D102

ДИССЕРТАЦИЯ

СПЕЦИАЛИЗИРАН НАУЧЕН СЪВЕТ ПО
ИНФОРМАТИКА И ПРИЛОЖНА МАТЕМАТИКА

Веска Пенчева Нончева

Един подход за изграждане на интелигентен статистически софтуер

ДИСЕРТАЦИЯ

за присъждане на образователна и научна степен
"Доктор"

Научна специалност: 01.01.13

Научни консултанти:

доц. д-р Димитър Вълдев
доц. д-р Станимир Стоянов

2000 г.

Съдържание:

Увод *3*

Глава 1. ***Експертни системи в статистиката***
1.1 Съветващи системи *7*
1.2 Системи за статистически анализ *8*
1.3 Празни експертни системи *10*
1.4 Генериране на текст и графика *11*
1.5 Извличане на знания от данни *11*

Глава 2. ***Мултиагентни системи***
2.1 Агенти *13*
2.2 Подходи за създаване на мултиагентни системи *14*
2.3 FORESEE - мултиагентна система за анализ и прогнозиране на временни редове *16*
2.4 Мултиагентна система за статистически изследвания . . . *18*

Глава 3. ***Концептуален модел на агент за проверка на статистически хипотези***
3.1 Бейсова мрежа *22*
3.2 Предназначение на агента *24*
3.3 Модел на поведението на агента *25*
3.4 База знания *26*
3.4.1 Общи положения *26*
3.4.2 Описание *29*
3.5 Автоматични изводи *33*
3.6 Представяне на външни въздействия *39*

Глава 4. ***Математически модел на вярата***
4.1 Нормалност при предположение за симетричност *41*
4.1.1 Постановка на задачата *41*
4.1.2 Модел *46*
4.1.3 Функция полезност *49*
4.1.4 Достатъчни условия за монотонност на слабите правила . *53*
4.1.5 Достатъчно условие за силна монотонност *63*
4.2 Симетричност *66*
4.3 Комбиниран модел за нормалност *68*
4.4 Обобщен модел *75*

Глава 5. ***Структури от данни***
5.1 Мрежа за вземане на решения *85*
5.2 ***U***-мрежа *88*
5.2.1 Дърво на събитията *88*
5.2.2 Фактори *90*

Глава 6. ***Теоретични основи на алгоритъма за оптимизиране на вярата***
6.1 Алгоритъм *91*
6.2 Основна теорема *107*
6.3 Пример *118*
6.4 Алгоритъм за генериране на достатъчни условия за монотонност *122*
6.5 Извод *129*

Глава 7. ***Алтернативни симулационни подходи***
7.1 Алгоритъм с полагане *130*
7.2 Алгоритъм на Гибс *132*
7.3 Неитеративен симулационен алгоритъм *133*
7.4 Приложения *135*

Заключение *137*

Литература *139*

Публикации на дисертанта по дисертацията *147*

Intelligence, as is well known, is easier to recognize,
than to define.
Shlomo Zilberstein,
Operational rationality through compilation of anytime algorithms

Увод

Доскоро се смяташе, че такива проблеми като доказване на теореми, разпознаване на образи и реч, игра на шах, прогнозиране поведението на сложни системи могат да бъдат решени само от човек, защото тяхното формулиране и решаване изискват способности, присъщи само на човека - способност да мисли, да наблюдава, да разговаря, да се научава и други. Но значимите научни изследвания през последните три десетилетия, направени от изследователи от различни области на науката, показаха, че много от тези сложни детерминистични и стохастични проблеми могат да бъдат формулирани и решени от компютъра. С такива проблеми, които изглеждат невъзможни за директно формализиране, се занимава тази част от науката Информатика, наречена Изкуствен интелект.

На въпроса "Какво могат да правят интелигентните системи?" Уинстон, който се счита за човека, положил основите на Изкуствения интелект, дава следния отговор:

- да помагат на експертите при решаване на сложни проблеми,
- да се самообучават,
- да разбират и да генерират текст на естествен език [*Winston, 93*].

Развитието на глобалните компютърни комуникации оказа голямо влияние върху развитието на Изкуствения интелект. Докато процесът на преминаване от централизирани към разпределени бази данни протича сравнително гладко, то за децентрализирането на интелигентните системи съществуват множество пречки. Основната причина се крие в същността на информацията, с която оперират тези системи. От една страна степента на семантична зависимост в знанията е много по-голяма в сравнение с тази в структурите на една база данни. От друга страна често при знанията се касае за полуструктурирана и неструктурирана информация.

Усилията за намиране на адекватни и ефективни решения за децентрализация на базите знания формират един нов подход на

Изкуствения интелект - подход, базиращ се на агенти и многоагентни системи. Този нов подход се радва на голям изследователски интерес през последните 6-7 години. Публикувани са забележителни работи, в които са дискутирани основните и желателните характеристики на мултиагентните системи ([*Russell,. Norving, 95*], [*Wooldridge, Jennings,95б*]), както и тяхната архитектура ([*Wooldridge, Jennings,95*]).

В процеса на развитие на статистическия софтуер се наблюдават следните етапи:

- *пакети статистически програми*, реализиращи изчислителни процедури;
- *статистически експертни системи*, притежаващи статистически знания и умения;
- *агенти за решаване на статистически задачи*, притежаващи статистически знания и интелигентно поведение;
- *мултиагентни системи за статистически изследвания*.

Развитието на глобалните комуникации наложи изискването интелигентните статистически системи да притежават освен статистическа експертиза и интелигентно поведение.

В дисертацията се атакува следният двустранен проблем: Какво може статистическият софтуер да спечели от методите и средствата на Изкуствения интелект и какво може стохастиката да предложи на системите с изкуствен интелект.

В работата се разглежда проблемът за изграждане на съвременен статистически софтуер и се предлага един начин за решаването му със средствата на Изкуствения интелект и по-специално чрез мултиагентни системи. Статистическият софтуер получава нова функционалност и качествено нови свойства, които разширяват използваемостта му и повишават неговата гъвкавост и мобилност.

В настоящата работа се разглежда и проблемът на Изкуствения интелект за моделиране на вярата и се предлага един начин за решаването му с методите на стохастиката и по-специално с Бейсовия подход.

В *Глава 1* са разгледани някои експертни системи в статистиката, техните характерни особености и степента им на завършеност, съгласно публикуваните резултати. В §*1.1* са дискутирани съветващите системи, в §*1.2* - системите, извършващи статистически анализ, в §*1.3* - "празните" експертни системи в статистиката. В §*1.4* е разгледан въпросът за генериране на текст и графика. В §*1.5* е представена една нова научна област, занимаваща се с извличане на знания от данни.

Глава 2 съдържа кратко представяне на мултиагентните системи. В §*2.1* е дадено неформално определение на понятието агент. В §*2.2* са представени top_down и bottom_up подходите за създаване на мултиагентни системи. В §*2.3* е дискутирана мултиагентната система за анализ и прогнозиране на временни редове FORESEE. В §*2.4* е предложена архитектура на Мултиагентна система за проверка на статистически хипотези.

В *Глава 3* е представен концептуален модел на агент за проверка на статистически хипотези. В §*3.1* е дискутиран един важен клас от вероятностни модели, които могат да се представят чрез ориентирани графи, известен като Бейсова мрежа. В §*3.2* е мотивирана необходимостта от изграждането на агент за проверка на статистически хипотези. В §*3.3* е мотивиран подходът за представяне на поведението на агента чрез Бейсова мрежа. В §*3.4* е представен модел на Базата статистически знания. В §*3.5* се обсъжда алгоритъм за автоматични изводи. В §*3.6* се дискутира един подход за представяне на външни въздействия в Бейсова мрежа.

В *Глава 4* е представен математически модел на вярата на агента по отношение вида на вероятностното разпределение на изследваната популация. Вярата се измерва с реално число в интервала [0,1], наречено сигурност. Казваме, че агентът вярва във верността на дадено твърдение, ако сигурността е по-голяма от известно, предварително фиксирано число. Т.е. вярата на агента е прагова функция на сигурността с множество от стойности - множеството от състоянията на вярата. В §*4.1* е разгледан модел на вярата за нормалност на изследваната популация при предположение за симетричност. В *4.1.1* е дадена постановката на задачата. В *4.1.2* е описан подходът към задачата за оптимизиране на вярата на агента. Този подход взема под внимание всяко априорно знание и е едно приложение на Бейсовия подход. Дефинирани са слаби и силни монотонни правила за вземане на решения за състоянието на вярата. Допуска се популацията на агентите потребители да бъде разбита на няколко подпопулации в зависимост от тяхната статистическа компетентност. В *4.1.3* са дефинирани две структури полезност - Гаусова и complementary log-log структура полезност, определени чрез съответните функции на разпределение. Функциите на разпределение притежават ред предимства, когато с тях се определят структури полезност:

1. Известни са голям брой многопараметрични семейства разпределения.
2. Получените за тях резултати могат да се използват и тук.

В *4.1.4* са доказани достатъчни условия за монотонност на слабите правила за вземане на решения. В *4.1.5* е изведено достатъчно условие за силна монотонност.

В §*4.2* е представен модел на вярата за симетричност на непрекъсната популация.

В §*4.3* е разгледан модел на вярата за нормалност при предположение за непрекъснатост.

В §*4.4* е даден обобщен модел на вярата на агента за вида на вероятностното разпределение на изследваната популация.

В *Глава 5* са представени структури от данни. В §*5.1* е дефинирана мрежа за вземане на решение за състоянието на вярата на агента. Представен е и алгоритъм за извод в нея. В §*5.2* е въведено понятието *U*-мрежа. В *5.2.1* е дефинирана структурата дърво на събитията. В *5.2.2* е дефинирано понятието фактор.

В *Глава 6* са дадени теоретичните основи на алгоритъма за оптимизиране на вярата на агента във вероятностния модел. В §*6.1* е представен алгоритъмът за оптимизиране на вярата. В §*6.2* е доказана теоремата, гарантираща верността на представения алгоритъм. В §*6.3* е разгледан пример. В §*6.4* е представен алгоритъм за генериране на достатъчни условия за монотонност в символен вид. Чрез три теореми са представени теоретичните основи на този алгоритъм. В §*6.5* е направен изводът, че проблемът за определяне на вярата на агента във вероятностния модел може да се представи чрез наредената тройка $(\mathcal{B}, \mathcal{U}, \mathcal{D})$, където $\mathcal{B}$ е Бейсова мрежа, представяща вероятностната структура на проблема, $\mathcal{U}$ е *U*-мрежа, представяща структурата полезност, $\mathcal{D}$ е мрежа за вземане на решение.

В *Глава 7* са разгледани алтернативни симулационни подходи. В §*7.1* е представен Substitution algorithm. В §*7.2* е разгледан Gibbs' sampling algorithm. В §*7.3* - Importance sampling algorithm. Някои приложения и програмни реализации на тези симулационни алгоритми са разгледани в §*7.4*.

В работата е приета следната система за номериране: Разделите (параграфите) се номерират последователно в рамките на главата. Определенията, лемите, теоремите, фигурите, таблиците и примерите също се номерират последователно в рамките на главата.

Глава 1

Експертни системи в статистиката

Expert Systems are machines that think and reason as an expert would in a particular domain. ... A true Expert System not only performs the traditional computer functions of handling large amounts of data, but it also manipulates that data so the output is a meaningful answer to a less than fully specified question.

L. Stevens,
Artificial Intelligence. The Search for the Perfect Machine

Статистическите експертни системи се развиха за да подпомогнат специалистите при извършване на сложни статистически анализи. Някои основни принципи при изграждането им са разгледани в [*Hand,85*], [*Hand,90*], [*Siriopoulos,87*], [*Gale,86*], [*Wittkowski,88*], [*Wittkowski,90*], [*Molenaar,88б*], [*Nelder,88*], [*Noncheva,95*].

Ще направим преглед на статистическите експертни системи като наблегнем на новостите, реализирани във всяка от тях. За всяка от разглежданите системи ще опишем класа от задачи, които тя решава, кръгът от потребители, за които е предназначена, нейният тип (съветваща, реализираща статистически анализ, shell) и вида на конкретната реализация (експериментална версия, прототип на статистическа експертна система, комерсиална версия). При това класификацията, която тук използваме, няма претенциите да бъде нито изчерпателна, нито пък абсолютна - конкретна система може да попадне в няколко категории.

1.1 Съветващи системи

С нарастване на броя на статистическите пакети и увеличаване на обема на всеки от тях възникна необходимостта от създаване на интелигентни програмни системи, съветващи как лесно и ефективно да се използват огромните библиотеки от подпрограми. Този клас съветващи експертни системи са известни като supervising expert systems. В [*Mikulesky,Kalas,Kelemen,86*] се разглеждат някои възможности за представяне на знанията в тези системи чрез фреймова структура.

Експертната система P-ASPAM, представена в [*Mikulesky,Chovatova,87*], е проектирана да дава съвет как лесно и

ефективно да се използва огромната библиотека NAG. NAG (Numerical Algorithms Group) е библиотека от подпрограми, решаващи широк кръг математически и статистически задачи. Системата P-ASPAM е проектирана на две нива. На първото ниво знанията за методите и подпрограмите, реализиращи съответните методи, са трансформирани в клаузи на Хорн с цел създаване на ядрото на базата знания. На второто ниво базата знания е разширена чрез добавяне на евристични метаправила, отразяващи опита и знанията на специалистите при използване на методите и подпрограмите. Тази експертна система е реализирана на Prolog, версия NU7 Edinburg, на PDP-11 съвместими компютри. Тя е съвместна разработка на Института по информатика в Братислава и Факултета по математика на Братиславския университет.

1.2 Системи за статистически анализ

Исторически първата система, демонстрираща използването на техниките на изкуствения интелект при автоматичното прилагане на избран статистически метод е REX (Regresion Expert). Системата е създадена от Daril Pregibon и William Gale и е описана в [*Pregibon,Gale,84*]. REX съветва потребителя при използването на регресионен анализ. Тя ръководи статистическия анализ като проверява предположенията на регресионния модел, препоръчва подходящи трансформации над данните, обяснява. REX е предназначена за потребители, имащи желанието да си изяснят статистическите концепции и термини. Системата приключва анализа с доклад за данните и резултатите. Реализирана е на VAX. REX използва S като статистическа програмна система, която извършва пресмятания и чертае графики.

Авторите на TESS (Tree Environment for Statistical Strategy) са се опитали да програмират интуицията на човека експерт.

GLIMSE е надстройка на широко използвания пакет GLIM (General Linear modelling). Целта е пакетът GLIM да се направи достъпен за повече потребители. Разработчиците на системата са потърсили решение на проблема "Как да се създаде система, от която начинаещият потребител да се научи и постепенно да стане добър потребител и която да задържи добрите потребители?". GLIMSE притежава команден език за специалиста - статистик и помощен режим на три нива, подпомагащ използването му.

Системата MULREG е лингвистичен препроцесор на произволен статистически пакет. В нея са реализирани линейни модели за

многомерни променливи с възможности за стъпков избор на променливи, робастни оценки, диагностициране на остатъците, трансформации на променливите, проверка на предположенията за разпределението, анализ на категорийни променливи, небалансирани експерименти и други. Чрез осветяване MULREG се опитва да подсказва на потребителя най-добрия избор, който той би могъл да направи. На всяка стъпка, когато от потребителя се изисква да вземе решение, MULREG предлага подходящо подбрано графично представяне на данните.

Системата STATPATH [*Portier, Lai, 83*] използва внимателно подредени въпроси за да реализира търсене по двойчното дърво на известните й методи. Базата знания е представена чрез множество от продукционни правила. Тази система представя най-добрият подход за избор на метод в областта на анализа на данни [*Gale, 86a*]. Ключовият проблем тук е, че когато потребителят няма достатъчно знания по статистика за да избере подходящия метод, той може би не разбира достатъчно статистиката за да отговори на въпросите правилно. В STATPATH са реализирани и различни други механизми за помощ. Може да бъде получена допълнителна информация за въпроса. Всеки въпрос може да бъде заменен със списък от подходящи въпроси.

Друг пример за система, демонстрираща възможността за избор на техника на базата на получената от потребителя - нестатистик информация, е MUSE [*Dambroise, 86*].

Системата RX ([*Blum, 82*], [*Gale, 86a*]) е предназначена автоматично да открие причинните връзки. Внедрена в Stanford Rheumatology Clinic RX симулира работата на учения, като предлага вида на зависимостта, която ще бъде изследвана, формулира проблема в термините на статистическия анализ на данни и изпълнява необходимия анализ. RX е интегрирана със статистическия пакет IDL и използва само методите на регресионния анализ.

Тази експертна система транслира целите на изследователя на езика на статистиката. Основният проблем е представянето на статистическите знания и знанията от приложната област. В RX и знанията от статистиката и знанията от медицината са представени чрез семантични мрежи.

RX издига идеята за експертна система, в която са представени всички знания, необходими за прилагане на статистическите методи, а знанията от приложната област се добавят от съответните експерти.

Друг успешен прототип на експертна система е WAMASTEX [*Dodra,90*]. Внедрена е в болницата на Виенския университет. Системата извършва анализ на данни с цел отстраняване на

съмнителни наблюдения, определяне на модалността на разпределението, избор на подходящо множество от описателни статистики, графично, числово и лингвистично представяне на резултатите. Системата WAMASTEX използва системата SAS за статистически пресмятания.

1.3 Празни експертни системи

THESEUS е прототип на празна експертна система, състояща се от редактор на правила и интерпретатор на правила. Предназначена е за използване от специалисти статистици с цел създаване на съветващи експертни системи, подпомагащи специалистите, експерти в своята приложна област, но с недостатъчни знания в областта на статистиката.

TAXSY е друг прототип на празна съветваща експертна система, използваща системата SAS. Построена на основата на SAS, TAXSY също има пълен достъп до всички мощни пресмятания, които SAS може да извърши.

Една възможност за организиране на интелигентно асистиране е предложена в CALCAS [*Димитров,92*]. CALCAS е българска разработка на празна експертна система. В CALCAS се поддържа дървовидна структура, в която всеки възел е един проблем и така е организирано интелигентно асистиране.

STUDENT [*Gale,90*] е проектирана така, че да позволи на специалиста - статистик да построи съветваща експертна система в областта на анализа на данни. STUDENT е изградена на основата на съществуващата среда за статистически анализ QPE (Quantitative Programming Environment). Единственото изискване при използването на STUDENT е статистиците добре да познават QPE и нейния език. Тази експертна система притежава способността да придобива нови знания чрез интервю и демонстрации. Статистическите знания в STUDENT са представени чрез семантични мрежи. За да построи една стратегия за анализ на данни STUDENT използва различни техники за самообучение.

Основната цел на този вид софтуер е да предостави на специалиста - статистик възможността сам да построи експертна система, предназначена за потребители, експерти в определена област на приложенията, но с малък опит в статистическите методи.

1.4 Генериране на текст и графика

PostGraphe е прототип на система за автоматично генериране на доклади, съдържащи текст и графика, които са резултат от направената описателна статистика на данните ([*Fasciano, Lapalme, 98*], [*Corio, Lapalme, 98*], [*Fasciano, 96*]). От таблично представените данни за които е известен типа на данните във всяка колонка и от множество от намерения на автора системата генерира доклад на LaTEX с графични файлове в Post Script формат. В тези доклади текстовото и графичното представяне на информацията взаимно се допълват. PostGraphe е реализирана на Prolog. Идеите на PostGraphe са адаптирани в най-новата версия на Microsoft Excel като Postgraphe Excel Wizard.

Полезността на тези разработки в областта на интелигентния статистически софтуер е несъмнена. Натрупан е богат изследователски опит. Но обща архитектура на статистическите експертни системи не е предложена.

В много статии ([*Van den Berg,90*], [*Molenaar,88*] и други) се описват опити да се построи множество от евристики, описващи статистическата стратегия на експерта-статистик. Но необходимо ли е да се следва механизма на човешкото разсъждение, след като съвременните компютри притежават огромни изчислителни възможности?

1.5 Извличане на знания от данни

The statistical literature has a wealth of technical procedures and results to offer data mining, but it also has a few simple methodological morals: prove that estimation and search procedures used in data mining are consistent under conditions reasonably thought to apply in applications; use and reveal uncertainty, don't hide it; calibrate the errors of search, both for honesty and to take advantages of model averaging...

C. Glymour, D. Madigan , D. Pregibon, P. Smyth
Statistical themes and lessons for data mining

Извличане на знания от данни (Data mining, DM) е съвременна наука, която се строи основно върху постиженията на статистиката и информатиката за да реши проблема за извличане на знания от огромни бази данни.

Основните задачи на DM са:

- Представяне на данни и на знания
- Редуциране на данни
- Моделиране на структурирана, текстова и мултимедиина информация
- Управление на несигурността
- Търсене на метрики за полезността на откритите знания
- Обясняване на откритите нови знания
- Статистика над огромни множества от данни.

Методите на DM включват класификация, клъстаризация, стохастично моделиране, прогнозиране и оценяване, анализ на зависимости, търсене и оптимизация, паралелни и разпределени техники, агенти.

Статистиката дава елегантни решения на някои от проблемите на DM. Трите основни въпроса от статистиката, които са от особена важност за DM са [*Glymour, Madigan , Pregibon, Smyth, 97*]:

- Яснота на целите

  Някои от целите на анализа на данни са да се намери вида на разпределението на данните от една база данни, да се предскажат основните черти на новите случаи в една динамична ситуация, да се планират следващите етапи на статистическия анализ.
- Оценка на сигурността
- Правилно определяне на източниците на случайност.

Съвременната компютърна техника активно подпомага анализа на данни.

На практика статистиката е необходима, но не и достатъчна компонента на DM. Тази нова наука се строи и върху техниките и теориите на много други области - бази данни, разпознаване на образи, самообучение, визуализиране на данни, моделиране несигурността, оптимизации.

Глава 2

Мултиагентни системи

Artificial intelligence is the discipline that is concerned with programming computers to do clever, humanoid things - but not necessarily to do them in a humanoid way.

Herbert Simon,
Models of bounded rationality

Идеята на мултиагентните системи е даден сложен изследователски проблем да се разбие на подпроблеми, които могат да се решат от специализирани агенти, които могат интелигентно да комуникират.

Концепцията на мултиагентните системи за първи път е предложена от Marvin Minsky [*Minsky, 85*]. Изследователите в MIT са едни от първите разработчици на мултиагентни системи.

2.1 Агенти

Под *агент* в най-общия случай се разбира програмна система, която притежава следните основни свойства ([*Wooldridge, Jennings,95б*], [*Russell, Norving, 95*]):

- *автономност* - агентите функционират без пряката намеса на човек или друг агент и имат собствен контрол над своите действия и вътрешно състояние,
- *социални способности* - агентите взаимодействуват с други агенти и с хора чрез езици за комуникация,
- *адекватни реакции* - агентите разбират обкръжаващата среда и своевременно и адекватно реагират на промените в нея,
- *инициативност* - агентите не просто действат в съответствие със своята заобикаляща среда, те са способни да проявяват целенасочено поведение като поемат инициатива,
- *способност да използват понятия, които са характерни за хората,* като зная, вярвам, имам намерение, поемам задължението и други.

Други, допълнителни свойства на агентите са:

- *мобилност* - способността на агента да се премества по електронната мрежа,
- *доброжелателност* - агентите нямат конфликтни цели и всеки агент се опитва да извърши това, за което е помолен,
- *рационалност* - агентът търси ефективни пътища за постигане на своите цели.

Един от начините да се опишат агентите е чрез техните *съзнателни локални състояния* (*mental states*), като например: *зная, вярвам, имам намерение, задължен съм* и други.

В сборника статии [*Connolly, Edmonds, 94*] са разгледани различни агенти. В [*Cohen, Levesque,90*] е представен един подход за представяне на локалното състояние *имам намерение*. Този подход е дискутиран в [*Singh,92*].

Следващият етап от развитието на интелигентния статистически софтуер е създаване на агенти за решаване на статистически задачи.

Ще се концентрираме върху проектирането на агент за проверка на статистически хипотези.

Традиционният път за решаване на този проблем води до някои нерешени стари проблеми на Изкуствения интелект, като например проблема за разбиране на текст на естествен език. Ние ще заобиколим този проблем като го заменим с проблема за извличане на съдържателна статистическа информация от текст на естествен език [*Noncheva,Stojanov,96*] и от данни, и ще третираме непълното разбиране на проблема на потребителя като източник на случайност. Независими кооперирящи се агенти ще построят стохастичен модел на проблема на потребителя, който модел те ще проверяват и подобряват във времето чрез въпроси към потребителя и чрез самообучение.

Като език за комуникация на агентите в мултиагентната система ще използваме езика на математическата статистика [*Noncheva,Stojanov,95*], в който степента на сигурност може да бъде представена. Ако C е твърдение, можем да запишем $P(C)=0.95$, което изразява степента на сигурност относно верността на твърдението C. т.е. Вероятността не е металингвистична конструкция, а понятие на самия език.

2.2 Подходи за създаване на мултиагентни системи

Съществуват два основни подхода за изграждане на мултиагентни системи - top_down и bottom_up.

Top_down подходът в литературата е описан от Burkhard. Този подход се състои от следните две стъпки ([*Burkhard,93*], [*Burkhard,93б*], [*Burkhard, 94*]):

- Дефинира се мултиагентната система и нейните свойства.

- Строят се различни агенти, които се дефинират в контекста на мултиагентната система.

В този подход поведението на един агент не се дефинира като самостоятелно поведение, а се разглежда в контекста на мултиагентната система. Не се правят предположения за поведението на един агент извън тази система.

Във формализма на Burkhard поведението на мултиагентната система се описва чрез префиксно затворен език над крайна азбука от атомарни действия.

Основното предимство на top_down подхода е единният подход за реализиране на кооперирането и комуникацията между агентите в една разпределена среда. Друго предимство е лесното поддържане на желаните свойства на агентите (например коректност).

Този подход се прилага към мултиагентни системи, компонентите на които тепърва ще се разработват.

Подходът, който ние ще следваме е bottom_up подход. Bottom_up подходът реализира следната схема:

1. Разработване на множество от отделни агенти със свои локални свойства, които реализират разпределената функционалност на системата.
2. При изпълнение на заявки на потребителите към системата, агентите могат да се кооперират за решаване на общи задачи. Така между тях възниква определена динамична зависимост.

Различни мултиагентни системи в различно време могат да използват едни и същи агенти. В даден момент един агент може да решава повече от една задача ако са налични подходящи ресурси (например многопроцесорна машина). При определени условия един агент може да се мултиплицира чрез самокопиране. От друга страна може да има повече от един агент, които решават определен клас задачи.

Агентът работи за системата, която от своя страна може да даде оценка за полезността на всяка свършена от него работа.

Агентът променя своето поведение в зависимост от това коя система му задава задачата. Това е съществено за някои от агентите. В този случай речниците, смисловите тегла ще бъдат различни.

3. След изпълнението на задачата динамичната зависимост се разпада и агентите остават отново напълно автономни.

Bottom_up подходът позволява изграждането на мултиагентни системи, работещи в разпределени среди и включващи в структурата съществуващи програмни модули и пакети. Т.е. съществуващи програми могат да бъдат "агентифицирани".

Недостатък на този подход е трудната реализация на механизми за комуникация, синхронизация и кооперирание.

2.3 **FORESEE** - *мултиагентна система за анализ и прогнозиране на временни редове*

Мултиагентната система FORESEE е резултат от над четиринадесет годишния опит в развитието на системи, решаващи сложни статистически проблеми на колектив от Политехническия университет в Каталуня, Испания.

Първо е развита идеологията на MEDISA (Multi agEnt DIStributed Architecture) и средствата за проектиране и изграждане на агентите, след което са построени двете мултиагентни системи - DESEXP и FORESEE. Този подход е дискутиран в [*Prat at al.,94*].

FORESEE е прототип на мултиагентна система за анализ и прогнозиране на временни редове. Тя може да определя, оценява и проверява ARIMA модели като използва методологията на Бокс и Дженкинс. Процесът на построяване на модел се състои от трите, циклично повтарящи се стъпки:

- идентификация - избор на подходящ параметричен модел за изучавания временен ред, т.е. избор на класа модели. Методите на идентификация използват автокорелационни и частни автокорелационни функции.
- оценяване на неизвестните параметри на модела.
- проверка за адекватност.

Системата може да генерира прогнози и като използва генетични алгоритми и невронни мрежи.

FORESEE може да сравни резултатите от различните статистически подходи и подходите на изкуствения интелект и да избере този, който предлага най-добрата прогноза. За сравняване на моделите се използва критерия за минимална дисперсия на грешките на прогнозите.

Тази система може да работи в два режима - автоматичен и стъпка по стъпка, в зависимост от предпочитанието на потребителя.

На *фигура 2.1* е показана архитектурата на системата. Нейните компоненти са следните агенти:

- *Агент, реализиращ интерфейса* с агенти и потребители

Интерфейсът, управляван от меню, не е подходящ за системи, решаващи сложни проблеми, където си взаимодействат много агенти и потребителя. Поради тази причина интерфейсът на FORESEE съдържа сцени и диалози, в които има много участници.

Множеството от всичките взаимодействия с потребителя в процеса на решаване на даден проблем се нарича история. Това множество е разбито на подмножества, наречени сцени. Сцената съдържа всички необходими взаимодействия при решаването на даден подпроблем. Тя

е множество от диалози. Диалогът е основната единица на взаимодействие с потребителя.

Агентът, реализиращ интерфейса, е програмно реализиран на Visual Basic.

- *Обучаващ агент*

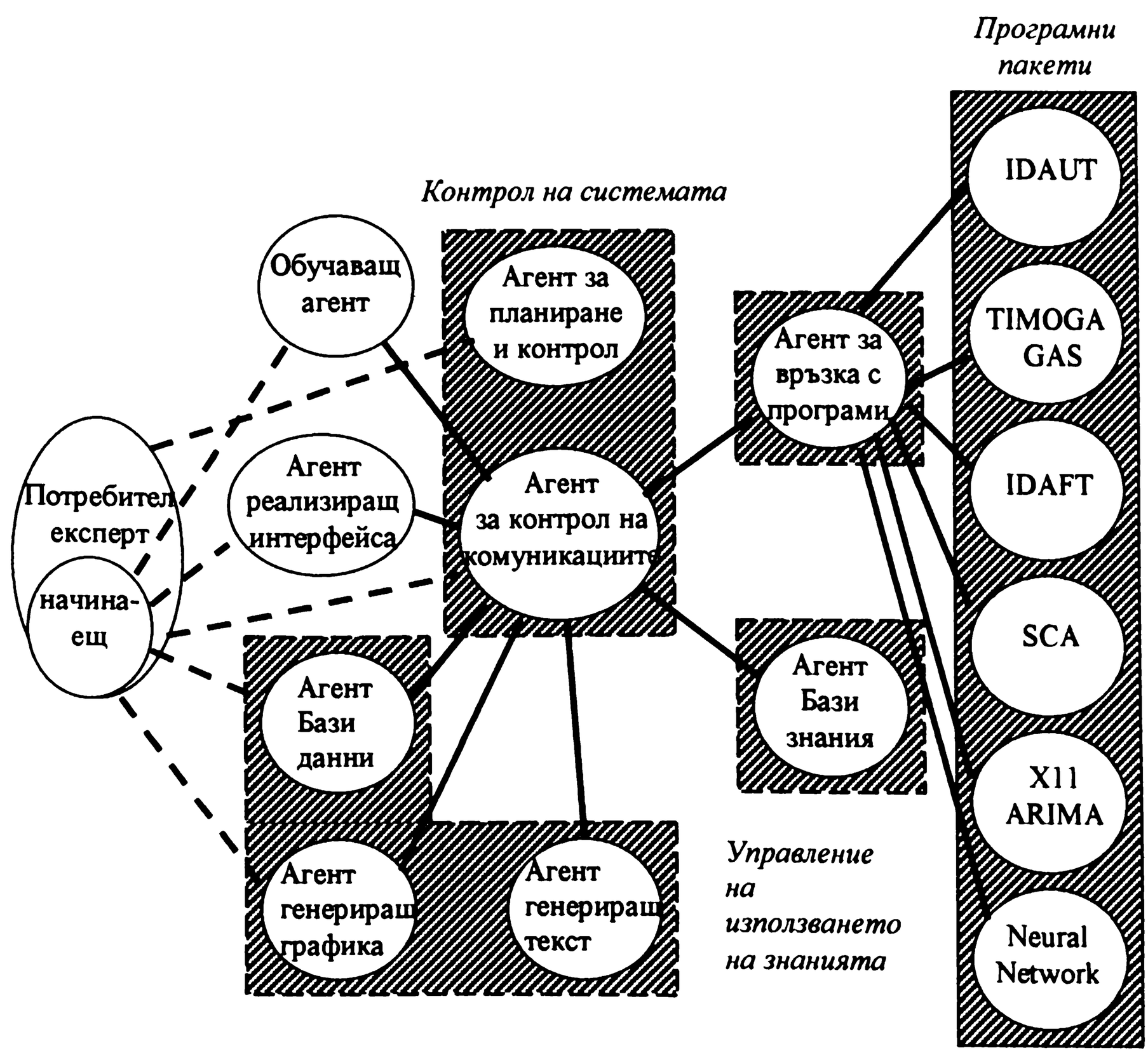
Обучаващият агент притежава знания на статистик - експерт в приложната област. Реализиран е със средствата на Visual Basic. Използват се и възможностите на Multimedia Viewer.

- *Агент База данни*

  Развит е със средствата на Access.
- *Агент, генериращ графика*

  За реализирането му е използван Excel.
- *Агент, генериращ текст* на естествен език

  За реализирането му е използван Winword.



Фигура 2.1 *Архитектура на мултиагентната система* **FORESEE**

- *Агент База знания*

Съдържа евристични знания, представящи поведението на статистика - експерт в областта на временните редове, в процеса на идентификация и проверка за адекватност. Знанията са представени чрез правила. Този агент е реализиран със средствата на KAPPA на Intercorp.

- *Агент, реализиращ интерфейса с пакети от програми*

Съществуват много програмни пакети и библиотеки, като NAG, BMDP и други, които системата би могла да използва за да реши даден подпроблем.

- *Агент за планиране и вътрешен контрол*

Този агент е отговорен за глобалната стратегия при решаване на проблема. Той притежава знания за стратегията и ръководи потребителя по време на решението. Във всеки момент той знае състоянието на процеса и получава информация от другите агенти за да има обща и пълна картина за всички действия.

Този агент може да извлича знания от експертите. За тази цел на експертите са предоставени визуални средства за да опишат своята стратегия при решаване на такива проблеми.

- *Агент за контрол на комуникациите*

FORESEE притежава следните видове знания:

- знания на консултант за цялостното решаване на проблема,
- знания на преподавател за методите за прогнозиране за да ги обясни,
- знания на експерт по методите на Box и Jenkins,
- знания на експерт за приложенията на системата.

2.4 Мултиагентна система за статистически изследвания

В този параграф ще представим архитектурата на Мултиагентната система за статистически изследвания. За построяването й избрахме мултиагентен bottom_up подход поради следните съображения:

- Искаме в системата да могат да се използват съществуващи и доказали своята ефективност статистически пакети.

- Искаме различните статистически методи да могат да се използват в разпределени среди и с възможност за отдалечен достъп.
- По-ефективно и гъвкаво адаптиране на съществуващите пакети в различни операционни среди.
- По-добро подпомагане (и евентуално контекстно-зависимо) на крайните потребители на системата.

Мултиагентната система за статистически изследвания се състои от следните, комуникиращи един с друг агенти (*фигура 2.2*):

- *Агент филтър*,
- *Агент, анализиращ текст на естествен език*,
- *Планиращ агент*,
- *Агент, генериращ текст на естествен език*,
- *Агент поща*,
- Различни *статистически агенти*,
- *Агент, управляващ статистическите пакети*.

Агентът филтър филтрира трафика на комуникациите между системата и реалния компютърен свят. Той пропуска към Мултиагентната система само тези съобщения, които е класифицирал като статистически задачи.

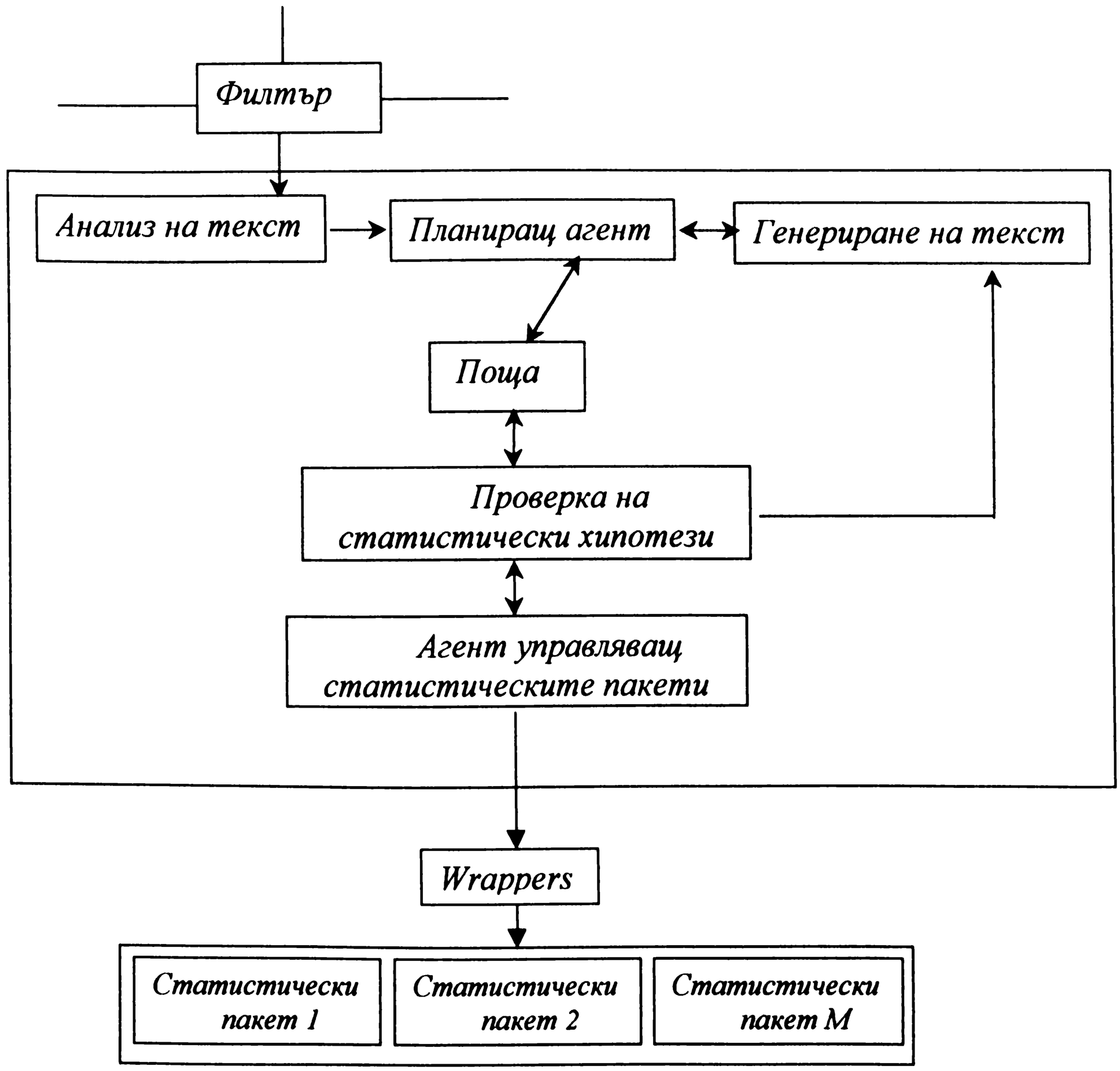
Агент, анализиращ текст на естествен език анализира проблема на потребителя, поставен на естествен език. Този агент е многоезиков - той анализира текстове на различни естествени езици. Един подход за "разбиране" на статистическия проблем чрез извличане на статистическа информация от текст на естествен език е предложен в [*Noncheva, Stojanov, 96*].

Планиращият агент анализира съобщението, изпратено от *Агента, анализиращ текст на естествен език*, определя към кой клас статистически задачи принадлежи задачата на потребителя и изпраща запитване към пощенската кутия на съответния статистически агент. Ако задачата не може да бъде решена, *Планиращият агент* изпраща съобщение за отказ към *Агента, генериращ текст на естествен език*.

Агентът, генериращ текст на естествен език осъществява връзката на системата с потребителя. Той може да генерира текст на езика на математическата статистика, текст на естествен език, съдържащ интерпретация на статистическите изводи и въпроси към потребителя в термините на самия потребител. Един подход за генериране на текст на естествен език е представен в [*Kerpedjiev,*

Noncheva,90], [*Нончева,90*]. Алгоритъм за форматиране на текст на български език е описан в [*Нончева,88*].

Всеки *статистически агент* решава определен клас статистически задачи и използва *Агента, управляващ статистическите пакети* за да реализира съответните изчисления.



Фигура 2.2 *Мултиагентна система за статистически изследвания.*

Агентът, управляващ статистическите пакети знае кои статистически пакети се намират в глобалната мрежа и могат да бъдат адресирани. Той решава кой пакет ще бъде използван,

подготвя входните данни в необходимия формат и връща получения резултат на статистическия агент.

Интелигентността на мултиагентната система се формира от интелигентността на отделните агенти и от интелигентността, породена от тяхното взаимодействие.

Подходът, който прилагаме при построяване на Мултиагентната система за статистически изследвания е bottom up подход, който се състои от следните два етапа:

1. Определяне на основното множество от агенти и изграждане на всеки агент. Основното множество от агенти на Мултиагентната система за статистически изследвания е *Агент филтър*, *Агент, анализиращ текст на естествен език*, *Планиращ агент*, *Агент, генериращ текст на естествен език*, *Агент поща*, *Агент за проверка на статистически хипотези*, *Агент, управляващ статистическите пакети*.
2. Построяване на мултиагентната система.

Разширяването на основното множество от агенти с *Обучаващ агент*, *Изпитващ агент* и *Агент за оценяване на знанията* води до нова мултиагентна архитектура, наречена Мултиагентна система за обучение по статистика [*Noncheva,96*].

Глава 3

Концептуален модел на агент за проверка на статистически хипотези

The only satisfactory description of uncertainty is probability. By this it is meant that every uncertainty statement must be in the form of a probability, that several uncertainties must be combined using the rules of probability, and that the calculation of the probabilities is adequate to handle all situations involving uncertainty. In particular, alternative descriptions of uncertainty are unnecessary.

Dennis V. Lindley,

The probability approach to the treatment of uncertainty in Artificial Intelligence

3.1 Бейсова мрежа

Нека $X=\{X_1, X_2, \ldots, X_n\}$ е множество от случайни величини.

Определение 3.1: Казваме, че върху множеството от случайни величини $X=\{X_1, X_2, \ldots, X_n\}$ е построен вероятностен модел, ако е определено съвместното им вероятностно разпределение $p(x_1,x_2,\ldots,x_n)$.

Нека ориентираният граф D представя зависимостите във вероятностния модел.

Определение 3.2: Перфектен граф на модел се нарича такъв граф, в който всяко твърдение за независимост в модела може да бъде представено и всяко представено в графа твърдение за независимост е в сила и за модела.

Могат ли моделите, които имат перфектен ориентиран граф, да бъдат напълно характеризирани с едно крайно множество от свойства, е един отворен въпрос.

Определение 3.3: Граф на независимостта в модела е такъв граф, в който всяко представено твърдение за независимост е в сила и за модела.

Графът на независимостта съдържа някои, но не всички твърдения за независимост в модела. Ние искаме да сме сигурни, че никакви неверни твърдения за независимост няма да се извлекат от графа и че броят на не представените в графа твърдения за независимост ще бъде сведен до минимум.

Когато не можем да построим ориентиран перфектен граф, естествено е да се опитаме да построим един минимален ориентиран граф на независимостта.

Определение 3.4: Граф на независимостта, който представя всяка зависимост в модела и престава да прави това ако премахнем поне една връзка, се нарича *минимален граф на независимостта*.

В сила е следната теорема:

Теорема 3.1: Нека $\boldsymbol{D}$ е ориентиран ацикличен граф и $\boldsymbol{p}(\boldsymbol{x}_1,\boldsymbol{x}_2, \ldots, \boldsymbol{x}_n)$ е вероятностен модел върху $\boldsymbol{D}$. Тогава следните две твърдения са еквивалентни:

- $\boldsymbol{p}(\boldsymbol{x}_1,\boldsymbol{x}_2, \ldots, \boldsymbol{x}_n) = \prod_{i=1}^{n} \boldsymbol{p}(\boldsymbol{x}_i/\pi_i)$, където π_i е родител на възела $\boldsymbol{X}_i$ в графа $\boldsymbol{D}$.
- $\boldsymbol{D}$ е граф на независимостта в модела $\boldsymbol{p}(\boldsymbol{x}_1,\boldsymbol{x}_2, \ldots, \boldsymbol{x}_n)$.

Теоремата е доказана в [*Castillo, Gutierrez, Hadi,97*].

Определение 3.5: Бейсова мрежа (Belief Networks) се нарича двойката $(\boldsymbol{D},\boldsymbol{P})$, където $\boldsymbol{D}$ е ориентиран ацикличен граф, $\boldsymbol{P} = \{\boldsymbol{p}(x_1/\pi_1), \boldsymbol{p}(x_2/\pi_2),\ldots, \boldsymbol{p}(\boldsymbol{x}_n/\pi_n)\}$ е множество от условни вероятностни разпределения на всяка случайна величина $\boldsymbol{X}_i$, π_i е множество от родителите на възела $\boldsymbol{X}_i$ в графа $\boldsymbol{D}$, $i = 1,\ldots n$. Множеството $\boldsymbol{P}$ дефинира съвместното вероятностно разпределение по следния начин:

$$\boldsymbol{p}(\boldsymbol{x}_1,\boldsymbol{x}_2, \ldots, \boldsymbol{x}_n) = \prod_{i=1}^{n} \boldsymbol{p}(\boldsymbol{x}_i/\pi_i).$$

$\boldsymbol{D}$ е минимален ориентиран граф на независимостта в модела $\boldsymbol{p}(\boldsymbol{x}_1,\boldsymbol{x}_2, \ldots, \boldsymbol{x}_n)$.

От *теорема 3.1* следва, че всяко твърдение за условна независимост, получено от $\boldsymbol{D}$ е в сила и в съответния модел $\boldsymbol{p}(\boldsymbol{x}_1,\boldsymbol{x}_2, \ldots, \boldsymbol{x}_n)$.

Бейсови мрежи с различни структури на графите, асоциирани с тях, са разгледани в [*Pearl,86*], [*Verma, Pearl,90*], [*Shafer,87*], [*Castillo, Gutierrez, Hadi,97*].

В [*Cox, Wermuth,93*] са дискутирани различни еквивалентни начини за графично представяне на линейни зависимости. Авторите демонстрират използването на графите като мощно и гъвкаво математическо средство и обещават да направят този формален графичен език широко разпространен в статистическия анализ.

Първоначалният замисъл за изграждане на експертна система за прилагане на подходящи статистически методи за проверка на хипотези в условията на непълна определеност постепенно прерасна в идея и конкретно виждане за създаване на статистически агент.

3.2 Предназначение на агента

Много потребителски задачи се свеждат до задачи за проверка на статистически хипотези. Пример:

Съгласно стандарта за производство на електрически крушки стандартното отклонение на живота на електрическите крушки трябва да е по-малко от 200 часа. Приема се, че животът на крушките е приблизително нормално разпределен. Направена е случайна извадка. Данните са записани във файл. Бихме ли могли да бъдем 99% сигурни, че производствения процес се контролира?

Тази задача се свежда до следната задача за проверка на статистически хипотези:

Направена е случайна извадка от нормална популация. Да се провери хипотезата:

H_0: $\sigma^2 = 40\ 000$ срещу

H_1: $\sigma^2 < 40\ 000$ при $\alpha=0{,}01$.

Агентът би могъл да се използва за проверка на статистически хипотези в многоагентна система за статистически анализ. Пример:

Дисперсионният анализ е статистически метод, изучаващ влиянието на една или повече групиращи променливи, наречени фактори, върху една количествена променлива, наречена отклик. Могат да се отделят факторите с най-голямо влияние върху отклика и да се оцени тяхното влияние.

Ще разгледаме постановката на задачата на еднофакторния дисперсионен анализ с постоянни ефекти. Наблюденията могат да се групират според различните стойности на факторите. Основната задача на еднофакторния дисперсионен анализ с постоянни ефекти се свежда до ***проверка на хипотезата дали съвпадат средните стойности на отклика в групите****. Ако хипотезата за равенство на средните се отхвърли, то следващата задача на дисперсионния анализ е оценка на средните стойности за всяка група.*

Агентът за проверка на статистически хипотези произвежда несигурни факти, които биха могли да допълват базата знания на други агенти, които могат да използват несигурни знания в своите разсъждения. Пример:

Резултатите от двуфакторния дисперсионен анализ на статистически език изглеждат така:

С ниво на значимост 0.99 хипотезата за незначимост на смесените ефекти се отхвърля, т.е. С увереност 0.99 може да се

приеме, че влиянието на типа на семената и торенето като цяло върху добива е съществено.

С със сигурност 99% можем да твърдим, че взаимодействието между факторите отсъства.

С със сигурност 99% можем да твърдим, че поотделно влиянието и на двата фактора е значимо и не може да бъде пренебрегнато.

Агентът за проверка на статистически хипотези може да се използва при определяне на мисловното състояние вяра, убеденост на различни статистически агенти. Проблемът за моделиране на вярата на агента във вида на популационното разпределение ще бъде разгледан подробно в настоящата работа.

3.3 Модел на поведението на агента

Нека X_t е множество от случайни величини, описващо състоянието на агента в момента t.

Нека E_t е възприетата от агента информация в момента t. Тогава E_1 , E_2 , ... , E_t е историята на възприетата информация до момента t, включително.

Нека A_t е действието, извършено от агента в момента t. Тогава A_1 , A_2 , ... , A_t са изпълнените действия до момента t.

Интересуваме се от $\mathbf{P}\{X_t / E_1, E_2, \ldots, E_t, A_1, A_2, \ldots, A_t\}$.

Ще направим следните предположения:

- $\mathbf{P}\{A_{t-1} / E_1, E_2, \ldots, E_{t-1}, A_1, A_2, \ldots, A_{t-2}\} = \mathbf{P}\{A_{t-1} / E_1, E_2, \ldots, E_{t-1}\}$.

Съгласно определението на агент действията на агента са резултат от неговите възприятия.

- $\mathbf{P}\{E_t / X_1, X_2, \ldots, X_t, E_1, E_2, \ldots, E_{t-1}, A_1, A_2, \ldots, A_{t-1}\} = \mathbf{P}\{E_t / X_t\}$.

Околната среда на агента е достъпна за агента. Но в една достъпна околна среда възприетата информация и състоянието могат да се приемат за еквивалентни. Т.е. състоянието да бъде направено достатъчно богато да отразява и всички особености на изменение на околната среда. С промяна на околната среда се променя съответно и състоянието на агента.

- $\mathbf{P}\{X_t / X_1, X_2, \ldots, X_{t-1}, A_1, A_2, \ldots, A_{t-1}\} = \mathbf{P}\{X_t / X_{t-1}, A_1, A_2, \ldots, A_{t-1}\}$

Основното предположение, което ще направим е, че проблемът притежава Марковското свойство, т.е. вероятностното разпределение на текущото състояние на агента зависи само от предишното състояние и от действията на агента, извършени до момента.

- $\mathbf{P}\{X_t / X_{t-1}, A_1, A_2, \ldots, A_{t-1}\} = \mathbf{P}\{X_t / X_{t-1}\}$

Получихме

$\mathbf{P}\{X_t / E_1, E_2, \ldots, E_t, A_1, A_2, \ldots, A_t\} = \mathbf{P}\{X_t / X_{t-1}\}$.

Следователно получихме Марковска верига. Състоянията на Марковската верига могат да се представят с Бейсова мрежа.

3.4 База знания

Базите знания могат да бъдат детерминистични и стохастични, според характера на проблемите, за които те са проектирани. Бази знания, които използват вероятността като мярка за несигурността, се наричат вероятностни. Трудностите при представянето на съвместното вероятностно разпределение забавиха развитието на вероятностните бази знания. С въвеждането на вероятностни мрежови модели тези пречки се преодоляха. Под вероятностен модел разбираме модел на зависимостта и едно съвместно вероятностно разпределение, асоциирано с този модел. Вероятностните мрежови модели се базират на представяне на зависимостите между променливите чрез граф. Това представяне води до ефективен начин на представяне на съвместното вероятностно разпределение и до ефективни алгоритми за автоматични изводи.

3.4.1 Общи положения

Още на първия етап на статистическото изследване трябва да се отговори на следните въпроси:

- Може ли да считаме, че данните, с които разполагаме, са резултат от независими наблюдения над някоя случайна величина?
- При наличието на няколко извадки може ли да считаме, че те са направени от една и съща популация?
- Кой вероятностен модел най-добре описва данните?
- Как да обработваме данните при наличието на "липсващи данни"?
- Как да се постъпи със "съмнителните наблюдения"?

От правилните отговори на тези въпроси зависи коректното и ефективното използване на методите на статистическия анализ.

Базата статистически знания на агента съдържа стандартните критерии за проверка на прости хипотези срещу сложни алтернативи.

Критерии за проверка на статистически хипотези са разгледани в [*Димитров, Янев, 90*], [*Винце,Върбанова,81*], [*Кендал, Стьюарт, 73*], [*Колев, 93*], [*Хог,Крейг,82*], [*Айвазян,Енюков,Мешалкин,83*], [*Афифи,*

Дизен, 82], [*Большев, 87*], [*Герасимович,83*], [*Daniel, Terrell, 89*], [*Hines, Montgomery,90*], [*Mendenhall, 87*]. В [*Logchem, 94*] е дадена една класификация на над 250 статистически критерии, предназначена за използване от изследователи в областта на социалните и поведенческите науки.

Ще разгледаме някои характерни особености на критериите, представени в базата знания.

В основата на някои от стандартните критерии за проверка на хипотези стои предположението за нормалност. Известно е, че t-критерият на Стюдънт за средното на непрекъснатата случайна величина, описваща популацията, е най-мощен. Този критерий е и робастен, т.е. той е нечувствителен към умерени отклонения от предположението за нормалност на изследваната популация. Но критерият χ^2 на Пирсън за дисперсията на непрекъснатата популация е силно чувствителен дори към слаби отклонения от предположението за нормалност. Асиметрията на извадковото разпределение насочва към подходяща трансформация на данните. Ако разпределението на трансформираните данни е по-близко до нормалното, то при проверка на хипотези за неизвестна дисперсия е целесъобразно да се използва новата променлива. Критерият χ^2 също е най-мощен.

Представени са и критерии за съгласие.

В Базата статистически знания е разгледан и случая, когато две различни непрекъснати случайни величини са определени върху една и съща популация. Тогава може да се изследва ковариацията и корелацията между тях. Случайната извадка е сдвоена. Ако двете случайни величини измерват еднакви характеристики, то можем да сравняваме средните им стойности чрез t-критерия за сдвоени извадки.

Ако една и съща непрекъсната случайна величина е определена върху две различни популации, тези популации могат да се разглеждат като подпопулации или слоеве (страти) на една по-голяма популация. В този случай популацията се разслоява на две подпопулации чрез дискретна случайна величина, която има две нива. Сега ковариацията и корелацията нямат смисъл. Правим две независими случайни извадки - по една от всяка популация. В този случай за проверка на хипотезите за равенство на средните се използва t-критерия за две независими извадки или критерия на Уелч. Т-критерият е най-мощен. Той предполага равенство на популационните дисперсии. Това предположение може да се провери чрез критерия на Фишер за отношение на дисперсиите. Той е най-мощен. Но този критерий е силно чувствителен към отклонения от предположението за нормалност. Затова при съмнения, че предположението за нормалност е нарушено, този критерий не бива да се използва.

Ако непрекъснатите случайни величини $X_1, X_2, \ldots X_p$ са определени върху една и съща популация можем да изследваме $p(p+1)/2$ различни коефициенти на ковариация и корелация между X_i и X_j, $i,j=1,2,\ldots,p$. Случайната извадка се състои от n p-мерни наблюдения, при което всичките p наблюдения при фиксирано n са направени над един и същ индивид от извадката.

Когато една непрекъсната случайна величина е определена върху p популации за проверката на хипотезата за равенство на популационните средни се използва еднофакторния дисперсионен анализ. Хипотезата за равенство на популационните дисперсии може да се провери чрез критерия на Бартлет, който е силно чувствителен към предположението за нормалност. Критерият на Бартлет при $p=2$ е с по-малка мощност от критерия на Фишер. Затова, ако всички извадки са с един и същ обем по критерия на Фишер се сравняват най-голямата и най-малката емпирични дисперсии. Ако тяхната разлика не е значима, то няма да е значима и разликата между останалите дисперсии.

Могат да се проверят и хипотези за линейни комбинации на средните. Сравнение между средните може да се извърши и с помощта на методите на множественото сравнение.

Статистически изводи за изследваната популация можем да направим и с помощта на честотни таблици. В случая на дихотомни наблюдения целта е да се проверят хипотези за пропорция. Използва се биномния критерий или негова нормална апроксимация.

Ако популацията се разбива на k непресичащи се класа, при което всеки елемент принадлежи само на един клас, то статистическите изводи се свеждат до изучаване на параметрите $p_1, p_2, \ldots, p_k$, където p_i е вероятността елементът да принадлежи на i-тия клас, $i=1,2,\ldots k$ и

$$\sum_{i=1}^{k} p_i = 1.$$

Ако всеки елемент от извадката едновременно се класифицира с помощта на два фактора или по два признака, съответно с r и c нива получаваме $r \times c$ таблица. Тогава можем да проверим хипотези за независимост на факторите.

Използването на различните критерии за независимост зависи от скалата на измерване на данните. Критерият за рангова корелация на Кендал се прилага за данни от порядковата, интервалната и относителната скала, но не и от номиналната. Критериите на Гудман и Крускал се прилагат за номиналната или порядковата скала, ако данните не са резултат от непрекъсната случайна величина.

За да избере правилно критерия за проверка на статистически хипотези статистикът трябва да си отговори на двата въпроса [*Motulsky,95*]:

- От какъв вид са данните, с които той разполага?
- Каква е целта?

За правилния автоматичен избор на статистически критерий е необходима следната информация:

- Брой на стратифициращите променливи- детерминистични променливи, които разбиват изследваната популация на подпопулации.
- Вероятностно разпределение на случайните величини, които описват подпопулациите.
- Скала на измерване на данните.
- Каква е целта?

3.4.2 Описание

С всеки от разгледаните проблеми ще свържем дискретна случайна величина X_i с краен брой стойности, описваща състоянието. Многообразието на методите за проверка на статистически хипотези ще представим чрез Бейсова мрежа (виж *фигура 3.1*), в която:

- Възелът X_1 представя случайната величина *стойност на стратифициращата променлива.*
- Възлите X_i, $i = 2,3,4$ представят съответно *брой стойности на* $(i-1)$*-вата дискретна случайна величина,*
- X_5 *- вид на дискретното разпределение,*
- X_6 *- брой на непрекъснатите случайни величини,*
- X_7 *- вид на непрекъснатото разпределение,*
- X_8 *- цел на потребителя,*
- X_9 *- вид на извадката,*
- X_{10} *- скала на измерване на данните.*
- Възлите $\{X_{2m+1}, m=5,6,\ldots,23\}$ представят класове статистически задачи. Например възлите X_{19} и X_{21} представят класа задачи, известен като проверка на статистически хипотези за съгласие при предположение съответно за дискретност и за непрекъснатост на вероятностния модел. Стойностите, които могат да приемат тези случайни величини, са две - "*приложим*" и "*неприложим*".
- Възлите $\{X_{2m+2}, m=5,6,\ldots,23\}$ представят статистически критерии. Стойностите, които могат да приемат тези случайни величини, са представени в *Таблица 3.2.*

При представяне на методите на математическата статистика чрез граф се взема под внимание съществуващата класификация на статистическите методи. Класовете параметрични и непараметрични

критерии за централната тенденция на непрекъсната популация се представят чрез два възела, независимо от това, че биха могли да бъдат представени чрез един. Критериите за съгласие и критериите за еднородност също биха могли да бъдат представени чрез един възел.

Така графът на Бейсовата мрежа представя исторически утвърдилата се класификация на статистическите методи.

"Качеството" на критериите се характеризира с нивото на значимост, с мощността и със свойството неизместеност. Известно е, че критерият на Абб за случайност е по-мощен от критерия на възходящите и низходящите серии. Известно е също, че при използването на критерия за съгласие на Колмогоров и на критерия на Крамер-Мизес-Смирнов към групирани данни нивото на значимост на критериите ще бъде по-малко от зададеното. Въпреки, че един и същ критерий се използва в различни ситуации, той може да има различни функции на мощността.

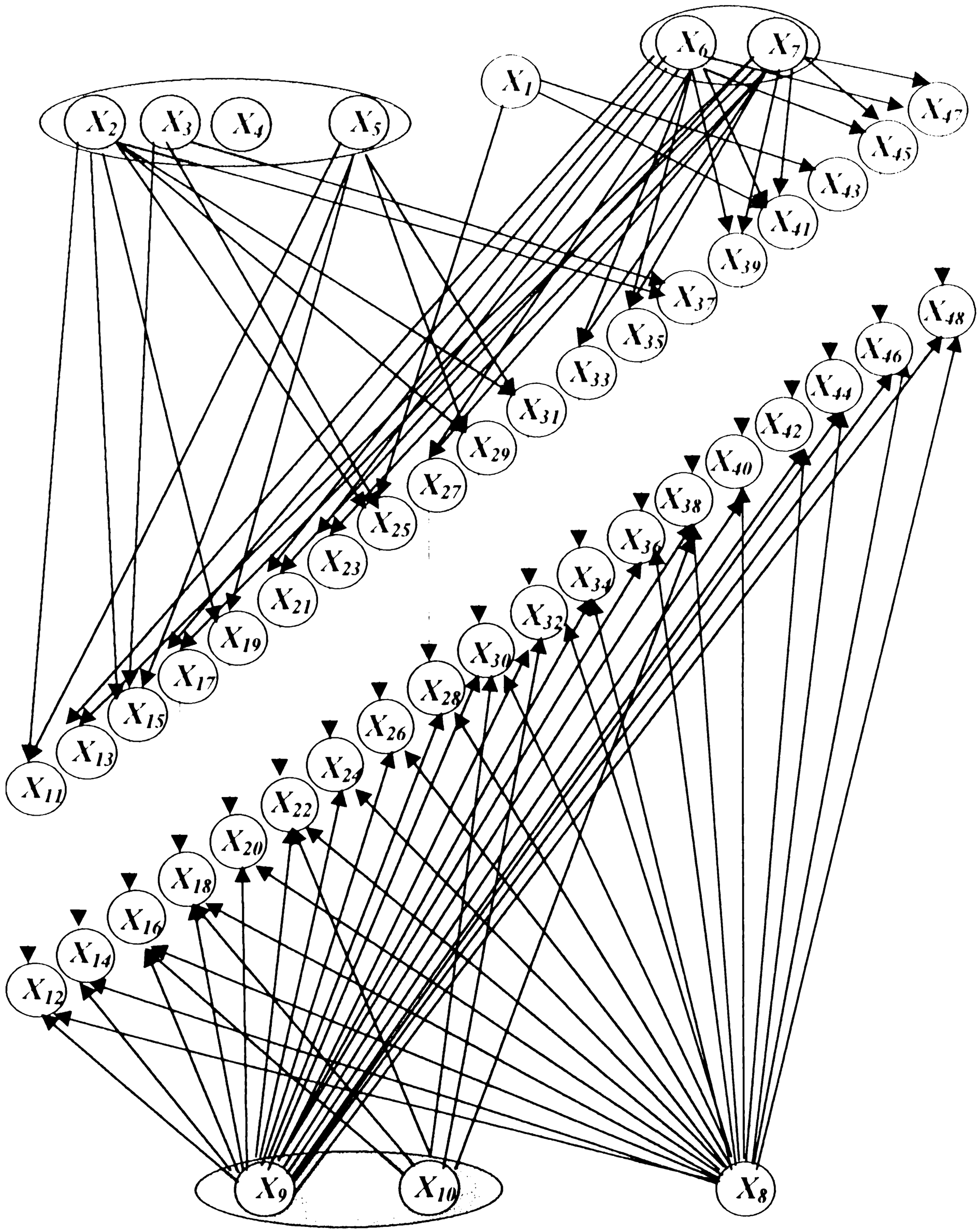
Поведението на интелигентния статистически агент при извършване на статистически анализ се управлява чрез съответните условни вероятности. Условните вероятностни разпределения отразяват предимствата и недостатъците на статистическите критерии.

Моделът на първоначалната база знания е графично представен на *фигура 3.1*. Полученият граф има poly-tree структура, т.е. той е дърво, в което някои възли имат повече от един родител. Случайните величини и тяхната интерпретация са представени в *Таблица 3.1*. Стойностите, които могат да приемат случайните величини, са представени в *Таблица 3.2*.

Структурата на зависимостта на дискретните случайни величини е определена и чрез минималното множество от променливи $Nbr(\boldsymbol{X}_i)$ за всяка променлива $\boldsymbol{X}_i$ за което е изпълнено $p\ (x_i\ /x\backslash x_i\ )=p(\ x_i\ /Nbr\ (\boldsymbol{X}_i\ ))$. Множеството $Nbr(\boldsymbol{X}_i)$ се нарича *съседи на* $\boldsymbol{X}_i$. Случайните променливи от модела и техните родители и съседи са представени в *Таблица 3.3*.

До сега разглеждахме само един вид статистически критерии - статистически критерии при които вероятността да отхвърлим вярната основна хипотеза е предварително определена (фиксирана). Развиват се и статистически критерии, отчитащи вероятността да се приеме основната хипотеза, когато тя не е вярна. Пример за такива критерии са критериите, построени с подхода на последователния анализ и критериите, построени с подхода на теорията на статистическите решения. Такива критерии се прилагат в радиотехниката при отделяне на сигналите на фона на шумовете и при разпознаване на образи.

Ако вероятността да се отхвърли вярната основна хипотеза и вероятността да се приеме основната хипотеза, когато тя не е вярна, са предварително фиксирани, може да се определи минималния обем на извадката, която трябва да се направи за да се вземе решение.



Фигура 3.1 *Модел на Базата статистически знания.*

Сл.в.	Определение
X_1	Стойност на стратифициращата променлива.
X_2	Брой стойности на първата дискретна случайна величина.
X_3	Брой стойности на втората дискретна случайна величина.
X_4	Брой стойности на третата дискретна случайна величина.
X_5	Вид на дискретното разпределение.
X_6	Брой на непрекъснатите случайни величини.
X_7	Вид непрекъснато разпределение.
X_8	Цел на потребителя.
X_9	Вид на извадката.
X_{10}	Скала на измерване на данните.
X_{11}	Проверка на хипотези за случайност при предположение за дискретност на случайната величина, която описва популацията
X_{12}	Критерий за случайност на извадката при предположение за дискретност.
X_{13}	Проверка на хипотези за случайност при предположение за непрекъснатост на случайната величина, която описва популацията.
X_{14}	Критерий за случайност на извадката при предположение за непрекъснатост.
X_{15}	Проверка на хипотези за еднородност при предположение за дискретност.
X_{16}	Критерий за еднородност при предположение за дискретност.
X_{17}	Проверка на хипотези за еднородност при предположение за непрекъснатост.
X_{18}	Критерий за еднородност при предположение за непрекъснатост.
X_{19}	Проверка на хипотези за съгласие при предположение за дискретност на вероятностния модел.
X_{20}	Критерий за съгласие при предположение за дискретност на вероятностния модел.
X_{21}	Проверка на хипотези за съгласие при предположение за непрекъснатост на вероятностния модел.
X_{22}	Критерий за съгласие при предположение за непрекъснатост на вероятностния модел.
X_{23}	Проверка на хипотези за съмнителни наблюдения.
X_{24}	Критерий за съмнителни наблюдения.
X_{25}	Проверка на хипотези за независимост при предположение за дискретност.
X_{26}	Критерий за независимост при предположение за дискретност.
X_{27}	Проверка на хипотези за коефициента на корелация.
X_{28}	Критерий за коефициента на корелация.
X_{29}	Проверка на хипотези за неизвестни параметри на дискретно разпределение.
X_{30}	Параметричен критерий за неизвестен параметър на дискретно разпределение.
X_{31}	Рангови критерии за централната тенденция.
X_{32}	Рангов критерий за централната тенденция.
X_{33}	Проверка на хипотези за неизвестни параметри на непрекъснати разпределения.
X_{34}	Параметричен критерий за неизвестен параметър на непрекъснато разпределение.
X_{35}	Непараметрични критерии за централната тенденция на непрекъснати разпределения.
X_{36}	Непараметричен критерий за централната тенденция.
X_{37}	Проверка на хипотези за равенство на неизвестни параметри на две дискретни разпределения.

X_{38}	*Параметричен критерий за равенство на параметри на две дискретни разпределения.*
X_{39}	*Проверка на хипотези за равенство на неизвестни параметри на две непрекъснати разпределения.*
X_{40}	*Параметричен критерий за равенство на параметри на две непрекъснати разпределения.*
X_{41}	*Еднофакторен дисперсионен анализ.*
X_{42}	*Методи на еднофакторния дисперсионен анализ.*
X_{43}	*Проверка на хипотези за линейни комбинации на средните.*
X_{44}	*Критерий за линейни комбинации на средните.*
X_{45}	*Проверка на непараметрични хипотези за разсейването.*
X_{46}	*Непараметричен критерий за разсейването.*
X_{47}	*Проверка на хипотези за симетричност.*
X_{48}	*Критерий за симетричност.*

Таблица 3.1 *Таблица на случайните величини.*

Обогатяването на Базата знания е един продължителен и по всяка вероятност безкраен процес. В този процес промените в графичната структура на Базата статистически знания се извършват от експерт- статистик. Вероятностната структура се обновява от статистик и чрез самообучение.

3.5 Автоматични изводи

Разглеждаме Бейсовия модел (D,P), определен върху множеството от дискретни случайни величини $\boldsymbol{X}=\{\boldsymbol{X_1}, \boldsymbol{X_2}, \ldots, \boldsymbol{X_n}\}$, представящ Базата статистически знания.

Преди да е дефинирана задачата на потребителя процесът на извод се състои в изчисляване на маргиналните вероятностни разпределения $p(\boldsymbol{X}_i=x_i)=p(x_i)$, за всяко $\boldsymbol{X}_i\in\boldsymbol{X}$. То дава информация за променливите в модела, преди да е настъпило някакво събитие.

Нека събитието $\boldsymbol{E}\subset\boldsymbol{X}$ е настъпило, т.е. дефинирана е задача на потребителя. Тогава някои променливи $\boldsymbol{X}_i$ приемат определени стойности, т.е. $\boldsymbol{X}_i=e_i$ за всяко $\boldsymbol{X}_i\in\boldsymbol{E}$. Например $\boldsymbol{X}_6=1$, $\boldsymbol{X}_7=$*нормално*. Настъпването на събитието $\boldsymbol{E}\subset\boldsymbol{X}$ ще означим с $\boldsymbol{E}=e$. В този случай процесът на извод се състои в изчисляване на условните вероятностни разпределения $p(x_i/e)$ за всяко $\boldsymbol{X}_i\in\boldsymbol{X}\backslash\boldsymbol{E}$, $\boldsymbol{E}=e$. Условните вероятностни разпределения измерват ефекта от настъпването на събитието $\boldsymbol{E}$.

Нашата задача е да предложим алгоритъм за извод, който отчита структурата на независимост между променливите, която е отразена в ориентирания ацикличен граф, представящ Базата статистически знания.

X_i	***Стойности на случайната величина X_i***
X_{12}	*Серийн критерий на Вало-Волфовиц за случайност, Нормална апроксимация на критерия на Вало-Волфовиц, Медианен серийн критерий.*
X_{14}	*Рангов критерий на Ман, Критерий на възходящите и низходящите серии, Медианен серийн критерий, Критерий на Абб за извадка от нормално разпределение.*
X_{16}	*χ^2 критерий за еднородност, Модифициран рангов критерий за групирани данни.*
X_{18}	*Критерий на Колмогоров-Смирнов, Критерий на Гнеденко-Королюк за извадки с еднакъв обем, χ^2-тест за еднородност, T^2 критерий за еднородност на многомерни нормални популации, Критерий на знаците на Фишер, Критерий на Большев-Никулин, Критерий на Крамер-Мизес, Линеен рангов критерий на Уилкоксън-Ман-Уитни, Линеен рангов критерий на Фишер-Йетс-Тери-Гефринг, Линеен рангов критерий на Ван дер Варден, Критерий на Клотс за равенство на дисперсиите, Критерий на Муд, Критерий на Крускал-Уолис при повече от две извадки.*
X_{20}	*χ^2-тест за вида на разпределението, Биномен критерий, χ^2 апроксимация на Биномния критерий, Критерий на Макнимар, Критерий на Кокрън, Критерий на знаците, Критерий на Рао-Робсън за биномност, Критерий на Большев-Мирвалиев за пуасоновост, Критерий на Большев-Мирвалиев за биномност, Критерий на Большев-Мирвалиев за отрицателна биномност.*
X_{22}	*Критерий на Колмогоров за извадка с голям обем, Критерий на Крамер-Мизес-Смирнов за извадка с голям обем, Модифициран критерий на Колмогоров за извадка с малък обем, Модифициран критерий на Крамер-Мизес-Смирнов за извадка с малък обем, χ^2-тест, Критерий на Рени, Критерий на Колмогоров при неизвестни параметри, оценени по извадка, Критерий на Крамер-Мизес-Смирнов при неизвестни параметри, оценени по извадка, Критерий на Колмогоров за нормалност при неизвестни параметри, оценени по извадка, Критерий на Крамер-Мизес-Смирнов за нормалност при неизвестни параметри, оценени по извадка, Проверка за нормалност по асиметрията и отношението на абсолютното отклонение и средноквадратичното отклонение, Проверка за нормалност по асиметрията и ексцеса при $n \geq 50$, Проверка за нормалност по асиметрията и ексцеса при $n > 1000$, Критерий на Шапиро-Уилк за нормалност при $n \leq 50$, Критерий на Шапиро-Франсия за нормалност при $n \geq 50$, Критерий на Большев за експоненциалност, Критерий на Шапиро-Уилк за експоненциалност, Критерий на Шьормер за експоненциалност.*
X_{24}	*Правило на Шовене, Модифицирано правило на Шовене, Критерий на Е. Пирсън, Критерий на Чандра Секара, Критерий на Смирнов, Критерий на Грубс, Критерий на Большев-Убайдуллова, Критерий на Титен-Мир.*
X_{26}	*χ^2 критерий за независимост при извадки с голям обем, χ^2 критерий с поправка на Уетс при $n < 20$, Биномен критерий, Точен критерий на Фишер, Критерий за рангова корелация на Кендал, Критерий на Гудман, Критерий на Крускал.*
X_{28}	*t-тест за коефициента на корелация, Тест за коефициента на корелация при голям обем на извадката, χ^2 критерий с преобразуване в порядковата скала.*
X_{30}	*Биномен критерий за пропорция, Нормална апроксимация на Биномния критерий, Критерий за параметъра на Пуасоновото разпределение, Критерий за параметъра на геометричното разпределение.*
X_{32}	*Критерий на Уилкоксън, Критерий на Ман-Уитни-Уилкоксън, Медианен критерий, Критерий на Ансари-Брадли, Критерий на Мозес.*
X_{34}	*t-критерий на Стюдънт, χ^2 критерий на Пирсън,*
X_{36}	*Рангов критерий на Уилкоксън за сдвоени извадки, Критерий на знаците на Фишер за сдвоени извадки, Рангов критерий на Уилкоксън за една извадка, Критерий на знаците на Фишер за една извадка, Критерий на ранговите суми на Уилкоксън за*

	две независими извадки, Критерий на Краскал-Уолис за k 2 независими извадки, Критерий на Темпстра-Джонкхир за наредена алтернатива за k 2 независими извадки, Критерий на Фридман- рандомизиран пълен блоков експеримент, Критерий на Пейдж за наредена алтернатива - рандомизиран пълен блоков експеримент, Критерий на Дърбин -балансиран непълен блоков експеримент.
X_{38}	*Биномен критерий, Нормална апроксимация на Биномния критерий.*
X_{40}	*t-тест, критерий на Уелч, Процедура на множественото сравнение, Критерий на Бартлет, Критерий на Кокрън.*
X_{42}	*Модел с фиксирани ефекти, Модел със случайни ефекти.*
X_{44}	*Процедура на множественото сравнение, Метод на Шефе, Метод на Тюки*
X_{46}	*Критерий на Ансари-Брадли при известни или еднакви медиани, Критерий на Мозес при неизвестни и различни медиани .*
X_{48}	*Критерий на Гупта за симетричност около неизвестната медиана, Критерий на Уилкоксън при известен център на симетрия, Критерий на Фрезер-Клетц при известен център на симетрия, Критерий на Ван дер Варден при известен център на симетрия, Аналог на критерия на Уилкоксън при известен център на симетрия, Аналог на критерия на Фрезер-Клетц, Аналог на критерия на Ван дер Варден.*

Таблица 3.2 *Таблица на стойностите на случайните величини.*

Целта на алгоритъма за автоматичен извод е да се изчислят условните вероятности $p(x_i/e)$ на всички възли $X_i \in X \setminus E$ при условие, че някое събитие $E \subset X$ е настъпило.

Но в много случаи ние се интересуваме само от определено множество от променливи $Y \subset X$ и нашата цел е да получим условните вероятности само на тези променливи, когато дадено събитие $E \subset X$ е настъпило. В този случай някои възли не са необходими при пресмятането на необходимите условни вероятности. Тези възли могат да бъдат отстранени от графа и да се получи подграф, съдържащ само необходимите възли.

Нека се интересуваме от множеството от променливи $Y \subset X$ и искаме да пресметнем условните вероятности $p(x_i/e)$ за всяко $X_i \in Y$ при $E=e$. Променливите от множеството Y се наричат *целеви променливи*. Възлите, необходими за изчисляване на $p(x_i/e)$ за всяко $X_i \in Y$ при $E=e$ се наричат *значими възли*.

Предложеният алгоритъм за извод в Базата статистически знания се състои от следните две стъпки:

- Редуциране на първоначалния граф до граф, съдържащ само значимите възли. Алгоритъмът за редуциране е представен в [*Castillo, Gutierrez, Hadi,97*] стр.379.
- Извод в дърво с възли с повече от един родител. Алгоритъмът за извод в ориентиран граф с poly-tree структура е представен в [*Castillo, Gutierrez, Hadi,97*] стр.328.

Ще поясним казаното до тук с един пример от математическата статистика.

Пример 3.1: Направени са независими наблюдения над случайната величина $\xi \in Ge(p)$. Да се провери хипотезата H_0: $p=1/2$ срещу алтернативата H_1:$p=2/3$ с ниво на съгласие $\alpha=0.01$. Данните са представени във файла G.dat.

Разглеждаме Бейсовия модел, представен с графа от *фигура 3.2*. Настъпило е събитието E={ $X_2=\propto$, $X_5=$Ge, $X_8=$*проверка на хипотези*, $X_9=$*случайна извадка*, $X_{10}=$*поредкова скала*, $X_{29}=$*проверка на хипотези за неизвестен параметър на дискретно разпределение*}. Интересуваме се от множеството $\boldsymbol{Y}=\{\boldsymbol{X}_{30}\}$, където $\boldsymbol{X}_{30}=$*критерий за неизвестен параметър на дискретно разпределение*. Целта на автоматичния извод е да се намерят условните вероятности $\boldsymbol{p}(x_{30}/e)$.

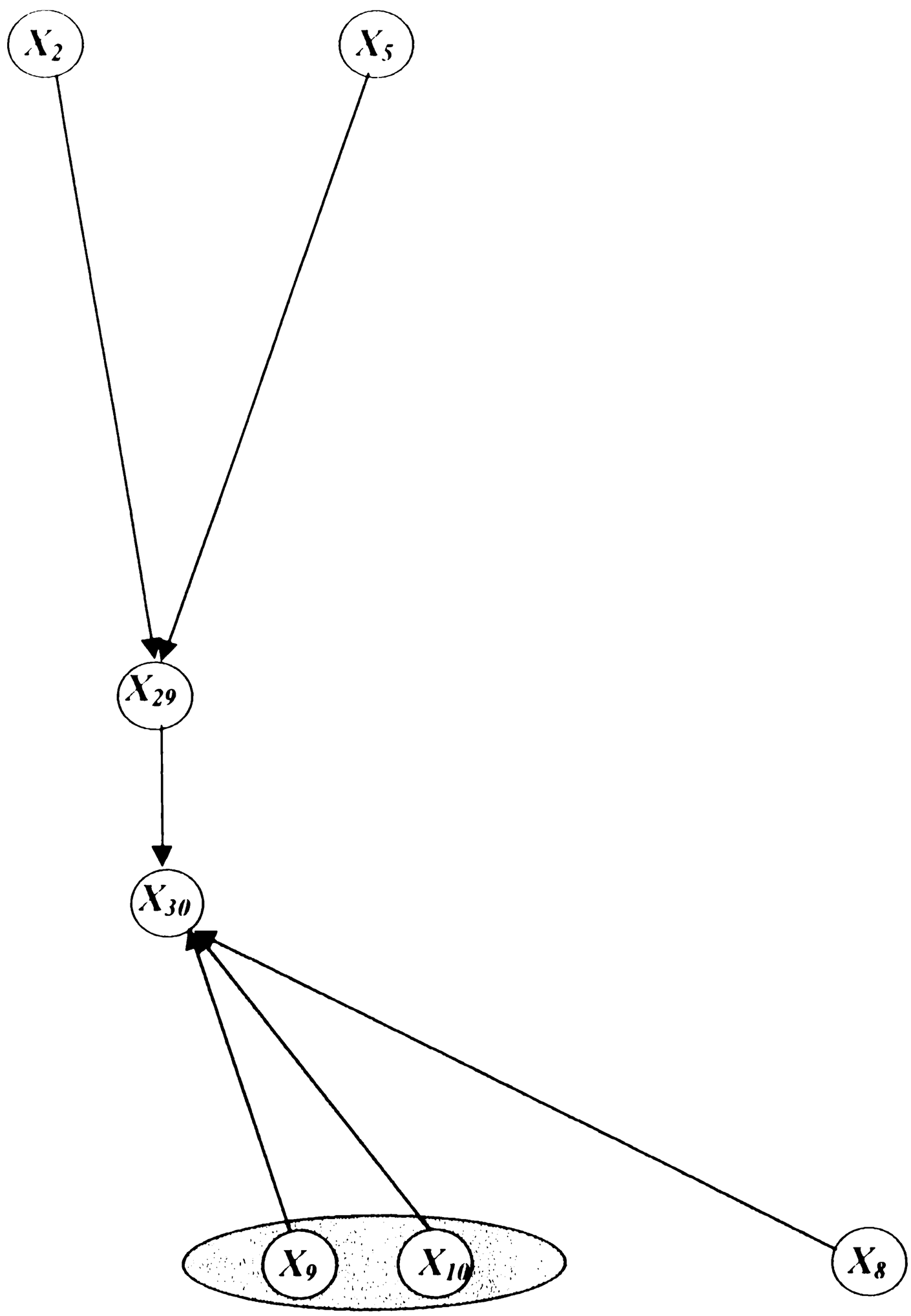
Графът, съдържащ само значимите възли, е представен на *фигура 3.2*.

Следователно опростихме структурата на първоначалния граф като премахнахме някои възли. По този начин избягваме излишните пресмятания на условни вероятности.

Прилагайки алгоритъмът за извод в ориентиран граф с poly-tree структура ще намерим търсеното условно вероятностно разпределение $\boldsymbol{p}(x_{30}/e)$.

Възли	Родители	Съседи
X_1	$\varnothing$	$\varnothing$
X_2	$\varnothing$	$\varnothing$
X_3	$\varnothing$	$\varnothing$
X_4	$\varnothing$	$\varnothing$
X_5	$\varnothing$	$\varnothing$
X_6	$\varnothing$	$\varnothing$
X_7	$\varnothing$	$\varnothing$
X_8	$\varnothing$	$\varnothing$
X_9	$\varnothing$	$\varnothing$
X_{10}	$\varnothing$	$\varnothing$
X_{11}	X_2, X_5	$X_2, X_5,$
X_{12}	X_8, X_9, X_{11}	$X_2, X_5, X_8, X_9, X_{11}$
X_{13}	X_6, X_7	X_6, X_7
X_{14}	X_8, X_9, X_{13}	$X_6, X_7, X_8, X_9, X_{13}$
X_{15}	X_2, X_3, X_5	X_2, X_3, X_5
X_{16}	X_8, X_9, X_{10}, X_{15}	$X_2, X_3, X_5, X_8, X_9, X_{10}, X_{15}$
X_{17}	X_6, X_7	X_6, X_7
X_{18}	X_8, X_9, X_{10}, X_{17}	$X_6, X_7, X_8, X_9, X_{10}, X_{17}$
X_{19}	X_2, X_5	X_2, X_5
X_{20}	X_8, X_9, X_{19}	$X_2, X_5, X_8, X_9, X_{19}$
X_{21}	X_6, X_7	X_6, X_7
X_{22}	X_8, X_9, X_{10}, X_{21}	$X_6, X_7, X_8, X_9, X_{10}, X_{21}$
X_{23}	X_6, X_7	X_6, X_7
X_{24}	X_8, X_9, X_{23}	$X_6, X_7, X_8, X_9, X_{23}$
X_{25}	X_1, X_2, X_3	X_1, X_2, X_3
X_{26}	X_8, X_9, X_{25}	$X_1, X_2, X_3, X_8, X_9, X_{25}$
X_{27}	X_6, X_7	X_6, X_7
X_{28}	X_8, X_9, X_{27}	$X_6, X_7, X_8, X_9, X_{27}$
X_{29}	X_2, X_5	X_2, X_5
X_{30}	X_8, X_9, X_{10}, X_{29}	$X_2, X_5, X_8, X_9, X_{10}, X_{29}$
X_{31}	X_2, X_5	X_2, X_5
X_{32}	X_8, X_9, X_{10}, X_{31}	$X_2, X_5, X_8, X_9, X_{10}, X_{31}$
X_{33}	X_6, X_7	X_6, X_7
X_{34}	X_8, X_9, X_{33}	$X_6, X_7, X_8, X_9, X_{33}$
X_{35}	X_6, X_7	X_6, X_7
X_{36}	X_8, X_9, X_{35}	$X_6, X_7, X_8, X_9, X_{35}$
X_{37}	X_2, X_3	X_2, X_3
X_{38}	X_8, X_9, X_{10}, X_{37}	$X_2, X_3, X_8, X_9, X_{10}, X_{37}$
X_{39}	X_6, X_7	X_6, X_7
X_{40}	X_8, X_9, X_{39}	$X_6, X_7, X_8, X_9, X_{39}$
X_{41}	X_1, X_6, X_7	X_1, X_6, X_7
X_{42}	X_8, X_9, X_{41}	$X_1, X_6, X_7, X_8, X_9, X_{41}$
X_{43}	X_1	X_1
X_{44}	X_8, X_9, X_{43}	X_1, X_8, X_9, X_{43}
X_{45}	X_6, X_7	X_6, X_7
X_{46}	X_8, X_9, X_{45}	$X_6, X_7, X_8, X_9, X_{45}$
X_{47}	X_6, X_7	X_6, X_7
X_{48}	X_8, X_9, X_{47}	$X_6, X_7, X_8, X_9, X_{47}$

Таблица 3.3 *Таблица на възлите и техните родители и съседи.*



Фигура 3.2 *Редуцираният граф от Пример 3.1.*

3.6 Представяне на външни въздействия

Определянето на състоянието на вярата на агента може да се представи като външно въздействие върху възела X_7 при което в даден момент от време на променливата X_7 й е наложено да приеме една фиксирана стойност x'_7. Това въздействие ще означим с $F_7 = set(X_7 = x'_7)$. Това външно въздействие е представено в графа D като допълнителен възел, родител на X_7 (виж *фигура 3.4*). По този начин получаваме един разширен граф $D' = D \cup \{F_7 \rightarrow X_7\}$, представен на *фигура 3.4*. Новата променлива F_7 приема стойности в множеството $\{set(x'_7), \varnothing\}$, където x'_7 принадлежи на множеството от възможните стойности на случайната променлива X_7, $\varnothing$ представя липсата на външно въздействие.

Разглеждаме външното въздействие $F_i = set(X_i = x'_i)$, кодирано чрез добавяне на връзката $F_i \rightarrow X_i$ към D.

Новото множество от родители на X_i в разширения граф $D' = D \cup \{F_i \rightarrow X_i\}$ е $\pi'_i = \pi_i \cup \{F_i\}$ и е свързано с X_i чрез условните вероятности

$$p(x_i/\pi'_i) = \begin{cases} p(x_i/\pi_i), & \text{ако } F_i = \varnothing \\ 0, & \text{ако } F_i = set(X_i = x'_i) \text{ и } x_i \neq x'_i \\ 1, & \text{ако } F_i = set(X_i = x'_i) \text{ и } x_i = x'_i \end{cases}$$

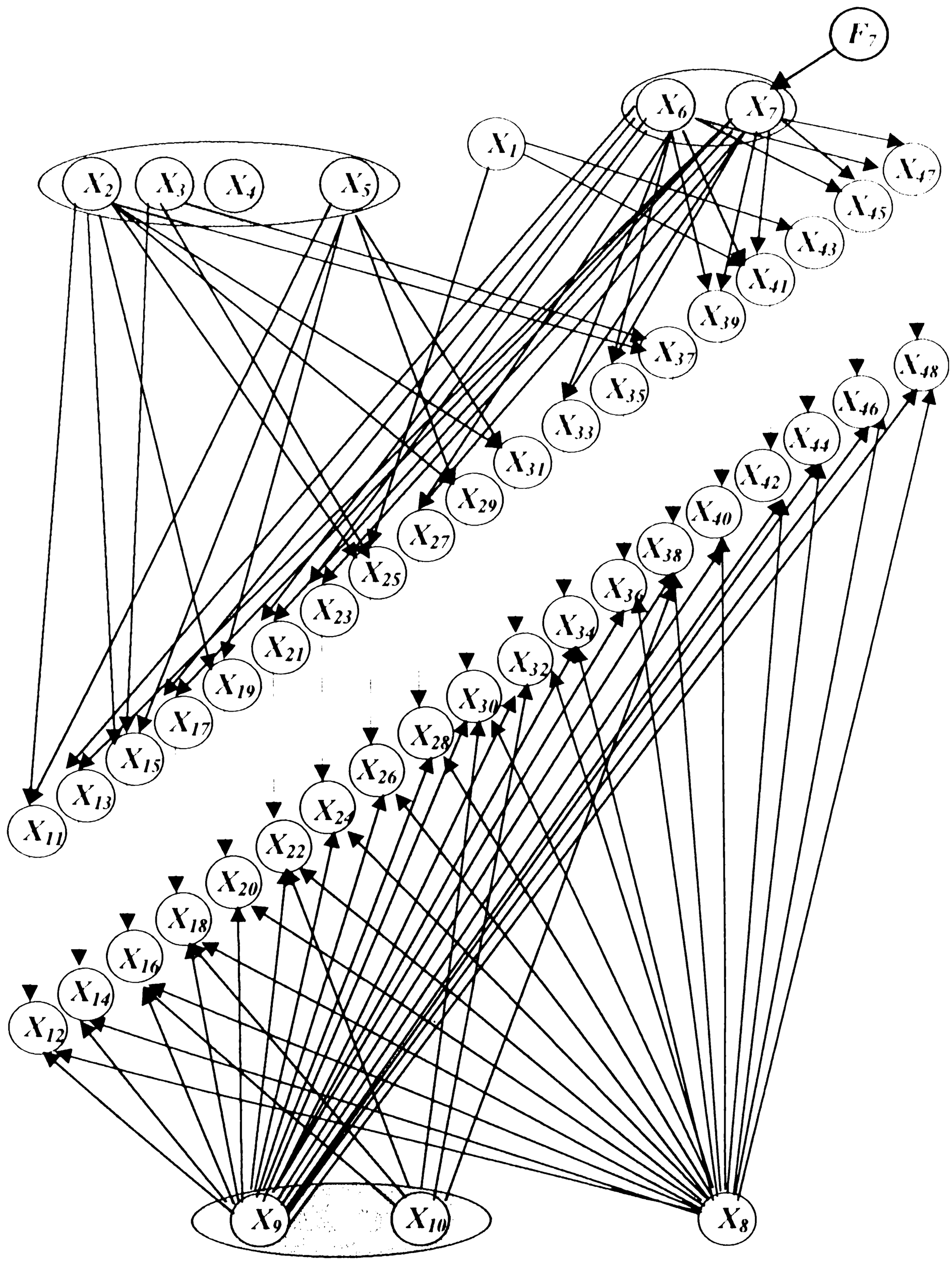
Съвместното вероятностно разпределение $p(x_1, x_2, \ldots, x_n)$ ще се трансформира в $p'(x_1, x_2, \ldots, x_n / F_i = set(x'_i))$, където

$$p'(x_1, x_2, \ldots, x_n / F_i = set(x'_i)) = \begin{cases} \dfrac{p(x_1, x_2, \ldots, x_n)}{p(x_i / \pi_i)}, & \text{при } x_i = x'_i \\ 0, & \text{в останалите случаи} \end{cases}$$

Ще разгледаме някои свойства на външните въздействия ([*Pearl,93*], [*Pearl,93б*]).

1. Външното въздействие $set(X_i = x'_i)$ може да повлияе само върху наследниците на X_i в D.
2. За всяко множество от променливи S е в сила твърдението за условна независимост $\mathbf{I}(S, F_i / (X_i, \pi_i))$.
3. Необходимото и достатъчно условие външното въздействие $set(X_i = x'_i)$ да има същото влияние върху X_j като събитието $X_i = x'_i$ е да е в сила твърдението за условна независимост $\mathbf{I}(X_j, \pi_i / X_i)$.

Следователно външното въздействие $F_7 = set(X_7 = x'_7)$ може да бъде представено в ориентиран ацикличен граф и да бъде анализирано чрез изводи в Бейсова мрежа когато събитието $X_7 = x'_7$ е настъпило.



Фигура 3.4 *Модел на Базата статистически знания с външно въздействие* $F_7 = set(X_7 = x'_7)$.

Глава 4

Математически модел на вярата

"Отражението на действителността (включително и вербалното) може да бъде истинно и неистинно. В комуникативния акт човек чувствува необходимост да изрази отношението си към факти, явления, действия и др. Понякога изрично се налага да подчертава своята *увереност* (особено ако има съмнения и колебания у събеседника) [*Кръстев,90, стр. 191*]. "

Вярата е убеждение в съществуването, осъществяването или истинността на нещо, основано на знания или опит [*Речник на българския език, 79*].

Вярата е увереност в научните изводи, хипотези, които в дадения момент още не могат да се докажат експериментално. Тя се опира на вече достигнатото и проверено на практика знание [*Философски речник, 78*].

Ще формализираме понятието вяра (увереност) на статистически агент във вероятностния модел по следния начин:

Вярата, увереността на агента ще измерваме с реално число, наречено *сигурност* (*степен на сигурност*). Казваме, че агентът вярва във верността на дадено твърдение, ако сигурността е по-голяма от известно, предварително фиксирано число. Т.е. вярата (увереността) на агента е прагова функция.

"*Подозрението* е предположение, свързано с някакво съмнение при разкриване на нещо неизвестно. То се проявява при търсене на истината [*Кръстев,90, стр. 283*]. "

"*Сигурността* е свързана с достоверността, истинността, възможността и увереността, че нещо е станало или ще стане. Тя е и състояние, и волева категория [*Кръстев,90, стр. 172*]. "

Формално, подозрението и сигурността са възможни стойности на праговата функция *вяра*, наречени *състояния на вярата на агента*.

4.1 Нормалност при предположение за симетричност

4.1.1 Постановка на задачата

Необходимостта от определяне на вярата на агента за вида на вероятностното разпределение произлиза от факта, че

популационното разпределение, стоящо в основата на много модели на математическата статистика, е неизвестно. Следствие от тези недостатъчни знания е неопределеността при избора на подходящия метод на статистически анализ, т.е. неопределеността при избора на най-доброто поведение на агента, реализиращ статистически анализ.

Нека T е непрекъсната случайна величина със стойности в Ω_t [0,1]. Нека t е възможна стойност (реализация) на случайната величина T. Случайната величина T ще наричаме *сигурност на агента*.

Нека $\{A_k \quad [t_k,t_{k+1}],\ t_k\in\Omega_t,\ t_{k+1}\in\Omega_t,\ k=0,1,\ldots,n-1\}$ е разлагане на Ω_t, т.е. $A_i\cap A_j=\varnothing$ при $i\neq j$ и $\sum_{k=0}^{n-1} A_k=\Omega_t$.

Нека δ е категорийна променлива, която може да приема стойности от множеството $D=\{a_0, a_1, \ldots, a_{n-1}\}$.

Ще дефинираме случайната величина Bel като функция на случайната величина T по следния начин:

$$Bel(T)=\begin{cases} a_0 & \text{при } t_0\le T<t_1 \\ a_1 & \text{при } t_1\le T<t_2 \\ \\ a_{n-1} & \text{при } t_{n-1}\le T<t_n \end{cases}$$

Случайната величина $Bel(T)$ ще наричаме *вяра на агента*. Стойностите a_0, a_1, ..., a_{n-1}, които може да приема случайната величина $Bel(T)$ ще наричаме *решения* или *състояния на вярата*. Множеството D ще наричаме *множество на решенията* или *множество на възможните състояния на вярата*.

Да разгледаме вярата на агента в нормалност на популационното разпределение.

В този случай $D=\{a_0, a_2\}$, където

a_0 – агентът отхвърля предположението за нормалност на разпределението, което описва изследваната популация.

a_2 - агентът е убеден (сигурен), че разпределението, което описва изследваната популация е нормално и ще използва това предположение в статистическия анализ.

Вярата на агента относно нормалност на популационното разпределение има вида:

$$Bel(T)=\begin{cases} a_0 & \text{при } 0\le T<t_c \\ a_2 & \text{при } t_c\le T<1, \end{cases}$$

където t_c е известна константа. Например t_c=0,95.

Целта е да се определи стойността на функцията $\delta = Bel(T)$. В този случай казваме, че агентът взема решение $\delta \in D$.

Но сигурността на агента T е ненаблюдаема случайна величина.

Вярата на агента във вероятностния модел (т.е. във вероятностното разпределение) трябва да се основава на знанията за механизма на изучаваното явление.

Но ако механизмът на изучаваното явление не е известен, агентът може да направи своя избор за вероятностното разпределение след като провери хипотеза за съгласие. Ще поясним изказаното твърдение.

Нека $X_1, X_2, \ldots, X_n$ е случайна извадка от разпределението на случайната величина, описваща популацията. Искаме да знаем разумно ли е да разглеждаме тази случайна величина като нормално разпределена.

Нека $x_1, x_2, \ldots, x_n$ са експерименталните стойности съответно на $X_1, X_2, \ldots, X_n$.

Издигаме статистическата хипотеза H_0, състояща се в това, че опитните данни са стойности на случайна величина с нормално $N(\mu,\sigma^2)$ разпределение, където μ и σ^2 са определени.

Статистически критерий (тест) се нарича правилото, по което вземаме решение да отхвърлим разглежданата статистическа хипотеза или да не я отхвърлим. Ниво на значимост α на статистическия критерий е максималната вероятност за отхвърляне на разглежданата основна хипотеза H_o при условие, че тя е вярна.

Нека $U=u(X_1, X_2, \ldots, X_n)$ е статистиката на оптималния статистически критерий, чието вероятностно разпределение при вярна основна хипотеза е известно. Нека $u=u(x_1, x_2, \ldots, x_n)$ е стойността на статистиката. Тогава $p(x_1,x_2,\ldots, x_n)=P\{U\geq u(x_1,x_2,\ldots, x_n)/H_o\}$ се нарича *p-стойност на статистиката на теста*. Статистическият критерий е следния: Ако $p(x_1,x_2,\ldots,x_n) < \alpha$, където α=1- t_c е желаното ниво на значимост на статистическия критерий, то основната статистическа хипотеза се отхвърля в полза на контра-хипотезата. В противен случай нямаме основание да отхвърлим основната хипотеза. т.е. Считаме, че данните, които имаме, не дават основание да се отхвърли хипотезата за нормалност.

След като провери хипотеза за съгласие при ниво на значимост α=1- t_c, агентът може да направи своя предварителен избор за вероятностното разпределение.

Но ако обемът на извадката не е голям или резултатите от измерванията (данните) са разположени в достатъчно малък интервал на изменение на случайната величина, описваща популацията, то

експерименталните данни могат добре да се съгласуват с различни вероятностни модели, т.е. с различни закони на разпределение. Затова не бива да се отдава голямо значение на положителния резултат от проверката на хипотези.

Ще подчертаем отново, че изборът на подходящия вероятностен модел трябва да се основава преди всичко на разбирането на механизма на изучаваното явление.

Когато агентът взема решение за вида на вероятностното разпределение той може да попита потребителя и да се основе на неговите знания, опит и интуиция.

Ако механизмът на изучаваното явление не е известен и на потребителя, то потребителят може да направи своя избор на разпределение като изхожда от следните съображения:

- вида на хистограмата.

  Но хистограмата може едновременно да напомня на няколко разпределения. Например по хистограмата практически е невъзможно да се различат логаритмически нормалното разпределение и разпределението на Вайбъл, даже при голям обем на извадката.

- графичното представяне на емпиричната функция на разпределение върху вероятностна хартия.

  Но няма количествен критерий за възможното отклонение на стойностите на емпиричната функция на разпределение от правата линия.

- извадковите коефициенти на асиметрия и ексцес.

  Ако за извадка с обем n са пресметнати точковите оценки на асиметрията $\overline{A} = \dfrac{\sum_{i=1}^{n}(x_i - \overline{x})^3}{ns^3}$ и на ексцеса $\overline{\varepsilon} = \dfrac{\sum_{i=1}^{n}(x_i - \overline{x})^4}{ns^4} - 3$ и техните средноквадратични отклонения $s_{\overline{A}} = \sqrt{\dfrac{6(n-1)}{(n+1)(n+3)}}$ и $s_{\overline{\varepsilon}} = \sqrt{\dfrac{24n(n-2)(n-3)}{(n-1)^2(n+3)(n+5)}}$, то считаме, че емпиричната функция на разпределение се съгласува с теоретичната, ако извадковите коефициенти на асиметрия и ексцес се различават по абсолютна стойност от своите математически очаквания не повече отколкото утроеното средноквадратично отклонение ([*Герасимович,83, стр. 221*]).

  За нормалното разпределение математическите очаквания на извадковите коефициенти на асиметрия и ексцес са равни на нула. Тогава правилото приема вида:

Ако са в сила неравенствата $\left|\overline{A}\right| < 3s_{\overline{A}}$ и $\left|\overline{\varepsilon}\right| < 3s_{\overline{\varepsilon}}$, то можем да считаме, че хипотезата за нормалност се съгласува с експерименталните данни. Ако поне едно от тези неравенства е нарушено, то основната хипотеза се отхвърля.

Но правилото на трите сигми е едно емпирично правило.

- резултатите от критерий за съгласие.

Но освен основните критерии за съгласие - χ^2 критерия и критерия на Колмогоров, има и други - критерия на Мизес-Смирнов, който за разлика от χ^2 критерия не изисква обединяване на числовите стойности в групи (в интервали), т.е. по-добре използва информацията, съдържаща се в извадката. Има и критерий за проверка за нормалност по достатъчно голям брой извадки с малък обем.

При вземане на решение за вида на вероятностното разпределение по един или повече от разгледаните по-горе начини опитният статистик може да заключи: "Аз съм 99% сигурен, че разпределението на изследваната популация е нормално." Друг потребител може да твърди, че само 80% е сигурен в нормалност на популационното разпределение. Трети разбира, че не може да определи вида на разпределението и твърди, че с една и съща вероятност то би могло да бъде нормално и да не бъде нормално. Степента на сигурност на тези трима потребители можем да представим съответно с числата (вероятностите) *0.99*, *0.8*, *0.5*.

Следователно при вземане на решение за нормалност на непрекъсната и симетрична популация агентът ще използва предварителен тест (пре-тест) и пост-тест. Обикновено като пре-тест се използва статистически критерий за вида на вероятностното разпределение, след което се иска и мнението на потребителя. Резултатите от тези два теста са съответно $1-\boldsymbol{p}$, където $\boldsymbol{p}$ е p-стойността на статистиката на теста, и степента на сигурност на потребителя, изразени като числа в интервала [0,1].

Понякога, например в регресионния анализ, първо се пита потребителя, а след това регресионните остатъци се проверяват за нормалност.

Естествено е да очакваме, че високите стойности на резултатите от пре-теста ще доведат до занижени изисквания към резултатите от пост-теста.

4.1.2 Модел

Нека с X да означим резултатите от пре-теста, с Y - резултатите от пост-теста. Нека да допуснем, че X и Y са непрекъснати случайни величини със стойности съответно в $\Omega_x=[0,1]$ и $\Omega_y=[0,1]$.

По резултатите от наблюденията x и y трябва да вземем решение за състоянието на вярата на агента $\delta=\delta(x,y)\in D$. Функцията $\delta(x,y)$, дефинирана върху множеството на възможните резултати от наблюденията и приемаща стойности в пространството D на възможните решения се нарича *решаваща функция* или *правило за вземане на решения*.

Нека T е сигурността на агента, т.е. T е непрекъсната случайна величина със стойности в $\Omega_t=[0,1]$ и която не може да се наблюдава.

Ще предположим още, че връзката между случайните величини X, Y и T може да бъде представена чрез съвместното вероятностно разпределение $f(x,y,t)$. Параметрите на това разпределение могат да се оценят като се извърши съответния експеримент.

Нека t е истинската стойност на сигурността на агента и е взето решение $d\in D$.

Нека с $u(d,t)$ да означим *полезността при вземане на решение* d.

Ако използваме правилото за вземане на решения $\delta(x,y)$, то очакваната полезност е $U(\delta/t) = \mathrm{E}u(\delta(x,y),t)$.

Функцията $U(\delta/t)$ се нарича *очаквана полезност* при използване на правилото за вземане на решения $\delta(x,y)$ при условие, че истинската стойност на сигурността на агента е t.

Изглежда разумно да изберем това правило за вземане на решение, ако това изобщо е възможно, на което съответства максимална очаквана полезност.

До тук предполагахме, че на всяко $x\in\Omega_x$ и на всяко $y\in\Omega_y$ съответства точно едно решение $\delta=\delta(x,y)\in D$. Но можем да допуснем избор на едно от няколко решения в съответствие с някакво разпределение на вероятността върху пространството на възможните решения. Можем да рандомизираме решенията.

Можем да смятаме, че при дадени x и y решението $\delta(x,y)$ е случайна величина със стойности от D и разпределение, което не зависи от t. Такова решаващо правило се нарича рандомизирано.

Ако $D=\{a_0, a_1, \ldots, a_{n-1}\}$, то решаващото правило ще се даде с вектора $\{p_0(x,y), p_1(x,y),\ldots, p_{n-1}(x,y)\}$, където $p_k(x,y)= p\{\delta(x,y)=d_k\}$, $k=0,1,\ldots,n-1$.

При рандомизация пространството на решенията D се заменя с изпъкналото множество от вектори $\{p_0(x,y), p_1(x,y),\ldots, p_{n-1}(x,y)\}$. Това е едно от основните преимущества на рандомизацията [*Климов,75*].

Търсим решаващо правило, което е най-добро в някакъв смисъл. За всяко решаващо правило δ означаваме

$$U(\delta) = \int_{\Omega_t} U(\delta / t)\mu(t)dt .$$

Дясната част на това равенство може да се интерпретира като претеглено средно на полезността. Тогава $\mu(t)$ показва степента на значимост, която се приписва на различните стойности на t [*Леман, 79*].

Решаващото правило δ^* се нарича *бейсовско*, ако

$$U(\delta^*) = \max U(\delta).$$

Да се върнем на задачата за определяне на вярата на агента относно нормалност на изследваната популация. Ще добавим нов елемент a_1 към множеството на възможните решения D, където състоянието a_1 има смисъла, че агентът предполага, подозира, че разпределението на изследваната популация е нормално.

Тогава $D=\{a_0, a_1, a_2\}$, където

a_0 – агентът отхвърля предположението за нормалност на разпределението, което описва изследваната популация,

a_1 – агентът предполага (подозира), че разпределението на изследваната популация е нормално. В процеса на статистически анализ той ще използва само критерии, които не са чувствителни към умерени отклонения от предположението за нормалност. Такива критерии се наричат робастни. Пример за такъв критерий е **t**-критерият.

a_2 - агентът е убеден (сигурен), че разпределението, което описва изследваната популация е нормално. Той ще използва и статистически критерии, които са чувствителни към отклонения от предположението за нормалност. Такива критерии са например критерият на Пирсън, критерият на Фишер за отношението на две дисперсии, критерият на Бартлет за равенство на **k** >2 дисперсии.

Популацията на потребителите ще разбием на g=2 подпопулации:

- подпопулация на потребителите със статистически знания и умения, която ще наричаме подпопулация на квалифицираните статистици и
- подпопулация на начинаещите статистици.

Ще предположим още, че връзката между измерените резултати от пре-теста X, измерените резултати от пост-теста Y и сигурността T в подпопулация i (i=1,...,g) може да бъде представена чрез съвместното вероятностно разпределение $f_i(x,y,t)$.

Правилото за вземане на решение $\delta(x,y)$ определя за всяка възможна реализация (x,y) на случайния вектор (X,Y) кое състояние a_j, $j=0,1,2$, ще бъде присвоено на вярата на агента.

Интуитивно е ясно, че високите стойности на резултатите от пре-теста водят до занижени изисквания към резултатите от пост-теста. Т.е. получената предварителна информация влияе на правилата за вземане на решения. Решенията са функция и от резултатите от пре-теста и от резултатите от пост-теста. Правила за вземане на решения, при които решенията от пост-теста са функции от конкретния резултат x от пре-теста ще наричаме *слаби правила*.

Слабото правило за вземане на решения δ може да бъде дефинирано по следния начин:

$$\begin{aligned} &\{(x,y): \delta(x,y)=a_0\}=A_i \times [0,1] \\ &\{(x,y): \delta(x,y)=a_1\}=A_i^c \times B_i(x) \\ &\{(x,y): \delta(x,y)=a_2\}=A_i^c \times B_i^c(x) \end{aligned} \tag{4.1}$$

където A_i, и A_i^c са множества от стойности x от пре-теста водещи съответно до отхвърляне и до приемане на твърдението за нормалност, $B_i(x)$ и $B_i^c(x)$ са множества от стойности y от пост-теста водещи съответно до отхвърляне и до приемане на твърдението за нормалност.

Правила за вземане на решения, при които решенията от пост-теста не са функции от конкретния резултат x от пре-теста ще наричаме *силни правила*.

Силното правило γ може да бъде дефинирано по следния начин:

$$\begin{aligned} &\{(x,y): \gamma(x,y)=a_0\}=A_i \times [0,1] \\ &\{(x,y): \gamma(x,y)=a_1\}=A_i^c \times B_i \\ &\{(x,y): \gamma(x,y)=a_2\}=A_i^c \times B_i^c . \end{aligned}$$

Естествено е да разглеждаме правила за вземане на решения, които имат *монотонна форма*. Например, агентът е убеден, че разпределението на популацията е нормално, ако стойността на променливата, представяща сигурността, е по-голяма от предварително определена прагова стойност (вероятност), и агентът отхвърля предположението за нормалност, ако стойността на променливата е по-малка от праговата стойност. Следователно, проблемът за вземане на решения относно вида на вероятностното разпределение на изследваната популация се свежда до намиране на две прагови точки (вероятности) x_c и y_c съответно за X и Y, които са оптимални в смисъла на Бейсовия подход. т.е $A_i=[0,x_c)$, $B_i=[0,y_c)$.

Наличието на популации от потребители с различна статистическа компетентност води до различни стойности на праговите точки x_c и y_c, които ще означаваме съответно с x_{ci} и y_{ci}. Праговата точка t_c на T е една и съща за всяка популация от потребители. Тя е предварително фиксирана и известна на агента, който взема решение. Например $t_c = 0.95$.

Тогава слабите монотонни правила за вземане на решение ще дефинираме по следния начин:

$$\delta(x,y)= \begin{cases} a_0 & \text{при } X<x_{ci} \\ a_1 & \text{при } X\geq x_{ci}\,,\ Y< y_{ci}(x) \\ a_2 & \text{при } X\geq x_{ci}\,,\ Y\geq y_{ci}(x), \end{cases} \qquad (4.2)$$

където a_j, $j=0,1,2$ са състоянията на вярата на агента, x_{ci} и y_{ci} са праговите точки за X и Y в подпопулацията i.

4.1.3 Функция полезност

Функция полезност се нарича функцията $u_{ji}(t)$, която описва полезността на резултата от присвояване на състоянието a_j, $(j = 0,1,2)$ на вярата на агента за случай с изследовател от подпопулация i когато сигурността на агента е t.

Видът и параметрите на функциите полезност могат да бъдат оценен по емпирични данни ([*Berhold, 67*], [*Davis, Hickman, Novick, 73*], [*Luce, Raiffa, 57*], [*Vrijhof, Mellenbergh, van den Brink, 83*]). Обикновено видът на функцията полезност е предварително определен и известен на този, който взема решение. Могат да бъдат разглеждани различни видове функции полезност [*Berhold, 73*], [*Vos, 90*], [*Vos, 91*], [*Vos, 93*], [*Vos, 94*].

Ние ще определим два вида функции полезност.

Нашата цел е да дефинираме реалистични функции полезност. Искаме в точките, близки до праговата стойност t_c, полезността от коректно към некоректно решение да се променя плавно, а не скокообразно. За тази цел ще използваме непрекъсната функция полезност. Един модел за вземане на решения относно вида на популационното разпределение, използващ линейна функция полезност е разгледан в [*Noncheva, 98*]. Естествено е да искаме функцията полезност да е ограничена и в отдалечените от праговата стойност t_c точки

полезността плавно да нараства и плавно да намалява към съответните граници. Ще разгледаме функция полезност, която е непрекъсната ограничена функция на истинската стойност на сигурността на агента. Probit моделите са подходящи за описване на такива зависимости. Тези модели имат вида $\pi(x) = F(\alpha + \beta x)$, където $F(.)$ е непрекъсната функция на вероятностно разпределение [*Agresti,90*].

В светлината на идеите на Berhold [*Berhold,73*] ще дефинираме функция полезност, която има следния вид:

$$u_{ji}(t) = \begin{cases} \phi[(\mu_{0i} - t)/\sigma_{0i}] & \text{при } j=0 \\ \phi[(t - \mu_{1i})/\sigma_{1i}] & \text{при } j=1 \\ \phi[(t - \mu_{2i})/\sigma_{2i}] & \text{при } j=2, \end{cases} \qquad (4.3)$$

където $\phi[.]$ е стандартната Гаусова функция на разпределение, а μ_{ji} и $\sigma_{ji} > 0$, $j=0,1,2$; $i=1,2$, са съответно средното и стандартното отклонение на Гаусовото разпределение.

Графиката на тази функция е представена на *Фигура 4.1*.

Тази структура ще наричаме *Гаусова структура полезност*.

Гаусовите функции на разпределение са симетрични относно 0,5 в смисъл, че $\phi(\mu + x) - 0.5 = 0.5 - \phi(\mu - x)$.

Ако искаме нашата функция полезност плавно да нараства от долната си граница и бързо да достига горната си граница ще използваме complementary log-log функцията $\pi(x) = 1 - \exp(-\exp(\alpha + \beta x))$. Графиката на тази функция е представена на *Фигура 4.2* при $\beta < 0$ и $\beta > 0$. Complementary log-log функцията е асиметрична функция, която при $\beta > 0$ се стреми по-бързо към горната си граница 1 отколкото към долната си граница 0. Втората й производна е $\dfrac{d^2\pi(x)}{d^2x} = \beta^2 \exp(\alpha + \beta x).\exp(-\exp(\alpha + \beta x)).[1 - \exp(\alpha + \beta x)]$.

Следователно инфлексната точка е $x = -\dfrac{\alpha}{\beta}$.

Следвайки този подход ще използваме функция полезност, която има следния вид:

$$u_{ji}(t) = \begin{cases} 1\text{-}exp(\text{-}exp(a_{0i}\text{-}b_{0i}t)) & \text{при } j=0 \\ 1\text{-}exp(\text{-}exp(a_{1i}+b_{1i}t)) & \text{при } j=1 \\ 1\text{-}exp(\text{-}exp(a_{2i}+b_{2i}t)) & \text{при } j=2, \end{cases} \qquad (4.4)$$

където $b_{0i} > 0$, $b_{2i} > 0$.

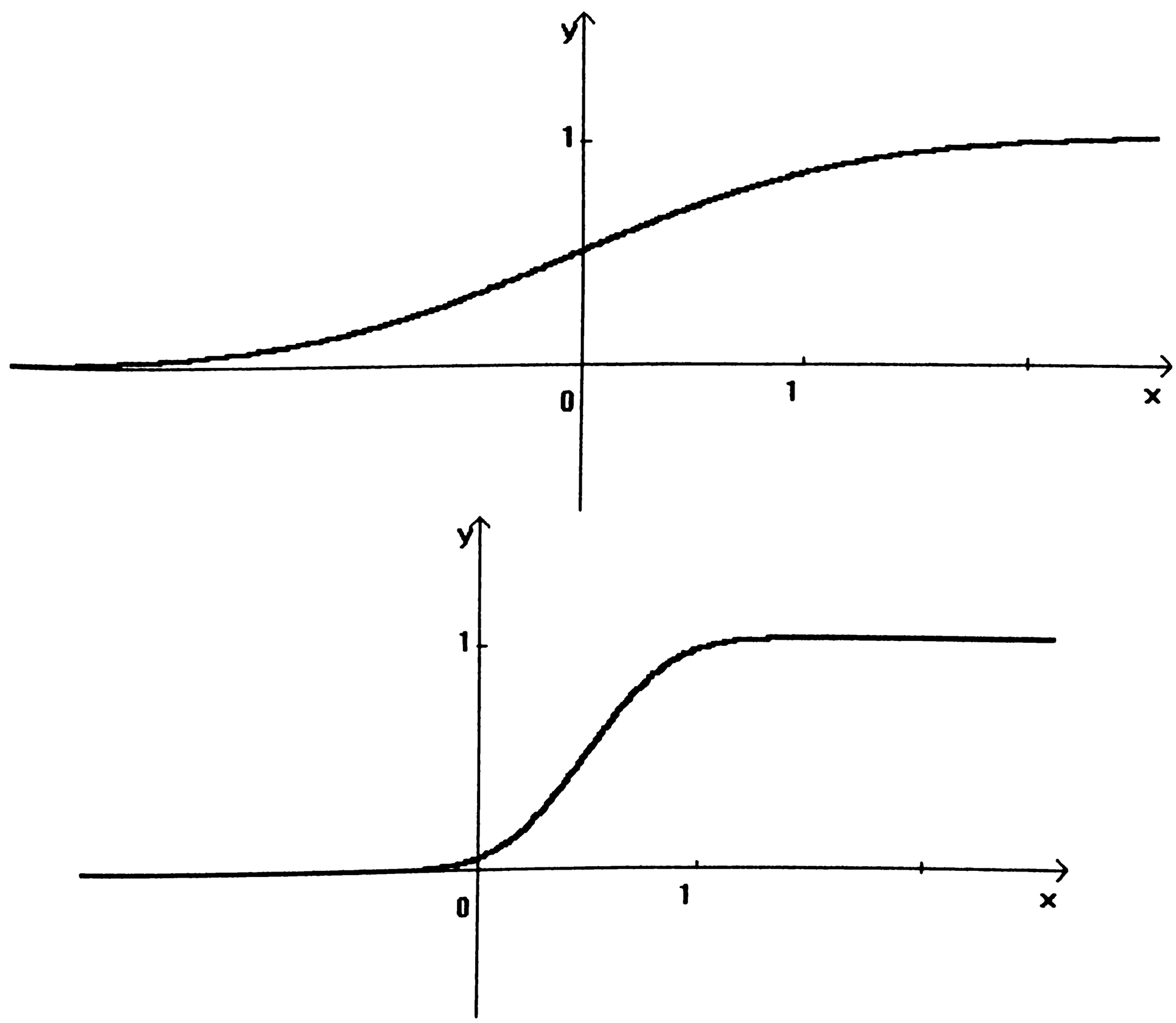
Параметрите a_{ji} и b_{ji}, j=0,1,2; i=1,2 ще определим така, че да е изпълнено условието: $-\frac{a_{ji}}{b_{ji}} = t_c$.

При това допълнително условие структурата (4.4) има вида:

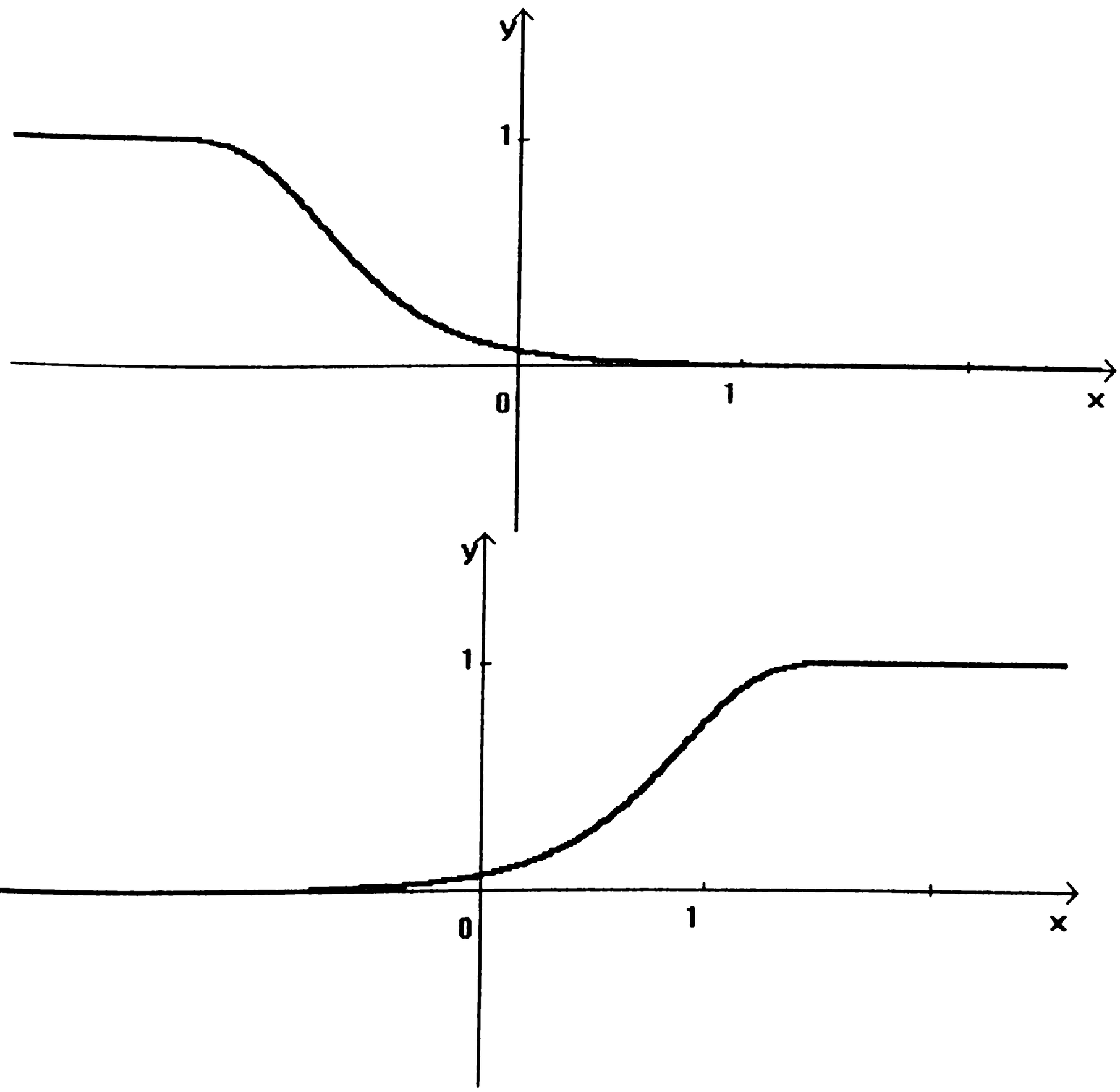
$$u_{ji}(t) = \begin{cases} 1-exp(-exp(b_{0i}(t_c-t)) & \text{при } j=0 \\ 1-exp(-exp(b_{1i}(t-t_c)) & \text{при } j=1 \\ 1-exp(-exp(b_{2i}(t-t_c)) & \text{при } j=2, \end{cases} \qquad (4.5)$$

където $b_{ji} > 0$, j=0,1,2 и $t_c \in [0,1]$ е фиксирано число.

Структурата полезност, дефинирана с (4.5) ще наричаме *complementary log-log структура полезност*.



Фигура 4.1 *Графика на стандартната Гаусова функция и на Гаусовата функция с параметри μ=0,5 и σ=0,3.*



Фигура 4.2 *Графика на complementary log-log функциите* ***1-exp(-exp(-2.7-3x))*** *и* ***1-exp(-exp(-2.7+3x))****.*

Стойностите на a_{ji} и b_{ji} могат да бъдат различни за различните подпопулации от потребители. Условията $b_{0i} > 0$ и $b_{2i} > 0$ са еквивалентни на твърдението, че полезността е строго намаляваща за случаите, в които агента отхвърля предположението за нормалност, и е строго растяща в случаите когато агентът е убеден, че разпределението, което описва изследваната популация е нормално.

Тъй като при по-тесни предположения за вида на изследваната популация се използват по-силни статистически методи, т.е. получават

се по-силни резултати от статистическия анализ, ще допуснем, че $u_{1i}(t)$ е растяща функция от t, т.е. $b_{1i} > 0$.

Да предположим, че h е популацията на потребителите, които са квалифицирани статистици, а i е популацията на потребителите, които са начинаещи статистици. Естествено е начинаещият статистик да разчита повече на мнението и съвета на агента, извършващ статистически анализ, отколкото квалифицираният статистик. Тогава некоректното решение ще има по-лоши последствия за популацията i отколкото за h. Или, което е почти същото, коректното решение е по-ценно за популацията i отколкото за популацията h. Затова ще изберем $b_{ji} > b_{jh}$, $j = 0,2$, $i = 1,\dots, g$, $h = 1,\dots, g$.

4.1.4 Достатъчни условия за монотонност на слабите правила

Оптималните правила за вземане на решение ще намерим като максимизираме очакваната полезност. Но ограничението до монотонни правила е коректно, ако няма немонотонни правила с по-голяма очаквана полезност.

В този параграф не се използва конкретната форма на структурата полезност. Достатъчно е да предполагаме за нея ограниченост, непрекъснатост и монотонност.

Разглеждаме i–тата подпопулация на потребителите. За правилата (4.1) ще намерим очакваната полезност $Eu_i(A_i^c, B_i^c(x))$ при фиксирана стойност на i.

$$Eu_i(A_i^c, B_i^c(x)) = \int_{A_i} \int_0^1 u_{0i}(t).w_i(x,t)dtdx +$$

$$+ \int_{A_i^c} \int_{B_i(x)} \int_0^1 u_{1i}(t).f_i(x,y,t)dtdydx +$$

$$+ \int_{A_i^c} \int_{B_i^c(x)} \int_0^1 u_{2i}(t).f_i(x,y,t)dtdydx,$$

където $w_i(x,t)$ е съвместната вероятностна плътност на X и T в подпопулацията i.

С $E_i(g(T)/x)$ ще означим функцията на регресия на $g(T)$ при фиксирано $X=x$. С $k_i(x,y)$ ще означим съвместната вероятностна плътност на X и Y в подпопулацията i. С $E_i(g(T)/x,y)$ ще означим функцията на регресия на $g(T)$ при фиксирани $X=x$ и $Y=y$. Тогава

$$Eu_i(A_i^c,B_i^c(x)) = E_i\{u_0(T)\} + \int_{A_i^c} E_i\{[u_1(T)-u_0(T)]/x\}\, q_i(x)dx +$$

$$+ \int_{A_i^c} \int_{B_i^c(x)} E_i\{[u_2(T)-u_1(T)]/x,y\}\, k_i(x,y)dydx \qquad (4.6)$$

Ще представим $k_i(x,y) = h_i(y/x).q_i(x)$. Тогава

$$Eu_i(A_i^c,B_i^c(x)) = E_i\{u_0(T)\} + \int_{A_i^c}\{E_i\{[u_1(T)-u_0(T)]/x\}+ \qquad (4.7)$$

$$+ \int_{B_i^c(x)} E_i\{[u_2(T)-u_1(T)]/x,y\}\, h_i(y/x)dy\}\, q_i(x)dx.$$

Следователно задачата се свежда до максимизиране на (4.7) за всяка стойност на i, $i = 1,2,\ldots,g$.

Ще намерим горна граница на очакваната полезност. За тази цел ще използваме почти очевидното твърдение:

***<u>Лема 1</u>**: Нека за функцията f(x) е изпълнено $\int |f(x)|dx < \infty$. Тогава за всяко множество S от стойности на x е в сила неравенството:*

$\int_S f(x)dx \le \int_{S_0} f(x)dx$, *където* $S_0 = \{x : f(x) \ge 0\}$.

От (4.7) следва, че за всяко $B^c(x)$ при фиксирано A^c е в сила неравенството

$$Eu(A^c,B^c(x)) \le E\{u_0(T)\} + \int_{A^c}\{E\{[u_1(T)-u_0(T)]/x\}+$$

$$+ \int_{B_0^c(x)} E\{[u_2(T)-u_1(T)]/x,y\}\, h\,(y/x)dy\}\, q(x)dx.$$

където

$$B_0^{\,c}(x) = \{y : E\{[u_2(T) - u_1(T)]/x, y\} \geq 0\}. \qquad (4.8)$$

От (4.7) след повторно прилагане на твърдението получаваме, че за всяко A^c е в сила неравенството:

$$Eu(A^c, B_0^{\,c}(x)) \leq E\{u_0(T)\} + \int_{A_0^c} \{E\{[u_1(T) - u_0(T)]/x\} +$$

$$+ \int_{B_0^c(x)} E\{[u_2(T) - u_1(T)]/x, y\}\, h(y/x)dy\}.q(x)dx.$$

където

$$A_o^{\,c} = \{x : E\{[u_1(T) - u_0(T)]/x\} + \int_{B_0^c(x)} E\{[u_2(T) - u_1(T)]/x, y\}\, h(y/x)dy \geq 0\}.$$

(4.9)

Ако функцията $E\{[u_2(T) - u_1(T)]/x, y\}$ от (6.8) е растяща по y за всяко x и функцията

$$E\{[u_1(T) - u_0(T)]/x\} + \int_{B_0^c(x)} E\{[u_2(T) - u_1(T)]/x, y\}\, h(y/x)dy$$

от (4.9) е растяща по x за всяко y, то (4.7) се максимизира върху множествата $A_0^{\,c} = [x_c, 1]$ и $B_0^{\,c}(x) = [y_c(x), 1]$, където x_c и $y_c(x)$ са корени съответно на уравненията (получени от (4.9) и (4.8)):

$$E\{[u_1(T) - u_0(T)]/x\} + \int_{B_0^c(x)} E\{[u_2(T) - u_1(T)]/x, y\}\, h(y/x)dy = 0$$

(4.10)

и

$$E\{[u_2(T) - u_1(T)]/x, y\} = 0. \qquad (4.11)$$

Едно друго по-слабо условие достатъчно за максимизиране на (4.7) върху множествата $A_0^{\,c} = [x_c, 1]$ и $B_0^{\,c}(x) = [y_c(x), 1]$ е функциите от (4.8) и (4.9) да сменят знака си от отрицателен към положителен точно един път.

Определение 1: *Казваме, че семейството от вероятностни плътности* ***f***(***x***/*t*), *където t е реален параметър, има монотонно частно на правдоподобие, ако за всяко* $t_2 < t_1$ *функциите на правдоподобие* $f(x/t_2)$ *и* $f(x/t_1)$ *са различни и частното* $\frac{f(x/t_1)}{f(x/t_2)}$ *е растяща (ненамаляваща) функция на* ***x***.

Членовете на еднопараметричното експоненциално семейство (нормалното, биномното, поасоновото и др.) и някои други широко използвани разпределения имат монотонно частно на правдоподобие.

Ще направим следната интерпретация:

Ако *t* е сигурността на агента, а ***x*** са наблюдаваните резултати от теста, при по-големи стойности на сигурността *t* наблюдаваните резултати ***x*** ще бъдат по-високи. Разумно е да искаме ако резултатите на един агент са по-високи от тези на друг, да можем да твърдим, че вероятно сигурността на първия е по-голяма.

В Бейсовия подход параметърът на разпределението се разглежда като случайна величина.

Определение 2: *Казваме, че класът от а-постериорни вероятностни плътности* ***f***(***t*** /***x***) *има монотонно частно на правдоподобие, ако за всяко* $x_2 < x_1$ *условните плътности* $f(t/x_2)$ *и* $f(t/x_1)$ *са различни и частното* $\frac{f(t/x_1)}{f(t/x_2)}$ *е растяща функция на* ***t***.

В сила е следната теорема (*виж* [*Chuang, Chen, Novick, 81*]):

Теорема: *Нека* ***f***(***t***) *е а-приорна плътността на сл.в.* ***T*** *и* ***f***(***x***/***t***) *има монотонно частно на правдоподобие. Тогава а-постериорната плътност* ***f***(***t*** /***x***) *има монотонно частно на правдоподобие.*

Следователно за всяко а-приорно разпределение на ***T*** наредбата, основана на ***X***, е същата, която би се получила от а-постериорното разпределение на ***T***.

До тук разгледахме две полезни свойства. За нашите цели ще използваме едно по-слабо условие.

Определение 3: *Казваме, че а-постериорното разпределение* ***f***(***t***/***x***) *е стохастически растящо в* ***x***, *ако от* $x_1 \geq x_2$ *следва че за функцията на разпределение е в сила* $F(t/x_1) \leq F(t/x_2)$, *за всяко* ***t***.

Всяко разпределение, имащо монотонно частно на правдоподобие е стохастически растящо [*Леман,79*]. Обратното твърдение не е вярно [*Chuang, Chen, Novick, 81*].

Лема 2: *Нека* ***T*** *и* ***Z*** *са случайни величини с реализации съответно* ***t*** *и* ***z***. *За всяка растяща функция* ***k***(***t***) *условното*

математическо очакване $E[k(T)/z]$ *е растяща функция в* z *тогава и само тогава, когато условната плътност* $f(t/Z=z)$ *е стохастически растяща в* z.

Достатъчното условие е доказано в [*Леман, 79*].

Необходимото условие е доказано в [*Chuang, Chen, Novick, 81*].

Лема 3: *Ако условната плътност* $v(t/x,y)$ *на* T *при условие* $X=x$ *и* $Y=y$ *е стохастически растяща в* x *и* y *и условната плътност* $h(y/x)$ *на* Y *при условие* $X=x$ *е стохастически растяща в* x, *то маргиналната условна плътност* $p(t/x)=\int_0^1 v(t/x,y).h(y/x)dy$ *е стохастически растяща в* x.

Доказателство:

$$1-P(t/x)=\int_t^1\int_0^1 v(z/x,y)h(y/x)dydz=\int_0^1[1-V(t/x,y)]h(y/x)dy$$

От условието и Лема 2 следва, че $1-P(t/x)$ е растяща в x, т.е. $P(t/x)$ е намаляваща в x. Следователно $p(t/x)$ е стохастически растяща в x.

Лема 4: *Ако функцията* $k(x,y)$ *е строго растяща по* x *и* y, *то зависимостта на* y *от* x, *определена чрез* $\{(x,y):k(x,y)=c,\ c\in R\}$ *е намаляваща функция по* x.

Лемата е доказана във [*Vos, 94, стр.51*].

Теорема 4.1: *Нека са в сила условията:*

- *функциите* $u_j(t)-u_{j-1}(t), j=1,2$ *са строго растящи по* t, (4.12)
- *условната функция на разпределение* $V(t/x,y)$ *е строго намаляваща в* x *и* y *за всяко* t, (4.13)
- *условната функция на разпределение* $H(y/x)$ *е строго намаляваща в* x *за всяко* y. (4.14)

Тогава оптималното слабо правило за вземане на решение за нормалност на вероятностното разпределение на изследваната популация е монотонно.

За частния случай, когато функцията полезност има линеен вид, теоремата е доказана във [*Vos, 94, стр.52-53*]. Ние ще я докажем по аналогичен начин за произволна непрекъсната и ограничена функция полезност.

Доказателство:

От условието на теоремата и Лема 2 следва, че

$E\{[u_2(T)-u_1(T)]/x,y\}$ е растяща в y за всяко x и е растяща в x за всяко y. Следователно множеството $B_0^c(x)$ приема формата $[y_c(x),1]$ за всички стойности на x.

От условието на теоремата, Лема 2 и Лема 3 следва, че $E\{[u_1(T)-u_0(T)]/x\}$ е растяща функция по x.

$E\{[u_2(T)-u_1(T)]/x,y\}$ е растяща по x и y и неотрицателна при $y \geq y_c(x)$ за всяко x (*виж* (4.8)), а $h(y/x) \geq 0$.

Нека $x_2 > x_1$. Тогава от Лема 4 следва, че

$$\int_{y_c(x_2)}^{1} E\{[u_2(T)-u_1(T)]/x_2,y\}h(y/x_2)dy -$$

$$- \int_{y_c(x_1)}^{1} E\{[u_2(T)-u_1(T)]/x_1,y\}h(y/x_1)dy >$$

$$\int_{y_c(x_1)}^{1} E\{[u_2(T)-u_1(T)]/x_2,y\}h(y/x_2)dy -$$

$$- \int_{y_c(x_1)}^{1} E\{[u_2(T)-u_1(T)]/x_1,y\}h(y/x_1)dy >$$

$$\int_{y_c(x_1)}^{1} E\{[u_2(T)-u_1(T)]/x_1,y\}h(y/x_2)dy -$$

$$- \int_{y_c(x_1)}^{1} E\{[u_2(T)-u_1(T)]/x_1,y\}.h(y/x_1)dy =$$

$$\int_{y_c(x_1)}^{1} E\{[u_2(T)-u_1(T)]/x_1,y\}.[h(y/x_2)-h(y/x_1)]dy =$$

$$\int_{0}^{1} I_{[y_c(x_1),1]}(y).E\{[u_2(T)-u_1(T)]/x_1,y\}.[h(y/x_2)-h(y/x_1)]dy$$

$$> 0,$$

където $I_{[y_c(x),1]}$ е индикаторът на множеството $[y_c(x),1]$. Използвахме, че плътността $h(y/x)$ е стохастически растяща в x (*виж* (4.14)).

Следователно

$$E\{[u_1(T)-u_0(T)]/x\}+\int_{y_c(x)}^{1} E\{[u_2(T)-u_1(T)]/x\ ,y\}h(y/x\ )dy$$

е растяща функция. Следователно множеството A_0^c приема формата $[x_c,1]$.

Следствие 4.1: *Ако*

- *функциите* $u_j(t)-u_{j-1}(t),\ j=1,2$ *са строго растящи в* t, (4.15)
- *условната плътност* $v(t/x,y)$ *има монотонно частно на правдоподобие в* x *и* y, (4.16)
- *условната плътност* $h(y/x)$ *има монотонно частно на правдоподобие в* x, (4.17)

то оптималното слабо правило за вземане на решение за нормалност на популационното разпределение е монотонно.

Разглеждаме Гаусовата структура полезност (4.3).

$$u_{ji}(t)=\begin{cases}\phi[(\mu_{0i}-t)/\sigma_{0i}] & \text{при } j=0\\ \phi[(t-\mu_{1i})/\sigma_{1i}] & \text{при } j=1\\ \phi[(t-\mu_{2i})/\sigma_{2i}] & \text{при } j=2,\end{cases}$$

където $\phi[.]$ е стандартната нормална функция на разпределение, а μ_{ji} и $\sigma_{ji}>0$, j=0,1,2; i=1,2, са съответно средното и стандартното отклонение на нормалното разпределение.

Тогава

$$u_1(t)-u_0(t)=\phi(\frac{t-\mu_1}{\sigma_1})-\phi(\frac{\mu_0-t}{\sigma_0}) \quad \text{и}$$

$$\frac{d}{dt}[u_1(t)-u_0(t)]=\frac{1}{\sigma_1}\varphi(\frac{t-\mu_1}{\sigma_1})+\frac{1}{\sigma_0}\varphi(\frac{\mu_0-t}{\sigma_0})\geq 0 \qquad \forall t\in[0,1],$$

където с $\varphi(.)$ сме означили плътността на стандартното нормално разпределение. Тогава

$$u_2(t)-u_1(t)=\phi(\frac{t-\mu_2}{\sigma_2})-\phi(\frac{t-\mu_1}{\sigma_1})$$ и

$$\frac{d}{dt}[u_2(t)-u_1(t)]=\frac{1}{\sigma_2}\varphi(\frac{t-\mu_2}{\sigma_2})-\frac{1}{\sigma_1}\varphi(\frac{t-\mu_1}{\sigma_1}).$$

Корените на уравнението

$$\frac{1}{\sigma_2^2}e^{-\frac{1}{2}\frac{1}{\sigma_2^2}(t-\mu_2)^2}-\frac{1}{\sigma_1^2}e^{-\frac{1}{2}\frac{1}{\sigma_1^2}(t-\mu_1)^2}=0$$

са

$$t_1=\frac{\sigma_1^{-2}\mu_1-\sigma_2^{-2}\mu_2-\sqrt{(\frac{\mu_1-\mu_2}{2})^2+(\sigma_2^2-\sigma_1^2)(\ln\sigma_2-\ln\sigma_1)}}{\sigma_1^{-2}-\sigma_2^{-2}}$$

$$t_2=\frac{\sigma_1^{-2}\mu_1-\sigma_2^{-2}\mu_2+\sqrt{(\frac{\mu_1-\mu_2}{2})^2+(\sigma_2^2-\sigma_1^2)(\ln\sigma_2-\ln\sigma_1)}}{\sigma_1^{-2}-\sigma_2^{-2}}$$

<u>Следствие 4.2:</u> *За Гаусовата структура полезност* (4.3)

$$u_{ji}(t)=\begin{cases}\phi[(\mu_{0i}-t)/\sigma_{0i}] & \text{при } j=0\\ \phi[(t-\mu_{1i})/\sigma_{1i}] & \text{при } j=1\\ \phi[(t-\mu_{2i})/\sigma_{2i}] & \text{при } j=2,\end{cases}$$

където $\phi[.]$ *е стандартната нормална функция на разпределение, а* μ_{ji} *и* $\sigma_{2i}>\sigma_{1i}>0$, $i=1,2$ *са съответно средните и стандартните отклонения, оптималното слабо правило за вземане на решение за нормалност на популационното разпределение е монотонно, ако са изпълнени условията:*

- $1<t_1$ *или* $t_2<0$, *където*

$$t_1=\frac{\sigma_1^{-2}\mu_1-\sigma_2^{-2}\mu_2-\sqrt{(\frac{\mu_1-\mu_2}{2})^2+(\sigma_2^2-\sigma_1^2)(\ln\sigma_2-\ln\sigma_1)}}{\sigma_1^{-2}-\sigma_2^{-2}}$$

$$t_2=\frac{\sigma_1^{-2}\mu_1-\sigma_2^{-2}\mu_2+\sqrt{(\frac{\mu_1-\mu_2}{2})^2+(\sigma_2^2-\sigma_1^2)(\ln\sigma_2-\ln\sigma_1)}}{\sigma_1^{-2}-\sigma_2^{-2}}$$

- *условната плътност* $v(t/x,y)$ *има монотонно частно на правдоподобие в* x *и* y,
- *условната плътност* $h(y/x)$ *има монотонно частно на правдоподобие в* x.

Следствие 4.3: За Гаусовата структура полезност (4.3)

$$u_{ji}(t) = \begin{cases} \phi[(\mu_{0i} - t)/\sigma_{0i}] & \text{при } j=0 \\ \phi[(t - \mu_{1i})/\sigma_{1i}] & \text{при } j=1 \\ \phi[(t - \mu_{2i})/\sigma_{2i}] & \text{при } j=2, \end{cases}$$

където $\phi[.]$ *е стандартната нормална функция на разпределение, а* μ_{ji} *и* $\sigma_{0i} > 0$, $\sigma_{1i} > \sigma_{2i} > 0$, $i=1,2$ *са съответно средните и стандартните отклонения, оптималното слабо правило за вземане на решение за нормалност на популационното разпределение е монотонно, ако са изпълнени условията:*

- $t_1<0<1<t_2$, *където*

$$t_1 = \frac{\sigma_1^{-2}\mu_1 - \sigma_2^{-2}\mu_2 - \sqrt{(\frac{\mu_1 - \mu_2}{2})^2 + (\sigma_2^2 - \sigma_1^2)(\ln\sigma_2 - \ln\sigma_1)}}{\sigma_1^{-2} - \sigma_2^{-2}}$$

$$t_2 = \frac{\sigma_1^{-2}\mu_1 - \sigma_2^{-2}\mu_2 + \sqrt{(\frac{\mu_1 - \mu_2}{2})^2 + (\sigma_2^2 - \sigma_1^2)(\ln\sigma_2 - \ln\sigma_1)}}{\sigma_1^{-2} - \sigma_2^{-2}}$$

- *условната плътност* $v(t/x,y)$ *има монотонно частно на правдоподобие в* x *и* y,
- *условната плътност* $h(y/x)$ *има монотонно частно на правдоподобие в* x.

Следствие 4.4: *За Гаусовата структура полезност* (4.3)

$$u_{ji}(t) = \begin{cases} \phi[(\mu_{0i} - t)/\sigma_{0i}] & \text{при } j=0 \\ \phi[(t - \mu_{1i})/\sigma_i] & \text{при } j=1 \\ \phi[(t - \mu_{2i})/\sigma_i] & \text{при } j=2, \end{cases}$$

където $\phi[.]$ *е стандартната нормална функция на разпределение, а* μ_{ji} *и* $\sigma_{0i} > 0$, $\sigma_i > 0$, $i=1,2$ *са съответно средните и*

стандартните отклонения, оптималното слабо правило за вземане на решение за нормалност на популационното разпределение е монотонно, ако:

- $\mu_{1i} < \mu_{2i}$ *и* $\frac{\mu_{1i}+\mu_{2i}}{2} < 0$, *или* $\mu_{1i} > \mu_{2i}$ *и* $\frac{\mu_{1i}+\mu_{2i}}{2} > 1$, $i=1,2$,
- *условната плътност* $v(t/x,y)$ *има монотонно частно на правдоподобие в* x *и* y,
- *условната плътност* $h(y/x)$ *има монотонно частно на правдоподобие в* x.

<u>Следствие 4.5:</u> *За структурата полезност* (4.5)

$$u_{ji}(t) = \begin{cases} 1-\exp(-\exp(b_{0i}(t_c-t)) & \text{при } j=0 \\ 1-\exp(-\exp(b_{1i}(t-t_c)) & \text{при } j=1 \\ 1-\exp(-\exp(b_{2i}(t-t_c)) & \text{при } j=2, \end{cases}$$

където $b_{ji} > 0$, $j=0,1,2$, и $t_c \in [0,1]$ *е фиксирано, оптималното слабо правило за вземане на решение за нормалност на популационното разпределение е монотонно, ако са изпълнени условията:*

- $b_{2i} > b_{1i}$,
- *условната плътност* $v(t/x,y)$ *има монотонно частно на правдоподобие в* x *и* y,
- *условната плътност* $h(y/x)$ *има монотонно частно на правдоподобие в* x.

Естествено е функцията полезност да приема по-големи стойности когато агентът е убеден, че разпределението на изследваната популация е нормално, отколкото когато агента подозира (предполага), че разпределението на изследваната популация е нормално. Следователно условието $b_2 > b_1$ е едно естествено ограничение.

Доказателството на *Следствие 5* непосредствено следва от това, че при $b_2 > b_1$ функцията $u_2(t) - u_1(t)$ сменя знака си от отрицателен към положителен точно един път в интервала [01] и от свойствата на математическото очакване.

Да резюмираме полученото:

Множествата $B_0^c(x)$ и A_0^c се определят по следния начин:

Първо, за всички стойности на x множествата $B_0^c(x)$ се определят чрез (4.8). След това множеството A_0^c се определя чрез (4.9).

4.1.5 Достатъчно условие за силна монотонност

Когато работим със силни правила множеството $B_0^c(x)$ не зависи от x. От (4.8) следва, че върху вероятностния модел трябва да наложим следните условия:

- за условната плътност на T при условие, че $X=x$ и $Y=y$, и за условната плътност на T при условие $Y=y$:
  $v(t/x,y)=k(t/y)$.

Т.е. След като е изпълнено $Y=y$ наблюдението $X=x$ не носи допълнителна информация.

- $h(y/x)=h(y)$.

Тогава

$$Eu(A^c,B^c) = E\{u_0(T)\} + \int_{A^c}\{E\{[u_1(T)-u_0(T)]/x\}+$$

$$+\int_{B^c}E\{[u_2(T)-u_1(T)]/y\}h(y)dy\}\,q(x)dx$$

Да предположим, че $h(y)>0$ и $q(x)>0$. Тогава

$$Eu(A^c,B^c) \le E\{u_0(T)\} + \int_{A_0^c}\{E\{[u_1(T)-u_0(T)]/x\}+$$

$$+\int_{B_0^c(x)}E\{[u_2(T)-u_1(T)]/y\}h(y)dy\}\,q(x)dx,$$

където $B_0^c=\{y: E[u_2(T)-u_1(T)/y]\ge 0\}$,

$$A_0^c=\{x: E[u_1(t)-u_0(t)/x]+\int_{B_0^c}E[u_2(t)-u_1(t)/y]h(y)dy\ge 0\}.$$

Тогава е в сила следната теорема:

Теорема 4.2: *Нека за вероятностния модел да са изпълнени условията:*

- $u_1(t)-u_0(t)$ *и* $u_2(t)-u_1(t)$ *са растящи функции на* t, (4.18)
- *условната плътност* $k(t/y)$ *е стохастически растяща в* y, (4.19)

- *условната плътност* $m(t/x)$ *е стохастически растяща в* x.

(4.20)

Тогава оптималното силно правило за вземане на решение за нормалност на популационното разпределение е монотонно.

<u>*Доказателство:*</u>

От (4.19) и Лема 2 следва, че условното математическо очакване $E[u_1(t)-u_0(t)/y]$ е растяща функция на y. Следователно множеството $B_0^c=\{y: E[u_2(T)-u_1(T)/y]\geq 0\}$ приема формата $B_0^c=[y_c,1]$, където y_c е решение на уравнението $E[u_2(t)-u_1(t)/y]=0$.

От (4.20) и Лема 2 следва, че условното математическо очакване $E[u_1(t)-u_0(t)/x]$ е растяща функция на x. Определеният интеграл $\int_{y_c}^{1} E[u_2(t)-u_1(t)/y].h(y)dy$ не зависи от x.

Следователно множеството

$$A_0^c=\{x: E[u_1(t)-u_0(t)/x]+\int_{y_c}^{1} E[u_2(t)-u_1(t)/y]h(y)dy\geq 0\}$$

приема формата $A_0^c=[x_c,1]$, където x_c е решение на уравнението $E[u_1(t)-u_0(t)/x]+\int_{y_c}^{1} E[u_2(t)-u_1(t)/y]h(y)dy=0$

Следователно оптималното силно правило за вземане на решение за нормалност на популационното разпределение е монотонно.

Оптималното решение на проблема е двойката (x_{ci}, y_{ci}) за която очакваната полезност достига своя максимум.

Следователно задачата се свежда до максимизиране на

$$Eq(T/x_{ci},y_{ci}) = E_i\{u_0(T)\}+\int_{x_{ci}}^{1} E_i\{[u_1(T)-u_0(T)]/x\}.q_i(x)dx+$$

(4.21)

$$+\int_{x_{ci}}^{1}\int_{y_{ci}}^{1} E_i\{[u_2(T)-u_1(T)]/y\}\, q_i(x)h_i(y)dydx.$$

за всяка стойност на i, $i = 1,2,\ldots,g$.

За да получим оптималните стойности на x_{ci} и y_{ci}, максимизиращи очакваната полезност ще диференцираме $Eu_i(T/x_{ci},y_{ci})$ по x_{ci} и y_{ci}.

$$\frac{\partial}{\partial y_{ci}} Eu_i(T/x_{ci},y_{ci})=$$

$$=\frac{\partial}{\partial y_{ci}}\int_{x_{ci}}^{1}\int_{y_{ci}}^{1} E_i\{[u_2(T)-u_1(T)]/y\}\, q_i(x)h_i(y)dydx =$$

$$=h_i(y_{ci})\int_{x_{ci}}^{1} E_i\{[u_2(T)-u_1(T)]/y_{ci}\}\, q_i(x)dx =$$

$$=E_i\{[u_2(T)-u_1(T)]/y_{ci}\}\,.\,h_i(y_{ci})\int_{x_{ci}}^{1} q_i(x)dx$$

Предполагаме, че $q_i(x)>0$ и $h_i(y)>0$.

Тогава след приравняване на 0 на производната получаваме равенството:

$$E_i\{[u_2(T)-u_1(T)]/y_{ci}\}=0 \qquad (4.22)$$

По аналогичен начин след диференциране по x_{ci} и предположението за маргиналната плътност на X $q_i(x)>0$ получаваме равенството:

$$E_i\{[u_1(T)-u_0(T)]/x_{ci}\}+\int_{y_{ci}}^{1} E_i\{[u_2(T)-u_1(T)]/y\}.h_i(y)dy=0 \qquad (4.23)$$

Решенията на системата интегрални уравнения (4.22)-(4.23) са търсените оптимални стойности на x_{ci} и y_{ci}. За дадени функции на регресия и вероятностни плътности, оптималната стратегия се представя от системата уравнения (4.22)-(4.23).

Да отбележим, че ако условията за монотонност не са строги или не е в сила допускането $h_i(y)>0$ или $q_i(x)>0$ в околност на решението, то оптималната стратегия може да не е единствена.

4.2 Симетричност

За да формира своята вяра за симетричност на вероятностното разпределение на изследваната популация агентът използва статистически критерий за симетричност. За случаите, за които хипотезата за симетричност не може да бъде отхвърлена, може да се провери и хипотеза за нормалност на разпределението. В случаите, в които хипотезата за симетричност се отхвърля, може да се продължи с премахване на съмнителни наблюдения и с трансформации на данните.

Нека X е $1-p$, където p е p-стойността на статистиката на критерия за симетричност, T е сигурността на агента. Очевидно X се наблюдава, а T не може да се наблюдава. Предполагаме, че X и T са непрекъснати случайни величини и връзката между тях може да се представи със съвместното им вероятностно разпределение $k(x,t)$.

Решението за симетричност зависи от прагова точка t_c, която е предварително определена и известна на агента. Например t_c=0,95. Ако истинската стойност T надвишава праговата стойност t_c, то агентът е убеден в симетричността на изследваната популация. В противен случай агентът е убеден в асиметричността на разпределението.

Наличието на популации от потребители с различна статистическа компетентност води до различни стойности на праговата точка x_c, които ще означаваме с x_{ci}. Праговата точка t_c на T е една и съща за всяка популация от потребители.

Тогава монотонните правила за вземане на решение ще дефинираме по следния начин:

$$\delta(x) = \begin{cases} a_0 & \text{при } X<x_{ci} \\ a_1 & \text{при } X\geq x_{ci}, \end{cases}$$

където състоянията a_0 и a_1 означават вярата на агента съответно в симетричността и в асиметричността на разпределението.

За всяко от възможните действия предполагаме, че функцията полезност има следния вид:

$$u_i(t) = \begin{cases} u_0(t) & \text{при } a=a_0 \\ u_1(t) & \text{при } a=a_1 \end{cases}$$

Тогава очакваната полезност е функция на праговата стойност x_c.

$$Eu_i(t) = \int_0^{x_c}\int_0^1 u_0(t)k(x,t)dtdx + \int_{x_c}^1\int_0^1 u_1(t)k(x,t)dtdx =$$

$$= \int_0^1\int_0^1 u_0(t)k(x,t)dtdx + \int_{x_c}^1\int_0^1 [u_1(t)-u_0(t)].k(x,t)dtdx$$

Първото събираемо не зависи от x_c. Следователно решението зависи само от второто събираемо. Нека $k(x,t) = p(t/x).q(x)$. Тогава

$$Eu_i(t) = E\{u_0(T)\} + \int_{x_c}^1 E\{[u_1(T)-u_0(T)]/x\}q(x)dx$$

За да намерим достатъчно условие за монотонност на правилата ще използваме *Лема 1*. Тогава е в сила следната теорема:

Теорема 4.3: *Ако са изпълнени условията:*

- *функцията* $u_1(t)-u_0(t)$ *е строго растяща,*
- *функцията* $P(t/x)$ *е строго намаляваща в* x *за всяко* t,

то правилото за вземане на решение за симетричност на популационното разпределение е монотонно.

Доказателството на *Теорема 4.3* е аналогично на доказателството на *Теорема 4.1* от *4.1.4*.

Следователно, при предположение, че $q(x_c) > 0$ решението на разглеждания проблем е решението спрямо x_c на уравнението $E\{[u_1(T)-u_0(T)]/x_c\}=0$.

4.3 Комбиниран модел

Ще разгледаме проблема за формиране на вярата на агента относно вида на вероятностното разпределение на изследваната популация при предположение, че то е непрекъснато. За тази цел се използват статистически тест за симетричност, статистически тест за нормалност (пре-тест за нормалност) и запитване за мнението на потребителя за вида на популационното разпределение (пост-тест за нормалност). Резултатите от тези тестове са съответно **1-p_1**, **1-p_2**, където p_1 и p_2 са p-стойностите на статистиките на двата теста, и степента на сигурност на потребителя, изразена като число в интервала [0,1].

Нека с X да означим наблюдавания резултат от теста за симетричност, с Y - наблюдавания резултат от пре-теста за нормалност, с Z - наблюдавания резултат от пост-теста за нормалност, а с T - сигурността на агента за вземане на решение за вида на разпределението, която не може да се наблюдава. Нека да предположим, че X, Y, Z и T са непрекъснати случайни величини със съвместно вероятностно разпределение $f(x,y,z,t)$.

Правилото за вземане на решение $\delta(X,Y,Z)$ определя за всяка възможна реализация (x,y,z) на случайния вектор (X,Y,Z) кое действие ще бъде извършено.

Най-общата форма на слабото правило за вземане на решение в този случай е:

$$\{(x,y,z):\delta(x,y,z)=a_0\}=A\times[0,1]\times[0,1]$$

$$\{(x,y,z):\delta(x,y,z)=a_{10}\}=A^c\times B(x)\times[0,1]$$

$$\{(x,y,z):\delta(x,y,z)=a_{110}\}=A^c\times B^c(x)\times C(x,y)$$

$$\{(x,y,z):\delta(x,y,z)=a_{111}\}=A^c\times B^c(x)\times C^c(x,y)$$

където:

A е множеството от стойности x, водещи до отхвърляне на хипотезата за симетричност на разпределението,

$B(x)$ е множеството от стойности на пре-теста y, водещи до отхвърляне на хипотезата за нормалност,

$C(x,y)$ е множеството от стойности z на пост-теста, водещи до отхвърляне на предположението за нормалност,

A^c, $B^c(x)$, $C^c(x,y)$ са допълненията съответно на множествата A, $B(x)$, $C(x,y)$.

Монотонното слабо правило за вземане на решение има вида:

$$\delta(x,y,z)=\begin{cases} a_0 \text{ при } & X<x_c,\ Y\in[0,1],\ Z\in[0,1] \\ a_{10} \text{ при } & X\geq x_c,\ Y<y_c(x),\ Z\in[0,1] \\ a_{110} \text{ при } & X\geq x_c,\ Y\geq y_c(x),\ Z<z_c(x,y) \\ a_{111} \text{ при } & X\geq x_c,\ Y\geq y_c(x),\ Z\geq z_c(x,y), \end{cases}$$

където

• a_j, j=0,10,110,111 са следните състояния на вярата на агента:

a_0 - агентът отхвърля предположението за симетричност на разпределението, което описва изследваната популация,

a_{10} - агентът отхвърля предположението за нормалност на разпределението, което описва изследваната популация,

a_{110} - агентът предполага, подозира, че разпределението на изследваната популация е нормално,

a_{111} - агентът е убеден (сигурен), че разпределението, което описва изследваната популация е нормално.

• x_c, y_c, z_c са праговите точки за X, Y и Z.

Предполагаме, че структурата полезност има следната адитивна форма:

$$u_{ijk}(t)=w_1u_i^{(s)}(t)+w_2u_j^{(n)}(t)+w_3u_k^{(u)}(t),$$

където $u_i^{(s)}(t)$, $u_j^{(n)}(t)$ и $u_k^{(u)}(t)$ са функциите полезност за вярата на агента съответно за симетричност и за нормалност, а w_1, w_2 и w_3 са неотрицателни теглови коефициенти.

Структурата полезност може да бъде представена и по следния начин:

$$u_{ijk}(t)=w_1u_i^{(s)}(t)+w_2u_j^{(n)}(t)+w_3u_k^{(u)}(t),$$

където $w_i \geq 0$, $i=1,2,3$, и $\sum_{i=1}^{3} w_i = 1$.

По този начин дефинирахме една по-реалистична структура полезност

$$u(t)=\begin{cases} u_0 \text{ при } & X<x_c,\ Y\in[0,1],\ Z\in[0,1] \\ u_{10} \text{ при } & X\geq x_c,\ Y<y_c(x),\ Z\in[0,1] \\ u_{110} \text{ при } & X\geq x_c,\ Y\geq y_c(x),\ Z<z_c(x,y) \\ u_{111} \text{ при } & X\geq x_c,\ Y\geq y_c(x),\ Z\geq z_c(x,y), \end{cases}$$

т.е.

$$u(t)=\begin{cases} w_1u_0^{(s)}(t) & \text{при } X<x_c,\ Y\in[0,1],\ Z\in[0,1] \\ w_1u_1^{(s)}(t)+w_1u_0^{(n)}(t) & \text{при } X\geq x_c,\ Y<y_c(x),\ Z\in[0,1] \\ w_1u_1^{(s)}(t)+w_1u_1^{(n)}(t)+w_1u_0^{(u)}(t) & \text{при } X\geq x_c,\ Y\geq y_c(x),\ Z<z_c(x,y) \\ w_1u_1^{(s)}(t)+w_1u_1^{(n)}(t)+w_1u_1^{(u)}(t) & \text{при } X\geq x_c,\ Y\geq y_c(x),\ Z\geq z_c(x,y). \end{cases}$$

Ще изчислим очакваната полезност.

$$\mathbf{E}u(A^c,B^c(x),C^c(x,y)) =$$

$$=\int_A \int_0^1 \int_0^1 \int_0^1 w_1 u_0^{(s)}(t).f(x,y,z,t)dt.dz.dy.dx +$$

$$+\int_{A^c} \int_{B(x)} \int_0^1 \int_0^1 u_{10}(t).f(x,y,z,t)dt.dz.dy.dx +$$

$$+\int_{A^c} \int_{B^c(x)} \int_{C(x,y)} \int_0^1 u_{110}(t).f(x,y,z,t)dt.dz.dy.dx +$$

$$+\int_{A^c} \int_{B^c(x)} \int_{C^c(x,y)} \int_0^1 u_{111}(t).f(x,y,z,t)dt.dz.dy.dx =$$

$$=w_1 Eu_0^{(s)}(T)+\int_{A^c} E\{[u_{10}(T)-w_1 u_0^{(s)}(T)]/x\}.q(x).dx +$$

$$+\int_{A^c} \int_{B^c(x)} E\{[u_{110}(T)-u_{10}(T)]/x,y\}.h(y/x).q(x).dy.dx +$$

$$+\int_{A^c} \int_{B^c(x)} \int_{C^c(x,y)} E\{[u_{111}(T)-u_{110}(T)]/x,y,z\}\, g(z/x,y)h(y/x)q(x)dzdy.dx =$$

$$=w_1 Eu_0^{(s)}(T)+\int_{A^c}\{E[u_{10}(T)-w_1 u_0^{(s)}(T)/x]+$$

$$+\int_{B^c(x)}\{E[u_{110}(T)-u_{10}(T)/x,y]+$$

$$+\int_{C^c(x,y)} E[u_{111}(T)-u_{110}(T)/x,y,z]g(z/x,y)dz\}\, h(y/x)dy\}\, q(x)dx$$

Прилагайки *Лема1* получаваме, че за всяко x и за всяко y една горна граница на очакваната полезност може да се намери върху множеството

$$C_0^c(x,y)=\{z: E[u_{111}(T)-u_{110}(T)/x,y,z]\geq 0\}.$$

Използвайки, че $h(y/x)\geq 0$ по аналогичен начин получаваме, че за всяко x една горна граница на очакваната полезност може да се намери върху множеството

$$B_0^c(x) = \{y : E[u_{110}(T) - u_{10}(T)/x, y] +$$

$$+ \int_{C_0^c(x,y)} E[u_{111}(T) - u_{110}(T)/x, y, z].g(z/x, y)dz \geq 0\}.$$

Използвайки, че $q(x) \geq 0$ аналогично получаваме, че горна граница на очакваната полезност може да се намери върху множеството

$$A_0^c = \{x : E[u_{10}(T) - w_1 u_0^{(s)}(T)/x] +$$

$$+ \int_{B_0^c(x)} \{E[u_{110}(T) - u_{10}(T)/x, y]_+$$

$$+ \int_{C_0^c(x,y)} E[u_{111}(T) - u_{110}(T)/x, y, z].g(z/x, y)dz\} h(y/x).dy \geq 0\}$$

Ще намерим достатъчно условие за слабата монотонност. В случая на слаби монотонни правила множествата $C_0^c(x, y)$, $B_0^c(x)$ и A_0^c приемат вида:

$$C_0^c(x, y) = [z_c(x, y), 1]$$

$$B_0^c(x) = [y_c(x), 1]$$

$$A_0^c = [x_c, 1].$$

Нека $v(t/x, y, z)$ и $r(t/x, y)$ са условните вероятностни плътности съответно на ***T*** при ***X***=*x*, ***Y***=*y*, ***Z***=*z* и на ***T*** при ***X***=*x*, ***Y***=*y*.

Нека $V(t/x, y, z)$, $G(z/x, y)$ и $H(y/x)$ са функциите на вероятностни разпределения съответно на ***T*** при ***X***=*x*, ***Y***=*y*, ***Z***=*z* , на ***Z*** при ***X***=*x*, ***Y***=*y* и на ***Y*** при ***X***=*x*.

Тогава следната теорема гарантира слабата монотонност на правилата за вземане на решения:

<u>Теорема 4.4</u>: *Нека са изпълнени следните условия:*

- *функциите* $u_{10}(t) - u_0(t)$, $u_{110}(t) - u_{10}(t)$, $u_{111}(t) - u_{110}(t)$ *са строго растящи.*
- $V(t/x, y, z)$ *е строго намаляваща в* ***x***, ***y*** *и* ***z*** *за всяко* ***t***.
- $G(z/x, y)$ *е строго намаляваща в* ***x*** *и* ***y*** *за всяко* ***z***.
- $H(y/x)$ *е строго намаляваща в* ***x*** *за всяко* ***y***.

Тогава съществува слабо монотонно решение на проблема за вземане на решение за нормалност на вероятностното разпределение на изследваната непрекъсната популация.

Доказателство:

Доказателството на теоремата е аналогично на доказателството на теоремата от 4.1.4.

Когато работим със силни монотонни правила множеството $B_0^c(x)$ не зависи от x, а множеството $C_0^c(x,y)$ не зависи от x и y. Затова върху вероятностния модел трябва да наложим следните условия:

- $v(t/x,y,z)=v(t/z)$,
- $m(t/x,y)=m(t/y)$,
- $g(z/x,y)=g(z)$,
- $h(y/x)=h(y)$.

Оптималното силно решение на проблема за вземане на решение за нормалност на популационното разпределение е тройката (x_c,y_c,z_c) за която очакваната полезност

$$Eu(t/x_c,y_c,z_c)=w_1Eu_0^{(s)}(T)+\int_{x_c}^{1}E\{[u_{10}(T)-u_0(T)]/x\}q(x)dx+$$

$$+\int_{x_c}^{1}\int_{y_c}^{1}E\{[u_{110}(T)-u_{10}(T)]/y\}h(y).q(x)dydx+$$

$$+\int_{x_c}^{1}\int_{y_c}^{1}\int_{z_c}^{1}E\{[u_{111}(T)-u_{110}(T)/z\}g(z).h(y).q(x).dzdydx$$

достига своя максимум.

За да получим оптималното решение ще диференцираме $Eu(t/x_c,y_c,z_c)$ по x_c, y_c и z_c.

$$\frac{\partial}{\partial z_c}Eu(t/x_c,y_c,z_c)=$$

$$=\int_{x_c}^{1}\int_{y_c}^{1}E\{[u_{111}(T)-u_{110}(T)/z_c\}g(z_c).h(y).q(x).dydx=$$

$$= g(z_c).E\{[u_{111}(T) - u_{110}(T)/z_c\}.\int_{x_c}^{1}\int_{y_c}^{1} h(y).q(x).dydx =$$

$$= g(z_c).E\{[u_{111}(T) - u_{110}(T)/z_c\}.\int_{y_c}^{1} h(y)dy.\int_{x_c}^{1} q(x).dx.$$

$$\frac{\partial}{\partial y_c} Eu(t/x_c, y_c, z_c) =$$

$$= \int_{x_c}^{1} E\{[u_{110}(T) - u_{10}(T)]/y_c\}.h(y_c).q(x)dx +$$

$$+ \int_{x_c}^{1}\int_{z_c}^{1} E\{u_{111}(T) - u_{110}(T)/z\}.h(y_c).q(x)r(z)dzdx =$$

$$= h(y_c).\{E\{[u_{110}(T) - u_{10}(T)]/y_c\}.\int_{x_c}^{1} q(x)dx +$$

$$+ \int_{x_c}^{1}\int_{z_c}^{1} E\{u_{111}(T) - u_{110}(T)/z\}.q(x)g(z)dzdx\} =$$

$$= h(y_c).\int_{x_c}^{1} q(x)dx.\{E\{[u_{110}(T) - u_{10}(T)]/y_c\} +$$

$$+ \int_{z_c}^{1} E\{u_{111}(T) - u_{110}(T)/z\}.g(z)dz\}.$$

$$\frac{\partial}{\partial x_c} Eu(t/x_c, y_c, z_c) = q(x_c)\ \{E\{[u_{10}(T) - u_0(T)]/x_c\} +$$

$$+ \int_{y_c}^{1} E\{[u_{110}(T) - u_{10}(T)]/y\}h(y).dy +$$

$$+ \int_{y_c}^{1}\int_{z_c}^{1} E\{[u_{111}(T) - u_{110}(T)/z\}g(z).h(y).dzdy\}$$

При следните предположения за маргиналните плътности на X, Y и Z: $q(x) > 0$, $h(y) > 0$, $g(z) > 0$, решенията на системата уравнения:

$$E\{[u_{111}(T) - u_{110}(T)/z_c\} = 0$$

$$E\{[u_{110}(T) - u_{10}(T)]/y_c\} + \int_{z_c}^{1} E\{[u_{111}(T) - u_{110}(T)]/z\}.g(z)dz = 0$$

$$E\{[u_{10}(T) - u_0(T)]/x_c\} + \int_{y_c}^{1} E\{[u_{110}(T) - u_{10}(T)]/y\}h(y).dy +$$

$$+\int_{y_c}^{1}\int_{z_c}^{1} E\{[u_{111}(T) - u_{110}(T)/z\}g(z).h(y).dzdy = 0$$

са търсените оптимални стойности на x_c, y_c и z_c.

4.4 Обобщен модел

За да формира своята вяра за вида на популационното разпределение агентът може да започне със статистически критерий за симетричност. В случаите, когато хипотезата за симетричност се отхвърли, може да се провери хипотеза за съмнителни наблюдения. Ако хипотезата за съмнителни наблюдения се отхвърли може да се продължи с трансформация на данните. Ако хипотезата за съмнителни наблюдения не може да бъде отхвърлена агентът трябва да попита потребителя дали има основания от нестатистически характер да се отстранят рязко отклоняващите се наблюдения. В случаите, за които хипотезата за симетричност не може да се отхвърли, агентът може да продължи с проверка за нормалност като използва пре-тест и пост-тест. Обикновено като пре-тест се използва статистически критерий за вида на вероятностното разпределение, след което се иска и мнението на потребителя. Понякога, например в регресионния анализ, първо се пита потребителя, а след това регресионните остатъци се проверяват за нормалност.

Ще представим вярата, убедеността на агента като прагова функция, която може да приеме едно от възможните пет състояния a_j, $j=0,1,\dots 4$, където:

a_0 – агентът е сигурен, че разпределението е асиметрично и че няма съмнителни наблюдения. Той ще предложи подходяща трансформация на данните.

a_1 – агентът е сигурен, че разпределението на популацията е асиметрично и че асиметрията се дължи на съмнителни наблюдения. Той ще попита потребителя дали да ги премахне.

a_2 – агентът е сигурен, че разпределението на популацията е симетрично, но не е нормално. Той ще използва предположението за симетричност в статистическия анализ.

a_3 – агентът е сигурен, че популационното разпределение е симетрично и предполага, че то е и нормално. Той ще използва робастни критерии за проверка на статистически хипотези.

a_4 – агентът е сигурен, че разпределението на изследваната популация е нормално. Той ще използва и статистически критерии, чувствителни към предположението за нормалност.

Нека с X да означим наблюдаваната p-стойност на статистиката на теста за симетричност, с Y да означим наблюдаваната p-стойност на статистиката на теста за съмнителни наблюдения, с Z - наблюдавания резултат от пре-теста за нормалност, с W - наблюдавания резултат от пост-теста за нормалност, с T - сигурността на агента, чиято стойност не може да се наблюдава.

Нека да допуснем, че X, Y, Z, W, T са непрекъснати случайни величини.

Да предположим, че връзката между измерените резултати от тестовете X, Y, Z, W и променливата T може да бъде представена чрез съвместното вероятностно разпределение $f(x,y,z,w,t)$.

Правилото за вземане на решение $\delta(x,y,z,w)$ определя за всяка възможна реализация (x,y,z,w) на случайния вектор (X,Y,Z,W) кое състояние a_j, $j=0,1,\ldots 4$, ще бъде присвоено на вярата на агента относно разпределението на изследваната популация.

Слабото правило за вземане на решение δ ще дефинираме по следния начин:

$$\{(x,y,z,w):\delta(x,y,z,w)=a_0\}=A\times B(x)\times[0,1]\times[0,1]$$

$$\{(x,y,z,w):\delta(x,y,z,w)=a_1\}=A\times B^c(x)\times[0,1]\times[0,1]$$

$$\{(x,y,z,w):\delta(x,y,z,w)=a_2\}=A^c\times[0,1]\times C(x)\times[0,1]$$

$$\{(x,y,z,w):\delta(x,y,z,w)=a_3\}=A^c\times[0,1]\times C^c(x)\times D(x,z)$$

$$\{(x,y,z,w):\delta(x,y,z,w)=a_4\}=A^c\times[0,1]\times C^c(x)\times D^c(x,z),$$

където A и A^c са множествата от стойности x от теста за симетричност, водещи съответно до отхвърляне и до приемане на твърдението за симетричност, $B(x)$ и $B^c(x)$ са множествата от стойности y от теста за съмнителни наблюдения, водещи съответно до отхвърляне и до приемане на твърдението за наличие на рязко отклоняващи се наблюдения, $C(x)$ и $C^c(x)$ са множествата от стойности z от пре-теста за нормалност, водещи съответно до отхвърляне и до приемане на твърдението за нормалност, $D(x,z)$ и $D^c(x,z)$ са множествата от стойности w от пост-теста за нормалност, водещи съответно до отхвърляне и до приемане на твърдението за нормалност.

Силното правило за вземане на решение γ ще дефинираме по следния начин:

$$\{(x,y,z,w):\gamma(x,y,z,w)=a_0\}=A\times B\times[0,1]\times[0,1]$$

$$\{(x,y,z,w):\gamma(x,y,z,w)=a_1\}=A\times B^c\times[0,1]\times[0,1]$$

$$\{(x,y,z,w):\gamma(x,y,z,w)=a_2\}=A^c\times[0,1]\times C\times[0,1]$$

$$\{(x,y,z,w):\gamma(x,y,z,w)=a_3\}=A^c\times[0,1]\times C^c\times D$$

$$\{(x,y,z,w):\gamma(x,y,z,w)=a_4\}=A^c\times[0,1]\times C^c\times D^c,$$

където A и A^c са множествата от стойности x от теста за симетричност, водещи съответно до отхвърляне и до приемане на твърдението за симетричност, B и B^c са множествата от стойности y от теста за съмнителни наблюдения, водещи съответно до отхвърляне и до приемане на твърдението за наличие на рязко отклоняващи се наблюдения, C и C^c са множествата от стойности z от пре-теста за нормалност, водещи съответно до отхвърляне и до приемане на твърдението за нормалност, D и D^c са множествата от стойности w от пост-теста за нормалност, водещи съответно до отхвърляне и до приемане на твърдението за нормалност.

Слабите монотонни правила за вземане на решение ще дефинираме по следния начин:

$$\delta(x,y,z,w) = \begin{cases} a_0 & \text{при } X<x_c,\ Y<y_c(x) \\ a_1 & \text{при } X<x_c,\ Y\geq y_c(x) \\ a_2 & \text{при } X\geq x_c,\ Z<z_c(x) \\ a_3 & \text{при } X\geq x_c,\ Z\geq z_c(x),\ W<w_c(x,z) \\ a_4 & \text{при } X\geq x_c,\ Z\geq z_c(x),\ W\geq w_c(x,z), \end{cases}$$

където a_j, j=0,1,2,3,4 са състоянията на вярата на агента, x_c, y_c, z_c, w_c са праговите точки съответно за X, Y, Z и W.

Нека структурата полезност има следния вид:

$$u(t) = \begin{cases} u_0(t) & \text{при } a_0 \\ u_1(t) & \text{при } a_1 \\ u_2(t) & \text{при } a_2 \\ u_3(t) & \text{при } a_3 \\ u_4(t) & \text{при } a_4 \end{cases}$$

Ще пресметнем очакваната полезност.

$$Eu(A^c, B^c(x), C^c(x), D^c(x,z)) =$$

$$= \int_A \int_{B(x)} \int_0^1 \int_0^1 \int_0^1 u_0(t) f(x,y,z,w,t)\,dt\,dw\,dz\,dy\,dx +$$

$$+ \int_A \int_{B^c(x)} \int_0^1 \int_0^1 \int_0^1 u_1(t) f(x,y,z,w,t)\,dt\,dw\,dz\,dy\,dx +$$

$$+ \int_{A^c} \int_0^1 \int_{C(x)} \int_0^1 \int_0^1 u_2(t) f(x,y,z,w,t)\,dt\,dw\,dz\,dy\,dx +$$

$$+\int_{A^c}\int_0^1\int_{C^c(x)}\int_{D(x,z)}\int_0^1 u_3(t)f(x,y,z,w,t)dtdwdzdydx+$$

$$+\int_{A^c}\int_0^1\int_{C^c(x)}\int_{D^c(x,z)}\int_0^1 u_4(t)f(x,y,z,w,t)dtdwdzdydx=$$

$$=E\{u_0(T)\}+$$

$$+\int_A\int_{B^c(x)} E\{[u_1(T)-u_0(T)]/x,y\}.k(x,y)dydx+$$

$$+\int_{A^c} E\{[u_2(T)-u_0(T)]/x\}.q(x)dx+$$

$$+\int_{A^c}\int_{C^c(x)} E\{[u_3(T)-u_2(T)]/x,z\}.l(x,z)dzdx+$$

$$+\int_{A^c}\int_{C^c(x)}\int_{D^c(x,z)} E\{[u_4(T)-u_3(T)]/x,z,w\}.m(x,z,w)dwdzdx=$$

$$=E\{u_0(T)\}+\int_A\int_{B^c(x)} E\{[u_1(T)-u_0(T)]/x,y\}.n(y/x)q(x)dydx+$$

$$+\int_{A^c}\{E\{[u_2(T)-u_0(T)]/x\}+\int_{C^c(x)}\{E\{[u_3(T)-u_2(T)]/x,z\}+$$

$$+\int_{D^c(x,z)} E\{[u_4(T)-u_3(T)]/x,z,w\}.r(w/x,z)dw\}.p(z/x)dz\}.q(x)dx=$$

$$=E\{u_0(T)\}+\int_{B^c(x)} E\{[u_1(T)-u_0(T)]/y\}.n(y)dy-$$

$$-\int_{A^C}\int_{B^c(x)} E\{[u_1(T)-u_0(T)]/x,y\}.n(y/x)q(x)dydx+$$

$$+\int_{A^c}\{E\{[u_2(T)-u_0(T)]/x\}+\int_{C^c(x)}\{E\{[u_3(T)-u_2(T)]/x,z\}+$$

$$+\int_{D^c(x,z)}E\{[u_4(T)-u_3(T)]/x,z,w\}.r(w/x,z)dw\}.p(z/x)dz\}.q(x)dx=$$

$$=E\{u_0(T)\}+\int_{B^c(x)}E\{[u_1(T)-u_0(T)]/y\}.n(y)dy+$$

$$+\int_{A^c}\{E\{[u_2(T)-u_0(T)]/x\}+\int_{C^c(x)}\{E\{[u_3(T)-u_2(T)]/x,z\}+$$

$$+\int_{D^c(x,z)}E\{[u_4(T)-u_3(T)]/x,z,w\}.r(w/x,z)dw\}.p(z/x)dz-$$

$$-\int_{B^c(x)}E\{[u_1(T)-u_0(T)]/x,y\}.n(y/x)dy\}q(x)dx$$

Една горна граница на очакваната полезност е

$$Eu(A^c,B^c(x),C^c(x),D^c(x,z))\leq$$

$$\leq E\{u_0(T)\}+\int_{B_0^c(x)}E\{[u_1(T)-u_0(T)]/y\}n(y)dy+$$

$$+\int_{A_0^c}\{E\{[u_2(T)-u_0(T)]/x\}+\int_{C_0^c(x)}\{E\{[u_3(T)-u_2(T)]/x,z\}+$$

$$+\int_{D_0^c(x,z)}E\{[u_4(T)-u_3(T)]/x,z,w\}.r(w/x,z)dw\}.p(z/x)dz-$$

$$-\int_{B_0^c(x)}E\{[u_1(T)-u_0(T)]/x,y\}.n(y/x)dy\}q(x)dx,$$

където

$$D_0^c(x,z)=\{w:E\{[u_4(T)-u_3(T)]/x,z,w\}\geq 0\} \qquad (4.22)$$

$$C_0^c(x) \quad = \{z : E\{[u_3(T) - u_2(T)]/x, z\} +$$

$$+ \int_{D_0^c(x,z)} E\{[u_4(T) - u_3(T)]/x, z, w\}.r(w/x, z)dw \geq 0\} \tag{4.23}$$

$$B_0^c(x) \;\; = \{y : E\{[u_1(T) - u_0(T)]/y\} \geq 0\}, \tag{4.24}$$

$$A_0^c \quad = \{\text{x}: \; E\{[u_2(T) - u_0(T)]/x\} + \int_{C_0^c(x)} \{E\{[u_3(T) - u_2(T)]/x, z\} +$$

$$+ \int_{D_0^c(x,z)} E\{[u_4(T) - u_3(T)]/x, z, w\} r(w/x, z)dw\}.p(z/x)dz -$$

$$- \int_{B_0^C(x)} E\{[u_1(T) - u_0(T)]/x, y\}.n(y/x)dy \geq 0\}. \tag{4.25}$$

Следните теореми дават достатъчни условия за слабата монотонност на правилата за вземане на решение.

<u>Теорема 4.5</u>: *Нека*

- *следните функции* $u_1(t) - u_0(t)$, $u_3(t) - u_2(t)$, $u_4(t) - u_3(t)$ *са строго растящи.*
- *функцията*

$$E\{[u_2(T) - u_0(T)]/x\} + \int_{C_0^c(x)} \{E\{[u_3(T) - u_2(T)]/x, z\} +$$

$$+ \int_{D_0^c(x,z)} E\{[u_4(T) - u_3(T)]/x, z, w\} r(w/x, z)dw\}.p(z/x)dz -$$

$$- \int_{B_0^C(x)} E\{[u_1(T) - u_0(T)]/x, y\}.n(y/x)dy$$

е строго растяща в ***x***.

- $V(t/x, z, w)$ *е строго намаляваща в* ***x***, ***z***, ***w*** *за всяко* ***t***.
- $U(t/x, y)$ *е строго намаляваща в* ***x*** и ***y*** *за всяко* ***t***.
- $R(w/x, z)$ *е строго намаляваща в* ***x*** *и* ***z*** *за всяко* ***w***.

Тогава съществува слабо монотонно решение на проблема за вземане на решение за вида на вероятностното разпределение на изследваната непрекъсната популация.

***<u>Теорема 4.6</u>**: Ако са в сила следните условия:*

- *функциите* $u_1(t)-u_0(t)$, $u_3(t)-u_2(t)$, $u_4(t)-u_3(t)$ *са строго растящи,*
- *функцията*

$$E\{[u_2(T)-u_0(T)]/x\}+\int_{C_0^c(x)}\{E\{[u_3(T)-u_2(T)]/x,z\}+$$

$$+\int_{D_0^c(x,z)}E\{[u_4(T)-u_3(T)]/x,z,w\}r(w/x,z)dw\}.p(z/x)dz-$$

$$-\int_{B_0^C(x)}E\{[u_1(T)-u_0(T)]/x,y\}.n(y/x)dy$$

е строго растяща в ***x****.*

- $V(t/x,y,z,w)$ *е строго намаляваща в* ***x, y, z, w*** *за всяко* ***t****,*
- $R(w/x,z)$ *е строго намаляваща в* ***x*** *и* ***z*** *за всяко* ***w****,*

то съществува слабо монотонно решение на проблема за вземане на решение за вида на вероятностното разпределение на изследваната непрекъсната популация.

Доказателството на тези две теореми е аналогично на доказателството на теоремата от 4.1.4.

Когато работим със силни правила върху вероятностния модел трябва да наложим следните условия:

- $v_1(t/x,y)=v_1(t/y)$,
- $v_2(t/x,z)=v_2(t/z)$,
- $v_3(t/x,z,w)=v_3(t/w)$,
- $g(z/x)=g(z)$,
- $h(y/x)=h(y)$,
- $k(w/x,z)=k(w)$.

Задачата за намиране на оптималните силни правила се свежда до максимизиране на очакваната полезност, която има следния вид:

$$Eu(t/x_c, y_c, z_c, w_c) = E\{u_0(T)\} +$$

$$+\int_{y_c}^{1}\int_{0}^{x_c} E\{[u_1(T)-u_0(T)]/y\}.q(x)h(y)dxdy +$$

$$+\int_{x_c}^{1} E\{[u_2(T)-u_0(T)]/x\}.q(x)dx +$$

$$+\int_{x_c}^{1}\int_{z_c}^{1} E\{[u_3(T)-u_2(T)]/z\}.q(x)g(z)dzdx +$$

$$+\int_{x_c}^{1}\int_{z_c}^{1}\int_{w_c}^{1} E\{[u_4(T)-u_3(T)]/w\}.q(x)g(z)k(w)dwdzdx =$$

$$= E\{u_0(T)\} + \int_{y_c}^{1} E\{[u_1(T)-u_0(T)]/y\}.h(y)dy -$$

$$-\int_{y_c}^{1}\int_{x_c}^{1} E\{[u_1(T)-u_0(T)]/y\}.q(x)h(y)dxdy +$$

$$+\int_{x_c}^{1} E\{[u_2(T)-u_0(T)]/x\}.q(x)dx +$$

$$+\int_{x_c}^{1}\int_{z_c}^{1} E\{[u_3(T)-u_2(T)]/z\}.q(x)g(z)dzdx +$$

$$+\int_{x_c}^{1}\int_{z_c}^{1}\int_{w_c}^{1} E\{[u_4(T)-u_3(T)]/w\}.q(x)g(z)k(w)dwdzdx =$$

За да получим оптималното решение на проблема ще диференцираме $Eu(t/x_c, y_c, z_c, w_c)$ по w_c, z_c, x_c и y_c и ще приравним на 0.

$$\frac{\partial}{\partial w_c} Eu(t/x_c, y_c, z_c, w_c) =$$

$$=\int_{x_c}^{1}\int_{z_c}^{1} E\{[u_4(T)-u_3(T)]/w_c\}.q(x)g(z)k(w_c)dzdx=$$

$$=k(w_c)E\{[u_4(T)-u_3(T)]/w_c\}.\int_{x_c}^{1}\int_{z_c}^{1} q(x)g(z)dzdx$$

$$\frac{\partial}{\partial z_c} Eu(t/x_c,y_c,z_c,w_c)=$$

$$=\int_{x_c}^{1} E\{[u_3(T)-u_2(T)]/z_c\}.q(x)g(z_c)dx+$$

$$+\int_{x_c}^{1}\int_{w_c}^{1} E\{[u_4(T)-u_3(T)]/w\}.q(x)g(z_c)k(w)dwdx=$$

$$=g(z_c).\{\int_{x_c}^{1} E\{[u_3(T)-u_2(T)]/z_c\}.q(x)dx+$$

$$+\int_{x_c}^{1}\int_{w_c}^{1} E\{[u_4(T)-u_3(T)]/w\}.q(x)k(w)dwdx\}=$$

$$=g(z_c).\int_{x_c}^{1} q(x)dx\{E\{[u_3(T)-u_2(T)]/z_c\}+$$

$$+\int_{w_c}^{1} E\{[u_4(T)-u_3(T)]/w\}.k(w)dw\}$$

$$\frac{\partial}{\partial x_c} Eu(t/x_c,y_c,z_c,w_c)=E\{[u_2(T)-u_0(T)]/x_c\}.q(x_c)+$$

$$+\int_{z_c}^{1} E\{[u_3(T)-u_2(T)]/z\}.q(x_c)g(z)dz+$$

$$+\int_{z_c}^{1}\int_{w_c}^{1} E\{[u_4(T)-u_3(T)]/w\}.q(x_c)g(z)k(w)dwdz-$$

$$-\int_{y_c}^{1} E\{[u_1(T)-u_0(T)]/y\}.q(x_c)h(y)dy=$$

$$= q(x_c)\{ E\{[u_2(T) - u_0(T)]/x_c\} + \int_{z_c}^{1} E\{[u_3(T) - u_2(T)]/z\}.g(z)dz +$$

$$+ \int_{z_c}^{1} \int_{w_c}^{1} E\{[u_4(T) - u_3(T)]/w\}.g(z)k(w)dwdz -$$

$$- \int_{y_c}^{1} E\{[u_1(T) - u_0(T)]/y\}.h(y)dy\}$$

$$\frac{\partial}{\partial y_c} Eu(t/x_c, y_c, z_c, w_c) = d(y_c)E\{[u_1(T) - u_0(T)]/y_c\}.\int_0^{x_c} q(x)dx$$

При предположение, че $q(x) > 0, d(y) > 0, g(z) > 0, h(w) > 0$ търсените оптимални стойности на x_c, y_c, z_c, и w_c са решенията на системата уравнения:

$$E\{[u_1(T) - u_0(T)]/y_c\} = 0$$

$$E\{([u_2(T) - u_0(T)]/x_c\} + \int_{z_c}^{1} E\{[u_3(T) - u_2(T)]/z\}.g(z)dz +$$

$$+ \int_{z_c}^{1} \int_{w_c}^{1} E\{[u_4(T) - u_3(T)]/w\}.g(z)k(w)dwdz -$$

$$- \int_{y_c}^{1} E\{[u_1(T) - u_0(T)]/y\}.h(y)dy = 0$$

$$E\{[u_3(T) - u_2(T)]/z_c\} + \int_{w_c}^{1} E\{[u_4(T) - u_3(T)]/w\}.k(w)dw = 0$$

$$E\{[u_4(T) - u_3(T)]/w_c\} = 0.$$

Глава 5

Структури от данни

5.1 Мрежа за вземане на решения

Ще разширим Бейсовата мрежа, описана в Глава 3 като добавим мрежа за вземане на решения с която се генерира събитието $E=\{X_7=x_7\}$, където случайната величина X_7 е *вид на непрекъснатото разпределение*.

Мрежите за вземане на решения са графични структури, представящи вероятностни зависимости и информационни потоци ([*Shachter,86*], [*Shachter,88*]).

Една мрежа за вземане на решения се състои от:

- Множество от елементарни действия
- Вероятностна структура
- Структура полезност
- Правило за вземане на решения
- Оптимално действие.

Определение 5.1: Мрежа за вземане на решения (**d**-мрежа, decision network, influence diagram) се нарича наредената петорка $((A,\mathbf{I}), (X,P), A^*, u, \delta)$, където

$(A,\mathbf{I})$ е ориентиран граф на елементарните действия, където $A=\{A_i,\ i=1,\dots,k\}$ е множеството от елементарни действия, $\mathbf{I}=\{\mathrm{I}_{l,m}=(A_l,A_m),\ l\neq m,\ l=1,\dots,k;\ m=1,\dots,k\}$ е множеството от ориентирани информационни дъги,

(X,P) е Бейсова мрежа,

A^* е множеството от оптималните действия (множеството от решенията),

$u: X\times A \rightarrow R$ е функция полезност, където R е реалната права,

$\delta: A \rightarrow A^*$ е правило за вземане на решение.

Определение 5.2: Мрежа за вземане на решение за състоянието на вярата на агента се нарича **d**-мрежа, в която

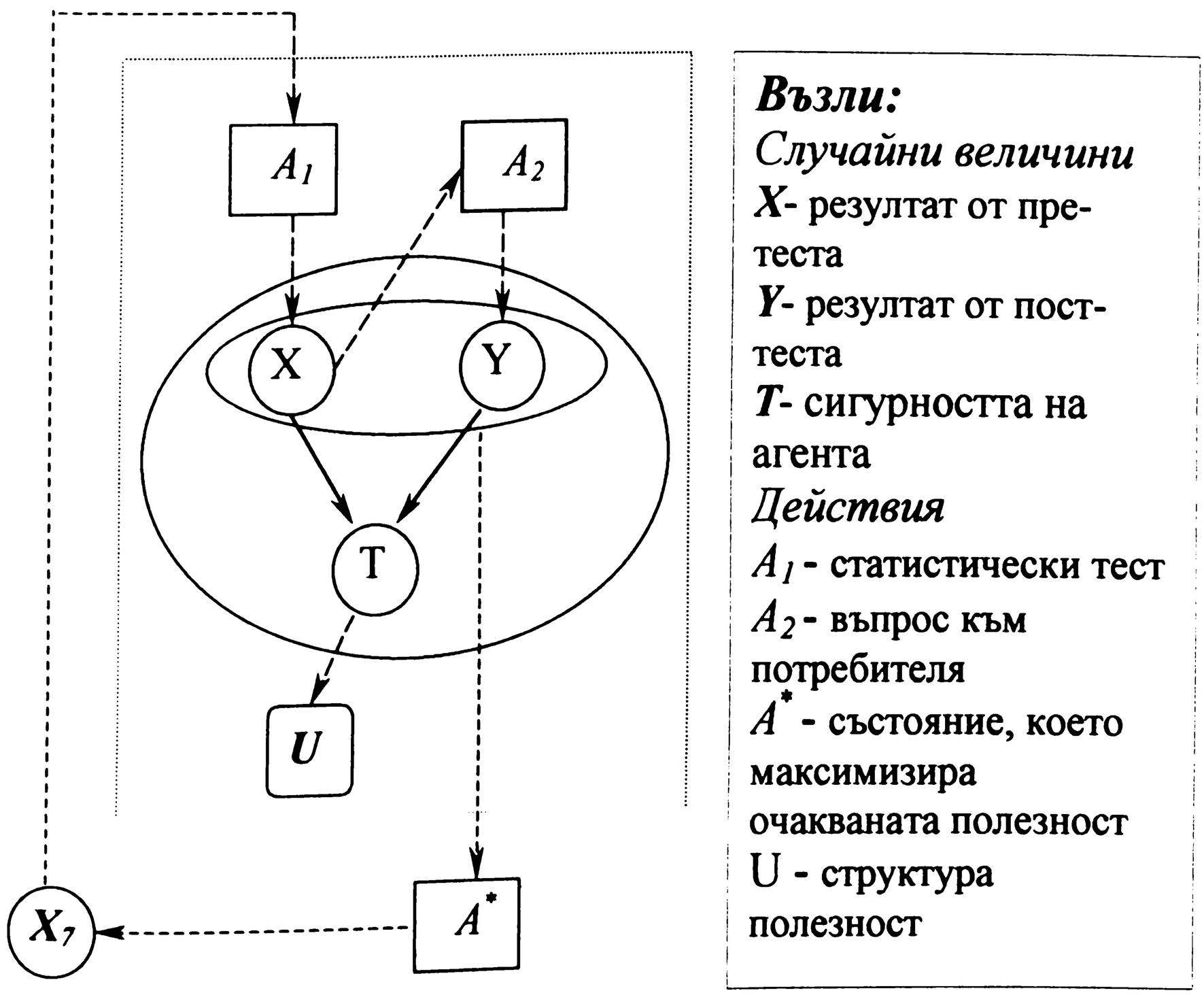
A е множество от (пре- и пост-) тестове,

A^* е множество от състоянията на вярата на агента,

$u: X\times A \rightarrow [0,1]$ е функция полезност,

$\delta: A \rightarrow A^*$ е монотонно слабо правило за вземане на решение за състоянието на вярата на агента.

На *фигура 5.1* е дадена мрежа за вземане на решение, представяща модела от *4.1.2*. В квадрати са визуализирани възлите-елементарни действия $A_i \in A$, i=1,2, реализиращи тестовете. В кръгове са визуализирани възлите-случайни величини. С пунктирани стрелки са представени информационните дъги, визуализиращи потока от информация. С непрекъснати стрелки са представени зависимостите във вероятностния модел $p(x,y,t)$. Правилото за вземане на решение δ е дефинирано чрез (4.1).



Фигура 5.1 *Мрежа за вземане на решение за състоянието на вярата на агента.*

Алгоритъм 5.1: *Алгоритъм за извод в мрежа за вземане на решение за състоянието на вярата на агента:*

Вход: мрежа за вземане на решение за състоянието на вярата на агента

Изход: оптималното състояние на вярата на агента

Стъпки:

1. Изпълни елементарните действия, в реда, определен от информационните дъги на **d**-мрежата.
2. Приложи правилото за вземане на решения към резултатите от елементарните действия за да получиш оптималното състояние на вярата на агента.
3. Край.

5.2 U-мрежа

5.2.1 Дърво на събитията

Нека $X=\{X_1, X_2, \ldots, X_n\}$ е множество от непрекъснати случайни величини.

Разглеждаме двоична дървовидна структура със следните свойства:

- в корена, възлите и листата са изобразени събития,
- в корена е изобразено достоверното събитие Ω,
- всеки възел има 0 или 2 наследника.
- Ако възлите A_j и A_k са съответно ляв и десен наследник на възела A_i, то A_j изобразява събитието $\{X_j < x_j^c\}$, а A_k изобразява противоположното събитие $\{X_j \geq x_j^c\}$, т.е.

  Означаваме $\overline{A_j} = A_k = \{X_j \geq x_j^c\}$.

Очевидно броят на елементите на тази структура е $2n+1$, броят на листата е $n+1$ и наследниците на всеки елемент са наредени.

Наредена двоична дървовидна структура от този вид ще наричаме ***дърво на събитията***.

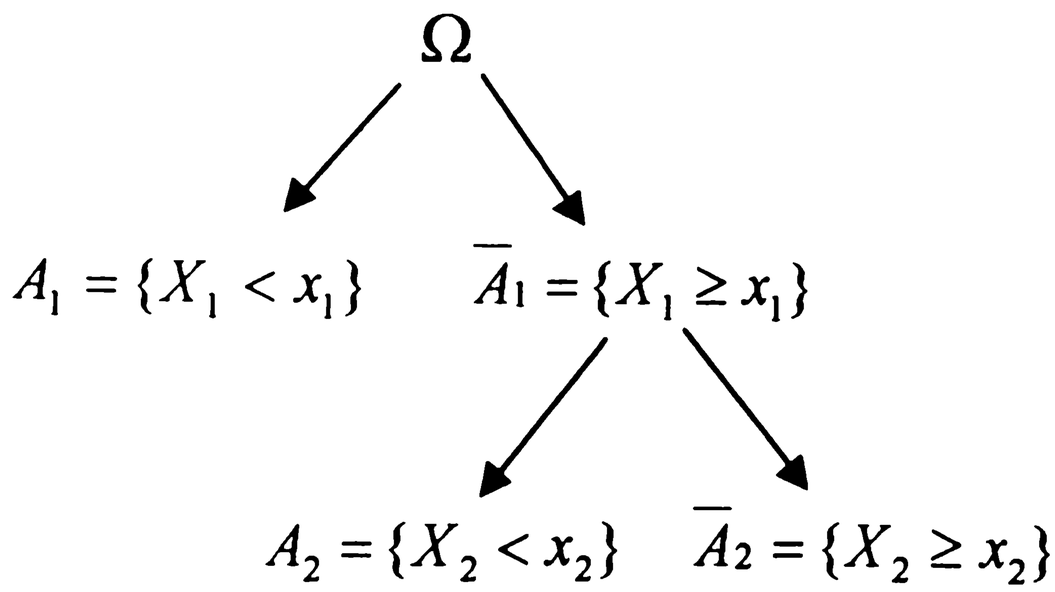
Определение 5.3: Казваме, че дървото на събитията е в *каноничен вид*, ако номерата на асоциираните случайни величини от корена към листата и от ляво на дясно съвпадат с първите ***n*** естествени числа.

Алгоритъмът за превеждане на произволно дърво в каноничен вид се състои в преномериране на асоциираните със събитията случайни величини в нарастващ ред в посока от корена към листата и от ляво на дясно.

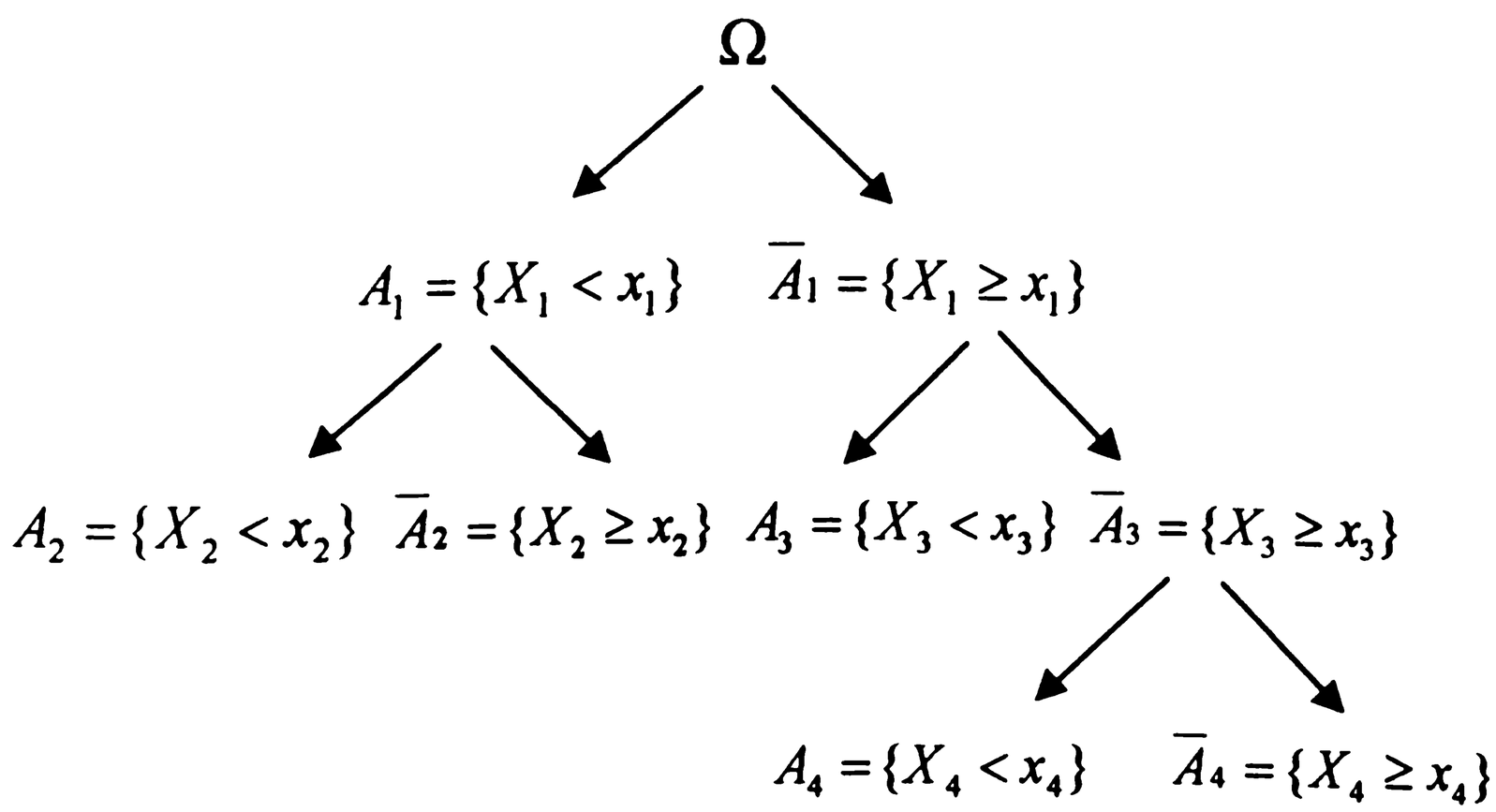
От тук нататък ще разглеждаме само дървета на събитията в каноничен вид.

Пример 5.1: На *фигура 5.2* е представено дървото на събитията с *5* възела от *4.1.2*.

Пример 5.2: На *фигура 5.3* е представено дървото на събитията с *9* възела от §*4.2*.



Фигура5.2 *Дървото на събитията от 4.1.2.*



Фигура5.3 *Дървото на събитията от §4.4.*

5.2.2 Фактори

Определение 5.4: Пътят от възел от първо равнище до листо в дървото на събитията ще наричаме *фактор*. Броят на възлите във фактора ще наричаме *дължина на фактора*.

Дължината на фактора F_i ще означаваме с $\#F_i$.

Очевидно броят на факторите в дървото на събитията е $n+1$. Факторът ще интерпретираме като едновременно настъпване на събитията, асоциирани с него, т.е. като сечение на асоциираните събития.

Нека $F = \{F_i, i = 1,2,\ldots,n+1\}$ е множеството от факторите на дървото на събитията.

Ще номерираме факторите от дървото на събитията от ляво на дясно. С всеки фактор F_i ще свържем една функция полезност u_i, $i = 1,\ldots,n+1$. Т.е. с всяко листо от дървото на събитията ще свържем възел-полезност u_i. Казваме, че с множеството от факторите на дървото на събитията е асоциирано множество от възли полезност.

С факторите можем да свържем и вероятности и да изчислим очакваната полезност.

Следователно с дървото на събитията свързахме множество от фактори и множество от функции полезност.

В светлината на идеите, представени в [*Shoham,97a*] и [*Shoham,97b*] ще дефинираме нова структура от данни.

Определение 5.5: Двойката (**F**,**U**), където **F** е множеството от факторите на дървото на събитията, а **U** е множеството от възлите полезност, асоциирани с него, се нарича **U**-*мрежа*.

За да можем да намерим очакваната полезност се нуждаем и от вероятностната структура на множеството от случайни величини $X = \{X_1, X_2, \ldots, X_n\}$.

Ще дефинираме следните операции над дървото на събитията.

Операцията $\psi(A_i)$ дава списък от случайните величини, свързани с фактора, асоциран с листото A_i.

Операцията $\varepsilon(A_i)$ дава индекса на елемента с най-голям индекс от списъка от случайните величини, свързани с фактора, асоциран с листото A_i.

Пример 5.3: Нека с листото A_i е асоциран фактора $F_i = \{X_1 < x_1\},\{X_2 \geq x_2\}$. Тогава $\psi(A_i) = X_1, X_2$, $\varepsilon(A_i) = 2$.

Операцията $U(A_i)$ дава функцията полезност, асоциирана с листото A_i.

Глава 6

Теоретични основи на алгоритъма за оптимизиране на вярата

6.1 Алгоритъм

Когато разглеждаме математическия модел с A_i и $\overline{A_i}$ ще означаваме съответно интервалите $[0, x_i^c)$ и $[x_i^c, 1]$. Когато разглеждаме структури от данни с A_i и $\overline{A_i}$ ще означаваме възли и листа.

Ще започнем с разглеждането на следните *три случая:*

Първи случай:

Нека $X = \{X_1, X_2\}$ е множество от непрекъснати случайни величини. Нека вероятностният модел се задава с Бейсовата мрежа от *фигура 6.4*. Нека с множеството $X=\{X_1, X_2\}$ е асоциирана произволна U-*мрежа* (*виж фигура 6.5.*). Оценявайки очакваната полезност постъпково ще намерим уравнения, чиито решения са оптималните точки x_i^c, $i = 1,2$, определящи факторите от множеството $F = \{F_i, i = 1,2,3\}$.

Разглеждаме двойката листа $(A_2, \overline{A_2})$ от второ равнище на дървото на събитията, представено на *фигура 6.5а*.

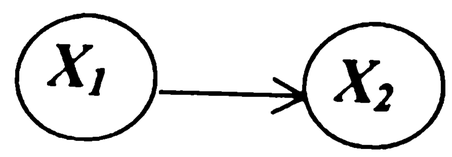
Нека с получената двойка листа $(A_2, \overline{A_2})$ са асоциирани възли полезност, които представят съответно функциите полезност $u_i(t)$ и $u_{i+1}(t)$.

Тогава разглеждаме очакваната полезност като функция на интервала $A_2 = [x_2^c, 1]$, т.е. на праговата точка x_2^c:

$$Eu(A_2) = \int_{A_2} E[u_i(T)/x_2, x_1]\, f_2(x_2/x_1)dx_2 +$$

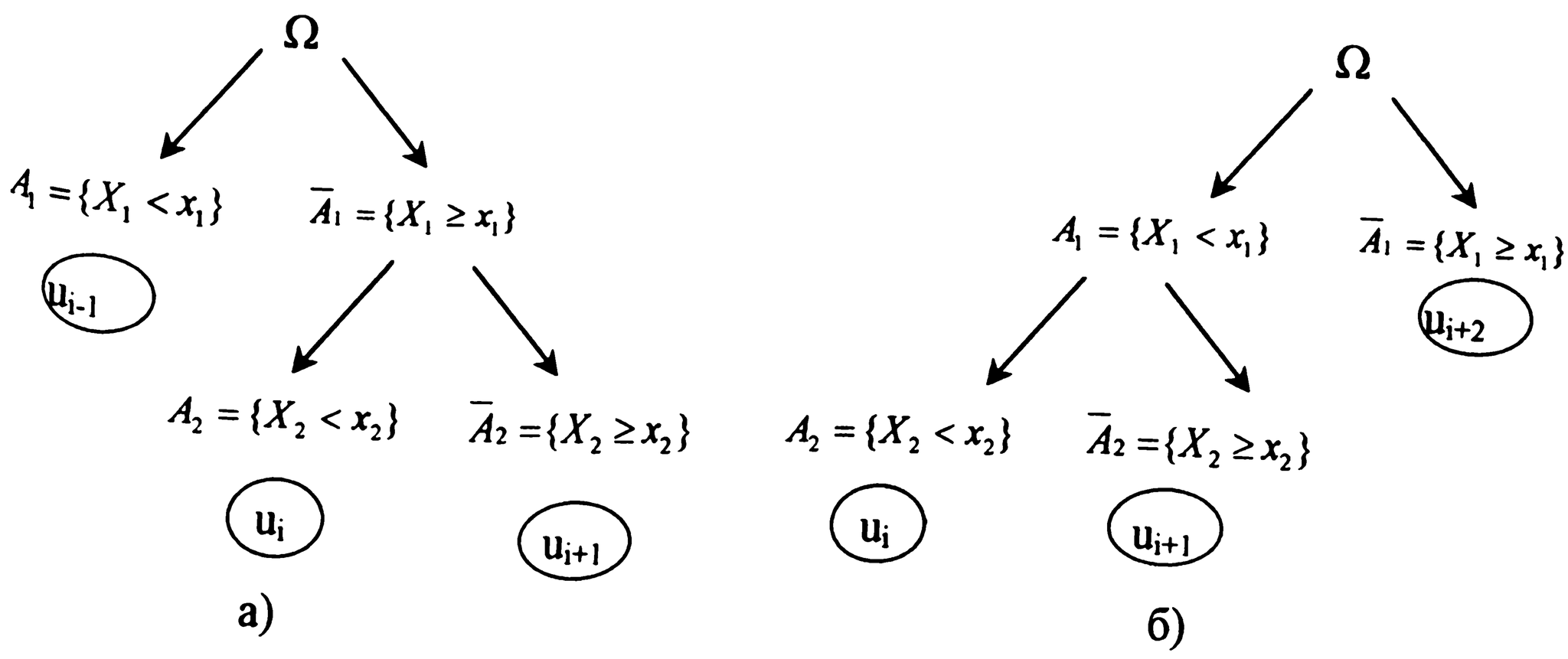
$$+ \int_{\overline{A_2}} E[u_{i+1}(T)/x_2, x_1]\, f_2(x_2/x_1)dx_2 =$$

$$= E\{u_i(T)/x_1\} + \int_{\overline{A_2}} E\{[u_{i+1}(T) - u_i(T)]/x_2, x_1\} . f_2(x_2/x_1) dx_2$$



$$f(x_1, x_2) = f_1(x_1).f_2(x_2 / x_1)$$

Фигура 6.4 *Бейсова мрежа от 4.1.2.*



Фигура 6.5 *U-мрежи.*

Една горна граница на очакваната полезност е:

$$Eu(\overline{A_2}) \leq E\{u_i(T)/x_1\} + \int_{\overline{A_2^0}} E\{[u_{i+1}(T) - u_i(T)]/x_2, x_1\} . f_2(x_2/x_1) dx_2,$$

където $\overline{A_2^0} = \{x_2 : E\{[u_{i+1}(T) - u_i(T)]/x_2, x_1\} \geq 0\}$.

Следователно оптималната прагова точка x_2^c е коренът на уравнението

$$E\{[u_{i+1}(T) - u_i(T)]/x_2, x_1\} = 0. \qquad (6.1)$$

Нека $\pi(A_2)$ е родителят на двойката листа $(A_2, \overline{A_2})$.

Над дървото на събитията ще извършим операцията изключване на елемент - листо. Ще изключим листата A_2 и $\overline{A_2}$. С новополученото листо $\pi(A_2)$ ще свържем получената горна граница на функцията полезност $Eu(\overline{A_2})$, т.е. с новополученото листо $\pi(A_2)$ ще асоциираме функцията полезност $u^*(t)$, която има следния вид:

$$u^*(t)=E\{u_i(T)/x_1\} + \int_{\overline{A_2}^0} E\{[u_{i+1}(T)- u_i(T)]/x_2, x_1\}.f_2(x_2/x_1)dx_2.$$

$u^*(t)=E\{u_i(T)/x_1\} + \overline{A_1}.tail$, където

$$\overline{A_1}.tail = \int_{\overline{A_2}^0} E\{[u_{i+1}(T)- u_i(T)]/x_2, x_1\}.f_2(x_2/x_1)dx_2.$$

Получихме ново дърво на събитията, което е поддървовидна структура на дървото на събитията, и ново множество от възли-полезност.

Разглеждаме двойката листа $(A_1, \overline{A_1})$.

Първо ще предположим, че $\overline{A_1} = \pi(A_2)$ (*виж фигура 6.5а*).

Знаем, че с листата A_1 и $\overline{A_1}$ са асоциирани възли полезност, които представят съответно функциите $u_{i-1}(t)$ и $u^*(t)$.

Тогава очакваната полезност е:

$$Eu(\overline{A_1}, \overline{A_2}^0) = \int_{A_1} E[u_{i-1}(T)/x_1].f_1(x_1)dx_1 + \int_{\overline{A_1}} E[u_i(T)/x_1)].f_1(x_1)dx_1 +$$

$$+ \int_{\overline{A_1}} \{ \int_{\overline{A_2}^0} E\{[u_{i+1}(T)- u_i(T)]/ x_2, x_1\}.f_2(x_2/ x_1)dx_2\}.f_1(x_1)dx_1 =$$

$$= E\{u_{i-1}(T)\} + \int_{\overline{A_1}} \{ E\{[u_i(T)- u_{i-1}(T)]/x_1\} +$$

$$+ \int_{\overline{A_2}^0} E\{[u_{i+1}(T)- u_i(T)]/ x_2, x_1\}.f_2(x_2/ x_1)dx_2\}.f_1(x_1)dx_1$$

Една горна граница на очакваната полезност е:

$$Eu(\overline{A_1}, \overline{A_2}^0) \le E\{u_{i-1}(T)\} + \int_{\overline{A_1}^0} \{ E\{[u_i(T)- u_{i-1}(T)]/x_1\} +$$

$$+ \int_{\overline{A_2}^0} E\{[u_{i+1}(T)- u_i(T)]/ x_2, x_1\} f_2(x_2/x_1)dx_2\}. f_1(x_1)dx_1 ,$$

където $\overline{A_1^0} = \{x_1 : E\{[u_i(T)- u_{i-1}(T)]/x_1\} +$

$$+ \int_{\overline{A_2}^0} E\{[u_{i+1}(T)- u_i(T)]/ x_2, x_1\} f_2(x_2/ x_1)dx_2\} \ge 0\}.$$

Т.е. горната граница е $Eu(\overline{A_1}^0, \overline{A_2}^0) = E\{u_{i-1}(T)\} + \overline{A_0}.tail$,
където

$$\overline{A_0}.tail = \int_{\overline{A_1}^0} \{ E\{[u_i(T)- u_{i-1}(T)]/x_1\} + \overline{A_1}.tail \}. f_1(x_1)dx_1$$

Следователно оптималната прагова точка x_1^c е коренът на уравнението

$$E\{[u_i(T)- u_{i-1}(T)]/x_1\} + \tag{6.2}$$

$$+ \int_{\overline{A_2}^0} E\{[u_{i+1}(T)- u_i(T)]/x_2,x_1\} f_2(x_2/x_1)dx_2\} = 0.$$

Т.е. $E\{[u_i(T)- u_{i-1}(T)]/x_1\} + \overline{A_0}.tail = 0$

Получихме система от две уравнения (6.1) и (6.2), корените на която са оптималните прагови точки x_i^c, i=1,2.

Да предположим сега, че $A_1 = \pi(A_2)$ (*виж фигура 6.5б*). Знаем, че с двойката листа $(A_1, \overline{A_1})$ е асоциирана структурата полезност

$$u(t) = \begin{cases} E\{u_i(T)/x_1\} + \int_{\overline{A_2}^0} E\{[u_{i+1}(T)- u_i(T)]/x_2,x_1\} f_2(x_2/x_1)dx_2 \\ u_{i+2}(t) \end{cases}$$

Тогава очакваната полезност е:

$$Eu(\overline{A_1}, \overline{A_2}^0) = \int_{A_1} E[u_i(T)/ x_1] f_1(x_1)dx_1 +$$

$$+\int_{A_1}\{\int_{\overline{A}_2^0} E\{[u_{i+1}(T)-u_i(T)]/x_2,x_1\}.f_2(x_2/x_1)dx_2\}.f_1(x_1)dx_1+$$

$$+\int_{\overline{A}_1} E[u_{i+2}(T)/x_1].f_1(x_1)dx_1=$$

$$=E\{u_i(T)\}+\int_{\overline{A}_2^0} E\{[u_{i+1}(T)-u_i(T)]/x_2\}.f_2(x_2/x_1)dx_2+$$

$$+\int_{\overline{A}_1} E\{[u_{i+2}(T)-u_i(T)]/x_1)\}.f_1(x_1)dx_1-$$

$$-\int_{A_1}\{\int_{\overline{A}_2^0} E\{[u_{i+1}(T)-u_i(T)]/x_2,x_1\}.f_2(x_2/x_1)dx_2\}.f_1(x_1)dx_1$$

Една горна граница на очакваната полезност е:

$$Eu(\overline{A}_1,\overline{A}_2^0)\le E\{u_i(T)\}+$$

$$+\int_{\overline{A}_2^0} E\{[u_{i+1}(T)-u_i(T)]/x_2\}.f_2(x_2/x_1)dx_2+\int_{\overline{A}_1^0}\{E\{[u_{i+2}(T)-u_i(T)]/x_1\}-$$

$$-\int_{\overline{A}_2^0} E\{[u_{i+1}(T)-u_i(T)]/x_2,x_1\}.f_2(x_2/x_1)dx_2\}.f_1(x_1)dx_1,$$

където

$$\overline{A}_1^0=\{x_1:E\{[u_{i+2}(T)-u_i(T)]/x_1-$$

$$-\int_{\overline{A}_2^0} E\{[u_{i+1}(T)-u_i(T)]/x_2,x_1\}.f_2(x_2/x_1)dx_2\}\ge 0\}.$$

Означаваме

$$A_0.tail=\int_{\overline{A}_1^0}\{E\{[u_{i+2}(T)-u_i(T)]/x_1\}-A_1.tail\}.f_1(x_1)dx_1$$

Следователно в този случай оптималната прагова точка x_1^c е коренът на уравнението:

$$E\{[u_{i+2}(T)-u_i(T)]/x_1\}-\int_{\overline{A}_2^0} E\{[u_{i+1}(T)-u_i(T)]/x_2,x_1\}f_2(x_2/x_1)dx_2=0 \qquad (6.3)$$

т.е. $E\{[u_{i+2}(T)-u_i(T)]/x_1\} - A_1.tail = 0$

В този случай оптималните прагови точки x_i^c, i=1,2 са решения на системата уравнения (6.1) и (6.3).

Втори случай:

Нека $X=\{X_1, X_2, X_3\}$ е множество от непрекъснати случайни величини. Нека вероятностният модел се задава с Бейсовата мрежа от *фигура 6.6* . Нека с множеството $X=\{X_1, X_2, X_3\}$ е асоциирана U-*мрежата от фигура 6.7.*

Разглеждаме двойката листа $(A_3, \overline{A_3})$ от трето равнище на дървото на събитията.

Една горна граница на очакваната полезност е:

$$Eu(\overline{A_3}) \le E\{u_1(T)/x_2,x_1\} + \int_{\overline{A_3^0}} E\{[u_2(T)-u_1(T)]/x_3,x_2,x_1\}.f_3(x_3/x_2,x_1)dx_3,$$

където $\overline{A_3^0}=\{x_3: E\{[u_2(T)-u_1(T)]/x_3,x_2,x_1\} \ge 0\}$.

Следователно оптималната прагова точка x_3^c е коренът на уравнението

$E\{[u_2(T)-u_1(T)]/x_3,x_2,x_1\} = 0$.

Разглеждаме двойката листа $(A_2, \overline{A_2})$.

Една горна граница на очакваната полезност е:

$$Eu(\overline{A_2}) \le E\{u_1(T)/x_1\} + \int_{\overline{A_3^0}} E\{[u_2(T)-u_1(T)]/x_3,x_2,x_1\}.f_3(x_3/x_2,x_1)dx_3 +$$

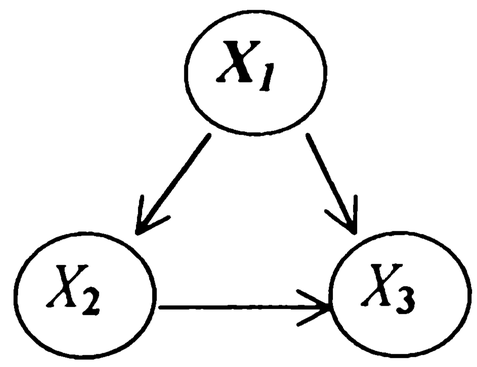
$$+ \int_{\overline{A_2^0}} \{E\{[u_3(T)-u_1(T)]/x_2,x_1\} - \int_{\overline{A_3^0}} E\{[u_2(T)-u_1(T)]/x_3,x_2,x_1\}.$$

$$f_3(x_3/x_2,x_1)dx_3\}f_2(x_2/x_1)dx_2,$$

където

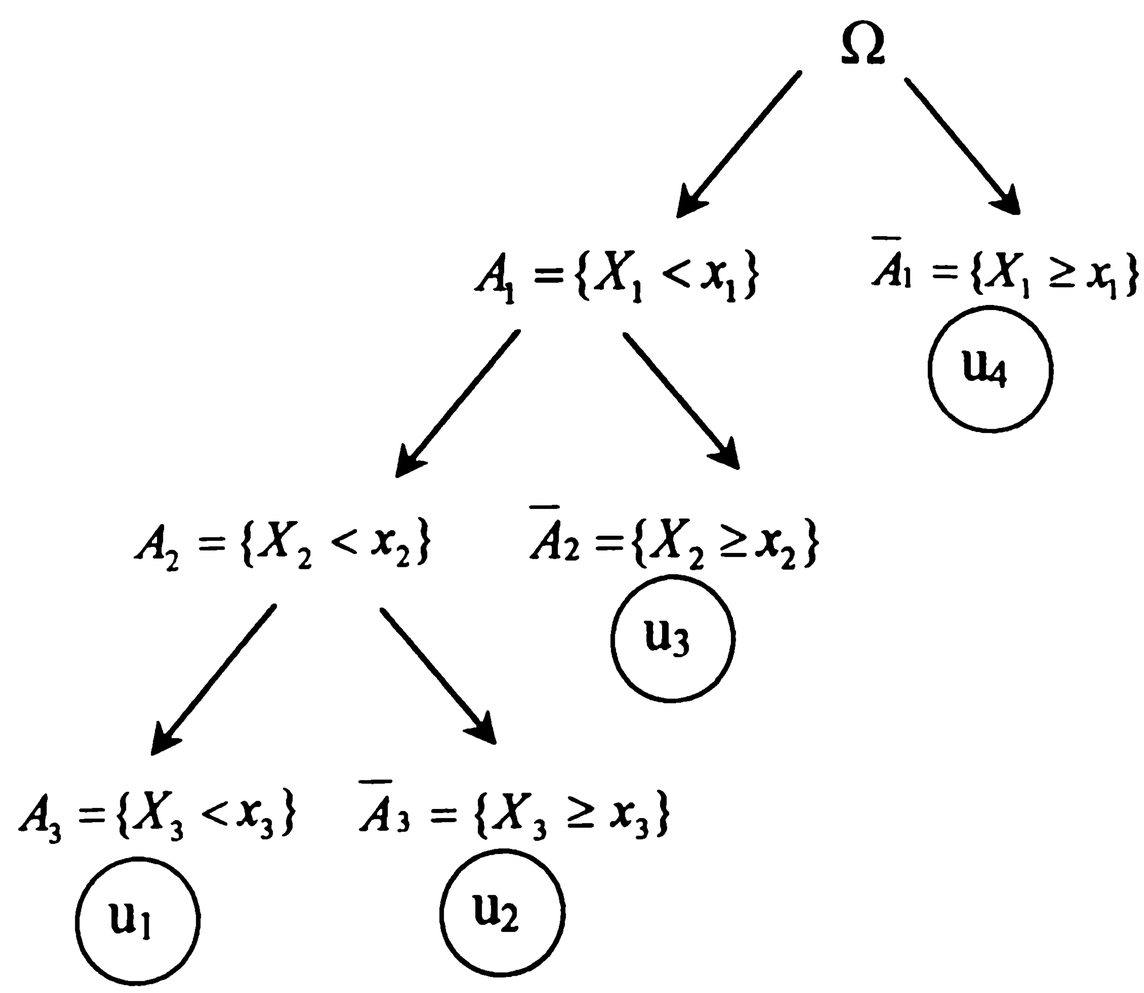
$$\overline{A_2^0}=\{x_2: E\{[u_3(T)-u_1(T)]/x_2,x_1\} - \int_{\overline{A_3^0}} E\{[u_2(T)-u_1(T)]/x_3,x_2,x_1\}$$

$$.f_3(x_3/x_2,x_1)dx_3 \ge 0\}.$$



$$f(x_1, x_2, x_3) = f_1(x_1).f_2(x_2 / x_1).f_3(x_3 / x_1, x_2)$$

Фигура 6.6 *Бейсова мрежа.*



Фигура 6.7 U-*мрежа.*

Означаваме

$$A_2.tail = \int_{\bar{A}_3^0} E\{[u_2(T) - u_1(T)] / x_3, x_2, x_1\}.f_3(x_3 / x_2, x_1)dx_3,$$

$$A_1.tail = \int_{\bar{A}_2^0} \{E\{[u_3(T) - u_1(T)] / x_2, x_1\} - A_2.tail\}.f_2(x_2 / x_1)dx_2 + A_2.tail.$$

Тогава горната граница на очакваната полезност е

$E\{u_1(T) / x_1\} + A_1.tail = 0.$

Следователно оптималната прагова точка x_2^c е коренът на уравнението

$$E\{[u_3(T)-u_1(T)]/x_2,x_1\}-\int_{\overline{A_3}^0} E\{[u_2(T)-u_1(T)]/x_3,x_2,x_1\}.$$

$$f_3(x_3/x_2,x_1)dx_3=0$$

т.е. $E\{[u_3(T)-u_1(T)]/x_2,x_1\}-A_2.tail=0.$

Разглеждаме двойката листа $(A_1,\overline{A_1})$.

По аналогичен начин получаваме, че оптималната прагова точка x_1^c е коренът на уравнението

$$E\{[u_4(T)-u_1(T)]/x_1\}-\int_{\overline{A_2}^0}\{E\{[u_3(T)-u_1(T)]/x_2,x_1\}-$$

$$-\int_{\overline{A_3}^0} E\{[u_2(T)-u_1(T)]/x_3,x_2,x_1\}.f_3(x_3/x_2,x_1)dx_3\}f_2(x_2/x_1)dx_2-$$

$$-\int_{\overline{A_3}^0} E\{[u_2(T)-u_1(T)]/x_3,x_2,x_1\}.f_3(x_3/x_2,x_1)dx_3=0$$

т.е. $E\{[u_4(T)-u_1(T)]/x_2,x_1\}-A_1.tail=0,$

където

$$A_2.tail=\int_{\overline{A_3}^0} E\{[u_2(T)-u_1(T)]/x_3,x_2,x_1\}.f_3(x_3/x_2,x_1)dx_3,$$

$$A_1.tail=\int_{\overline{A_2}^0}\{E\{[u_3(T)-u_1(T)]/x_2,x_1\}-A_2.tail\}f_2(x_2/x_1)dx_2+A_2.tail.$$

Получихме, че оптималните прагови точки са съответно корените на уравненията:

$$E\{[u_2(T)-u_1(T)]/x_3,x_2,x_1\}=0 \tag{6.4}$$

$$E\{[u_3(T)-u_1(T)]/x_2,x_1\}-\int_{\overline{A_3}^0} E\{[u_2(T)-u_1(T)]/x_3,x_2,x_1\}.$$

$$.f_3(x_3/x_2,x_1)dx_3=0 \tag{6.5}$$

$$E\{[u_4(T)-u_1(T)]/x_1\}-\int_{\overline{A_2}^0}\{E\{[u_3(T)-u_1(T)]/x_2,x_1\}-$$

$$-\int_{\bar{A}_3^0} E\{[u_2(T)-u_1(T)]/x_3,x_2,x_1\}.f_3(x_3/x_2,x_1)dx_3\}f_2(x_2/x_1)dx_2-$$

$$-\int_{\bar{A}_3^0} E\{[u_2(T)-u_1(T)]/x_3,x_2,x_1\}.f_3(x_3/x_2,x_1)dx_3 = 0 \qquad (6.6)$$

т.е. оптималните прагови точки са съответно корените на уравненията:

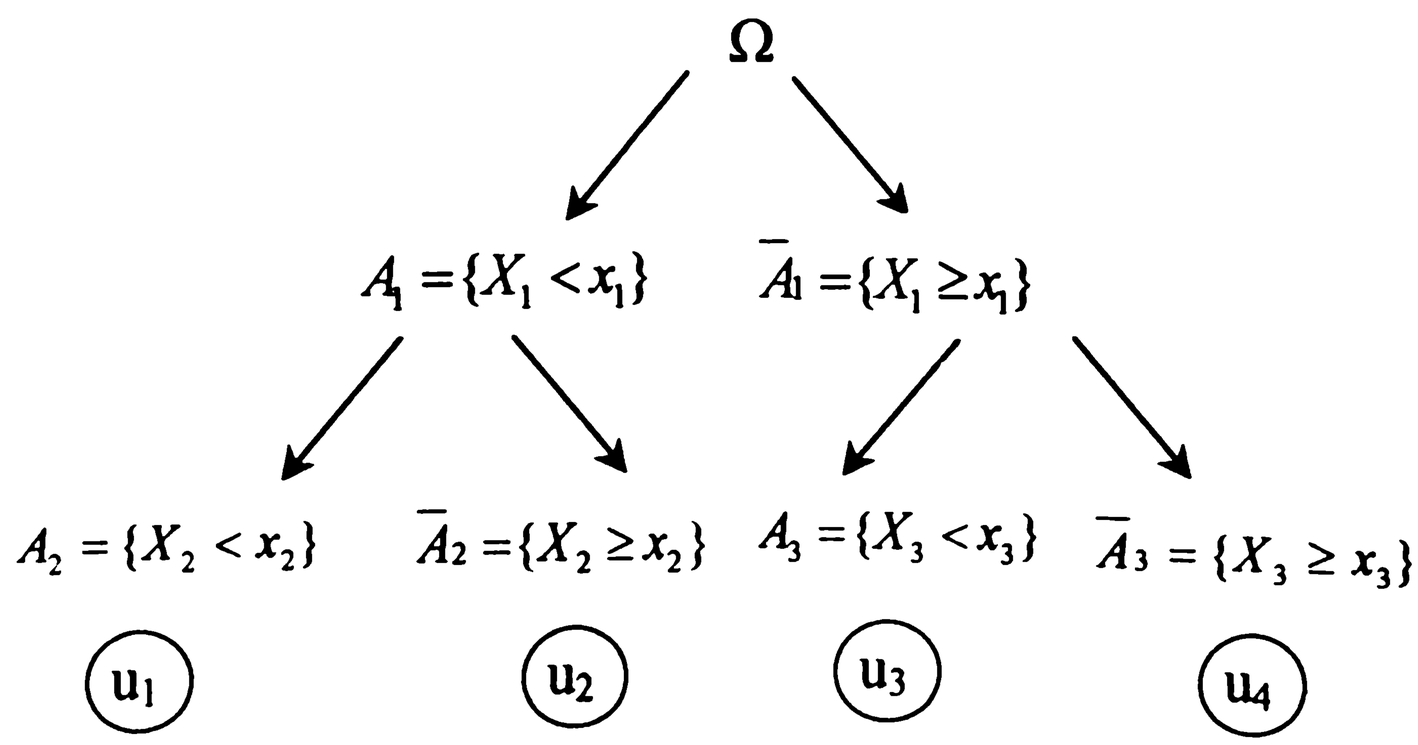
$$E\{[u_2(T)-u_1(T)]/x_3,x_2,x_1\} = 0$$

$$E\{[u_3(T)-u_1(T)]/x_2,x_1\} - A_2.tail = 0$$

$$E\{[u_4(T)-u_1(T)]/x_1\} - A_1.tail = 0.$$

Трети случай:

Нека $X=\{X_1, X_2, X_3\}$ е множество от непрекъснати случайни величини. Нека вероятностният модел се задава с Бейсовата мрежа от *фигура 6.6* . Нека с множеството $X=\{X_1, X_2, X_3\}$ е асоциирана U-*мрежата от фигура 6.8.*



Фигура 6.8 *U-мрежа.*

По аналогичен начин получаваме, че оптималните прагови точки са съответно корените на уравненията

$$E\{[u_4(T)-u_3(T)]/x_3,x_1\} = 0 \qquad (6.7)$$

$$E\{[u_2(T)-u_1(T)]/x_2,x_1\} = 0 \qquad (6.8)$$

$$E\{[u_3(T)-u_1(T)]/x_1\}+\int_{\bar{A}_3^0} E\{[u_4(T)-u_3(T)]/x_3,x_1\}.f_3(x_3/x_1)dx_3-$$

$$- \int_{\overline{A}_2^0} E\{[u_2(T)- u_1(T)] /x_2,x_1\} f_2(x_2/x_1)dx_2 = 0. \qquad (6.9)$$

т.е. $E\{[u_3(T)- u_1(T)] /x_1\} + \overline{A}_1.tail - A_1.tail = 0,$

където

$$\overline{A}_1.tail = \int_{\overline{A}_3^0} E\{[u_4(T)- u_3(T)] /x_3,x_1\} f_3(x_3/x_1)dx_3,$$

$$A_1.tail = \int_{\overline{A}_2^0} E\{[u_2(T)- u_1(T)] /x_2,x_1\} f_2(x_2/x_1)dx_2.$$

Ще разгледаме *общия случай*:

Нека $X=\{X_1, X_2, \ldots, X_n\}$ е множество от непрекъснати случайни величини.

Разглеждаме произволна U-*мрежа*, асоциирана с множеството $X=\{X_1, X_2, \ldots, X_n\}$. Ще оценяваме очакваната полезност постъпково.

Разглеждаме множеството от двойките листа $(A_i, \overline{A_i})$ с максимален номер на равнище в дървото на събитията. От това множество избираме елемента с най-голям индекс i.

Нека с получената двойка листа $(A_i, \overline{A_i})$ са асоциирани възли полезност, които представят съответно функциите полезност $u_i(t)$ и $u_{i+1}(t)$.

Нека с листото A_i е асоциран фактора F_i. Следователно с двойката листа $(A_i, \overline{A_i})$ е асоциран списъкът от случайни величини $\Psi(F_i)$.

Тогава разглеждаме очакваната полезност като функция на праговата точка x_i^c:

$$Eu(\overline{A_i}) = \int_{A_i} E[u_i(T)/\Psi(F_i)] f_i(x_i/[\Psi(F_i)\setminus x_i])dx_i +$$

$$+ \int_{\overline{A}_i} E[u_{i+1}(T)/\Psi(F_i)] f_i(x_i/[\Psi(F_i)\setminus x_i])dx_i =$$

$$= E\{u_i(T)/ [\Psi(F_i)\setminus x_i]\} +$$

$$+\int_{\overline{A_i}} E\{[u_{i+1}(T)-u_i(T)]/\Psi(F_i)\}\, f_i(x_i/[\Psi(F_i)\setminus x_i])dx_i$$

Една горна граница на очакваната полезност е:

$$Eu(\overline{A_i}) \le E\{u_i(T)/[\Psi(F_i)\setminus x_i]\} +$$

$$+\int_{\overline{A_i^0}} E\{[u_{i+1}(T)-u_i(T)]/\Psi(F_i)\}\, f_i(x_i/[\Psi(F_i)\setminus x_i])dx_i,$$

където $\overline{A_i^0} = \{x_i: E\{[u_{i+1}(T)-u_i(T)]/\Psi(F_i)\} \ge 0\}$.

Следователно оптималната прагова точка x_i^c е коренът на уравнението

$E\{[u_{i+1}(T)-u_i(T)]/\Psi(F_i)\} = 0$.

Нека $\pi(A_i)$ е родителят на двойката листа $(A_i, \overline{A_i})$.

Над дървото на събитията ще извършим операцията изключване на елемент - листо. Ще изключим листата A_i и $\overline{A_i}$. С новополученото листо $\pi(A_i)$ ще свържем получената горна граница на функцията полезност $Eu(\overline{A_i})$, т.е. с новополученото листо $\pi(A_i)$ ще асоциираме функцията полезност $u^*(t)$, която има следния вид:

$$u^*(t) = E\{u_i(T)/[\Psi(F_i)\setminus x_i]\} +$$

$$+\int_{\overline{A_i^0}} E\{[u_{i+1}(T)-u_i(T)]/\Psi(F_i)\}\, f_i(x_i/[\Psi(F_i)\setminus x_i])dx_i.$$

Получихме ново дърво на събитията, което е поддървовидна структура на дървото на събитията, и ново множество от възли-полезност, т.е. получихме нова U-мрежа.

По аналогичен начин разглеждаме множеството от двойките листа с максимален номер на равнище в полученото дърво на събитията и от това множество избираме елемента с най-голям индекс.

Първо ще предположим, че тази двойка е $(A_{i-1}, \overline{A_{i-1}})$, където $\overline{A_{i-1}} = \pi(A_i)$. Нека с листото A_{i-1} е асоциран фактора F_{i-1}.

Следователно с всяко листо от двойката листа $(A_{i-1}, \overline{A_{i-1}})$ е асоциран списъкът от случайни величини $\Psi(F_{i-1})$. Тогава $\Psi(F_{i-1}) = \Psi(F_i) \backslash X_i$.

Знаем, че с листата A_{i-1} и $\overline{A_{i-1}}$ са асоциирани възли полезност, които представят съответно функциите $u_{i-1}(t)$ и $u^*(t)$.

Тогава очакваната полезност е:

$$Eu(\overline{A_{i-1}}, \overline{A_i}^0) =$$

$$= \int_{A_{i-1}} E[u_{i-1}(T)/\Psi(F_{i-1})].f_{i-1}(x_{i-1}/[\Psi(F_{i-1}) \backslash x_{i-1}])dx_{i-1} +$$

$$+ \int_{\overline{A}_{i-1}} E[u_i(T)/\Psi(F_{i-1})].f_{i-1}(x_{i-1}/[\Psi(F_{i-1}) \backslash x_{i-1}])dx_{i-1} +$$

$$+ \int_{\overline{A}_{i-1}} \{ \int_{\overline{A}_i^0} E\{[u_{i+1}(T) - u_i(T)]/\Psi(F_i)\}.f_i(x_i)dx_i\}.$$

$$.f_{i-1}(x_{i-1}/[\Psi(F_{i-1}) \backslash x_{i-1}])dx_{i-1} =$$

$$= E\{u_{i-1}(T)/[\Psi(F_{i-1}) \backslash X_{i-1}]\} +$$

$$+ \int_{\overline{A}_{i-1}} \{ E\{[u_i(T) - u_{i-1}(T)]/\Psi(F_{i-1})\} +$$

$$+ \int_{\overline{A}_i^0} E\{[u_{i+1}(T) - u_i(T)]/\Psi(F_i)\}.f_i(x_i/[\Psi(F_i) \backslash x_i])dx_i\}.$$

$$f_{i-1}(x_{i-1}/[\Psi(F_{i-1}) \backslash x_{i-1}])dx_{i-1}$$

Една горна граница на очакваната полезност е:

$$Eu(\overline{A_{i-1}}, \overline{A_i}^0) \le E\{u_{i-1}(T)/[\Psi(F_{i-1}) \backslash X_{i-1}]\} +$$

$$+ \int_{\overline{A}_{i-1}^0} \{ E\{[u_i(T) - u_{i-1}(T)]/\Psi(F_{i-1})\} +$$

$$+ \int_{\overline{A}_i^0} E\{[u_{i+1}(T) - u_i(T)]/\Psi(F_i)\}.f_i(x_i/[\Psi(F_i) \backslash x_i])dx_i\}.$$

$$f_{i-1}(x_{i-1}/[\Psi(F_{i-1})\setminus x_{i-1}])dx_{i-1},$$

където

$$\overline{A}^0_{i-1} = \{x_{i-1}: E\{[u_i(T)- u_{i-1}(T)]/\Psi(F_{i-1})\} +$$

$$+ \int_{\overline{A}_i^0} E\{[u_{i+1}(T)- u_i(T)]/\Psi(F_i)\} f_i(x_i/[\Psi(F_i)\setminus x_i])dx_i \geq 0\}.$$

Следователно оптималната прагова точка x_{i-1}^c е коренът на уравнението

$$E\{[u_i(T)- u_{i-1}(T)]/\Psi(F_{i-1})\} +$$

$$+ \int_{\overline{A}_i^0} E\{[u_{i+1}(T)- u_i(T)]/\Psi(F_i)\} f_i(x_i/[\Psi(F_i)\setminus x_i])dx_i = 0.$$

Да предположим сега, че двойката възли с най-голям номер на равнище и с най-голям индекс е $(A_{i-1}, \overline{A_{i-1}})$, където $A_{i-1} = \pi(A_i)$. Знаем, че с тази двойка листа е асоциирана структурата полезност

$$u(t)= \begin{cases} E\{u_i(T)/ [\Psi(F_i)\setminus x_i]\} + \int_{\overline{A}^0} E\{[u_{i+1}(T)- u_i(T)]/\Psi(F_i)\} f_i(x_i/\Psi(F_{i-1}))dx_i \\ u_{i+2}(t) \end{cases}$$

Тогава очакваната полезност е:

$$Eu(\overline{A_{i-1}}, \overline{A_i}^0) = \int_{A_{i-1}} E[u_i(T)/\Psi(F_{i-1})] f_{i-1}(x_{i-1}/[\Psi(F_{i-1})\setminus x_{i-1}])dx_{i-1} +$$

$$+ \int_{A_{i-1}} \int_{\overline{A}_i^0} E\{[u_{i+1}(T)- u_i(T)]/\Psi(F_i)\} f_i(x_i/\Psi(F_{i-1}))dx_i.$$

$$. f_{i-1}(x_{i-1}/[\Psi(F_{i-1})\setminus x_{i-1}])dx_{i-1}+$$

$$+ \int_{\overline{A}_{i-1}} E[u_{i+2}(T)/\Psi(F_{i-1})] f_{i-1}(x_{i-1}/\Psi(F_{i-1}\setminus x_{i-1}))dx_{i-1} =$$

$$= E\{u_i(T)/[\Psi(F_{i-1})\setminus x_{i-1}]\} +$$

$$+ \int_{\overline{A}_i^0} E\{[u_{i+1}(T)- u_i(T)]/\Psi(F_i)\setminus x_{i-1})\} f_i(x_i/\Psi(F_{i-1})\setminus x_{i-1})dx_i -$$

$$- \int_{A_{i-1}} \{ \int_{\overline{A_i}^0} E\{[u_{i+1}(T)- u_i(T)]/[\Psi(F_i)\backslash x_{i-1}]\} . f_i(x_i/[\Psi(F_{i-1})\backslash x_{i-1}])dx_i .$$

$$. f_{i-1}(x_{i-1}/[\Psi(F_{i-1})\backslash x_{i-1}])dx_{i-1} +$$

$$+ \int_{\overline{A_{i-1}}} E\{[u_{i+2}(T)-u_i(T)]/\Psi(F_{i-1})\} . f_{i-1}(x_{i-1}/[\Psi(F_{i-1})\backslash x_{i-1}])dx_{i-1}$$

Една горна граница на очакваната полезност е:

$$Eu(\overline{A_{i-1}}, \overline{A_i}^0) \leq E\{u_i(T)/ [\Psi(F_{i-1})\backslash x_{i-1}]\} +$$

$$+ \int_{\overline{A_i}^0} E\{[u_{i+1}(T)- u_i(T)]/[\Psi(F_i) \backslash x_{i-1}]\} . f_i(x_i/[\Psi(F_{i-1}) \backslash x_{i-1}])dx_i +$$

$$+ \int_{\overline{A_{i-1}}^0} \{E\{[u_{i+2}(T)-u_i(T)]/\Psi(F_{i-1})\} -$$

$$- \int_{\overline{A_i}^0} E\{[u_{i+1}(T)- u_i(T)]/\Psi(F_i)\} . f_i(x_i/\Psi(F_{i-1}))dx_i \} .$$

$$. f_{i-1}(x_{i-1}/[\Psi(F_{i-1})\backslash x_{i-1}])dx_{i-1} ,$$

където

$$\overline{A_{i-1}^0} = \{x_{i-1}: E\{[u_{i+2}(T)-u_i(T)]/\Psi(F_{i-1})\} -$$

$$- \int_{\overline{A_i}^0} E\{[u_{i+1}(T)- u_i(T)]/\Psi(F_i)\} . f_i(x_i/\Psi(F_{i-1}))dx_i \} \geq 0\} .$$

Следователно в този случай оптималната прагова точка x_{i-1}^c е коренът на уравнението:

$$E\{[u_{i+2}(T)-u_i(T)]/\Psi(F_{i-1})\} -$$

$$- \int_{\overline{A_i}^0} E\{[u_{i+1}(T)- u_i(T)]/\Psi(F_i)\} . f_i(x_i/\Psi(F_{i-1}))dx_i = 0 .$$

Премахваме листата A_i и $\overline{A_i}$. С новополученото листо $\pi(A_i)$ ще свържем получената горна граница на функцията полезност $Eu(\overline{A_i})$.

Повтаряме тази процедура докато получим изродено дърво на събитията, състоящо се само от един елемент - корена на първоначалното дърво на събитията. С този елемент е асоциирана горна граница на очакваната полезност. Но нашата цел е постъпково генериране на системата уравнения, корените на които са оптималните прагови точки x_i^c, i=1,2,…,n. Затова в процеса на генериране на системата уравнения с всяко новополучено листо ще асоциираме символни низове, необходими за генерирането на търсените уравнения.

След n-тата стъпка ще получим система от n уравнения, корените на която са праговите точки x_i^c, i=1,2,…,n.

<u>Алгоритъм 6.1:</u> *<u>Алгоритъм за генериране на уравнения в символен вид:</u>*

Вход:

- U-мрежа
- Бейсова мрежа

Изход: n уравнения в символен вид

Инициализация: С всяко листо A_i на дървото на събитията асоциираме две символни променливи

$A_i.head := U(A_i)$

$A_i.tail := NIL$

Стъпки:

1. Намираме елемента $(A_i, \overline{A_i})$ с най-голям индекс i от множеството от двойките листа с максимален номер на равнище в дървото на събитията.
2. Намираме $\psi(A_i)$ - списъкът от случайните величини, асоциирани с листото A_i, и $i = \varepsilon(A_i)$.
3. Намираме лявата страна на търсеното уравнение

   $left := \underline{E}\{[\overline{A_i}.head\,\underline{(T)} - A_i.head\,\underline{(T)}]/\psi(A_i)\} + \overline{A_i}.tail - A_i.tail$
4. Генерираме уравнението:

   $left = 0$
5. Намираме родителя на двойката листа $(A_i, \overline{A_i})$ $A_k = \pi(A_i)$. Ако това е коренът на дървото, премини към *стъпка* **9**. В противен случай обозначи $\overline{A_i^0} = [x_i^c;1]$ и намери списъка на случайните величини $\psi(A_k) = \psi(A_i) \backslash x_i$.

6. Ако $(A_i, \overline{A_i})$ е десен наследник на $\pi(A_i)$, т.е. $\pi(A_i)=\overline{A_k}$, присвояваме

$\overline{A_k}$.**head**:= $\boldsymbol{A_i}$.**head**,

$$\overline{A_k}\ .\mathbf{tail} := \int_{\overline{A_i}^0} \{\mathbf{left}\}.\underline{f_i(x_i/\psi(A_k))\mathrm{d}\,x_i} \underline{+} A_i\ .\mathbf{tail}$$

В противен случай, т.е. ако $(A_i, \overline{A_i})$ е ляв наследник на $\pi(A_i)$, т.е. $\pi(A_i)=A_k$, присвояваме

$\boldsymbol{A_k}$.**head**:= $\boldsymbol{A_i}$.**head**,

$$A_k\ .\mathbf{tail} := \int_{\overline{A_i}^0} \{\mathbf{left}\}.\underline{f_i(x_i/\psi(A_k))\mathrm{d}\,x_i} \underline{+} A_i\ .\mathbf{tail}$$

7. Премахваме листата A_i и $\overline{A_i}$.
   Получихме уравнение за x_i^c и нова U-мрежа, чието дърво е поддърво на дървото на събитията.
8. Повтаряме горните стъпки, като започнем от *стъпка* 1.
9. Край.

В описания алгоритъм със знака := сме означили операцията присвояване стойност на символна променлива, а операцията присвояване стойност на числова променлива сме означили със знака =. Операцията конкатенация не е зададена явно. За прегледност символните низове в операцията конкатенация са подчертани с права линия, а символните променливи - с по-тъмен шрифт.

6.2 Основна теорема

Разглеждаме произволно дърво на събитията с $n+1$ листа. Нека да намерим горна граница на очакваната полезност стъпка по стъпка. На всяка стъпка k, $k=1,2,\dots,n$, ще извършим следните действия:

Разглеждаме една двойка листа и асоциираните с нея два възела полезност. Намираме една горна граница на очакваната полезност. Премахваме двойката листа, като получената горна граница на очакваната полезност асоциираме с техния родител. Но нашата цел не е да намерим горна граница на очакваната полезност, а да генерираме уравнения за оптималните прагови стойности. За тази цел се оказва, че е достатъчно в две символни променливи по подходящ начин да запазим необходимата информация и да свържем тези две променливи с родителя на листата. Нещо повече, информацията, съдържаща се в тези две променливи е достатъчна за да възстановим горната граница на очакваната полезност на всяка стъпка.

С всяко листо A_k, $k=1,..,n+1$, на това дърво асоциираме две символни променливи $A_k.head$ и $A_k.tail$, в които запазваме информацията, необходима за генериране на системата уравнения.

Инициализираме дървото, като за всяко листо A_k, $k=1,..,n+1$, тези две променливи приемат стойности съответно $A_k.head = u_k(t)$, $A_k.tail = \phi$, където $u_k(t)$ е функцията полезност, асоциирана с листото A_k, а ϕ е празния низ.

Разглеждаме U-мрежа с инициализирано дърво на събитията. Нека $A_{j-1} = \pi(A_j)$ и $\overline{A_{j-1}} = \pi(A_i)$. Прилагаме *Алгоритъм 6.1*. Тогава е в сила следната теорема:

Теорема 6.1: Нека за всяко $j=1,2,\dots,n$ функцията

$$E\{[\overline{A_{j-1}}.head(T) - A_{j-1}.head(T)]/\psi(A_{j-1})\} + \overline{A_{j-1}}.tail - A_{j-1}.tail$$

сменя знака си от отрицателен към положителен точно един път, където

$$A_{j-1}.head = A_j.head,$$

$$A_{j-1}.tail = \int_{\overline{A_j^0}} \{E\{[\overline{A_j}.head(T) - A_j.head(T)]/\psi(A_j)\} +$$

$$+ \overline{A_j}.tail - A_j.tail\} f_j(x_j/\psi(A_{j-1}))dx_j + A_j.tail,$$

$$\overline{A_j^0} = \{x_j : E\{[\overline{A_j}.head(T) - A_j.head(T)]/\psi(A_{j-1})\} + \overline{A_j}.tail - A_j.tail \geq 0\};$$

$$\overline{A_{j-1}}.head = A_i.head,$$

$$\overline{A_{j-1}}.tail = \int_{\overline{A_i^0}} \{E\{[\overline{A_i}.head\,(T) - A_i.head\,(T)]/\psi(A_i)\} +$$

$$+ \overline{A_i}.tail - A_i.tail\}.f_i(x_i/\psi(A_{j-1}))dx_i + A_i.tail,$$

$$\overline{A_i^0} = \{x_i : E\{[\overline{A_i}.head(T) - A_i.head(T)]/\psi(A_i)\} + \overline{A_i}.tail - A_i.tail \geq 0\}.$$

Тогава за всяко *j, j=1,...,n* праговата стойност x_{j-1}^c, асоциирана с двойката възли $(A_{j-1}, \overline{A_{j-1}})$, удовлетворява уравнението

$$E\{[\overline{A_{j-1}}.head(T) - A_{j-1}.head(T)]/\psi(A_{j-1})\} + \overline{A_{j-1}}.tail - A_{j-1}.tail = 0$$

и *Алгоритъм 6.1* правилно ще генерира уравненията.

Доказателство:

Основните единици, от който е съставена една U-мрежа ще наричаме U-единици. Трите вида U-единици са *симетрична, дясно асиметрична и ляво асиметрична* и са представени на *фигура 6.9*.

1. Отначало ще разгледаме **частния случай** когато U-единиците са U-мрежи.

1.1 Разглеждаме U-единицата от *фигура 6.9а* с инициализирано дърво на събитията. Ще оценяваме очакваната полезност постъпково.

$$Eu(\overline{A_j}) = E\{u_j(T)/\psi(F_{j-1})\} + \int_{\overline{A_j}} E\{[u_{j+1}(T) - u_j(T)]/\psi(F_j)\}.f_j(x_j/\psi(F_{j-1}))dx_j$$

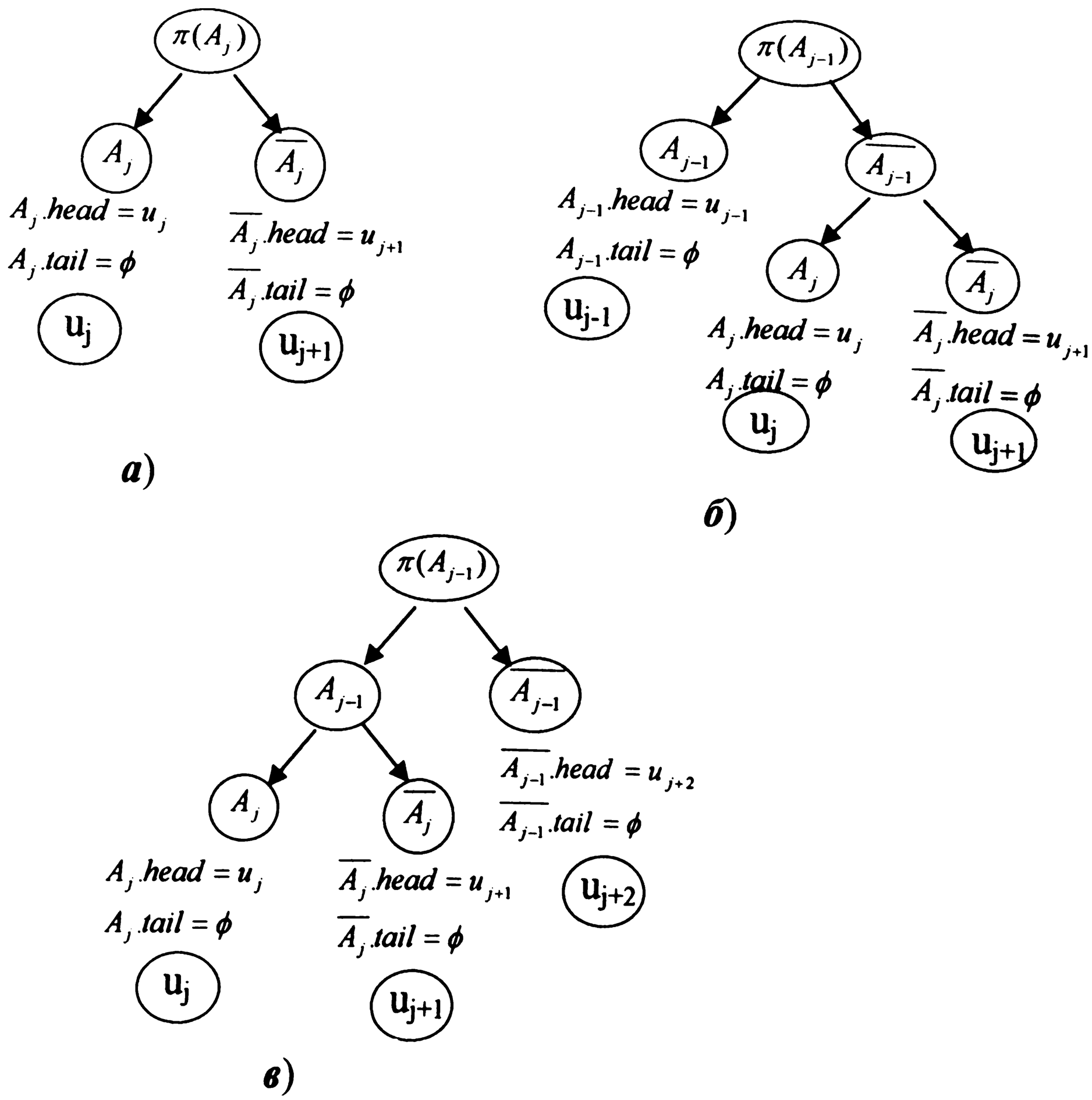
От направената инициализация следва, че

$$A_j.head = u_j(t),\quad A_j.tail = \phi,\quad \overline{A_j}.head = u_{j+1}(t),\quad \overline{A_j}.tail = \phi$$

и очакваната полезност може да се представи във вида:

$$Eu(\overline{A_j}) = E\{u_j(T)/\psi(F_{j-1})\} + \int_{\overline{A_j}} \{E\{[\overline{A_j}.head(T) - A_j.head(T)]/\psi(F_j)\} +$$

$$+ \overline{A_j}.tail - A_j.tail\} f_j(x_j/\psi(F_{j-1}))dx_j =$$



Фигура 6.9 *Инициализирани U-единици:*
а) симетрична U-единица
б) дясно асиметрична U-единица
в) ляво асиметрична U-единица.

$$= E\{A_j.head(T)/\psi(F_{j-1})\} + \int_{\overline{A_j}} \{E\{[\overline{A_j}.head(T) - A_j.head(T)]/\psi(F_j)\} +$$

$$+ \overline{A_j}.tail - A_j.tail\} f_j(x_j/\psi(F_{j-1}))dx_j .$$

Една горна граница на очакваната полезност е:

$$Eu(\overline{A_j^0}) =$$

$$= E\{A_j.head(T)/\psi(F_{j-1})\} + \int_{\overline{A_j^0}} \{E\{[\overline{A_j}.head(T) - A_j.head(T)]/\psi(F_j)\} +$$

$$+ \overline{A_j}.tail - A_j.tail\} f_j(x_j/\psi(F_{j-1}))dx_j ,$$

където

$$\overline{A_j^0} = \{x_j : E\{[\overline{A_j}.head(T) - A_j.head(T)]/\psi(F_j)\} + \overline{A_j}.tail - A_j.tail \geq 0\}.$$

От условието на теоремата знаем, че функцията от лявата страна на горното неравенство сменя знака си от отрицателен към положителен точно един път.

Тогава уравнението за x_j^c има вида:

$$E\{[\overline{A_j}.head(T) - A_j.head(T)]/\psi(F_j)\} + \overline{A_j}.tail - A_j.tail = 0$$

(*виж Стъпка 3* и *Стъпка 4* на алгоритъма).

1.2 Разглеждаме **U**-единицата от *фигура 6.9б* с инициализирано дърво на събитията. Използваме резултата от **1.1**. Нека в променливите, асоциирани с възела $\overline{A_{j-1}} = \pi(A_j)$ да запазим необходимата информация по следния начин:

$$\overline{A_{j-1}}.head = A_j.head = u_j$$ и

$$\overline{A_{j-1}}.tail = \int_{\overline{A_j^0}} \{E\{[\overline{A_j}.head(T) - A_j.head(T)]/\psi(F_j)\} +$$

$$+ \overline{A_j}.tail - A_j.tail\}.f_j(x_j/\psi(F_{j-1}))dx_j + A_j.tail$$, където

$$\overline{A_j^0} = \{x_j : E\{[\overline{A_j}.head(T) - A_j.head(T)]/\psi(F_j)\} + \overline{A_j}.tail - A_j.tail \geq 0\}$$

(*виж Стъпка 6*).

Тогава получената горна граница на очакваната полезност може да се представи по следния начин:

$$Eu(\overline{A_j^0}) = E\{\overline{A_{j-1}}.head(T)/\psi(F_{j-1})\} + \overline{A_{j-1}}.tail \qquad (6.15)$$

Премахваме двойката листа $(A_j, \overline{A_j})$ (*виж Стъпка 7* на *Алгоритъм 6.1*). Възелът $\pi(A_j) \equiv \pi(\overline{A_j})$ носи цялата информация, необходима за генериране на следващите уравнения.

Продължаваме да оценяваме очакваната полезност постъпково.

$$Eu(\overline{A_{j-1}}, \overline{A_j^o}) = E\{u_{j-1}(T)/\psi(F_{j-2})\} +$$

$$+ \int_{\overline{A_{j-1}}} \{E\{[u_j(T) - u_{j-1}(T)]/\psi(F_{j-1})\} + \int_{\overline{A_j^0}} E\{[u_{j+1}(T) - u_j(T)]/\psi(F_j)\}$$

$$.f_j(x_j/\psi(F_{j-1}))dx_j\}.f_{j-1}(x_{j-1}/\psi(F_{j-2}))dx_{j-1}$$

От направената инициализация следва, че

$$A_{j-1}.head = u_{j-1}(t), \quad A_{j-1}.tail = \phi$$

и очакваната полезност може да се представи във вида:

$$Eu(\overline{A_{j-1}}, \overline{A_j^o}) = E\{u_{j-1}(T)/\psi(F_{j-2})\} +$$

$$+ \int_{\overline{A_{j-1}}} \{E\{[A_{j-1}.head(T) - A_{j-1}.head(T)]/\psi(F_{j-1})\} +$$

$$+ \overline{A_{j-1}}.tail - A_{j-1}.tail\}.f_{j-1}(x_{j-1}/\psi(F_{j-2}))dx_{j-1} =$$

$$= E\{A_{j-1}.head(T)/\psi(F_{j-2})\} +$$

$$+ \int_{\overline{A_{j-1}}} \{E\{[A_{j-1}.head(T) - A_{j-1}.head(T)]/\psi(F_{j-1})\} +$$

$$+ \overline{A_{j-1}}.tail - A_{j-1}.tail\}.f_{j-1}(x_{j-1}/\psi(F_{j-2}))dx_{j-1}.$$

Една горна граница на очакваната полезност е:

$$Eu(\overline{A_{j-1}^0}, \overline{A_j^o}) =$$

$$= E\{A_{j-1}.head(T)/\psi(F_{j-2})\} +$$

$$+ \int_{\overline{A_{j-1}^0}} \{E\{[A_{j-1}.head(T) - A_{j-1}.head(T)]/\psi(F_{j-1})\} +$$

$$+ \overline{A_{j-1}}.tail - A_{j-1}.tail\}.f_{j-1}(x_{j-1}/\psi(F_{j-2}))dx_{j-1},$$

където

$$\overline{A_{j-1}^o} = \{x_{j-1} : E\{[\overline{A_{j-1}}.head(T) - A_{j-1}.head(T)]/\psi(F_{j-1})\} +$$

$$+ \overline{A_{j-1}}.tail - A_{j-1}.tail \geq 0\}.$$

Да предположим, че функцията от лявата страна на горното неравенство сменя знака си от отрицателен към положителен точно един път.

Тогава уравнението за x^c_{j-1} има вида:

$$E\{[\overline{A_{j-1}}.head(T) - A_{j-1}.head(T)]/\psi(F_{j-1})\} + \overline{A_{j-1}}.tail - A_{j-1}.tail = 0$$

1.2.1 Нека $\pi(A_{j-1}) = \overline{A_{j-2}}$. В променливите, асоциирани с възела $\overline{A_{j-2}} = \pi(A_{j-1})$ запазваме необходимата информация по следния начин:

$$\overline{A_{j-2}}.head = A_{j-1}.head$$

и

$$\overline{A_{j-2}}.tail = \int_{\overline{A^0_{j-1}}} \{E\{[\overline{A_{j-1}}.head(T) - A_{j-1}.head(T)]/\psi(F_{j-1})\} + \overline{A_{j-1}}.tail - A_{j-1}.tail\}.f_{j-1}(x_{j-1}/\psi(F_{j-2}))dx_{j-1} + A_{j-1}.tail,$$

където

$$\overline{A^0_{j-1}} = \{x_{j-1} : E\{[\overline{A_{j-1}}.head(T) - A_{j-1}.head(T)]/\psi(F_{j-1})\} + \overline{A_{j-1}}.tail - A_{j-1}.tail \geq 0\}$$

Тогава една горна граница на очакваната полезност е:

$$Eu(\overline{A^0_{j-1}}, \overline{A^0_j}) = E\{\overline{A_{j-2}}.head(T)/\psi(F_{j-2})\} + \overline{A_{j-2}}.tail$$

Видът на тази горна граница е аналогичен на (6.15).

1.2.2 Нека $\pi(A_{j-1}) = A_{j-2}$.

В променливите, асоциирани с възела $A_{j-2} = \pi(A_{j-1})$ запазваме необходимата информация по следния начин:

$$A_{j-2}.head = A_{j-1}.head$$

и

$$A_{j-2}.tail = \int_{\overline{A^0_{j-1}}} \{E\{[\overline{A_{j-1}}.head(T) - A_{j-1}.head(T)]/\psi(F_{j-1})\} + \overline{A_{j-1}}.tail - A_{j-1}.tail\}.f_{j-1}(x_{j-1}/\psi(F_{j-2}))dx_{j-1} + A_{j-1}.tail$$

(*виж Стъпка 6*).

Тогава една горна граница на очакваната полезност е:

$$Eu(\overline{A^0_{j-1}}, \overline{A^0_j}) = E\{A_{j-2}.head(T)/\psi(F_{j-2})\} + A_{j-2}.tail \quad . \qquad (6.16)$$

Премахваме двойката листа $(A_{j-1}, \overline{A_{j-1}})$ (*виж Стъпка 7* на *Алгоритъм 6.1*). Възелът $\pi(A_{j-1}) \equiv \pi(\overline{A_{j-1}})$ носи цялата информация, необходима за генериране на следващите уравнения.

1.3 Разглеждаме **U**-единицата от *фигура 6.9в* с инициализирано дърво на събитията. По аналогичен начин ще оценим очакваната полезност постъпково.

Направената инициализация има вида:

$$A_j.head = u_j, \quad A_j.tail = \phi, \quad \overline{A_j}.head = u_{j+1}, \quad \overline{A_j}.tail = \phi.$$

Тогава на първата стъпка очакваната полезност може да се представи във вида:

$$Eu(\overline{A_j}) = E\{u_j(T)/\psi(F_{j-1})\} + \int_{\overline{A_j}} E\{[u_{j+1}(T) - u_j(T)]/\psi(F_j)\} f_j(x_j/\psi(F_{j-1}))dx_j =$$

$$= E\{A_j.head(T)/\psi(F_{j-1})\} + \int_{\overline{A_j}} \{E\{[\overline{A_j}.head(T) - A_j.head(T)]/\psi(F_{j-1})\} +$$

$$+ \overline{A_j}.tail - A_j.tail\} f_j(x_j/\psi(F_{j-1}))dx_j$$

В променливите, асоциирани с възела $A_{j-1} = \pi(A_j)$ запазваме необходимата информация по следния начин:

$$A_{j-1}.head = A_j.head$$

и

$$A_{j-1}.tail = \int_{\overline{A_j^0}} \{E\{[\overline{A_j}.head(T) - A_j.head(T)]/\psi(F_j)\} +$$

$$+ \overline{A_j}.tail - A_j.tail\}.f_j(x_j/\psi(F_{j-1}))dx_j + A_j.tail$$

(*виж Стъпка 6*).

Тогава получаваме следния резултат за горна граница на очакваната полезност:

$$Eu(\overline{A_j^0}) = E\{A_{j-1}.head(T)/\psi(F_{j-1})\} + \overline{A_{j-1}}.tail.$$

Видът на тази горна граница е аналогичен на **(6.16)**.

Уравнението за x_i^c има вида:

$$E\{[\overline{A_j}.head(T) - A_j.head(T)]/\psi(F_{j-1})\} + \overline{A_j}.tail - A_j.tail = 0$$

Премахваме двойката листа $(A_j, \overline{A_j})$ (*виж Стъпка 7* на *Алгоритъм 6.1*) и с техния родител A_{j-1} свързваме получената горна граница на очакваната полезност:

$$Eu(\overline{A_j^0}) = E\{A_{j-1}.head(T)/\psi(F_{j-1})\} + A_{j-1}.tail$$

1.3.1 Нека $\pi(A_{j-1}) = A_{j-2}$.

Тогава на втората стъпка очакваната полезност е:

$$Eu(\overline{A_{j-1}}, \overline{A_j^0}) = E\{A_{j-1}.head(T)/\psi(F_{j-2})\} + A_{j-1}.tail +$$

$$+ \int_{\overline{A_{j-1}}} \{E\{[\overline{A_{j-1}}.head(T) - A_{j-1}.head(T)]/\psi(F_{j-1})\} +$$

$$+ \overline{A_{j-1}}.tail - A_{j-1}.tail\} f_{j-1}(xj_{i-1}/\psi(F_{j-2}))dx_{j-1}.$$

Уравнението за x_{j-1}^c е:

$$E\{[\overline{A_{j-1}}.head(T) - A_{j-1}.head(T)]/\psi(F_{j-1})\} + \overline{A_{j-1}}.tail - A_{j-1}.tail = 0.$$

Присвояваме

$$A_{j-2}.head = A_{j-1}.head$$

$$A_{j-2}.tail = \int_{\overline{A_{j-1}^0}} \{E\{[\overline{A_{j-1}}.head(T) - A_{j-1}.head(T)]/\psi(F_{j-1})\} +$$

$$+ \overline{A_{j-1}}.tail - A_{j-1}.tail\} f_{j-1}(x_{j-1}/\psi(F_{j-2}))dx_{j-1} + A_{j-1}.tail,$$

където

$$\overline{A_{j-1}^0} = \{x_{j-1} : E\{[\overline{A_{j-1}}.head(T) - A_{j-1}.head(T)]/\psi(F_{j-1})\} + \overline{A_{j-1}}.tail - A_{j-1}.tail \geq 0\}.$$

Тогава една горна граница на очакваната полезност е:

$$Eu(\overline{A_{j-1}^0}) = E\{A_{j-2}.head(T)/\psi(F_{j-2})\} + A_{j-2}.tail.$$

Видът на тази горна граница е аналогичен на (6.16).

1.3.2 Нека $\pi(A_{j-1}) = \overline{A_{j-2}}$.

Тогава на втората стъпка очакваната полезност е:

$$Eu(\overline{A_{j-1}}, \overline{A_j^0}) = E\{A_{j-1}.head(T)/\psi(F_{j-2})\} + A_{j-1}.tail +$$

$$+\int_{\overline{A_{j-1}^0}}\{E\{[\overline{A_{j-1}}.head(T)-A_{j-1}.head(T)]/\psi(F_{j-1})\}+$$

$$+\overline{A_{j-1}}.tail+A_{j-1}.tail\}f_{j-1}(x_{j-1}/\psi(F_{j-2}))dx_{j-1}$$

Уравнението за x_{j-1}^c е:

$$E\{[\overline{A_{j-1}}.head(T)-A_{j-1}.head(T)]/\psi(F_{j-1})\}+\overline{A_{j-1}}.tail-A_{j-1}.tail=0.$$

Присвояваме

$$\overline{A_{j-2}}.head=A_{j-1}.head$$

и

$$\overline{A_{j-2}}.tail=\int_{\overline{A_{j-1}^0}}\{E\{[\overline{A_{j-1}}.head(T)-A_{j-1}.head(T)]/\psi(F_{j-1})\}+$$

$$+\overline{A_{j-1}}.tail-A_{j-1}.tail\}.f_{j-1}(x_{j-1}/\psi(F_{j-1}))dx_{j-1}$$

$$+A_{j-1}.tail.$$

Тогава една горна граница на очакваната полезност е:

$$Eu(\overline{A_{j-1}^0})=E\{A_{j-2}.head(T)/\psi(F_{j-2})\}+\overline{A_{j-2}}.tail.$$

Видът на тази горна граница е аналогичен на (6.15).

Премахваме двойката листа $(A_{j-1},\overline{A_{j-1}})$ (*виж Стъпка 7* на *Алгоритъм 6.1*). Така възелът $\pi(A_{j-1})=\pi(\overline{A_{j-1}})$ носи цялата информация, необходима за генериране на следващите уравнения.

2. В **общия случай** листата на U-единиците са възли в U-мрежата (виж *фигура 6.10*).

Да допуснем, че с листото $\overline{A_j}$ е асоциирана функция полезност със следния вид:

$u^*(t)=E\{\overline{A_j}.head(T)/\psi(F_j)\}+\overline{A_j}.tail$ (виж (6.15)).

Да допуснем, че с листото A_j е асоциирана функцията полезност, чийто вид е:

$u^{**}(t)=E\{A_j.head(T)/\psi(F_j)\}+A_j.tail$ (виж (6.16)).

Тогава очакваната полезност е

$$Eu(\overline{A_j})=E\{A_j.head(T)/\psi(F_{j-1})\}+A_j.tail+$$

$$+\int_{\overline{A_j}}\{E\{[\overline{A_j}.head(T)-A_j.head(T)]/\psi(F_j)\}+$$

$$+\overline{A_j}.tail-A_j.tail\}\ f_j(x_j/\psi(F_{j-1}))dx_j\ .$$

Една горна граница на очакваната полезност е

$$Eu(\overline{A_j^0})=E\{A_j.head(T)/\psi(F_{j-1})\}+A_j.tail+$$

$$+\int_{\overline{A_j^0}}\{E\{[\overline{A_j}.head(T)-A_j.head(T)]/\psi(F_j)\}+$$

$$+\overline{A_j}.tail-A_j.tail\}\ f_j(x_j/\psi(F_{j-1}))dx_j\ ,$$

където

$$A_j^0=\{x_j:E\{[\overline{A_j}.head(T)-A_j.head(T)]/\psi(F_j)\}+\overline{A_j}.tail-A_j.tail\}\geq 0\ .$$

Да предположим, че функцията от лявата страна на горното неравенство сменя знака си от отрицателен към положителен точно един път.

Тогава уравнението за x_j^c има вида:

$$E\{[\overline{A_j}.head(T)-A_j.head(T)]/\psi(F_j)\}+\overline{A_j}.tail-A_j.tail=0\ .$$

2.1 Нека $\pi(A_j)=A_{j-1}$.

В променливите, асоциирани с възела $A_{j-1}=\pi(A_j)$ запазваме необходимата информация по следния начин:

$$A_{j-1}.head=A_j.head$$

и

$$A_{j-1}.tail=\int_{\overline{A_j^0}}\{E\{[\overline{A_j}.head(T)-A_j.head(T)]/\psi(F_j)\}+$$

$$+\overline{A_j}.tail-A_j.tail\}.f_j(x_j/\psi(F_{j-1}))dx_j+A_j.tail$$

(*виж Стъпка 6*).

Тогава получаваме следния резултат за горна граница на очакваната полезност:

$$Eu(\overline{A_j^0})=E\{A_{j-1}.head(T)/\psi(F_{j-1})\}+A_{j-1}.tail$$ (виж (6.16)).

Премахваме двойката листа $(A_j, \overline{A_j})$ (*виж Стъпка 7* на *Алгоритъм 6.1*) и с техния родител A_{j-1} свързваме получената горна граница на очакваната полезност:

$$Eu(\overline{A_j^0}) = E\{A_{j-1}.head(T) / \psi(F_{j-1})\} + A_{j-1}.tail.$$

2.2 Нека $\pi(A_j) = \overline{A_{j-1}}$

В променливите, асоциирани с възела $\overline{A_{j-1}} = \pi(A_j)$ запазваме необходимата информация по следния начин:

$$\overline{A_{j-1}}.head = A_j.head$$

и

$$\overline{A_{j-1}}.tail = \int_{\overline{A_j^0}} \{E\{[\overline{A_j}.head\,(T) - A_j.head\,(T)] / \psi(F_j)\} +$$

$$+ \overline{A_j}.tail - A_j.tail\}.f_j(x_j / \psi(F_{j-1}))dx_j + A_j.tail,$$

където

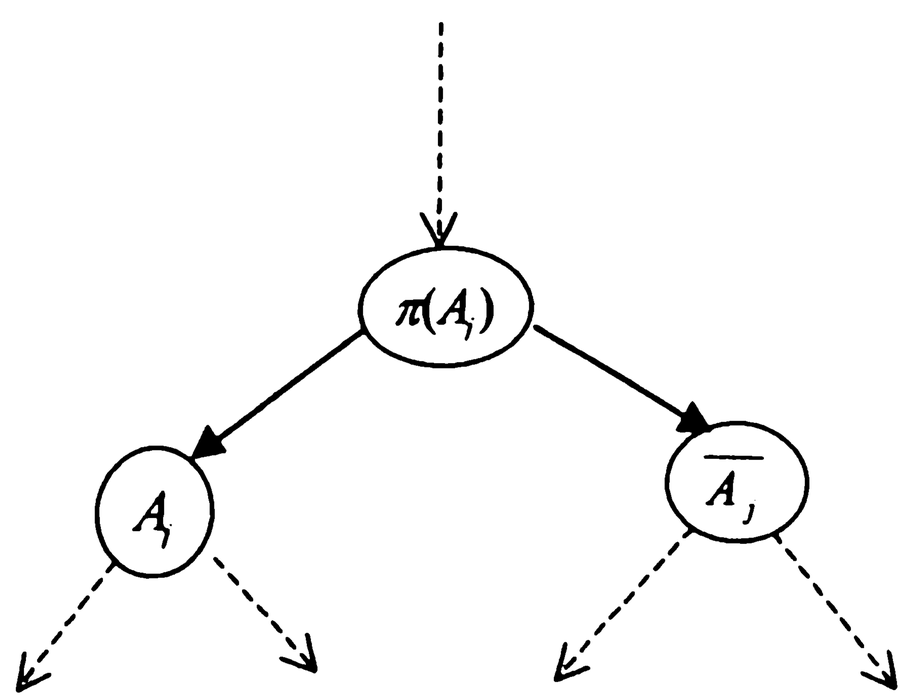
$$\overline{A_j^0} = \{x_j : E\{[\overline{A_j}.head(T) - A_j.head(T)] / \psi(F_j)\} + \overline{A_j}.tail - A_j.tail \geq 0\}$$

(*виж Стъпка 6*).

Тогава една горна граница на очакваната полезност е:

$$Eu(\overline{A_j^0}) = E\{\overline{A_{j-1}}.head(T) / \psi(F_{j-1})\} + \overline{A_{j-1}}.tail$$ (виж (6.15)).

Теоремата е доказана.



Фигура 6.10 *Симетрична U-единица в U-мрежа.*

6.3 Пример

Пример 6.4: Ще генерираме интегралните уравнения на задачата за вземане на решения, която може да се представи с **U**-мрежата от *фигура 6.11* и Бейсовата мрежа от *фигура 6.12*.

Инициализация:

$\boldsymbol{A_2.head}:=u_1$, $\boldsymbol{A_2.tail}:=NIL$,

$\overline{A_2}$ $\boldsymbol{.head}:=u_2$, $\overline{A_2}$ $\boldsymbol{.tail}:=NIL$,

$\boldsymbol{A_3.head}:=u_3$, $\boldsymbol{A_3.tail}:=NIL$,

$\boldsymbol{A_4.head}:=u_4$, $\boldsymbol{A_4.tail}:=NIL$,

$\overline{A_4}$ $\boldsymbol{.head}:=u_5$, $\overline{A_4}$ $\boldsymbol{.tail}:=NIL$.

Стъпки:

1. Намираме елемента $(A_4, \overline{A_4})$ с най-голям индекс от множеството от двойките листа с максимален номер на равнище в дървото на събитията.
2. Намираме $\psi(A_4) = \boldsymbol{X}_4, \boldsymbol{X}_3, \boldsymbol{X}_1$ и $\boldsymbol{i}=\varepsilon(A_4)=4$.
3. Намираме лявата страна на търсеното уравнение
   $\boldsymbol{left}:= \mathrm{E}\{[\boldsymbol{u}_5(\mathrm{T})- \boldsymbol{u}_4(\mathrm{T})]/ \boldsymbol{X}_4, \boldsymbol{X}_3, \boldsymbol{X}_1\}$
4. Генерираме уравнението
   $\mathrm{E}\{[\boldsymbol{u}_5(\mathrm{T})- \boldsymbol{u}_4(\mathrm{T})]/ \boldsymbol{X}_4, \boldsymbol{X}_3, \boldsymbol{X}_1\}=0$
5. Намираме $\overline{A_4}^0 = [x_4{}^c, 1]$ и родителя на двойката листа $(A_4, \overline{A_4})$
   $\overline{A_3} = \pi(A_4)$. $\Psi(\overline{A_3}) = X_3, X_1$
6. Присвояваме $\overline{A_3}$ $\boldsymbol{.head}:=\boldsymbol{u}_4$,
   $$\overline{A_3}\ \boldsymbol{.tail}:= \int_{\overline{A_i}^0} \{\mathrm{E}\{[\boldsymbol{u}_5(\mathrm{T})-\boldsymbol{u}_4(\mathrm{T})]/\boldsymbol{X}_4,\boldsymbol{X}_3,\boldsymbol{X}_1\}\}.f_4(\boldsymbol{x_4}/\boldsymbol{X}_3,\boldsymbol{X}_1)\mathrm{d}\boldsymbol{x_4}$$
7. Премахваме листата A_4 и $\overline{A_4}$.
   Следователно получихме уравнението
   $\mathrm{E}\{[\boldsymbol{u}_5(\mathrm{T})- \boldsymbol{u}_4(\mathrm{T})]/ \boldsymbol{X}_4, \boldsymbol{X}_3, \boldsymbol{X}_1\}=0$, което е еквивалентно на уравнението от (4.22) на §*4.4*.
8. Повтаряме горните стъпки, започвайки от *стъпка* **1**.

1. Намираме елемента $(A_3, \overline{A_3})$ с най-голям индекс от множеството от двойките листа с максимален номер на равнище в дървото на събитията.
2. Намираме $\psi(A_3) = X_3, X_1$ и $i = \varepsilon(A_4) = 3$.
3. Намираме лявата страна на търсеното уравнение

$$\textit{left} := E\{[u_4(T) - u_3(T)]/X_3, X_1\} + \int_{\overline{A_4}^0} \{E\{[u_5(T) - u_4(T)]/X_4, X_3, X_1\}\}.f_4(x_4/X_3, X_1)dx_4$$

4. Генерираме уравнението

$$E\{[u_4(T) - u_3(T)]/X_3, X_1\} + \int_{\overline{A_4}^0} \{E\{[u_5(T) - u_4(T)]/X_4, X_3, X_1\}\}.f_4(x_4/X_3, X_1)dx_4 = 0$$

5. Намираме $\overline{A_3}^0 = [x_3^c, 1]$ и родителя на двойката листа $(A_3, \overline{A_3})$ $\overline{A_1} = \pi(A_3)$. $\Psi(\overline{A_1}) = X_1$
6. Присвояваме $\overline{A_1}.\textit{head} := u_3$,

$$\overline{A_1}.\textit{tail} := \int_{\overline{A_3}^0} \{E\{[u_4(T) - u_3(T)]/X_3, X_1\} + \int_{\overline{A_4}^0} \{E\{[u_5(T) - u_4(T)]/X_4, X_3, X_1\}\}.f_4(x_4/X_3, X_1)dx_4\} f_3(x_3/X_1)dx_3$$

7. Премахваме листата A_3 и $\overline{A_3}$.

Следователно получихме уравнението

$$E\{[u_4(T) - u_3(T)]/X_3, X_1\} + \int_{\overline{A_4}^0} \{E\{[u_5(T) - u_4(T)]/X_4, X_3, X_1\}\}.f_4(x_4/X_3, X_1)dx_4 = 0,$$

което е еквивалентно на уравнението от (4.23) на §4.4.

8. Повтаряме горните стъпки, започвайки от *стъпка* **1**.

1. Намираме елемента $(A_2, \overline{A_2})$ с най-голям индекс от множеството от двойките листа с максимален номер на равнище в дървото на събитията.
2. Намираме $\psi(A_2) = X_2, X_1$ и $i = \varepsilon(A_4) = 2$.
3. Намираме лявата страна на търсеното уравнение
   $\textit{left} := E\{[u_2(T) - u_1(T)]/X_2, X_1\}$
4. Генерираме уравнението
   $E\{[u_2(T) - u_1(T)]/X_2, X_1\} = 0$

5. Намираме родителя на двойката листа $(A_2, \overline{A_2})$

$\overline{A_1} = \pi(A_2).\ \Psi(\overline{A_1}) = X_1$

6. Присвояваме $\overline{A_2}^0 = [x_2^c, 1]$, $A_1.\boldsymbol{head} := \boldsymbol{u}_1$,

$$A_1.\boldsymbol{tail} := \int_{\overline{A_2}^0} \{ E\{[\boldsymbol{u}_2(T) - \boldsymbol{u}_1(T)]/X_2, X_1\}\}.f_2(x_2/X_1)dx_2$$

7. Премахваме листата A_2 и $\overline{A_2}$.

Следователно получихме уравнението $E\{[\boldsymbol{u}_2(T) - \boldsymbol{u}_1(T)]/X_2, X_1\} = 0$, което е еквивалентно на уравнението от (4.24) на §4.4.

8. Повтаряме горните стъпки, като започваме от *стъпка 1.*

1. Намираме елемента $(A_1, \overline{A_1})$ с най-голям индекс от множеството от двойките листа с максимален номер на равнище в дървото на събитията.

2. Намираме $\psi(A_1) = X_1$ и $i = \varepsilon(A_4) = 1$.

3. Намираме лявата страна на търсеното уравнение

$$\boldsymbol{left} := E\{[\boldsymbol{u}_3(T) - \boldsymbol{u}_1(T)]/X_1\} + \int_{\overline{A_3}^0} \{E\{[\boldsymbol{u}_4(T) - \boldsymbol{u}_3(T)]/X_3, X_1\}$$

$$+ \int_{\overline{A_4}^0} \{E\{[\boldsymbol{u}_5(T) - \boldsymbol{u}_4(T)]/X_4, X_3, X_1\}\}.f_4(x_4/X_3, X_1)dx_4\} f_3(X_3/X_1)dx_3 -$$

$$\int_{\overline{A_2}^0} \{ E\{[\boldsymbol{u}_2(T) - \boldsymbol{u}_1(T)]/X_2, X_1\}\}.f_2(x_2/X_1)dx_2$$

4. Генерираме уравнението

$$E\{[\boldsymbol{u}_3(T) - \boldsymbol{u}_1(T)]/X_1\} + \int_{\overline{A_3}^0} \{E\{[\boldsymbol{u}_4(T) - \boldsymbol{u}_3(T)]/X_3, X_1\}$$

$$+ \int_{\overline{A_4}^0} \{E\{[\boldsymbol{u}_5(T) - \boldsymbol{u}_4(T)]/X_4, X_3, X_1\}\}.f_4(x_4/X_3, X_1)dx_4\} f_3(X_3/X_1)dx_3 -$$

$$\int_{\overline{A_2}^0} \{ E\{[\boldsymbol{u}_2(T) - \boldsymbol{u}_1(T)]/X_2, X_1\}\}.f_2(x_2/X_1)dx_2 = 0$$

5. Намираме родителя на двойката листа $(A_1, \overline{A_1})$ $\Omega = \pi(A_2)$.

9. Край.

Следователно получихме уравнението

$$E\{[\boldsymbol{u}_3(T)-\boldsymbol{u}_1(T)]/\boldsymbol{X}_1\} + \int\limits_{\bar{A}_3^0} \{E\{[\boldsymbol{u}_4(T)-\boldsymbol{u}_3(T)]/\boldsymbol{X}_3, \boldsymbol{X}_1\}$$

$$+ \int\limits_{\bar{A}_4^0} \{E\{[\boldsymbol{u}_5(T)-\boldsymbol{u}_4(T)]/\boldsymbol{X}_4,\boldsymbol{X}_3,\boldsymbol{X}_1\}\}.f_4(\boldsymbol{x}_4/\boldsymbol{X}_3,\boldsymbol{X}_1)d\boldsymbol{x}_4\}f_3(\boldsymbol{X}_3/\boldsymbol{X}_1)d\boldsymbol{x}_3$$

$$- \int\limits_{\bar{A}_2^0} \{ E\{[\boldsymbol{u}_2(T)-\boldsymbol{u}_1(T)]/\boldsymbol{X}_2, \boldsymbol{X}_1\}\}.f_2(\boldsymbol{x}_2/\boldsymbol{X}_1)d\boldsymbol{x}_2 = 0,$$

което е еквивалентно на уравнението от (4.25) на §4.4.

6.4 Алгоритъм за генериране на достатъчно условие за монотонност

Ще разгледаме алгоритъм за генериране на достатъчно условие за монотонност в символен вид.

Алгоритъм 6.2: *Алгоритъм за генериране на достатъчно условие за монотонност:*

Вход:

- U-мрежа с инициализирано дърво на събитията
- Бейсова мрежа

Изход: условия на достатъчното условие за монотонност

Инициализация: С всяко листо A_i на дървото на събитията асоциираме две символни променливи

A_i.***head***:=$U(A_i)$

A_i.***tail***:=NIL

Стъпки:

1. Намираме списъкът от случайните величини, участващи в модела, т.е. $\psi_k := x_1, x_2, \ldots, x_n$. Тогава генерираме първото условие за монотонност:
   - ***Условната плътност*** $f(t/\psi_k)$ ***да е стохастически растяща.***
2. Намираме елемента $(A_i, \overline{A_i})$ с най-голям индекс ***i*** от множеството от двойките листа с максимален номер на равнище в дървото на събитията.
3. Намираме $\psi(A_i)$ - списъкът от случайните величини, асоциирани с листото A_i, ***i***$=\varepsilon(A_i)$ и $\psi_k := \psi(A_i) \setminus x_i$.
4. Ако A_i.***tail*** $=\phi$ то изпълняваме *стъпка* **5**, в противен случай преминаваме към *стъпка* **9**.
5. Намираме функцията ***left***:= $\overline{A_i}$.***head*** (t) - A_i .***head*** (t) и генерираме следващото условие за монотонност:
   - ***Функцията left да е растяща за всяко t.***
6. Ако $\overline{A_i}$.***tail*** $\neq \phi$ то преминаваме към *стъпка* **7**, в противен случай изпълняваме *стъпка* **8**.
7. Генерираме следващото условие за монотонност:
   - ***Условната плътност*** $\overline{A_i}$ ***.density да е стохастически растяща.***

8. Намираме функцията

$$\textbf{\textit{left}}:=\underline{E}\{[\overline{A_i}.\textbf{\textit{head}}\,\underline{(T)}-A_i.\textbf{\textit{head}}\,\underline{(T)}]/\psi(A_i)\}+\overline{A_i}.\textbf{\textit{tail}}$$

Преминаваме към *стъпка* **10**.

9. Намираме функцията

$$\textbf{\textit{left}}:=\underline{E}\{[\overline{A_i}.\textbf{\textit{head}}\,\underline{(T)}-A_i.\textbf{\textit{head}}\,\underline{(T)}]/\psi(A_i)\}+\overline{A_i}.\textbf{\textit{tail}}-A_i.\textbf{\textit{tail}}$$

и генерираме следващото условие за монотонност:

- ***Функцията left да е растяща за всяко x_i.***

10. Намираме родителя на двойката листа $(A_i, \overline{A_i})$ $A_k=\pi(A_i)$. Ако това е коренът на дървото, преминаваме към *стъпка* **14**. В противен случай обозначаваме $\overline{A_i^0}=[x_i^c;1]$ и изпълняваме *стъпка* **11**.

11. Ако $(A_i, \overline{A_i})$ е десен наследник на $\pi(A_i)$, т.е. $\pi(A_i)=\overline{A_k}$, присвояваме

$$\overline{A_k}.\textbf{\textit{head}}:=A_i.\textbf{\textit{head}},$$

$$\overline{A_k}.\textbf{\textit{tail}}:=\int_{\overline{A_i}^0}\{\textbf{\textit{left}}\}.f_i(x_i/\psi_k)\,\underline{d}\,x_i+A_i.\textbf{\textit{tail}},$$

$$\overline{A_k}.\textbf{\textit{density}}:=f_i(x_i/\psi_k)$$

В противен случай, т.е. ако $(A_i, \overline{A_i})$ е ляв наследник на $\pi(A_i)$, т.е. $\pi(A_i)=A_k$, присвояваме

$$A_k.\textbf{\textit{head}}:=A_i.\textbf{\textit{head}},$$

$$A_k.\textbf{\textit{tail}}:=\int_{\overline{A_i}^0}\{\textbf{\textit{left}}\}.f_i(x_i/\psi_k)\,\underline{d}\,x_i+A_i.\textbf{\textit{tail}}$$

12. Премахваме листата A_i и $\overline{A_i}$.

Получихме част от условията за монотонност и нова **U**-мрежа, чието дърво е под-дърво на дървото на събитията.

13. Повтаряме горните стъпки, като започнем от *стъпка* **2**.

14. Край.

Разглеждаме **U**-мрежа с инициализирано дърво на събитията. Нека $A_{j-1}=\pi(A_j)$ и $\overline{A_{j-1}}=\pi(A_i)$. Оценяваме очакваната полезност постъпково и на всяка стъпка намираме част от условията на достатъчното условие за монотонност. По аналогичен начин както *Теорема 6.1* може да се докаже следната теорема:

Теорема 6.2: Нека на всяка стъпка $j=1,2,\dots,n$ функцията

$$E\{[\overline{A_{j-1}}.head(T)-A_{j-1}.head(T)]/\psi(A_{j-1})\}+\overline{A_{j-1}}.tail-A_{j-1}.tail$$

да е растяща за всяко t,

където

$$A_{j-1}.head = A_j.head,$$

$$A_{j-1}.tail = \int_{\overline{A_j^0}}\{E\{[\overline{A_j}.head(T)-A_j.head(T)]/\psi(A_j)\}+$$

$$+\overline{A_j}.tail-A_j.tail\}f_j(x_j/\psi(A_{j-1}))dx_j+A_j.tail,$$

$$\overline{A_j^0}=\{x_j : E\{[\overline{A_j}.head(T)-A_j.head(T)]/\psi(A_j)\}+\overline{A_j}.tail-A_j.tail\geq 0\};$$

$$\overline{A_{j-1}}.head = A_i.head,$$

$$\overline{A_{j-1}}.tail = \int_{\overline{A_i^0}}\{E\{[\overline{A_i}.head(T)-A_i.head(T)]/\psi(A_i)\}+$$

$$+\overline{A_i}.tail-A_i.tail\}.f_i(x_i/\psi(A_{j-1}))dx_i+A_j.tail,$$

$$\overline{A_i^0}=\{x_i : E\{[\overline{A_i}.head(T)-A_i.head(T)]/\psi(A_i)\}+\overline{A_i}.tail-A_i.tail\geq 0\}.$$

Тогава правилото за определяне състоянието на вярата на агента е монотонно и *Алгоритъм 6.2* правилно ще генерира вида на достатъчното условие за монотонност.

Като се използват *Лема 2*, *Лема 3* и *Лема 4* от параграф *4.1.4* може да се докаже, че е в сила следната теорема:

Теорема 6.3: Нека на стъпката j ($j=1,..n$) са изпълнени условията:

- $\overline{A_{j-1}}.head(t)-A_{j-1}.head(t)$ да е растяща за всяко t, където.

$$\overline{A_{j-1}}.head = A_i.head,\quad A_{j-1}.head = A_j.head.$$

- условната плътност $f(t/\psi(A_j))$ да е стохастически растяща,
- условната плътност $f_{j+1}(x_{j+1}/\psi(A_j))$ да е стохастически растяща.

Тогава функцията

$$E\{[\overline{A_{j-1}}.head(T)-A_{j-1}.head(T)]/\psi(A_{j-1})\}+\overline{A_{j-1}}.tail$$

е растяща за всяко t.

Алгоритъм 6.2 с успех може да бъде използван за генериране на нови твърдения за монотонност на правилата за вземане на решения, които твърдения в последствие да бъдат доказани. Той не може да замести доказателствата на условията за монотонност.

Да отбележим, че едно по-общо условие на условието

- За всяко $j=1,\dots,n$ условната плътност $f(t/\psi(A_j))$ да е стохастически растяща,

е условието:

- Условната плътност $f(t/x_1,\dots,x_n)$ да е стохастически растяща.

В сила е и следната теорема:

Теорема 6.4: Нека за всяко $j=1\dots,n$ функцията

$$E\{[\overline{A_{j-1}}.head(T)-A_{j-1}.head(T)]/\psi(A_{j-1})\}+\overline{A_{j-1}}.tail-A_{j-1}.tail$$

да сменя знака си от отрицателен към положителен точно един път, където

$$A_{j-1}.head = A_j.head,$$

$$A_{j-1}.tail = \int_{\overline{A_j^0}}\{E\{[\overline{A_j}.head(T)-A_j.head(T)]/\psi(A_j)\}+$$

$$+\overline{A_j}.tail-A_j.tail\}f_j(x_j/\psi(A_{j-1}))dx_j+A_j.tail,$$

$$\overline{A_j^0}=\{x_j : E\{[\overline{A_j}.head(T)-A_j.head(T)]/\psi(A_j)\}+\overline{A_j}.tail-A_j.tail\geq 0\};$$

$$\overline{A_{j-1}}.head = A_i.head,$$

$$\overline{A_{j-1}}.tail = \int_{\overline{A_i^0}}\{E\{[\overline{A_i}.head(T)-A_i.head(T)]/\psi(A_i)\}+$$

$$+\overline{A_i}.tail-A_i.tail\}.f_i(x_i/\psi(A_{j-1}))dx_i+A_j.tail,$$

$$\overline{A_i^0}=\{x_i : E\{[\overline{A_i}.head(T)-A_i.head(T)]/\psi(A_i)\}+\overline{A_i}.tail-A_i.tail\geq 0\}.$$

Тогава правилото за определяне състоянието на вярата на агента е монотонно.

По аналогичен начин може да се построи алгоритъм, който генерира условията от *Теорема 6.4* в символен вид. Верността на получените твърдения трябва да бъде доказана.

Пример 6.2: Ще приложим *Алгоритъм 6.2* към проблема, описан в параграф 4.4:

Вход:

- **U**-мрежата от *фигура 6.11* с инициализирано дърво на събитията,
- Бейсовата мрежа от *фигура 6.12.*

Изход: условия на достатъчното условие за монотонност

Инициализация:

A_2.***head***:=u_1, $\overline{A}_2$.***head***:=u_2, A_3.***head***:=u_3, A_4.***head***:=u_4, $\overline{A}_4$.***head***:=u_5,

A_2.***tail***:=*NIL*, $\overline{A}_2$.***tail***:=*NIL*, A_3.***tail***:=*NIL*, A_4.***tail***:=*NIL*, A_5.***tail***:=*NIL*.

Стъпки:

1. Намираме списъка от случайните величини, участващи в модела, т.е. $\psi_k := x_1, x_2, x_3, x_4$. Тогава първото условие е:
   - Условната плътност $f(t/x_1, x_2, x_3, x_4)$ да е стохастически растяща.
2. Намираме елемента $(A_4, \overline{A_4})$.
3. Намираме списъка от случайните величини $\psi(A_4) = x_4, x_3, x_1$, ***i***=4 и $\psi_k := x_3, x_1$.
4. A_4.***tail*** $=\phi$.
5. ***left***:= $u_5(t) - u_4(t)$ и следващото условие има вида:
   - Функцията $u_5(t) - u_4(t)$ да е растяща за всяко t,
6. $\overline{A_i}$.***tail*** $=\phi$.

8. ***left***:=$E\{[u_5(T) - u_4(T)]/ x_4, x_3, x_1\}$

10. Намираме родителя на $(A_4, \overline{A_4})$ $\overline{A}_3 = \pi(A_4)$. Обозначаваме $\overline{A_4^0} = [x_4^c; 1]$ и изпълняваме *стъпка* 11.

11. $\pi(A_4) = \overline{A_3}$. Следователно присвояваме

$\overline{A_3}$.***head***:= u_4,

$$\overline{A_3}\ .\mathbf{tail} := \int_{\overline{A_4}^0} \{E\{[u_5(T) - u_4(T)]/ x_4, x_3, x_1\}.f_4(x_4/x_3, x_1)\,d\,x_4$$

$\overline{A_3}$.***density*** := $f_4(x_4/x_3, x_1)$

12. Премахваме листата A_4 и $\overline{A_4}$.
Получихме част от условията за монотонност и нова U-мрежа, чието дърво е под-дърво на дървото на събитията.
13. Повтаряме горните стъпки, като започнем от *стъпка* **2**.

2. Намираме елемента $(A_4, \overline{A_4})$.

3. Намираме списъка от случайните величини $\psi(A_3) = x_3, x_1$, ***i***=3 и $\psi_k := x_1$.

4. A_3.***tail*** $=\phi$.

5. ***left***:= u_4(t) - u_3 (t) и следващото условие има вида:
   - Функцията u_4(t) - u_3 (t) да е растяща за всяко t,

6. $\overline{A_i}$.***tail*** $=\phi$.

8. ***left***:=E{[u_4(T) - u_3 (T)]/ x_3, x_1 }

10. Намираме родителя на $(A_3, \overline{A_3})$ $\overline{A_1}$ =π(A_3). Обозначаваме $\overline{A_4^0} = [x_4^c;1]$ и изпълняваме *стъпка* 11.

11. π(A_4)= $\overline{A_3}$. Следователно присвояваме

$\overline{A_1}$.***head***:= u_4,

$$\overline{A_1}\ .\mathbf{tail} := \int_{\overline{A_3}^0} \{ E\{[u_4(T) - u_3 (T)]/ x_3, x_1 \} + \int_{\overline{A_4}^0} \{ E\{[u_5(T) - u_4 (T)]/ x_4, x_3, x_1 \}.f_4(x_4/x_3,x_1)d\, x_4\} f_3(x_3/x_1)d\, x_3$$

12. Премахваме листата A_3 и $\overline{A_3}$.
Получихме част от условията за монотонност и нова U-мрежа, чието дърво е под-дърво на дървото на събитията.
13. Повтаряме горните стъпки, като започнем от *стъпка* **2**.

2. Намираме елемента $(A_2, \overline{A_2})$.

3. Намираме списъка от случайните величини $\psi(A_2) = x_2, x_1$, ***i***=2 и $\psi_k := x_1$.

4. A_2.***tail*** $=\phi$.

5. ***left***:= u_2(t) - u_1 (t) и следващото условие има вида:
   - Функцията u_2(t) - u_1 (t) да е растяща за всяко t,

6. $\overline{A_i}$.***tail*** $=\phi$.

8. ***left***:=E{[u_2(T) - u_1 (T)]/ x_2, x_1 }

10. Намираме $A_1=\pi(A_2)$. Обозначаваме $\overline{A_4^0}=[x_4^c;1]$ и изпълняваме *стъпка* 11.

11. Присвояваме $A_k.\mathbf{head}:=A_i.\mathbf{head}$,

$$A_k.\mathbf{tail}:=\int_{\overline{A_i}^0}\{\mathbf{left}\}.f_i(x_i/\psi_k)\,\mathrm{d}\,x_i \pm A_i.\mathbf{tail}$$

12. Премахваме листата A_2 и $\overline{A_2}$. Получихме част от условията за монотонност и нова U-мрежа, чието дърво е под-дърво на дървото на събитията.

13. Повтаряме горните стъпки, като започнем от *стъпка* **2**.

2. Намираме елемента $(A_1,\overline{A_1})$.

3. Намираме списъка от случайните величини $\psi(A_1)=x_1$, i=1 и $\psi_k:=\phi$.

4. $A_i.\mathbf{tail}\neq\phi$

9. Следващото условие има вида: Функцията

$$E\{[u_3(T)-u_1(T)]/x_1\}+\int_{\overline{A_3}^0}\{E\{[u_4(T)-u_3(T)]/x_3,x_1\}+\int_{\overline{A_4}^0}\{E\{[u_5(T)-u_4(T)]/x_4,x_3,x_1\}.f_4(x_4/x_3,x_1)\,\mathrm{d}\,x_4\}f_3(x_3/x_1)\,\mathrm{d}\,x_3-\int_{\overline{A_2}^0}\{E\{[u_2(T)-u_1(T)]/x_2,x_1\}.f_2(x_2/x_1)\,\mathrm{d}\,x_2$$

да е растяща.

10. Намираме $\pi(A_2)=\Omega$ и изпълняваме *стъпка* 14.

14. Край.

Следователно получихме условията от *Теорема 4.6.*

6.5 Извод

Проблемът за определяне на вярата на агента във вероятностния модел може да се представи чрез наредената тройка $(\mathcal{B}, \mathcal{U}, \mathcal{D})$, където $\mathcal{B}$ е Бейсова мрежа, представяща вероятностната структура на проблема, $\mathcal{U}$ е U-мрежа, представяща структурата полезност, $\mathcal{D}$ е мрежа за вземане на решение.

Чрез автоматични изводи над $\mathcal{D}$ може да се вземе решение за състоянието на вярата.

Чрез автоматични изводи над $\mathcal{U}$ може да се оптимизира вярата.

Чрез автоматични изводи над $\mathcal{B}$ и $\mathcal{U}$ може да се получат достатъчните условия за слаба монотонност на правилата за определяне състоянието на вярата.

Този подход може да се приложи за моделиране на вярата въобще.

Глава 7

Алтернативни симулационни подходи

Разглеждаме отново реални случайни величини, имащи съвместно вероятностно разпределение, чиято плътност е строго положителна върху извадковото пространство. Предполагаме съществуването на плътности, свързани или с Лебеговата или с изброима мярка за всички маргинални и условни разпределения.

7.1 Алгоритъм с полагане

Идеята на този алгоритъм е да се използва полагане при намиране на точкови решения на клас интегрални уравнения.

Нека случайната величина X има плътност $r(x)$, а случайната величина Y има плътност $k(y)$, които са неизвестни. Известни са условните плътности $g(x/y)$ и $q(y/x)$. Тогава

$$r(x)=\int g(x/y)k(y)dy \tag{7.1}$$

$$k(y)=\int q(y/x)r(x)dx \tag{7.2}$$

Полагаме (7.2) в (7.1) и получаваме:

$$r(x)=\int g(x/y)\int q(y/x')r(x')dx'dy=$$

$$=\int h(x,x')r(x')dx', \tag{7.3}$$

където

$$h(x,x')=\int g(x/y)\, q(y/x')dy.$$

Нека $r_i(x)$ е оценка на $r(x)$, получена на i-тата стъпка на един итеративен процес.

Тогава от (7.3) следва, че

$$r_{i+1}(x)=\int h(x,x')r_i(x')dx'=\mathbf{I}_h\, r_i(x),$$

където $\mathbf{I}_h$ е интегрален оператор, свързан с функцията h.

Tanner и Wong са показали, че за този итеративен процес са в сила следните твърдения [*Tanner, Wong, 87*]:

T1: *Истинската маргинална плътност $r(x)$ е единственото решение на интегралното уравнение* (7.3)

$$r(x)=\int h(x,x')r(x')dx'.$$

T2: *За почти всички $r_0(x)$ редицата $r_1(x)$, $r_2(x)$, $r_3(x)$, ..., определена с равенството $r_{i+1}(x)=\mathbf{I}_h\, r_i(x)$, $i=0,1,2,\ldots$, клони монотонно към $r(x)$ в $\mathbf{L}_1$.*

T3: $\int |r_i(x)-r(x)|dx \to 0$ при $i\to\infty$.

Ще представим алгоритъма.

Нека $g(x/y)$ и $q(y/x)$ са известни. Нека $r_0(x)$ е произволна оценка на плътността. Нека случайната величина X^0 има плътност $r_0(x)$, което ще означаваме $X^0 \in r_0(x)$.

Нека случайната величина Y^1 има плътност $q(y/x^0)$, т.е. $Y^1 \in q(y/x^0)$. Тогава от (7.2) следва, че маргиналното разпределение на Y^1 е $k_1(y)=\int q(y/x)r_0(x)dx$.

Нека случайната величина X^1 има плътност $g(x/y')$. От (7.1) и (7.3) следва, че

$$X^1 \in r_1(x)=\int g(x/y)k_1(y)dy=\int h(x,x')r_0(x')dx'=\mathbf{I}_h\, r_0(x).$$

Повтаряйки горните стъпки след i-тата итерация получаваме случайния вектор (X^i,Y^i), за който е в сила твърдение **T2**, т.е.

$X^i \xrightarrow{d} X \in r(x)$ и $Y^i \xrightarrow{d} Y \in k(y)$.

Повтаряме този итерационен процес m пъти и на i-тата итерация получаваме m случайни вектора $(X^i_j,Y^i_j), j=1,2,\dots,m$.

Тази схема на генериране на случайни величини се нарича *алгоритъм с полагане* (*substitution algorithm*).

Ако прекъснем всеки от тези итерационни процеси на i-тата итерация, по метода на Монте-Карло получаваме следната оценка на плътността $r(x)$:

$$\overline{r_i(x)} = \frac{1}{m}\sum_{j=1}^{m} g(x/y_j^i) \tag{7.4}$$

Да отбележим, че случайните величини X_j^i не се използват в (7.4).

Glick (1974) е показал, че $\overline{r_i(x)} \xrightarrow{P} r_i(x)$ при $m \longrightarrow \infty$ и $i \longrightarrow \infty$.

Ще разгледаме този алгоритъм за три случайни величини X, Y и Z. Тогава

$$r(x)=\iint f_1(x,z/y)k(y)dydz \tag{7.5}$$

$$k(y)=\iint f_2(y,x/z)l(z)dzdx \tag{7.6}$$

$$l(z)=\iint f_3(z,y/x)r(x)dxdy \tag{7.7}$$

Полагайки (7.7) в (7.6) и след това (7.6) в (7.5) получаваме уравнение, подобно на уравнение (7.3) и нова функция h, асоциирана с интегралния оператор $\mathbf{I}_h$.

Започваме с произволна начална маргинална плътност $r_0(x)$. Последователно генерираме следващите случайни величини:

$X^0 \in r_0(x)$,

$(Z^{0\prime}, Y^{0\prime}) \in f_3(z,y/x^0)$,

$(Y^1, X^{0\prime}) \in f_2(y,x/z^{0\prime})$,

$(X^1, Z^1) \in f_1(x,z/y^1)$.

Пълният цикъл за генериране на X^1 от X^0 изисква да се генерират шест случайни величини. Повтаряйки този цикъл i пъти получаваме (X^i, Y^i, Z^i). В сила са следните твърдения за сходимост по разпределение:

$$X^i \xrightarrow{d} X \in r(x),\ Y^i \xrightarrow{d} Y \in k(y),\ Z^i \xrightarrow{d} Z^i \in l(z).$$

Ако повторим този итерационен процес m пъти ще получим (X_j^i, Y_j^i, Z_j^i), $j=1,2,\dots m$, които са независими при различни стойности на j и не са независими при фиксирано j.

Тогава оценката на маргиналната плътност $r(x)$ е

$$\overline{r_i(x)} = \frac{1}{m}\sum_{j=1}^{m} g(x / y_j^i, z_j^i).$$

Ще направим обобщение на този алгоритъм за k случайни величини $U_1, U_2, \dots, U_k$.

Ако стартираме генерирането на m редици по описания по-горе начин на i-тата итерация получаваме k-мерния случаен вектор $(U_{1j}^i, U_{2j}^i, \quad, U_{kj}^i)$, $j=1,2,\dots,m$. За оценка на плътността $r(u_s)$ на случайната величина U_s получаваме:

$$\overline{r_i(u_s)} = \frac{1}{m}\sum_{j=1}^{m} g(u_s / u_{tj}^i, t \neq s).$$

Да отбележим, че прилагането на този алгоритъм не изисква съвместното вероятностно разпределение да е известно.

7.2 Алгоритъм на Гибс

Нека

$r(x) = \iint p(x/z,y).q(z/y).k(y)dydz$

$k(y) = \iint m(y/z,x).n(x/z).l(z)dxdz$

$l(z) = \iint a(z/y,x).b(y/x).r(x)dxdy$

За да приложим *алгоритъма с полагане* (*substitution algorithm*) е необходимо да са известни всичките шест условни разпределения, което рядко се случва в действителност.

Условните разпределения $p(x/z,y)$, $m(y/z,x)$ и $a(z/y,x)$ напълно определят съвместното вероятностно разпределение на случайните величини X, Y и Z [*Besag,74*].

Един алгоритъм за извличане на маргиналното разпределение от тези условни разпределения е формално въведен от Geman и Geman в [*Geman, Geman, 84*] и е известен като *алгоритъм на Гибс* (*Gibbs' sampling algorithm*).

Разглеждаме множеството от случайни величини $U_1, U_2, \ldots, U_k$. Предполагаме, че са известни условните разпределения

$f(u_i/u_j, j\neq i)= f(u_i/u_j, j\in S_i)$, $i=1,2,\ldots,k$, където S_i е множеството от съседите на X_i.

Алгоритъм на Гибс е Марковска схема за обновяване, която работи по следния начин:

Нека е дадено произволно начално множество от стойности $u_1^0, u_2^0, \ldots, u_k^0$.

Нека случайната величина U_1^1 има вероятностно разпределение $f_1(u_1/u_2^0, u_3^0, \ldots, u_k^0)$, т.е.

$$U_1^1 \in f_1(u_1/u_2^0, u_3^0, u_4^0, \ldots, u_k^0).$$

Нека $U_2^1 \in f_2(u_2/u_1^1, u_3^0, u_4^0, \ldots, u_k^0)$.

Нека $U_3^1 \in f_3(u_3/u_1^1, u_2^1, u_4^0, \ldots, u_k^0)$.

. . . . . .

Нека $U_k^1 \in f_k(u_k/u_1^1, u_2^1, u_3^1, \ldots, u_{k-1}^1)$.

Така направихме преход от $u^0 = (u_1^0, u_2^0, \ldots, u_k^0)$ към $u^1 = (u_1^1, u_2^1, \ldots, u_k^1)$. След i такива итерации ще получим $(U_1^i, U_2^i, \ldots, U_k^i)$, т.е. Повтаряйки този процес получаваме редицата $u^0, u^1, \ldots, u^i, \ldots$, която е реализачия на Марковска верига с вероятност на прехода от u^i към u^{i+1}, която се дава с равенството [*Smith, Roberts, 93*]:

$$K_G(u^i, u^{i+1}) = \prod_{l=1}^{k} f_l(x_l^{i+1} / x_j^{i+1}, j < l, x_j^i, j > l).$$

Geman и Geman в [*Geman, Geman, 84*] показаха, че са в сила следните резултати:

T4: $(U_1^i, U_2^i, \ldots, U_k^i) \xrightarrow{d} (U_1, U_2, \ldots, U_k)$ при $i \longrightarrow \infty$.

T5: За всяка измерима функция T на $U_1, U_2, \ldots, U_k$, за която съществува $ET(U_1, U_2, \ldots, U_k)$, е в сила:

$$\lim_{i\to\infty} \frac{1}{i}\sum T(u_i^l, \ldots, u_k^l) \xrightarrow{n.c.} ET(u_1, \ldots, u_k).$$

Очевидно в двумерния случай двата алгоритъма са еднакви. В k-мерния случай *алгоритъма на Гибс* предполага, че са известни k условни разпределения, а *алгоритъма с полагане* предполага, че са известни $k.(k-1)$ условни разпределения.

7.3 Неитеративен симулационен алгоритъм

Rubin (1987) предложи неитеративен Монте-Карло метод за генериране на маргинални разпределения. Идеята е следната:

Предполагаме, че търсим маргиналното разпределение на случайната величина X при следните условия:

- Известен е функционалният вид на съвместното вероятностно разпределение $f(x,y)$.
- Известно е условното разпределение $g(x/y)$.
- Не е известно маргиналното разпределение на случайната величина Y.

Избираме плътност $k_s(y)$ на случайната величина Y, получена с някакъв *sampling* алгоритъм. Тогава $g(x/y).k_s(y)$ е *sampling* разпределение на (X,Y).

Нека $(X_l,Y_l) \in g(x/y).k_s(y)$, $l=1,2,\dots,N$. Например, $Y_l \in k_s(y)$, $X_l \in g(x/y_l)$.

Идеята на Rubin е да се пресметне

$r_l = \dfrac{f(x_l,y_l)}{g(x_l/y_l)}.k_s(y_l)$, при $l=1,2,\dots,N$, и да се оцени маргиналната плътност $r(x)$ по следния начин

$$\overline{r(x)} = \frac{\sum_{l=1}^{N} g(x/y_l).r_l}{\sum_{l=1}^{N} r_l}.$$

Тогава е в сила следното твърдение:

T6: $\overline{r(x)} \xrightarrow[N\to\infty]{} r(x)$ с вероятност 1 за почти всички X.

Ако условната плътност $q(y/x)$ е известна, веднага можем да намерим една оценка на маргиналното разпределение на случайната величина Y:

$$\overline{k(y)} = \frac{\sum_{l=1}^{N} q(y/x_l).r_l}{\sum_{l=1}^{N} r_l}.$$

Tanner и Wong в [*Tanner, Wong, 87*] предлагат за $k_s(y)$ да се използва оценката, получена след i итерации на sampling алгоритъма.

Ще обобщим идеята на този алгоритъм, известен като *importance sampling algorithm*, за $k=3$ случайни величини.

Предполагаме, че търсим маргиналното разпределение на случайната величина X при следните условия:

- Известен е функционалният вид на съвместното вероятностно разпределение $f(x,y,z)$.

•Известно е условното разпределение $g(x/y,z)$.

Не е известно маргиналното разпределение на случайния вектор (Y,Z) и трябва да определим двумерно разпределение, получено с някакъв *sampling* алгоритъм. Да отбележим, че ако $p(y/z)$ е известно ние трябва да определим само $q_s(z)$.

Получаваме $(X_l, Y_l, Z_l) \in f(x,y,z)$, $l=1,2,\dots,N$, и изчисляваме

$r_l = \dfrac{f(x_l, y_l, z_l)}{g(x_l / y_l, z_l)} . k_s(y_l, z_l)$. Тогава оценката на маргиналната плътност на случайната величина X е

$$\overline{r(x)} = \frac{\sum_{l=1}^{N} g(x / y_l, z_i).r_l}{\sum_{l=1}^{N} r_l}.$$

Да отбележим, че в k -мерния случай този алгоритъм изисква $N.k$ симулации, докато *алгоритъма на Гибс*, спрян след i-тата итерация, изисква $m.i.k$ симулации. На базата на емпирични резултати, представени в [*Gelfand, Smith, 90*] авторите твърдят, че е за предпочитане е да се използват итерационни алгоритми със малки извадки, отколкото една огромна извадка. Итеративните методи (substitution, Gibbs) дават по-добри резултати.

7.4 Приложения

Нека m случайни (зависими) извадки (x_i, y_i), $i=1,2,\dots,m$, са направени от едно съвместно вероятностно разпределение чрез алгоритъма *Gibbs sampler*. Тези извадки могат да се използват за да се пресметнат характеристиките (средното, дисперсията) на разпределението $P(x)$ чрез метода Монте-Карло.

Интегралът $u=\int g(x)dP(x)$ може да се апроксимира с

$\overline{u} = \dfrac{1}{m}\sum_{i=1}^{m} g(x_i)$, където $g(x)$ е произволна характеристика на разпределението $P(x)$. Когато $m \to \infty$ $\overline{u} \xrightarrow{n.c.} u$ [*Geyer,92*].

Друг начин да се изчислят характеристиките на разпределението $P(x)$ е като първо се оцени плътността $p(x)$ и след това се получат статистиките от оценената плътност. Една оценка на $p(x)$ е средното на m условни плътности $p(x/y_i)$ [*Gelfand,Smith,90*]

$$\overline{p}(x) = \frac{1}{m}\sum_{i=1}^{m} p(x / y_i).$$

Тази оценка не използва извадките x_i, $i=1,2,\dots,m$, но използва извадките y_i, $i=1,2,\dots,m$ чрез условната плътност $p(x/y)$.

Друга алтернативна оценка на плътността $p(x)$, която използва само извадките x_i, $i=1,2,\dots,m$, е

$$\overline{p(z)} = \frac{1}{m.h}\sum_{i=1}^{m} K[(z-x_i)/h],$$

където $\overline{p(z)}$ е оценената плътност в точката z, $K(.)$ е функция, h е фиксирана константа, която определя гладкостта на оценената крива. Ако използваме функцията на плътността на нормалното разпределение, получаваме

$$\overline{p(z)} = \frac{1}{m.h}\sum_{i=1}^{m} \frac{K}{\sqrt{2\pi}} \exp\{-\frac{1}{2}[(z-x_i)/h]^2\}.$$

След като е намерена оценка на плътността $p(x)$ по един от двата описани по-горе начини, необходимите статистики (средното, дисперсията) могат да бъдат изчислени.

Използването на *Gibbs sampler* за намиране на маргинални разпределения в прагов модел е обсъдено в [*Sorensen, Andersen, Gianola, Korsgaard, 95*].

Друг симулационен подход за анализ на прагов модел с фиксирани и случайни ефекти от гледна точка на обобщените линейни смесени модели (Generalized linear mixed models) е обсъден в [*Engel, Buist, Visscher, 95*].

Използването на *Gibbs sampler* за оценяване на параметрите на обобщени линейни модели с Гаусови случайни ефекти е показано в [*Zeger, Karim, 91*].

MTGSAM (Multiple-Trait Gibbs SAMpler) е множество от програми, написани на Fortran, реализиращи алгоритъма *Gibbs sampleling*. Тяхното приложение за оценка на компонентите на модели в животновъдството е дискутирано в [*Tassell, Vleck, 96*].

BUGS е програма, написана на Modula-2, която реализира изводи в Бейсови мрежи. За тази цел тя използва симулационни *Gibbs sampling* алгоритми. В BUGS са реализирани език, на който се описва моделът на потребителя и множество от команди.

BUGS е предназначен за проблеми, за които или не съществуват точни аналитични решения, или стандартните приближени методи не могат да се приложат. За този клас проблеми се използва симулационния подход Монте-Карло (Markov Chain Monte Carlo).

В [*Spiegelhalter, Thomas, Best, Gilks, 96б*] са описани примери, които илюстрират използването на BUGS при решаването на проблеми от различни приложни области. [*Spiegelhalter, Thomas, Best, Gilks, 96a*] е ръководство за работа с тази програма.

Заключение

Основните научни приноси на дисертацията са:

- Предложен е нов вид архитектура на статистически системи.
- За първи път при реализиране на интелигентен статистически софтуер е приложена идеята за изграждане на вероятностна база знания.
- Формализирано е понятието вяра на агента във вероятностния модел. Намерени са достатъчни условия за монотонност на слабите и силните правила за вземане на решение за състоянието на вярата.
- В работата функциите полезност за първи път са представени чрез функции на вероятностни разпределения. Намерени са достатъчни условия за слаба монотонност на правилата за вземане на решения при използването на тези две конкретни функции полезност - Гаусовата и complementary log-log функция.
- Дефинирана е U-мрежа, представяща събития и полезност.
- Проблемът за определяне на вярата на агента във вероятностния модел се представя чрез наредената тройка $(\mathcal{B},\mathcal{U},\mathcal{D})$, където $\mathcal{B}$ е Бейсова мрежа, представяща вероятностната структура на проблема, $\mathcal{U}$ е U-мрежа, представяща структурата полезност, $\mathcal{D}$ е мрежа за вземане на решение. Построена е серия от алгоритми - за вземане на решение за състоянието на вярата, за оптимизиране на вярата, за генериране на достатъчните условия за монотонност на правилата за определяне състоянието на вярата.

Най-важната особеност на агента за проверка на статистически хипотези е неговата способност да изгражда планове за действие, когато задачата на потребителя не е коректно дефинирана, и да функционира разумно при изпълнението на тези планове. Чрез своите *намерения* агентът предсказва, до известна степен, бъдещото си поведение. Намеренията пораждат *задължения*. Преди да се пристъпи към изграждане на прототип на агент за проверка на статистически хипотези трябва да се формализират локалните му състояния *намерение* и *задължение*.

Основен въпрос при изграждането на мултиагентните системи е въпросът за кооперирането на агентите. Естествено е този проблем да се атакува със стохастичен подход.

Очаква се бъдещият интелигентен статистически софтуер да се базира на $(\mathcal{B}, \mathcal{U}, \mathcal{D})-$мрежи.

Литература:

1. [*Agresti,90*] Agresti A. *Categorical data analysis*, Wiley Interscience, 1990.
2. [*Berhold,67*] Berhold M.H.An Analysis of Contractual Incentives, Ph.D. dissertation, University of California, Los Angeles, 1967.
3. [*Berhold,73*] Berhold M.H. The use of distribution functions to represent utility functions, *Management science,* Vol. 19, No.7, pp. 825-829, 1973.
4. [*Besag,74*] Besag, J. Spatial Interaction and the Statistical Analysis of Lattice Systems (with discussion), Journal of the Royal Statistical Society, Ser. B, 36, 192-326, 1974.
5. [*Blum, 82*] Blum R. L. Discovery and Representation of Causal Relationships from a Large Time-Oriented Clinical Database: The RX Project, Springer-Verlag, New York, 1982.
6. [*Burkhard,93*] Burkhard Hans-Dieter, Haw to define agent properties - or: What is a fair agent?, In K. Ghedira and F. Sprumont (ads.): MAAMAW'93, 1993.
7. [*Burkhard,93б*] Burkhard Hans-Dieter, Liveness and Fairness properties in multi-Agent Systems, Proc. 13th ICAI 1993, pp. 325-330, 1993.
8. [*Burkhard,94*] Burkhard Hans-Dieter, On Fair Controls in Multi-Agent Systems, ECAI 1994, 11th European Conference on Artificial Intelligence Edited by A. Cohn, Published in 1994 by John Wiley & Sons, Ltd., pp. 254-258, 1994.
9. [*Castillo, Gutierrez, Hadi,97*] Castillo E., Gutierrez J.M., Hadi A. Expert Systems and Probabilistic Network Models, Springer, 1997, p. 605.
10. [*Chuang, Chen, Novick, 81*] Chuang, D. T., J. J. Chen, M. R. Novick, Theory and practice for the use of cut-scores for personnel decisions. *Journal of Educational Statistics,* 6, pp.129-152, 1981.
11. [*Cohen, Levesque,90*] Cohen P. R., H. J. Levesque Intention is choice with commitment, *Artificial Intelligence*, 42, pp. 213-261, 1990.
12. [*Connolly, Edmonds, 94*] Connolly J. H., E. A. Edmonds (eds) Computer Supported Cooperative work and Artificial Intelligence, Springer Verlag, 1994.

13. [*Corio, Lapalme,98*] Corio M, G. Lapalme Integrated generation of graphics and text: a corpus study, submitted to CVIR'98, 1998.
14. [*Cox, Wermuth,93*] Cox D. R., N. Wermuth, Linear dependencies represented by chain graphs, Statistical Science, Vol. 8, No. 3, pp. 204-283, 1993.
15. [*Dambroise, 86*] Dambroise E. Muse, an Expert System in Statistics, in *COMPSTAT 1986: Proceedings in Computational Statistics,* F. De Antoni, N. Lauro, and A. Rizzi, editors, Physica-Verlag, pp. 271-276, 1987.
16. [*Daniel, Terrell, 89*] Daniel W.W., J.C. Terrell Business Statistics for Management and Economics, Houghton mifflin company, 1989.
17. [*Davis, Hickman, Novick, 73*] Davis C. E., J. Hickman, M. R. Novick, A primer on decision analysis for individually prescribed instruction, Iowa City: Publication of the Research and Development Division, The American College Testing Program, 1973.
18. [*Decker, Lesser, 95*] Decker K.S., V.R. Lesser Designing a family of coordination algorithms, ICMAS-95, pp. 73-80, 1995.
19. [*Dodra,90*] Dodra W., K. A. Froeschl, W. Grossmann WAMASTEX - Heuristic guidance for statistical analysis, COMPSTAT 1990, pp. 93-98.
20. [*Engel, Buist, Visscher, 95*] Engel B., W. Buist, A. Visscher, Inference for threshold models with variance components trom the generalized linear mixed model pespective, *Genet Sel Evol*, 27, pp. 15-32, 1995.
21. [*Fasciano,96*] Fasciano M. Generation integree de textes et de graphiques statistiques. Ph.D. thesis, Universite de Montreal, 1996.
22. [*Fasciano,Lapalme,98*] Fasciano M., G. Lapalme, Intentions in the coordinated generation of graphics and text from tabular data, submitted to Knowledge and Information System, p.23, 1998.
23. [*Gale,86a*] Gale W. A. Statistical applications of Artificial Intelligence in *Handbook of Statistics, Volume 9, editor C. R. Rao,* pp. 237-285, 1986.
24. [*Gale,86*] Gale W. A. (ed.) Artificial Intelligence and Statistics, Addison Wesley, Reading, Massachusetts, 1986.
25. [*Gale,90*] Gale W. A. Student: A tool for constructing consultation systems in data analysis, Invited Paper in 46th Session of the ISI, pp. 1-17, 1990.
26. [*Gelfand, Smith, 90*] Gelfand A. E., A. F. M. Smith Sampling-Based Approaches to Calculating Marginal Densities, *Journal of the*

American Statistical Association, Vol. 85, No. 410, pp. 398-408, 1990.

27. [*Geman, Geman, 84*] Geman S., D. Geman Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images, in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6, pp. 721-741, 1984.
28. [*Geyer,92*] Geyer C. J. Practical Markov chain Monte Carlo, *Stat. Sci.*, 7, pp. 467-511, 1992.
29. [*Glymour, Madigan , Pregibon, Smyth, 97*] Glymour C., D. Madigan, D. Pregibon, P. Smyth Statistical themes and lessons for data mining, Data Mining and Knowledge Discovery, 1, pp. 11-28, 1997.
30. [*Goldstein, 1997*] Goldstein R. Statistical packages, in *Encyclopedia of statistical sciences* [edited by Kotz S., Read C., Banks D.], Vol. 1, A Wiley-Interscience publication, pp. 504–510, 1997.
31. [*Hand,85*] Hand D. J. Statistical Expert Systems: necessary attributes, Journal of Applied Statistics, Vol. 12, No 1, 1985, pp. 19-27.
32. [*Hand,90*] Hand D. J. Practical experience in developing statistical knowledge enhancement systems, Annals of Mathematics and Artificial Intelligence 2, 1990, pp. 197-208.
33. [*Hines, Montgomery,90*] Hines W.W., D.C. Montgomery Probability and Statistics in Engineering and Management Science, John Wiley & Sons, 1990.
34. [*Howard, 66*] Howard R.A. Information value theory , IEER Transactions on Systems Science and Cybernetics, SSC-2(1) 22-26, 1966.
35. [*Kerpedjiev, Noncheva,90*] Kerpedjiev S., V. Noncheva Intelligent Handling of Weather Forecasts, 13th *International Conference on Computational Linguistics* COLING -90, Helsinki, Vol. 3, pp. 379-381, 1990.
36. [*Lindley,87*] Lindley Dennis V. The probability approach to the treatment of uncertainty in Artificial Intelligence and Expert Systems, *Statistical Science*, 1987, Vol.2, No.1, 17-24.
37. [*Logchem, 94*] *Maja van* Logchem, Compilation and classification of statistical tests, Rotterdam: SRM-Documentation Centre, pp. 41, 1994.
38. [*Luce, Raiffa, 57*] Luce R. D., H. Raiffa, Games and decisions. New York: John Wiley and Sons, 1957.

39. [*Mendenhall, 87*] Mendenhall W. Introduction to Probability and Statistics, PWS Publishers, 1987.
40. [*Mikulesky,Chovatova,87*] Mikulesky P., I. Chovatova A knowledge-based tool for a scientific profram library, in Artificial Intelligence and Information-Control Systems of Robots, I. Plander (editor), Elsevier Science Publishers, 1987, pp. 355-358.
41. [*Mikulesky,Kalas,Kelemen,86*] Mikulesky P., I. Kalas, J. Kelemen Representation of meta-knowledge in expert systems, Computers and Artificial Intelligence, 5, No. 3, 1986, pp. 223-234.
42. [*Minsky, 85*] Minsky M. The Society of Mind, Simon and Shuster, 1985.
43. [*Molenaar,88*] Molenaar I. W., Formal statistics and informal data analysis or why laziness should be discouraged, Statistica Nederlandica 42, 1988, No. 2.
44. [*Molenaar,88б*] Molenaar I. W., Statistical consultants and statistical expert systems, COMPSTAT 1988, pp. 187-192, 1988.
45. [*Motulsky,95*] Motulsky H. J. Intuitive Biostatistiks: Choosing statistical tests; Oxford University Press Inc., 1995.
46. [*Nelder,88*] Nelder J. A. How should the statistical expert system and its user see each other?, Proc. of COMPSTAT 1988, pp.107-116, 1988,
47. [*Noncheva,95*] Noncheva V. Statistical expert systems - a stage of statistical software evolution, Journal of Technical University of Plovdiv, Vol. 1 "Fundamental Sciences and Applications", pp. 101-106, 1995.
48. [*Noncheva,96*] Noncheva V. An approach to learning statistics with modern artificial intelligence tools, *Applications of Mathematics in Engineering and Business*, 1996.
49. [*Noncheva,98*] Noncheva V. An approach to optimizing the agent's belief about the probability distribution of a population, Научни трудове на Пловдивски Университет "Паисий Хилендарски", т.33, кн.3, 1998.
50. [*Noncheva,Stojanov,95*] Noncheva V., S. Stojanov, What should a future multiagent system for statistical investigations be?, SDA95, Proc. of Statistical data analysis 96, 1996, pp. 69-75.
51. [*Noncheva,Stojanov,96*] Noncheva V., S. Stojanov, Extracting statistical information from text in Multiagent system for statistical investigations, Proc. of *Statistical data analysis 96*, 1996, pp. 176-182.

52. [*Pearl,86*] Pearl J. Fusion, Propagation and Structuring in Belief Networks. *Artificial Intelligence*, 29, pp. 241-288, 1986.
53. [*Pearl,93*] Pearl J. From Bayesian Networks to Causal Networks, 49th session of the International statistical Institute, florence, Italy, august 25 - september 3, 1993.
54. [*Pearl,93б*] Pearl J. Graphical models, causality, and intervention, *Comment in Statistical Science*, Vol. 8, No. 3, 203, pp. 266-269, 1993.
55. [Poh, Horvitz, 96] Poh K.L., E. Horvitz A Graph-Theoretic Analysis of Information Value, Proceedings of the Twelfth Annual Conference on *Uncertainty in Artificial Intelligence*, 1996.
56. [*Portier, Lai, 83*] Portier K. M., P. Lai, A Statistical Expert System for Analysis Determination, in *Proceedings of the Statistical Computing Section,* American Statistical Association, Virginia, pp. 309-311, 1983.
57. [*Prat at al.,94*] Prat A., Lores J., Sanjeevan K., Sole I., Catot J., Application of a multiagent distributed architecture for time seris forecasting, COMPSTAT-94, 1994, pp. 301-316.
58. [*Pregibon,Gale,84*] Pregibon D., W. Gale REX: an expert system for regression analysis, COMPSTAT 1984, Prague, 1984, pp. 242-248.
59. [*Russell, Norving, 95*] Russell S., P. Norving, Artificial Intelligence. A Modern Approach, p. 932, 1995.
60. [*Shachter,86*] Shachter R.D. Evaluating influence diagrams, *Operations Research,* No.34, pp. 871-882, 1986.
61. [*Shachter,88*] Shachter R.D. Probabilistic Inference and Influence Diagrams, *Operations Research,* 36, pp. 591-604, 1988.
62. [*Shafer,87*] Shafer G. Probability judgement in Artificial Intelligence and Expert system, Statistical Science, Vol. 2, No. 1, pp. 3-44, 1987.
63. [*Shoham, 91*] Shoham Y. AGENT0: A simple agent language and its interpreter. In Proceedings of the Ninth National Conference on Artificial Intelligence, 1991, pp. 704-709.
64. [*Shoham,97a*] Shoham Y. A symmetric view of probabilities and utilities. In *Proc. of IJCAI-97,* 1997.
65. [*Shoham,97б*] Shoham Y. Two senses of conditional utility. In *Proc. of UAI-97,* 1997.
66. [*Singh,92*] Singh M. P. A critical examination of the Cohen-Levesque theory of intentions, *ECAI 92*, 1992.
67. [*Siriopoulos,87*] Siriopoulos C. Two future directions in statistical research, The 17th European Meeting of Statisticians, 1987, Greece.

68. [*Smith, Roberts, 93*] Smith A. F. M. , G. O. Roberts Bayesian Computation via the Gibbs Sampler and Related Markov Chain Monte Carlo Methods, *J. R. Statist. Soc.* B, 55, No. 1, pp. 3-23, 1993.
69. [*Sorensen, Andersen, Gianola, Korsgaard, 95*] Sorensen D. A., S. Andersen, D. Gianola, I. Korsgaard, Bayesian inference in threshold models using Gibbs sampling, *Genet Sel Evol*, 27, pp. 229-249, 1995.
70. [*Spiegelhalter, Thomas, Best, Gilks,96a*] Spiegelhalter D., A. Thomas, N. Best, W. Gilks, BUGS 0.5, *Institute of Public Health, Cambridge, ftp:ftp.mrc-bsu.cam.ac.uk*, 1996.
71. [*Spiegelhalter, Thomas, Best, Gilks,96б*] Spiegelhalter D., A. Thomas, N. Best, W. Gilks, BUGS 0.5 Examples, *Institute of Public Health, Cambridge, ftp:ftp.mrc-bsu.cam.ac.uk*, 1996.
72. [*Stevens,84*] Stevens L. Artificial Intelligence. The Search for the Perfect Machine, Hayden Book Company, Hasbrouck Heights, N.J., 1984.
73. [*Tanner, Wong, 87*] Tanner M. A., Wong W. H. The calculation of posterior distributions by data augmentation, in *J Am Stat Assoc* 82, pp. 528-540, 1987.
74. [*Van den Berg,90*] Van den Berg G.M., R. A. Visser Knowledge Modelling for statistical consultation system - Two emperical studies, COMPSTAT 1990, pp. 75-80.
75. [*Van Tassell, Van Vleck, 96*] Van Tassell C. P., L. D. Van Vleck, Multiple-Trait Gibbs Sampler for Animal Models: Flexible Programs for Bayesian and Likelihood-Based (Co) Variance Component Inference, *J. Anim. Sci.*, 74, pp. 2586-2597, 1996.
76. [*Verma, Pearl,90*] Verma T., J. Pearl Causal Networks: Semantics and Expressiveness. In Shachter, Levitt, Kanal and Lemmer, editors, *Uncertainty in Artificial Intelligence 4.* North Holland, Amsterdam, 1990, pp. 69-76.
77. [*Vos, 90*] Vos H.J. Simultaneous Optimization of Decisions Using a Linear Utility Function, Journal of Educational Statistics, 1990, Vol. 15, No. 4, pp. 309-340.
78. [*Vos, 91*] Vos H.J. Simultaneous Optimization of the Aptitude Treatment Interaction Decision Problem with Mastery scores, in M. Wilson (Ed.), *Objective measurement: Theory into practice* (Vol.1, pp. 313-331) , Norwood, New Yersey: Alex Publishing Corporation, 1991.
79. [*Vos, 93*] Vos H.J. Simultaneous Optimization of Selection-placement Decisions followed by a Mastery Decision, in F. J. Maarse et al.

(Eds.), *Computers in psychology: Tools for experimental and applied psychology* (Vol.4, pp. 235-244) , Lisse, The Netherlands: Swets & Zeitlinger, 1993.

80. [*Vos, 94*] Vos H.J. Simultaneous Optimization of Test-based Decisions in Education, Ph.D. dissertation, University of Twente, Netherlands, 1994.
81. [*Vrijhof, Mellenbergh, van den Brink, 83*] Vrijhof, B. J., G. J. Mellenbergh, W. P. van den Brink, Assessing and studying utility functions in psychometric decision theory. *Applied Psychological Measurement*, 7(3), 341-357, 1983.
82. [*Winston, 1993*] Winston P. Artificial Intelligence, 1993.
83. [*Wittkowski,88*] Wittkowski K. M. Building a statistical expert system with knowledge Bases of different levels of application, COMPSTAT 1988, pp. 129-134.
84. [*Wittkowski,90*] Wittkowski K. M. Statistical knowledge-based systems - critical remarks and requirements for Approval, COMPSTAT 1990, pp. 49-56.
85. [*Wooldridge, Jennings,95*] Wooldridge M., Jennings N. Agent Theories, Architectures, and Languages: A Survey, In M. Wooldrige and N. R. Jennings, editors, Intelligent Agents - Proceedings of the 1994 Workshop on Agent Theories, Architectures, and Languages, 1995.
86. [*Wooldridge, Jennings,95б*] Wooldridge M., Jennings N. Intelligent Agents: Theory and Practice, submitted to Knowledge Engineering Review 1995.
87. [*Zeger, Karim, 91*] Zeger S. L., R. Karim, Generalized Linear Models With Random Effects; A Gibbs Sampling Approach, *Journal of the American Statistical Association, Vol. 86, No. 413*, pp. 79-86, 1991.
88. [*Zilberstein,93*] Zilberstein S., Operational rationality through compilation of anytime algorithms, Ph.D. Thesis, University of Massachusetts, 1993.
89. [*Айвазян,Енюков,Мешалкин,83*] Айвазян С. А., И. С. Енюков, Л. Д. Мешалкин, Прикладная статистика. Основы моделирования и первичная обработка данных, Финансы и статистика, Москва, 1983.
90. [*Афифи, Эйзен, 82*] Афифи А., С. Эйзен, Статистический анализ. Подход с использованием ЭВМ, Мир, Москва, 1982.
91. [*Большев, 87*] Большев Л. Н. Избранные труды: Теория вероятностей и математическая статистика, Наука, Москва, 1987.

92. [*Винце,Върбанова,81*] Винце И., Върбанова М., Непараметрични статистически методи, Наука и изкуство, София, 1981.
93. [*Герасимович,83*] Герасимович А. И. Математическая статистика, Минск, Вышэйшая школа, 1983.
94. [*Димитров,92*] Димитров Е. Една възможност за организиране на интелигентно асистиране (работен материал), 1992.
95. [*Димитров, Янев, 90*] Димитров Б., Н. Янев Вероятности и статистика, Университетско издателство "Климент Охридски", София, 1990.
96. [*Кендалл, Стьюарт, 73*] Кендалл М., А. Стьюарт Статистические выводы и связи, Наука, Москва, 1973.
97. [*Климов,75*] Климов Г. П. Приложна математическа статистика, Издателство "Наука и изкуство", София, 1975.
98. [*Колев, 93*] Колев Н. Приложна статистика, София, 1993.
99. [*Кръстев,90*] Кръстев Б. Граматика на комуникацията, Университетско издателство "Св. Климент Охридски", София, 1990.
100. [*Леман,79*] Леман Э. Проверка статистических гипотез, "Наука", Москва, 1979.
101. [*Нончева,88*] Нончева В. Алгоритъм за автоматично пренасяне на думи в българския език, Математика и математическо образование, БАН, София, 479-482, 1988.
102. [*Нончева,90*] Нончева В. Синтез на текст на метеорологична прогноза, Математика и математическо образование, БАН, София, 269-274, 1990.
103. [*Речник на българския език, 79*] Речник на българския език, Главен редактор Кристалина Чолакова, Издателство на Българската академия на науките, том 2, 1979.
104. [*Философски речник, 78*] Философски речник, Под редакцията на М. Бъчваров, М. Драганов, С. Стоев, София, 1978.
105. [*Хог,Крейг,82*] Хог Р., А. Крейг Увод в математическата статистика, Издателство "Техника", 1982.

Публикации на дисертанта по дисертацията

1. Kerpedjiev S., V. Noncheva Intelligent Handling of Weather Forecasts, 13th *International Conference on Computational Linguistics* COLING -90, Helsinki, Vol. 3, pp. 379-381, 1990.
2. Noncheva V. Statistical expert systems - a stage of statistical software evolution, *Journal of Technical University of Plovdiv, Vol. 1* "Fundamental Sciences and Applications", pp. 101-106, 1995.
3. Noncheva V., S. Stojanov, What should a future multiagent system for statistical investigations be?, *SDA95*, Proc. of Statistical data analysis, pp. 69-75, 1996.
4. Noncheva V. An approach to learning statistics with modern artificial intelligence tools, *Applications of Mathematics in Engineering* (Proceedings of the XXII Summer School, Sozopol, Bulgaria), pp. 262-267, 1996.
5. Noncheva V., S. Stojanov, Extracting statistical information from text in multiagent system for statistical investigations, *SDA96*, Proc. of *Statistical data analysis*, pp. 176-182, 1996.
6. Noncheva V. An approach to optimizing the agent's belief about the probability distribution of a population, Научни трудове на Пловдивски Университет "Паисий Хилендарски", т. 33, кн.3, 1998.
7. Noncheva V. Agent Belief: Presentation, Propagation, and Optimization, *INFORMATICA, Vol.10, No.4*, 1999.
8. Noncheva V. A Statistical Agent's Belief in Gaussian Model, *Mathematica Balkanica, Vol. 14, Fasc. 3-4*, 2000 (to appear).
9. Нончева В. Алгоритъм за автоматично пренасяне на думи в българския език, Математика и математическо образование, БАН, София, 479-482, 1988.
10. Нончева В. Синтез на текст на метеорологична прогноза, Математика и математическо образование, БАН, София, 269-274, 1990.