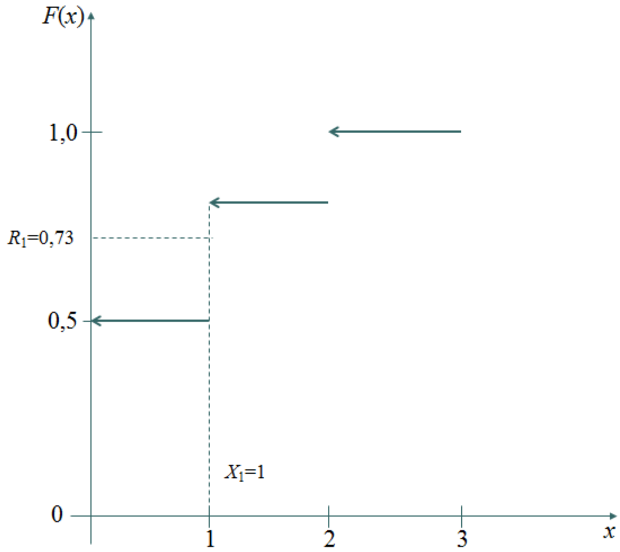
Табела 3.2 Распределба на бројот на пратки X.

X е дискретно распределена случајна променлива , со веројатности:

$$P(0) = P(X = 0) = 0.5,$$
$$P(1) = P(X = 1) = 0.3,$$
$$P(2) = P(X = 2) = 0.2.$$

Додека густината на распределба $F(x) = P(X \leq x)$, е дадена со:

$$F(x) = \begin{cases} 0, & x < 0 \\ 0.5, & 0 \leq x < 1 \\ 0.8, & 1 \leq x < 2 \\ 1.0, & 2 \leq x \end{cases}$$



Фигура 3.4 Функција на распределба за бројот на пратки

i	Влез - r_i	Излез - x_i
1	0.50	0
2	0.80	1
3	1.00	2

Табела 3.3 Табела за генерирање на дискретна променлива X

За генерирање на дискретните случајни променливи се користи техниката *table-lookup* процедурата. За илустрирање на ова процедура, претпоставуваме дека R_1 = 0.73. Графички , прво на y- оската се поставува R_1 потоа се поставува хоризонтална линија се наоѓа пресекот со функцијата на распределба, и од оваа точка се поставува вертикална линија до пресек со x - оската, добиената вредност на x - оската е генерираната вредност. Значи за R_1 = 0.73, се добива X_1 = 1.Со помош на процедурата *table-lookup,* кога вредноста R_1 = 0.73 е генерирана, прво го наоѓаме интервалот во кој R_1 лежи. Во општ случај, $R = R_1$, ако

$$F(x_{i-1}) = r_{i-1} < R \leq r_i = F(x_i), \tag{3.10}$$

и тогаш поставуваме $X_1 = x_i$. Овде $r_0 = 0$, $x_0 = -\propto$, додека $x_1, x_2, \ldots, x_n$, се можни вредности на случајната променлива, и $r_k = p(x_1) + \ldots + p(x_k)$, $k = 1, 2, \ldots, n$. За овој пример, $n = 3$, $x_1 = 0$, $x_2 = 1$, $x_3 = 2$ и овде $r_1 = 0.5$, $r_2 = 0.8$ и $r_3 = 1$.

Значи r_1=0.5 <R_1 = 0.73 ≤r_2 = 0.83, и поставуваме X_1 = x_2 = 1. Или пократко запишано:

$$X = \begin{cases} 0, & R \le 0.5 \\ 1, & 0.5 < R \le 0.8 \\ 2, & 0.8 < R \le 1 \end{cases}$$

Пример 3.3 (дискретна рамномерна распределба) Се разгледува дискретна рамномерна распределба со следните веројатности и функцијата на распределба:

$$P(x) = 1/k \text{ , за } x = 1,2, \dots, k,$$

и

$$F(x) = \begin{cases} 0, & x < 1 \\ \frac{1}{k}, & 1 \le x < 2 \\ \frac{2}{k}, & 2 \le x < 3 \\ \vdots & \vdots \\ \frac{k-1}{k}, & k-1 \le x < k \\ 1, & k \le x \end{cases},$$

Нека $x_i = i$, и $r_i = p(1) + \dots + p(x_i) = F(x_i) = i/k$, за i=1, 2, . . . k. Прво се наоѓа интервалот во кој лежи точката R

$$r_{i-1} = i - 1/k < R \le r_i = i/k, \tag{3.11}$$

потоа X се генерира така што се поставува $X = i$, од каде

$$i - 1 < Rk \le i,$$

$$Rk \le i < Rk + 1, \tag{3.12}$$

Со $\lceil y \rceil$ се означува најмалиот цел број $\ge y$. Значи:

$$X = \lceil Rk \rceil. \tag{3.12}$$

За примерот, се разгледува генерирањето на случајната променлива X, рамномерно распределена на [1,2, . . . , 10] . Ако се земе k = 10, за некои случајни броеви R, може да се добијат следниве вредности на случајната променлива X.

$R_1 = 0.78 \quad X_1 = \lceil 7.8 \rceil = 8,$

$R_1 = 0.03 \quad X_1 = \lceil 0.3 \rceil = 1,$

$R_1 = 0.23 \quad X_1 = \lceil 2.3 \rceil = 3,$

$R_1 = 0.97 \quad X_1 = \lceil 9.7 \rceil = 10,$

Пример 3.4 (геометриска распределба): Геометриската распределба ја има следната распределба:

$$p(x) = p(1- p)^x, \; x = 0, 1, 2\dots, \; 0 < p < 1,$$

Функцијата на распределба е дадена со :

$$F(x) = \sum_{j=0}^{x} p(1-p)^x = \frac{p\left\{1-(1-p)^{x+1}\right\}}{1-(1-p)} = 1-(1-p)^{x+1} .$$

Повторно се користи техниката за инверзна трансформација, и во овој случај:

$$F(x - 1) = 1 - (1 - p)^x < R \leq 1 - (1 - p)^{x+1} = F(x).$$

Каде R е генериран случаен број помеѓу 0 и 1, со решавање на горното равенство се добива:

$$(1 - p)^{x+1} \leq 1 - R < (1 - p)^x ,$$

$$(x + 1)\ln(1 - p) \leq x < x\ln(1 - p).$$

Бидејќи, $1 - p < 1$, следува дека $\ln(1 - p) < 0$,

$$\ln(1 - R)/\ln(1 - p) - 1 \leq x < \ln(1 - R)/\ln(1 - p). \tag{3.14}$$

Значи, $X = x$, за некој цел број x, што го задоволува горното равенство се добива.

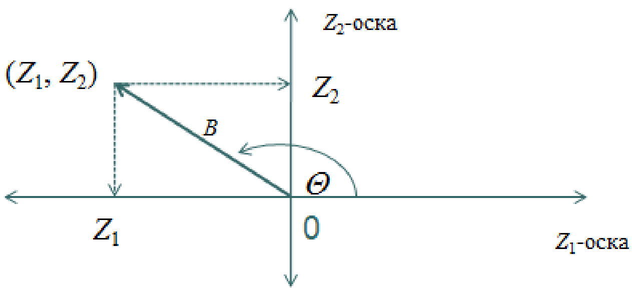
$$X = \left\lceil \frac{\ln(1-R)}{\ln(1-p)} - 1 \right\rceil. \tag{3.15}$$

3.1.7 Директна трансформација за нормална и логнормална распределба

За генерирање на нормалната распределба биле развивани многу методи. Техниката за инверзна трансформација во овој случај не може лесно да се искористи. Нормалната распределба ја има следната густина на распределба.

$$F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt, \qquad -\infty < x < \infty$$

Прво се опишува директната трансформација, со која се генерира нормирана нормална случајна променлива, со математичко очекување 0 и дисперзија 1. Со помош на други трансформации од нормираната нормална случајна променлива се добива нормална случајна променлива со математичко очекување μ и дисперзија σ^2. Ако постои метод за генерирање на случајна променлива X со $N(\mu, \sigma^2)$ распределба, тогаш може да се генерира и логнормална случајна променлива Y со параметри μ и σ^2, со директна трансформација $Y = e^X$.



Фиг. 3.5 Поларна репрезентација на пар од стандардна нормална распределба

Прво се разгледуваат две нормирано - нормално распределени случајни променливи Z_1 и Z_2, дадени како на Фигура 2.6 , и претставени со поларни координати :

$$Z_1 = B\cos\theta,$$
$$Z_2 = B\sin\theta. \tag{3.16}$$

Каде $B = Z_1^2 + Z_2^2$ има χ^2- квадрат распределба со два степени на слобода, што е еднакво на експоненцијална распределба со математичко очекување 2. Како што е покажано во делот за генерирање на експоненцијална случајна променлива со функција на распределба $F(x) = 1 - e^{-\lambda x}$, $x \geq 0$, се генерираат вредности $X_i = -1/\lambda \ln(1 - R_i)$ каде што R_i се генерирани случајни броеви, па за радиусот B ќе имаме:

$$B = (-2\ln R)^{1/2}. \tag{3.17}$$

За добивање на нормалната распределба, може да претпоставиме дека аголот θ е рамномерно распределен меѓу 0 и 2π радијани. Исто така радиусот B и аголот θ се независни. Со комбинирање на двете горни равенства се добива директен метод за генерирање на две независни нормални нормирани вредности Z_1 и Z_2, за два независни случајни броеви R_1 и R_2 :

$$Z_1 = (-2\ln R_1)^{1/2} \cos(2\pi R_2),$$
$$Z_2 = (-2\ln R_1)^{1/2} \sin(2\pi R_2). \tag{3.18}$$

За илустрирање на шемата нека $R_1 = 0.1758$ и $R_2 = 0.1489$. Тогаш се добива:

$$Z_1 = (-2\ln(0.1758))^{1/2} \cos(2\pi 0.1489) = 1.11,$$
$$Z_2 = (-2\ln(0.1758))^{1/2} \sin(2\pi 0.1489) = 1.50.$$

За добивање на нормални вредности X_i со математичко очекување μ и дисперзија σ^2, се применува следната трансформација:

$$X_i = \mu + \sigma Z_i.$$

Во горниот пример за математичко очекување $\mu = 10$ и дисперзија $\sigma^2 = 4$, се добива:

$$X_1 = 10 + 2(1.11) = 12.22,$$
$$X_2 = 10 + 2(1.5) = 13.$$

3.2 Метод на конволуција

Распределбата на веројатностите на сума од две или повеќе независни случајни променливи е наречена конволуција на распределбите на оригиналните распределби. Со овој метод со помош на две или повеќе случајни променливи се добива нова случајна променлива со саканата распределба. Оваа техника се користи за добивање на Ерлангова и биномна распределба
.

3.2.1 Ерлангова распределба

Случајната променлива X со Ерлангова распределба со параметри (K, θ), може да се добие од K независни експоненцијално распределени случајни променливи, X_i (i = 1,…, K), кои имаат математичко очекување $1/K\theta$

$$X = \sum_{i=1}^{K} X_i ,$$

Значи секој X_i може да биде генериран со равенката (4,3), со $1/\lambda = 1/K\theta$, и се добива:

$$X = \sum_{i=1}^{K} -\frac{1}{K\theta} \ln R_i = -\frac{1}{K\theta} \ln\left(\prod_{i=1}^{K} R_i\right). \tag{3.19}$$

3.3 Техника на прифаќање и отфрлање

Се претпоставува дека аналитичарот има потреба да добие метод за генерирање на случајна променлива X, рамномерно распределена меѓу 1/4 и 1. Еден начин за постигнување на оваа цел е опишан со следниве чекори:

Чекор 1. Се генерира случаен број R,

Чекор 2а. Ако $R \geq 1/4$, се дозволува $X = R$, и се оди на чекор 3,

Чекор 2б. Ако $R < 1/4$, се отфрла R, и се оди на чекор 1,

Чекор 3. Ако е потребна друга вредност рамномерно распределена на [1/4,1], се повторува целата постапка тргнувајќи од чекор 1, ако не овде се завршува постапката.

Значи со извршување на чекор 1, се генерира нов случаен број R. Со чекорот 2а овој број се прифаќа или со чекорот 2б, се отфрла. За завршување на постапката, случаен број R(во овој случај рамномерно распределена на [0,1]), се генерира се додека не се добие случаен број кој ќе задоволи некој услов (во случајот $R \geq 1/4$). Кога условот ќе биде задоволен, случајната променлива X (овде рамномерна на [0.1/4]), ќе биде пресметана ($X = R$). Значи случајната променлива R, ја нема распределбата која е потребно да се добие, но ако се земе дека R задоволува одредени услови (во случајот, $R \geq$ ¼), тогаш се добива посакуваната распределба. За да го покажеме ова се зеде, $1/4 \leq a < b \leq 1$, тогаш:

$$P\left(a < R \leq b \mid 1/4 \leq R \leq 1\right) = \frac{P\left(a < R \leq b\right)}{P\left(1/4 \leq R \leq 1\right)} = \frac{b-a}{3/4} \quad , \tag{3.20}$$

што е точно веројатноста на рамномерно распределена случајна променлива на [1/4,1]. Бидејќи ,$1/4 \leq R \leq 1$ се поставува $X = R$.

3.3.1 Пуасонова распределба

Пуасоновата распределба е дадена со распределбата на веројатностите:

$$p(n) = P(N = n) = e^{-\alpha}\alpha^n/n!, \quad n = 0,1,2, \ldots$$

каде N може да се интерпретира како број на пристигнувања на настани од паусоновиот влезен поток за единичен временски период. Нека времињата меѓу две последователни пристигнувања на клиентите A_1, A_2, ... се експоненцијално распределени со интензитет α. (α е средниот број на пристигнување во единица време). Во претходниот дел ја видовме постапката за генерирање на експоненцијалната распределба. Постои релација меѓу дискретната пуасонова распределба и непрекината експоненцијална распределба. Нека

$$N = n, \tag{3.21}$$

ако и само ако

$$A_1 + A_2 + \ldots + A_n \leq 1 < A_1 + A_2 + \ldots + A_{n+1}. \tag{3.22}$$

Равенката (3.21), $N = n$, значи дека има точно n пристигнувања за единечен временски период. Но следната релација значи дека n - тото доаѓање ќе се случи пред време 1, а n+1-то доаѓање ќе се случи после време 1. Овие две равенства се еквивалентни. Значи се става $N = n$.

За оваа цел ќе се искористи равенството за генерирање на експоненцијална распределба, применето на A_i (бидејќи тие се експоненцијално распределени). Значи $A_i = (-1/\alpha)\ln R_i$

Се изразува ова неравенство преку α и се добива:

$$\sum_{i=1}^{n} -\frac{1}{\alpha}\ln R_i \le 1 < \sum_{i=1}^{n+1} -\frac{1}{\alpha}\ln R_i ,$$

од овде се добива:

$$\ln \prod_{i=1}^{n} R_i = \sum_{i=1}^{n} \ln R_i \ge -\alpha > \ln \prod_{i=1}^{n+1} R_i = \sum_{i=1}^{n+1} \ln R_i ,$$

на крај со користење на релацијата $e^{\ln x} = x$, се добива:

$$\prod_{i=1}^{n} R_i \ge e^{-\alpha} > \prod_{i=1}^{n+1} R_i \tag{3.23}$$

Значи процедурата за генерирање на Пуасонова распределба ги содржи следниве чекори:

Чекор 1. Се поставува $n = 0$, $P = 1$,

Чекор 2. Се генерира случаен број R_{n+1} и се заменува P со $P \cdot R_{n+1}$,

Чекор 3. Ако $P<e^{-\alpha}$ тогаш се зема $N= n$, во спротивно се отфрла моменталното n, се зголемува за 1, и се оди на чекор 2.

Пример 3.5 (Пуасонова распределба): Се генерира Пуасонова распределба со параметар α=0.2. Прво се пресметува $e^{-\alpha} = e^{-0.2} =0.8187$. Потоа се генерираат случајни броеви и се следат чекорите од 1 до 3:

Чекор 1. Се поставува $n = 0$, $P = 1$,
Чекор 2.$R_1 = 0.4357$, $P =1\cdot R_1=0.4357$,
Чекор 3.$P = 0.4357 <e^{-\alpha} =0.8187$, па се прифаќа $N = 0$,

Чекор 1-3 (на ист начин се прифаќа $N = 0$, и за R_1=0.4146).

Чекор 1. Се поставува $n = 0$, $P = 1$,
Чекор 2. R_1=0.8353, $P=1\cdot R_1=0.8353$,
Чекор 3. $P=0.0.8583 \ge e^{-\alpha} =0.8187$, па се отфрла n=0 и се враќаме на чекор 2, со n=1

Чекор 2. $R_2=0.9952$, $P=R_1\cdot R_2=0.8313$
Чекор 3. $P=0.0.8313 \ge e^{-\alpha} =0.8187$, па се отфрла n=1 и се враќаме на чекор 2, со n=2

Чекор 2. $R_3=0.8004$, $P=R_1\cdot R_2\cdot R_3=0.6654$
Чекор 3. $P=0.6654< e^{-\alpha} =0.8187$, па се прифаќа N=2

4. МОДЕЛИРАЊЕ НА ВЛЕЗНИ ПОДАТОЦИ

Влезните податоци се еден од најважните елементи на симулациониот модел. Во симулирањето на системите за масовно опслужување, влезни податоци се распределбата на времето помеѓу две последователни пристигнувања на клиентите и времето на опслужување на клиентите. Во моделот на залиха, влезен податок е распределбата на побарувањето. Да нагласиме дека во апликациите кои симулираат реални системи, определувањето на соодветните распределби на влезните податоци е многу значајно, бидејќи и при најдобро изработен модел, ако влезните податоци и претпоставки се лошо избрани, тогаш и интерпретацијата на излезните податоци ќе биде погрешна, а со тоа и добиениот резултат ќе биде лош. Најчесто доколку се има добар симулационен софтвер, делот за собирање на податоци побарува многу време и познавање на проблемот.

Постојат четири чекори за развој на успешен модел на влезните податоци:

1. Собирање податоци од реалниот систем што е цел на проучувањето. Најчесто се собираат податоци за времето и ресурсите. Понекогаш, во некои системи не можат да се соберат податоци (на пример, времето е екстремно ограничено, или кога влезниот процес не постои, таков пример е ако сакаме да градиме мрежа на патишта не може да се обезбеди примерок од влезни податоци бидејќи системот не постои, или кога некои правила или законитости го спречуваат собирањето на податоци. Кога податоците не се достапни, се користат мислењата и знаењата на експертите.

2. Идентификација на веројатностите распределби (функцијата на распределба и функцијата на густината на распределба). Кога податоците се достапни, овој чекор типично започнува со определување на распределбата на фреквенциите или хистограмите на податоците. Врз база на распределбата на фреквенциите и структурните познавања на процесот се бираат соодветните фамилии распределби. Најчесто се користат добро познатите стандардни теориски распределби како добра апроксимација во пракса.

3. Избор на параметрите со што се определуваат специфични инстанци на фамилиите распределби. Кога влезните податоци се достапни, врз нивна база овие параметри може да бидат оценети.
4. Оценување на избраните распределби и параметрите на распределбите со тестови за поклопување со теориските распределби (*goodness-of -fit*). *Hi*-квадрат тестот и Колмогоров-смирнов тестот се стандардните *goodness-of – fit* тестови. Ако избраната распределба не е добра апроксимација на податоците тогаш анализата се враќа на вториот чекор и се повторува постапката.

4.1 Собирање на податоци

Собирањето на податоците е една од најважните задачи во решавањето на реалните проблеми. Тоа е еден од најзначајните и најтешките проблеми во процесот на симулирање. Исто така кога податоците се достапни, многу ретко тие се во форма така што можат директно да се користат за влезен модел во симулацијата. Ако влезните податоци и претпоставки се лоши, тогаш и резултатот е лош- „*GIGO*” или „*garbage-in , garbaгe-out*”. *„GIGO”* е основниот концепт во компјутерската наука што се применува во симулирањето на дискретни системи. Дури и кога структурата на моделот е валидна, ако влезните податоци не се точно избрани, не соодветно анализирани, или не се репрезентативни за околината, излезните податоци ќе бидат погрешни.

Пример 4.1 (систем за перење и сушење):Системот за перење и сушење е еден пример на поедноставен систем, и се состои од 10 машини за перење и 6 машини за сушење. Меѓутоа од аспект за собирање на податоците проблемот брзо станува прилично голем. Распределбата меѓу две последователни пристигнувања не е хомогена, таа се менува во тек на денот, и во истиот период во различни денови. Сервисот е отворен 7 дена во неделата по 16 часа дневно, или 112 часа неделно. Да претпоставиме дека нема доволно ресурси за да системот постојано се набљудува (на пример само две личности се вклучени во процесот на симулирањето) а исто така постои и временска ограниченост (симулацијата треба да се заврши за 4 недели). Распределбата меѓу две пристигнувања во текот на една недела не мора да биде иста со таа во следната недела. Како компромис, избран е одреден временски интервал. Интервалот меѓу две последователни пристигнување беше определен и класифициран во согласност со интензитетот на доаѓање (можеби несоодветно) како „висок", „среден " и „низок". Распределбата на времето на опслужување го прави проблемот потежок од повеќе аспекти. Опсегот на барањата за опслужување на корисниците може да биде различен(некои да бараат само перење, некои и перење и сушење, некои само сушење итн.), па поради тоа е потребно набљудување и регистрирање. Наједноставен случај е кога корисникот бара една машина на перење , а веднаш потоа една машина за сушење. Меѓутоа корисниците може да побаруваат две машини за перење и една машина за сушење, или само една машина за сушење. Поради тоа што барања за перење и барањата за сушење зависат од самите клиенти, од каде што произлегува дека перењето и сушењето се независни операции во овој процес, па поради тоа и времето на перење и времето на сушење се земаат како независни случајни променливи. Исто така постои разлика и меѓу времето на задржување на клиентите во системот, некои клиенти чекаат во системот се додека нивните алишта не бидат готови, други пак ги оставаат своите алишта и потоа кога ќе бидат готови се враќаат и ги земаат. Исто така и машините може да се расипат во тек на време. Должина на периодот кога машината ќе биде расипана, може да се разликува (на пример ако машината се расипе во петок навечер, проблемот не може да биде фиксиран до понеделник).

Следните сугестии може да го зголемат и олеснат процесот на собирање на податоци:

1. Корисно е да се испланира колку време ќе се потроши за собирањето на податоците. Тоа може да се направи во текот на собирањето на податоците или пред да започне самото собирање. Значи потребен е план за собирање на податоците и како тоа се остварува во текот на времето. Планирањето е исто така значајно ако податоците се добиваат автоматски (на пример преку компјутерско собирање), поради сигурност дека соодветни податоци се достапни. Исто така кога податоците ќе бидат собрани, треба со сигурност да се знае дека ќе има доволно време да се конвертираат податоците во формат во кој што ќе можат да бидат искористени.

2. Обид за анализа на податоци додека се собираат. Да се определи дека податоците што ќе бидат собрани ќе ја дадат соодветната распределба за влезот на симулацијата. Исто така треба да се определи дали некои податоци не се потребни за симулацијата.

3. Проверка на хомогеноста на множеството податоци. Хомогеноста да се провери во последователни временски период и за време на ист временски период во последователни денови. На пример, да се провери хомогеноста на податоците од два до три часот, и од три до четири часот, исто така потребно е да се види дали податоците се хомогени од два до три часот во

четврток и петок. Со проверка на хомогеноста, се гледа дали распределбите (на пример времето на пристигнувањето) се еднакви. За ова цел двапати t-тестот може да биде искористен.

4. Можност да се земат само дел од податоците, да се отфрлат оние кои не се од потполн интерес на испитувањето. Овој проблем многу често се случува кога аналитичарот го интересира времето потребно за некој процес (на пример времето на преглед на пациентот но не и времето потребно за пополнување на документацијата), но започнува пред или завршува после комплетирањето на испитуваниот период.

5. Определување дали постои релација меѓу две променливи, изградба на scatter дијаграм. Понекогаш скенирањето на *sctter* дијаграмот ќе иницира дали постои релација меѓу две променливи кои се од интерес на испитувањето.

6. Оценување, на веројатноста за автокорелација на низата податоци во последователни моменти на време. На пример времето на опслужувањето на i-от клиент може да влијае на времето на опслужување на (i+n)-от клиент.

7. Да се разликуваат влезните и излезните податоци. Во системите за масовно опслужување времето меѓу пристигнувањето на клиентите е вообичаено влезен податок, додека задоцнувањето на почнувањето на опслужување на клиентот е излезен податок.

4.2 Идентификација на распределби преку податоци

Во овој дел се разгледани методите за избор на фамилии од влезни распределби кога податоците се достапни.

4.2.1 Хистограми

Фреквентната распределба или хистограмот е корисна за идентификација на формата на распределбата. Хистограмот се конструира на следниот начин:

1. Се реди рангот на податоци во интервали -интервалите се вообичаено со еднакви должини, нееднакви интерлвали може да се користат ако се прилагодуваат висините на фреквенциите.

2. На хоризонталната оска се означуваат селектираните интервали.

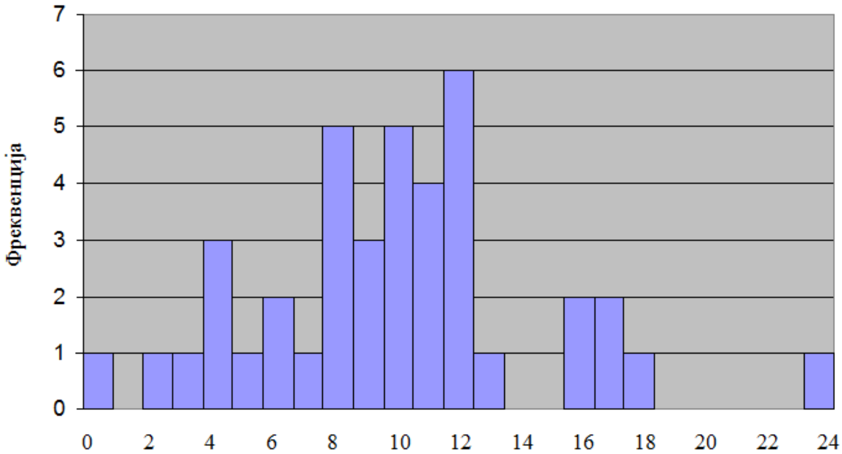
3. Се определуваат фреквенциите на настаните во сите интервали.

4. Се означува вертикалната оска така да може вкупните случувања да бидат претставени за секој интервал.

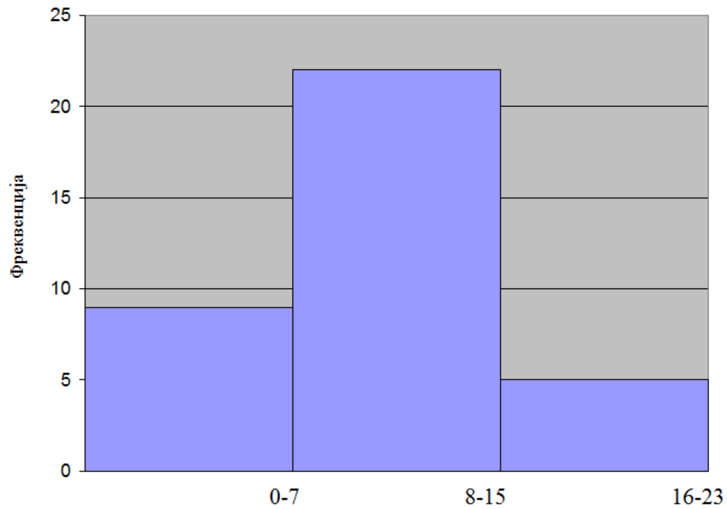
5. Се означуваат фреквенциите на вертикалната оска.

Бројот на интервали зависи од бројот на набљудувања и од расфрлегоста или диспрезиите на податоците. *Hines и Montgomery* (1990) воочиле дека изборот на бројот на интервали апроксимативно е еднаков на квадратниот корен од големината на примерокот со кој се работи во пракса. Ако интервалите се многу широки хистограмите ќе бидат необработливи и нивната форма и другите детали нема да бидат добро прикажани. Ако интервалите се многу тесни, хистограмот ќе биде искинат (расцепкан) и податоците не ќе бидат израмнети. Пример на расцепкан, необработли и соодветен хистограм со користење на истите податоци е прикажан во Фигурата 4.1

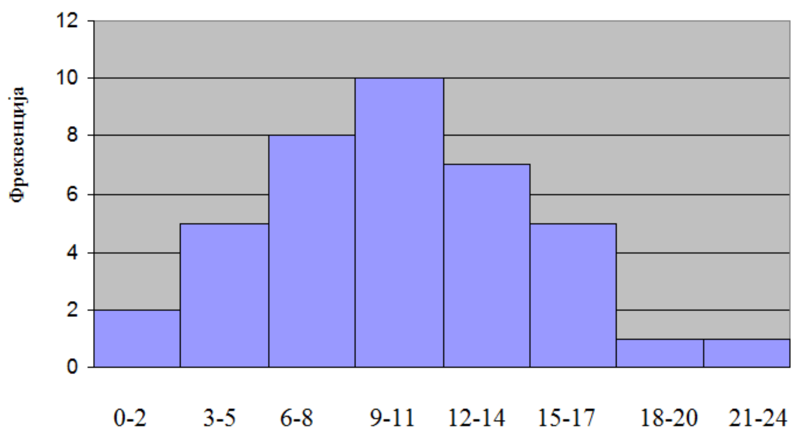
Современите софтвери често овозможуваат големината на интервалите да се менува лесно и интерактивно се додека не се најде добар избор



a)



(b)



(c)

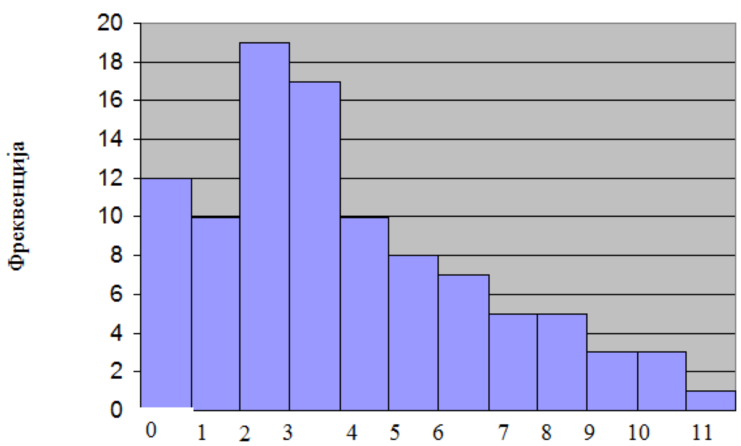
Фиг.4.1 Расцепкан, необработлив и соодветен хистограм; (a)оригинални податоци - многу расцепкани;(b)комбинирање соседни келии-многу необработлив;(c) комбинирае соседни келии –соодветен

Хистограмите за непрекинатите податоци кореспондираат на функцијата на густината на распределба на теориската распределба. Затоа е потребно да се определи оптимална ширина на интервалите што ќе даде најдобра претстава на соодветната теориска функција на густината на распределбата.

Хистограмите за дискретните податоци, во случај кога е достапен голем број од податоци е потребно ќелија за секоја вредност од рангот на вредности. Меѓутоа кога има помалку податоци, тогаш може да е неопходно комбинирање на соседните келии за елиминирање на расцепканоста од изгледот на хистограмот.

Пример 4.2 (дискретни податоци): Бројот на коли што пристигнуваат во северозападниот агол во период од 5 минути од 7 до 7:05 часот, бил набљудуван сите 5 работни дена во текот на 20 недели. Од податоците се забележува дека во 12периоди со траење од 5 минути не пристигнува кола, во 10 пристигнува 1 кола итн. Значи бројот од автомобили е дискретна случајна променлива, постои еден примерок од податоци, и хистограмот има ќелија за секоја можна вредност од рангот на податоците. Податоците се прикажани на Фиг. 4.2.

Пристигнувања во период	Фреквенција
0	12
1	10
2	19
3	17
4	10
5	8
6	7
7	5
8	5
9	3
10	3
11	1



Фигура 4.2 Хистограм за бројот на пристигнувања во период

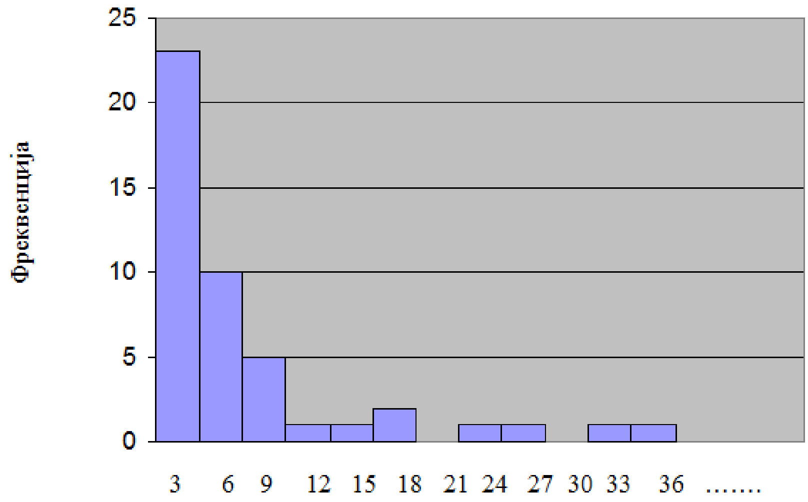
Пример 4.3 (неприкинати податоци)Се разгледува времето на живот на случајно избран примерок од електонски чипови. Нивните времиња на живот се:

79.919	3.081	0.062	1.961	5.845
3.027	6.505	0.021	0.013	0.123
6.769	59.899	1.192	34.760	5.009
18.387	0.141	43.565	24.420	0.433
144.695	2.663	17.967	0.091	9.003
0.941	0.878	3.371	2.157	7.579
0.624	5.380	3.148	7.078	23.960
0.590	1.928	0.300	0.002	0.543
7.004	31.764	1.005	1.147	0.219
3.217	14.382	1.008	2.336	4.562

Времето на живот на чиповите е непрекината случајна променлива, и дадена е со точност со 3 децимали. Хистограмот е направен со сместување на податоците во класи од интервали.

Време на живот на чиповите (во денови)	Фреквенција
$0 \le x_j < 3$	23
$3 \le x_j < 6$	10
$6 \le x_j < 9$	5
$9 \le x_j < 12$	1
$12 \le x_j < 15$	1
$15 \le x_j < 18$	2
$18 \le x_j < 21$	0
$21 \le x_j < 24$	1
$24 \le x_j < 27$	1
$27 \le x_j < 30$	0
$30 \le x_j < 33$	1
$33 \le x_j < 36$	1
.	.
.	.
.	.
$42 \le x_j < 45$	1
.	.
.	.
.	.
$57 \le x_j < 60$	1
.	.
.	.
.	.
$78 \le x_j < 81$	1
.	.
.	.
.	.
$144 \le x_j < 147$	1

Табела 4.1 Податоци за животот на електронските чипови



Фигура 4.3 Хистограм на времето на живот на електронските чипови

4.2.2 Селектирање на фамилии од распределби

Целта на хистограмот е да се избере позната функција на густината на распределбата или функција на распределба. Функцијата на распределби се избира врз основа на формата на хистограмот. Најчесто се зема една од следниве распределби:

Биномна распределба која го дава бројот на успеси во n-независни обиди. Веројатноста за успех е *p*. Пример: бројот на расипани компјутерски чипови најдени во *n*-чипови.

Негативна биномна распределба (ја вклучува геометриската распределба) Модели за бројот на обиди за да се добијат *k*-успеси. Пример: бројот на обиди што треба да се извршат за да се добијат *k*-расипани чипови.

Пуасонова распределба, бројот на независни настани што се случуваат во однапред познат интервал од време или фиксиран простор. Пример: бројот на корисници што пристигнуваат во продавницата за време од еден саат, или бројот од дефекти најден во 30m^2 метална жица.

Нормална распределба, ја моделира распределбата на процес што може да биде претставен како сума од потпроцеси . Пример: времето за составување продукт што е сума од потребните времиња за сите операции. Да се забележи дека ако нормалната распределба прима негативни вредности тогаш таа не може да се користи за распределба на времето на некој процес.

Логнорормалн распределба, ја моделира распределбата на процес што може да биде претставен како производ од потпроцеси.

Екпоненцијна распределба го моделира времето меѓу независни настани , во овој случај времето на процесирање го има својството „*memoryless*" – без меморија, (ако се знае колку време трае процесот, тоа не дава никаква информација уште колку време е потребно за да се комплетира процесот). Ова е најчесто користена распределба. Да се напомене дека ако времето меѓу пристигнувањето на клиентите е експоненцијално распределено, тогаш бројот на клиенти што ќе пристигнат во фиксиран интервал од време има Пуасонова распределба.

***Gamma* распределба,** екстремно приспособлива распределба која се користи како модел за ненегативни случајни променливи.

***Beta* распределба,** екстремно приспособлива распределба која се користи како модел за ограничени случајни променливи.

Ерлангова распределба, се користи за моделирање на процеси што може да бидат разгледувани како сума од неколку експоненцијално распределени процеси. Ерланговата е специјален случај од *Gamma* распределбата.

Weilbull распределба, за моделирање на времето на расипување на компонентите. На пример, времето на оштетување на хард диск. Експоненцијалната распределба е специјален случај од Weilbull распределбата.

Дискретна или непрекината рамномерна распределба, за моделирање на комплетно неизвесни процеси, кога сите излези се еднакво веројатни.

Триаголна распределба, за моделирање на процеси кога максималната најверојатна и минималната вредност на процесот се познати.

Емпириска распределба, примероци од актуелно собраните податоци се користат кога теориските распределби не се применливи.

Исто така не треба да се игнорираат физичките карактеристики на процесите кога се избираат распределбите. Дали процесот прима дискретни вредности? Дали е ограничен или не? Овие знаења, кои не зависат од податоците може да помогнат при изборот на фамилии од распределби.

4.2.3 Q-Q постапка (*Quantile- Quantile plots*)

Конструкцијата на хистограми и препознавањето на формата на распределбата се неопходен составен дел за селектирање на фамилијата од распределби за дадениот примерок од податоци. Меѓутоа хистограмите не се користат за оценување дали е соодветна избраната распределба. Q-Q постапката е корисна алатка со која се оценува дали распределбата е добро избрана. Ако X е случајна променлива со функција на распределба F, тогаш q-квантил на X е вредноста y, за која $F(y)=P(X\leq y)=q$, за $0<q<1$, кога F има инверзна функција, $y=F^{-1}(q)$.

Нека $\{x_i,\ i=1,2,\ldots,n\}$ е примерок од податоци од X. Ги подредуваме елементите од примерокот по големина во растечки редослед, $\{y_j,\ j=1,2,\ldots,n\}$, каде $y_1 \leq y_2 \leq \ldots \leq y_n$, со j се означува редоследот на елементот во примерокот. Q-Q постапката е базирана на фактот дека y_j е оценка на $(j-1/2)/n$ квантил на X, со други зборови y_j е апроксимација на $F^{-1}\left(\frac{j-1/2}{n}\right)$.

Нека е избрана теориска распределба со функција на распределба F, како можна репрезентација на X. Ако Fе член од соодветната фамилија распределби тогаш со цртање на кривата $\{y_j,\ F^{-1}\left(\frac{j-1/2}{n}\right)\}$, треба да се добие права линија.

За развојот на Q-Q постапката треба да се каже следново:

1. Набљудуваните вредности никогаш нема да дадат точно права линија.

2. Подредените податоци не се независни, значи тие се рангирани (класифицирани). Ако една точка е над рамната линија тогаш веројатно следната точка ќе биде исто така над линијата. И исто така е неверојатно дека точките ќе бидат расфрлани околу линијата.

3. Дисперзиите на екстремните вредности (најмалата и најголемата) се многу поголеми од дисперзиите во средината на рамната површина. Поголеми дисперзии може да бидат прифатени за екстремните вредности. Линеарноста на точките во средината е многу позначајна од линеарноста на екстремите.

Модерните софтвери за анализа на податоци вклучуваат алатки за генерирање на Q-Qplots, посебно за нормална распределба. Q-Qplots исто така може да се користат за споредба дали два примерока од податоци може да се репрезентираат со иста распределба.

4.3 Оценување параметри

Откако ќе бидат избрани фамилиите распределби, следен чекор е да се оценат параметрите на распределбите. Многу софтверски пакети (во некои од нив се вградени симулациони јазици) се сега возможни за пресметување оценувачи.

4.3.1 Прелиминарни статистики: Средна вредност на примерокот и дисперзија на примерокот

Во голем број од примероците, средната вредност на примерокот или средната вредност на примерокот и дисперзијата на примерокот се користат за да се оценат параметрите на хипотетичката распределба.

Равенките (4.1) и (4.2) се користат кога дискретни или непрекинати необработени податоци се достапни.

Ако се разгледува примерок со големина n, X_1, X_2,..., X_n, средната вредност $(\overline{X})$ е дефинирана со:

$$\overline{X} = \sum_{i=1}^{n} \frac{X_i}{n}, \tag{4.1}$$

и дисперзијата на примерокот е дефинирана со:

$$S^2 = \frac{\sum_{i=1}^{n} X_i^2 - n\overline{X}}{n-1}. \tag{4.2}$$

Равенките (4.3) и (4.4) се користат кога податоците се дискретни и се групирани по распределби на фреквенциите. Во овој случај средната вредност на примерокот се пресметува со:

$$\overline{X} = \frac{\sum_{j=1}^{k} f_j X_j}{n}, \tag{4.3}$$

и дисперзијата на примерокот со:

$$S^2 = \frac{\sum_{j=1}^{k} f_j X_j^2 - n\overline{X^2}}{n-1}, \tag{4.4}$$

каде k е бројот од различни вредности на X и f_j е набљудуваната фреквенција на вредностите X_j на X.

Пример 4.4 (Групирање на податоци): Податоците од примерот 4.2 може да се анализираат при што се добива n= 100, f_1=12,X_1=0, f_2= 10, X_2=1, $\sum_{j=1}^{k} f_j X_j = 364$ и $\sum_{j=1}^{k} f_j X_j^2 = 2080$. За равенката 4.3 се добива:

$$\overline{X} = \frac{364}{100} = 3.64,$$

и за равенката 4.4:

$$S^2 = \frac{2080 - 100(3.64)^2}{99} = 7.63.$$

стандардната девијација на примерокот σе квадратен корен од дисперзијата, во овој случај $\sigma = \sqrt{7.63} = 2.76$

Равенките 4.5 и 4.6 се користат кога податоците се дискретни или непрекинати и се распоредени во интервали. Овие равенки се апроксимативни и може да се користат кога необработените податоци не се достапни, и со нив не е можно да се добијат точни вредности за средната вредност и дисперзијата:

$$\overline{X} = \frac{\sum_{j=1}^{c} f_j m_j}{n}, \tag{4.5}$$

$$S^2 = \frac{\sum_{j=1}^{c} f_j m_j - n\overline{X^2}}{n-1}. \tag{4.6}$$

каде f_jе фреквенцијата на j-от интервал,n_j е средната точка на j-от интервал, и c е бројот на интервали.

4.3.2 Предложени оценувачи

Нумеричките оценувачи на параметрите на распределбите се потребни за редуцирање на фамилијата распределби до специфична распределба и за тестирање на хипотези. Во табелата 4.2 се дадени оценувачите за распределбите кои најчесто се користат во процесот на симулирање, базирани на необработливи податоци.

Овие оценувачи се максимални најверојатни оценувачи, со исклучок на оценувачот на дисперзијата σ^2 на нормалната распределба.

Нека $f(x;\theta)$ е функција на густината на распределба со непознат параметар (вектор) θ. Ако X= (X_1,X_2,…, X_n) е случаен примерок за f, тогаш:

$$L(\theta, X) = \prod_{i=1}^{n} f(X_i, \theta), \tag{4.7}$$

доколку $\hat{\theta}$ го задоволува равенството:

$$L(\hat{\theta}, X) = \sup_{\theta} L(\theta, X), \tag{4.8}$$

тогаш $\hat{\theta}$ е максимален најверојатен оценувач за θ.

Во долната табела 4.2 се дадени оценувачите за најкористените распределби во процесот на симулирање

Распределба	Параметар	Оценувач

Пуасонова	α	$\hat{\alpha} = \overline{X}$
Екпоненцијална	λ	$\hat{\lambda} = \frac{1}{\overline{X}}$
Gamma	$\beta,\ \theta$	β од таблица $\hat{\theta} = \frac{1}{\overline{X}}$
Нормална	μ , σ^2	$\hat{\mu} = \overline{X}, \hat{\sigma}^2 = S^2$
Логнормална	μ , σ^2	$\hat{\mu} = \overline{X}, \hat{\sigma}^2 = S^2$
Weibull	α, β	$\hat{\beta}_0 = \frac{X}{S}$ $\hat{\beta}_j = \hat{\beta}_{j-1} - \frac{f\left(\hat{\beta}_{j-1}\right)}{f'\left(\hat{\beta}_{j-1}\right)}$, $f(\beta)$ и $f'(\beta)$ се добиваат од следниве равенки: $f(\beta) = \frac{n}{\beta} + \sum_{i=1}^{n} \ln X_i - \frac{n \sum_{i=1}^{n} X_i^{\beta} \ln X_i}{\sum_{i=1}^{n} X_i^{\beta}}$ $f'(\beta) = -\frac{n}{\beta^2} - \frac{n \sum_{i=1}^{n} X_i^{\beta} (\ln X_i)^2}{\sum_{i=1}^{n} X_i^{\beta}} + \frac{n(\sum_{i=1}^{n} X_i^{\beta} \ln X_i)^2}{(\sum_{i=1}^{n} X_i^{\beta})^2}$ $\hat{\alpha} = \left(\frac{1}{n} \sum_{i=1}^{n} X_i^{\hat{\beta}}\right)^{1/\hat{\beta}}$

Табела 4.2 Оценувачи на најкористени распределби

4.4 Селектирање влезни модели без податоци

Понекогаш во пракса е потребно да се развие симулирачки модел (можеби за демонстративни цели или прелиминарни студии) пред било какви податоци да се достапни:

„Инженерски податоци” - често некои информации за продуктите или процесите може да бидат добиени од производството. Пример: животниот век на диск е 5000 часови; брзината на ласерскиот печатач е 4 стр./min.

Мислења на експертите –станува збор за луѓе кои имаат искуство со процесот или со други слични процеси. Тие можат да дадат песимистички, оптимистички или најверојатни времиња на траење на некоја активност.

Физички или други ограничувања. - Многу реални процеси имаат физички граници на перформансите. Пример: Внесување на компјутерските податоци не може да биде побрзо отколку брзината на куцање на личноста која внесува податоци.

Природа на процесот- кога податоците не се достапни, рамномерната, триаголната и *Beta* распределбата најчесто се користат како влезен модел.

Пример 5.7(Планирање производство): Продавачот тврди дека од продуктот XYZ-123 ќе бидат продадени не помалку од 1000 производи и не повеќе од 5000 производи. Врз база на своето искуство, тој тврди дека постојат повеќе од 90% шанси продажбата да биде поголема од 2000 производи, 25% да е поголема од 3500 производи и само 1% да е поголема од 4500 производи. Резултатите се сумирани во Табела 4.3

I	Интервали (во часови)	Кумулативни фреквенции c_i
1	$1000 \le x \le 2000$	0.10
2	$2000 \le x \le 3500$	0.75
3	$3500 \le x \le 4500$	0.99
4	$4500 \le x \le 5000$	1.00

Табела 4.3 Сумирање на информациите од продавачот

4.5 Зависни податоци, мултиваријантен влезен модел и влезни модели од временски серии

Доколку променливите кои се појавуваат како влезови во симулирачкиот модел се во релација, тогаш треба да се определат овие релации и да се земат во предвид. Кога имаме зависни променливи се користат мултиваријантниот влезен модел и временските серии.

Мултиваријантен модел се користи во случај на конечен број на случајни променливи кои не се меѓусебно независни.

Временски серии се користат ако концептуално постојат бесконечен број меѓусебно зависни случајни променливи.

4.5.1 Коваријанса и корелација

Нека X_1 и X_2 се две случајни променливи и $\mu_i=E(X_i)$ и $\sigma_i^2 = Var(X_i)$ се математичкото очекување и дисперзијата на X_i соодветно.

Коваријансата и корелацијата се мерки на линеарната зависност меѓу X_1 и X_2. Со други зборови коваријансата и корелацијата укажуваат колку добро релацијата меѓу X_1 и X_2 е опишана со моделот:

$$(X_1 - \mu_1) = \beta(X_2 - \mu_2) + \varepsilon,$$

каде ε е случајна променлива со математичко очекување 0, која е независна од X_2.
Ако $\varepsilon = 0$, моделот е перфектен.
Ако X_1 и X_2 се независни тогаш $\beta = 0$, и во овој случај немаме модел.
Ако $\beta > 0$, тогаш и X_1 и X_2 се над или под нивните средни вредности, а кога $\beta < 0$, тогаш X_1 и X_2 ќе бидат на спротивната страна од средната вредност.

Коваријансата меѓу X_1 и X_2 е дефинирана со:

$$cov(X_1, X_2) = E[(X_1 - \mu_1)(X_2 - \mu_2)] = E(X_1, X_2) - \mu_1\mu_2 \quad (4.9)$$

ако $co\ v(X_1, X_2)=0$ следи $\beta = 0$, додека ако $cov(X_1, X_2)<0$ (>0) следи $\beta < 0$ (> 0)
Коваријансата може да прими вредности од $-\infty$ до $+\infty$. Корелацијата ја стандардизира коваријансата да биде помеѓу -1 и 1.

$$\rho = corr(X_1, X_2) = \frac{cov(X_1, X_2)}{\sigma_1\sigma_2} \quad (4.10)$$

Повторно ако $corr(X_1,X_2) = 0$ следи $\beta = 0$, додека ако $corr(X_1,X_2) < 0$ (>0) следи $\beta < 0$ (> 0).
Ако $\rho = -1$ или $\rho = 1$ постои строга линеарна зависност меѓу X_1 и X_2.

Нека има низа од случајни променливи X_1, X_2, . . . што се идентично распределени (имаат исти математички очекувања и дисперзии, но можат да бидат зависни). Се разгледува оваа низа како временска серија. Со $cov(X_t, X_{t+h})$ и $corr(X_t, X_{t+h})$ се означуваат lag-*h*ау автокорелацијата и lag-*h* автодисперзијата. Ако вредностите на оваа автодисперзија зависат само од *h* а не и од *t*, тогаш се нарекува стационарна коваријанса. За временски серии со стационарна автоковаријанса се користи кратенката:

$$\rho_h = corr(X_t, X_{t+h})$$

за lag-*h* автоковаријанса, и со ρ_hсе мери зависноста меѓу случајните променливи од временската серија што се одделени со други *h*-1 случајни променливи.

4.5.2 Мултиваријантни влезни модели

Ако X_1 и X_2 се нормално распределени случајни променливи тогаш зависноста меѓу нив може да биде моделирана со биваријантна нормална распределба со параметри μ_1, μ_2, σ_1^2 , σ_2^2 и $\rho = corr(X_1, X_2)$.

За оценка на ρ, претпоставуваме дека имаме *h* независни и еднакво распределени парови (X_{11},X_{21}), (X_{21},X_{22}), . . . , (X_{1n},X_{2n}). Тогаш коваријансата на примерокот е

$$\hat{\text{cov}}(X_1, X_2) = \frac{1}{n-1}\sum_{j=1}^{n}\left(X_{1j} - \overline{X_1}\right)\left(X_{2j} - \overline{X_2}\right) = \frac{1}{n-1}\left(\sum_{j=1}^{n} X_{1j}X_{2j} - n\overline{X_1}\,\overline{X_2}\right), \tag{4.11}$$

каде $\overline{X_1}$ и $\overline{X_2}$ се математичките очекувања на примероците. Корелацијата е оценета со:

$$\hat{\rho} = \frac{\hat{\text{cov}}(X_1, X_2)}{\hat{\sigma}_1 \hat{\sigma}_2}, \tag{4.12}$$

каде $\hat{\sigma}_1$ и $\hat{\sigma}_2$ се стандардни девијации на примерокот.

4.5.3 Временски серии

Нека X_1, X_2, . . . е низа од идентични но зависни и со стационарни коваријанси случајни променливи.

Опишани се два модели, за кои автокорелација ќе има форма:

$$\rho_h = corr\,(X_t, X_{t+h}) = \rho^h,$$

Каде *h*=1,2,… Треба да се забележи дека lag-*h* автокорелацијата се намалува геометриски со зголемувањето на lag-от, со што набљудувањата после долг временски интервал се скоро независни. Во првиот модел што ќе се разгледува , случајните променливи се нормално распределени ,а во вториот модел се експоненцијално распределени.

AR(1) модел

Се разгледува временската серија:

$$X_t = \mu + \phi(X_{t-1} - \mu) + \xi_t \tag{4.13}$$

за t=2,3, . . . , каде ξ_2, ξ_3, ... се независни и еднакво распределени случајни променливи со математичко очекување 0 и дисперзија $\hat{\sigma}_\varepsilon$ и -1 <ϕ<1. Ако иницијалната вредност X_1 е избрана апроксимативно тогаш X_1,X_2,. . . се нормално распределени случајни променливи со математичко очекување μ и дисперзија $\sigma_\varepsilon^2 / (1-\phi^2)$ и $\rho_h = \phi^h$ за h =1,2, . . .

Ваквиот модел е наречен авторегресивен модел од ред 1 или кратко AR(1). Оценката на параметарот ϕ, може да биде дефинирана со

$$\phi=\rho^1 = corr(X_t, X_{t+1})$$

со lag-1 автокорелацијата. Поради тоа за да се оцени ϕ, прво се оценува lag -1 автоковаријансата со:

$$\hat{\mathrm{cov}}(X_t, X_{t+1}) = \frac{1}{n-1}\sum_{t=1}^{n}(X_t - \overline{X})(X_{t+1} - \overline{X}) = \frac{1}{n-1}\left(\sum_{t=1}^{n-1} X_t X_{t+1} - (n-1)\overline{X^2}\right), \quad (4.14)$$

и дисперзијата $\sigma^2 = Var\ (X)$со вообичаената оценка $\hat{\sigma}^2$. Тогаш

$$\hat{\phi} = \frac{\hat{\mathrm{cov}}(X_t, X_{t+1})}{\hat{\sigma}^2},$$

конечно се оценуваат μ и σ^2 со $\hat{\mu} = \overline{X}$ и $\hat{\sigma}_\varepsilon^2 = \hat{\sigma}^2(1-\hat{\phi}^2)$.

Со следниот алгоритам е дадена постапка за генерирање на стационарна AR(1) временска серија, дадена со вредностите на параметрите ϕ, μ и $\hat{\sigma}_\varepsilon^2$.

1. Се генерира X_1 со нормална распределба со матетатичко очекување μ и дисперзија $\sigma_\varepsilon^2 / (1-\phi^2)$. Се поставува t=2,
2. Се генерира ε_t со нормална распределба со математичко очекување 0 и дисперзија $\hat{\sigma}_\varepsilon^2$,
3. Се поставува $X_t = \mu + \phi(X_{t-1} - \mu) + \xi_t$,
4. Се поставува t=t+1 и се оди на чекор 2.

EAR(1) модел

Се разгледува следниот модел на временска серија:

$$X_t \begin{cases} \phi X_{t-1} & \text{со веројатност } \phi \\ \phi X_{t-1} + \varepsilon_t & \text{со веројатност } 1-\phi \end{cases} \quad (4.15)$$

за t=2,3, . . . каде ε_2, ε_3 се независни и еднакво распределени со експоненцијална распределба со математичко очекување $1/\lambda$ и $0 \le \phi \le 1$. Ако иницијалната вредност X_1 е избрана апроксимативно, тогаш X_1, X_2, . . . се екпоненцијално распределени со математичко очекување $1/\lambda$ и $\rho_h = \phi^h$ за h=1,2, . . .

Оваа временска серија е наречена екпоненцијална авторегресивна од ред 1 или кратко EAR(1). Само автокорелациите поголеми од 0 може да бидат претставени со овој модел. Оценката на параметарот продолжува со примена на процедурата кај

AR(1), со поставување $\hat{\phi} = \hat{\rho}$, оценување на lag-1 автокорелацијата и поставување $\hat{\lambda} = 1/\bar{X}$.

Со следниот алгоритам е дадена постапка за генерирање на стационарни EAR(1) временска серија дадена со вредностите на параметрите ϕ и λ:

1. Се генерира X_1 со експоненцијална распределба со математичко очекување $1/\lambda$. Се поставува t=2,
2. Се генерира случајна променлива U рамномерно распределена на [0,1]. Ако $U \le \phi$, тогаш се поставува

$$X_t = \phi X_{t-1}$$

инаку се генерираат ε_t, со експоненцијална распределба со математичко очекување $1/\lambda$ и се поставува:

$$X_t = \phi X_{t-1} + \varepsilon_t,$$

3. Се поставува $t=t+1$, и се оди на чекор 2.

5. ВЕРИФИКАЦИЈА И ВАЛИДАЦИЈА

Една од најзначајните и најтешките задачи во развојот на моделот е верификацијата и валидацијата на симулациониот модел. Секој симулирачки код кој се развива е само апстракција на реалниот систем, па затоа секогаш треба да се остави место за сомневање на коресподенцијата или односот меѓу реалниот систем и моделот. Процесите на верификација и валидација ги опфаќаат следниве компоненти:

1. Верификација е концентрирана околу изградбата на правилен модел. Таа се користи за споредба на концептуалниот модел со компјутерската репрезентација што го интерпретира тој концепт. Дава одговор на прашањата: Дали моделот има правилна софтверска репрезентација? Дали се влезните параметри и логичката структура на моделот коректно репрезентирани?
   Значи накратко кажано, верификацијата се однесува на проверка дали софтверската програма за симулацијата, која го имплементира моделот, е без грешки и е конзистентна со моделот.

2. Валидацијата се користи да се определи дали моделот е прецизна репрезентација на реалниот систем. Најчесто е тоа итеративна процедура во која однесувањето на моделот, се споредува со однесувањето на реалниот систем и се воочуваат разликите кои понатаму се користат за доградување и исправка на моделот. Постапката за подобрување на моделот се продолжува се додека не се одлучи дека добиената точност на моделот е задоволена.

Поголемиот дел од методите што се користат во процесот на верификација и валидација се субјективни споредби, а помал број се формални статистички процедури.

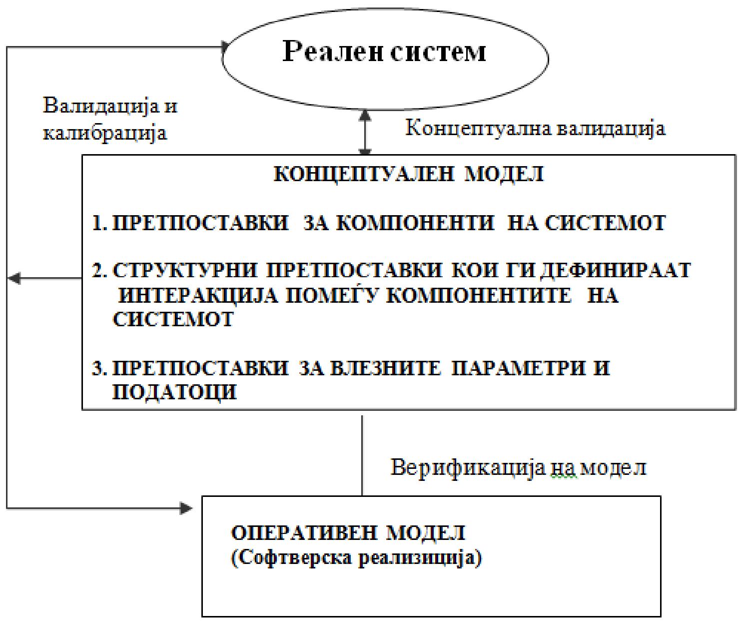
5.1 Изградба на моделот, верификација и валидација

Првиот чекор во изградбата на моделот се состои од набљудување на реалниот систем, интеракциите помеѓу неговите компоненти и собирањето на податоците. Но само со набљудување многу ретко се добиваат доволно информации за однесувањето на системот. Личности кои се блиски на системот (кои го проучуваат) или на потсистемите треба да се вклучат во развојот на симулацијата за да се земе во предвид нивното знаење. Операторите, техничарите, луѓето од одржување, инженерите, супервизорите и менаџерите, имаат познавања од одредени аспекти од системот кои не се блиски на другите личности вклучени во проектот. Во текот на изградбата на моделот, нови прашања може да произлезат , и развивачите на моделот ќе се вратат на овој чекор за проучување на структурата и однесувањето на реалниот систем.

Вториот чекор од изградбата на моделот е конструкцијата на концептуалниот модел: собирање претпоставки за компонентите и структурата на системот, структурни претпоставки кои дефинираат итерации меѓу компонентите на системот и претпоставки за влезните параметри и податоци.

Како што може да се види од Фигурата 5.1, валидацијата е споредба на реалниот систем со концептуалниот модел.

Третиот чекор е преведување на оперативниот модел во компјутерска препознатлива форма – софтверска реализација.



Фиг. 5.1 Изградба на моделот, верификација и валидација

Во пракса градењето на моделот не е линеарен процес кој се состои од овие три чекори. Наместо тоа моделарите се навраќаат на секој од овие чекори многу пати за време на изградбата, верификацијата и валидацијата на моделот.

5.2 Верификација на симулациониот модел

Целта на верификацијата на моделот е осигурување дека концептуалниот модел е правилно рефлектиран во компјутерска реализација. Концептуалниот модел многу често воведува неколку степени на апстракција за операциите на системот, или пак нивно поедноставување.

Верификацијата дава одговор на прашањето: Дали е концептуалниот модел (претпоставките на системските компоненти и структурата на системот, вредностите на параметрите, поедноставувањето и апстракциите) точно претставен со операциониот модел (компјутерска респрезентација)?

Следниве сугестии се користат во процесот на верификација:

1. Компјутерската репрезентација треба да биде контролирана од страна на девелоперот (развивачот)
2. Изработка на flow – дијаграм кој ги вклучува сите можни логички акции што се преземат кога се случува некој настан, и го следат логичкиот модел за секоја акција за секој тип на настан.
3. Испитување на излезот на моделот со различни множества влезни податоци.
4. Компјутерската репрезентација треба да ги печати влезните параметри на крајот на симулацијата, заради осигурување дека вредностите на овие параметри нема да се променат случајно во текот на програмата.
5. Да се направи документација во која се дава прецизна дефиниција на сите променливи кои што се користат и опис на целта на секој дел од кодот.
6. Претставување на текот на симулацијата, со графичко претставување, визуелизација и испис на текот на симулацијата.
7. *Interactiveruncontroller* или дебагерот е значајна компонента за успешна изградба на симулациониот модел. Дури и најдобрите аналитичари прават грешки кога моделот е во изградба. IRC асистира во наоѓањето и поправките на овие грешки.

8. Графичкиот интерфејс се препорачува во процесот на валидација и верификација. Графичката репрезентација треба да биде дадена со документација за полесно да може да се разбере системот.

Овие сугестии се дадени врз база на софтверското инженерство. Две статистики кои може да дадат брза индикација за оправданоста на моделот се: *currentcontents* и *totalcount. Currentcontents* укажува на бројот на елементи во секоја компонента во системот во даден момент. *Totalcount* укажува на вкупниот број на компоненти што влегуваат во секоја компонента на системот во дадено време. Во некои симулациони софтвери како *GPSS/H и Automod*, овие статистики се автоматски зачувани и може да бидат прикажани во секоја точка за време на симулацијата. Во други симулациони софтвери едноставни бројачи се додаваат за пресметување и прикажување на соодветните времиња. Ако *currentcontents* во некој дел од системот е висок, тоа укажува дека голем број од влезовите се задоцнети. Додека пак ако *totalcount* за некој потсистем е нула, тоа укажува дека нема влезни елементи во тој потсистем.

Внимателниот развој на овие статистики може да помогне во откривањето на грешки во логичкиот модел и некои погрешно специфицирани податоци. За некои модели можно е да се оцени дали поедината статистика е оправдана. Постои можност за пресметување на *longrun* мерите на перформансите. Пример: - Аналитичарот може да ја пресмета *longrun* искористеноста на серверот за поголем број од системите за масовно послужување, без специјални претпоставки што се однесуваат на распределбите на интервалот меѓу две последователни доаѓања и времето на опслужување. Потребни информации се само конфигурацијата на мрежата, интензитетот на доаѓање и интензитетот на опслужување.

Било која мера на перформансите која може да се пресмета аналитички, и да се спореди со резултатите добиени со симулацијата претставува скапоцена алатка во процесот на верификацијата.

Друга значајна помош во процесот на верификацијата е често запоставуваната фаза на документација. Ако во текот на изградбата на моделот се пишуваат кратки коментари во компјутерскиот програм, плус се даде дефиниција за сите променливи и параметри, и опис на секој главен дел од моделот, тогаш ќе биде полесно покасно да се изврши процесот на верификација.

Многу пософистицирана техника е користењето на *trace* (линија). Во општо, *trace* е детален компјутерски приказ, кој ја дава вредноста на секоја променлива во компјутерскиот програм, секогаш кога вредноста на променливата ќе се промени.

Со верификација се врши и проверка во однос на евентуално познатото решение на системот. Моделот се поставува така што влезните податоци би требало да дадат познато решение. Исто така се врши тестирање на сензитивноста и тестирање на пореметувањето.

Тестирање на сензитивност: со промена на еден параметар се одредува дали останатите остануваат непроменети, и дали однесувањето на моделот е сензитивно во однос на тој параметар.

Тестирање на пореметување: - Се поставуваат параметрите на моделот на неприродни вредности и се проверува однесувањето на моделот.

5.3 Валидација на моделот

Валидација е споредба на моделот и неговото однесување со реалниот систем и неговото однесување. Проблемот на валидација на моделот настанува бидејќи во фазата на изградба на моделот се внесуваат многу апроксимации на реалниот систем. На таков начин моделот се ограничува (се поставуваат граници во однос на околината), со игнорирање на се што не е експлицитен влез за системот, надвор од него и отфрлање на факторите кои се смета дека не се битни.

Видови на ограничувања:

1. Функционална апроксимација,
2. Апроксимација на распределби,
3. Апроксимација на независност,
4. Апроксимација на агрегација,
5. Апроксимација на стационарност.

Функционална апроксимација: исклучиво нелинеарни функции често се апроксимираат со некоја едноставна функција, на пример линеарна. Важна претпоставка е дека едноставната функција треба да биде приближно „оригинална” на функцијата во областа во која системот најверојатно ќе функционира. Ако програмата мора да функционира и во области каде поклопувањето со функциите е слабо, потребно е обезбедување на пораки за претпазливост.

Апроксимација на распределби: - Реалните распределби кои се и сами познати апроксимативно, често се апроксимираат со едноставни распределби, како што се експоненцијалната и нормалната. Најекстремни примери за апроксимација се кога случајната променлива се заменува со константа.

Апроксимација на независност: - Моделот често се поедноставува така што се претпоставува дека различните компоненти (опишани со случајни промненливи) се статистички независни.

Апроксимација на агрегација: - Под агрегација се подразбираат состојби каде повеќе елементи се разгледуваат како една целина. Типични примери на агрегација се:

Временска агрегација: Интервал од време, како што е ден се третира како поединечен период. За сите настани кои настанале во текот на денот се претпоставува дека се случиле истовремено.

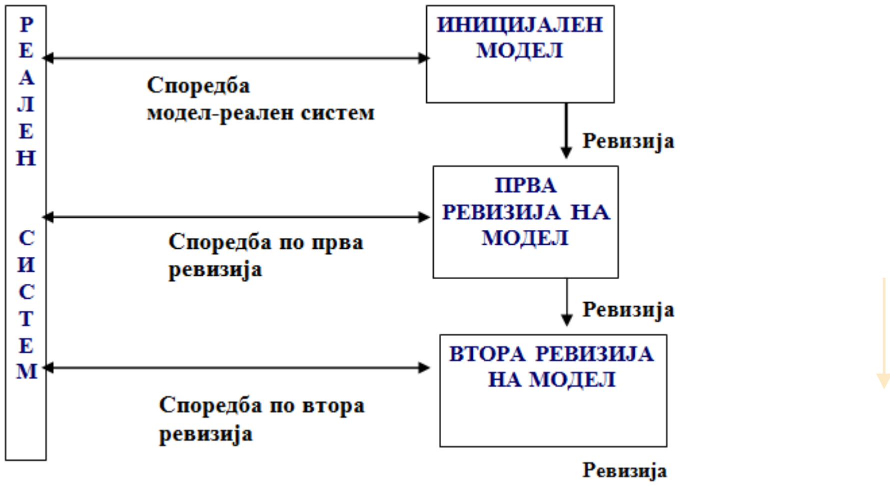
Помеѓу секторска агрегација: - Повеќе оддели, фирми, производни линии и итн. се гледаат како една целина.

Агрегација на помошни средства: - На пример, сметачки систем со два паралелни процесори, со агрегација сметаме дека имаме еден, два пати побрз централен процесор.

Апроксимација на стационарност: - Кај некои системи параметрите и другите карактеристики се менуваат со време. Но кај некои физички процеси стационарноста може да биде прифатлива апроксимација. Меѓутоа во политичките, економски, организациони и др. системи, секојдневното искуство покажува дека многу појави се нестационарни тако што оваа апроксимација е неприфатлива.

„Калибрација” е итеративна постапка за споредба на моделот со реалниот систем, се вршат прилагодувања на моделот, се споредува испитуваниот модел со реалниот систем, се вршат додатни модификации и ова се повторува се додека не се добие задоволителен модел.

Во фигурата 5.2 е прикажана релацијата на моделот на „калибрација” за процесот на валидација:



Фиг. 5. 2 Итеративен процес на калибрација на моделот

За споредба на моделот со реалниот систем се користат тестови, некои од нив се субјективни останатите се објективни. Субјективните тестови вообичаено воведуваат личности кои имаат познавања од еден или повеќе аспекти на системот даваат решенија што се однесуваат на системот и неговиот излез. Објективните тестови секогаш бараат податоци за однесување на системот и овие податоци да бидат кореспондентни со податоците добиени од моделот. Споредбите се вршат со помош на статистички тестови. Еден можен недостаток на фазата на „калибрација" е тоа што завршната фаза на валидација се врши само со едно множество податоци. Поради тоа е потребно да се избере ново множество од системски податоци кои ќе се користат во завршната фаза на валидацијата. Ако и за новото множество податоци моделот останува стабилен и близок до реалниот систем постапката завршува, ако пак се јават поголеми разлики системот се враќа во фазата на „калибрација" и моделот се модифицира се додека не стане прифатлив. Валидација не е претпоставка и моделот не е тотална репрезентација на реалниот систем што се проучува. Исто така секоја ревизија на моделот вклучува трошоци, време и обиди. Моделарот не може да гарантира зголемено ниво на точност на модел,наспроти тоа што трошоците се зголемуваат при секој обид на валидација. Исто така тој има одредено максималната разлика меѓу моделот и реалниот систем. Ако ова ниво на точност не може да се добие, тогаш или точноста на моделот мора да биде намалена или моделот да биде отфрлен.

Во процесот на валидација препорачливо е да се следат следниве чекори:

1. Градење на модел кој точно ќе ги претставува особините на реалниот систем.
2. Потврда на претпоставките на моделот.
3. Споредба на влезно-излените трансформации на моделот со влезна излезните трансформации на реалниот систем.

5.3.1 Вистинско претставување

Првата цел на моделарот е да конструира модел што ќе биде задоволителен за корисниците и останатите кои имаат одредени познавања околу реалниот систем што се симулира. Потенцијалните корисници треба да бидат вклучени во целокупната конструкција на моделот. Корисниците исто така можат да бидат вклучени во процесот „калибрација" кога моделот итеративно се подобрува. За проверка на моделот може да се користи анализа на чувствителност (*sensitivityanalysis*). Тоа е постапка со која се тестира однесувањето на системот кога еден или повеќе влезни параметри ќе се подобрат. На пример, во систем за масовно опслужување ако интензитетот на доаѓање на корисниците се зголеми што треба да се очекува за искористеноста на серверите, должината на редицата итн. Базирајќи се на искуството и испитувањето на реалниот систем, корисниците на моделот и изведувачот на моделот веројатно имаат некои забелешки како се менува излезот на моделот кога влезните параметри се зголемени или намалени. Изведувачот на моделот треба да внимава околу изборот на променливи така што ќе бидат избрани најкритичните променливи бидејќи тестирањето е доста скап процес и побарува многу време во случај да се земаат сите влезни променливи. Ако податоците на реалниот систем се достапни за најмалку две нагодувања на влезните параметри, може да бидат конструирани објективни научни тестови за чувствителност со користење на соодветни статистички техники.

5.3.2 Претпоставки за моделот

Претпоставките се поделени во две главни категории: претпоставки за структурата и претпоставки за податоците.

Претпоставките за структурата се однесуваат на прашања околу функционирањето на системот и најчесто воведуваат поедноставувања и апстракција на реалниот систем. Пример: Редица на корисници при опслужување во банка. Корисниците може да чекаат во една редица или да има повеќе редици за сите шалтери. Ако има повеќе редици опслужувањето може да биде *FIFO (firstin—firstout*), или клиентите може да ги менуваат редиците доколку некоја редица е побрза. Бројот на шалтери може да биде фиксен или да се менува. Претпоставките за структурата треба да го потврдат актуелното набљудување за време на дадениот временски период, заедно со дискусиите околу менаџирањето и политиките кои ги има банката и можноста за соодветно подобрување.

Претпоставките за податоците се базираат на доверливо множество податоци и коректна статистичка анализа на овие податоци. Во примерот со банката споменат погоре се собираат следниве податоци:

1. Време меѓу пристигнување на корисници во неколку двочасовни „*rushhour*" периоди , т.е. периоди кога доаѓаат најмногу корисници.
2. Време меѓу пристигнување на корисниците во померен период(период кога пристигнуваат помалку корисници).
3. Време на опслужување за комерцијални трансакции.
4. Време на опслужување за лични трансакции.

Сигурноста на податоците треба да се потврди со консултација на менаџерот на банката кој ќе ги идентификува периодите во кои најмногу доаѓаат корисници и помирните периоди. Кога се комбинираат две или повеќе множества собрани податоци во различни временски периоди, сигурноста на податоците може да се зголеми, со објективни статистички тестови за хомогеност на податоците. Дали две множества податоци $\{X_i\}$и $\{Y_i\}$(времиња на опслужување собрани во различни временски период) се примероци од иста популација. Ако се дали може да се комбинираат. Исто така може да се вршат и тестови за корелација. Ако се утврди дека нема корелација меѓу податоците, аналитичарот во тој случај име случаен примерок и може да започне со статистичка анализа. Статистичката анализа како што е опишано во глава 5 се состои од:

1. Идентификација на соодветните веројатностни распределби;
2. Оценка на параметрите на хипотетичката распределба;
3. Потврдување на дадениот модел статистички со *goodness–of- fit* тестови како *Hi*- квадрат тестот, Колмогоров-Смирнов тестот и графички методи.

5.3.3 Валидација на влезно - излезни трансформации

Последниот тест за моделот и единствениот објективен тест на моделот разгледуван како целина кој се однесува на можноста на моделот да го предвиди идното однесување на моделот кога влезните податоци на моделот одговараат на влезните податоци на реалниот систем. Во иднина ако нивото на некоја влезна променлива се зголеми или намали, моделот треба точно да знае што ќе се случи со реалниот систем под истите околности. Со други зборови структурата на моделот треба да биде доволно прецизна, за да моделот дава добри предвидувања, не само за едно множество од влезни податоци туку за целиот ранг на множествата од влезни податоци кои се од интерес. Во оваа фаза на валидација моделот се разгледува како влезно-излезна трансформација. Моделот ги прима вредностите на влезните параметри и ги трансформира овие влезови во излезни мери на перформансите.

Наместо валидација на влезно-излезни трансформации со предвидување на иднината, моделарот може да користи минати историски податоци кои се резервирани единствено за целите на валидацијата така што, ако едно множество податоци се користи за развој и „калибрација” на моделот, тогаш се препорачува друго множество податоци да се користи за финалниот тест на валидација. Во овој случај „прецизно предвидување на минатото” може да го замени предвидувањето на иднината за целите на валидација моделот.

Моделот вообичаено се развива така што од примерен интерес е да се добијат вредностите на множество од излезни променливи под даден опсег на влезни услови. Пример во систем за масовно опслужување излезни променливи се искористеноста на серверите, и времето на задоцнување на опслужување на клиентите, додека опсегот на влезни услови може да вклучува два или три сервери и избор на правила за издавање (дисциплина на редицата). Моделарот треба да ги користи најбитните излезни променливи кои се разгледуваат како критериуми за валидација на моделот. Ако моделот покасно се користи за цели различни од првобитната тогаш треба да се ревалидира во термини на нови излезни променливи под ново множество од влезни услови.

Потребен услов за валидација на влезно-излезните трансформации е реалниот систем што се проучува да постои, за да можат податоци од најмалку едно множество од влезни услови да бидат собрани за споредба со предвидувањето на моделот.

Ако системот е во изградба не можат системски оперативни податоци да се соберат, тогаш комплетна влезно-излезна трансформација не е можна. Но доколку некој од потсистемите постои, тогаш е можно да се пристапи кон негова влезно-излезна трансформација.

Пример 5.1 (Банкарски шалтер за услуга на клиенти во кола). Еден канал за опслужување и еден прозорец за коли. Една или две трансакции се можни на прозорецот за коли. Се претпоставува дека секое време за опслужување е случајно. Времињата на опслужување $\{S_i| \ i=1,90\}$,и интервалите меѓу пристигнување на клиентите $\{A_i| \ i=1,90\}$,биле собрани за 90 корисници што доаѓале меѓу 10 и 13 часот во петок. Анализата на податоците покажува дека станува збор за Пуасонов поток (времињата меѓу две последователни пристигнувања на клиенти се експоненцијално распределени) со интензите *0* =45/час и времето на опслужување е распределено со апкорсимативна нормална распределба со математичко очекување еднакво на 1.1 мин. и стандардна девијација од 0.2 мин. Значи има две влезни променливи:

1. Времињата помеѓу пристигнувањата на клиентите се експоненцијално распределени со λ=45/час.
2. Времето за опслужување е приближно $N(1.1,(0.2)^2)$

Интервалите меѓу пристигнувањата на клиентите не се контролабилни (не може да се контролираат). Времињата на опслужување исто така се неконтролабилни променливи, меѓутоа ниво од времиња на опслужување може да биде парцијално контролабилно.

На пример, средното време на опслужување може да се намали на 0.9мин. Со користење на понапредни технологии ниво од променливата време на опслужување може да стане решлива променлива и контролабилен параметар. На пример моментално има еден шалтер (D_1=1), средното време на опслужување (D_2=1.1) и една редица на чекање (D_3=1). D_1, D_2, D_3се променливи кои може да се контролираат, додека времињата меѓу пристигнувањето на клиентите и времето на пристигнување на клиенти се променливи што не може да се контролираат. Моделот на банка е развиен и верифициран во затворени услови со менаџмент и вработени. Моделот се разгледува како „црна кутија" која дадените влезни променливи ги трансформира во множество од излезни променливи. Излезните променливи $\{Y_1,Y_2,\ldots,Y_n\}$ се разни статистики од интерес кои го оценуваат системот. Тоа се Y_1 зафатеност на шалтерот, Y_2 просечно време на чекање на опслужување, Y_3 максимална должина на редицата овие три излезни променливи се променливи од примерен интерес, но исто така се разгледуваат и следниве излезни променливи, Y_4 набљудувања на влезниот интензитет, Y_5 средно време на опслужување, Y_6 стандардна девијација на опслужувањето и Y_7 средна должина од време.

Овие променливи се детално дадени во следната табела:

Влезни променливи	Излезни променливи, Y
D=решливи променливи	Променливи од примарен интерес (Y_1,Y_2,Y_3)
X=други променливи	Y_1 искористеност на шалтерот
Пуасонов влезен поток со интензитет 45/час	Y_2 средно време чекање во редицата
Времиња на опслужување,$N(D_2,0.2^2)$	Y_3 максимална должина на редицата
X_{21},X_{22}	Други променливи од секундарен интерес:
	Y_4 набљудуван интензитет на пристигнувања
D_1=1(еден шалтер)	Y_5 средно време на опслужување
D_2=1.1 min.(средно време на опслужување)	Y_6 стандардна девијација на опслужување
D_3=1(една редица)	Y_7 средна должина на време

Табела 5.1 Влезни и излезни променливи во моделот на банка

Влезните променливи што не може да се контролираат се означени со X, оние што може да се контролираат со D, и излезните со Y. Значи во моделот „црна кутија" се земаат X и D како влез и се добива излезот Y.

$$(X,D)\xrightarrow{f}Y\ ,$$

$$f(X,D)=Y.$$

Со X_{1n} се означува генерираниот интервал меѓу пристигнувањето на n-1-от и n-от клиент, а со X_{2n} генерираното време на опслужување.

За валидација на влезно-излезните трансформации на банкарскиот систем треба да се соберат податоци од реалниот систем под истите услови. Резултатите од излезните променливи треба да бидат собрани за време на истиот временски период(од 10:00 до 13:00 часот во петок), при што влезни податоци се:$\{A_i,S_i\}$. Тоа е битно бидејќи ако податоците за излезните променливи се собрани во некој "побавен" период, кога интензитетот на доаѓање е 40/час, тогаш излезните променливи како искористеност на шалтерот Z_1, средно време на чекање во редицата Z_2 и максимална должина на редицата Z_3, очекувано ќе бидат помали отколку истите променливи кога интензитетот на доаѓање е 45/час, како што се претпоставува во моделот. За средното време на чекање е добиено Z_2=4.3 min, во петок меѓу 10 и 13 часот.

Кога моделот е извршен со генерираните влезни случајни променливи X_{1n} и X_{2n} се очекува вредноста што ќе се добие за променливата Y_2 средно време на чекање на клиентите во редицата да е блиско до добиената вредност Z_2=4.3min. Меѓутоа не може да се очекува генерираните влезни променливи $\{X_{1n},X_{2n}\}$ да бидат точно еднакви со случајните променливи $\{A_i,S_i\}$, но тие треба да претставуваат статистички примерок на актуелните влезни променливи. Се очекува симулираната генерирана вредност на Y_2 да биде конзистентна со набљудуванат вредност на Z_2. Се прават мал број од статистички независни повторувања на моделот. Статистичката независност се гарантира со непреклопени множества од случајни броеви добиени со генератор на случајни броеви.

Резултатите добиени од шест независни повторувања секое во траење од 2 часа, се дадени со табелата 5.2

Повторување	Y_4(пристигнувања/час)	Y_5(min)	Y_2(min)
1	51	1.07	2.79
2	40	1.12	1.12
3	45.5	1.06	2.24
4	50.5	1.10	3.45
5	53	1.09	3.13
6	49	1.07	2.38
Математичко очекување на примерокот			2.51
Стандардна девијација			0.82

Табела 5.2 Резултати од шест повторувања во моделот на банка

Добиениот интензитет на доаѓање Y_4 и средното време на опслужување при секое повторување на моделот се споредуваат со соодветните теориски вредности 45/час и 1.1мин. Тестот на валидација се состои од споредба на излезните променливи Y_2 и Z_2. Се разгледува нултата хипотеза

$$H_0: E(Y_2)=4.3\text{min}$$

Наспроти алтернативната хипотеза

$$H_1: E(Y_2)\neq 4.3\text{min} \quad (6.1)$$

Ако H_0 не се отфрла нема причина да се оцени дека моделот не е валиден. Ако H_0 се отфрли тогаш моменталната верзија на моделот се отфрла, и тековниот модел треба да се подобри.

Тест што овде се применува е *t*- тестот, кој е конструиран на следниот начин:

Се избира ниво на значајност α примерок со големина n. Во моделот на банка се зема:

$$\alpha = 0.05 \text{ и } n=6$$

Се пресметуваат средната вредност на примерокот $\overline{Y_2}$ и стандардната девијација на примерокот S со следниве формули:

$$\overline{Y}_2 = \frac{1}{n}\sum_{i=1}^{n} Y_{2i} = 2.51\,\text{min}, \quad (5.2)$$

$$S=\left[\frac{\sum_{i=1}^{n}\left(Y_{2i}-\overline{Y}_2\right)^2}{n-1}\right]^{1/2}=0.82\ \text{min}. \tag{5.3}$$

Се чита критичната вредност од табела $t_{\alpha/2,n-1}=t_{0.25,5}=2.571$, потоа се пресметува тест статистиката:

$$t_0=\frac{\overline{Y}_2-\mu_0}{S/\sqrt{n}}, \tag{5.4}$$

каде $\mu_0=4.3$, и во случајот:

$$t_0=\frac{2.51-4.3}{S0.82/\sqrt{6}}=-5.34 \tag{5.5}$$

Во општ случај ако $|t_0|>t_{\alpha/2,n-1}$ се отфрла H_0. И во овој случај се отфрла H_0 бидејќи $|t|=5.34>t_{0.25,5}=2.571$

Значи се отфрла H_0 и моделот е неадекватен за предвидување на средното време во редицата.
Веројатноста за отфрлање на хипотеза е силна,, бидејќи

$$P(H_0 \text{ да е одбиена} \mid H_0 \text{ да е вистинита})=\alpha=0.05$$

Според ова веројатноста да се направи грешка при отфрлањето на нулата хипотеза е многу мала.

Оправдување за користење на *t*-тестот се добива бидејќи променливите Y_{2i}се нормално распределени и независни променливи.

1. Во *i*-тата опсервација, Y_{2i} е средно време на чекање за сите клиенти кои доаѓаат во временскиот интервал од 10:00 до 13:00 часот во петок. Според централната гранична теорема може да се смета дека Y_{2i} е приближно нормално распределено ако бројот на клиенти што пристигнуваат во овој интервал не е мал.
2. Y_{2i}, $i=1,..6$ се статистички независни, бидејќи при секое повторување се генерираат нови случајни броеви.
3. *T*- статистиката може да се примени и во случај кога Y_{2i}, i=1,...,6 не се нормално распределени, т.е. *T* статистиката е „робусна".

Бидејќи не се добива потврда за валидноста на моделот, прашање е што да се направи понатаму. За понатамошно испитување се разгледуваат следниве две претпоставки:

1. Кога клиентот со кола доаѓа на празен прозорец, службеникот непосредно го започнува неговото опслужување.
2. Нема чекање помеѓу две последователни опслужувања, кога еден клиент завршува со опслужување веднаш започнува да се опслужува друг клиент што чека во редицата.

Втората претпоставка е коректна.
Првата претпоставка не се прифаќа, бидејќи службеникот може да опслужува клиент што дошол пешки ако нема клиент со кола, и секогаш го завршува започнатото опслужување пред да започне ново опслужување.

За корекција на овој неадекватен модел, структурата на моделот се променува со вклучување дополнителни барања за времето на опслужување, т.е. се собираат времиња на опслужување на клиенти што доаѓаат пешки. Со анализа на овие времиња на опслужување, се добива дека тие се распределени експоненцијално со математичко очекување еднакво на 3 мин.

За ревидираниот модел резултати се дадени во табела 5.3

Повторување	Y_4 (пристигнувања/час)	Y_5 (min.)	Y_2
1	51	1.07	5.37
2	40	1.11	1.98
3	45.5	1.06	5.29
4	50.5	1.09	3.82
5	53	1.08	6.74
6	49	1.08	5.49
Математичко очекување на примерокот			4.78
стандардна девијација			1.66

Табела 5.3 Резултати од шест повторувања на ревидираниот модел на банка

Повторно се тестира нултата хипотеза H_0: $E(Y_2)=4.3$ min. Се зема $\alpha=0.05$, $n=6$ и се пресметува $\overline{Y_2}=4.78$ min. , $S=1.66$ min и $t_{0.025,5}=2.571$.

За T статистиката се добива:

$$t_0=\frac{\left(\overline{Y_2}-\mu_0\right)}{\left(S/\sqrt{n}\right)}=0.710\,.$$

Во овој случај, $|t_0|<t_{0.025,5}=2.571$, и нултата хипотеза H_0 не се отфрла, што значи дека овој модел се прифаќа како валиден. Не отфрлањето на нултата хипотеза H_0 се смета за слаб заклучок доколку:

P(не е отфрлена $H_0|H_1$е вистинита)=β е големо

β пак зависи од големината на примерокот n и од δ:

$$\delta=\frac{|E(Y_2-\mu_0)|}{\sigma},$$

во овој случај $\mu_0=4.3$ мин. и σ е стандардната девијација на популацијата која е оценета со S. Ако се претпостави дека $\delta=|E(Y_2)\ \mu_0|<1$, тогаш δ се оценува со $\bar{\delta}$ и за оценувачот се добива:

$$\bar{\delta}=\frac{|E(Y_2-\mu_0)|}{S}=\frac{1}{1.66}=0.66\,.$$

За t- тестот со α = 0.05 и n = 6, од табела се чита:

$$\beta\left(\bar{\delta}\right)=\beta(0.66)=0.75\,.$$

Ако се бара $\beta\left(\bar{\delta}\right)\le 0.10$, тогаш бидејќи на β при дадено α, влијае бројот на повторувања, се зема n=30.

Во разгледуваниот случај n=6 и стандардната девијацијата е еднаква на 1.66, веројатноста де се прифати нултата хипотеза H_0, кога моделот не е валиден е доста висока 0.75. За да се добие $\beta<0.1$ треба да се извршат 30 повторувања. Значи секогаш е подобро да се контролира и β или грешката од II тип (како попрецизна грешка) со избор на бројот на повторувања и специфицирање на критично δ.

5.3.4 Влезно излезна валидација со користење на историски влезни податоци

Во претходниот пример за банка, вештачките влезни променливи $\{X_{1n}, X_{2n}, n=1,2,...\}$ за интервалите меѓу пристигнувањата на клиентите и времето на опслужување беа генерирани и вредностите добиени за излезната променлива Y_2 при сите повторувања беа споредени со вредностите на истата излезна променлива добиени со набљудување на реалниот систем. Една алтернатива за генерирање на влезни податоци е користење на историски податоци $\{A_n, S_n, n=1,2,..\}$, и потоа да се спореди излезот на моделот со резултатите добиени за излезот при разгледувањето на реалниот систем. За имплементација на оваа техника во моделот на банка податоците $A_1, A_2, . . .$ и $S_1, S_2, . . .$ треба да бидат внесени во моделот во низа, или складирани во file за да може да се прочитаат кога ќе биде потребно. После доаѓањето на n-тиот клиент во моментот $t_n = \sum_{i=1}^{n} A_i$, n+1-от клиент треба да биде распореден на листата од идни настани во некое идно време $t_n + A_{n+1}$(без било какво генерирање на случајни броеви). Ако n-тиот клиент дојде на опслужување во моментот $t_n^{'}$, тогаш неговото опслужување треба да заврши во моментот $t_n^{'} + S_n$. Ова распоредување без генерирање на случајни броеви, може да се имплементира доста лесно во многу симулациони јазици со користење на низи за складирање на податоците.

При користењето на оваа техника, дизајнерот на моделот се надева дека симулација ќе ги копира најзначајните настани што се случуваат во реалниот систем колку што е можно подобро. Во моделот на банка, времињата меѓу пристигнувањето на клиентите и времињата на опслужување ќе бидат точна копија на соодветните времиња што се добиваат во реалниот систем во временски интервал од 10:00 до 13:00 часот во петок. Ако моделот е доволно прецизен, тогаш времето на чекање во редицата, должината на редицата и искористеноста на серверот ќе бидат блиски до соодветните времиња што се добиваат во реалниот систем. Секако, моделарот и корисниците на моделот имаа прецизирано одредено ниво на точност.

Во раководењето на валидациониот тест со користењето на историски влезни податоци, значајно е сите влезни податоци $\{A_n, S_n, ...\}$ и сите излезни променливи како време на чекање во редицата Z_2 да бидат собрани во ист временски период. Имплементацијата на оваа техника ќе биде тешка за големи системи, бидејќи е потребно симултано собирање на податоци за сите влезни променливи и за сите излезни променливи кои се од примерен интерес. Не е можно утврдување на валидноста на моделот со само едно множество од влезни и излезни податоци. Но ако K историски влезни податоци се собрани, и K набљудувања $Z_{i1}, Z_{i2}, ..., Z_{iK}$ од некоја излезна променлива Z_is е собрани, така што Z_{ij} кореспондира на j - тото множество влезови, објективни тест статистики може да се применат. Со K множества од влезни податоци, моделарот може да го изврши моделот K пати и да ги спореди резултатите $W_{i1}, W_{i2}, ..., W_{iK}$ добиени од симулации со соодветните излезни историски податоци $Z_{i1}, Z_{i2}, ..., Z_{iK}$. Начинот на кој се врши оваа споредба е даден во табела 5.4

Множество влезни податоци	Излези од системот	Излези од моделот	Добиени разлики	Квадратна девијација за средната вредност $\left(d_j - \bar{d}\right)^2$
1	Z_{i1}	W_{i1}	$d_1 = Z_{i1} - W_{i1}$	$\left(d_1 - \bar{d}\right)^2$
2	Z_{i2}	W_{i2}	$d_2 = Z_{i2} - W_{i2}$	$\left(d_2 - \bar{d}\right)^2$
3	Z_{i3}	W_{i3}	$d_3 = Z_{i3} - W_{i3}$	$\left(d_3 - \bar{d}\right)^2$
.				
.				
.				
K	Z_{iK}	W_{iK}	$d_K = Z_{iK}\ W_{iK}$	$\left(d_K - \bar{d}\right)^2$
			$\bar{d} = \frac{1}{K}\sum_{j=1}^{K} d_j$	$S_d^2 = \frac{1}{K-1}\sum_{j=1}^{K}\left(d_j - \bar{d}\right)^2$

Табела 5.4 Споредување на излезните мерки на моделот и системот, кога се користат историски влезни податоци

Ако K–те множества од влезни податоци се прилично хомогени, тогаш е разумно да се претпостави дека K-те добиени разлики $d_j = Z_{ij} - W_{ij}$, за $j=1,2,..,K$ се идентично распределени. Ако собирањата на K-те множества од влезни податоци било во одделни временски периоди, тогаш е разумно да се претпостави дека K-те разлики d_1, d_2..., d_k се статистички независни, од каде може да се заклучи дека разликите d_1, d_2, ..., d_k даваат случен примерок. Во многу случаи Z_i и W_i е среден примерок над корисниците, тогаш се користење на централната гранична теорема, разликите $d_j = Z_{ij} - W_{ij}$ се приближно нормално распределени со математичко очекување μ_d и дисперзија σ_d^2 .Се тестира хипотезата дека нема разлика меѓу Z_{ij} и W_{ij} , т.е. μ_d :

$$H_0: \mu_d = 0,$$

наспроти алтернативната хипотеза дека има значајни разлики:

$$H_1: \mu_d \neq 0.$$

за тестирање е искористена и - статистиката:

$$t_0 = \frac{\bar{d} - \mu_d}{S_d / \sqrt{K}} \qquad (5.6)$$

каде $\bar{d}$ е средна вредност на примерокот, S_d^2 е дисперзија на примерокот, и дадена критичната вредност $t_{\alpha/2,K-1}$, каде α е нивото на значајност и K-1 е бројот на степени на слобода. Ако $|t_0| > t_{\alpha/2,K-1}$, се отфрла хипотезата H_0 дека нема разлики меѓу набљудуваните вредности, и се заклучува дека моделот не е адекватен. Ако $|t_0| < t_{\alpha/2,K-1}$, не се отфрла хипотезата H_0 дека нема разлики меѓу набљудуваните вредности, и се заклучува дека моделот е адекватен.

5.3.5 Влезно излезна валидација: Користење на Туринг тест

Како додаток на статистичките тестови или кога статистичкиот тест не е применлив, знаењата за однесувањето на системот на експертите може да се користат за споредба на излезните податоци на моделот со излезните податоци на системот.

Пример 5.2: Пет извештаи на системските перформанси во пет различни денови се изготвени, и излезните податоци од симулациите се користат за добивање на пет „фалсификувани" извештаи. Сите десет извештаи треба да имаат ист формат и да користат исти информации. Потоа сите десет извештаи се случајно распоредуваат, и се даваат на инженер да одлучи кои од нив се реални а кои „фалсификати". Ако инженерот точно ги идентификува „фалсификуваните" извештаи, тогаш моделот не е адекватен и треба да се подобри. Доколку пак инженерот не може да воочи разлика меѓу вистинските и „фалсификуваните" извештаи, тогаш моделарот ќе заклучи дека овој тест не дава информации дека моделот е неадекватен. Овој тест е наречен Туринг тест. Неговото користење во развојот на моделот може да биде скапоцена алатка за откривање дали моделот е неадекватен и евентуално за зголемување на кредибилноста на моделот, како и за подобрување и прочистување на моделот.

5.4 Формални критериуми за утврдување на валидноста на моделот

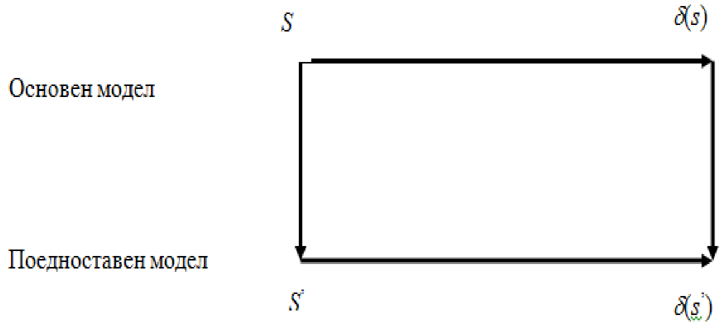
Хомоморфизам е формална постапка која се состои во наоѓање на пресликување *H* со помош на кое е можно од секоја состојба на основниот модел (на пример *s*) да се премине во соодветната состојба на упростениот модел (на пример *s*$^{'}$) Ако ваквото пресликување постои, се заклучува дека влезно излезните однесувања на основниот и поедноставениот модел се исти за дадени експериментални услови.

За да се утврди хомоморфизам треба да се исполнат условите:

1. Инваријантност на функцијата на време;
2. Инваријантност на функцијата на премин на состојби;
3. Инваријантност на излезните функции.

Инваријантност на функцијата на време: - Состојбата во основниот модел и соодветната состојба во поедноставениот модел мора да имаат иста функција на време. Тоа значи дека состојбите на основниот и упростениот модел се менуваат во исти моменти на време.

Инваријантност на функцијата на премин на состојби: - Ако на состојбата s на основниот модел одговара состојбата s$^{'}$ на упростениот модел (т.е. со примена на пресликување H од состојбата s се добива состојбата s$^{'}$) тогаш и следните состојби во кои моделите ќе преминат во наредниот момент ќе одговараат една на друга. Следна состојба на основниот модел ќе биде δ(s), а на упростениот δ(s$^{'}$), што значи дека со примена на пресликувањето H од состојбата δ(s) треба да се добие состојбата δ(s$^{'}$).



Фиг. 5.3 Инваријантност на функцијата на премин на состојби

Инваријантност на излезна функција: Нека е *s* состојба на системот и нека е можно со примена на пресликување *H* да премине во состојба $s^{'}$ на поедноставениот модел. Примената на излезната функција на основниот модел (за овие експериментални услови за кое е извршено упростувањето (λ) за да се добие поедноставен модел за состојбата *s*) дава ист излез како и примената на излезната функција на поедноставениот модел ($\lambda^{'}$) на состојбата $s^{'}$. Тука јасно се уочува дека валидноста на поедноставениот модел се утврдува и во однос на зададените поедноставни услови. Ако пресликувањето H, е инјекција и субјекција т.е. е од тип 1:1, тогаш станува збор за **изоморфизам**.

6. СИСТЕМИ ЗА МАСОВНО ОПСЛУЖУВАЊЕ

6.1 Вовед во системи за масовно опслужување

Симулациите често се користат за анализа на системите на масовно опслужување. Задачите и методите на теоријата на масовно опслужување заземаат значајно место во проучувањето на реалните системи. Тие имаат голема примена во науката и техниката. Теоријата на масовно оспсужување го користи апаратот на теоријата на веројатност и ги проучува веројатносните ситуации поврзани со елементите на случајност во системите за масовно опслужување.

Моделите на системите на масовно опслужување, било да се решаваат математички или да се анализираат со помош на симулации, го снабдуваат аналитичарот со моќна алатка за дизајнирање и развој на перформансите на системите за масовно опслужување. Типичните мерки (величини) на перформансите на системот ја вклучуваат искористеноста на серверот, должината на редицата на чекање, и задржувањето на клиентите во системот. Теоријата на масовно опслужување како и анализата со помош на симулации се користат за да ги предвидат мерките (величините) на системските перформанси како функција од влезните параметри. Влезните параметри се: интензитетот на доаѓање на клиенти, интензитетот на опслужување на серверите, бројот на сервери и др.

За релативно едноставните системи, овие мерките на перформансите на системот може да бидат пресметани математички, но за поголема сигурност добиените резултати може да се споредат со резултатите добиени со користење на симулации. Но за покомлексни системи вообичаено се користата симулациите. Еден систем за масовно опслужуање се смета за изучен, ако се определени величините кои го карактеризираат и кои во општ случај се случајни променјливи.

На Фигура 6.1 се дадени основните елементи на еден систем за масовно опслужување:

Фигура 6.1 Карактеристики на системот

Клучни елементи на системите за масовно опслужување се клиентите и серверите (опслужувачи). Поимот "клиент" може да се однесува на луѓе, автомобили, машини, пациенти, авиони, односно на се што побарува опслужување (услуга). Поимот "сервер" се однесува на ресурси кои го извршуваат опслужувањето, тоа може да бидат рецепционер, механичар, мајстор, медицински персонал, процесор на компјутер итн.

6.1.1 Карактеристики на системи за масовно опслужување

Популација од клиенти може да биде конечна или бесконечна. Бесконечна популација е теориски модел на систем со голем број на корисници, на пример клиенти кои побаруваат услуга од банка, продавница, бензинска пумпа, ресторант и слично. Како примери на конечна популација може да се сметаат бројот на процеси што треба да ги изврши компјутерот, или бројот на машини што треба да бидат поправени од сервисерот. Главната разлика меѓу конечните и бесконечните популации од клиенти е начинот на кој интензитетот на доаѓање (среден број на пристигнувања во единица време) е дефиниран. Кај бесконечните популации од клиенти интензитетот на доаѓање не зависи од бројот на клиенти кои се наоѓаат во процес на опслужување (се опслужуваат или чекаат за опслужување). Од друга страна, кај конечните популации од клиенти интензитетот на доѓање зависи од бројот на клиенти кои се опслужуваат или чекаат за опслужување. Ако на пример, популацијата се состои од само еден клиент, тогаш ако клиентот се опслужува, интензитетот на доаѓање ќе биде нула.

Влезниот поток го карактеризира начинот на кој клиентите влегуваат во системот. Пристигнувањата можат да бидат во одредено време или во случајно време. Кога пристигнувањата се случајни, тогаш времето помеѓу две последователни пристигнувања се карактеризира со веројатносни распределби. Нека t_1, $t_{2,\ldots}$ - се моменти на пристигнување на клиенти во системот. Тогаш $T_i = t_i - t_{i\text{-}1}$, i = 1,2,... е должината на временскиот интервал помеѓу пристигнувањето на i-1-тата и i-тата група од клиенти. Нека со Z_t е означен бројот на клиенти пристигнати во системот во моментот t.

Влезниот поток се наракува **поток со ограничено последејство**, ако случајните променливи T_1, T_2,……. и Z_1,Z_2,…. се независни во целина.

Влезниот поток е **ординарен**, ако не е возможно (со позитивна веројатност) да се појават два или повеќе клиенти во еден ист момент на време.

Влезниот поток се нарекува **поток без последејствие**, ако веројатноста да се појават k клиенти во интерлавот [T, $T + t$] не зависат од тоа колку клиенти се појавиле пред тој интервал.

Влезниот поток е **стационарен** ако веројатноста да се појават k клиенти во интервалот [T, $T + t$] не зависи од T, туку од t и k.

$$P\{k \text{ клиенти во } [T,T + t]\}= P\{k \text{ клиенти во } [0,t]\}$$

Ако T_i се независни и еднакво распределени случајни променливи и потокот е ординарен тогаш потокот се нарекува **рекурентен поток**.

Потокот од клиенти кој е:

- стационарен
- без последејствие
- ординарен

се нарекува прост **(Пуасонов) поток.**

Редицата претставува точен број на клиенти кои треба да бидат опслужени (се разбира редицата може да биде и празна). Обично клиентот што се опслужува се смета дека не е во редицата. Максималната должина на редицата е вкупниот број на клиенти кои чекаат во редицата (плус оние кои се опслужуваат). Во практичните симулации редицата е ограничена, но некои теориски модели претпоставуваат дека редицата е со неограничена должина. Ако редицата е со ограничена должина некои клиенти ќе мора да бидат вратени без да бидат опслужени.

.Дисциплината го претставува начинот на кој се организира редицата (правила според кои клиентите влегуваат или излегуваат од редицата). Постојат следниве начини:

- FIFO (first in first out),
- LIFO(last in first out),
- SIRO (опслужување со случаен редослед),

- Редици со приоритет (прво се опслужуваат клиентите со приоритет),
- Многу други посложени методи на редење кои ја менуваат положбата на клиентите во зависност од времето кое веќе го поминале во редицата, очејувано време на опслужување и приоритет.

Опслужувањето (или сервисот) преставува активноста за која клиентите чекаат. Обично времетраењето на опслужувањето е случајно. Теоретските модели се базираат на случајна распределба на времето на опслужување. Исто така важен параметер е и бројот на сервери. Системот со еден сервер е едноканален систем, ситемот со повеќе сервери е повеќеканален систем.

Излезот (output) претставува начинот на клиентите го напуштаат системот. Излезот е честопати игнориран од теоретските модели, но понекогаш клиентите кои го напуштаат сервисот може повторно да влезат во редицата.

6.1.2 Кендалова класификација на системи за масовно опслужување

Постојат неколку модификации на Кенделовата класификација на системите за масовно опслужување. Најкористената класификаци користи шест симболи:

$$A\,/\,B\,/\,s\,/\,q\,/\,c\,/\,p$$

Овие симболи ги претствавуваат следниве карактеристики на системот:

- распределба на интервал меѓу две пристигнувања на клиенити
- распределба на време на опслужување
- *s* - број на сервери
- *q* - дисциплина на редење (FIFO, LIFO…)
- *c* - капацитет на системот
- *p* - големина на популација (број на можни клиенти)

За *A* и *B* можат да се земат следниве ознаки:

- *M* - Пуасонов поток
- E_m - Ерлангова распределба
- *D* - симбол за детерминистичко (познато) пристигнување и константно време на опслужување
- *G* - општа (било која) распределба
- *GI* - општа (било која) распределба со независни случајни променливи

6.1.3 Прост поток на настани

Како што е стомната погоре, стационарен, ординарен поток без последејствие се нарекува **прост поток на настани**.
Дадените услови дозволуваат наоѓање на веројатноста $P_k(t)$ со точност до еден параметар.

Теорема 6.1. Прост поток на настани ја има следнава особина:

$P_1(\Delta t) = \lambda \Delta t + \sigma(\Delta t)$ каде $\lambda > 0$ е константа, односно

$$\lim_{\Delta t \to 0} \frac{P_1(\Delta t)}{\Delta t} = \lambda \quad \text{или} \qquad P_0(\Delta t) = e^{-\lambda \Delta t}.$$

Доказ: Нека $t = 1$ и се означува со $\theta = P_0(t)$ (веројатноста во единичен временски интервал да не се случи ниеден настан). Интервалот t = 1 е поделен на n подинтервали со должина 1/n. За да не настапи настан на целиот единичен интервал потребно и доволно е да не настапи настан на ниту еден од подинтервалите. Потокот е стационарен и без последејствие, па следи дека

$$\theta = \left[P_0\left(\frac{1}{n}\right)\right]^n \quad \text{или} \quad P_0\left(\frac{1}{n}\right) = \theta^{\frac{1}{n}}.$$

Слично се добива и дека $P_0\left(\frac{k}{n}\right) = \theta^{\frac{k}{n}}$.

Нека е $t > 0$. За дадено n постои цел позитивен број k така што $\frac{k-1}{n} \le t \le \frac{k}{n}$.

Бидејќи $P_0(t)$ е нерастечка функција од времето ($P_0(0) = 1$, $P_0(\infty)=0$) следи:

$$P_0\left(\frac{k-1}{n}\right) \ge P_0(t) \ge P_0\left(\frac{k}{n}\right) \quad \text{и} \quad \theta^{\frac{k-1}{n}} \ge P_0(t) \ge \theta^{\frac{k}{n}},$$

$$\lim_{n\to\infty} \frac{k}{n} = \lim_{n\to\infty} \frac{k-1}{n} = t \text{ (} k \text{ зависи од } n\text{)} \quad \text{и} \quad \lim_{n\to\infty} P_0\left(\frac{k-1}{n}\right) = \lim_{n\to\infty} P_0\left(\frac{k}{n}\right) = \theta^t.$$

Значи:

$$P_0(t) = \theta^t$$

Бидејќи $0 \le P_0(t) \le 1$ може да се случи една од следниве ситуации:

1. $\theta = 0$,
2. $\theta = 1$.
3. $0 < \theta < 1$;

Првите два случаи ќе се отфрлат:

1. Во првиот случај $P_0(t)$ =0 за произволно t од каде непосредно следи дека за произволен временски интервал настапува со веројатност 1 барем еден настан. Од овде веднаш следи дека во интервал со произволна должина со веројатност 1 настапуваат бесконечно многу настани, но во тој случај потокот не е ординарен.

2. Во вториот случај $P_0(t) = 1$ - нема поток на настани.

3. Бидејќи $0 < \theta < 1$ се става $\theta = e^{-\lambda}$, каде $\lambda > 0$, па следи $P_0(t) = e^{-\lambda t} = 1 - \lambda t + \sigma(t)$ (од Тејлоровиот развој на функцијата). Бидејќи важи $P_0(t) + P_1(t) + P_{>1}(t) = 1$ и услов на ординарност, следи дека $1 - \lambda t + \sigma(t) + P_1(t) + \sigma(t) = 1$, односно $P_1(t) = \lambda t + \sigma(t)$.

Да се определи сега $P_k(t)$. Се разгледува системот на интервалот $[0,\ t + \Delta t]$ и се определува $P_k(t+\Delta t)$. Во интервалот $[0,\ t+\Delta t]$ можат да се случат k настани на следниов начин:

1. k настани во интервалот $[0,\ t]$ и 0 настани во $[t,\ t + \Delta t]$;

2. k-1 настани во интервалот $[0, t]$ и 1 настан во $[t, t + \Delta t]$;
...
3. 0 настани во интервалот $[0, t]$ и k настани во $[t, t + \Delta t]$;

Од претходното следи дека

$$P_k(t+\Delta t) = P_k(t)P_0(\Delta t) + P_{k-1}(t)P_1(\Delta t) + \sum_{i=0}^{k-2} P_i(t)P_{k-i}(\Delta t).$$

Нека $\qquad I = \sum_{i=0}^{k-2} P_i(t)P_{k-i}(\Delta t).$

Бидејќи $P_i(t) \leq 1$ следи дека:

$$I < \sum_{i=0}^{k-2} P_{k-i}(\Delta t) = \sum_{j=2}^{k} P_j(\Delta t) \leq \sum_{j=2}^{\infty} P_j(\Delta t) = P_{>1}(\Delta t) = \sigma(\Delta t)$$

(се воведува смена $j = k - i$).
Според тоа:

$$P_k(t+\Delta t) = P_{k-1}(t)P_1(\Delta t) + P_k(t)P_0(\Delta t) + \sigma(\Delta t).$$

$$P_k(t+\Delta t) = P_{k-1}(t)\big(\lambda\Delta t + \sigma(\Delta t)\big) + P_k(t)\big(1 - \lambda\Delta t + \sigma(\Delta t)\big) + \sigma(\Delta t).$$

$$P_k(t+\Delta t) = \lambda P_{k-1}(t)\Delta t + P_k(t) - \lambda P_k(t)\Delta t + \sigma(\Delta t).$$

$$\frac{P_k(t+\Delta t) - P_k(t)}{\Delta t} = \lambda P_{k-1}(t) - \lambda P_k(t) + \frac{\sigma(\Delta t)}{\Delta t}.$$

Кога $\Delta t \to 0$

$$P_k'(t) = \lambda P_{k-1}(t) - \lambda P_k(t).$$

Секако, последново важи за $k \geq 1$.

За $k=0$

$$P_0(t+\Delta t) = P_0(t)P_0(\Delta t) = P_0(t)\big(1 - \lambda\Delta t + \sigma(\Delta t)\big).$$

$$\frac{P_0(t+\Delta t) - P_0(t)}{\Delta t} = -\lambda P_0(t) + \frac{\sigma(\Delta t)}{\Delta t}.$$

$$P_0'(t) = -\lambda P_0(t).$$

За решавање на овој систем диференцијални равенки се воведува следнава смена:

$$P_k(t) = u_k(t)\, e^{-\lambda t}.$$

$k \geq 1$: $\qquad u_k'(t)e^{-\lambda t} - u_k(t)\lambda e^{-\lambda t} = \lambda u_{k-1}(t)e^{-\lambda t} - \lambda u_k(t)e^{-\lambda t},$

$k = 0$: $\qquad u_0'(t)e^{-\lambda t} - u_0(t)\lambda e^{-\lambda t} = -\lambda u_0(t)e^{-\lambda t},$

$$u_0'(t)e^{-\lambda t} = 0 \Rightarrow u_0'(t) = 0 \Rightarrow u_0(t) = c,$$

$$u_0(t) = c \Rightarrow P_0(t) = ce^{-\lambda t},$$
$$P_0(0) = 1 \Rightarrow ce^{-0} = 1 \Rightarrow c = 1.$$

Значи $u_0(0) = P_0(0)$, односно $u_0(t) = 1$.

$$u_1'(t) = \lambda u_0(t) = \lambda \Rightarrow u_1(t) = \lambda t + c$$

$P_k(0)=0 \Rightarrow 0 = u_k(0)e^{-0} = u_k(0)$; $u_1(0) = \lambda 0 + c = 0 \Rightarrow c = 0$

$u_1(t) = \lambda t$;

$u_2'(t) = \lambda u_1(t) = \lambda\lambda t = \lambda^2 t \Rightarrow u_2(t) = \lambda^2 t^2/2 + c_2$; c_2 е исто така 0; итн...

$$u_k(t) = \frac{(\lambda t)^k}{k!}$$

или

$$P_k(t) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}$$

Бројот на настани кои се случуваат во одредено време кај прост поток на настани има пуасонова распределба.

Математичкото очекување или средниот број на настапувања на настани во интервал со должина t е еднаков на

$$E(X(t)) = \bar{k} = \sum_{k=1}^{\infty} k P_k(t) = \sum_{k=1}^{\infty} k \frac{(\lambda t)^k}{k!} e^{-\lambda t} = e^{-\lambda t} \sum_{k=1}^{\infty} k \frac{(\lambda t)^k}{k!} =$$

$$= e^{-\lambda t} \lambda t \sum_{k=1}^{\infty} k \frac{(\lambda t)^{k-1}}{k(k-1)!} = e^{-\lambda t} \lambda t \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} = e^{-\lambda t} \lambda t e^{\lambda t} = \lambda t;$$

$$\bar{k} = \lambda t.$$

Слично може да се покаже и дека $D(X(t)) = \lambda t$.

Дефиниција 6.1 Интензитетот I на поток на настани е математичко очекување на бројот на настани во единица време.
Значи, за прост поток $I = \lambda$.
Со $\pi(t)$ се означи веројатноста за време t да се случи барем еден настан.

$$\pi(t) = \sum_{k=1}^{\infty} p_k(t) = 1 - p_0(t).$$

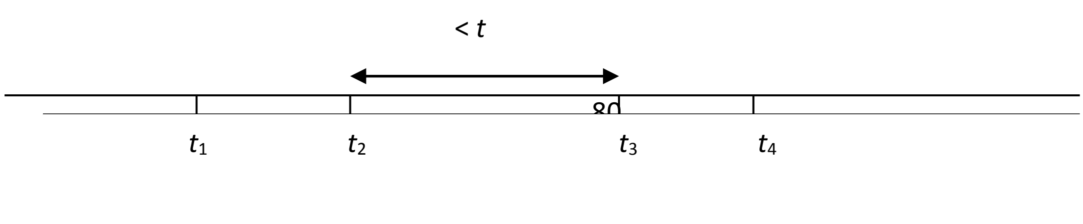
Дефиниција 6.2 Параметар на поток е границата

За прост поток на $\lim\limits_{t\to 0} \frac{\pi(t)}{t} = \lim\limits_{t\to 0} \frac{\sum_{k=1}^{\infty} P_k(t)}{t} = \lim\limits_{t\to 0} \frac{1 - P_o(t)}{t}$ настани

$$\lim_{t\to 0} \frac{1 - P_o(t)}{t} = \lim_{t\to 0} \frac{1 - (1 - \lambda t + \sigma(t))}{t} = \lambda.$$

Кај прост поток на настани параметарот на потокот има значење на среден број на настапувања на настани во единица време.

Се определува законот на распределба на должината на интервалот помеѓу случувања на два последователни настани.
Се бара $P(T < t) = F(t)$ (веројатноста дека интервалот T помеѓу настапувања на два настани е помал од t).



Фиг. 6.2 Времиња помеѓу две последователни пристигнување

Веројатноста дека во интервал со должина t нема да се случи ниту еден настан е

$$P_0(t) = e^{-\lambda t},$$

а таа е еднаква на веројатноста дека интервалот помеѓу случувања на два последователни настана е поголем од t.

$$P_0(t) = e^{-\lambda t} = P(T \geq t) = 1 - F(t).$$

Од тука

$$F(t) = P(T < t) = 1 - P(T \geq t) = 1 - e^{-\lambda t}.$$

Густината на распределба е

$$p(t) = F'(t) = \lambda e^{-\lambda t}.$$

Значи, пуасоновиот поток има експоненцијална распределба на должината на интервалот помеѓу настапувања на два последователни настана.

Значењето на простиот поток на настани во теоријата на масовното опслужување е слично на значењето на нормалната распределба во теоријата на веројатаност и статистика поради тоа што кај потоците на настани делува исто како централната гранична теорема во теоријата на веројатност.

Сумарниот поток на множество потоци на прости настани е, исто така, прост поток.

6.1.4 Времетраење на опслужување

Времетраењето на опслужување е најчесто случајна променлива. Овде се разгледани примери кога тоа е случајна променлива со експоненцијална распределба:

$$F(t) = P(T < t) = 1 - e^{-\mu t}, \quad t \geq 0.$$

Овде параметарот $1/\mu$ има значење на средна должина на времетраењето на опслужување, додека дисперзијата на оваа променлива е $1/\mu^2$. ($E(T) = 1/\mu$, $D(T) = 1/\mu^2$).

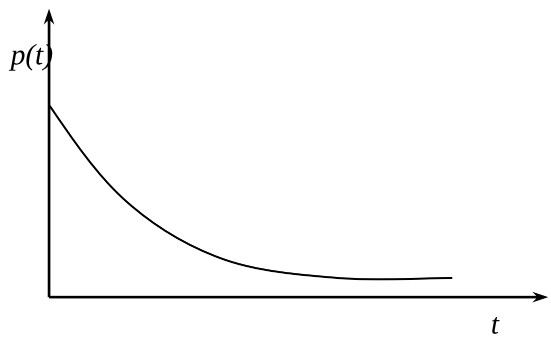
Оваа распределба овозможува наоѓање аналитички решенија за поединечни системи за масовно опслужување, но таа одговара и на многу практични примери на опслужување. Освен тоа, ја поседува и следнава значајна особина: распределбата на преостанатото време на опслужување не зависи од тоа колку време опслужувањето веќе траело. Ова својство на експоненцијалната распределба особено доаѓа до израз кај пуасоновиот поток на настани, каде времето на доаѓање на следниот клиент не зависи од тоа кога дошол последниот клиент.

Доказ на својството на експоненцијалната распределба:

Се воочува дека $F(t) = P(T < t) = 1 - e^{-\mu t}$; $1 - F(t) = P(T \geq t) = e^{-\mu t}$; $P(t \geq a) = e^{-\mu a}$.
Очигледно $P(T > t + a) = e^{-\mu(t+a)}$.
Важи следново: $P(T > t + a) = P(T \geq a)\, P_a(T \geq t)$, каде $P_a(T \geq t)$ е условна веројатност дека опслужувањето ќе трае уште најмалку t временски единици ако пред тоа траело a единици.

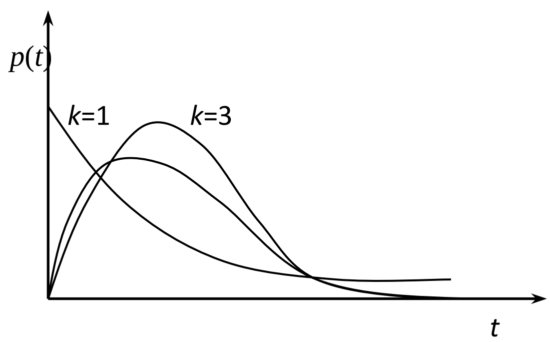
$$P_a(T \geq t) = \frac{P(T \geq t + a)}{P(T \geq a)} = \frac{e^{-\mu(t+a)}}{e^{-\mu a}} = e^{-\mu t} = P(T \geq t)$$

Густина на распределба на оваа случајна променлива е $p(t) = F'(t) = \mu e^{-\mu t}$.



Фиг. 6.3 Густина на распределба за променлива со експоненцијална распределба

Се гледа дека модалната вредност на распределбата е за $t = 0$, односно дека експоненцијалната распределба добро ги опишува примерите во кои времетраењето на опслужување за најголем број на клиенти е многу кратко.
За случаи кога оваа претпоставка не е основана времетраењето на опслужување се опишува со ерланговa распределба.



Фиг. 6.4 Функција на густина на распределба за променлива со ерлангова распределба за различни вредности на k.

Густината на распределба е од следниов облик:

$$f(t) = \frac{(\mu k)^k}{(k-1)!} t^{k-1} e^{-\mu k t}, t \geq 0$$

$E(T)=1/\mu$, $D(T)=1/(k\mu^2)$.

6.1.5 Опслужување со чекање

Се разгледува следниов процес на масовно опслужување:
Системот содржи n канали за опслужување. Секој канал во даден момент опслужува само еден клиент и еден клиент има потреба од опслужување од само еден канал. Ако сите канали се зафатени клиентот застанува во редица и чека да биде опслужен. Системот може да прими неограничен број на клиенти. Тие немаат ограничено време на чекање. Влезниот поток на клиенти е прост а времетраењето на опслужување е

распределено со експоненцијална распределба на секој од каналите, при што тие се меѓусебно независни.

Во произволен временски момент во системот се наоѓаат k клиенти (k = 0,1,2,3,...).
Ако $k \leq n$ тогаш $s = n - k$ е број на слободни канали;
Ако $k > n$ тогаш $L_Q = k - n$ е број на клиенти во редицата.

Со $p_k(t)$ се означува веројатноста дека во системот се наоѓаат k клиенти во временскиот момент t.
Важи:

$$\sum_{k=0}^{\infty} p_k(t) = 1$$

Веројатноста дека во интервалот Δt во системот ќе дојде еден клиент е еднаква на

$$P_1(\Delta t) = \lambda\Delta t + \sigma(\Delta t)$$

Веројатноста дека за време Δt нема да дојде клиент е еднаква на

$$P_0(\Delta t) = 1 - \lambda\Delta t + \sigma(\Delta t)$$

Веројатноста дека во интервалот Δt ќе заврши опслужувањето на одреден канал е еднаква на $P(T < \Delta t) = 1 - e^{-\mu\Delta t} = 1 - (1 - \mu\Delta t + \mu^2\Delta^2 t/2! - ...) = \mu\Delta t + \sigma(\Delta t)$.

Веројатноста дека во интервалот нема да заврши опслужувањето на одреден канал е еднаква на $P(T \geq \Delta t) = e^{-\mu\Delta t} = 1 - \mu\Delta t + \mu^2\Delta^2 t/2! - ... = 1 - \mu\Delta t + \sigma(\Delta t)$.

Да се пресмета веројатноста дека во моментот $t + \Delta t$ сите канали се слободни. Тоа може да се случи на еден од следниве начини:

a. Во временскиот момент t сите канали биле слободни а во Δt не дошол ниту еден клиент:

$$p_0(t)\,(1 - \lambda\Delta t + \sigma(\Delta t)) = p_0(t) - \lambda p_0(t)\Delta t + \sigma(\Delta t);$$

b. Во моментот t точно еден канал бил зафатен, но во Δt опслужувањето завршило и ниту еден клиент не дошол во Δt.

$$p_1(t)\,(1 - \lambda\Delta t + \sigma(\Delta t))\,(\mu\Delta t + \sigma(\Delta t)) = p_1(t)\mu\Delta t + \sigma(\Delta t).$$

$$p_0(t + \Delta t) = p_0(t) - \lambda p_0(t)\,\Delta t + \sigma(\Delta t) + p_1(t)\mu\Delta t + \sigma(\Delta t)$$
$$= p_0(t) - \lambda p_0(t)\Delta t + p_1(t)\mu\Delta t + \sigma(\Delta t).$$

$$\frac{p_0(t+\Delta t) - p_0(t)}{\Delta t} = -\lambda p_0(t) + \nu p_1(t) + \frac{\sigma(\Delta t)}{\Delta t}$$

со $\Delta t \to 0$

$$p_0'(t) = -\lambda p_0(t) + \mu p_1(t).$$

Сега да се најде веројатноста дека во моментот $t + \Delta t$ во системот има k клиенти,каде $1 \leq k < n$. Тоа е можно во следниве случаи:

a) Во моментот t во системот имало k клиенти, во интервалот Δt во системот ниту дошол клиент, ниту се завршило опслужување
b) Во моментот t во системот имало k-1 клиенти, во интервалот Δt во системот дошол еден клиент а не завршило ниедно опслужување;

4. Во моментот t во системот имало k+1 клиенти, во интервалот Δt во системот не дошол клиент, а завршило едно опслужување (тоа е можно на $\binom{k+1}{1}$ начини, останатите k не завршиле со опслужување);

5. Во моментот t во системот имало k клиенти, во интервалот Δt во системот дошол еден клиент а и еден завршил (тоа е можно на $\binom{k}{1}$ начини, останатите k-1 не завршиле со опслужување)

Сите останати промени во интервалот Δt се пропорционални на $\sigma(\Delta t)$(доаѓање на два клиента, завршување на опслужување на два клиента итн.). Затоа:

Со
$$p_k(t+\Delta t) = p_k(t)(1-\lambda\Delta t)(1-\mu\Delta t)^k + p_{k-1}(t)\lambda\Delta t(1-\mu\Delta t)^{k-1} + $$
$$+ p_{k+1}(t)(1-\mu\Delta t)\binom{k+1}{1}\mu\Delta t(1-\mu\Delta t)^k + p_k(t)\lambda\Delta t\binom{k}{1}\mu\Delta t(1-\mu\Delta t)^{k-1}$$
средувањето се добива

$$p_k(t+\Delta t) = p_k(t)(1-\lambda\Delta t)(1-\mu k\Delta t+\sigma(\Delta t)) + p_{k-1}(t)\lambda\Delta t(1-\mu(k-1)\Delta t+\sigma(\Delta t))$$
$$+ p_{k+1}(t)(1-\lambda\Delta t)(k+1)\mu\Delta t + (1-\mu k\Delta t+\sigma(\Delta t)) + p_k(t)\lambda\Delta t k\mu\Delta t(1-\mu(k-1)\Delta t+\sigma(\Delta t)).$$

$$p_k(t+\Delta t) = p_k(t)(1-\lambda\Delta t-\mu k\Delta t+\sigma(\Delta t)) + p_{k-1}(t)(\lambda\Delta t+\sigma(\Delta t)) + p_{k+1}(t)(k+1)(\mu\Delta t+\sigma(\Delta t)).$$

$$\frac{p_k(t+\Delta t)-p_k(t)}{\Delta t} = -\lambda p_k(t) - \mu k p_k(t) + p_{k-1}(t)\lambda + p_{k+1}(t)(k+1)\mu + \frac{\sigma(\Delta t)}{\Delta t}.$$

$\Delta t \to 0$

$$p_k'(t) = -(\lambda+\mu k)p_k(t) + \lambda p_{k-1}(t) + (k+1)\mu p_{k+1}(t).$$

Аналогно за $k \geq n$ ќе се добие

$$p_k'(t) = -(\lambda+\mu n)p_k(t) + \lambda p_{k-1}(t) + n\mu p_{k+1}(t).$$

Овие диференцијални равенки се прилично тешки за решавање. Ќе се најдат стационарни решенија, т.е. кога веројатностите се независни од времето t, тие се од поголем интерес во теоријата на масовно опслужување. Постоењето на такви решенија е обезбедено од таканаречените ергодични теореми Во овој случај постојат граничните (стационарни) веројатности, кои се означени со p_k,односно овие се гранични веројатности на $p_k(t)$ кога $t\to\infty$. Тогаш $p_k'(t)\to 0$ кога $t\to\infty$. Горниот систем преминува во систем линеарни равенки:

$$-\lambda p_0 + \mu p_1 = 0$$
$$\lambda p_{k-1} - (\lambda + k\mu)p_k + (k+1)\mu p_{k+1} = 0 \quad ; \ 1 \leq k < n$$
$$\lambda p_{k-1} - (\lambda + n\mu)p_k + n\mu p_{k+1} = 0 ; \qquad k \geq n.$$
$$\sum_{k=0}^{\infty} p_k = 1$$

За решавање на системот воведуваме смена $z_k = \lambda p_{k-1} - k\mu p_k$ за $1 \leq k < n$ и

$$z_k = \lambda p_{k-1} - n\mu p_k \text{ за } k \geq n.$$

Системот се трансформира во

$$z_1 = 0, \ z_k - z_{k+1} = 0 \text{ за } k \geq 1.$$

Значи, за сите $k \geq 1$, $z_k = 0$, односно за $1 \leq k < n$ $\lambda p_{k-1} = k\mu p_k$ а за $k \geq n$, $\lambda p_{k-1} = n\mu p_k$. Ако означиме $\rho = \lambda/\mu$, од претходниве равенки следи дека за $1 \leq k \leq n$:

$p_k = \frac{\rho^k}{k!} p_0$, а за $k > n$ $p_k = \left(\frac{\rho}{n}\right)^{k-n} p_n$, односно $p_k = \frac{\rho^k}{n! n^{k-n}} p_0$,

Останува уште да се одредили p_0. Во сумата $\sum_{k=0}^{\infty} p_k = 1$ за p_k се заменуваат изразите од последните равенства и се добива

$$p_0 \left[\sum_{k=0}^{n} \frac{\rho^k}{k!} + \frac{n^n}{n!} \sum_{k=n+1}^{\infty} \left(\frac{\rho}{n}\right)^k \right] = 1.$$

Бидејќи бесконечниот ред во последниот израз конвергира само за $\rho < n$, под претпоставка дека ова е точно, се добива:

$$p_0 = \left[\sum_{k=0}^{n} \frac{\rho^k}{k!} + \frac{n^n}{n!} \sum_{k=n+1}^{\infty} \left(\frac{\rho}{n}\right)^k \right]^{-1}.$$

Ако е $\rho \geq n$ сите p_k мора да бидат 0. Тоа значи дека не постои позитивна гранична веројатност дека во системот се наоѓаат конечен број клиенти, односно кога $t \to \infty$ редицата по веројатност станува бесконечна. Практично, тоа значи дека за да нема бесконечно натрупување во редицата потребно е просечниот број клиенти кои доаѓаат во единица време да биде помал од просечниот број клиенти кои се опслужуваат во единица време. Од овие резултати понатаму може да се определат повеќе значајни параметри за системот за масовно опслужување.Од интерес се следниве параметри: просечен број клиенти во редицата просечен број на слободни канали (за да се знае искористеноста на серверите), веројатноста дека се зафатени сите канали итн.

Со L_Q ќе се означува просечниот број клиенти во редицата.

$$L_Q = \sum_{k=n+1}^{\infty} (k-n) p_k = \sum_{k=n+1}^{\infty} (k-n) \frac{\rho^k}{n! n^{k-n}} p_0 = \frac{n^n}{n!} p_0 \sum_{k=n+1}^{\infty} \frac{(k-n)\rho^k}{n^k} =$$

$$= \frac{n^n}{n!} \left(\frac{\rho}{n}\right)^{n+1} \frac{1}{\left(1-\frac{\rho}{n}\right)^2} p_0 = \frac{\rho^{n+1}}{n! n} \frac{1}{\left(1-\frac{\rho}{n}\right)^2} p_0 = \frac{\rho}{n} \frac{1}{\left(1-\frac{\rho}{n}\right)^2} p_n$$

Образложение:

$$\sum_{k=n+1}^{\infty} \frac{(k-n)\rho^k}{n^k} = \sum_{j=1}^{\infty} j \left(\frac{\rho}{n}\right)^{j+n} =$$

$$= \left(\frac{\rho}{n}\right)^n \sum_{j=1}^{\infty} j \left(\frac{\rho}{n}\right)^j = \left(\frac{\rho}{n}\right)^n \frac{\rho}{n} \sum_{j=1}^{\infty} j \left(\frac{\rho}{n}\right)^{j-1} = \left(\frac{\rho}{n}\right)^n \frac{\rho}{n} \frac{1}{\left(1-\frac{\rho}{n}\right)^2}$$

Да се пресмета просечниот број слободни канали.

$$\bar{s} = \sum_{k=0}^{n} (n-k) p_k = \sum_{k=0}^{n-1} (n-k) \frac{\rho^k}{k!} p_0 = \sum_{k=0}^{n-1} n \frac{\rho^k}{k!} p_0 - \sum_{k=0}^{n-1} k \frac{\rho^k}{k!} p_0 =$$

$$= n p_0 \sum_{k=0}^{n-1} \frac{\rho^k}{k!} - p_0 \rho \sum_{k=1}^{n-1} \frac{\rho^{k-1}}{(k-1)!} = n p_0 \sum_{k=0}^{n-1} \frac{\rho^k}{k!} - p_0 \rho \sum_{k=1}^{n} \frac{\rho^{k-1}}{(k-1)!} + \rho p_0 \frac{\rho^{n-1}}{(n-1)!} =$$

$$= (n-\rho) p_0 \sum_{k=0}^{n-1} \frac{\rho^k}{k!} + \frac{p_0 \rho^n}{(n-1)!} = p_0 (n-\rho) \left(\sum_{k=0}^{n-1} \frac{\rho^k}{k!} + \frac{\rho^n}{(n-\rho)(n-1)!} \right) =$$

$$= p_0 (n-\rho) \frac{1}{p_0} = n - \rho.$$

Просечен број на зафатени канали (клиенти што се опслужуваат):

$$\bar{\rho} = n - \bar{s} = n - n + \rho = \rho$$

Просечен број на клиенти во системот:

$$L = L_Q + \overline{p} = L_Q + n - \overline{s} = L_Q + \rho$$

Веројатност дека се зафатени сите канали:

$$P(k \geq n) = \sum_{k=n}^{\infty} p_k = \sum_{k=n}^{\infty} \frac{\rho^k}{n! n^{k-n}} p_0 = \frac{p_0 n^n}{n!} \sum_{k=n}^{\infty} \left(\frac{\rho}{n}\right)^k = \frac{p_0 n^n}{n!} \left(\frac{\rho}{n}\right)^n \sum_{k=0}^{\infty} \left(\frac{\rho}{n}\right)^k =$$

$$= \frac{p_0 \rho^n}{n!} \frac{1}{1-\rho/n} = \frac{p_n}{1-\rho/n}.$$

6.1.6 Распределба на веројатност за должина на чекање кај систем $M|M|n|\infty$

Должината на чекање во редицата е случајна променлива, која ќе се означува со W_Q.

Нека е:

$P(W_Q > t)$ - веројатност дека чекањето во редицата трае најмалку t временски единици

$P_k(W_Q > t)$ - условна веројатност дека клиентот што влегува во системот ќе чека најмалку t временски единици, ако во системот има k клиенти.

Според формулата за тотална веројатност

$$P(W_Q > t) = \sum_{k=n}^{\infty} P_k P_k(W_Q > t)$$

(бидејќи $P_k(W_Q > t) = 0, k = 1,2,..n-1$ ако во системот има повеќе канали отколку клиенти, клиентот не чека.)

Ако во моментот на доаѓање на клиентот во редицата се наоѓаат k-n клиенти и опслужувањето е по редоследот на доаѓање, новодојдениот клиент треба да чека да бидат опслужени k-n+1 клиенти.

Нека $q_s(t)$ е веројатност дека во интервалот со должина t по доаѓањето на клиентот ќе завршат со опслужување точно s клиенти.

За $k \geq n$ важи равенство

$$P_k(W_Q > t) = \sum_{s=0}^{k-n} q_s(t)$$

Веројатноста дека во интервалот за време t нема да заврши ниту едно опслужување (не се ослободи ниту еден канал) е еднаква на

$$q_0(t) = P(T > t)^n = (e^{-\mu t})^n = e^{-\mu tn}$$

Ако сите сервери се зафатени и сеуште има доволно долга редица на клиенти, потокот на клиенти кои заминуваат после опслужувањето ќе биде прост. Всушност, во овој случај сите три услови (стационарност, ординарност и отсуство на последејствие) се задоволени. Средно време на опслужување е $1/(n\mu)$, односно среден број на опслужени клиенти во единица време е $n\mu$. Веројатноста дека точно s сервери ќе бидат ослободени во временскиот интервал t е

$$q_s(t) = \frac{(n\mu t)^s}{s!} e^{-n\mu t}$$

или

$$P_k(W_Q > t) = \sum_{s=0}^{k-n} \frac{(n\mu t)^s}{s!} e^{-n\mu t}.$$

Од тука

$$P(W_Q > t) = \sum_{k=n}^{\infty} \left(\frac{\rho}{n}\right)^{k-n} p_n \sum_{s=0}^{k-n} \frac{(n\mu t)^s}{s!} e^{-n\mu t}$$

$$P(W_Q > t) = p_n e^{-n\mu t} \sum_{k=n}^{\infty} \left(\frac{\rho}{n}\right)^{k-n} \sum_{s=0}^{k-n} \frac{(n\mu t)^s}{s!};$$

$$P(W_Q > t) = p_n e^{-n\mu t} \sum_{s=0}^{\infty} \frac{(n\mu t)^s}{s!} \sum_{k=s+n}^{\infty} \left(\frac{\rho}{n}\right)^{k-n};$$

$$P(W_Q > t) = p_n e^{-n\mu t} \sum_{s=0}^{\infty} \frac{(n\mu t)^s}{s!} \left(\frac{\rho}{n}\right)^{s} \frac{1}{1-\rho/n};$$

$$P(W_Q > t) = \frac{p_n e^{-n\mu t}}{1-\rho/n} \sum_{s=0}^{\infty} \frac{(\lambda t)^s}{s!};$$

$$P(W_Q > t) = p_n e^{-n\mu t} \sum_{s=0}^{\infty} \sum_{k=s+n}^{\infty} \left(\frac{\rho}{n}\right)^{k-n} \frac{(n\mu t)^s}{s!};$$

$$P(W_Q > t) = p_n e^{-n\mu t} \sum_{s=0}^{\infty} \frac{(n\mu t)^s}{s!} \left(\frac{\rho}{n}\right)^{s} \sum_{k=0}^{\infty} \left(\frac{\rho}{n}\right)^{k};$$

$$P(W_Q > t) = \frac{p_n e^{-n\mu t}}{1-\rho/n} \sum_{s=0}^{\infty} \frac{(n\mu t)^s}{s!} \left(\frac{\rho}{n}\right)^{s};$$

$$P(W_Q > t) = \frac{p_n e^{-n\mu t + \lambda t}}{1-\rho/n} = \pi e^{-t(n\nu - \lambda)}.$$

(π е веројатноста дека сите сервери се зафатени.)

Јасно, за $t < 0$, $P(W_Q > t) = 1$.

Функцијата $P(W_Q > t)$ има прекин во $t = 0$; скокот е еднаков на веројатноста дека во моментот на доаѓање на клиентот сите сервери ќе бидат зафатени.

Со користење на претходниот резултат можеме да ги најдеме сите потребни нумерички карактеристики на времето на чекање. Конкретно, математичкото очекување на времето на чекање (или средно време на чекање) е еднакво на

$$a = E(W_Q) = \int_0^{\infty} t dP(W_Q > t) = \pi \int_0^{\infty} t(n\mu - \lambda) e^{-(n\mu - \lambda)t} dt.$$

Со едноставно интегрирање се добива

$$a = \frac{\pi}{\mu(n-\rho)}.$$

Да се пресмета и средната загуба на време за сите клиенти кои пристигнуваат во системот за временски интервал со должина T. Во временскиот интервал T просечно λT клиенти пристигнуваат во системот, средната загуба на време за тие клиенти е еднаква на $a\lambda T$, односно

$$a\lambda T = \frac{\pi\lambda}{\mu(n-\rho)} = \frac{\pi\rho T}{n-\rho}.$$

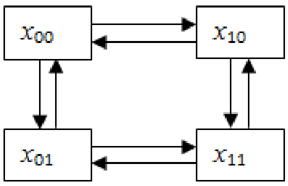
6.2 Прости системи за масовно опслужување и Маркови процеси

Прости системи за масовно опслужување обично се нарекуваат системите кај кои влезниот поток е пуасонов а распределбата на времетраењето на опслужување е експоненцијална. Овие процеси се специјален вид на марковите процеси, па за полесно изучување на овие системи потребни се одредени предзнаења од процесите на Марков, а особено теоремата за определување веројатностите на премин.

6.2.1 Случајни процеси

Нека (Ω, F, P) е даден простор на веројатност, T е дадено множество. Тогаш фамилијата случајни променливи $\{X(t) \mid t \in T\}$од дадениот простор се нарекува случаен процес. Множеството T може да биде дискретно или непрекинато. Ако t е фиксирано, тогаш $X(t)$ е случајна променлива. Множеството состојби на случајната променлива $X(t)$ кога t е фиксирано може да биде дискретно или непрекинато. Во зависност од тоа процесот може да биде дискретен или непрекинат. Овде не интересираат дискретни процеси со непрекинато време, бидејќи со нив може да се опишуваат системите за масовно опслужување.

Пример 6.1. Систем за масовно опслужување со два сервера може да се опишува со дискретен процес со непрекинато време. Во фиксиран временски момент t системот може да се најде во една од следниве 4 состојби: x_{00}, x_{10}, x_{01}, x_{11}. Индексите означуваат незафатеност, односно зафатеност на серверите (пр: состојбата x_{10} означува дека првиот сервер е зафатен, а вториот е слободен). Може се нацртат премините кои можат да настанат за мал временски момент.



Фиг. 6.5 Граф на состојби и премини за двоканален систем.

Премини од состојба x_{00} во состојба x_{11} и обратно за мал временски интервал се многу малку веројатни (како и од x_{10} во состојба x_{01} и обратно), па не се означени.

Основна задача е да се најдат веројатностите дека во даден момент процесот се наоѓа во дадена состојба, односно $p_i(t) = P\{X(t) = x_i\}$.

6.2.2 Дискретен Марков процес со непрекинато време

Марковиот процес спаѓа во групата дискретни случајни процеси, а го има својството:

$$P\{X(t_n) = x_j \mid X(t_1)=x_{i1}, X(t_2) = x_{i2},\dots X(t_{n-1}) = x_i\} = P\{X(t_n) = x_j \mid X(t_{n-1}) = x_i\}$$

(Условната веројатност дека во даден момент t_n процесот се наоѓа во состојба x_j ако во претходните моменти се наоѓал во состојби x_i, ... x_{i2}, x_{i1} зависи само од состојбата во која се наоѓал во претходниот момент.)
Со $p_{ij}(t_0, t)$ се означува веројатноста $P\{X(t_0 + t) = x_j \mid X(t_0) = x_i\}$.

Својства на дискретни Маркови процеси: Еден марков процес е **хомоген** ако $p_{ij}(t_0, t)$ не зависи од t_0 , туку само од должината t на интервалот помеѓу двете состојби, односно $p_{ij}(t_0, t) = p_{ij}(t)$. Таков процес е без последејствие и веројатностите $p_{ij}(t)$ образуваат матрица на премини $\mathbf{P}(t) = [p_{ij}(t)]$, во случај на конечен број на состојби.

Важат равенствата на Чепмен-Колмогоров

$$p_{ij}(t+\tau) = \sum_{k\in K} p_{ik}(t)\, p_{kj}(\tau)$$

i, $j \in K$, K е множеството индекси на можните состојби.

Во матричен облик последните равенства се запишуваат како: $\mathbf{P}(t+\tau) = \mathbf{P}(t)\mathbf{P}(\tau)$.

Ако со $p_k(t)$ ја означиме веројатноста дека во моментот t процесот се наоѓа во состојба x_k, независно од каде дошол во таа состојба, тогаш

$$p_k(t) = \sum_{k \in K} p_{ik}(t)$$

$$p_k(t+\tau) = \sum_{i \in K} p_i(t) p_{ik}(\tau)$$

Овде разгледуваме хомогени дискретни Маркови процеси.

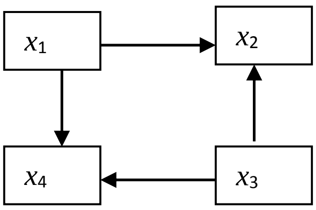
Дискретен Марков процес со непрекинато време е **ергодичен** ако постојат гранични вредности за веројатностите системот да се наоѓа во одредена состојба кога времето тежи кон бесконечност, односно

$$\lim_{\tau \to \infty} p_{ij}(t_0, \tau) = p_j, \text{koga procesot ne e homogen}$$

$$\lim_{\tau \to \infty} p_{ij}(\tau) = p_j, \text{koga procesot e homogen}.$$

Еден процес е **транзитивен** ако за произволен пар состојби x_i, x_j постои $t > 0$ така што ако процесот е нехомоген тогаш $p_{ij}(t_0, t) > 0$, а ако процесот е хомоген тогаш $p_{ij}(t) > 0$ (всушност, за произволни две состојби можно е да се стигне од едната до другата).

Пример 6.2. Од Фигура 6.6 (граф на еден процес) се гледа дека тој не е транзитивен(не може да се премине од состојба x_1 во x_3, x_1 и x_3 се состојби без влез, а x_2 и x_4 се состојби без излез)



Фиг. 6.6 Пример на нетранзитивен процес

Теорема 3.2 (Теорема на Марков) Произволен транзитивен хомоген Марков процес со конечен број на состојби е ергодичен.

Равенствата на Чепмен-Колмогоров овозможуваат да се дефинираат диференцијални равенки за произволен марков процес под услови да:

- Познати се сите можни состојби.
- Може да се нацрта графот на можни премини.
- За секој можен премин од една состојба x_i до друга x_j се знае интензитетот на премин $\lambda_{ij}(t)$.
- Познатите се веројатностите за почетни состојби.

Од равенствата на Чепмен-Колмогоров се добиваат следниве диференцијални равенки:

$$p_k'(t) = -\sum_{j \in R} \lambda_{kj}(t) p_k(t) + \sum_{i \in S} \lambda_{ik}(t) p_i(t),$$

R е множество индекси на состојби кон кои постои излезен лак од состојбата x_k, а S е множество индекси на состојби од кои постои влезен лак до состојбата x_k.

За ергодичен процес $\lambda_{ij}(t) = \lambda_{ij}$, односно интензитетите на премин не зависат од времето после доволно долг временски интервал.

Всушност, кај ергодичните процеси временскиот интервал се дели на два дела:

$(0, t_p)$ - интервал со нестационарен режим на работа и (t_p,∞) - интервал со стационарен режим на работа,.

Инаку, секој ергодичен процес е стационарен, а стационарен процес не мора да е ергодичен.

Интервалот (t_p,∞) се наоѓа така што за доволно мало ε се избира t_p за кое

$| p_j(t) - p_j| < \varepsilon$ кога $t \in (t_p,\infty)$.

За овој интервал (t_p,∞) системот преминува во систем линеарни равенки

$$0 = -\sum_{j\in R}\lambda_{kj}p_k + \sum_{i\in S}\lambda_{ik}p_i ,$$

$$\sum_{i\in K}p_i = 1.$$

6.2.3 Маркови Вериги

Случајниот процес $\{X_t,\ t \in T\}$ се нарекува Марков процес, ако за секој $n \in N$, за секои $t_1 < t_2 < ...< t_n$, $t_i \in T$, $i = 1,..., n$ и за секои $x_1, x_2,..., x_n \in R_X$, точно е равенството:

$$P\{X_{t_n} < x_n \mid X_{t_{n-1}} = x_{n-1}, X_{t_{n-2}} = x_{n-2},..., X_{t_1} < x_1,\} = P\{X_{t_n} < x_n \mid X_{t_{n-1}} = x_{n-1}\}.$$

Ако случајниот процес е дискретен, тогаш Марковото својство, добива облик:

$$P\{X_{t_n} = x_n \mid X_{t_{n-1}} = x_{n-1}, X_{t_{n-2}} = x_{n-2}, \ldots, X_{t_1} = x_1\} = P\{X_{t_n} = x_n \mid X_{t_{n-1}} = x_{n-1}\}.$$

Дискретен Марков процес со дискретно множество состојби и дискретно параметарско множество се нарекува *верига на Марков*, т.е.

Дефиниција 6.3 Дискретен случаен процес $X_1, X_2, \ldots$ се нарекува *верига на Марков*, ако за секој $n \in N$, и за секои $x_1, x_2,..., x_{n+1} \in R_X$, точно е равенството:

$$P\{X_{n+1}= x_{n+1}| X_n= x_n, X_{n-1}= x_{n-1},\ldots, X_1= x_1\}= P\{X_{n+1}= x_{n+1}| X_n= x_n\}.$$

Означуваме $P\{X_n= j| X_{n-1}= i\} = p_{ij}^{(n)}$. Ако $\{X_i\}$ е верига на Марков, тогаш X_n се нарекува состојба на системот во момент n.

Веројатностите $p_{ij}^{(n)}$ се нарекуваат преодни веројатности од состојба i во состојба j во n-тиот момент на промена. Овие веројатности ја формираат матрица од преодни веројатности $\boldsymbol{P}^{(n)} = [p_{ij}^{(n)}]$. Согласно, дефиницијата на Марков процес, ако процесот има s состојби, елементите на оваа матрица ги задоволуваат следните својства:

$$0 \le p_{ij}^{(n)} \le 1, \quad i, j = 1,2,\ldots,s,$$

$$\sum_{j=1}^{s}p_{ij}^{(n)} = 1, \quad i = 1,2,\ldots,s.$$

Нека $p_j^{(n)} = P\{X_n = j\}$ е веројатноста дека во n-тиот момент, системот ќе се најде во состојба j. Тогаш:

$$\sum_{j=1}^{s} p_{ij}^{(n)} = 1, \quad n = 1,2,...$$

Со тоа е определен векторот $\boldsymbol{p}^{(n)} = (p_1^{(n)}, p_2^{(n)},\dots, p_s^{(n)})$.

За $p_j^{(n)}$ добиваме: $p_j^{(n)} = P\{X_n = j\} = \sum_{i=1}^{s} P\{X_{n-1} = i\} P\{X_n = j \mid X_{n-1} = i\} = \sum_{i=1}^{s} p_i^{(n-1)} p_{ij}^{(n)}$,

или во матрична форма $\boldsymbol{p}^{(n)} = \boldsymbol{p}^{(n-1)} \boldsymbol{P}^{(n)}$, $n = 1,2,\dots$

Нека $\boldsymbol{p}^{(0)} = (p_1^{(0)}, p_2^{(0)},\dots, p_s^{(0)})$ е вектор на веројатности на почетната состојба на системот. Со последователна примена на $\boldsymbol{p}^{(n)} = \boldsymbol{p}^{(n-1)} \boldsymbol{P}^{(n)}$, за $n = 1,2,\dots$, се добива:

$$\boldsymbol{p}^{(n)} = \boldsymbol{p}^{(n-1)} \boldsymbol{P}^{(n)} = \boldsymbol{p}^{(n-2)} \boldsymbol{P}^{(n-1)} \boldsymbol{P}^{(n)} = \dots = \boldsymbol{p}^{(0)} \boldsymbol{P}^{(1)} \boldsymbol{P}^{(2)} \dots \boldsymbol{P}^{(n-1)} \boldsymbol{P}^{(n)}.$$

Дефиниција 6.4. Веригата на Марков се нарекува *хомогена* (или *временски инваријантна*), ако условните веројатности $p_{ij}^{(n)}$ не зависат од n, т.е. за $n = 1,2,\dots$

$$P\{X_{n+1}= j \mid X_n= i\} = P\{X_2= j \mid X_1= i\} = p_{ij}, \text{ за сите } i, j \in R_X$$

Веројатностите p_{ij} се нарекуваат преодни веројатности од состојба i во состојба j за еден чекор. Соодветната матрица се означува со $\boldsymbol{P} = [p_{ij}]$. Во овој случај, $\boldsymbol{p}^{(n)} = \boldsymbol{p}^{(0)} \boldsymbol{P}^n$.

Во хомогена верига на Марков, означуваме:

$$p_{ij}(n) = P\{X_n= j \mid X_0= i\}.$$

Тоа е веројатноста дека системот ќе помине од состојба i во состојба j за n чекори.

Согласно равенствата на Чепмен-Колмогоров за хомоген Марков процес, имаме:

$$p_{ij}(n+1) = \sum_{k=1}^{s} p_{ik}(n) p_{kj}, \quad i, j = 1, 2,\dots,s, \quad n = 1,2,\dots$$

Овој систем равенства може да се запише во матрична форма

$$\boldsymbol{P}(n+1) = \boldsymbol{P}(n) \cdot \boldsymbol{P}, \text{ за секој } n = 1,2,\dots$$

Оттука, се добива дека $\boldsymbol{P}(n) = \boldsymbol{P}^n$. Според ова, секоја конечна хомогена верига на Марков е определена, ако е познат векторот $p^{(0)}$ на веројатности на почетната состојба и матрицата $\boldsymbol{P} = [p_{ij}]$ на преодни веројатности за еден чекор.

Ако е можно со позитивна веројатност од произволна состојба на веригата да се стигне во која било друга состојба со конечен број чекори, тогаш велиме дека веригата е иредуцибилна.

Случаен процес е (строго) стационарен, ако заедничката распределба на било кое подмножество случајни променливи е инваријантна во однос на транслација на временскиот индекс, т.е. во дискретен случај

$$P\{X_1 = x_1, X_2 = x_2,\dots, X_n = x_n\} = P\{X_{1+l} = x_1, X_{2+l} = x_2,\dots, X_{n+l} = x_n\}.$$

Може да се покаже точноста на следната теорема:

Теорема 6.3. Конечна хомогена верига на Марков е строго стационарен процес, акко

$$\boldsymbol{p}^{(n)} = \boldsymbol{p}^{(0)},$$

за секој n = 1, 2, …

Значи, верига на Марков е стационарен процес, ако почетната состојба е распределена согласно стационарната распределба.

Бидејќи $\boldsymbol{p}^{(n)} = \boldsymbol{p}^{(n-1)}\boldsymbol{P}$, означувајќи $\boldsymbol{p}^{(n)} = \boldsymbol{p}^{*}$, за секој n = 1,2,…, равенството добива облик:

$$\boldsymbol{p}^{*} = \boldsymbol{p}^{*}\,\boldsymbol{P}.$$

Според Теорема 3.3, конечна хомогена верига на Марков е стационарна, ако векторот $\boldsymbol{p}^{(0)} = (p_1^{(0)}, p_2^{(0)},\ldots, p_s^{(0)})$ на веројатности на почетната состојба е решение на претходната равенка. Таа равенка е, всушност, хомоген систем од s линеарни равенки:

$$p_j^* = \sum_{k=1}^{s} p_k^* p_{kj} \qquad j = 1,2,\ldots s.$$

За да решението на системот биде распределба на веројатност, мора да биде задоволен условот:

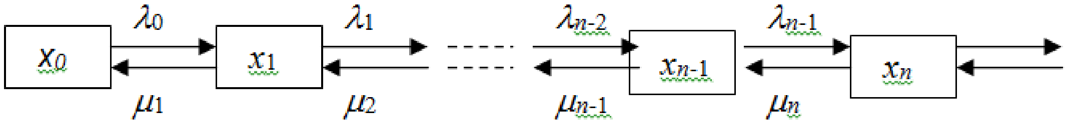
$$\sum_{j=1}^{s} p_j^* = 1.$$

6.3 Процеси на раѓање и умирање

Да се претпостави дека во било која точка од времето системот може да се наоѓа во една од конечно или преброиво многуте состојби x_1, x_2, x_3, ... Од некакви причини состојбата на системот се менува во текот на времето и интензитетот на премин од состојба x_n во состојба x_{n+1} е λ_n, а од состојба x_n во состојба x_{n-1} е μ_n. Веројатностите дека системот ќе премине од состојба x_n во состојба x_{n+k} и од x_n во состојба x_{n-k} за мал временски интервал, каде k > 1, се незначително мали. Од овде, веројатноста дека за мал временски интервал Δt системот ќе остане во состојба x_n е еднаква на $1 - \lambda_n\Delta t - \mu_n\Delta t$. Се претпоставува дека интензитетите на премин не зависат од времето

Овие случајни процеси се нарекуваат **процеси на раѓање и умирање.** Ако x_n го означува настанот дека една популација се состои од n единки, тогаш преминот $x_n \to x_{n+1}$ означува раѓање на нова единка, а преминот $x_n \to x_{n-1}$ означува умирање на една единка. Ако за секое $n \geq 1$ важи μ_n=0, т.е. само премини $x_n \to x_{n+1}$ се возможни, тогаш процесот е чист процес на раѓање, а ако за секое $n \geq 1$ важи $\lambda_n = 0$, т.е. само премини $x_n \to x_{n-1}$ се возможни, тогаш процесот е чист процес на умирање.

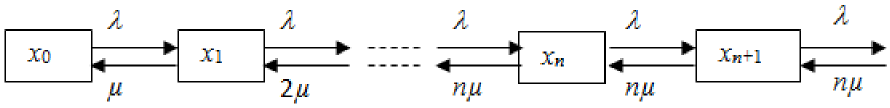
За да се решат системите диференцијални равенки потребно е да се знаат веројатностите на состојбите во време 0.



Фиг. 6.7 Процес на раѓање и умирање

6.3.1 Систем $M|M|n|\propto$

Системот за масовно опслужување $M|M|n|\propto$ може да се опише со марков процес како на Фигура 6.8



Фиг. 6.8 Систем $M|M|n|\propto$

Всушност, индексот на состојбата го покажува бројот на клиенти во системот; интензитетот на премин од една состојба во следна е еднаков на интензитетот на доаѓање на клиентите λ, а интензитетот на премин од една состојба во нејзината претходна е еднаков на интензитетот на заминување на клиентите, кој, ако бројот на клиенти k е помал од n, е еднаков на $k\mu$ (k сервери работат), а ако $k \geq n$ е еднаков на $n\mu$ (n сервери работат).

Значи:

$$\lambda_k = \lambda,$$
$$\mu_k = k\mu,\ k < n,$$
$$\mu_k = n\mu,\ k \geq n.$$

За овој систем се добиваат следниве системи диференцијални равенки:
Ако бројот на состојби е конечен (m):

$$p_0'(t) = -\lambda_0 p_0(t) + \mu_1 p_1(t)$$
$$\text{за } k < m: p_k'(t) = -(\lambda_k + \mu_k)\, \underline{p_k}(t) + \lambda_{k-1} p_{k-1}(t) + \mu_{k+1} p_{k+1}(t)$$
$$\text{за } k = m: p_m'(t) = \lambda_{m-1} p_{m-1}(t) - \mu_m p_m(t);$$

$$\sum_{k=0}^{m} p_k(t) = 1.$$

Ако бројот на состојби е бесконечен:

$$p_0'(t) = -\lambda_0 p_0(t) + \mu_1 p_1(t);$$
$$p_k'(t) = -(\lambda_k + \mu_k)\, p_k(t) + \lambda_{k-1} p_{k-1}(t) + \mu_{k+1} p_{k+1}(t)$$

$$\sum_{k=0}^{\infty} p_k(t) = 1$$

Претходниот пример е пример на прост процес на раѓање и умирање бидејќи интензитетите не зависат од времето. За простиот процес на раѓање и умирање важи дека е хомоген, транзитивен и ако е со конечен број на состојби, тогаш е и ергодичен. Се претпоставува дека системот е конечен (во пракса тоа значи дека нема бесконечно натрупување во редицата, што е исклучително важно); тогаш по истек на доволно голем временски интервал процесот влегува во стационарен режим на работа и системот диференцијални равенки преминува во систем обични линеарни равенки:

$$0 = -\lambda_0 p_0 + \mu_1 p_1$$
$$\text{за } k < m: 0 = -(\lambda_k + \mu_k) p_k + \lambda_{k-1} p_{k-1} + \mu_{k+1} p_{k+1}$$
$$\text{за } k = m: 0 = \lambda_{m-1} p_{m-1} - \mu_m p_m;$$

$$\sum_{k=0}^{m} p_k = 1$$

Начин за решавање на овие равенки: Се воведува замена $z_k = -\lambda_k\ p_k + \mu_{k+1} p_{k+1}$. Добиваме: $z_k = 0$, $k = 0,1,...,\ m-1$, односно $p_{k+1}=(\lambda_k\ /\mu_{k+1})\ p_k$ или, со последователна замена, за $k = 0,1,...m-1$

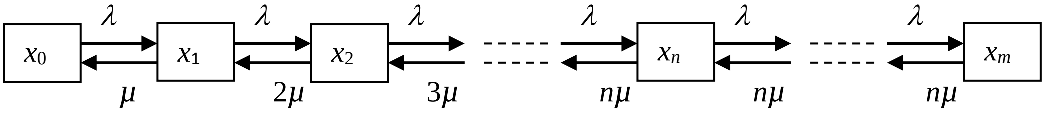
$$p_{k+1} = p_0 \prod_{i=0}^{k} \frac{\lambda_i}{\mu_{i+1}}$$

$$\sum_{k=0}^{m} p_k = p_0 + \sum_{k=1}^{m} p_k = p_0 + \sum_{k=0}^{m-1} p_0 \prod_{i=0}^{k} \frac{\lambda_i}{\mu_{i+1}} = 1$$

$$p_0 = \frac{1}{1 + \sum_{k=0}^{m-1} \prod_{i=0}^{k} \frac{\lambda_i}{\mu_{i+1}}}$$

6.3.2 Систем *M*|*M*|*n*|*m*

Се разгледува систем сличен на претходните кој се карактеризира со прост влезен и излезен поток на настани, *n* сервери и *m* места во системот(редицата на чекање е ограничена) . Значи, освен влезен поток на клиенти постои и поток на опслужени клиенти и поток на клиенти кои претрпеле отказ. Графот на состојби на системот е даден на Фигура 6.19.



Фиг. 6.9 Систем *M*|*M*|*n*|*m*

Ако во системот има најмногу *n* клиенти, тогаш тие веднаш се опслужуваат па потокот на опслужени клиенти е со интензитет $k\mu$, каде k е бројот на клиенти. Ако бројот на клиенти е поголем од бројот на сервери n, тогаш клиентите се опслужуваат со интензитет $n\mu$.

Системот диференцијални равенки гласи:

$$p_0'(t) = -\lambda p_0(t) + \mu p_1(t);$$

$$\text{за } k < n: \quad p_k'(t) = -(\lambda+k\mu)\, p_k(t) + \lambda\, p_{k-1}(t) + (k+1)\mu p_{k+1}(t)$$

$$\text{за } n \le k \le m: \quad p_k'(t) = -(\lambda+n\mu)\, p_k(t) + \lambda\, p_{k-1}(t) + n\mu\, p_{k+1}(t)$$

$$p_m'(t) = \lambda p_{m-1}(t) - n\mu p_m(t);$$

$$\sum_{k=0}^{m} p_k(t) = 1$$

Процесот е хомоген, со конечен број на состојби и е транзитивен, па е ергодичен. Ова значи дека може да се разгледува стационарен режим на работење. Тогаш, за доволно големо t, $p_k'(t) = 0$ и $p_k(t) = p_k$. Равенките се сведуваат на:

$$0 = -\lambda p_0 + \mu p_1;$$

$$\text{за } k < n: \quad 0 = -(\lambda+k\mu)p_k + \lambda p_{k-1} + (k+1)\mu p_{k+1}$$

$$\text{за } n \le k \le m: \quad 0 = -(\lambda+n\mu)p_k + \lambda p_{k-1} + n\mu p_{k+1}$$

$$0 = \lambda p_{m-1} - n\mu p_m;$$

$$\sum_{k=0}^{m} p_k = 1.$$

Со решавање на системот се добива:

$$p_k = \frac{\lambda}{n\mu} p_{k-1}, \quad n \le k \le m$$

$$p_k = \frac{\lambda}{k\mu} p_{k-1}, \quad 1 \le k < n$$

$$p_k = \frac{\lambda^k}{\mu^k k!} p_0 = \frac{\rho^k}{k!} p_0, \quad 1 \le k < n$$

$$p_k = \frac{\rho^{k-n}}{n^{k-n}} p_n = \frac{\rho^{k-n}}{n^{k-n}} \frac{\rho^n}{n!} p_0 = \frac{\rho^k}{n^{k-n} n!} p_0, \quad n \le k \le m.$$

$$\sum_{k=0}^{m} p_k = 1$$

Од последните равенки добиваме:

$$p_0 = \left(\sum_{k=0}^{n} \frac{\rho^k}{k!} + \frac{\rho^m}{n!} \sum_{k=1}^{m-n} \frac{\rho^k}{n^k} \right)^{-1}$$

$$p_0 = \left(\sum_{k=0}^{n} \frac{\rho^k}{k!} + \frac{\rho^m \rho}{n! n} \frac{1-(\rho/n)^{m-n}}{1-(\rho/n)} \right)^{-1}.$$

или:

Доколку ρ/n = 1 (што во случај на конечен број места во системот е можно и сумата е конечна), ќе важи:

$$p_0 = \left(\sum_{k=0}^{n} \frac{\rho^k}{k!} + \frac{\rho^m}{n!} (m-n) \right)^{-1}.$$

Веројатноста клиент што доаѓа во системот да добие отказ е еднаква на:

$$p_m = \frac{\dfrac{\rho^m}{n! n^{m-n}}}{\displaystyle\sum_{k=0}^{n} \frac{\rho^k}{k!} + \frac{\rho^{n+1}}{n!} \frac{1-(\rho/n)^{m-n}}{n-\rho}} = \frac{\rho^m}{n! n^{m-n}} p_0.$$

Оттука, веројатноста дека клиентот ќе биде опслужен е еднаква на

$$p_{opsl} = P(k < m) = 1 - p_m.$$

Веројатноста дека се зафатени сите канали е еднаква на:

$$P(k \geq n) = \sum_{k=n}^{m} p_k = \sum_{k=n}^{m} \frac{\rho^n}{n!} p_0 \frac{\rho^{k-n}}{n^{k-n}} = \frac{p_0}{n!} \rho^n \sum_{k=n}^{m} \frac{\rho^{k-n}}{n^{k-n}} = p_n \frac{1-(\rho/n)^{m-n+1}}{1-(\rho/n)}.$$

Ако ρ/n =1 тогаш $P(k \geq n) = p_n(m-n+1)$.

Останува уште да ги разгледаме слевните параметри.
Среден број клиенти во редицата:

$$L_Q = \sum_{k=n}^{m} (k-n) p_k = p_n \sum_{k=n}^{m} (k-n) \frac{\rho^{k-n}}{n^{k-n}}.$$

со интегрирање член по член , и сумирање се добива:

$$L_Q = p_n \frac{\rho/n}{(1-\rho/n)^2} \left[1-(m-n+1)(\rho/n)^{m-n} + (m-n)(\rho/n)^{m-n+1}\right]$$

Ако ρ/n=1 тогаш

$$L_Q = p_n \frac{(m-n+1)(m-n)}{2}.$$

Среден број на зафатени канали:

$$\begin{aligned}\bar{p} &= \sum_{k=1}^{n} k p_k + \sum_{k=n+1}^{m} n p_k = \sum_{k=1}^{n} k \frac{\rho}{k} p_{k-1} + \sum_{k=n+1}^{m} n \frac{\rho}{n} p_{k-1} = \\ &= \sum_{k=1}^{n} \rho p_{k-1} + \sum_{k=n+1}^{m} \rho p_{k-1} = \rho \sum_{k=0}^{n-1} p_k + \rho \sum_{k=n}^{m-1} p_k = \rho \sum_{k=0}^{m-1} p_k = \rho(1-p_m).\end{aligned}$$

Среден број на слободни канали:

$$\bar{s} = n - \bar{p}.$$

Среден број на клиенти во системот:

$$L = \sum_{k=0}^{m} k p_k = \sum_{k=0}^{n} k \frac{\rho^k}{k!} p_0 + \sum_{k=n+1}^{m} k \frac{\rho^k}{n! n^{k-n}} p_0.$$

На крај да ја најдеме и распределбата на веројатноста на време на чекање. Се претпоставува дека клиентот што пристигнува не добива отказ. Тогаш:

$$P(W_Q > t) = \sum_{k=n}^{m-1} p_k P_k(W_q > t) = \sum_{k=n}^{m-1} p_k \sum_{s=0}^{k-n} q_s(t)$$

$$q_s(t) = \frac{(n\mu t)^s}{s!} e^{-n\mu t}$$

веројатноста дека во интервал t, s клиенти ќе завршат со опслужување.

За да се упрости изразот за веројатноста, се менува редоследот на сумирање:

$$P(W_Q > t) = \sum_{s=0}^{m-1-n} \sum_{k=s+n}^{m-1} p_k q_s(t) = \sum_{s=0}^{m-1-n} \sum_{j=s}^{m-1-n} q_s(t) p_{n+j} = \sum_{s=0}^{m-1-n} \frac{(n\nu t)^s}{s!} e^{-n\nu t} \sum_{j=s}^{m-1-n} \left(\frac{\rho}{n}\right)^j p_n =$$

$$= \frac{p_n e^{-n\nu t}}{1-\rho/n} \sum_{s=0}^{m-1-n} \frac{(n\nu t)^s}{s!} \left(\frac{\rho}{n}\right)^s \left(1-(\rho/n)^{m-n-s}\right) = \frac{p_n e^{-n\nu t}}{1-\rho/n} \sum_{s=0}^{m-1-n} \frac{(n\nu t)^s}{s!} \left((\rho/n)^s - (\rho/n)^{m-n}\right) (*)$$

$$P(W_Q < t) = 1 - P(W_q > t).$$

Равенката (*) не може да се сумира конечно. Посебно се разгледува случајот кога $m = n$. Поедноставен израз се добива за математичкото очекување на времето на чекање:

$$W = E(W_Q) = \int_0^\infty t \cdot dP(W_Q < t) = \int_0^\infty [t \cdot \frac{p_n n\mu e^{-n\mu t}}{1-\rho/n} \sum_{s=0}^{m-1-n} \frac{(n\mu t)^s}{s!} \left((\rho/n)^s - (\rho/n)^{m-n}\right) - \frac{p_n e^{-n\mu t}}{1-\rho/n} \cdot$$

$$\cdot \sum_{s=0}^{m-1-n} \frac{(nt)^s}{s!} s t^{s-1} \left((\rho/n)^s - (\rho/n)^{m-n}\right)] dt$$

Заменувајќи го редоследот на сумирање и интегрирање и интегрирајќи добиваме:

$$W_Q = \frac{p_n}{1-\rho/n} \sum_{s=0}^{m-n-1} \left((\rho/n)^s - (\rho/n)^{m-n}\right) = \frac{p_n}{n\mu(1-\rho/n)^2} \left(1 - (m-n-1)\cdot(\rho/n)^{m-n} + (m-n)(\rho/n)^{m-n-1}\right)$$

Важи:

$$L_Q = \lambda W_Q \qquad W_Q / L_Q = \frac{1/(n\mu)}{\rho/n} = \frac{1}{\lambda};$$

Од овде следи дека средно време на престој во системот:

$$W = \frac{L}{\lambda} = \frac{L_Q}{\lambda} + \frac{\bar{p}}{\lambda} = W_Q + \frac{\rho(1-p_m)}{\lambda} = W_Q + \frac{(1-p_m)}{\mu}.$$

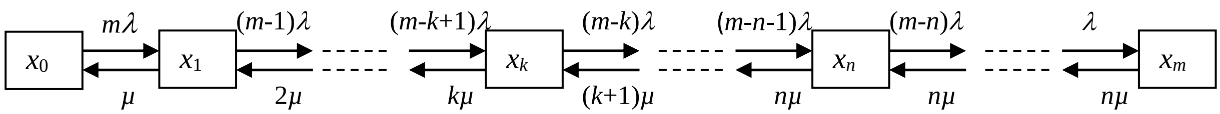
6.3.3 Затворен систем за масовно опслужување

Пример 6.3 Се разгледува возен парк на автомобили кои се расипуваат во тек на време. Бројот на автомобили е фиксен (не доаѓаат нови автомобили). Постојат и сервисери за поправка на автомобилите. Ако бројот на расипани автомобили е поголем од бројот на сервисери тогаш некои автомобили чекаат во редицата за опслужување. Значи, во овој систем бројот на канали е еднаков на бројот на сервисери, а бројот на клиенти е еднаков на бројот на автомобили.
Со $h(t)$ ќе го означуваме бројот на исправни автомобили, со $r(t)$ бројот на автомобили во редицата на чекање, а со $p(t)$ бројот на автомобили што се поправаат.

Ако m е бројот на автомобили, тогаш $m = h(t) + r(t) + p(t)$. Ако системот е ергодичен, тогаш

$$m = \bar{r} + \bar{p} + \bar{h}.$$

Состојбите можат да се опишат на начин прикажан на Фигура 6.10 (x_k - број на откажани возила, n- канали за опслужување)



Фиг. 6.10 Затворен систем за масовно опслужување

Системот е транзитивен со конечен број на состојби, од каде следува дека е ергодичен. Во стационарен режим на работа решенија на системот се:

$$\begin{cases} p_k = \dfrac{m-(k-1)}{k}\rho \cdot p_{k-1}, & 1 \le k < n \\ p_k = \dfrac{m-(k-1)}{n}\rho \cdot p_{k-1}, & n \le k \le m \end{cases} \qquad \begin{cases} p_k = \dbinom{m}{k}\rho^k p_0, & 1 \le k < n \\ p_k = \dbinom{m}{k}\dfrac{k!}{n!n^{k-n}}\rho^k p_0, & n \le k \le m \end{cases}$$

$$\sum_{k=0}^{n} p_k + \sum_{k=n+1}^{m} p_k = 1.$$

Од претходните равенки следи:

$$p_0 = \left(\sum_{k=0}^{n} \binom{m}{k}\rho^k + \sum_{k=n+1}^{m} \binom{m}{k}\frac{k!}{n!n^{k-n}}\rho^k \right)^{-1},$$

Математичко очекување за број на зафатени канали е:

$$\overline{p} = \sum_{k=0}^{n} kp_k + \sum_{n=n+1}^{m} np_k$$

$$\overline{r} = \sum_{k=n+1}^{m} (k-n)p_k;$$

$$\overline{s} = \sum_{k=0}^{n-1} (n-k)p_k$$

Веројатноста дека клиент ќе чека во редица:

$$P(k \ge n) = \sum_{k=n}^{m} p_k.$$

Коефициент на искористување на системот: се пресметува средниот број на нерасипани елементи во однос на вкупниот број елементи:

$$\eta = \frac{\overline{h}}{m} = \frac{m - \overline{r} - \overline{p}}{m} = 1 - \frac{\overline{r} + \overline{p}}{m}.$$

КОРИСТЕНА ЛИТЕРАТУРА

[1] A.O. Allen, “Probability, Statistics and Queueing Theory with Computer Science Applications”, Accademic Press, 1978.

[2] Donald Gross, John F. Shortle, James M. Thompson, Carl M. Harris,” Fundamentals of Queueing Theory”, 4th edition, Wiley Series in Probability and Statistics, 2008

[3] George S. Fishman, “Discrete-Event Simulation: Modeling, Programming, and Analysis**”,** Springer, Series in Operations Research,2001.

[4] Jerry Banks, John S. Carson II, Barry L. Nelson, David M. Nicol ,” Discrete-Event System Simulation”, 5th Edition, Prentice Hall, 2009.

[5] Магдалена Георгиева, Верица Бакева “Случајни процеси – Предавање и вежби”, 2007, Скопје.

[6] Sheldon M. Ross, ” Introduction to Probability Models ,”, 6th edition, ACADEMIC PRESS, 1997.

[7] Taylor M. Howard, Karlin Samuel , “An Introduction to Stochastic Modeling ”, ACADEMIC PRESS, 1998.

БИОГРАФСКИ ПОДАТОЦИ

Благој Делипетрев е продекан и доцент на Факултетот за Информатика, Универзитет Гоце Делчев во Штип. Докторирал на Факултетот за Електротехника и Информатички Технологии, Универзитет Св Кирил и Методиј во Скопје во 2011 година, и магистрирал на истиот факултет во 2007 година. Неговите истражувања се во областите: нови алгоритми за оптимизација, симулација и моделирање, машинско учење и развивање на Облак апликации. Д-р. Делипетрев има објавено повеќе трудови во престижни интернационални списанија и учествува во неколку меѓународни и домашни проекти.

Наташа Максимова е лаборант на Факултетот за Информатика, на Универзитетот Гоце Делчев во Штип. Дипломира на Природно-математичкиот факултет на Универзитетот Св Кирил и Методиј во Скопје каде и магистрира во 2009 година. Нејзини области на истражување се теорија на графови и моделирање и симулации. М-р Максимова има објавено повеќе трудови во меѓународни списанија и конференции.

Зоран Утковски е доцент на Факултетот за информатика на Универзитетот Гоце Делчев во Штип и раководител на Катедрата за комуникациски технологии и процесирање на сигнали. Дипломира во 2000 година на Електротехничкиот Факултет, на Универзитетот “Св. Кирил и Методиј” во Скопје, на насоката Електроника и Телекомуникации. Магистрира (со дистинкција) во 2004 година на Техничкиот Универзитет Чалмерс во Шведска. Докторира (со дистинкција) на Универзитетот во Улм во 2010. Д-р. Утковски е добитник е на Фулбрајтовата стипендија и на ДААД стипендијата за Академска размена. Неговиот галавен истражувачки интерес е во областите на безжичните комуникации и процесирањето на сигнали.

ISBN 978-608-244-250-1