УНИВЕРЗИТЕТ „ГОЦЕ ДЕЛЧЕВ“ ШТИП
ТЕХНОЛОШКО-ТЕХНИЧКИ ФАКУЛТЕТ

Д-р. Димко Димески
Д-р. Винета Сребренкоска

Дизајн и анализа на експерименти

Сефакторен план на експерименти

Штип, 2014

Предговор

Решавањето на мноштво проблеми во хемијата и хемиската технологија, како и во останатите гранки од инженерството, често е поврзано со изведување на сложени и скапоцени експерименти. Оттука е разбирливо значењето од постоењето на начини и методи за оптимално планирање на експериментите, кои во низа случаи овозможуваат суштествено да се скратат времето и материјалните трошоците при извршување на истражувањето. Долго време редоследот на извршување на експериментите бил базиран на личното искуство и интуиција на истражувачот. Првите обиди да се применат математички методи за оптимално планирање на експериментите биле од страна на англискиот математичар Р. Фишер во почетокот на XX век. Теоријата за планирање на експериментите, особено брзо почнала да се развива по 1951 година, а е поврзана со работата на Г. Бокс и К. Вилсон. Методите на оптимално планирање на експериментите овозможуваат користење на математички методи не само за обработка на резултатите од тестовите, туку и во фазата на подготовката и спроведувањето на експериментите. Работата на истражувачите кои ги користат тие методи значително се олеснува, затоа што се изведува по логично дефинирана, редоследна постапка.

Во современата математичка теорија за оптимално планирање на експериментите постојат два основни оддели:

1. Планирање на експериментите заради изучување на механизмите на сложените процеси и својствата на повеќе-компонентните системи.

2. Планирање на експериментите заради оптималност на технолошките процеси и својствата на повеќе-компонентните системи.

За пошироко распростирање на методите за планирање (дизајн) и анализа на експериментите меѓу истражувачите неопходно е постоење на методолошки упатства напишани на разбирлив, јасен и едноставен начин. Токму од таа мисла се водевме при пишувањето на овој е-учебник; да биде концизен, јасен, лесно разбирлив, и применлив.

Авторите *Штип,февруари, 2014*

Содржина

Краток историјат на експериментирањето

Најголемиот придонес на теоријата и практиката за статистичкото донесување заклучоци во XX век е во дизајнирањето на експерименти. Постојат четири ери во современиот развој на статистичкото планирање на експериментите. *Агрикултурната ера* е во знакот на Роланд Фишер (Roland A. Fisher).Во текот на неговата пионерска работа при Rothmasted Experimental Station во близината на Лондон во 1920-те и раните 1930-ти години, тој работел на статистика и анализа на податоците. Фишер го проучувал приносот во земјоделството во зависност од видот и количината на ѓубриво, климата и врнежите. Забележал дека грешките кои се прават во експериментите кои генерирале податоци често оневозможуваат анализа на податоците од системот. Соработувајќи со научници и истражувачи од многу полиња,Фишер развива гледиште кое доведува до трите основни принципи на планираниот експеримент: рандомизација, репликација и блокирање.

а)

б)

Слика 1. Роналд Фишер (а) и Џорџ Бокс (б)

Втората или *индустриската ера* е катализирана од развојот на методологијата за одзивната површина од страна на Џорџ Бокс (George Box) и неговите соработници во 1951г. Тој сфатил и го искористил фактот дека многу индустриски експерименти фундаментално се разликуваат од тие во земјоделството од два аспекти: (1) одзивната варијабла е видлива многу брзо, и (2) експериментаторот многу брзо може да ги добие круцијалните информации од мала група на изведени експерименти кои можат да се искористат за планирање на следниот експеримент. Во следните 30 години техниката на одзивната површина (види поглавје за факторен експеримент) и други експериментални техники се проширувааt во хемиската и други процесни индустрии, главно во истражувачка и развојна работа. Меѓутоа, примената на статистичкиот дизајн на ниво на погон или производен процес сѐ уште не била доволно раширена.

Во текот на оваа втора или индустриска ера Кифер (Kiefer, 1959, 1961) предлага формален пристап на селектирање на дизајн кој е специфичен за целта на експериментирањето,за да се постигне оптималност. Овој пристап не наоѓа некоја голема примена заради недостаток на компјутерски алатки за негова имплементација.

Третата ера од статистичкиот дизајн е одбележана со работата на Геничи Тагучи (Genichi Taguchi) во доцните 1970-ти години. Неговата работа има огромно влијание врз проширувањето на интересот за планираните експерименти. Тагучи е поборник на користење планирани експерименти, како што вели, за робустен дизајн. Тоа значи:

- Правење на процесот да биде неосетлив на факторите од амбиентот (околината) или од други фактори кои тешко се контролираат.
- Правење производите да бидат неосетливи на варијациите пренесени од компонентите.
- Наоѓање на нивоа на процесните варијабли кои ја доведуваат средната вредност на саканото ниво а истовремено редуцирајќи ја варијабилноста околу тоа ниво.

Тагучи предлага високо-фракциски фактор дизајни и воведува нови статистички методи за остварување на робустноста.

Како резултат од работата на Тагучи, произлегуваат неколку позитивни последици. Прво, примената на планирање на експеримент многу се раширува и во дискретните делови од индустријата како што е: автомобилската, авио, електрониката, индустријата за полуспроводници, и многу други, кои претходно воопшто не ја користеле оваа техника. Второ, почнува *четвртата ера* на статистичкиот дизајн на експерименти. Оваа ера се карактеризира со побуден голем интерес за статистички дизајн и од страна на истражувачите и од страна на производителите и со развој на многу нови корисни пристапи кон експерименталните проблеми. Новите пристапи се алтернативи на техниките методи на Тагучи што овозможуваат неговиот инженерски концепт да биде спроведен во практика ефикасно и ефективно, пред сѐ во функција на подобрување на квалитетот.Трето, има огромно подобрување во компјутерскиот софтвер за дизајн и евалвација на експерименти со многу нови карактеристики и способности. Четврто, формалната едукација за статистичките експериментални дизајни станува составен дел од додипломските и постдипломските студии на многу универзитети. Успешното интегрирање на добриот експериментален дизајн во инженерството и науката е клучен фактор во идната индустриска конкурентност.

§1. ВОВЕД

1.1. Стратегија на експериментирањето

Посматрањето или опсервацијата на некој систем или процес кога е во функција, е важен дел од процесот на стекнување знаења, и е интегрален дел на учењето и разбирањето, како системите и процесите функционираат. Меѓутоа, за да се сфати што се случува со процесот ако се променат некои влезни (инпут) фактори не е доволно само да се посматра – потребно е да се менуваат факторите. За навистина да се разбере релацијата: *причина-и-последица* на некој систем, мора намерно да се менуваат влезните варијабли во системот и да се посматраат промените на излезот (аутпут) од системот кои што тие варијабли ги предизвикуваат. Со други зборови, потребно е да се извршат **експерименти** врз системот. Опсервациите на системот или процесот водат кон теории и хипотези за тоа што го прави системот да функционира, но за да се покаже дали тие теории се точни, потребни се експерименти.

Истражувачите вршат експерименти, буквално, во сите полиња од интерес, особено за да откријат нешто посебно за даден систем или процес. Секој изведен експеримент е всушност тест. Експериментот формално, можеме да го дефинираме како тест или серија од тестови кај кој/кои намерно вршиме промени на влезните варијабли од процесот или системот за да можеме да ги посматраме и идентификуваме промените кои настануваат на излезот или **одзивот** на системот. Можно е ние да сакаме да одредиме кои влезни варијабли се одговорни за опсервираните промени во одзивот, да развиеме модел кој ќе го поврзе одзивот со главните влезни варијабли и да го искористиме тој модел за подобрување на системот или процесот или за донесување на одлуки.

Овој учебник е наменет за планирање и изведување на експерименти и за анализирање на добиените резултати со цел да се добијат валидни и објективни заклучоци.Експериментирањето игра голема улога во технологијата, комерцијализацијата и во активностите за реализирање на нови продукти. Тоа опфаќа дизајн на новиот продукт, развој на производен процес и подобрување на процесот. Целта, во многу случаи, е да се развие **робустен** процес т.е. процес чија променливост минимално зависи од надворешните услови. Исто така постои голема примена на планираните експерименти во непроизводни гранки, како што се: маркетингот, услужните дејности, информатиката и општите бизнис активности.

Еве пример за експеримент. Да претпоставиме дека еден металург е заинтересиран за проучување на ефектот на два процеси за зголемување на тврдината на алуминиумска легура – калење во раствор на масло и калење во раствор на солена вода. Овде целта на експериментаторот (металургот) е да утврди кој од растворите дава максимална тврдина за оваа легура. Металургот одлучува да подложи повеќе тест примероци на калење во двата раствори и да ја мери нивната тврдина по калењето. Просечната тврдина на примероците ќе се искористи за да се утврди кој раствор е подобар.

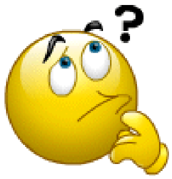
Имајќи го во предвид овој едноставен експеримент можат да се поставаат низа важни прашања:

1. Дали само овие два раствори се од потенцијален интерес?
2. Постојат ли и други фактори кои влијаат врз тврдината и кои треба да се истражат или контролираат во текот на експериментот (пр. температурата на растворите)?
3. Колку примероци треба да се тестираат во секој од растворите?
4. По кој редослед треба да се калат примероците и по кој редослед треба да се соберат податоците?
5. Каков метод за анализа на податоците треба да се употреби?
6. Која разлика во просечната тврдина на примероците ќе се смета за важна?

Сите овие прашања, и можеби многу други, треба задоволително да се одговорат пред да се изврши експериментот.

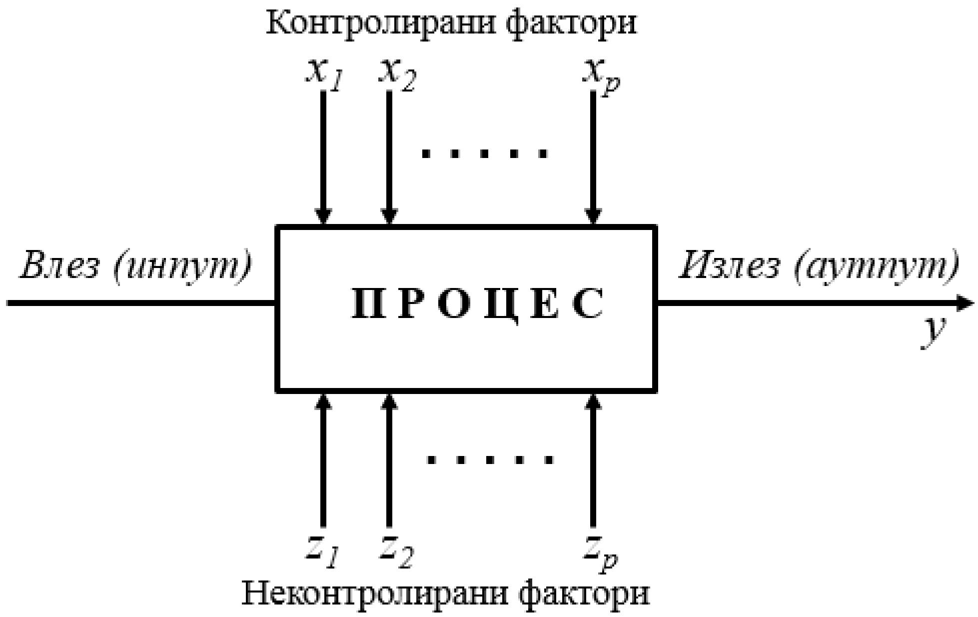
Експериментирањето е витален дел од научните или инженерските методи. Има низа ситуации каде научниот феномен е толку познат што можат да се развијат математички модели од (добро) познатите принципи. Моделите на вакви феномени кои произлегуваат директно од физичкиот механизам обично се нарекуваат **механички модели**. Прост пример за ова е познатиот Омов закон кој ги поврзува отпорот, јакоста и напонот на струјата, $U = RI$. Меѓутоа, повеќето проблеми во науката и инженерството бараат опсервација на системот во функција и експериментирање за да се разјасни зошто и како тој функционира? Добро дизајнираните експерименти можат често да водат кон модел за перформансите на системот; вака експериментално определените модели се нарекуваат **емпириски модели**. Во понатамошниот текст ќе бидат презентирани техники за конвертирање на резултатите од планираниот експеримент во емпириски модел на проучениот систем. Овие емпириски модели можат да бидат манипулирани од страна на научниците или инженерите,исто како што можат да бидат манипулирани и механичките модели.

Доброто дизајнираниот (планираниот) експеримент е важен заради фактот што резултатите и заклучоците што ќе се изведат врз негова база, во голем дел, зависат од начинот на кој тие податоци се добиени.



Дизајн или план?
*Во контекст на овој курс „**план**“ и „**дизајн**“ се користат како синоними т.е. како термини со исто значење. Во литература на англиски јазик се користи терминот „дизајн“. Така на пример „планирање на експерименти“ е “design of experiments”; „сефакторен план на експеримент“ е “full factorial experimental design”. Во германската, француската, руската и источно-европската терминологија се користи терминот „**план**“.*

За да го илустрираме ова, ќе претпоставиме дека металургот, од претходниот експеримент, земал примероци од еден третман во масло и примероци од друг третман во вода. Сега, ако се спореди средната тврдина, металургот не е во можност да каже колку многу од забележаната разлика се должи на различниот медиум, а колку на различните третмани. На тој начин, методата на собирање на податоци има контрадикторен ефект врз заклучокот кој треба да се донесе од експериментот.



Слика: 1.1. Генерален модел за систем или процес

Генерално, експериментите се користат за проучување на перформансите на процесите или системите. Процесите или системите можат да бидат претставени со модел прикажан на сликата 1.1. Обично процесот го замислуваме како комбинација од операции, машини, методи, луѓе и други елементи која трансформира некаков инпут (често материјал) во аутпут кој има една или повеќе видливи одзивни варијабли. Некои од процесните варијабли и особини на материјалите $x_1, x_2, \ldots, x_p$ се контролибилни, додека други, $z_1, z_2, \ldots, z_q$ се неконтролибилни.

Целта на експериментот може да го опфати следново:

1. Одредување кои варијабли се највлијателни врз одзивот y.
2. Одредување како да се поставaт влијателните x-ови, така што y ќе биде скоро секогаш блиску до номиналната вредност.
3. Одредување каде да се поставaт влијателните x-ови, така што варијабилноста на y да биде мала.
4. Одредување каде да се поставaт влијателните x-ови, така што ефектот на неконтролибилните варијабли $z_1, z_2, \ldots, z_q$ да биде минимизиран.

Како што може да се види од претходната дискусија, експериментите често вклучуваат неколку фактори. Обично, цел на експериментаторот е да го утврди влијанието што го имаат факторите врз одзивот на системот. Општиот пристап кон планирањето и изведувањето на експериментите се нарекува стратегија на експериментирањето. Еден експериментатор може да употреби неколку стратегии. Некои од тие стратегиите се следниве:

1.1.1. Стратегија на најдобро нагаѓање (*Best-guess approach*)

Оваа стратегија на експериментирање се заснова на пристапот *проба-и-грешка*. Често во практиката ја користат научниците и инженерите. Многу пати функционира задоволително добро, затоа што експериментаторите што ја користат имаат големо техничко и теоретско познавање за системот што го проучуваат како и добро практично искуство. Користејќи го своето знаење експериментаторите избираат т.е. нагаѓаат кој е најдобриот сет од факторите за даден процес/систем. Врз база на резултатите од еден тест со менување на еден/повеќе фактори се прави следниот, па следниот, па... и така може да се експериментира до бесконечност. Оваа стратегија има, најмалку, два недостатоци. *Прво*, да претпоставиме дека првиот тест не дава задоволителни резултати. Потоа експериментаторот треба да направи друго нагаѓање со променето ниво на факторите. Ова може да трае долго време без каква било гаранција за успех. *Второ*, да претпоставиме дека првото нагаѓање дава прифатлив резултат. Во ваков случај експериментаторот го прекинува тестирањето, иако нема гаранција дека го нашол најдоброто решение.

1.1.2. Стратегија на еден-по-еден-фактор

Друга стратегија на експериментирање, која интензивно се користи во практика, е пристапот на **еден-по-еден-фактор** (*one-factor-at-a-time* или OFAT). OFAT пристапот се состои од избор на стартна точка или основна линија на нивото на секој од факторите, и потоа сукцесивно се менува нивото на секој фактор во рамките на неговиот опсег, додека останатите фактори се држат константни на ниво од основната линија. По изведувањето на сите тестови, најчесто, се конструираат серија графикони кои покажуваат како варијаблата на одзивот се менува со промената на секој фактор поединечно додека другите фактори се држат константни.

Главен недостаток на овој пристап е тоа што не ја зема во предвид интеракцијата. *Интеракцијата* претставува ефект кој се манифестира со тоа што еден фактор не врши еднакво влијание врз одзивот при различно ниво на друг фактор. Интеракцијата меѓу факторите е многу честа. И доколку ја има, стратегијата на еден-по-еден-фактор ќе продуцира лоши резултати. Многу луѓе не се свесни за ова и OFAT експериментите многу често се изведуваат во практика. Еден-по-еден-фактор експериментите секогаш се помалку ефикасни од останатите методи базирани на статистички пристап кон планирањето на експериментите.

1.1.3. Факторна стратегија

Правилен пристап за справување со неколку фактори е да се изведува експеримент на факторот. Ова е експериментална стратегија кај која факторите се менуваат ***заедно***, место еден по еден. Фактор концептот на дизајн на експерименти е екстремно битен и затоа му е посветено најголемо внимание во овој учебник.

1.2. Типична примена на експерименталниот дизајн

Методите на експерименталниот дизајн наоѓаат широка примена во многу гранки од науката и техниката. Како што веќе претходно спомнавме, експериментирањето можеме да го гледаме како дел од научниот процес и како еден од начините со кој учиме, како системот или процесот функционира. Генерално, учиме низ серија од активности во кои правиме претпоставки за процесот, изведуваме експерименти за да добиеме податоци од процесот, и потоа ги користиме информациите од експериментот за да воспоставиме нови претпоставки, кои водат до нови експерименти, итн.

Експерименталниот дизајн е критички важна алатка во научната и инженериската дејност за подобрување на процесот за реализација на производ. Критичните компоненти на овие активности се состојат во: нов дизајн на производниот процес и развојот, како и во менаџирање на процесот. Примената на техниките на експерименталниот дизајн во раната фаза на развој на процесот може да резултира во:

- Зголемена продуктивност на процесот.
- Редуцирана варијабилност и поблиско совпаѓање со номиналните или целните барања.
- Намалено време за развој.
- Редуцирани севкупни трошоци

Методите на експерименталниот дизајн исто така се од фундаментално значење во техничкиот дизајн, кадешто се развиваат нови производи, а постоечките – се подобруваат. Некои од апликациите на експерименталниот дизајн во техничкиот дизајн вклучуваат:

- Евалвација и споредба на основната конфигурација на дизајнот.
- Евалвација на алтернативни материјали.
- Избор на параметрите на дизајнот со цел производот да служи во широк опсег на услови т.е. да биде робустен.
- Одредување на основните параметри на дизајнот на производот што влијаат врз неговите перформанси.
- Создавање на нови производи.

Примената на експерименталниот дизајн во создавањето на производи резултира во производи кои се полесни за производство и имаат зголемени функционални перформанси и надлежност (сигурност), помали производни трошоци и пократко време за дизајнирање, како и време за развој. Дизајнираните експерименти, исто така, имаат екстензивна примена во маркетингот, истражувањето на пазарот, трансакциските и сервисните активности и генералните деловни операции.

1.3. Основни принципи

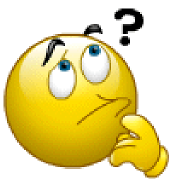
За еден експеримент нај ефективно да се изведе потребно е да се изведе на научна основа. Под *статистички дизајн на експерименти* се подразбира процес на планирање на експериментот така што соодветните податоци ќе бидат собрани и анализирани со статистички методи што ќе резултира во валидни и објективни заклучоци. Статистичкиот пристап кон експерименталниот дизајн е неопходен ако сакаме да извлечеме разумни заклучоци од податоците. Ако проблемот вклучува податоци кои се подложни на експериментална грешка, статистичките методи се единствениот објективен пристап за анализа. Така, се јавуваат два аспекти на секој експериментален проблем: дизајн на експериментот и статистичка анализа на податоците. Овие два аспекти се тесно поврзани бидејќи методата на анализа директно зависи од применетиот дизајн.

Трите основни принципи на експерименталниот дизајн се: рандомизација, репликација и блокирање.

1.3.1. Рандомизација

Рандомизацијата е фундаменталната основа врз која почива примената на статистичките методи во експерименталниот дизајн. *Рандомизација значи случајно одредување на местото кадешто се наоѓа експерименталниот материјал и на редоследот по кој се вршат индивидуалните тестови на експериментот*. Статистичките методи бараат опсервациите (или грешките) да бидат независно распределени, случајни варијабли. Со рандомизацијата овој услов се исполнува. Со адекватна рандомизација на експериментот ќе овозможиме екстерните фактори (кои можеби се присутни) да ги доведеме на „просечна вредност“. На пример, да претпоставиме дека примероците, од претходниот експеримент за тврдина, незначително се разликуваат во дебелината и дека ефектот на медиумот за калење можеби зависи од дебелината на примероците. Ако сите примероци кои се калат во масло се подебели од тие што се калат во солена вода, можеме да внесеме систематска грешка во експерименталните резултати. Оваа грешка го хендикепира еден од медиумите и следствено ги прави нашите резултати невалидни. Случајното земање на примероци за калење во соодветниот медиум, го смалува тој можен негативен ефект.

Компјутерските софтверски програми нашироко се користат при изведувањето на планираните експерименти. Овие програми често го презентираат изведувањето на опитите по случаен редослед.



Што подразбираме под „проба“ во контекст на овој текст?

Пробата е дел од експериментот и подразбира изведување на два или повеќе тестови при еднакво ниво (вредност) на променливите, а има за цел овозможување на статистичка обработка на резултатите. На пример, притисокот на одредено количество идеален гас во затворен сад, со клип зависи од температурата на гасот и неговиот волумен. За да ја одредиме неговата вредност на пример при 65 °C и волумен од 2,5 dm^3, вршиме проба (два или повеќе тестови) при овие (непроменети) услови, и го одредуваме притисокот (одзивот) како средна вредност од сите тестови.

Случајниот (*random*) распоред се добива со генератор на случајни броеви (компјутерски) или, во недостаток на генератор, се зема од табела со случајни броеви. Дури и со вакви компјутерски програми, често пати е неопходно, сами да го одредиме распоредот, во зависност од условите на опитите, редоследот на нивното изведување (видете пример за експеримент со два фактори, во глава §3).

Понекогаш експериментаторите се среќаваат со ситуација кога, од одредени причини, рандомизацијата е тешко да се изведе. На пример, во даден хемиски процес каде што температурата е еден од факторите и има нивоа на пример: 300 °C и 400 °C тешко е да се изврши рандомизација ако редоследот бара едно ниво на факторите при 400 °C, па следното при 300 °C, па следното при 400 °C. Во овој случај целата реакциона смеса треба да ја загрееме на 400 °C, па да ја изладиме на 300 °C, па повторно да ја загрееме на 400 °C, итн. што претставува трошок и губење на време. Во вакви случаи комплетна рандомизација на факторите е непрактична. Логично е прво да се спроведе експериментот кога факторот температура е на 300 °C, менувајќи ги сите останати фактори, а потоа на 400 °C. За справување со вакви случаи постојат посебни статистички процедури кои ги земаат во предвид рестрикциите при рандомизацијата.

1.3.2. Репликација

Под репликација подразбираме независно спроведен тест за секој комбинација од факторите. Репликацијата (повторувањето) има две важни особини. Прво му овозможува на експериментаторот да направи процена за експерименталната грешка. Оваа процена на грешката станува основна мерна единица за одредување дали забележаните разлики во податоците се навистина *статистички* различни. Второ, ако сакаме да ја одредиме вистинската средна вредност на одзивот за едно ниво на факторите во експериментот, репликацијата му овозможува на експериментаторот да добие по прецизна процена на овој параметар. На пример, ако σ^2 е варијацијата од една индивидуална опсервација и ако има n реплики, варијантата на средната вредност на примерокот ќе биде,

$$\sigma_y^2 = \frac{\sigma^2}{n}$$

Практичното значење на тоа е - ако имаме n = 1 повторувања (реплики) и ако добиеме $y_1 = 145$ (калење во масло) и $y_2 = 147$ (калење во вода) нема да бидеме во можност да направиме задоволителен заклучок за влијанието на медиумот за калење, т.е. забележаните разлики можат да бидат резултат на експериментална грешка. Поентата на тоа е дека без репликација нема начин да дознаеме зошто опсервациите се различни. Од друга страна, ако n е големо и ако експерименталната грешка е доволно мала, и ако добиеме средна вредност $\bar{y}_1 < \bar{y}_2$ ќе бидеме оправдано сигурни во заклучокот дека калењето во вода дава поголема тврдина.

1.3.3. Блокирање

Блокирањето е техника која се користи за да се подобри прецизноста со која се прави споредба меѓу факторите. Често блокирањето се користи за да се елиминира или редуцира влијанието на непријатните фактори (*nuisance factors*). Непријатни се оние фактори кои можат да влијаат врз одзивот и за кои не сме директно заинтересирани. На пример, за еден експеримент во хемиски процес можно е да се потребни две шаржи на суровини за да се изведат сите опити. Меѓутоа, можно е да има варијабилност меѓу шаржите на пример ако се од различни производители. Ако посебно не сме заинтересирани за овој ефект, тогаш земаме дека шаржите на суровини се непријатен фактор. Генерално, под блок подразбираме сет на релативно хомогени експериментални услови. Во горниот пример, секоја шаража на суровини би претставувала еден блок затоа што варијабилноста во рамките на една шаржа се очекува да биде помала отколку варијабилноста меѓу различни шаржи. Ако експериментот се врши во повеќе работни смени, секоја смена ќе претставува посебен блок, итн.

Под блокирање не подразбираме буквално – спречување, туку разделување во различни блокови.

Потоа, експериментаторот опсервациите ги разделува од статистичкиот дизајн во групи според тоа на кој блок му припаѓаат.

1.4. Процедура за дизајнирање на експерименти

За да се употреби статистички пристап кон дизајнирање и анализа на некој експеримент, неопходно е сите што се вклучени во експериментот да имаат однапред јасна претстава што треба да се проучува, како податоците ќе се собираат и барем квалитативно знаење за тоа, како податоците ќе се анализираат. На сликата 1.2 даден е концизен преглед на процедурата која во продолжение ќе ја елаборираме точка по точка.

Слика: 1.2. Процедура за дизајнирање на експеримент
(првите три се подготвителни т.е. пред-експериментални фази)

1.4.1. Претстава за суштината на проблемот

Оваа можеби изгледа многу очигледна точка, но во практика често пати не е ниту едноставно да се сфати дали проблемот за кој е потребно експериментирање навистина постои, ниту е едноставно да се стекне суштинска претстава за проблемот. Неопходно е да се разгледаат сите аспекти за целите на експериментот. Многу е важно да се настојува да се добие мислење од сите релевантни страни: техничкиот сектор, секторот за обезбедување на квалитет, производниот сектор, маркетингот, менаџментот, муштериите и оперативниот персонал. Заради тоа, за дизајнирање на експерименти се препорачува тимски пристап.

Корисно е да се подготви листа со специфичните проблеми или прашања на кои треба да се обрне внимание во експериментот. Јасното разбирањето на суштината на проблемот често суштествено придонесува за подобро разбирање на феноменот кој се проучува и за негово конечно решение.

Има неколку поважни причини за изведување на експеримент, а секој вид на експеримент генерира своја листа на специфични прашања со кои треба да се справи. Некои од причините за изведување експерименти се:

а) Скрининг на факторите или карактеризација

Кога системот или процесот се нови, важно е да се знае кои фактори имаат најголемо влијание на одзивот/одзивите од интерес. Често се работи за многу фактори, што значи дека експериментаторите не знаат многу за системот, па скринингот е есенцијален ако сакаме ефикасно да ги добиеме саканите перформанси на системот. Скринингот на експериментите е екстремно важен кога се работи со нови системи или технологии. Под скрининг, во овој контекст подразбираме: раздвојување на битно од небитно, важно од неважно.

б) Оптимизација

Откако системот е карактеризиран и кога сме сигурни дека важните фактори се идентификувани, следната цел обично е оптимизација т.е. наоѓање на позицијата или нивото на важните фактори кои ќе доведат до посакуваните вредности на одзивот. На пример, ако скринингот на некој хемиски процес резултира со идентификување на времето и температурата како два најважни фактори, оптималноста на експериментот може да има за цел наоѓање на нивото на времето и температурата кои го максимизираат приносот. Оптимизација на експериментот обично следи по скринингот. Многу е необично скринингот на експериментот да продуцира оптимални позиции на важните фактори.

в) Потврдување

Со потврдувачки (*confirmation*) експеримент, експериментаторот обично настојува да потврди или да се увери дека системот функционира или се однесува, на начин кој е во склад со теоријата или минатото искуство. На пример, ако теоријата или искуството индицира дека одреден нов материјал е еднаков со тој што моментално се користи и дека новиот материјал е пожелен (заради цената; полесен за манипулација или друго поволно својство) тогаш се изведува потврдувачки експеримент за верификација дека замената со новиот материјал нема да резултира во промена на карактеристиките што ќе влијаат врз неговата апликација. Трансферот на новиот процес во индустриски услови на производство, базиран на резултатите добиени со експериментирање на пилот - постројка, е уште еден пример кога се прави потврдувачки експеримент. Со експериментот се потврдува дали истите фактори и истите нивоа кои се утврдени во текот на развојот се соодветни и за сериско производство.

г) Откривање

Во експериментите за откривање, експериментаторите обично настојуваат да утврдат што се случува кога се истражуваат нови материјали, нови фактори или нови опсези за факторите. Во фармацевтската индустрија, научниците постојано изведуваат откривачки (*discovery*) експерименти за да најдат нови материјали или комбинација на материјали кои ќе бидат ефикасни во третирање на болести.

д) Робустност

Кај овие експерименти секогаш се поставува прашањето: под кои услови одзивните варијабли од интерес сериозно деградираат? Или, кои услови ќе водат до неприфатлива варијабилност на одзивните варијабли? Постигнувањето на робустност значи поставување на контролираните фактори во системот на начин што ќе ја минимизира варијабилноста на одзивот предизвикана од неконтролираните фактори.

Очигледно е дека специфичните прашања на кои треба да даде одговор експериментот се поврзани со севкупните цели. Важен аспект при формулација на проблемот е сфаќањето дека еден голем, детален експеримент тешко може да даден задоволителен одговор на сите прашања. Единствениот, детален експеримент бара експериментаторите да го знаат одговорот на многу прашања, и ако грешат, резултатите ќе бидат разочарувачки. Тоа води кон трошење време, материјали и на други ресурси и како резултат може никогаш да не се добијат одговори на оригиналните истражувачки прашања во задоволителна форма. Секвенцијалниот (последователен) пристап, на изведување серија мали експерименти, секој со специфична цел, како што е скринингот на факторите, е подобра стратегија.

1.4.2. Избор на одзивната варијабла

При селекцијата на одзивната варијабла, експериментаторот треба да е уверен дека таа варијабла навистина дава корисни информации за процесот кој се проучува. Најчесто, просечната или стандардната девијација (или двете) на мерената карактеристика претставува одзивна варијабла. Повеќекратните одзиви не се необичност. Експериментаторите треба да одлучат како секој од одзивите ќе биде мерен; како мерниот систем ќе биде калибриран и како да се направи калибрацијата да остане непроменета во текот на експериментот.

Доверливоста во способноста на мерниот систем (или мерната грешка) е исто така важен фактор. Ако мерниот систем е неадекватен, само големите ефекти на факторот ќе бидат детектирани со експериментот. Од критичка важност е да се дефинираат одзивите од интерес и тоа како ќе бидат мерени *пред* да се изведи експериментот.

1.4.3. Избор на фактори, нивоа и опсег

Кога се разгледуваат факторите кои можат да имаат влијание врз перформансите на процесот или системот, експериментаторите откриваат дека тие фактори можат да се класифицираат како *потенцијални дизајн фактори* или како непријатни фактори. Потенцијалните дизајн фактори се оние фактори кои експериментаторот сака да ги варира во експериментот. Често се открива дека има многу потенцијални дизајн фактори и препорачливо е кај нив да се изврши понатамошна класификација пример на *дизајн фактори* и *константно-држани фактори*. Дизајн факторите се всушност факторите кои се избрани да се проучуваат во експериментот. Константно-држани фактори се варијабли кои можат да имаат извесно влијание врз одзивот, но во контекст на актуелниот експеримент, овие фактори не се од интерес и се држат константни на извесно ниво.

Непријатните фактори, од друга страна, можат да имаат голем ефект кој треба да се земе во предвид, но сепак може да не сме заинтересирани за нив во контекст на актуелниот експеримент. Непријатните фактори често се класифицираат како *контролирани, неконтролирани* или *пречки*. Контролирани, непријатни фактори се оние чие ниво може да го постави експериментаторот. На пример, експериментаторот може да избери различни шаржи или различни денови од седмицата за изведување на експериментот. Принципот на блокирање е често корисен за справување со контролираните непријатни фактори. Ако непријатниот фактор е неконтролибилен во експериментот, но може да се измери, за да се компензира неговиот ефект често се користи процедура за анализа наречена *анализа на коваријансата*. На пример, релативната влажност во амбиентот во кој се одвива процесот може да влијае врз перформансите на процесот, и ако влажноста не може да се контролира, таа веројатно може да се измери и да се третира како коваријат. Ако некој фактор, кој природно и неконтролирано варира во процесот, а може да се контролира (заради експериментот), често го нарекуваме пречка фактор (*noise factor*). Во ваков случај, наша цел е да најдеме таква поставка на контролираните дизајн фактори која ќе го минимизира дејството на пречка факторите.

Штом експериментаторот ќе ги селектира дизајн факторите, потребно е да ги избере опсезите во кои факторите ќе бидат варирани, како и

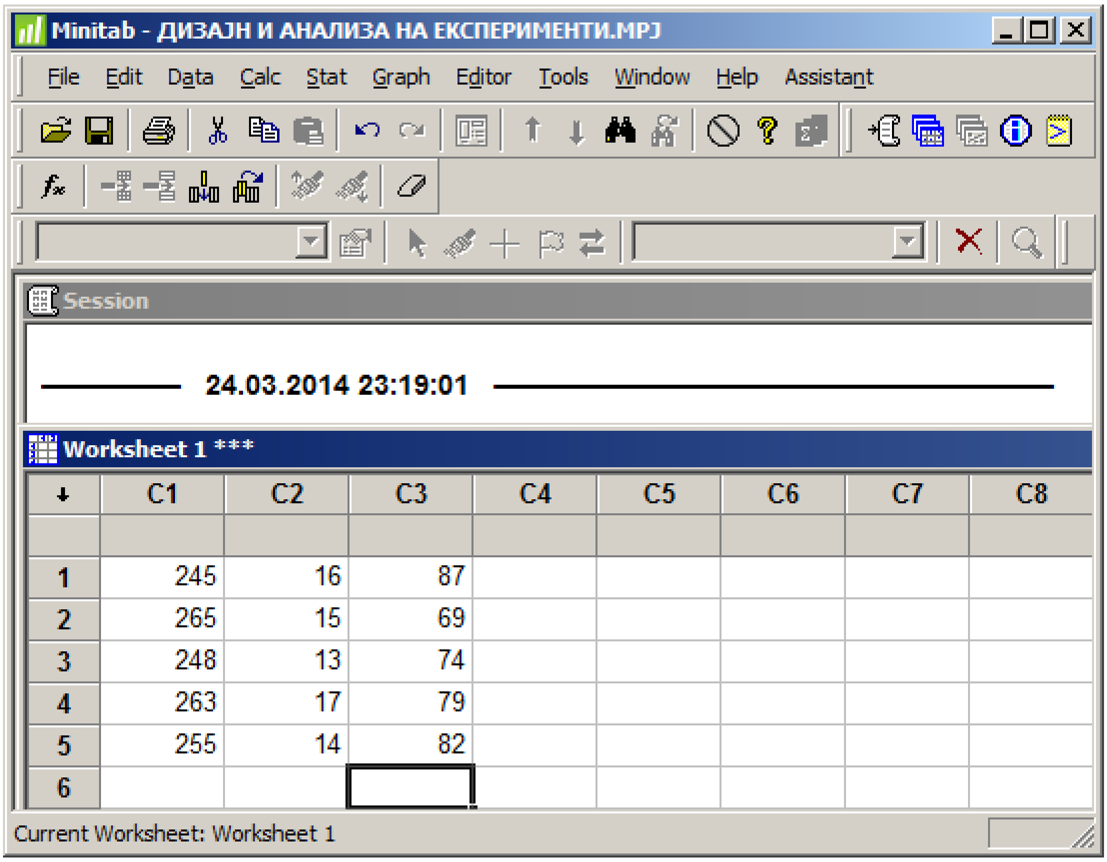
специфичните нивоа при кои ќе се прават опитите. При тоа треба да се размисли: како ќе бидат контролирани овие фактори на саканата вредност и како ќе се врши нивно мерење? Експериментаторот, исто така треба да одлучи за опсегот во кој секоја варијабла посебно ќе варира. За да се направи тоа потребно е познавање на процесот. Познавањето на процесот обично е комбинација на практично искуство и теоретско знаење.
Кога целта на експериментот е скрининг на факторите или карактеризација на процесот најдобро е, бројот на нивото на факторите да се држи низок. Генерално, две нивоа се доволни при скрининг на нивото на факторите. Изборот на регионот од интерес, е исто така важен. За скрининг на факторите, регионот од интерес треба да е релативно голем т.е. опсегот на варирање на факторите треба да е широк. Колку што повеќе учиме за тоа кои варијабли се важни и кои нивоа продуцираат најдобри резултати, регионот од интерес, при следните експерименти, обично станува потесен.

Чекорите од 1 до 3 , слика 1.2, кои претходно ги елабориравме, претставуваат пред-експериментално планирање. Тешко е да се претпостави дека еден човек ќе го има сето потребно знаење за да го направи сето тоа во сите (разни) ситуации. Затоа силно се препорачува тимска работа во подготовка на експериментот. Успехот во голем дел, зависи од тоа колку добро ќе се одработи пред-експерименталното планирањето.

1.4.4. Избор на експериментален дизајн

Ако пред-експерименталното планирање е коректно извршено, овој чекор е релативно лесен. Изборот на дизајн, вклучува размислување за големината на примерокот (број на реплики), избор на погоден редослед на опитите за експерименталните проби и утврдување дали има рестрикции за блокирањето или рандомизацијата.

Постојат неколку интерактивни статистички софтверски пакети за поддршка на оваа фаза од експерименталниот дизајн. Кај нив експериментаторот може да ги внесе податоците за бројот на факторите, нивото и опсезите и овие програми ќе презентираат листа на дизајни за избор или ќе препорачаат одреден дизајн. Повеќето софтверски пакети даваат извесни дијагностички информации за тоа, како, кој дизајн ќе функционира. Ова е корисно за евалвација, за разни дизајнерски алтернативи на експериментот. Овие програми, обично даваат табела (worksheet) со редослед на рандомизираните опити по кои треба да се изведе експериментот. Еден таков популарен комерцијален софтвер е Minitab, слика 1.3.



Слика 1.3. Интерфејс на софтверскиот пакет Minitab 14.1

Изборот на дизајн, исто така, вклучува размислување за избор на прелиминарен емпириски модел за толкување на резултатите. Моделот всушност претставува квантитативна поврзаност (равенка) меѓу одзивот и значајните дизајн фактори. Во многу случаи полиномен модел од понизок ред се зема за соодветен. Моделот од прв ред со две варијабли е

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$$

каде y е одзивот, x-ите се факторите, со β е означен непознат параметар кој ќе биде одреден од податоците на експериментот и ε е експерименталната грешка на системот кој се проучува. Моделот од прв ред понекогаш се нарекува и модел на главните ефекти. Моделите од прв ред екстензивно се користат при скринингот или карактеризацијата на експериментот. Вообичаена екстензија на моделот од прв ред е додавањето на интерактивен член, т.е.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \varepsilon$$

каде што членот x_1x_2 ја претставува интеракцијата на два фактори. Бидејќи интеракцијата на факторите е релативно честа, моделот од прв ред со интеракција многу широко се користи. Интеракција од повисок ред исто така може да биде вклучена кај експериментите со повеќе од два фактори ако е неопходно. Друг широко користен модел е моделот од втор ред,

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \varepsilon$$

Моделите од втор ред често се користат кај оптимизационите експерименти.

При изборот на дизајнот, императив е да се имаат во мисла целите на експериментот. При многу технички експерименти уште на самиот почеток знаеме дека некои нивоа на факторите ќе резултираат во различни вредности на одзивот. Консеквентно, ние сме заинтересирани да идентификуваме со тоа кои фактори ја прават таа разлика и да го процениме интензитетот на промената на одзивот.

1.4.5. Изведување на експериментот

Кога се изведува експериментот од витално значење е процесот внимателно да се посматра за да се утврди дали сѐ оди според планот. Грешки во експерименталната процедура во оваа фаза можат да го направат експериментот невалиден. Една од најчестите грешки што се јавува при некои тестови е не поставувањето на варијаблите на соодветно ниво. Затоа, пред секој тесте е потребно да се проверува поставеноста на варијаблите.

Препорачливо е пред да се изведе експериментот да се направи пробно изведување. Овие проби даваат корисни информации за конзистентноста на експерименталниот материјал, функционирањето на мерниот систем, груба идеа за експерименталната грешка и шанса да се извежба целата експериментална процедура. Тоа, исто така, дава можност да се коригираат одлуките донесени во чекорите 1 – 4 ако е неопходно.

1.4.6. Статистичка анализа на податоците

Статистички методи треба да се користат за анализа на податоците за резултатите и заклучоците да бидат објективни. Ако експериментот е добро дизајниран и изведен според планот, нема да биде тешко да се применат статистичките методи. Постојат многу одлични софтверски пакети наменети за анализа на податоци, всушност многу од програмите кои се користат во чекорот 4 вклучуваат и статистичка анализа на податоците. Често пати се уверуваме дека едноставните графички методи играат значајна улога за анализа и интерпретација на податоците. Исто така, многу е корисно да се презентираат резултатите од многу експерименти во вид на емпириски модел т.е. во вид на равенка изведена од податоците која ја изразува поврзаноста меѓу одзивот и важните дизајн фактори. Дополнителната анализа за репродуктивноста на опитите и адекватноста на моделот, исто така се важни техники за анализа. Овој начин детално ќе го дискутираме кај фактор експериментот.

Треба да се има во предвид дека статистичките методи не можат да потврдат дали некој фактор (фактори) има одреден ефект. Тие само претставуваат водич за доверливоста и валидноста на резултатите.

1.4.7. Заклучоци и препораки

Штом податоците ќе бидат анализирани, експериментаторот треба да извлече *практични* заклучоци за резултатите и да препорача соодветна акција. Графичките методи се често корисни во оваа фаза, особено, за да им се презентираат резултатите на други лица. Следно, потребно е да се спроведе валидно тестирање за да се види дали се валидни заклучоците од експериментот.

Низ целиот овој процес, важно е да се има во мисла дека експериментирањето е важен сегмент од процесот на учење, во кој ние правиме провизорни хипотези за системот, изведуваме експерименти за да ги испитаме тие хипотези, и врз база на резултатите правиме нови хипотези, итн. Тоа значи дека експериментирањето е итеративно (репетитиво). Голема грешка е да се дизајнира единичен, голем, детален експеримент на почеток од истражувањето. Успешниот експеримент бара познавање на важните фактори, опсезите низ кои овие фактори ќе варираат, соодветен број на нивоа кои ќе се користат и правилни единици за мерење на варијаблите. Генерално, совршено не го знаеме одговорот на сите овие прашања, но учиме во текот на експериментот. Како што експерименталната програма напредува, често, отфрламе некои влезни варијабли, додаваме други, го менуваме опсегот на проучување на некои фактори или додаваме нови одзивни варијабли. Консеквентно, обично експериментираме секвенцијално, и според непишаното правило, за првиот експеримент не треба да се потрошат повеќе од 25 % од предвидените средства. Тоа овозможува да останат доволно средства за да се направат потврдувачки проби и за крајно исполнување на целите од експериментот.

§2. СТАТИСТИЧКИ МЕТОДИ

Повеќето истражувања во инженерството, науката и индустријата се емпириски и екстензивно се користат при експериментирањето. Статистичките методи многу можат да ја зголемат ефикасноста на овие експерименти и често ги зајакнуваат добиените заклучоци. За правилно користење на статистичките техники во експериментирањето од вас како експериментатори се бара да ги земете следниве точки во предвид:

1. Користи го своето не статистичко познавање на проблемот

Експериментаторите, обично, се добро верзирани во своето поле. Во некои научни полиња постои големо теоретско познавање на кое треба да се повикаме за да ја објасниме релацијата меѓу факторите и одзивите. Овој вид на не статистичко знаење е драгоцено при изборот на факторите, одредување на нивното ниво, одлучувањето колку реплики на опитот да се направат, интерпретација на резултатите од анализата итн. Користењето на планираниот експеримент не е замена за познавањето на проблемот.

2. Направи ги дизајнот и анализата што е можно поедноставни

Не бидете премногу амбициозни во користењето на комплексни, софистицирани статистички техники. Релативно, едноставниот дизајн и анализа скоро секогаш се најдобри. Ако подготвителната фаза (пред-експерименталното планирање) се изврши внимателно и ако се избере разумен дизајн, анализата скоро секогаш ќе биде релативно едноставна. Всушност, добро дизајниран експеримент понекогаш скоро самиот себе ќе се анализира. Меѓутоа, ако пред-експерименталното планирање е лошо изведено и ако се изведе експерименталниот дизајн лошо, тогаш ниту нај комплексната статистика не може да помогне.

3. Препознај ја разликата меѓу практичната и статистичката значајност

Ако два експериментални услови дадат средни одзиви кои се статистички различни,тоа не мора да значи дека таа разлика е толку голема што ќе има каква било практична вредност. На пример, еден инженер може да одлучи да го замени уредот за вбризгување гориво во автомобилот со што просечно ќе помине 1 km повеќе со 10 литри бензин или 0,1 km/l, што е статистички значаен резултат. Меѓутоа, ако замената кошта 800 €, тогаш разликата 0,1 km/l е веројатно премала да има практична вредност.

4. Експериментите се итеративни

Спомнавме дека,на почетокот од проучувањето во најголем број случаи, не е препорачливо да се оди со опширен и голем експеримент. Успешниот дизајн бара знаење на важните фактори, опсези преку кои факторите ќе варираат, соодветен број на нивоа за секој фактор и соодветни методи и единици за мерење на секој фактор и одзив. Генерално, не сме доволно спремни да одговориме на сите овие барања на почетокот од експериментот, но како што напредуваме со експериментирањето ние учиме. Тоа е образложението во полза на итеративниот (повторливиот) или секвенцијалниот пристап, за кој дискутиравме претходно. Се разбира, има ситуации кога големиот, опширен експеримент е сосема прикладен, но генерално, повеќето експерименти се итеративни. Консеквентно, не треба да инвестираме повеќе од 25 % од средствата за почетниот експеримент. Често, овие почетни напори се искуство на стекнување знаење и потребно е да оставиме средства за исполнување на крајните цели на експериментот.

Во продолжение, од два технички примери, ќе видиме како практично се користат статистичките методи за анализа на податоците.

2.1. Статистичка анализа на податоци

2.1.1. Вовед

При тестовите кои се вршат на одредени карактеристики на материјалите често пати се случува да има прилично расејување на резултатите. Во таков случај потребно е да се изврши статистичка анализа на резултатите за тие да бидат поразбирливи и покорисни. Двата чекори кои ќе ги користиме при таа анализа го покриваат дефинирањето и справувањето со податоците што отскокнуваат (*outliers*) во однос на другите и статистичката анализа на доверливоста во резултатите од тестот.

Прво, податоците што отскокнуваат треба да се елиминираат.

Податоци што отскокнуваат се оние, кои значително се разликуваат од останатите или можеме да кажеме дека, на некој начин, се екстремни вредности во однос на останатите вредности од истиот тест.

Вредноста што отскокнува може да се јави како резултат на експериментален или проблем при пресметувањето. Штом ќе се идентификува вредност што отскокнува потребно е да се проверат постапката на тестирање и постапката на пресметување. Ако кај нив се најде грешка, тогаш податокот треба да се коригира и може понатаму да се вклучи во преостанатите пресметки. Во спротивно, истиот треба да се елиминира од понатамошните пресметки.

Штом податоците што отскокнуваат ќе бидат коригирани или елиминирани, може да се изврши едноставна статистичка анализа на преостанатите податоци. Цел на статистичката анализа е да се определи средната или просечната вредност на резултатите од тестот т.е. 95 % -от интервал на доверба за средната вредност и опсегот во кој сите останати податоци од тестот ќе се најдат. Овие пресметки му покажуваат на експериментаторот колку може да има доверба во пресметаната средна вредност и му даваат добра процена за опсегот на расејување на сите резултати од тестот.

Со цел да се добие разумна доверба во резултатите од тестот, и да се користат методите кои се објаснети во продолжение, погребно е првобитно на располагање да се имаат минимум пет податоци. Бројот на податоци може да се редуцира на четири, со отфрлање на податокот што отскокнува, и сепак да се искористат во пресметките на статистичка анализа.Помалку од четири вредности не треба да се користат за статистичка анализа. Ако се бара поголема доверба во резултатите од тестот, пожелно е да се имаат десет или повеќе податоци.
Во продолжение ќе ја проследиме постапката чекор по чекор користејќи ги следниве симболи:

2.2. Листа на симболи

Симбол	Значење
x_i	Индивидуална вредност т.е. податок
$\overline{x}$	Средна (просечна) вредност од индивидуалните вредности т.е. податоци
$x_{\min}$	Минимална вредност
$x_{\max}$	Максимална вредност
s	Стандардна девијација
$T_{S\min}$	Тест критериум за минималната вредност
$T_{S\max}$	Тест критериум за максималната вредност

$T_{5\%}$	Максимално дозволена вредност на тест критериумот за 5% ниво на значајност.
N, n	Број на податоци
f	Број на степени на слобода (f = N – 1).
$t_{0.05}$	Вредност на Студентова t – дистрибуција за 95% ниво на доверба (се зема од табела).
95% C.I.	95% ниво на доверба за средната вредност.
95% D.R.	Опсег во кој 95% од сите податоци, досегашни и идни, ќе се најдат

2.3. Справување со податоци што отскокнуваат

Штом сите резултати од тестот (минимум 5) ќе се соберат и ќе се стават во табела во конечните инженериски единици,може да се спроведе следна анализа. Таа анализа најдобро се прави на лист во форма како што е прикажано на табелата - 1 во продолжение.

Прво се врши проценка т.е. одредување на средната вредност и проценка на стандардната девијација, кои се одредуваат на следниов начин:

- ***Средна вредност:***

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

- ***Стандардна девијација на примерокот:***

$$s = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}}$$

каде n = број на податоци (резултати), додека x_i претставува вредност на секој индивидуален податок.

Табела 1 . Табела за стстистичка анализа на композитен материјал

Карактеристика	Темпер.	Ориентација	N	$\bar{x}$	S	Минимум	Максимум	95% C.I.	95% D.R.
Јакост	Собна	Надолжна	28	289	44,4	183	359	±17	±91
Јакост	150 °C	Попречна	5	43,3	5,16	34,5	47,9	Минималната е премала	
Јакост	150 °C	Попречна	4	45,5	1,78	44,1	47,9	±2,8	±5,6

Потребно е да се пресметаат минималната и максималната вредност на податоците за да се утврди дали некој од нив отскокнува. *Податоци што отскокнуваат* ги нарекуваме оние кои имаат екстремни вредности т.е. оние кои се многу помали или поголеми од останатите. Тоа се прави со нивно споредување со средната вредност во однос на стандардната девијација користејќи го тест критериумот T_S:

$$T_{S_{\max}} = \frac{(x_{\max} - \bar{x})}{s}$$

$$T_{S_{\min}} = \frac{(\bar{x} - x_{\min})}{s}$$

Тест критериумот треба да биде доволно мал за да индицира дека минималната и максималната вредност се наоѓаат во опсегот во кој се очекува да бидат 95 % од сите вредности, а се пресметува преку стандардната девијација и бројот на опсервации (податоци). Максималните вредности на тест критериумот, како функција на бројот на податоците (или бројот на опсервациите т.е. бројот на тестови) се прикажани во табелата 3 (*Критични вредности за t (едностран тест) кога стандардната девијација се пресметува од ист примерок)* која е преземена од стандардот ASTM E178 – *Standard Practice for Dealing With Outlying Observations*.

Колоната, горното 5 % ниво на значајност треба да се земе за определување на максимално дозволената вредност за тест критериумот т.е. $T_{5\%}$. Ако тест критериумот е поголем од $T_{5\%}$, податокот спаѓа во оние што отскокнуваат и треба да биде коригиран или елиминиран, т.е.

Ако $T_{S_{min}} > T_{5\%}$, x_{min} е податок што отскокнува

Ако $T_{S_{max}} > T_{5\%}$, x_{max} е податок што отскокнува

Ако се покаже дека максималната или минималната или двете вредности отскокнуваат, тогаш целата, погоре опишана, процедура мора да се повтори со коригирани податоци или со елиминација на лошите податоци. Потребно е да се пред калкулираат $\bar{x}$, s, $T_{S_{min}}$ и $T_{S_{max}}$ и да се одреди дали $T_{S_{min}}$ и $T_{S_{max}}$ се помали од максимално дозволените вредности според тест критериумот. Ако се помали, се продолжува со постапката понатаму, додека ако не се, горната процедура треба да се повторува се додека не се добијат помали вредности. Во тој случај ќе ги имаме вредностите за N (бројот на податоците т.е. мерењата), Табела 2. Критични вредности за Т

TABLE 2 Critical Values for T (One-Sided Test) When Standard Deviation Is Calculated from the Same Sample[A]

Number of Observations, n	Upper 0.1 % Significance Level	Upper 0.5 % Significance Level	Upper 1 % Significance Level	Upper 2.5 % Significance Level	Upper 5 % Significance Level	Upper 10 % Significance Level
3	1.155	1.155	1.155	1.155	1.153	1.148
4	1.499	1.496	1.492	1.481	1.463	1.425
5	1.780	1.764	1.749	1.715	1.672	1.602
6	2.011	1.973	1.944	1.887	1.822	1.729
7	2.201	2.139	2.097	2.020	1.938	1.828
8	2.358	2.274	2.221	2.126	2.032	1.909
9	2.492	2.387	2.323	2.215	2.110	1.977
10	2.606	2.482	2.410	2.290	2.176	2.036
11	2.705	2.564	2.485	2.355	2.234	2.088
12	2.791	2.636	2.550	2.412	2.285	2.134
13	2.867	2.699	2.607	2.462	2.331	2.175
14	2.935	2.755	2.659	2.507	2.371	2.213
15	2.997	2.806	2.705	2.549	2.409	2.247
16	3.052	2.852	2.747	2.585	2.443	2.279
17	3.103	2.894	2.785	2.620	2.475	2.309
18	3.149	2.932	2.821	2.651	2.504	2.335
19	3.191	2.968	2.854	2.681	2.532	2.361
20	3.230	3.001	2.884	2.709	2.557	2.385
21	3.266	3.031	2.912	2.733	2.580	2.408
22	3.300	3.060	2.939	2.758	2.603	2.429
23	3.332	3.087	2.963	2.781	2.624	2.448
24	3.362	3.112	2.987	2.802	2.644	2.467
25	3.389	3.135	3.009	2.822	2.663	2.486
26	3.415	3.157	3.029	2.841	2.681	2.502
27	3.440	3.178	3.049	2.859	2.698	2.519
28	3.464	3.199	3.068	2.876	2.714	2.534
29	3.486	3.218	3.085	2.893	2.730	2.549
30	3.507	3.236	3.103	2.908	2.745	2.563

$\overline{x}$ (средната вредност на податоците), и s (стандардна девијација на примерокот) како податоци за тестот со исклучени сите лоши податоци од понатамошно разгледување.

2.4. Статистичка анализа

Интервалот на доверба ($C.I.$ или *Confidance Interval*) од 95 % за средната вредност е одличен индикатор за довербата во резултатите од тестот. Тој ни го покажува опсегот во кој со 95 % сигурност можеме да кажеме дека вистинската средна вредност на карактеристиката од материјалот се наоѓа. Тоа значи дека вистинската вредност на карактеристиката од материјалот има шанса од 19:1 да се најде во тој интервал [(19/20) x 100 = 95 %]. Го пресметуваме според релацијата:

$$95\% C.I = \overline{x} \pm t_{0,05} \cdot \frac{s}{\sqrt{N}}$$

Каде $t_{0,05}$ се наоѓа во α = 0,05 колоната во табелата 4.1 како функција од степените на слобода ($f = N - 1$). Ова е Студентската (Student) t дистрибуција за мали примероци. Останатите променливи од горната формула $\overline{x}$, s, и N веќе претходно беа дефинирани.

95% интервал на доверба може да се изрази на следниве начини:

$$95\%\ C.I. = \pm \text{ инженериски единици, или}$$

$$95\%\ C.I. = \pm\ \%\text{ од средната вредност}$$

Начинот на изразување во инженерски единици е корисен за графичко претставување, додека начинот на изразување во % од средната вредност - за писмена или вербална комуникација со резултатите од тестот.

Опсегот во кој 95% од сите резултати (95% *Data Range*) на тестот (сегашни и идни) ќе се најдат, претставува добра индикација за расејувањето на одредена карактеристика на соодветниот материјал. Шансите се 1 : 20 дека некој резултат ќе се најде надвор од овој опсег:

$$95\%\ Data\ Range = \overline{x} \pm t_{0,05} \cdot s$$

Вредностите за $\overline{x}$, $t_{0,05}$ и s, веќе беа претходно одредени.

Табела 3 .Студентова t дистрибуција – вредности за t_α

f	0,1	0,05	0,025	0,01	0,005	0,0025	0,001	0,0005
1	3,077685	6,313749	12,70615	31,82096	63,6559	127,3211	318,2888	636,5776
2	1,885619	2,919987	4,302656	6,964547	9,924988	14,08916	22,32846	31,59977
3	1,637745	2,353363	3,182449	4,540707	5,840848	7,4532	10,21428	12,92443
4	1,533206	2,131846	2,776451	3,746936	4,60408	5,59754	7,17293	8,610077
5	1,475885	2,015049	2,570578	3,36493	4,032117	4,773319	5,893526	6,868504
6	1,439755	1,943181	2,446914	3,142668	3,707428	4,316826	5,207548	5,958718
7	1,414924	1,894578	2,364623	2,997949	3,499481	4,029353	4,785252	5,408074
8	1,396816	1,859548	2,306006	2,896468	3,355381	3,832538	4,500762	5,041366
9	1,383029	1,833114	2,262159	2,821434	3,249843	3,689638	4,29689	4,780886
10	1,372184	1,812462	2,228139	2,763772	3,169262	3,581372	4,143658	4,586764
11	1,36343	1,795884	2,200986	2,718079	3,105815	3,496607	4,024769	4,436879
12	1,356218	1,782287	2,178813	2,68099	3,054538	3,428431	3,929599	4,317844
13	1,350172	1,770932	2,160368	2,650304	3,012283	3,372479	3,852037	4,220929
14	1,345031	1,761309	2,144789	2,624492	2,976849	3,325695	3,787427	4,140311
15	1,340605	1,753051	2,131451	2,602483	2,946726	3,286041	3,732857	4,07279
16	1,336757	1,745884	2,119905	2,583492	2,920788	3,251989	3,686146	4,014873
17	1,333379	1,739606	2,109819	2,56694	2,898232	3,222449	3,645764	3,965106
18	1,330391	1,734063	2,100924	2,552379	2,878442	3,196583	3,610476	3,921741
19	1,327728	1,729131	2,093025	2,539482	2,860943	3,1737	3,579335	3,883324
20	1,325341	1,724718	2,085962	2,527977	2,845336	3,1534	3,551831	3,849564
21	1,323187	1,720744	2,079614	2,517645	2,831366	3,13521	3,527093	3,819296
22	1,321237	1,717144	2,073875	2,508323	2,818761	3,118839	3,504974	3,792229
23	1,319461	1,71387	2,068655	2,499874	2,807337	3,103996	3,484965	3,767636
24	1,317835	1,710882	2,063898	2,492161	2,796951	3,090536	3,466776	3,745372
25	1,316346	1,70814	2,059537	2,485103	2,787438	3,078203	3,450186	3,725145
26	1,314972	1,705616	2,055531	2,478628	2,778725	3,066889	3,43498	3,706664
27	1,313704	1,703288	2,051829	2,472661	2,770685	3,056521	3,42101	3,689493
28	1,311435	1,699127	2,045231	2,46202	2,756387	3,03804	3,396271	3,659516
29	1,311435	1,699127	2,045231	2,46202	2,756387	3,03804	3,396271	3,659516
30	1,310416	1,69726	2,04227	2,457264	2,749985	3,029782	3,385212	3,645982

95% опсег на податоци ($D.R$ = *Data Range*) може да се изрази во инженериски единици или како % од средната вредност:

$$95\%\ D.R. = \pm \text{ инженериски единици}$$
$$95\%\ D.R. = \pm \%\text{ од средната вредност}$$

Скратена содржина на горните процедури е дадена во продолжение, а примери се дадени на крајот од ова поглавје.

За полесно и побрзо спроведување, оваа постапка може да биде програмирана на компјутер или на програмиран калкулатор.

2.5. Скратена содржина на процедурите за статистичка анализа

Во продолжение, скратено, чекор по чекор, се дадени процедурите за статистичка анализа на податоците.

Чекори:

1. Сведуваме најмалку пет податоци од испитувањето на материјалот во нивните инженериски единици. Овие податоци се означуваат x_i (i = 1, 2, 3, …n).

2. Ја пресметуваме средната вредност:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

3. Ја пресметуваме стандардната девијација:

$$s = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}}$$

4. Го пресметуваме тест - критериумот за минималната и максималната вредност на податоците.

$$T_{S_{\max}} = \frac{(x_{\max} - \bar{x})}{s}$$

$$T_{S_{\min}} = \frac{(\bar{x} - x_{\min})}{s}$$

5. Го споредуваме тест - критериумот со максимално дозволениот тест - критериум земен од табела 3 ($T_{5\%}$). Оваа вредност треба да се земе од колоната на горното 5% ниво на значајност (*upper 5% level of significance*) во зависност од бројот на податоците што се обработуваат (**n**). Ако двете вредности за T_s се помали од $T_{5\%}$ продолжуваме кон чекор 7. Ако која било од овие вредности е поголема од онаа во табелата, продолжуваме кон чекор 6, т.е.

Ако $(T_{S_{min}}$ и $T_{S_{max}}) < T_{5\%}$, одиме на чекор 7

Ако $(T_{S_{min}}$ и $T_{S_{max}}) > T_{5\%}$, одиме на чекор 6

6. Ги коригираме погрешните податоци или ги елиминираме од пресметките. Се враќаме на чекор 2 или со коригирана вредност x_i или со нова вредност на n која ќе ја добиеме со елиминација на еден или два податоци.

7. Сега ги имаме вредности за N, $\bar{x}$, и s. Го пресметуваме 95% интервал на доверба.

$$95\%C.I = \pm t_{0,05} \cdot \frac{s}{\sqrt{N}}$$

и/или

$$95\%C.I = \pm \frac{t_{0,05} \cdot \frac{s}{\sqrt{N}}}{\bar{x}} \cdot 100\%$$

Каде $t_{0,05}$ се наоѓа во α = 0,05 колоната во табелата 4.1 како функција од степените на слобода (f = N - 1).

8. Го пресметуваме 95% опсег на податоци.

$$95\%\ Data\ Range = \pm\, t_{0,05} \cdot s$$

и/или

$$95\%D.R = \pm \frac{t_{0,05} \cdot s}{\bar{x}} \cdot 100\%$$

2.6. Примери

2.6.1. Пример 1

Јакост на затегање на слоевит композит врз база на стаклени влакна/синтетичка смола при собна температура

Следниве податоци се добиени од 28 проби за јакоста на затегање на стакло/синтетичка смола композит во лонгитудинална насока при собна температура

Чекор 1. Сведуваме 28 податоци во инженериски единици (треба најмалку пет), табела 4.

Табела 4. Јакост на затегање на слоевит композит, (MPa)

Тест	Јакост	Тест	Јакост
1	220	15	321
2	316	16	257
3	325	17	239
4	292	18	272
5	292	19	276
6	358	20	259
7	297	21	248
8	354	22	359
9	303	23	285
10	313	24	183
11	326	25	222
12	345	26	283
13	339	27	287
14	253	28	278

Чекор 2. Ја пресметуваме средната вредност од овие 28 податоци тргнувајќи од следнава релација:

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Збирот на сите 28 податоци изнесува:

$$\sum_{i=1}^{28} x_i = 8102$$

Од тука средната аритметичка вредност е:

$$\bar{x} = \frac{\sum_{i=1}^{28} x_i}{n} = \frac{8102}{28} = 289{,}4MPa$$

Чекор 3. Ја пресметуваме стандардната девијација на овие 28 податоци тргнувајќи од следнава релација:

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$$

За полесно да ја пресметаме стандардната девијација, бидејќи поединечните податоци се цели броеви, ќе земеме, без да правиме голема грешка, заокружено, дека средната вредност е цел број т.е. 289 MPa. Сега од секоја поединечна вредност ја вадиме средната вредност и разликата ја дигаме на квадрат, како што следи:

$$\sum_{i=1}^{28}(x_i - \bar{x})^2 =$$

$$(220-289)^2+(316-289)^2+(325-289)^2+(292-289)^2+(292-289)^2+(358-289)^2+$$
$$(297-289)^2+(354-289)^2+(303-289)^2+(313-289)^2+(326-289)^2+(345-289)^2+$$
$$(339-289)^2+(253-289)^2+(321-289)^2+(257-289)^2+(239-289)^2+(272-289)^2+$$
$$(276-289)^2+(259-289)^2+(248-289)^2+(359-289)^2+(285-289)^2+(183-289)^2+$$
$$(222-289)^2+(283-289)^2+(287-289)^2+(278-289)^2$$

$$=$$

$$(-69)^2+27^2+36^2+3^2+3^2+69^2+8^2+65^2+14^2+24^2+37^2+56^2+50^2+(-36)^2$$
$$+32^2+(-32)^2+(-50)^2+(-17)^2+(-13)^2+(-30)^2+(-41)^2+70^2+(-4)^2+(-106)^2$$
$$+(-67)^2+(-6)^2+(-2)^2+(-11)^2$$

$$=$$

$$4761+729+1296+9+9+4761+64+4225+196+576+1369+3136+2500+1296+1024+$$
$$1024+2500+289+169+900+1681+4900+16+11236+4489+36+4+121= 53316$$

или

$$\sum_{1}^{28}(x_i - 289)^2 = 53316$$

ако замениме во равенката ќе добиеме:

$$s = \sqrt{\frac{53316}{28-1}} = \sqrt{\frac{53316}{27}} = \sqrt{1974{,}6} = 44{,}4MPa$$

Чекор 4. Ако внимателно ја погледнеме табелата 6.1.1 ќе заклучиме дека податоци со минимална и максимална вредност се 183 и 359, респективно.

$$T_{S_{\max}} = \frac{(x_{\max} - \bar{x})}{s}$$

$$T_{S_{\max}} = \frac{(359 - 289{,}4)MPa}{44{,}4MPa} = 1{,}57$$

$$T_{S_{\min}} = \frac{(\bar{x} - x_{\min})}{s}$$

$$T_{S_{\min}} = \frac{(289{,}4 - 183)MPa}{44{,}4MPa} = 2{,}40$$

Чекор 5. Од табелата 3.1, во пресекот на колоната, горно 5% ниво на значајност и редот n = 28 наоѓаме дека максималната вредност на $T_S = 2{,}714$. Двете наши T_S вредности се помали од $2{,}714$ што значи дека нема податоци што отскокнуваат, па продолжуваме кон чекор 7.

Чекор 7

$$95\%C.I = \pm t_{0,05} \cdot \frac{s}{\sqrt{N}}$$

$$95\%C.I = \pm 1{,}703 \cdot \frac{44{,}4MPa}{\sqrt{28}} = \pm 14{,}3MPa$$

или

$$95\%C.I = \pm \frac{14{,}3MPa}{289MPa} \cdot 100\% = \pm 4{,}9\%$$

Вредноста 1,703 (вредноста 1,703288 е заокружена на три децимали) ја зедовме од табелата 4.1, а се наоѓа во пресекот на 27- иот $(f = 28 - 1)$ ред и колоната α = 0,05.

Чекор 8

$$95\%\ Data\ Range = \pm\, t_{0,05} \cdot s$$

$$95\%\ Data\ Range = \pm\, 1{,}703 \cdot 44{,}4MPa = \pm\, 75{,}6\ MPa \quad \textit{или}$$

$$95\%D.R = \pm \frac{75{,}6MPa}{289MPa} \cdot 100\% = \pm 26{,}1\%$$

Значи, најдовме дека просечната јакост на затегање е 289 MPa со 95% интервал на доверба од ± 14,3 MPa и ± 75,6 MPa 95% опсег на податоците.

95% интервал на доверба покажува дека средната вредност се наоѓа во опсегот 289 ± 14,3 MPa со веројатност од 20 :1, или можеме да кажеме дека 95% интервал на доверба во овој случај е ± 4,9 %.

95% опсег на податоци се користи за да го покаже расејувањето на податоците од тестот. Расејувањето може да биде изразено во графичка или во писмена форма. Во овој случај 95 % од сите резултати на тестот треба да се најдат во опсегот 289 ± 75,6 MPa или 95% опсег на податоци е ± 26,1 % од средната вредност.

2.6.2. Пример 2

Јакост на затегање на слоевит композит врз база на стаклени влакна/синтетичка смола при 150 °C

Следниве податоци се добиени од 5 тестови за јакоста на затегање на стакло/синтетичка смола, композит во трансверзална насока при 150 °C.

Чекор 1.

Табела 5. Јакост на затегање, (MPa)

Тест	Јакост
1	47,9
2	44,2
3	45,8
4	34,5
5	44,1

Чекор 2

$$\bar{x} = \frac{47{,}9 + 44{,}2 + 45{,}8 + 34{,}5 + 44{,}1}{5} = \frac{216{,}5}{5} = 43{,}3MPa$$

Чекор 3

$$s = \sqrt{\frac{\sum (47{,}9 - 43{,}3)^2 + (44{,}2 - 43{,}3)^2 + (45{,}8 - 43{,}3)^2 + (34{,}5 - 43{,}3)^2 + (44{,}1 - 43{,}3)^2}{5 - 1}}$$

$$s = 5{,}16\ MPa$$

Чекор 4

$$T_{S_{\min}} = \frac{(43{,}3 - 34{,}5)MPa}{5{,}16MPa} = 1{,}71$$

$$T_{S_{\max}} = \frac{(47{,}9 - 43{,}3)MPa}{5.16MPa} = 0{,}89$$

Чекор 5

Од табелата 3 гледаме дека максималната вредност за T_s е 1,672 (за 5 мерења и горно 5 % ниво на значајност). Бидејќи $T_{S_{min}}$ има повисока вредност од оваа, минималната вредност треба да биде коригирана или елиминирана.

Чекор 6

Четвртиот податок од тестот ќе биде елиминиран од понатамошните пресметки. Сега, новиот број на опсервации е n = 4.

Чекор 2

Новата средна вредност е:

$$\bar{x} = \frac{47{,}9 + 44{,}2 + 45{,}8 + 44{,}1}{4} = \frac{182}{4} = 45{,}5MPa$$

Чекор 3

Новата стандардна девијација е:

$$s = \sqrt{\frac{\sum (47{,}9 - 43{,}3)^2 + (44{,}2 - 43{,}3)^2 + (45{,}8 - 43{,}3)^2 + (44{,}1 - 43{,}3)^2}{4 - 1}} = 1{,}78MPa$$

Чекор 4

Новиот тест критериум за минимална и максимална вредност е:

$$T_{S_{\min}} = \frac{(45{,}5 - 44{,}1)MPa}{1{,}78MPa} = 0{,}79$$

$$T_{S_{\max}} = \frac{(47{,}9 - 45{,}5)MPa}{1.78MPa} = 1{,}35$$

Чекор 5

Од табелата 3 гледаме дека максималната вредност за T_s е 1,463 (за 4 мерења и горно 5 % ниво на значајност). Бидејќи двете T_S вредности се помали одиме на чекор 7.

Чекор 7

$$95\%C.I = \pm 2{,}353 \cdot \frac{1{,}78}{\sqrt{4}} = \pm 2{,}09MPa$$

или

$$95\%C.I = \pm \frac{2{,}09}{45{,}5} \cdot 100 = \pm 4{,}6\%$$

Чекор 8

$$95\%\ Data\ Range = \pm\ 2{,}353\ \cdot 1{,}78 = \pm\ 4{,}2\ MPa$$

или

$$95\%D.R = \pm \frac{4{,}2MPa}{45{,}5MPa} \cdot 100\% = \pm 9{,}2\%$$

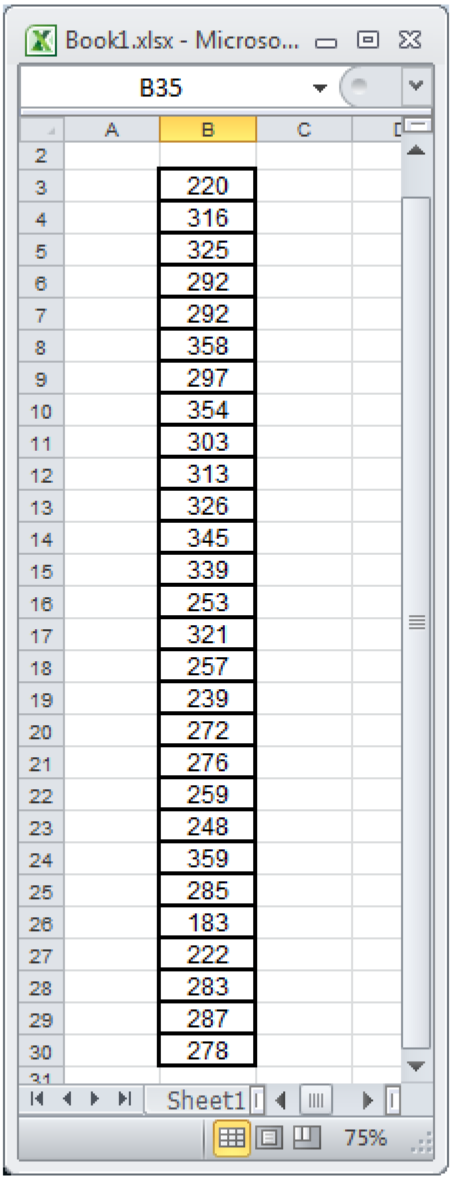
Најдовме дека средната јакост на затегање е 45,5 MPa со 95% интервал на доверба на средната вредност од ±2,09 MPa или ±4,6 %. 95%-иот опсег на податоците покажува дека расејувањето на податоците е ±4,2 MPa од средната вредност или во проценти, ±9,2%.

2.7. Користење на Microsoft Excel за статистичка анализа

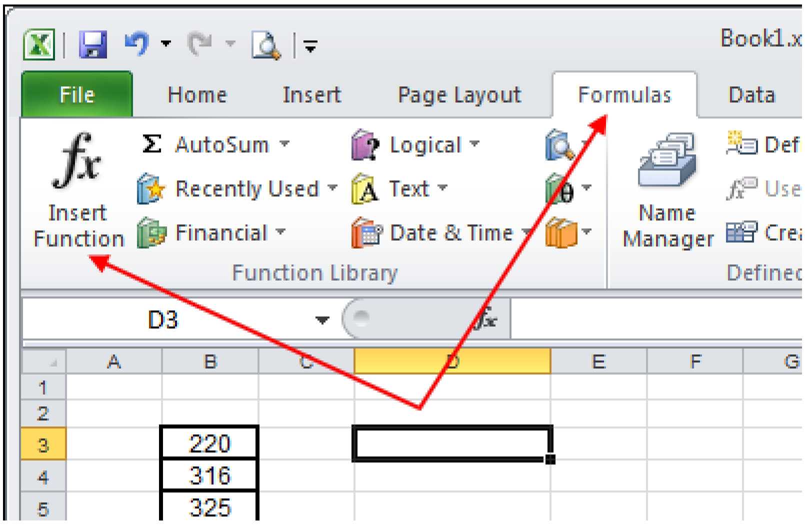
2.7.1. Функција COUNT

Без да користиме специјална компјутерска програма или програмиран калкулатор, за да ги извршиме процедурите објаснети во претходните поглавја, многу можеме да си ја олесниме работата ако ја користиме *Microsoft Excel 2010* програмата која е дел од *Microsoft Office*.

Прво што треба да направиме откако сме ја отвориле програмата е да ги внесеме податоците. Нив можеме да ги внесеме било во колона било во ред, се едно е за постапката на обработка на резултатите. Воопшто не е важно од која ќелија сме почнале да ги внесуваме податоците, но при тоа е многу важно тие да бидат во соседни ќелии. На сликата 2.1 внесени се податоците од примерот во колона почнувајќи од ќелијата B3. Тоа значи дека почетната ќелија се наоѓа во пресекот меѓу колоната B и третиот ред.За да ја пресметаме средната вредност од сите резултати на тестот покрај поединечната вредност на секој податок треба да знаеме колку вкупно податоци имаме. Во конкретниов пример имаме 28 податоци и релативно лесно можеме да ги изброиме, но може да се случи да имаме и неколку стотици податоци и во тој случај, со броење, многу лесно можеме да направиме грешка. Ако податоците сме ги внесле во ќелијата A1 нивниот број можеме лесно да го определиме погледнувајќи го бројот на редот во кој се наоѓа последниот податок. Но во практика, најчесто, пред листата со податоци ставаме наслов или текстуално појаснување па ретко се случува да ставаме податоци за обработка во првата ќелијата A1. Од друга страна ако податоците ги ставиме во ред (место во колона) оваа предност не ќе можеме да ја искористиме. Затоа ќе ја употребиме, веќе во *Excel*, вградената функцијата ***COUNT*** *(изброи)*.

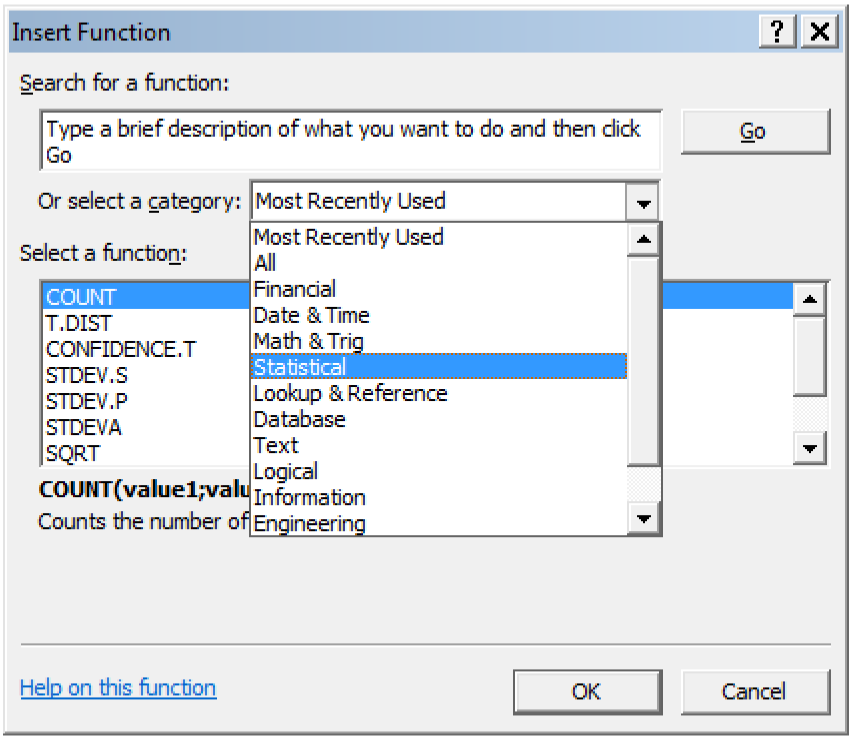


Слика 2.1 Резултати од испитувањето внесени во колона

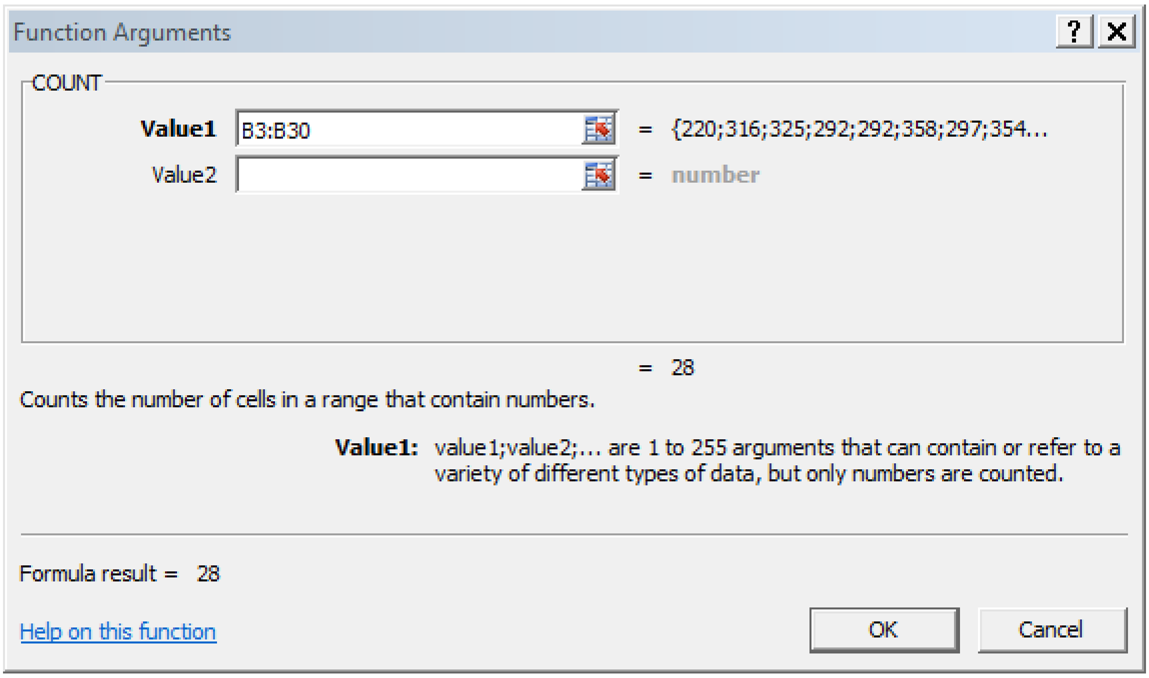


Слика 2.2. Инсертирање на функција во ќелија

Прво ја избираме и селектираме ќелијата во која што сакаме да ни се прокаже резултатот (во овој случај тоа е D3) потоа („⇒“ овој знак ќе го користиме во иднина место терминот „потоа“) во менито го избираме табот *Formulas ⇒Insert Function*, слика 2.2. Во дијалог боксот кој ќе се појави, во паѓачкото мени *Category*, избираме *Statistical* ⇒ *Count*, како што е прикажано на сликата 2.3.



Слика 2.3. Инсертирање на функцијата Count



Слика 2.4. Внесување на аргументи за функцијата Count

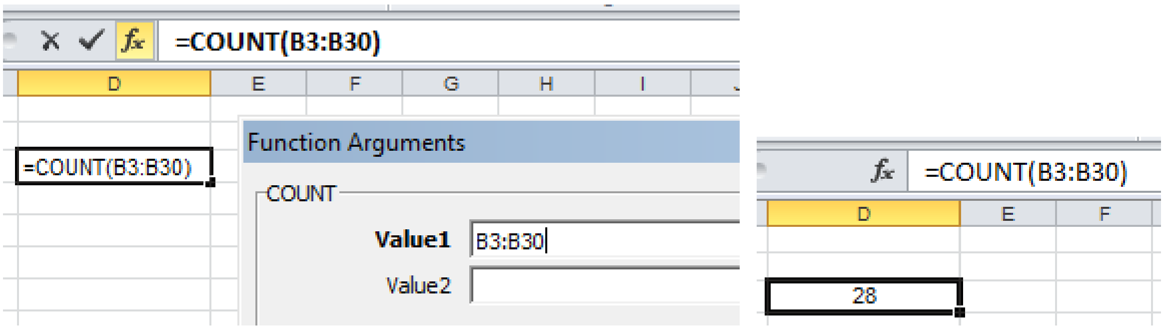
Ни се појавува бокс како на сликата 2.4 каде во полето **Value 1** го внесуваме опсегот на ќелии во кои ни се наоѓаат податоците. Во нашиов случај тоа се од B3 до B30 (видете слика 2.1). Тоа го правиме на тој начин што прво ја пишуваме почетната ќелија (B3) проследена со две точки (B3:) и потоа крајната ќелија (B3:B30).

Наместо да ги пишувате аргументите во полето **Value 1** *по појавата на боксот, слика 2.4, селектирајте ги сите ќелии и опсегот автоматски ќе се внесе. Ова е посигурен начин од мануелното внесување затоа што се избегнуваат грешки во синтаксата.*

Програмата *Excel* е многу паметна програма, дури и пред да кликнеме на копчето *OK* за да ги потврдиме внесените аргументи, веќе ни го дава резултатот = 28, слика 2.4. Во ќелијата D3 можеме да ја видеме инсертираната функција, како и во лентата за функции (десно од иконата f_x), слика 2.5а. Синтаксата на инсертираната функција е:

$$f_x = \text{COUNT(B3:B30)}$$

Оваа синтакса значи „изброи колку податоци има од ќелијата B3 до ќелијата B30“. По кликањето на копчето *OK*, *Excel* ни го дава резултатот во ќелијата D3, слика 2.5б.



(а) (б)

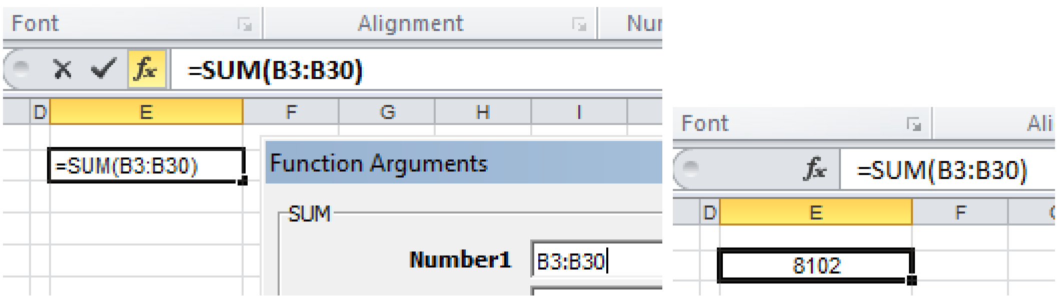
Слика 2.5. Внесување на аргументи за функцијата Count

Функција можете да инсертирате и ако кликнете на иконата f_x прикажана на сликата 2.5а.

Можеме да заклучиме дека со помош на функцијата COUNT го определивме бројот (n) на податоци кои ги имаме за обработка.

2.7.2. Функција SUM

Со функцијата ***SUM*** (собери) вршиме собирање на нумерички податоци од ќелиите. Неа ја користиме на следниов начин: прво ја селектираме ќелијата во која што сакаме резултатот да ни се прикаже (пр. E2) ⇒ f_x ⇒ категорија *Math & Trig* ⇒ *SUM* ⇒ ги селектираме сите ќелии од B3 до B30 ⇒ *Excel* во полето **Number 1** го внесува опсегот на податоци, а во ќелијата E2 функцијата, слика 2.6а ⇒ *OK* ⇒ *Excel* го внесува резултатот во ќелијата E2 (8102), слика 2.6б.



(а) (б)

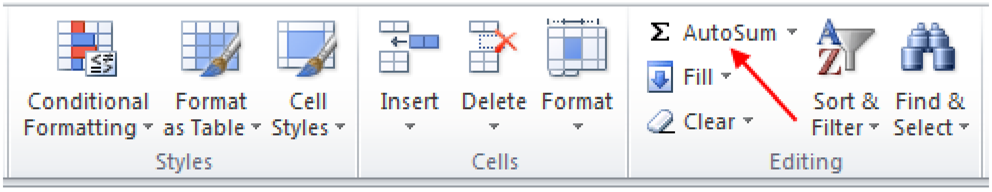
Слика 2.6. Пресметување на збирот на сите податоци

Синтаксата на функцијата SUM е:

$$f_x = \mathrm{SUM(B3:B30)}$$

Оваа синтакса значи „собери ги сите податоци почнувајќи од ќелијата B3 па се до ќелијата B30“.

Полесен начин за собирање на броеви во Excel е ако при селектиран Home таб во менито ја избереме иконата **ΣAutoSum** *која се наоѓа на крајот од рибонот во групата Editing, слика 2.7.*



Слика 2.7.. Иконата ΣAutoSum се наоѓа во групата Editing

Сега веќе го имаме вкупниот број на податоци, како и нивната сума, па можеме да ја пресметаме средната вредност, но пред тоа, да се запознаеме подетално со функциите и формулите во *Excel.*

2.7.3. Формули и функции

Многу корисници на *Excel* го користат термините „формула“ и „функција“ како синоними, но тие не се тоа. Овие термините немаат исто значење, туку постои дистинктивна разлика меѓу нив.

Формулата е наредба која ја внесуваме во ќелија и која пресметува одредена вредност. Во *Excel* секогаш ја почнуваме формулата со знак на еднаквост, „=“, со тоа му се кажува на програмата дека она што следи по „=“ е формула. Формулите, се занимаваат со основни аритметички операции (собирање, вадење, делење, множење и степенување) и се погодни за користење само за едноставни калкулации со многу мал број на податоци. Во многу случаи, е многу полесно, да се користи функција.

Функцијата е вградена серија на постапки, или *алгоритам*, која *Excel* ја следи за да го пресмета резултатот. И функцијата почнува со знак на еднаквост, „=“, потоа следи името на функцијата и во мала заграда се внесуваат аргументите. Во горниот пример ја употребивме функцијата **SUM** за да ги собереме сите броеви во дефинираниот опсег на ќелии а *Excel* ни ја прикажа синтаксата како $\boldsymbol{f}_x$ = SUM(B3:B30)

Генералната синтакса за функција е:

f_x = Име_на_Функцијата(......)

Редослед на операциите кај формулите и функциите

Да разгледаме еден многу едноставен пример:

Кој е точниот одговор на следниов израз ?

2 + 3 x 4 = ?

Дали е тоа : 20 ?
Дали е тоа : 14 ?

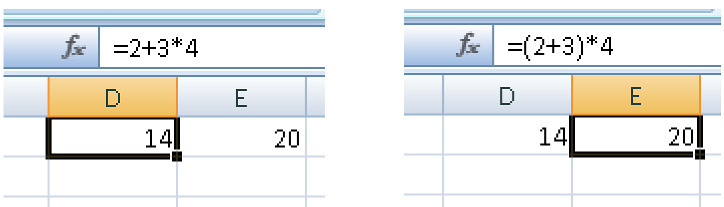
За програмерите, одговорот е веќе познат, но за тие што не се, точниот одговор е 14.
Зошто е тоа така? Затоа што кај *Excel* (и не само кај оваа програма) постои приоритет т.е. редослед на операциите по кој тие се извршуваат и кој е претставен во табелата 6.

Табела 6. Приоритет на операциите

Приоритет	Операција	Ознака
Прв	Заграда	**( )**
Втор	Експонент	**^**
Трет	Множење Делење	***** **/**
Четврт	Собирање Вадење	**+** **−**

Ако ја погледнеме табелата 6 ќе видиме дека множењето има повисок приоритет од собирањето, па кај горниот израз *Excel* прво се множи 3*4 (=12) и потоа врши собирање, т.е. продуктот го собира со 2 (=14).

Да видиме како тоа изгледа во практика:



(а) *(б)*

Слика 2.8. Приоритет на операциите

На сликата 2.8а, во ќелијата D1, *Excel* ни го презентира точниот резултат од горниот израз, а во лентата за функции ја гледаме формулата (претходно треба да ја селектираме ќелијата D1), онака како што е напишана погоре и според која е пресметан резултатот.

Но што ќе се случи ако формулата ја напишеме на следниов начин:

(2 + 3) x 4 = ?

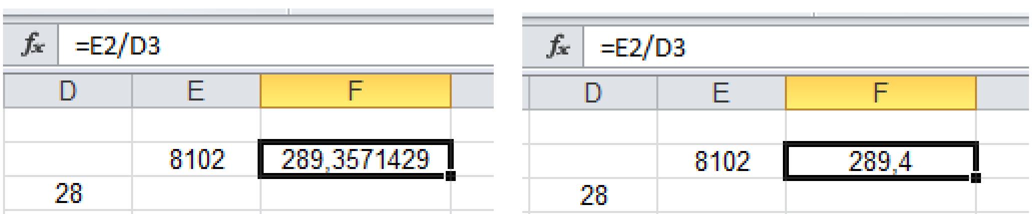
Во склад со табелата 6 заградата има повисок приоритет од множењето, па *Excel* прво ќе го пресмета изразот во заградата и резултатот ќе го помножи со 4, и крајниот резултат е 20, слика 2.8б. Ако кликнеме во ќелијата E1, каде што ја внесовме формулата, во лентата за формули ќе ја видиме формулата според која е пресметан резултатот.

Кога се работи со голем број на податоци, т.е. со многу колони и редови (ќелии), со стотици, па и повеќе податоци, користењето формула е непрактично. Прво пишувањето на формулата одзема многу време, а второ, секогаш постои можност да се направи грешка во синтаксата, па место резултат од *Excel* да добиеме порака за грешка.
Иако, постои разлика во синтаксата и дефиницијата кај формулите и функциите, некои автори функциите ги дефинираат како „пред-дефинирани“ или „вградени“ формули во *Excel* (кои патем ги има повеќе од 300).

2.7.4. Пресметување на средната вредност

- Со формула

Претходно, го презентиравме вкупниот број на податоци и нивниот збир во ќелиите D3 и E2, соодветно, сега треба да одбереме ќелија (која било)



(а) (б)

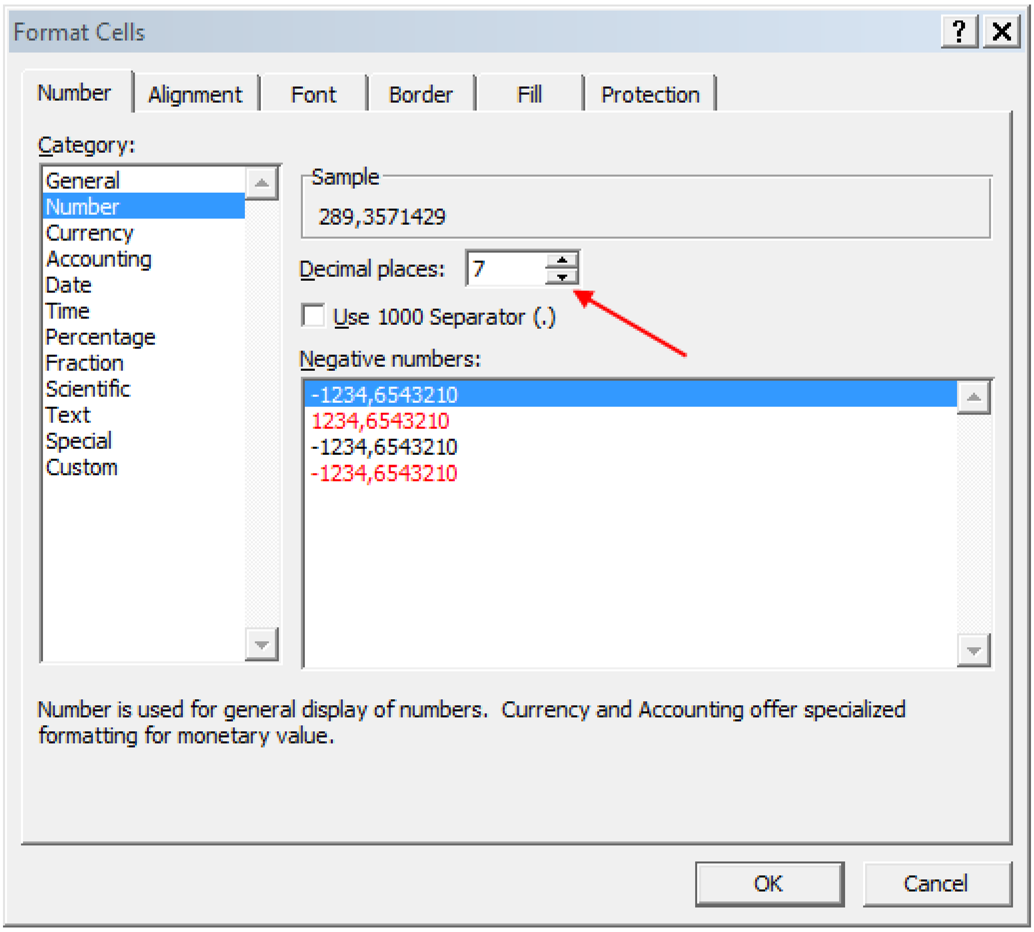
Слика 2.9. Користење на формула за пресметување на средна вредност

во која што сакаме да ја презентираме средната вредност. Нека биде тоа ќелијата F2. Во селектираната ќелија ја внесуваме следнава формула (од табелата 6 - гледаме дека за делење го користиме знакот „ **/** “):

=E2/D3

Што направивме со тоа? Со тоа му наредивме на *Excel* *„подели ја, која и да е вредност во ќелијата E2 со било која и да е вредност во ќелијата D3 и презентирај го резултатот во ќелијата F2*“. Ако потоа кликнеме на било која празна ќелија *Excel* во ќелијата F2 ќе ни ја презентира пресметаната средна вредност. Ако повторно кликнеме во ќелијата F2, во линијата за функции ќе ја видиме внесената формула, слика 2.9а.

Како што можеме да видиме од ќелијата F2 презентираниот резултат е 289,3571429. Од математика е познато, во најголем број на случаи, дека ако некоја децимала е ≥ 5 претходната децимала можеме да ја



Слика 2.10. Одредување на децимални места во ќелија

заокружиме за +1. Во нашиов случај втората децимала е „5“ па следи дека претходната децимала (3) можеме да ја заокружиме на 4 (3+1) така, како што ја пресметавме мануелно средната вредност во претходното поглавје (289,4).

Но, како можеме да постигнеме *Excel* сам да го направи тоа за нас т.е. да ни презентира резултат заокружен на една децимала. За тоа потребно е да му

„кажеме“ дека сакаме резултат со една децимала, а му „кажуваме“ на следниов начин:

Ја селектираме ќелијата F2 , слика 2.9а ⇒ кликаме со десното копче од глувчето врз селектираната ќелија ⇒ со левото копче кликаме на *Format Cells* ⇒ кликаме на *Number* ⇒ во дијалог боксот што ќе ни се појави, кликаме на стрелката надолу во полето *Decimal places* сѐ додека не дојдеме до бројот 1, слика 2.10 ⇒ *ОК*.
Сега резултатот во ќелијата F2 ни е презентиран со една децимала, 289,4, како на сликата 2.9б.

- Со функција AVERAGE

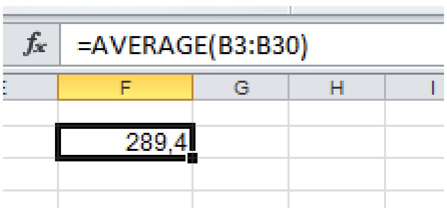
Ако ја користиме функцијата **AVERAGE** (просечна вредност) воопшто не ни треба да го знаеме бројот на податоците и нивниот збир. *Excel* нив ги зема, ги пресметува во позадина и во ќелијата ни го презентира само крајниот резултат.

Во категоријата *Statistical*, слика 2.3 под *Select a function* избираме *AVERAGE* ⇒ ги селектираме сите податоци од сликата 2.1 ⇒ *ОК*

Excel ни го презентира резултатот во ќелијата F2, додека во лентата за функции ја гледаме инсертираната функција, слика 2.11.

Синтаксата на функцијата е:

$$f_x = \text{AVERAGE(B3:B30)}$$



Слика 2.11. Користење на функцијата AVERAGE

Оваа синтакса значи „пресметај ја средната вредност на сите податоци кои се наоѓаат почнувајќи од ќелијата B3 па, се до ќелијата B30“.

2.7.5. Функција STDEV.S

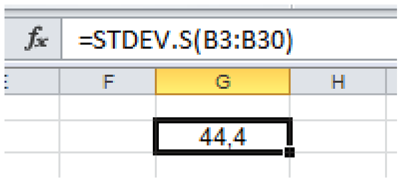
Оваа функција ја користиме за да ја определиме стандардната девијација на примерокот. Ја користиме на сличен начин како претходните. Прво ја селектираме ќелијата во која сакаме да ни се прикаже резултатот, пр. G2.

Од статистичката категорија на функции ја избираме функцијата *STDEV.S*, слика 2.3 ⇒ ги селектираме сите податоци ⇒ *ОК*.

Во ќелијата G2 ни се покажува вредноста за стандардната девијација (44,4) додека во линијата за функции – инсертираната функција, слика 2.12.

Синтаксата на функцијата е:

$$f_x = \mathrm{STDEV.S(B3{:}B30)}$$



Слика 2.12. Употреба на функцијата STDEV.S.

Оваа синтакса значи „пресметај ја стандардната девијација на сите податоци почнувајќи од ќелијата B3 па се до ќелијата B30“.

2.7.6. Функциите MAX и MIN

На идентичен начин како претходно, со користење на функциите MAX и MIN, ги одредуваме максималната и минималната вредност на податоците, слика 2.13.

Слика 2.13. Одредување на максималната и минималната вредност

Синтаксата на овие функции е:

$$f_x = \mathrm{MAX(B3{:}B30)}$$

$$f_x = \mathrm{MIN(B3{:}B30)}$$

2.7.7. Одредување на $T_{S_{min}}$ и $T_{S_{max}}$

За пресметување на $T_{S_{min}}$ и $T_{S_{max}}$ ќе ни биде потребно да ги знаеме вредностите за максималната , минималната и средната вредност, како и стандардната девијација. Овие податоци, да се потсетиме, ги внесовме во ќелиите претставени во табелата 7.

Табела 7. Податоци во Excel ќелиите

Податок	**Ќелија**
Средна вредност	F2
Стандардна девиација	G2
Максимална вредност	F1
Минимална вредност	G1

За да пресметаме $T_{S_{min}}$ потребно е да ја употребиме следнава равенка:

$$T_{S_{\min}} = \frac{(\bar{x} - x_{\min})}{s}$$

Во *Excel,* резултатот ќе го презентираме во ќелијата G3, во која ќе ја напишеме горната формула на следниов начин:

$T_{S_{min}}$=(F2-G1)/G2

Соодветно за $T_{S_{max}}$ формулата е:

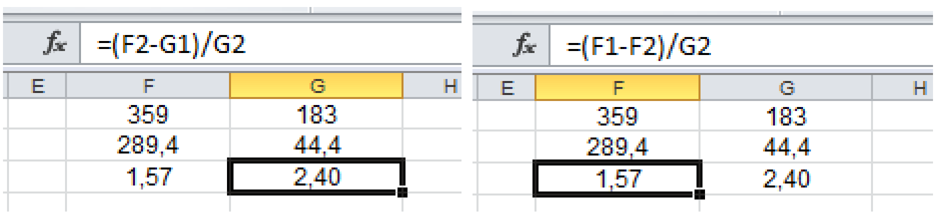
$$T_{S_{\max}} = \frac{(x_{\max} - \bar{x})}{s}$$

а во *Excel* би изгледала:

$T_{S_{max}}$=(F1-F2)/G2

Оваа формула ќе ја внесеме во ќелијата F3.

На сликата 2.14 ни се презентирани соодветните резултати.



$T_{S_{min}}$ $T_{S_{max}}$

Слика 2.14. Пресметување на $T_{S_{min}}$ *и* $T_{S_{max}}$

При оваа пресметка користевме формула, затоа што во *Excel* не постои вградена функција за тоа.

За пресметките за кои не постојат вградени функции во Excel користиме формули.

2.7.8. Одредување на 95% C.I.

За одредување на 95% интервал на доверба ја користиме равенката:

$$95\%C.I = \pm t_{0,05} \cdot \frac{s}{\sqrt{N}}$$

Стандардната девијација ни беше пресметана во ќелијата G2, додека вкупниот број на податоци во D3, па овие променливи во *Excel* ќе ги замениме со соодветните ќелии, додека вредноста за $t_{0,05}$ ја земеме од табела 3 во пресекот на колоната 0,05 и редот 27 (f = 28 – 1) и изнесува 1,703. Горната формула сега можеме да ја напишеме како:

$$95\%C.I = \pm 1{,}703x\frac{G2}{\sqrt{D3}}$$

или во *Excel* оваа формула би изгледала:

$$f_x = 1{,}703*(G2/D3\text{^}0{,}5)$$

Оваа формула ќе ја внесеме пример во ќелијата F4 каде ќе ни се презентира резултатот, =14,3 слика 2.15а.

fx =1,703*(G2/D3^0,5)

F	G
359	183
289,4	44,4
1,57	2,4
14,3	4,9

(а)

fx =(F4/F2)*100

F	G
359	183
289,4	44,4
1,57	2,4
14,3	4,9

(б)

Слика 2.15. Пресметување на 95% C.I.

Доколку во ќелијата G4 сакаме да го презентираме 95% C.I. во % ќе ја напишеме формулата како што е претставено на сликата 2.15б.

2.7.9. Одредување на *95% D.R.*

За да го одредиме 95% опсег на податоци во ќелијата F5 ќе ја напишеме формулата:

$$95\%\ Data\ Range = 2{,}052*G2$$

а за да одредиме %, во ќелијата G5 ја пишуваме формулата:

$$95\%\ Data\ Range=(F5/F2)*100$$

Резултатите ни се презентирани на сликата 2.16.

fx =1,703*G2

F	G
359	183
289,4	44,4
1,57	2,4
14,3	4,9
75,6	26,1

fx =(F5/F2)*100

F	G
359	183
289,4	44,4
1,57	2,4
14,3	4,9
75,6	26,1

Слика 2.16. Пресметување на 95% C.I.

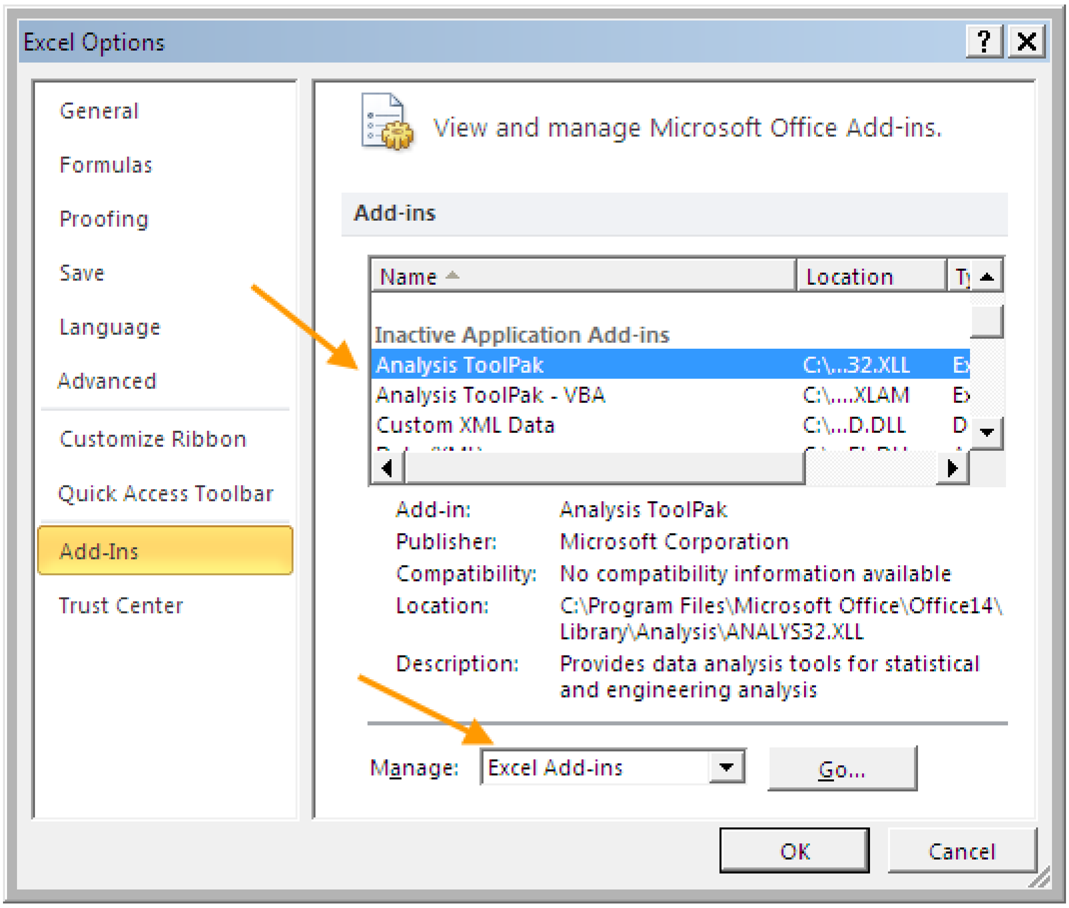
Домашна задача

Пробајте да ги обработите податоците од примерот 2, со *Excel,* користејќи ги коригираните 4 податоци.

2.7.10. Додаток во *Excel* за анализа на податоци

За да можеме да ги користиме можностите на *Excel 2010* за статистичка анализа потребно е да го инсталираме додатокот за анализа, кој по дифолт инсталација на програмата, не се инсталира. Постапката е следна:

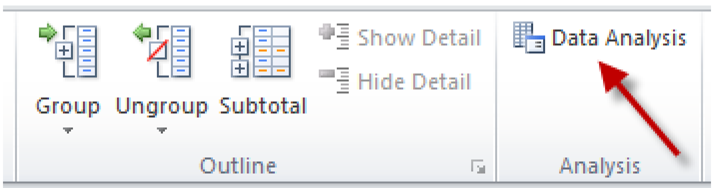
Кликаме на *File* ⇒ *Options* ⇒ *Add-Ins* и ни се појавува прозорец со активни и неактивни додатоци (*add-ins*), слика 2.17. Од сликата, можеме да видиме дека *Analysis ToolPak* не ни е активен. Во паѓачкото мени *Manage* (долу , лево)



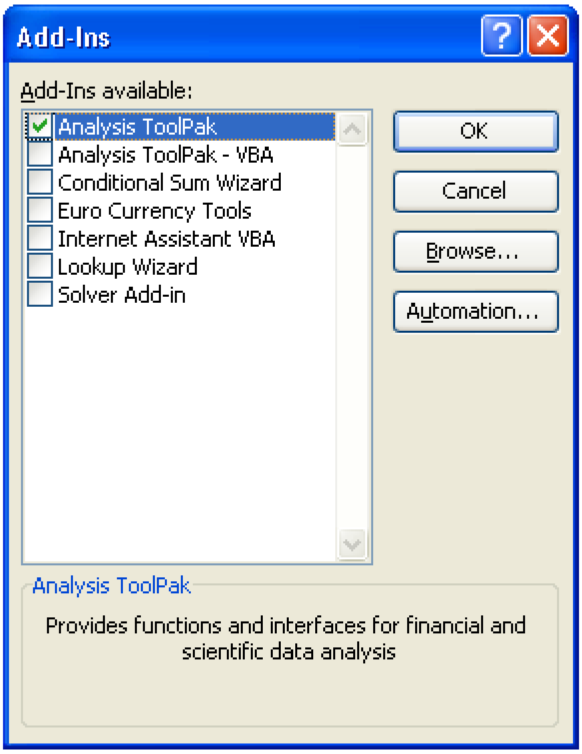
Слика 2.17. Инсталација на Analysis ToolPak во Excel 2010

избираме *Excel Add-Ins* ⇒ кликаме на копчето *Go*…. ⇒ ни се појавува прозорец како оној од сликата 2.19 ⇒ го чекираме боксот пред *Analysis ToolPak* ⇒ кликаме *ОК* ⇒ YES

Ако е инсталацијата успешна, кога ќе го кликнеме *Data* табот, во менито на крајот на рибонот ќе ни се појави групата *Analysis* и во неа командата *Data Analysis*, слика 2.18.

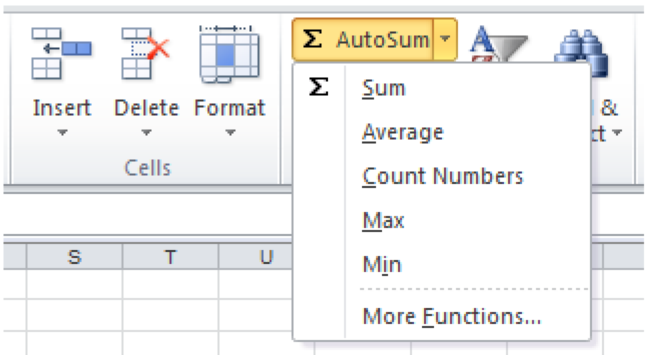


Слика 2.18. Доказ за инсталација на алатките за анализа на податоци

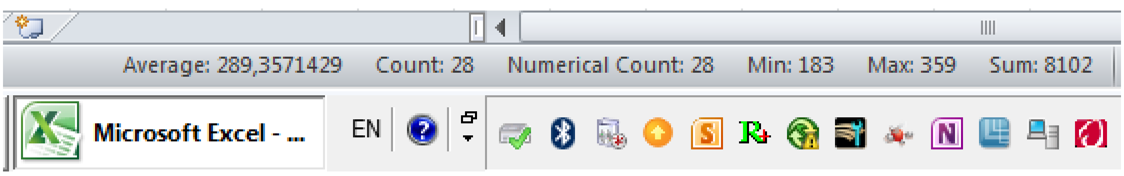


Слика 2.19. Додатоци на Excel 2010

И конечно, по инсталацијата на додатокот за анализа, ако ги селектираме ќелиите со податоците, *Excel 2010* автоматски ни ги презентира основните статистички податоци, слика 2.20а, во паѓачкото мени на *AutoSum* командата, а исто така ги презентира во најдолната лента од интерфејсот, слика 2.20б.



(а)



(б)

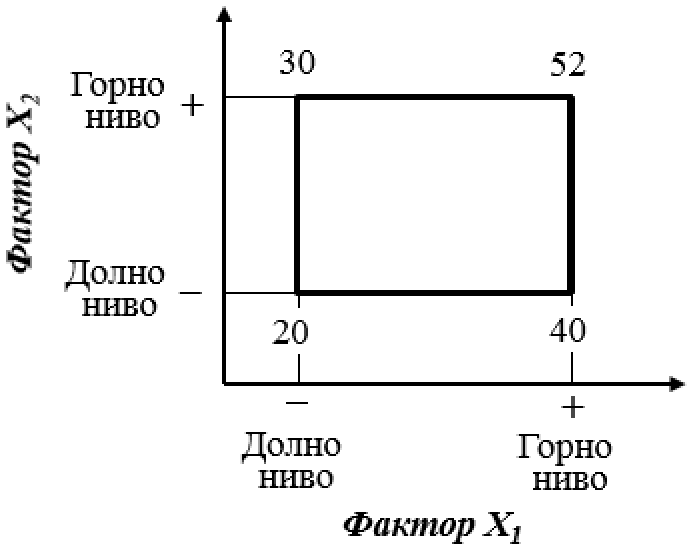
Слика 2.20. Основните статистички податоци Excel 2010 автоматски ни ги презентира за избраниот опсег на ќелии

§3. ВОВЕД ВО ФАКТОРЕН ДИЗАЈН НА ЕКСПЕРИМЕНТИ

3.1. Основни дефиниции и принципи

Многу експерименти вклучуваат проучување на ефектите од два или повеќе фактори. Во принцип,фактор дизајнот е факторните дизајни најефикасен од овој тип на експерименти. Под фактор дизајн факторен дизајн подразбираме дека во еден комплетен експеримент сите можни комбинации од фактори се испитани. На пример, ако имаме ***а*** ниво на факторот X_1 и ***б*** ниво на факторот X_2 комплетниот експеримент ќе ги содржи сите ***аб*** комбинации. Кога факторите се подредени по фактор дизајн често се вели дека се **вкрстени**.

Ефектот на факторот го дефинираме како промена на одзивот предизвикана со промена на нивото на факторот. Ова често се нарекува **главен ефект** бидејќи се однесува на примарните фактори од интерес во експериментот. На пример, да разгледаме едноставен експеримент на сликата 3.1. Ова е експеримент со два фактори кај кој двата фактори се на две нивоа.



Слика 3.1.Експеримент со два фактори - одзивот (y) прикажан во ќошевите

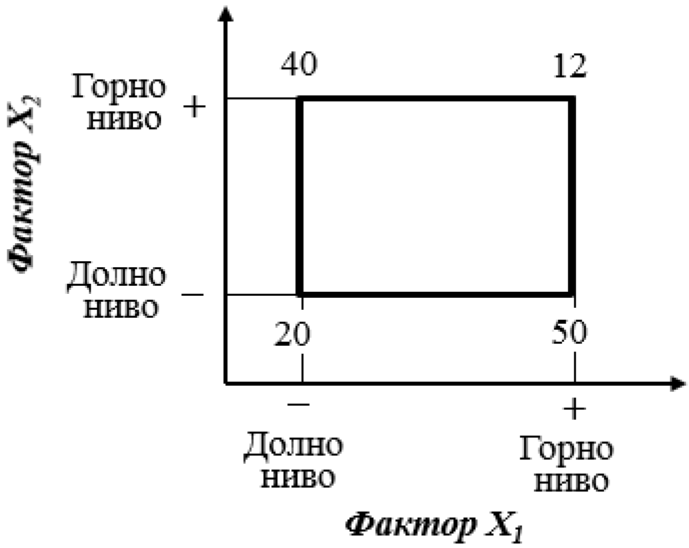
Овие нивоа ги нарекуваме „**долно**“ и „**горно**“ и ги означуваме со „–“ и „**+**“ соодветно. Во овој дизајн со два фактори главниот ефект на факторот X_1 го посматраме како разлика на средниот (просечниот) одзив на долното ниво од X_1 и средниот одзив на горното ниво од X_1. Нумерички, тој е:

$$X_1 = \frac{40+52}{2} - \frac{20+30}{2} = 21$$

Тоа значи дека зголемувањето на факторот X_1 од долно на горно ниво предизвикува средно зголемување на одзивот од 21 единица. На сличен начин, главниот ефект на X_2 е:

$$X_2 = \frac{30+52}{2} - \frac{20+40}{2} = 11$$

Кај некои експерименти, може да се забележи дека разликата на одзивот меѓу нивото на еден фактор не е иста на сите нивоа од другиот фактор. Ако ова е случај тоа значи дека постои **интеракција** меѓу факторите. На пример, да го разгледаме експеримент со два фактори прикажан на слика 3.2.



Слика 3.2. Експеримент содва фактори - интеракција

На долното ниво од факторот X_2 ефектот на X_1 е:

$$X_1 = 50 - 20 = 30$$

додека на горното ниво од факторот X_2, ефектот на $X1$ е:

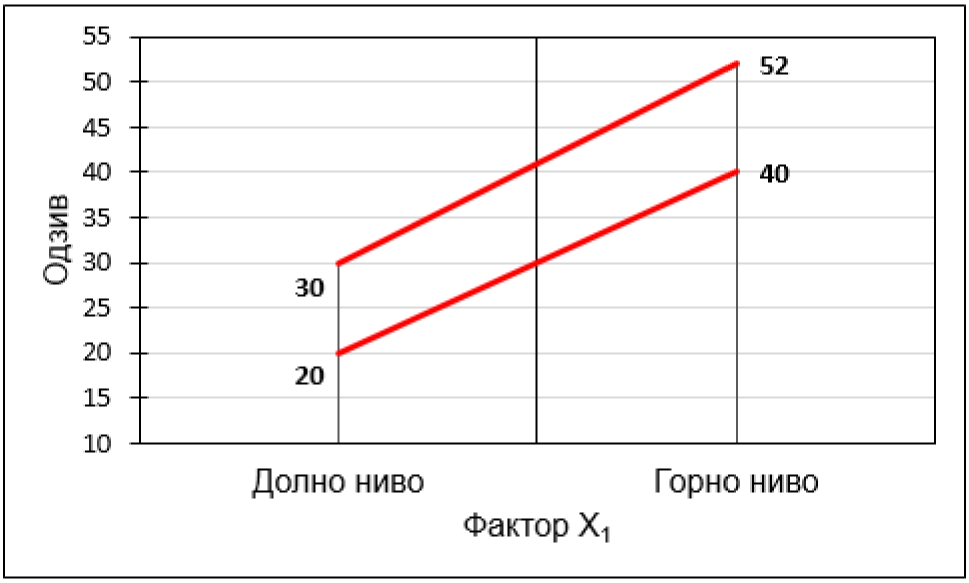
$$X_1 = 12 - 40 = -28$$

Бидејќи ефектот на X_1 зависи од нивото на факторот X_2 велиме дека постои интеракција меѓу X_1 и X_2. Интензитетот на интеракцијата е средната *разлика* во овие два X_1 ефекти, или:

$$X_1 X_2 = \frac{-28-30}{2} = -29$$

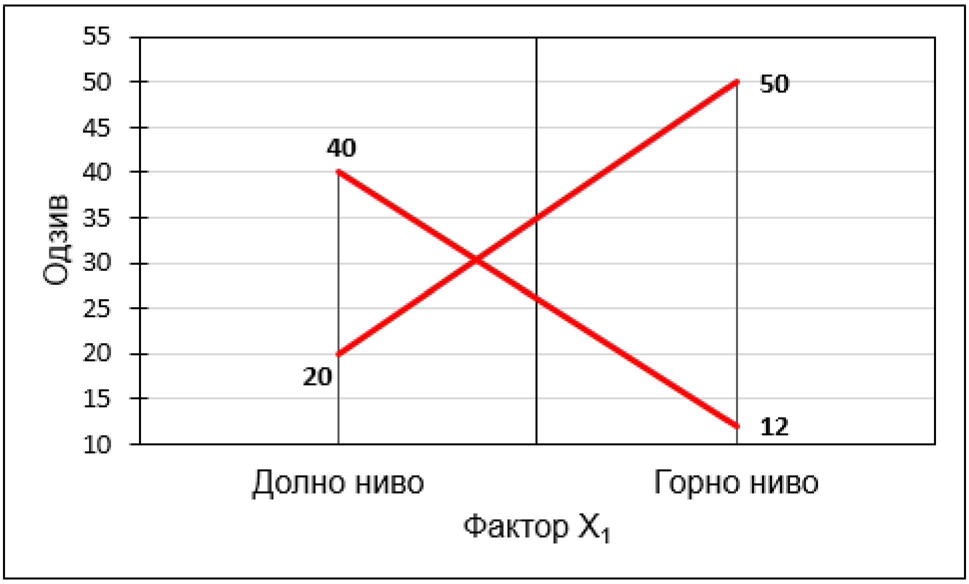
Очигледно е дека во овој експеримент интеракцијата е голема.

Ова објаснување можеме графички да го илустрираме. На сликата 3.3 прикажан е одзивот од сликата 3.1 за факторот X_1 при двете нивоа на факторот X_2. Може да се забележи дека $-X_2$ и $+X_2$ линиите се скоро паралелни што индицира непостоење на интеракција меѓу факторите X_1 и X_2. Слично, на



Слика 3.3. Експеримент од два фактори без интеракција

на сликата 3.4 се прикажани податоците од слика 3.2. Овде гледаме дека $-X_2$ и $+X_2$ линиите не се паралелни. Тоа е индикација за интеракција меѓу факторите X_1 и X_2.



Слика 3.4. Експеримент со два фактори - интеракција

Постои и друг начин да се илустрира интеракцијата. Да претпоставиме дека двата горни фактори се квантитативни (температура, притисок, време, итн.).

Тогаш тие можат да се презентираат преку регресивниот модел на експеримент со два фактори кој може да се напиши како:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \varepsilon$$

y е одзивот, x-те се факторите (варијаблите), со β е означен непознат параметар кој ќе биде одреден од податоците на експериментот и ε е експерименталната грешка на системот што го проучуваме. Во кодиран облик варијаблите X_1 и X_2 се дефинирани во опсегот од -1 до +1 (долно и горно ниво), додека X_1X_2 ја претставува интеракцијата меѓу X_1 и X_2. Одредувањето на вредноста на параметарот β во регресивниот модел е поврзано со ефектот на факторите. За експериментот прикажан на сликата 3.1 најдовме дека главниот ефект на X_1 и X_2 соодветно е 21 и 11. Од тука, имајќи во предвид дефиницијата дека β_1 и β_2 се еднакви на една половина од соодветниот главен ефект, следи дека се:

$$\beta_1 = \frac{21}{2} = 10{,}5$$

$$\beta_2 = \frac{11}{2} = 5{,}5$$

Ефектот на интеракцијата од сликата 3.1 е $X_1\ X_2 = 1$, па следи:

$$\beta_{12} = \frac{1}{2} = 0{,}5$$

Параметарот β_0 го одредуваме како средна вредност од сите четири одзиви:

$$\beta_0 = \frac{20+40+30+52}{4} = 35{,}5$$

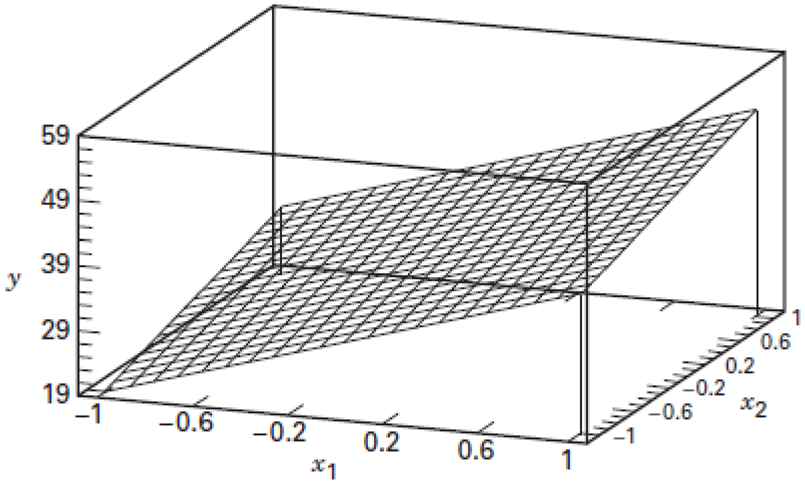
Според тоа, поставениот регресивниот модел е:

$$y = 35{,}5 + 10{,}5x_1 + 5{,}5x_2 + 0{,}5x_1x_2$$

Коефициентот на интеракција (β_{12} = 0,5) е релативно премал, споредeн со коефициентите на главните ефекти β_1 и β_2 и затоа може да се игнорира. Со отфрлање на членот $0{,}5x_1x_2$ го добиваме моделот:

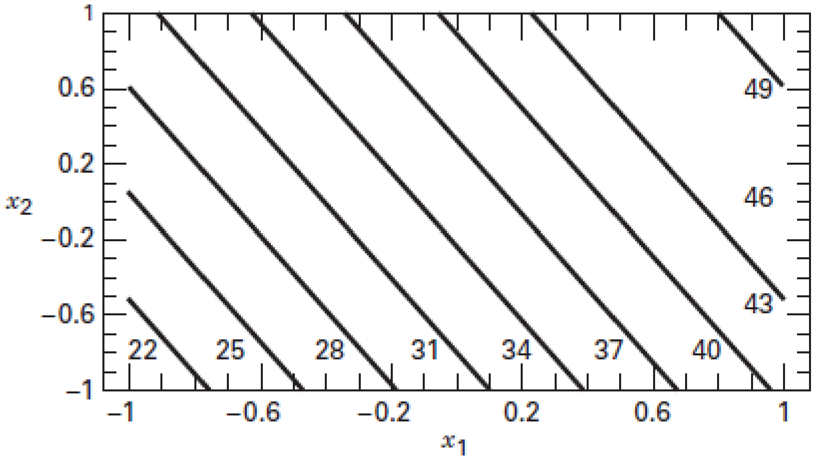
$$y = 35{,}5 + 10{,}5x_1 + 5{,}5x_2$$

На сликата 3.5 е дадена графичка презентација на овој модел. Тука имаме



Слика 3.5 Одзивна површина

графички приказ на рамнина од y вредности генерирани за различни комбинации на X_1 и X_2. Овој графички приказ во три димензии го нарекуваме графички приказ на одзивна површина. Сликата 3.6 ги прикажува контурните линии на константен одзив y во X_1, X_2 рамнината. Бидејќи одзивната површина е рамнина, графичкиот приказ на контурите содржи паралелни прави линии.



Слика 3.6. Графички приказ на контурите

Што всушност претставува регресивниот модел? Регресивниот модел претставува математичка форма на зависност на одзивот од вредноста на факторите во испитуваниот опсег од -1 до +1. Врз база само на четирите вредности на одзивот кои ги добивме од експериментот ја одредивме равенката на одзивот во функција од факторите. Сега за која било вредност на факторите можеме да го одредиме одзивот без да правиме дополнителни експерименти. Ќе го провериме моделот на тој начин што во равенката ќе ги замениме вредностите на X_1 и X_2 при кои е изведен експериментот односно стандардната комбинација на факторите за дизајнирање на експерименти.

Така на пример ако во моделот

$$y = 35{,}5 + 10{,}5x_1 + 5{,}5x_2 + 0{,}5x_1x_2$$

замениме:

- $X_1 = -1, X_2 = -1$, ќе добиеме:

$$y = 35{,}5 + 10{,}5(-1) + 5{,}5(-1) + 0{,}5(-1)(-1) = 20$$

- $X_1 = +1, X_2 = -1$, ќе добиеме:

$$y = 35{,}5 + 10{,}5(+1) + 5{,}5(-1) + 0{,}5(+1)(-1) = 40$$

- $X_1 = -1, X_2 = +1$, ќе добиеме:

$$y = 35{,}5 + 10{,}5(-1) + 5{,}5(+1) + 0{,}5(-1)(+1) = 30$$

- $X_1 = +1, X_2 = +1$, ќе добиеме:

$$y = 35{,}5 + 10{,}5(+1) + 5{,}5(+1) + 0{,}5(+1)(+1) = 52$$

Како што гледаме од моделот ги добивме точно истите вредности кои што ги добивме од експериментот, што значи дека моделот математички точно го опишува процесот кој го проучуваме. Сега можеме за која било комбинација на променливите да го пресметаме одзивот без да правиме експеримент. На пример да земеме дека е $X_1 = 0{,}5$ и $X_2 = 0{,}5$, тогаш одзивот ќе биде:

$$y = 35{,}5 + 10{,}5(0{,}5) + 5{,}5(0{,}5) + 0{,}5(0{,}5)(0{,}5) = 43{,}6$$

Ако правевме пресметка на одзивите според моделот во кој ја игнориравме интеракцијата ќе ги добиевме вредностите презентирани во табелата 7.

Табела 7. Споредбени вредности на одзивот за модел со и без интеракција

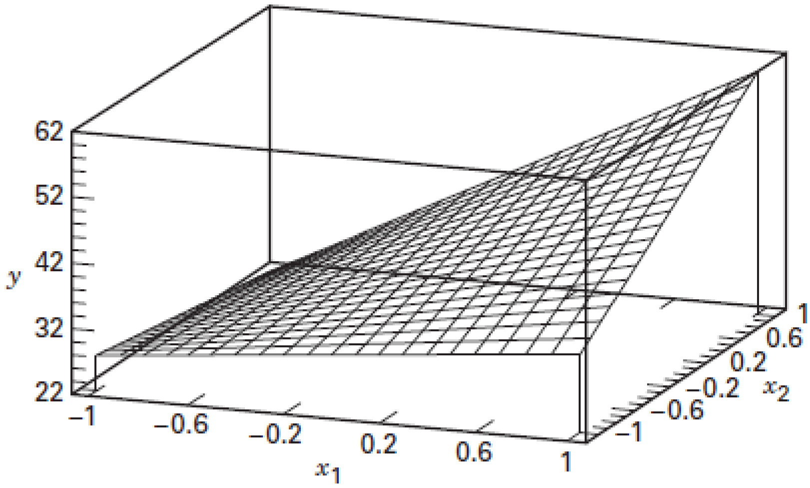
$y = 35{,}5 + 10{,}5x_1 + 5{,}5x_2 + 0{,}5x_1x_2$	$y = 35{,}5 + 10{,}5x_1 + 5{,}5x_2$
20	19,5
40	40,5
30	30,5
52	51,5

Како што можеме да видиме од табелата 7 влијанието на интеракцијата врз одзивот е незначително и оправдано е нејзиното игнорирање.

Но сега да претпоставиме експеримент во кој интеракцијата не е занемарлива, односно експеримент во кој β_{12} не е мал. Нека тоа биде моделот:

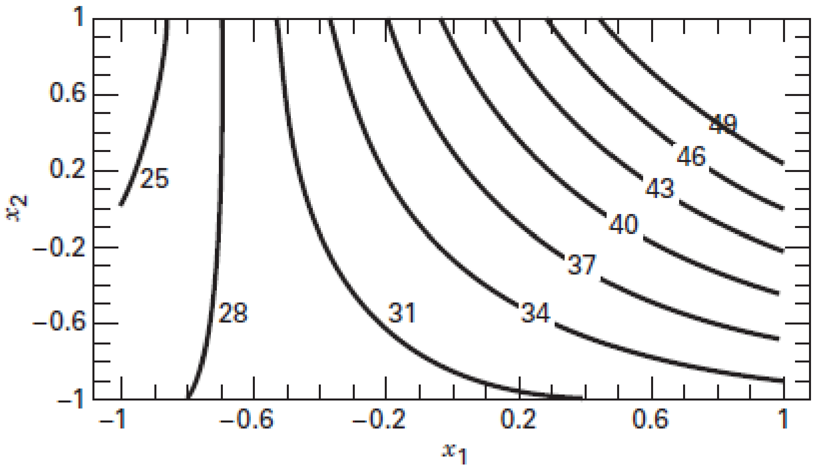
$$y = 35{,}5 + 10{,}5x_1 + 5{,}5x_2 + 8x_1x_2$$

На сликите 3.7 и 3.8 презентирани се графички прикази на одзивната површина и контурите.



Слика 3.7. Одзивна површина за модел $y = 35{,}5 + 10{,}5x_1 + 5{,}5x_2 + 8x_1x_2$

Можеме да забележиме дека значајните интерактивни ефекти ја „усукуваат" рамнината на слика 3.7. Ова усукување на одзивната површина во закривени контури - линии со константен одзив во X_1, X_2 рамнината е прикажано на сликата 3.8.

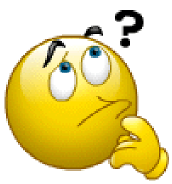


Слика 3.8. Графички приказ на контурите за модел
$y = 35{,}5 + 10{,}5x_1 + 5{,}5x_2 + 8x_1x_2$

Од тоа можеме да заклучиме дека интеракцијата претставува форма на закривена одзивна површина.Генерално, кога интеракцијата е голема, соодветните главни ефекти немаат големо практично значење. За експериментот на слика 3.2 главниот ефект на X_1 ќе биде:

$$X_1 = \frac{50+12}{2} - \frac{20+40}{2} = 1$$

што е многу мала вредност и сме доведени во забуна да заклучиме дека X_1 нема ефект. Меѓутоа, кога ќе ги провериме ефектите на X_1 на *различни нива од факторот* X_2 ќе видиме дека тоа не е случај. Факторот X_1 има ефект, но тој *зависи од нивото на факторот* X_2. Според тоа, познавањето на X_1X_2 ефектот е покорисно отколку познавањето на главниот ефект.



Задача 3.1.

Одреди го моделот за експериментот презентиран на сликата 3.4.

<u>Решение:</u>
Прво го определуваме ефектот на факторите:

$$X_1 = \frac{50+12}{2} - \frac{40+20}{2} = 1$$

$$X_2 = \frac{40+12}{2} - \frac{50+20}{2} = -9$$

$$X_1X_2 = \frac{-28-30}{2} = -29$$

Сега можеме да ги определиме коефициентите на регресивната равенка:

$$\beta_0 = \frac{20+50+40+12}{4} = 30{,}5$$

$$\beta_1 = \frac{1}{2} = 0{,}5$$

$$\beta_2 = \frac{-9}{2} = -4{,}5$$

$$\beta_{12} = \frac{-29}{2} = -14{,}5$$

Од тука моделот на овој процес ќе биде:

$$y = 30{,}5 + 0{,}5x_1 - 4{,}5x_2 - 14{,}5x_1x_2$$

Како што можеме да видиме од оваа равенка коефициентот на интеракцијата (β_{12} = 14,5) е прилично голем и не можеме да го игнорираме како во претходната равенка. И кај овој модел, ако во равенката ги замениме соодветните вредности за X_1 и X_2 ќе констатираме дека моделот веродостојно го опишува процесот или системот.

Значајната интеракција често го засенува значењето на главните ефекти. Тоа јасно е покажано на интерактивниот графички приказна слика 3.5. Во случај на присуство на голема интеракција, експериментаторот треба да ги испита ефектите на еден од факторите (пр. X_1) при константно ниво на другите фактори како би извлекол заклучоци за главниот ефект на X_1.

§4. ДВОФАКТОРЕН ПЛАН НА ЕКСПЕРИМЕНТИ

4.1. Вовед

Секогаш кога треба да се изврши некаков експеримент тој треба добро да се осмисли. Императив при тоа е да се извршат само оние проби чии резултати ќе ни дадат најдобри информации и увид во она што сакаме да го дознаеме. Методите и начините на изведување на експериментите треба да ни овозможат добивање максимален број на информации при минимален број на проби. Планот на експериментот е тесно поврзан со проучувањето на процесот или системот за кој го изведуваме експериментот т.е. со стекнување сознанија за нив.

4.2. Процес на стекнување знаења

Стекнувањето на знаења за процесот или системот можеме да го сметаме како одговор на внимателно поставени прашања.

На пример, ако еден земјоделец сака да дознае како вештачкото ѓубриво влијае врз приносот на пченка, следниве (и уште многу други) се логични прашања кои треба да си ги постави:

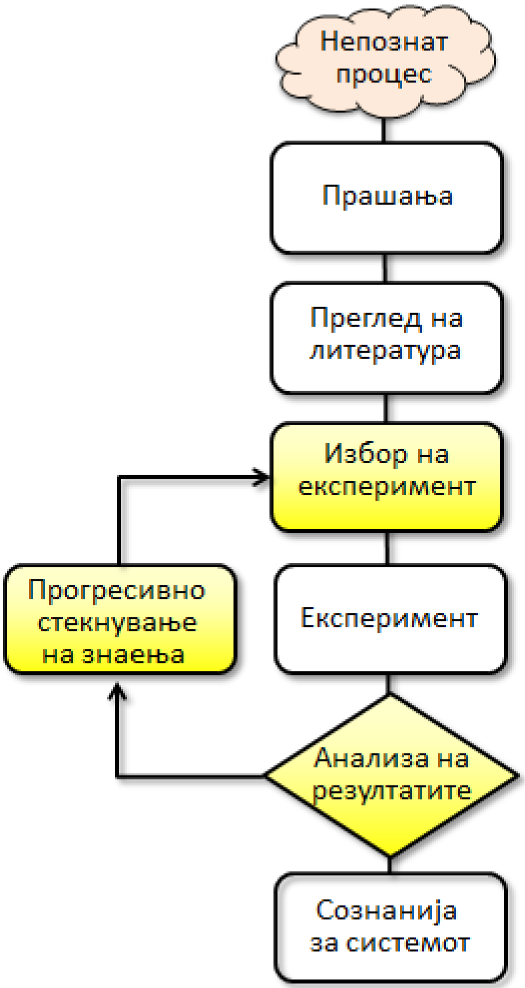
- Дали ќе добијам 100 kg повеќе пченка по хектар ако ја зголемам количината на ѓубриво?
- Како количината на дожд влијае врз ефикасноста на ѓубривото?
- Дали ќе го задржам квалитетот на пченката ако употребам одреден вид на ѓубриво?
- Колку ѓубриво треба да употребам за да добијам најголем принос?

Овие прашања го опфаќаат и проблемот и решението. Тие, исто така специфицираат кои активности треба да се преземаат. Затоа е многу важно да се поставаат прашања кои се навистина репрезентативни за проблемот.

Се разбира, пред да се почне со експериментот треба да се провери дали можеби веќе постои решение за проблемот којшто сакаме да го решиме. Затоа е потребно да се прегледа соодветна литература или да се консултираат стручни лица, да се извршат некои теоретски калкулации, симулации или да се направи нешто друго што можеби ќе даде одговор на прашањето без да се експериментира. Ако во прелиминарното истражување се најде одговор, тогаш нема потреба да се изведува експеримент.

За прашањата за кои не постои одговор неопходно е да се изведат експерименти.
Како треба да се планираат експериментите за да одговорат на следниве цели?

- Брзо пристигнување до најдобриот можен резултат.
- Избегнување на непотребните проби.
- Добивање на резултати со најголема можна прецизност.
- Изведување без неуспех.
- Создавање на модел за проучуваниот феномен.
- Откривање на оптимално решение.



Слика 4.1. Планот на експериментот ги оптимизира трите осветлени фази на процесот на стекнување на знаења

Трите основни аспекти на процесот за стекнување сознанија се прикажани на сликата 4.1, тоа се :

- Избор на експериментален метод.
- Анализа на резултатите.
- Прогресивно стекнување на знаења.

Да ги погледнеме овие аспекти подетално имајќи во предвид дека:

Експериментите се изведуваат за да ја олеснат анализата на резултатите и да овозможат прогресивно стекнување на знаења

Избор на експеримент

Експерименталниот метод треба да ја потпомогне интерпретацијата на резултатите. Меѓутоа, тој исто така, треба да го минимизира бројот на проби без да се жртвува квалитетот. Користењето на дизајнирани експерименти овозможува стекнување на предуслови кои даваат најдобра можна прецизност со најмал број на опити. Планираните експерименти даваат максимална ефикасност при најмал број на проби, а со тоа и при минимални трошоци.

Сефакторен план на експерименти

Сефакторниот план е важен метод за одредување на ефектот на повеќе варијабли врз одзивот. Традиционално, експериментите се планираат за да го одредат ефектот на ЕДНА варијабла врз ЕДЕН одзив. R.A. Fisher покажа дека има многу предности ако се врши комбинирано проучување на повеќе влезни варијабли во ист фактор експеримент. Фактор планот го смалува бројот на проби што треба да се извршат со проучување на влијанието на повеќе фактори симултано (истовремено). Додатно на тоа, тој може да се употреби за да се најдат главните ефекти, не само од секој независен фактор, туку и од интеракцијата меѓу факторите. Тоа е корисен метод за планирање на експерименти во лабораториски и индустриски услови. Фактор планот врши тестирање на сите можни комбинации на влијателни фактори врз процесот/системот. Бидејќи, овие планови, можат да доведат до голем број на проби, кои можат да бидат скапи и долготрајни, овие планови се нај погодни за мал број на варијабли со неколку нивоа (1 – 3). Фактор плановите се многу пригодни кога интеракцијата меѓу варијаблите е јака и важна и каде што секоја варијабла поединечно има важен придонес врз

Анализа на резултатите

Анализирањето на резултатите од експериментот е поврзано со изборот на експерименталниот метод. Ако експериментите се добро подготвени, резултатите лесно се толкуваат, а исто така се богати со корисни информации. Благодарение на компјутерите и софтверите, планирањето на експериментите и неопходните калкулации за анализа станаа многу едноставни и рутински. Овие алатки, исто така, овозможуваат графичка презентација која ги илустрира резултатите на јасен начин со што се олеснува разбирањето на процесот.

Прогресивно стекнување на знаења

Експериментаторот, кој го врши проучувањето на процесот, на почетокот не ги знае конечните резултати ниту што тие ќе покажат. Затоа е препорачливо постепено да се оди со експериментирањето за да би се овозможило да се прилагодат експерименталните опити врз база на иницијалните резултати и сознанија. Иницијалните резултати , на пример, овозможуваат да се насочат експериментите само во насока на интересните аспекти на процесот и да се избегне неуспех при експериментирањето.

Иницијалната серија на експерименти води кон прелиминарни, провизорни заклучоци. Користејќи ги овие иницијални заклучоци, експериментаторот може да изведе нова серија на подобрени тестови. Двете серии на експерименти се користат за да се добијат прецизни резултати. На овој начин, експериментаторот акумулира сознанија кои му се потребни, и има флексибилност да прекине со експериментот кога ќе добие задоволителни резултати.

4.3. Проучување на процесот или системот

Проучувањето на процесот често се смета како фокусирање на напорите кон одредено мерење: пример потрошувачката на гориво на една кола, цената на хемикалиите на големо, приносот на пченка по хектар. Овие мерења (потрошувачка, цена, принос) зависат од многу варијабли.

На пример, потрошувачката на гориво зависи од брзината на возење, снагата на моторот, стилот на возење, правецот на ветрот, притисокот во гумите, присуството/отсуството на багажник на покривот, бројот на патници, типот на колата итн. Цената на хемикалиите зависи од квалитетот на суровините, продуктивноста на секој производен погон, квалитетот што се бара, условите на производство и од многу други фактори. Приносот на пченка, исто така, зависи од квалитетот на почвата, количината на употребеното ѓубриво, изложеноста на сонце, климата, сортата на пченка, итн.

Математички, она што не интересира да го дознаеме го пишуваме со y и го нарекуваме ***одзив*** и е функција од неколку варијабли x_i кои ги нарекуваме ***фактори***. Општо, таа поврзаност ја пишуваме на следниов начин:

$$y = f(x_1, x_2, x_3, \ldots, x_k)$$

Проучувањето на процесот се сведува на одредување на функцијата f која го поврзува одзивот y, со факторите x_1, x_2, $x_3, \ldots,$ x_k.
Но пред да почнеме со одредување на функцијата потребно е да се запознаеме со начините како и со терминологијата која е специфична за планирање на експерименти.

4.4. Терминологија

Во продолжение ќе биде објаснета терминологијата поврзана со факторите и нивното означување.

4.4.1. Видови на фактори

Елементите на планот од експериментот како и интерпретацијата на неговите резултати многу зависи од видот на факторите кои се јавуваат во процесот. Статистиката разликува неколку видови на фактори. Во продолжение ќе разгледаме четири видови на фактори:

- **Континуирани фактори**
  Притисокот претставува пример за континуиран фактор. За даден интервал која било вредност од интервалот може да биде избрана. Други примери се бранова должина, температура или концентрацијата на раствор. Значи *континуирани се оние фактори кои можат да имаат било која вредност во даден интервал*.

- **Дискретни фактори**
  За разлика од континуираните, дискретните фактори можат да имаат само одредени вредности. Овие вредности не мора секогаш да се нумерички. Да земеме, на пример, дека површинската маса на ткаенина, според спецификацијата на даден процес, треба да биде од 220 до 240 g/m^2. Тоа значи дека ткаенината може да биде: 225; 237; 228,5; 236,87 g/m^2 итн. Значи, ткаенината може да има вредност на кој било цел и кој било децимален број во дадениот интервал. Во процесот во кој се преработува таа ткаенина, во зависност од потребите, секогаш работат, да речеме, меѓу 50 и 60 работници. Бројот на работниците некогаш е 52, некогаш – 58 но никогаш не може да биде 54,7 или 58,5. Што значи, бројот на работниците може да биде дефиниран само со одредени вредности т.е. цели броеви. *Целите броеви се еден пример за дискретни фактори*, додека површинската маса на ткаенината е пример за континуиран фактор.

Редни фактори

Редните фактори се дискретни фактори кои можат да се стават во логичен редослед. На пример, големината на конфекциските делови може да се означи како мала, средна, голема или екстра голема. Хиерархијата на одлучување е исто така реден фактор: прво, второ, трето... ниво на одлучување на пример во една фирма или институција.

- **Буленови фактори**

  Буленовите фактори се дискретни фактори кои можат да имаат само две вредности: вклучен/исклучен, горе/долу, отворен/затворен, +1/-1, итн.

Понекогаш, разликата меѓу овие различни видови на фактори не може јасно да се разграничи затоа што тие можат да се сфатат контекстуално како еден или друг вид на фактори. Континуиран фактор, како што е брзината може да се трансформира во реден или дискретен фактор: голема или мала, или брзина А и брзина В. Оваа можност не е недостаток – таа претставува дополнителна флексибилност која експериментаторот може да ја искористи при интерпретација на резултатите. Всушност, изборот понекогаш го олеснува истакнувањето на некои аспекти од проучувањето. Промената на типот на варијаблата, исто така, значи адаптирање на одговорот кон целта на проучувањето. На пример, да претпоставиме дека ја проучуваме возраста на извесна популација од луѓе, и ако сакаме да го најдеме бројот на луѓе на извесна возраст, варијаблата „возраст“ е редна варијабла бидејќи постојат неколку возрасни категории: деца, тинејџери, возрасни и старци. Но, ако проучуваме колку има млади луѓе, помлади од 18 години, варијаблата „возраст“ станува Буленова: помлади од 18 и постари од 18.

4.4.2. Експериментален простор

За да се сфати дизајнот на експерименти потребно е да се знае што е тоа *експериментален простор* на варијаблите од процесот. Овој фундаментален концепт детално ќе го објасниме користејќи континуирани фактори, бидејќи се тоа најчесто користените фактори.

За графички да го илустрираме експерименталниот простор ќе употребиме површина со две димензии која, по потреба, лесно може да се прошири на мулти димензионален простор.

Еден континуиран фактор може да се претстави на насочена и градуирана оска. Ако има и втор континуиран фактор и тој се претставува на слична оска. Оваа втора оска се црта ортогонално (нормално) во однос на првата (т.е. тие формираат агол од 90°). Површината меѓу оските се вика експериментален простор, слика 4.2. Експерименталниот простор се состои од точки на површината (фактор 1) x (фактор 2), во кој секоја точка претставува експериментална проба.

Слика 4.2. Експериментален простор

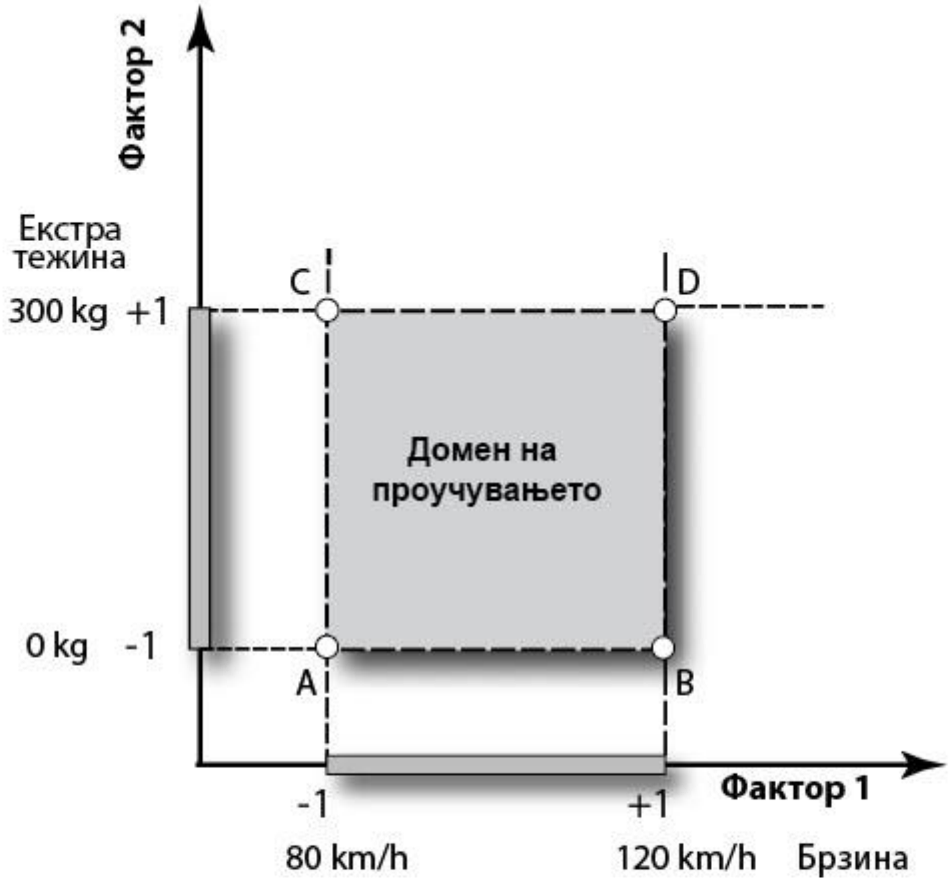
Ако постои трет фактор, и тој ќе биде претставен со насочена и градуирана оска поставена нормално на првите две.

4.4.3. Домен на факторите

Вредноста која му се доделува на факторот додека се прави експерименталната проба се вика ***ниво***. Кога го проучуваме влијанието на даден фактор ги ограничуваме неговите вредности меѓу два лимити. Експериментаторот ги дефинира овие два лимити според спецификите на проучувањето за кое се прави експериментот. Долниот лимит го нарекуваме ***долно ниво*** додека горниот лимит – ***горно ниво***. На пример, ако го проучуваме влијанието на брзината со која се движи една кола врз потрошувачката на гориво можеме да ја менуваме брзината на движење од 80 до 120 km/h. Брзината од 80 km/h е долно ниво додека брзината од 120 km/h – горно ниво. Множеството на сите вредности меѓу долното и горното ниво, кои факторот може да ги има, се вика ***домен на варирање на факторот*** или поедноставно ***домен на факторот***.

Слика 4.3. Горно и долно ниво на факторите

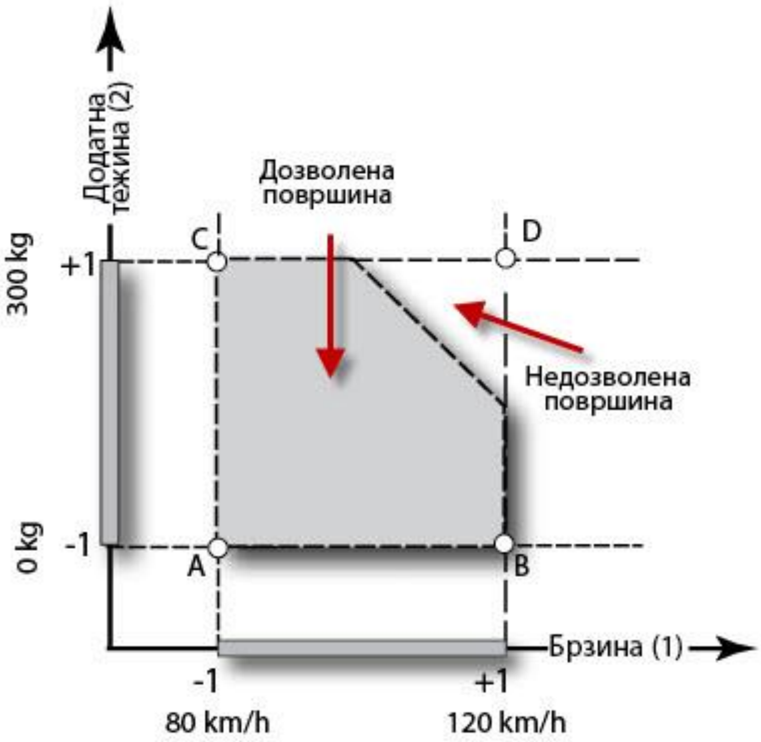
Ако има повеќе фактори, секој има свој домен. Бидејќи различните фактори можат да имаат различни мерни единици корисно е да се има заеднички репрезент за сите нив. При планирањето на експериментите вообичаено е долното ниво да се означи со **– 1** а горното со **+ 1**. Во горниот пример брзината од 80 km/h ја означуваме како – 1 ниво, додека 120 km/h како + 1 ниво, слика 4.3. Внатрешноста на доменот од факторот ги содржи сите вредности кои тој теоретски може да ги има.



Слика 4.4. Доменот на проучување е дефиниран како унија од домените од различните фактори. Овде нема ограничување

4.4.4. Домен на проучување

Во практика, експериментаторот избира дел од експерименталниот простор за да го изврши проучувањето. Оваа посебна зона од експерименталниот простор ја нарекуваме ***домен на проучувањето***, слика 4.4. Овај домен е дефиниран од горните и долните нивоа на сите фактори, а понекогаш и од ограничувањата меѓу факторите. Да земеме дека втор фактор е екстра (додатната) тежина на возилото која ја дефинираме како екстра маса, покрај онаа од возилото и возачот (без багаж). Долното ниво на оваа додатна тежина е 0 kg - кога нема додатна тежина. Во возилото можат да се возат уште три лица од по пр. 75 kg кои носат по 25 kg багаж. Во тој случај екстра тежината ќе биде (3 x 75) + (3 x 25) = 300 kg, па горното ниво го дефинираме дека е 300 kg. Ако нема ограничувања, доменот на проучување е претставен со сите точки меѓу 0 и 300 kg екстра тежина и со 80 до 120 km/h брзина, слика 4.4.



Слика 4.5. Домен на проучување со ограничување

Можно е и да постои ограничување во однос на доменот на проучување. На пример, можно е колата да не може да постигне брзина од 120 km/h кога е премногу оптоварена со додатна тежина. На сликата 4.5 е илустрирано ова посебно ограничување на почетниот домен на проучување.

4.5. Центрирани и скалирани варијабли

Кога долното ниво на факторите е претставено со – 1 а горното со +1 се случуваат две важни промени:

1. **Центарот на мерење се поместува**
   Во нашиов пример центарот на интервалот [- 1, 1] е нула и одговара на вредноста 100 km/h [(80+120)/2]. Нумеричката вредност на точката каде сега се наоѓа нулата т.е. новата точка на координатниот почеток се разликува од координатниот почеток (нулата) кога брзината ја меревме во оригиналните експериментални единици, кои понекогаш ги нарекуваме ***инженериски единици*** или ***натурални единици*** (пр. km/h).
2. **Единиците на мерење се променуваат**
   Во нашиов пример, долното ниво на факторот брзина е 80 km/h а горното ниво 120 km/h па меѓу овие две вредности има разлика од 40 km/h (120 – 80

= 40) т.е. 40 по множено по единицата за брзина (km/h). Меѓу – 1 и 1 има разлика „2“ која е нова единица [1 – (–1) = 2]: ново дефинираната единица (2) одговара на 40 km/h и се нарекува ***интервал***.

Овие две промени доведуваат до воведување на нови варијабли наречени *центрирани и скалирани варијабли*. „Центрирани“ се однесува на промената на координатниот почеток, а „скалирани“ се однесува на промената на мерните единици. Новите варијабли обично се нарекуваат ***кодирани варијабли*** или ***кодирани единици***.

Конверзијата на оригиналните варијабли A во кодирани варијабли x (и обратно) е дадена со следнава релација:

$$x = \frac{A - A_0}{Интервал} \tag{5.1}$$

A_0 – е централната вредност во инженериски единици.

Предноста од користењето на кодирани единици лежи во нивната можност да презентираат планирани експерименти на ист начин, без разлика од избраниот домен на проучување и без разлика на видот на факторите т.е. нема не компатибилност меѓу нив заради нивните различни мерни единици.

Кодираните варијабли претставуваат однос (дропка) на две исто-димензионални физички единици, па според тоа тие се без димензионални. Предноста од отсуството на. инженериски единици се должи на фактот што сите фактори имаат ист домен на варирање (тоа се двете кодирани единици), со што се овозможува директна споредба на ефектот кој што се јавува од меѓусебната интеракцијата (содејството) на самите фактори.

Пример 1.

Еден експериментатор го избира факторот брзина на долното ниво да биде 80 km/h а на горното - 120 km/h. Која брзина во кодирани единици одговара на 95 km/h?

$$Интервал = \frac{A_{+1} - A_{-1}}{2} = \frac{120 - 80}{2} = 20$$

A_0 е централната вредност, во инженериски единици, меѓу горното и долното ниво, што значи претставува половина од сумата на горното и долното ниво:

$$A_0 = \frac{A_{+1} + A_{-1}}{2} = \frac{120 + 80}{2} = 100$$

Ако го замениме во равенката за конверзија (5.1) ќе добиеме:

$$x = \frac{A - A_0}{Интервал} = \frac{95 - 100}{20} = -\ 0{,}25$$

Брзината од 95 km/h одговара на вредност од – 0,25 во кодирани единици т.е. во кодирана вредност.

Пример 2.

Вредноста можеме да ја пресметаме и во инженериски единици ако ја знаеме кодираната вредност. Која е вредноста на факторот брзина, во инженериски единици, што одговара на вредност од +0,5 во кодирани единици?

Ако тргнеме од претходната равенка, со замена ќе добиеме:

$$+0{,}5 = \frac{A - 100}{20}$$

Следи: $$A = 100 + (0{,}5x20) = 110$$

Кодираната брзина од 0,5 одговара на вистинска брзина од 110 km/h.

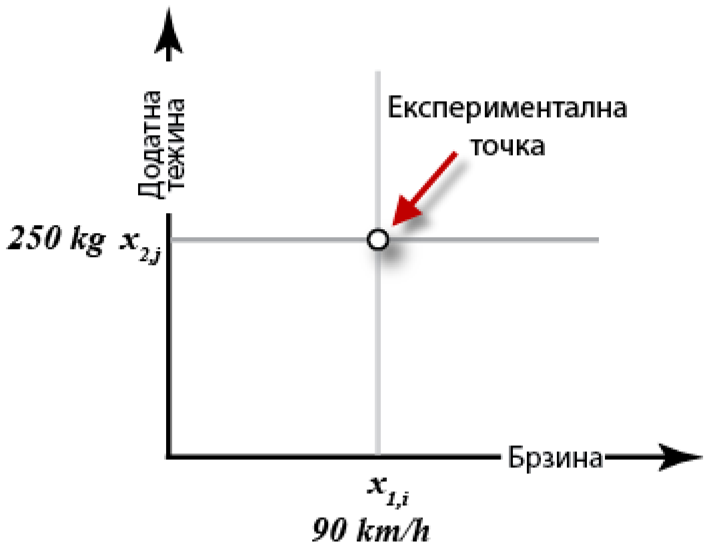
Кодираните варијабли не можат да имаат помала вредност од – 1, ниту поголема од + 1.

4.6. Експериментални точки

Во просторот со две димензии, нивото i на факторот 1, напишано како $x_{1,i}$ и нивото j на факторот 2 напишано како $x_{2,j}$, можат да се сметаат како координати на точка во експерименталниот простор или во доменот на проучување, слика 4.6. На пример, ако нивото на брзината е 90 km/h, а додадената тежина е 250 kg, координатите на експерименталната точка се:

$$x_{1,i} = 90\ km/h$$

$$x_{2,j} = 250\ kg$$

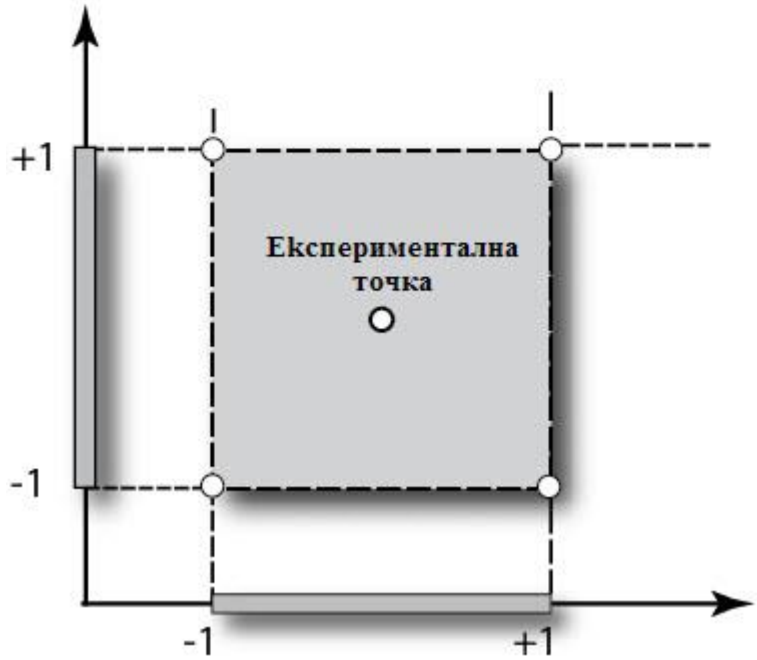


Слика 4.6. Експериментална точка

Една експериментална проба може да биде претставена во координатниот систем со една точка. Тоа е причина што пробата често ја нарекуваме ***експериментална точка*** или само ***точка***. Според тоа, планираниот експеримент, е претставен со множество на експериментални точки сместени во експерименталниот простор. Во нашиов конкретен пример, експериментот се изведува врз кола која се движи со 90 km/h со додатна тежина од 250 kg.

4.7. Дизајн на експериментот

Изборот на бројот и положбата на експерименталните точки е фундаментален проблем на дизајнираните експерименти. Во идеален случај, би сакале да направиме минимален број на проби при тоа редуцирајќи го влијанието на експерименталната грешка врз моделите кои ќе се користат за да се донесат одлуки. Оваа цел се постигнува со разгледување на математичките и статистичките карактеристики кои го поврзуваат одзивот со факторите. Ако нема ограничувања во однос на доменот на проучување, постојат класични планови кои имаат одлични статистички квалитети и кои овозможуваат моделирање на одзивот при најповолни услови, слика 4.7.



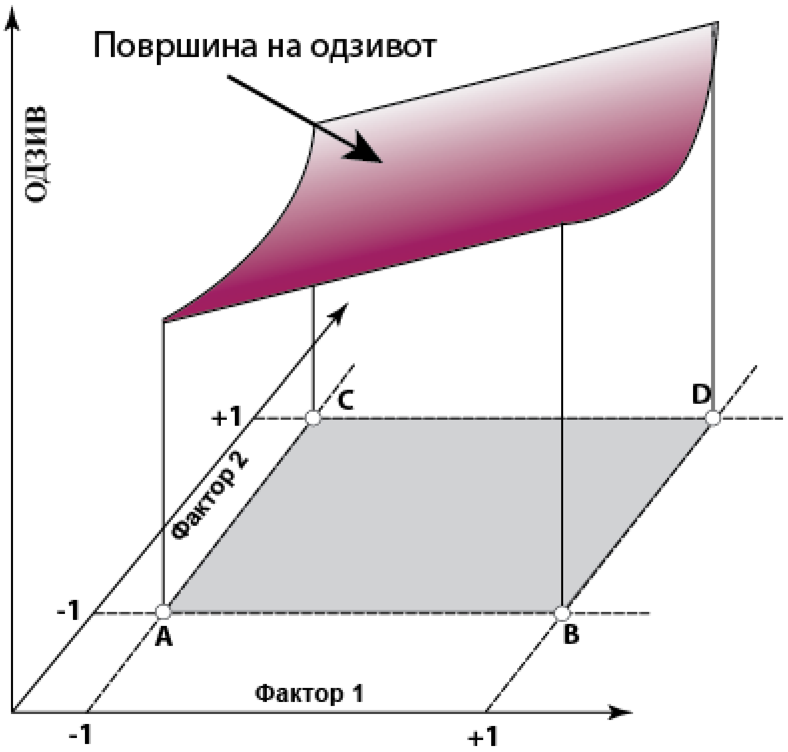
Слика 4.7. Експериментална точка

Ако има ограничувања во однос на планот, треба да креираме посебни планови со наоѓање на позициите на точките кои доведуваат, на ист начин, до квалитетна статистика и добро моделирање на одзивот.

4.7.1. Површина на одзивот

Секоја точка во доменот на проучување има свој одзив. Заедно сите точки во доменот на проучување имаат множество на одзиви кои се лоцирани на некоја површина. Оваа површина ја нарекуваме ***површина на одзивот***, слика 4.8.

Генерално, само неколку одзиви се познати: оние кои одговараат на експерименталните точки избрани од експериментаторот. Да се добие површината на одзивот неопходно е да се изврши интерполација користејќи математички модел.



Слика 4.8. Површина на одзивот

Оние точки кои се избрани со користење на планот на експериментот овозможуваат најдобро можна прецизност на формата и позицијата на површината на одзивот.

4.7.2. Математички модел на одзивот

При изведувањето на некој експеримент ние намерно менуваме една или повеќе процесни варијабли (или фактори) со цел да видиме како ќе се одразат врз една или повеќе одзивни варијабли. Статистички осмислените експерименти претставуваат ефикасна постапка за планирање на експериментите, така што добиените податоци можат да бидат анализирани за да доведат до валидни и објективни заклучоци.

Планирањето на експериментите почнува со определување на целите на експериментот и со избор на процесните фактори кои ќе се проучуваат. Планирањето на експериментите е како нацрт т.е. проект на детално испланирани експериментални постапки пред да се почне со самиот експеримент. Добро испланираните експерименти ја максимизираат „количината“ на добиени информации за дадена „количина“ на експериментална работа.

Планираниот експеримент ги максимизира информациите и од економска гледна точка.

Статистичката теорија, која е во основата на планирањето на експериментите, почнува со *моделите на процесот*. Вообичаено е да се почне со процесниот модел од типот „црна кутија“ со неколку дискретни или континуирани инпут фактори кои можат да бидат контролирани т.е. варирани по желба од страна на експериментаторот, и еден или повеќе мерени аутпут одзиви. Се зема дека аутпут одзивите се континуирани. Експерименталните податоци се користат за да се извлече емпириски (апроксимативен или приближен) модел кој ги поврзува аутпутите со инпутите. Овие емпириски модели обично содржат членови од прв и втор ред.

Често, експериментот треба да се направи и со мноштво на неконтролирани фактори, кои можат да бидат дискретни (тука спаѓаат различните машини, опрема или оператори), и/или континуирани како што се амбиенталната температура или влажност.

Најчестите емпириски модели за експерименталните податоци имаат или *линеарна* или *квадратна* форма.

*Линеарниот модел,*со два фактори како што веќе претходно дискутиравме може да се напише како:

$$y = a_0 + a_1x_1 + a_2x_2 + a_{12}x_1x_2 + \textit{експериментална грешка}$$

Овде, y е одзивот за дадените нивоа на главните ефекти x_1 и x_2 и членот $x_1\, x_2$, кој е вклучен за да го земе во предвид ефектот од можната интеракција меѓу x_1 и x_2. Константата a_0 е всушност одзивот y, кога двата главни фактори се нула.

За посложен процес каде има три главни фактори x_1, x_2, x_3 и еден одзив, y, линеарниот модел изгледа вака:

$$y = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + a_{12}x_1x_2 + a_{13}x_1x_3 + a_{23}x_2x_3 + a_{123}x_1x_2x_3 +$$
$$+ \textit{експерименталната грешка}$$

Трите члена, каде има само по еден x (a_1x_1, a_1x_2, a_1x_3), се членови на главните ефекти, потоа има три члена со двострана интеракција ($a_{12}x_1x_1$, $a_{13}x_1x_3$, $a_{23}x_2x_3$) и еден член со тространа интеракција, последниот ($a_{123}x_1x_2x_3$), кој често се отфрла за да се упрости моделот е интеракцијата на сите три фактори. Кога експерименталните податоци се анализираат сите непознати a параметри се проценуваат (одредуваат) а коефициентите на членовите што се со по еден x (x_1, x_2, x_3), се тестираат за да се одреди кои од нив значително се разликуваат од нула.

Квадратниот модел (или модел од втор ред) не го вклучува членот со тространата интеракција, но затоа има три додатни членови:

$$a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2$$

Овој модел главно се користи во експериментите кои се нарекуваат *експерименти на одзивната површина (response surface)* и кај кои се претпоставува накривеност на одзивната површина..

За да го апроксимираме одзивот во нашиот пример, вршиме негово математичко моделирање со помош на линеарен (од прв ред) модел.

$$y = a_0 + a_1x_1 + a_2x_2 + a_{12}x_1x_2$$

Каде што:

$\mathbf{y}$ – е одзивот или потрошувачката на гориво

$\boldsymbol{x_1}$ и $\boldsymbol{x_2}$ се факторите брзина и додатна тежина, соодветно

$\boldsymbol{a_0, a_1, a_2, a_{12}}$, - се полиномни коефициенти.

Овој модел се нарекува ***априори*** модел или ***постулиран*** (претпоставен) модел. Предодредените модели се валидни претпоставени модели во доменот на проучување кои секогаш треба прецизно да се воспостават. Тоа не се теоретски модели базирани на физичко-хемиски или механички законитости туку на предубедување.

4.8. Реализација на експериментот

4.8.1. Вовед

Со стекнатите знаења од претходната дискусија можеме да пристапиме кон планирање на првиот наш експеримент. Ќе го следиме принципот чии фази се:

- Подготовка за проучувањето на проблемот
- Избор на план за експериментот
- Експериментирање (изведување на експериментот)
- Интерпретација на резултатите
- Одлука да се прекине или продолжи со проучувањето

Секој од овие чекори ќе ги разгледаме поединечно користејќи го примерот за потрошувачката на гориво на колата што се движи со различни брзини и различно оптоварување.

4.8.2. Подготовка на нашиот експеримент

Оваа фаза може да се раздели во неколку чекори од кои најважните се објаснети во продолжение.

a. Дефинирање на целта на проучувањето

Дефинирањето на целта од нашиот пример е сосема јасна. Целта е да се дознае потрошувачката на гориво на кола кога се вози брзо или бавно со - или без екстра товар.

b. Опис на предметот на експериментот

Експерименталните проби на проучувањето се вршени врз една единствена кола. Колата, само со возачот, е возена по автопат, иако, генерално, е многу полесно да се вршат експериментални проби на патека, во лабораторија, во прототипска работилница или на кое било друго место, кое е специјално подготвено за експериментирање.

c. Избор на соодветен одзив

Одзивот во нашиов пример е потрошувачката на гориво на 100 километри поминат пат.

d. Истражување на факторите што можат да влијаат врз одзивот

Кои се факторите кои влијаат врз потрошувачката на гориво?

Тоа се, сѐ разбира, двата фактори кои ќе ги проучуваме: додатната тежина и брзината на колата. Меѓутоа, исто така, постојат и други фактори. На пример, типот на колата и притисокот во гумите, присуство/отсуство на багажник на покривот, правецот на ветрот и дождот, штелованоста на моторот, бројот на сопирања и тргнувања, и квалитетот на патот. Препорачливо е да се запишат сите можни фактори. За да се комплетира листата можно е да се изврши посета на автомеханичар кој можеби ќе наведе уште некој фактор кој може да влијае врз потрошувачката на гориво. Секогаш треба да настојуваме комплетно да го проучиме проблемот. Не сме принудени да го проучуваме секој предложен фактор, но сме заинтересирани да дознаеме кои фактори се највлијателни. Факторите што не се проучуваат во планираниот експеримент, генерално, се фиксирани на константно ниво во текот на целиот експеримент.

e. Дефинирање на нивоата на факторите

Прво ги избираме долното и горното ниво на секој фактор.

Брзината не треба да биде многу мала или многу голема бидејќи пробите се правени на автопат и не смее да се премине максимално дозволената граница. Во нашиов пример, малата брзина е 80 km/h, додека големата – 120 km/h.

Тежина на празната кола е дефинирана со сопствената тежина на колата и со тежината на возачот без багаж. Додатната тежина е дефинирана со било која тежина која ја надминува тежината на празната кола. За овој експеримент колата е оптоварена со 300 kg додатна тежина.

Долното ниво на факторот дополнителната тежина е тежината само на колата и возачот (ништо друго) т.е. нула, екстра тежина.

Горното ниво на факторот дополнителна тежина е дефинирано со тежините на колата, возачот и дополнителниот товар од 300 kg кој е рамномерно распореден во колата.

Добро е да се сумираат овие информации како што е направено во табелата 4.1.

Табела 4.1. Фактори и опсег на проучување

Фактор	Долно Ниво (-)	Горно Ниво (+)
Брзина (1)	80 km/h	120 km/h
Додатна тежина (2)	0 kg	300 kg

Оваа табела е важна затоа што заклучоците од проучувањето се валидни само во доменот на проучувањето. Ако ја оптовариме колата со 350 kg и возиме 130 km/h, не би можеле да ги користиме резултатите од ова проучување за да ја предвидиме потрошувачката на гориво.

f. Проверка на ограничувањата

При возењето количеството на гориво во резервоарот се смалува. Според тоа, тежината на возилото во текот на експериментот се менува. Па, потребно е секоја проба да почне со полн резервоар за овој непосакуван фактор драстично да не повлијае врз резултатите.

Исто така, неопходно е времето (температура, ветар, дожд) да остане константно во текот на сите проби. Па најдобро е да се изведат сите проби во ист ден (без ветар и дожд).

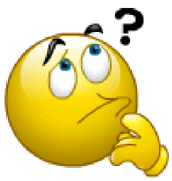
За експериментот е земена иста делница од патот и иста насока за сите проби за падината на патот да не би влијаела врз резултатите. Исто така, пред секој тест, се проверува притисокот во гумите кој треба да е ист за секоја проба.

Сето ова значи, во суштина, преземање на секакви мерки на претпазливост за контрола на експериментот со што се овозможува добивање што е можно поточни резултати.

4.8.3. Избор на план

Имаме два фактори за проучување. Долното и горното ниво за секој фактор веќе ги дефиниравме. Факторите кои треба да останат константни низ целиот експеримент се: растојанието (т.е. бројот на поминати километри), насоката на возење, стартната точка, крајната точка и притисокот во гумите.

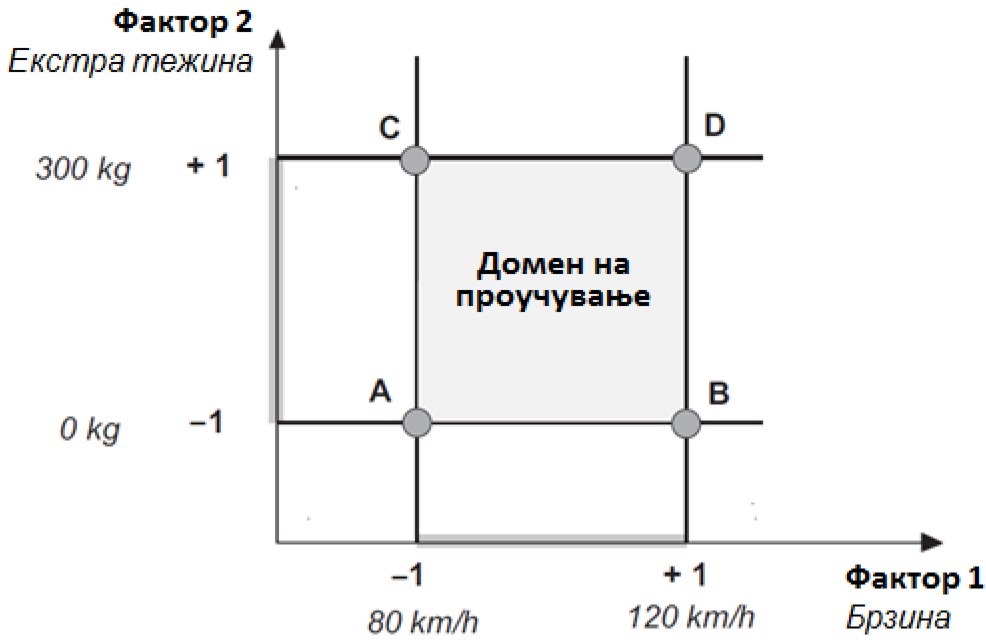
Со оглед дека имаме два фактори, и дека секој има две нивоа, наједноставен план кој можеме да избереме е ***сефакторен*** $\boldsymbol{2^2}$ ***план*** (*full factorial* 2^2 *experimental design*). ***Под „сефакторен“ подразбираме „сите можни фактори“.*** Сефакторниот план на експеримент е само еден од многуте планови кои се користат за планирање на експериментите.



Зошто терминот „сефакторен“?
Овој вид на експеримент е развиен во англиското говорно подрачје (R. Fisher, Англија) и оригиналниот назив му е *full factorial experimental design*. Во преводите на јазиците вон англиско говорно подрачје обично се преведува **буквално** како „полн факторен дизајн/план на експеримент“. Во оригиналниот назив придавката “full” не е употребена во смисла на „полн“, затоа што факторите на можат да бидат полни/празни, туку во смисла на „сите“ (можни фактори). Во англискиот назив “full” е употребен во смисла на *all; all-inclusive; complete* (кои се синоними на full). Од тие причини во овој текст го користиме терминот „сефакторен“ како соодветен превод за „full factorial”. Патем, колоквијално и во нашиот јазик велиме „колата е со full опрема“ во смисла „колата е со сета можна опрема”.

Ознаката $\mathbf{2^2}$ го има следново значење: <u>*експонентот 2*</u> ни го покажува бројот на фактори, додека <u>*2 во основата*</u> ни го покажува бројот на нивото кое секој фактор ги зазема. Овој план е многу прикладен за поставениот проблем бидејќи точно кореспондира на двата избрани фактори кои имаат по две нивоа. Експерименталните точки имаат координати кои се горно и долно ниво на факторите.

Постојат многу, подеднакво добри, начини за претставување на овој план. На пример, можеме прво да го нацртаме доменот на проучување во експерименталниот простор и потоа да ги внесеме експерименталните точки користејќи ги нивните нивоа како координати, слика 4.9.



Слика 4.9. Претставување на планот на експериментот

Експерименталните проби можат да се подредат табеларно користејќи или инженериски единици (km/h и kg) или кодирани единици (- 1 и +1). Табелата со инженериските единици се нарекува ***експериментална табела*** , додека табелата со кодираните единици се нарекува ***план на експериментот*** или ***експериментална матрица***.

Во првата колона на експерименталната табела се назначуваат пробите, табела 4.2. Тие едноставно, можат да бидат нумерирани или да им се даде соодветен назив. Во втората колона се наоѓа првиот фактор со неговите назначени нивоа,подредени по ред. Во трета колона се наоѓа вториот фактор; и во неа експерименталните проби се подредени по ред.. Првата проба , проба 1 или проба A, се извршува со брзина 80 km/h и без додатна тежина. Втората проба, проба 2 или проба B, се извршува со брзина од 120 km/h и без додатна тежина. Третата проба, проба 3 или проба C, се извршува со брзина 80 km/h и додатна тежина од 300 kg. Четвртата проба, проба 4 или проба D, се извршува со брзина 120 km/h и додатна тежина од 300 kg.

Табела 4.2.. Експериментална табела

Проба	**Брзина**	**Екстра тежина**
	Фактор 1	*Фактор 2*
1 (A)	80 km/h	0 kg
2 (B)	120 km/h	0 kg
3 (C)	80 km/h	300 kg
4 (D)	120 km/h	300 kg

Табелата е многу корисно средство во текот на изведувањето на експериментот.

Првата колона кај експерименталната матрица, табела 4.3., се користи за да се покажат називите на пробите. Во втората колона се наоѓа првиот фактор со наведени експериментални нивоа во форма на код т.е. со единици и минус

Табела 4.3. Експериментална матрица

Проба	Брзина	Екстра тежина
	Фактор 1	*Фактор 2*
1 (A)	- 1	- 1
2 (B)	+ 1	- 1
3 (C)	- 1	+ 1
4 (D)	+ 1	+ 1
- 1 Ниво	80 km/h	0 kg
+ 1 Ниво	120 km/h	300 kg

единици. Во третата колона се сите код нивоа од вториот фактор. Последните два реда на дното од табелата го покажуваат значењето на – 1 и + 1 нивоата за секој фактор. Оваа табела се користи при креирањето на планот. За интерпретација на резултатите од пробите може да се користи или експериментална табела или експериментална матрица во зависност кои факти не интересираат.

4.8.4. Изведување на експериментот

Изведувањето на самиот експеримент е техничкиот дел од проучувањето. Пробите треба да бидат внимателно изведени за да се обезбедат квалитетни резултати. Кај овој пример, редоследот на пробите не е многу битен. Која проба прво, второ... ќе ја извршиме теоретски е се исто, но практично не е. Имено, можеме прво да ја изведеме пробата 4 (D) па потоа првата, 1 (A), па третата, 3 (C), па втората (B). Во овој случај треба колата прво да ја натовариме, па да ја растовариме, па повторно да ја натовариме и пак да ја растовариме. За да го избегнеме тоа пробите ќе ги направиме редоследно.

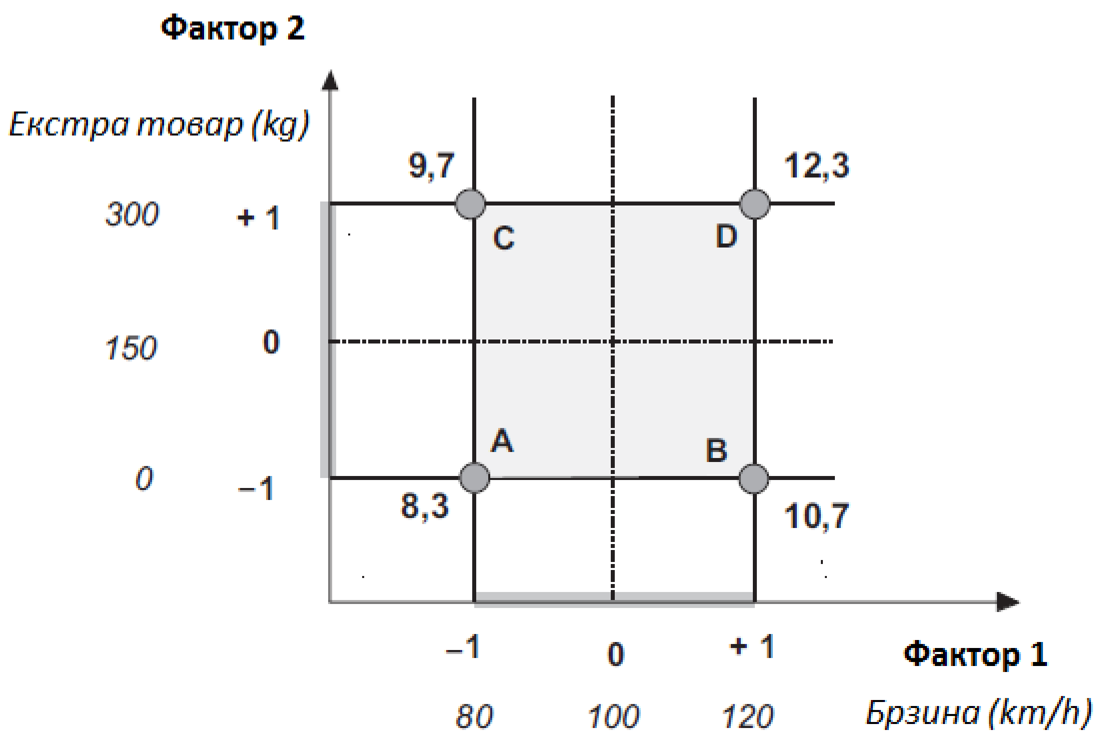
Кај планираните експерименти, не е секогаш се' исто по кој редослед се прават тестовите. Постојат планови на експерименти кај кои строго е одреден редоследот на изведување на тестовите.

Делницата која се вози е долга 120 km. По извршувањето на првиот тест (80 km/h, без додатен товар) ја запишуваме количината на потрошено гориво. За делницата од 120 km се потрошени 10 литри или 8,3 литри за 100 километри [(100/120) x 10]. За вториот тест (120 km/h, без додатна тежина) потрошени се 10,7 литри за 100 километри. За останатите два теста (тест 3 и 4) потрошени се 9,7 литри и 12,3 литри за 100 километри, соодветно. Резултатите ги запишуваме во четвртата колона од експерименталната матрица, табела 4.4.

Табела 4.4. Експериментална матрица со резултати

Проба	**Брзина**	**Екстра тежина**	**Потрошувачка**
	Фактор 1	*Фактор 2*	*l/100km*
1 (A)	- 1	- 1	8,3
2 (B)	+ 1	- 1	10,7
3 (C)	- 1	+ 1	9,7
4 (D)	+ 1	+ 1	12,3
- 1 Ниво	80 km/h	0 kg	
+ 1 Ниво	120 km/h	300 kg	

Ќе ги илустрираме овие резултати и во доменот на проучувањето, како што е презентирано на сликата 4.10.



Слика 4.10. Вредност на одзивот за различни точки во проучуваниот домен

4.9. Значење на коефициентите

Кога ќе ја погледнеме табелата за измерената потрошувачка не можеме, на прв поглед, да најдеме едноставно толкување на резултатите. Би сакале да имаме одговор на прашањата како што се:

1. „Колкава е потрошувачката на возилото ако е оптоварено со 100 kg и ако возиме 90 km/h?“ или
2. „Колкава е дополнителната потрошувачка кога возиме 100 km/h наместо 80 km/h со екстра товар од 150 kg?“

Одговори на овие прашања секако дека постојат. Одговорите всушност се „скриени“ во резултатите кои ги добивме, но во моментов не можеме да ги видиме. Треба да ги трансформираме нашите сурови податоци во разбирливи резултати за да би добиле јасни и прецизни одговори. Тоа е суштината на фазата на нивната интерпретација.

За 2^2 сефакторен експеримент претпоставениот модел е:

$$y = a_0 + a_1x_1 + a_2x_2 + a_{12}x_1x_2 \qquad (2.5.1)$$

каде што:

- y - е одзивот т.е. потрошувачката на гориво во нашиов експеримент.
- x_1 - го претставува нивото на факторот брзина (фактор 1), 80 km/h (- 1) или 120 km/h (+ 1), за нашиот експеримент.
- x_2 - го претставува нивото на факторот додатна тежина (фактор 2), 0 kg (- 1) или 300 kg (+1), за нашиот експеримент .
- x_1x_2 - е продукт на нивоата на факторите брзина и додатна тежина (т.е. *фактор 1* x *фактор 2*). Во нашиов пример, со кодирани единици, овој продукт е еднаков на -1 [x_1x_2 = - 1 x (+1) = +1 x (- 1) = - 1] или +1 [(x_1x_2 = - 1 x (- 1) = (+ 1 x (+ 1) = +1].
- a_0 - е константен коефициент на моделот.
- a_1 – е коефициент на факторот брзина (фактор 1).
- a_2 – е коефициент на факторот додатна тежина (фактор 2).
- a_{12} – е коефициент на членот x_1x_2 (интеракцијата меѓу факторите)

Овој модел се нарекува ***модел од прв ред со интеракција***, кај кој значењето на секој член посебно ќе го разгледаме.

4.9.1. Значење на константниот коефициент

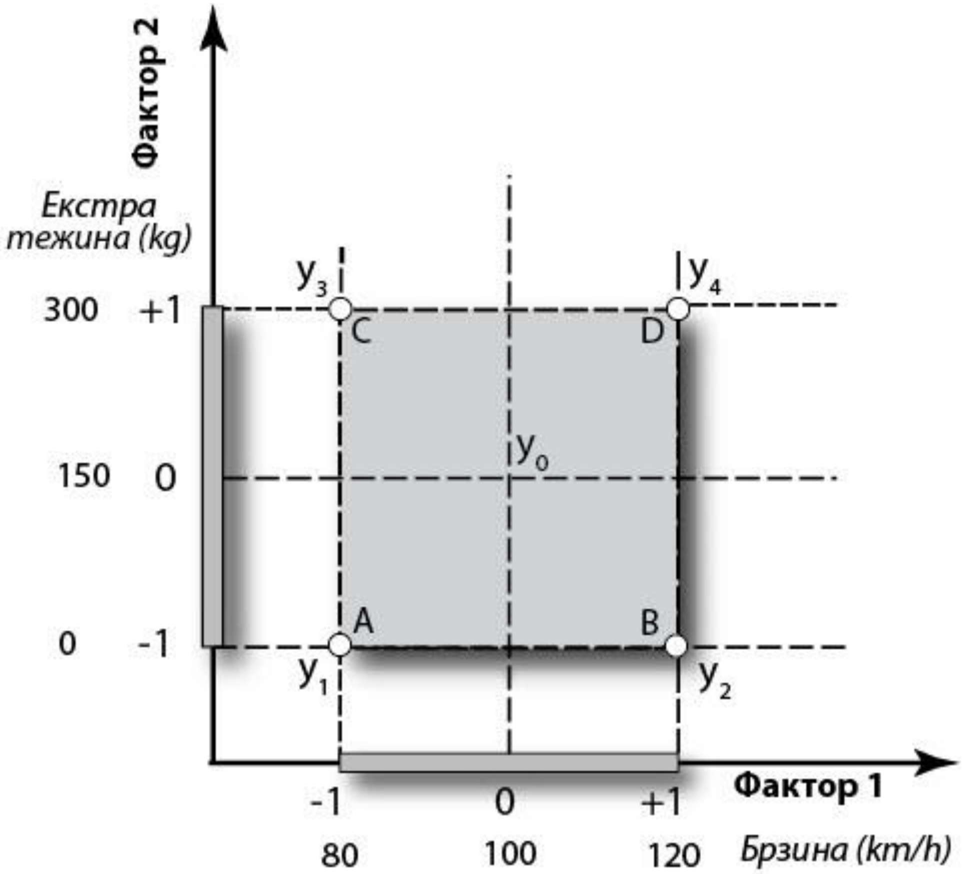
За да се најде значењето на константниот коефициент a_0 , на двата фактори едноставно им доделуваме вредност 0 (во кодирани единици) како ниво. Оваа репрезентативна експериментална точка тогаш се совпаѓа со центарот на проучуваниот домен, слика 4.11, а одзивот во оваа точка има вредност која ја означуваме со y_0 .

Равенката 2.5.1 станува:

$$y = a_0 + a_1 \cdot 0 + a_2 \cdot 0 + a_{12} \cdot 0 \cdot 0$$

$$y_0 = a_0 \tag{2.5.2}$$

Вредноста на коефициентот a_0 е еднаква на одзивот во центарот на доменот на проучување.



Слика 4.11. Вредноста на одзивот во пресечната точка е во центарот на доменот

4.9.2. Значење на коефициентот на брзината (Фактор 1)

Да ги разгледаме двете точки, B и D, кои се наоѓаат на горното ниво од факторот 1 (брзината), прикажани на сликата 4.11. Координатите, во кодирани единици, на овие точки се:

$$\mathrm{B}\left|\begin{matrix} x_1 = +1 \\ x_2 = -1 \end{matrix}\right. \qquad \mathrm{D}\left|\begin{matrix} x_1 = +1 \\ x_2 = +1 \end{matrix}\right.$$

Одзивот во точката B, означен со y_2 ,ќе го добиеме користејќи ги нивоата на соодветните фактори во нивните кодирани единици:

$$y_2 = a_0 + a_1 \times (+1) + a_2 \times (-1) + a_{12} \times (+1) \times (-1) = a_0 + a_1 - a_2 - a_{12}$$

Слично, ќе го добиеме одзивот во точката D, означен со y_4 , користејќи ги нивоата на соодветните фактори во нивните кодирани единици:

$$y_4 = a_0 + a_1 \times (+1) + a_2 \times (+1) + a_{12} \times (+1) \times (+1) = a_0 + a_1 + a_2 + a_{12}$$

Ако ги собереме двата одзиви y_2 и y_4 ќе добиеме:

$$y_2 + y_4 = 2(a_0 + a_1) \tag{4.1}$$

Следно, ја повторуваме истата постапка за точките A и C т.е. за долното ниво на факторот брзина, за кое одзивот е y_1 и y_3, соодветно.
Тоа дава:

$$y_1 + y_3 = 2(+a_0 - a_1) \tag{4.2}$$

Ако од равенката (4.1) ја извадиме равенката (4.2) ќе добиеме:

$$4a_1 = -y_1 + y_2 - y_3 + y_4 \tag{4.3}$$

Оваа равенка можеме да ја напишеме како:

$$a_1 = \frac{1}{2}\left[\frac{y_2 + y_4}{2} - \frac{y_1 + y_3}{2}\right] \tag{4.4}$$

Од оваа равенка можеме да видиме дека:

$$\frac{y_2 + y_4}{2}$$

е средна вредност од одзивот при горното ниво на факторот брзина. Оваа средна вредност ќе ја означиме со $\bar{y}_+$,

додека изразот

$$\frac{y_1 + y_3}{2}$$

е средна вредност на одзивот при долното ниво на факторот брзина, која ќе ја означиме со $\bar{y}_-$. Ако направиме замена во равенка 4.4 ќе добиеме:

$$a_1 = \frac{1}{2}\left[\bar{y}_+ - \bar{y}_-\right] \tag{4.5}$$

Според тоа, коефициентот a_1 претставува една половина од разликата на средната вредност на одзивот на факторот 1 при горното ниво и средната вредност на одзивот на долното ниво од истиот фактор 1. Со промена од долно на горно ниво, одзивот варира, во просек, како разликата

$\left[\bar{y}_+ - \bar{y}_-\right]$ Ако оваа разлика е голема, одзивот прилично многу варира и обратно, ако оваа разлика е мала, одзивот малку варира. Според тоа, ова ни овозможува да знаеме како варира одзивот во зависност од факторот 1. Тоа е и причината зошто коефициентот a_1 е наречен ***ефект на факторот 1.***

Примена

Бидејќи ги знаеме четирите одзиви, лесно можеме да го пресметаме коефициентот a_1 :

$$a_1 = \frac{1}{4}[-y_1 + y_2 - y_3 + y_4]$$

$$a_1 = \frac{1}{4}[-8{,}3 + 10{,}7 - 9{,}7 + 12{,}3] = 1{,}25$$

Ефектот на брзината (ефектот од факторот 1) е 1,25 литри на 100 километри. Тоа значи дека ако брзината се промени од 80 km/h на 100 km/h , потрошувачката се зголемува, во просек, за 1,25 литри на 100 километри . Ако брзината се промени од 80 km/h на 120 km/h потрошувачката се зголемува за 2,5 литри на 100 километри.

4.9.3. Значење на коефициентот на факторот 1

Средната вредност на одзивите на горното ниво на брзината, $\bar{y}_+$, лежи на површината од одзивот вертикално над точката **M₊**, во средината на сегментот BD, слика 4.12, и затоа ги има координатите:

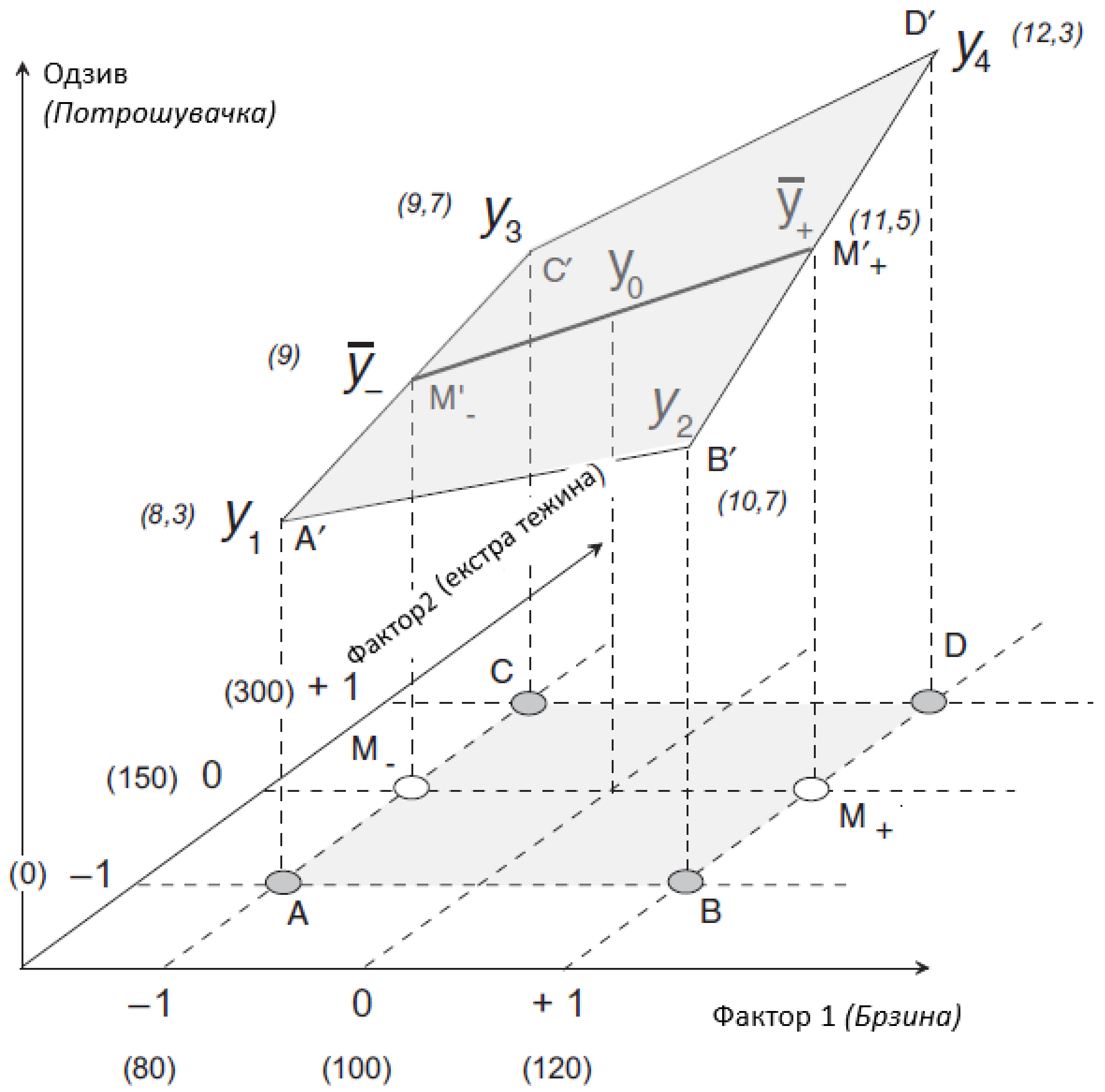
$$M_+ \left| \begin{array}{l} x_1 = +1 \\ x_2 = 0 \end{array} \right.$$

Средната вредност на одзивите на долното ниво од факторот 1 лежи на површината на одзивот вертикално над точката, **M₋** , во средината на сегментот AC, и затоа ги има координатите:

$$M_- \left| \begin{array}{l} x_1 = -1 \\ x_2 = 0 \end{array} \right.$$

Разликата во одзивот меѓу средната вредност на одзивите на горно ниво на факторот 1, $\bar{y}_+$, и средната вредности на одзивите на долно ниво на истиот фактор 1, $\bar{y}_-$, е $\bar{y}_+ - \bar{y}_-$, т.е. е двојно поголема од вредноста на коефициентот a_1.

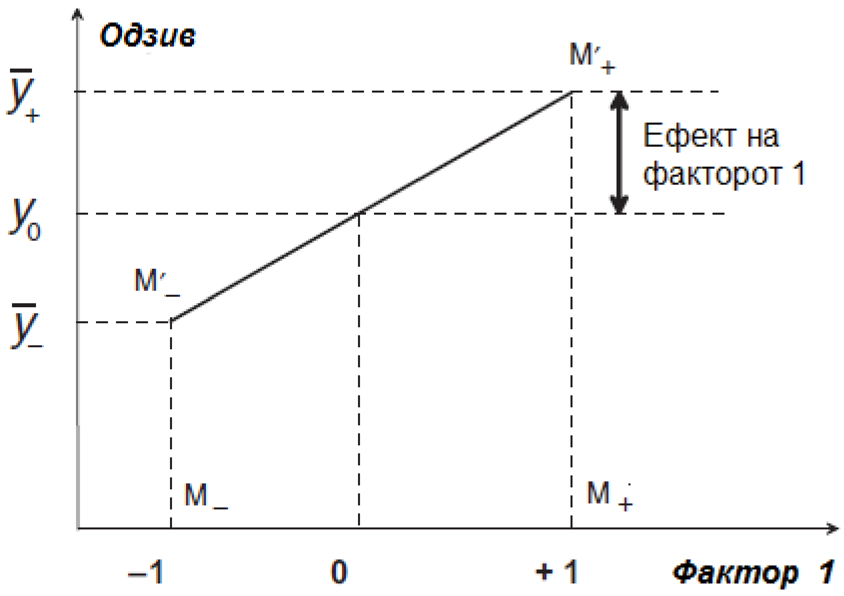
Според тоа, коефициентот a_1 е еднаков на разликата во одзивот меѓу y_0, одзивот во центарот на доменот на проучување , и $\bar{y}_+$, средната вредност на одзивите на горното ниво на факторот 1. Соодветно, коефициентот a_1 можеме да го посматраме како наклон на линијата $M'_- M'_+$. Исто така можеме да речеме дека коефициентот a_1 е еднаков на просечната разлика на одзивите кога факторот 1 преминува од ниво нула на горно ниво. На тој начин, тој го претставува влијанието на брзината во доменот на проучувањето.



Слика 4.12. Коефициентот на факторот брзина е претставен со падината на линијата која ги поврзува одзивите $\bar{y}_+ и\ \bar{y}_-$

.4.9.4. Илустрација на ефектот на брзината

Со цел јасно да ја видиме линијата $M'_-M'_+$ и да го илустрираме ефектот на факторот 1, ја екстрахираме рамнината $M'_-M'_+M_+M_-$ од сликата 4.12. Тоа ни ја дава сликата 4.13 која е многу полесна за сфаќање и интерпретирање.



Слика 4.13. Илустрација на ефектот на факторот 1

Ефектот на факторот 1 може да биде позитивен или негативен, исто како и предзнакот на коефициентот.

Примена

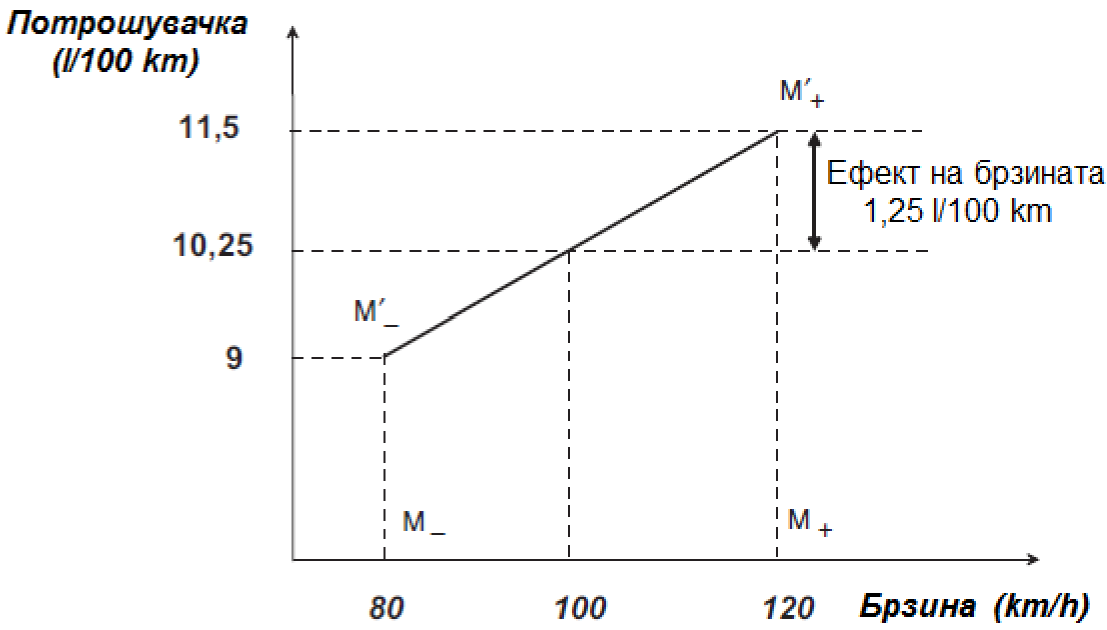
Средната потрошувачка на горното ниво од брзината (фактор 1) е:

$$\bar{y}_{+} = \frac{y_2 + y_4}{2} = \frac{10{,}7 + 12{,}3}{2} = 11{,}5$$

Средната потрошувачка на долното ниво од брзината (фактор 1) е:

$$\bar{y}_{-} = \frac{y_1 + y_3}{2} = \frac{8{,}3 + 9{,}7}{2} = 9$$

Потрошувачката се менува во просек од 9 на 11,5 литри на 100 километри, кога брзината ќе се зголеми од 80 km/h на 120 km/h. Потрошувачката во центарот е половина од двете средни вредности т.е. 10,25 литри на 100 километри. Потрошувачката преминува од 10,25 на 11,5 литри на 100 километри кога брзината се менува од 100 на 120 km/h. Ова зголемување од 1,25 литри на 100 километри го претставува ефектот на факторот „брзина“, слика 4.14.



Слика 4.14. Илустрација на ефектот на брзината

4.9.5. Значење на коефициентот на факторот 2

На идентичен начин, може да се забележи дека коефициентот a_2 е еднаков на просечната разлика на одзивот кога факторот 2 преминува од ниво 0 кон горното ниво. Тој го претставува влијанието на факторот 2 во доменот на проучување. Него го нарекуваме „ефект на факторот 2".

Кога избраниот модел е полиномен, коефициентите на членовите од прв ред се ефекти на факторите.

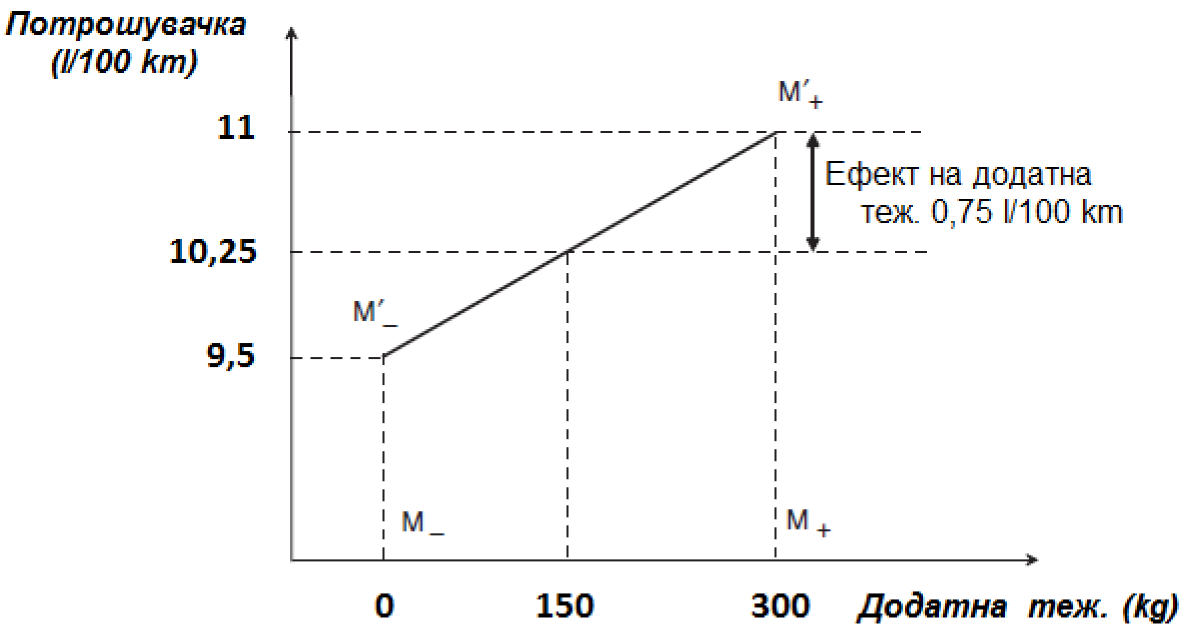
Примена

Знаејќи ги четирите одзиви лесно можеме да го пресметаме коефициентот a_2

$$a_2 = \frac{1}{4}[-y_1 - y_2 + y_3 + y_4].$$

$$a_2 = \frac{1}{4}[-8{,}3 - 10{,}7 + 9{,}7 + 12{,}3] = 0{,}75$$

Ефектот на додатната тежина (фактор 2) е 0,75 литри на 100 километри. Тоа значи дека кога додатната тежина преминува од 0 на 150 kg, потрошувачката просечно се зголемува за 0,75 l на 100 km. Ако додатната тежина премине од 0 на 300 kg потрошувачката се зголемува во просек за 1,5 l на 100 km.



Слика 4.15. Илустрација на ефектот од додатната тежина

Просечната потрошувачка на гориво за горното ниво на додатната тежина (фактор 2) е:

$$\bar{y}_+ = \frac{y_3 + y_4}{2} = \frac{9{,}7 + 12{,}3}{2} = 11$$

Просечната потрошувачка на гориво за долното ниво на додатната тежина (фактор 2) е:

$$\bar{y}_- = \frac{y_1 + y_2}{2} = \frac{8{,}3 + 10{,}7}{2} = 9{,}5$$

Потрошувачката преминува, во просек, од 9,5 на 11 литри на 100 километри кога додатната тежина преминува од 0 на 300 kg. Потрошувачката во центарот е половина од збирот на двете просечни вредности т.е. 10,25 [(9,5 + 11)/2] литри на 100 километри.

Потрошувачката преминува, во просек, од 10,25 на 11 l на 100 km кога додатната тежина преминува од 150 kg на 300 kg. Ова зголемување од 0,75 l на 100 km е ефектот на факторот додатна тежина, слика 4.15.

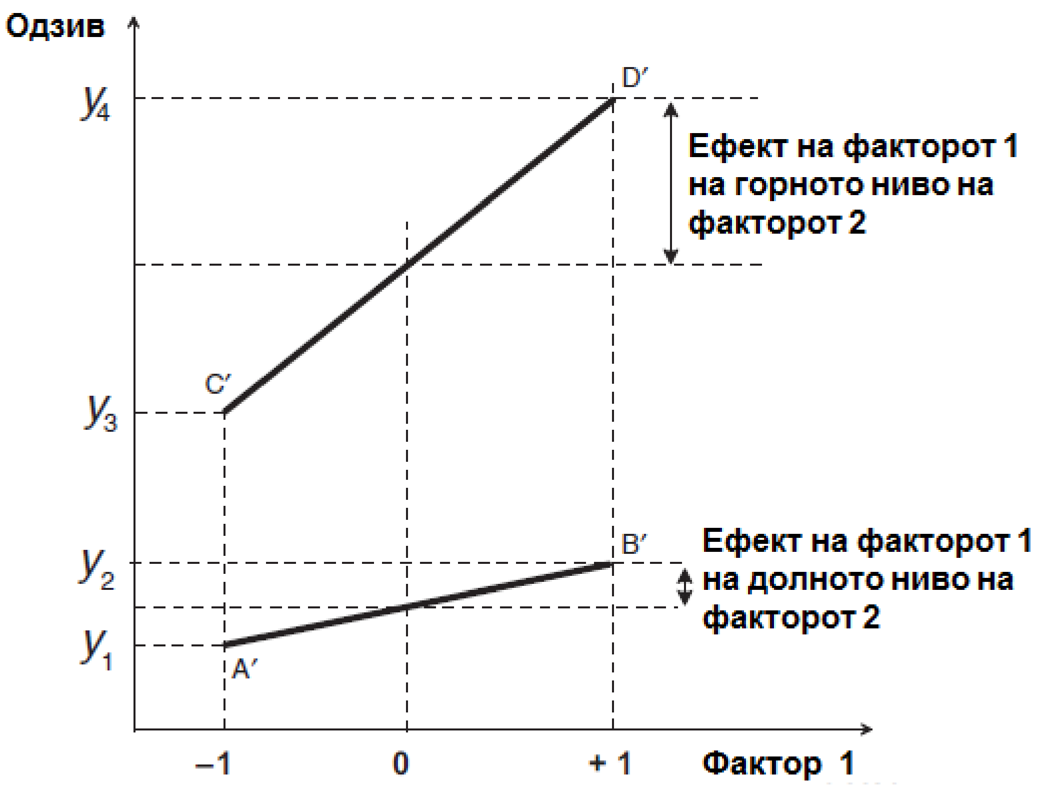
4.9.6. Значење на коефициентот a_{12}

Можеме да го пресметаме коефициентот a_{12} на аналоген начин, како што тоа го направивме за коефициентите a_1 и a_2.

Коефициентот a_{12} е еднаков на:

$$a_{12} = \frac{1}{2}\left[\frac{y_4 - y_3}{2} - \frac{y_2 - y_1}{2}\right]$$

$\frac{y_4 - y_3}{2}$ е ефектот од факторот 1 кога факторот 2 е на горното ниво. Тоа е половината од разликата на одзивите меѓу y_4 и y_3. Овој ефект е илустриран со падина на кривата C'D', слика 4.12 и 4.16.



Слика 4.16. Илустрација на интеракцијата меѓу двата фактори

$\frac{y_2 - y_1}{2}$ е ефектот од факторот 1 кога факторот 2 е на долното ниво. Тоа е половината од разликата на одзивите меѓу y_2 и y_1. Овој ефект е илустриран со падина на кривата A'B', слика 4.16.

Коефициентот a_{12} е половина од разликата меѓу овие два ефекти.

Според тоа, коефициентот a_{12} ја мери промената на ефектот на факторот 1 кога нивото на факторот 2 се менува. Може исто така да се забележи, дека истиот коефициент a_{12} ја мери и разликата во ефектот на факторот 2 кога нивото на факторот 1 се менува на ист начин.

Коефициентот a_{12} се вика ***интеракција*** меѓу факторите 1 и 2.

Интеракцијата меѓу двата фактори може да се илустрира со екстрахирање на рамнината ABA'B' (долното ниво на факторот 2) и CDC'D' (горното ниво на факторот 2) од сликата 4.12 и со проектирање на овие рамнини на иста рамнина, слика 4.16.

Ако постои интеракција меѓу двата фактори наклоните на двете претходни линии нема да бидат исти. Колку повеќе се разликуваат овие наклони дотолку е посилна интеракцијата.

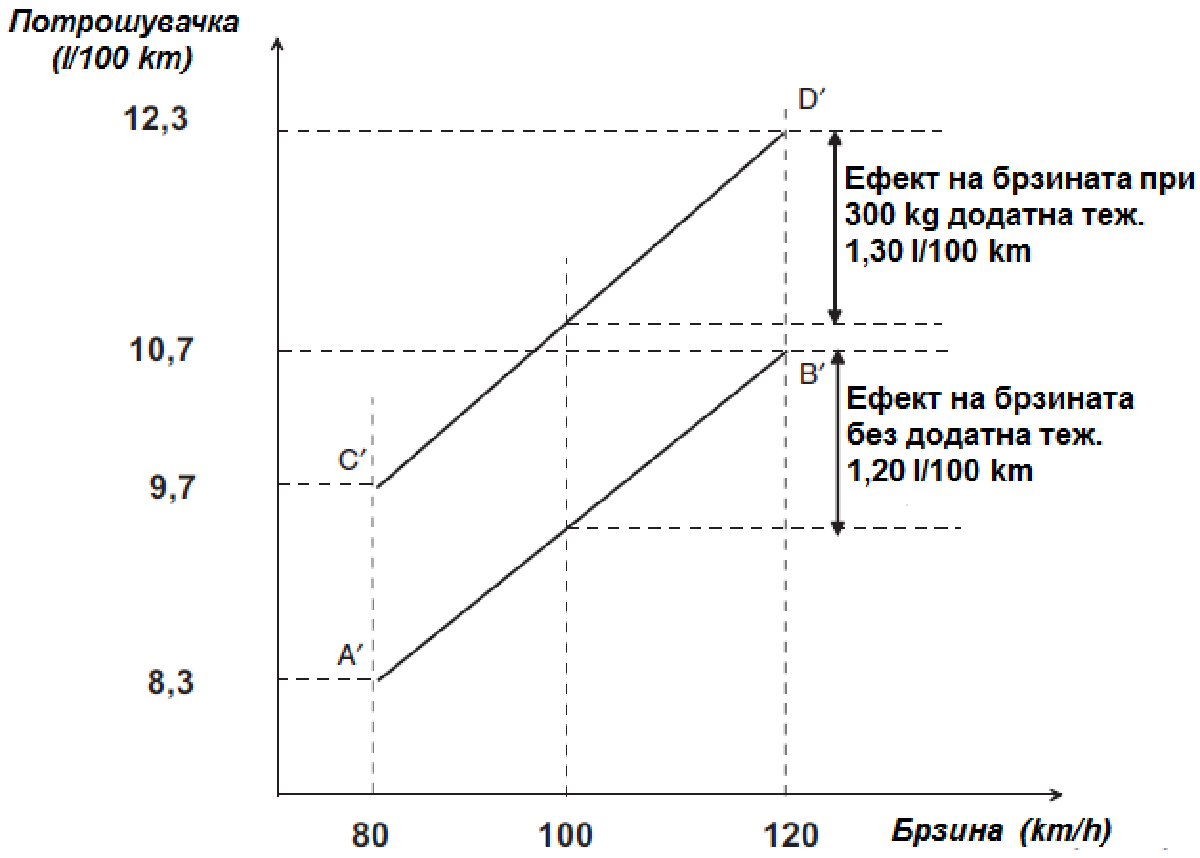
Според тоа, можеме да ја пресметаме интеракцијата преку релацијата:

$$a_{12} = \frac{1}{4}[+y_1 - y_2 - y_3 + y_4]$$

$$a_{12} = \frac{1}{4}[+8,3 - 10,7 - 9,7 + 12,3] = 0,05$$

Интеракцијата меѓу факторите 1 и 2 изнесува 0,05 литри на 100 километри. Тоа покажува дека ефектот на брзината е незнатно поголем кога има додатна тежина. Кога брзината е 80 km/h, ефектот на додатната тежина е 0,7 l на 100 km. Кога брзината е 120 km/h ефектот на додатната тежина е 0,8 l на 100 km.

Тоа, исто така, ни покажува дека ефектот на додатната тежина е позначаен кога се вози побрзо. Кога нема додатна тежина, ефектот на брзината е 1,2 l на 100 km. Кога додатната тежина е 300 kg ефектот на брзината е 1,3 l на 100 km, слика 4.17.



Слика 4.17. Илустрација на интеракцијата меѓу брзината и додатната тежина. Оваа интеракција е многу слаба затоа што двете прави имаат скоро идентична падина

4.10. Интерпретација на резултатите

Сега ни се познати следниве вредности:

- константниот коефициент: a_0 = 10,25
- коефициентот на факторот 1 (брзината): a_1 = 1,25
- коефициентот на факторот 2 (екстра тежина): a_2 = 0,75
- интеракцијата a_{12} меѓу брзината и екстра тежината: a_{12} = 0,05

Овие вредности можеме да ги замениме во равенката 2.5.1 за постулираниот модел:

$$y = 10{,}25 + 1{,}25x_1 + 0{,}75x_2 + 0{,}05x_1x_2 \qquad 4.6$$

Со овој модел, се подразбира дека ќе можеме да ги пресметаме сите одзиви во доменот на проучувањето. Доволно е само да се внесат вредности за нивото x_1 и x_2 и одма да се добие потрошувачката. Бидејќи моделот е во кодирани единици, мораме пресметките да ги извршиме во овие единици и потоа да ги конвертираме добиените резултати во натурални (инженериски) единици. Доколку сакаме директно да пресметуваме во натурални единици потребно е релацијата 4.6 да ја трансформираме користејќи ја конверзионата релацијата:

$$x = \frac{A - A_0}{Интервал}$$

Фактор 1 (брзина):

$$x_1 = \frac{A_1 - A_0}{Интервал_1} = \frac{A_1 - 100}{20}$$

Фактор 2 (екстра тежина):

$$x_2 = \frac{A_2 - A_0}{Интервал_2} = \frac{A_2 - 150}{150}$$

$$y = 10{,}25 + 1{,}25\left(\frac{A_1}{20} - \frac{100}{20}\right) + 0{,}75\left(\frac{A_2}{150} - \frac{150}{150}\right) + 0{,}05\left(\frac{A_1}{20} - \frac{100}{20}\right)\left(\frac{A_2}{150} - \frac{150}{150}\right)$$

$$y = 10{,}25 - \frac{125}{20} - 0{,}75 + \frac{1{,}25A_1}{20} + \frac{0{,}75A_2}{150} + \left(\frac{0{,}05A_1}{20} - \frac{5}{20}\right)\left(\frac{A_2}{150} - 1\right)$$

$$y = 3{,}25 + \frac{5}{20} + \left[\frac{1{,}25}{20} - \frac{0{,}05}{20}\right]A_1 + \left[\frac{0{,}75}{150} - \frac{5}{20 \times 150}\right]A_2 + \frac{0{,}05A_1}{20}\ \frac{A_2}{150}$$

Од тука ја добиваме равенката во натурални единици:

$$y = 3{,}5 + 0{,}06A_1 + 0{,}0033A_2 + 0{,}0000167A_1A_2 \qquad 4.7$$

Сега лесно можеме да одговориме на прашањата кои си ги поставивме во поглавјето 4.9.

Примена

Прашањето гласеше: „Колкава е потрошувачката на возилото ако е оптоварено со 100 kg и ако возиме 90 km/h?“.

Ќе ја примениме равенката 4.7.

$$y = 3{,}5 + 0{,}06 \times 90 + 0{,}00333 \times 100 + 0{,}0000167 \times 90 \times 100$$
$$y = 3{,}5 + 5{,}4 + 0{,}333 + 0{,}150 = 9{,}383$$

Одговорот гласи: Ако возиме 90 km/h со додатна тежина од 100 kg ќе потрошиме 9,38 литри гориво на 100 километри.

Примена

Второто прашањето гласеше: „Колкава е дополнителната потрошувачка кога возиме 100 km/h наместо 80 km/h со екстра товар од 150 kg?“

Ќе почнеме со трансформација на натуралните единици во кодирани:

80 km/h = ниво – 1
100 km/h = ниво 0
150 kg = ниво 0

и со примена на равенката 4.6.

$$y = 10{,}25 + 1{,}25 \times (-1) + 0{,}75 \times (0) + 0{,}05(-1 \times 0)$$
$$y = 10{,}25 - 1{,}25 = 9$$

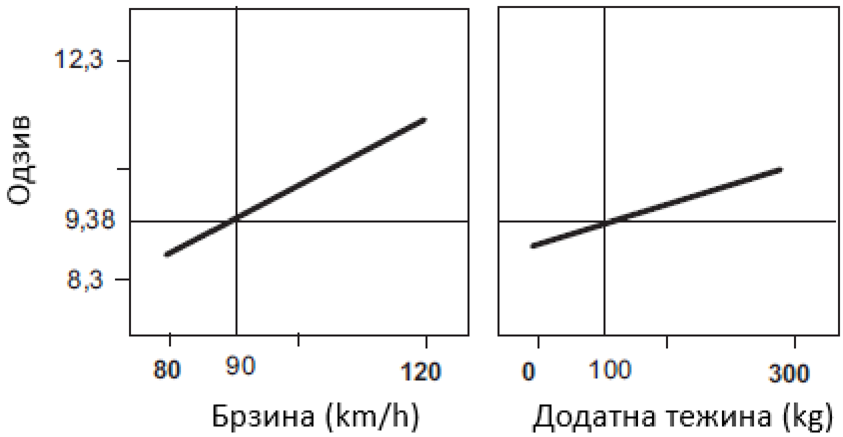
Кога возиме на ниво - 1 (80 km/h) со додатен товар на ниво 0 (150 kg) трошиме 9 литри гориво на 100 километри.

Кога возиме на ниво 0 (100 km/h) со додатен товар на ниво 0 (150 kg) трошиме 10,25 литри гориво на 100 километри. Според тоа одговорот е: со додатен товар од 150 kg дополнителната потрошувачка е 1,25 (10,25 – 9) литри на 100 килограми кога наместо 80 km/h возиме 100 km/h.

$$y = 10{,}25 + 1{,}25 \times (0) + 0{,}75 \times (0) + 0{,}05(0 \times 0)$$
$$y = 10{,}25$$

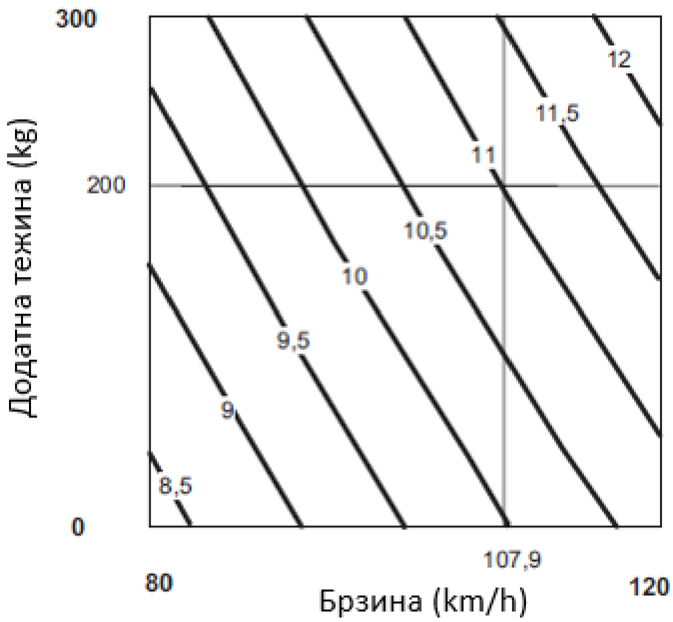
На овој начин можеме да одговориме и на многу други прашања кои опфаќаат различни комбинации на оптоварувања и брзини. Меѓутоа, овие калкулации се репетитивни, тривијални и заморни и не нѐ учат за процесот или системот. Нашите напори треба да бидат насочени во попродуктивни активности одошто во прости калкулации. Затоа силно се препорачува користење на софтвер за дизајнирање на експериментите.

На пример, следниве два дијаграми ја демонстрираат моќта на софтверите. Двата дијаграми се софтверски генерирани. Првиот дијаграм, слика 4.18 го покажува ефектот на двата фактори, брзината и додатната тежина, врз вредноста на одзивот. Овој дијаграм е интерактивен, можеме да избереме кое било ниво во доменот на проучувањето. За секоја - двојка на ниво можеме на ординатата да ја прочитаме вредноста на одзивот. Како што можеме да видиме од него многу лесно и без да правиме долготрајни калкулации можеме да го видиме одговорот на нашето прво прашање: потрошувачката е 9,38 литри на 100 km ако возиме со брзина од 90 km/h со додатна тежина од 100 kg. Овој дијаграм, исто така, ни овозможува да одговориме и на второто прашање.



Слика 4.18. Одредување на одзивот на различни точки во доменот на проучување

Вториот дијаграм, слика 4.19, ни ги покажува контурните линии на **константен одзив** во доменот на проучувањето. На пример, контурната линијата 10 ни ги покажува сите комбинации на брзината и додатната тежина кои одговараат на потрошувачка од 10 литри гориво на 100 km. Од друга страна, ако не сакаме да ја надминеме потрошувачката од 11 литри на 100 km дијаграмот ни покажува како да ја прилагодиме брзината за да го постигнеме тоа. На пример, ако додатната тежина е 200 kg, гледаме, дека брзината не смее да биде поголема од 107,9 km/h.



Слика 4.19. Одредување на максималната брзина за дадена вредност на додатната тежина

4.11. Решени примери

4.11.1. Пример 1:

Да се определи дали температурата и притисокот влијаат врз специфичниот волумен на амонијак, ако температурата се менува во доменот од 40 °C до 140 °C, а притисокот од 500 кPa до 1000 кPa.

Решение :

Од задачата можеме да заклучиме дека имаме два фактори (температура и притисок) и дека секој од нив има по две вредности (долно и горно ниво). Според тоа ќе употребиме сефакторен план за да го изведеме експериментот.

Бројот на проби, *N*, кој треба да го направиме го одредуваме од релацијата:

$$N = n^k$$

каде се:

N = бројот на проби што треба да ги направиме
n = број на нивоа на факторите
k = број на фактори

Во нашиов пример бројот на нивоа на секој фактор е 2, горно и долно; бројот на фактори е исто така 2, притисок и температура; според тоа бројот на проби кои што треба да ги направиме е:

$$N = 2^2 = 4$$

Сефакторниот план на експерименти е таков план кај кој сите можни неповторувачки комбинации на факторите и нивоата се користат во експериментот.

Тоа значи дека се можни 4 различни комбинации на сите фактори при сите нивоа. Тие комбинации се презентирани во матрицата на експериментот, табела 4.5.

Табела 4.5. Матрица на експериментот

Проба	**T, °C**	**P, кPa**	$T_{кодирано}$	$P_{кодирано}$	**V, m³/kg**	**Одзив**
1	40	500	- 1	- 1	0,29	y_1
2	140	500	+ 1	- 1	0,40	y_2
3	40	1000	- 1	+ 1	0,14	y_3
4	140	1000	+ 1	+ 1	0,19	y_4

Во претпоследната колона, ***V***, ни се дадени измерените вредности од експериментот при сите можни различни комбинации на фактори и нивоа.

Избраниот модел, равенка 2.5.1, ни е:

$$y = a_0 + a_1 x_1 + a_2 x_2 + a_{12} x_1 x_2$$

Го наоѓаме коефициентот a_1 од равенката:

$$a_1 = \frac{1}{4}[-y_1 + y_2 - y_3 + y_4]$$

$$a_1 = \frac{1}{4}[-0{,}29 + 0{,}40 - 0{,}14 + 0{,}19] = 0{,}04$$

Го наоѓаме коефициентот a_2 од равенката:

$$a_2 = \frac{1}{4}[-y_1 - y_2 + y_3 + y_4]$$

$$a_2 = \frac{1}{4}[-0{,}29 - 0{,}40 + 0{,}14 + 0{,}19] = -0{,}09$$

Го наоѓаме коефициентот a_{12} од равенката:

$$a_{12} = \frac{1}{4}[+y_1 - y_2 - y_3 + y_4]$$

$$a_{12} = \frac{1}{4}[0{,}29 - 40 - 0{,}14 + 0{,}19] = -0{,}015$$

Го наоѓаме коефициентот a_0 од равенката за y_2 или y_4:

$$y_2 = a_0 + a_1 - a_2 - a_{12}$$

$$0{,}40 = a_0 + 0{,}04 + 0{,}09 + 0{,}015 \Rightarrow a_0 = 0{,}255$$

Ако замениме во равенката 2.5.1, ќе добиеме:

$$y = 0{,}255 + 0{,}04 T_{код.} - 0{,}09 P_{код.} - 0{,}015 T_{код.} P_{код.}$$

Од оваа равенка можеме да заклучиме дека постои влијание и на притисокот и на температурата врз специфичниот волумен (y) на амонијакот.

Како се прави тоа во Excel ?

Користејќи *Excel* многу можеме да си ја олесниме работата при решавањето на овој пример. Еден од пристапите може да биде следниов: Во колоната C ќе ги внесеме ознаките за одзивите и соодветно на нив, во колоната D, нивните вредности, слика 4.18. Во колоната E можеме да ги внесеме ознаките за коефициентите додека во колоната F, соодветно на нив, ќе ги инсертираме формулите за пресметка на коефициентите. На сликата 4.18 за секоја ќелија во колоната F е прикажана инсертираната формула. По инсертирањето на

fx =1/4*(-D1+D2-D3+D4)

C	D	E	F
y1	0,29	a1	0,04
y2	0,40	a2	-0,09
y3	0,14	a12	-0,015
y4	0,19	a0	0,255

fx =1/4*(-D1-D2+D3+D4)

C	D	E	F
y1	0,29	a1	0,04
y2	0,40	a2	-0,09
y3	0,14	a12	-0,015
y4	0,19	a0	0,255

fx =1/4*(D1-D2-D3+D4)

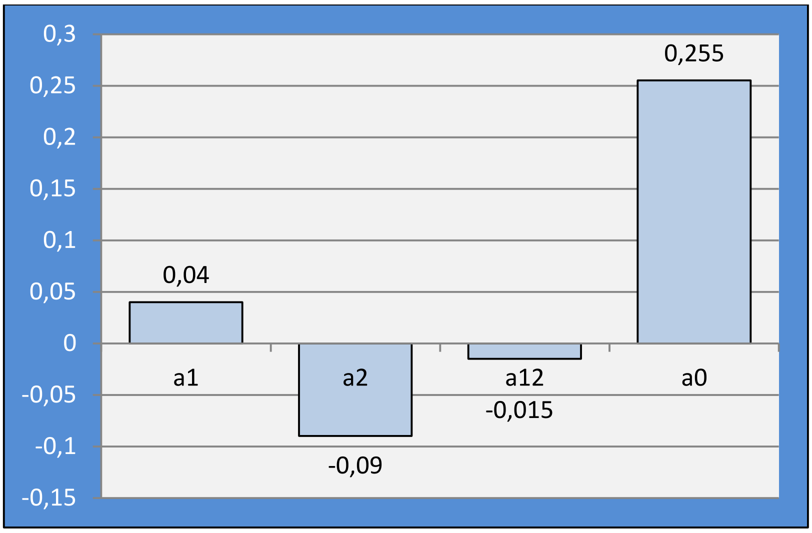
C	D	E	F
y1	0,29	a1	0,04
y2	0,40	a2	-0,09
y3	0,14	a12	-0,015
y4	0,19	a0	0,255

fx =D2-F1+F2+F3

C	D	E	F
y1	0,29	a1	0,04
y2	0,40	a2	-0,09
y3	0,14	a12	-0,015
y4	0,19	a0	0,255

Слика 4.18. Инсертирање на формули за пресметка на коефициентите

формулите, како што може да се виде, програмата автоматски ги внесува соодветните вредности на коефициентите. Ако сакаме да добиеме појасна претстава за влијанието на секој коефициент нив можеме графички да ги претставиме, повторно користејќи *Excel*, слика 4.19.



Слика 4,19. Графичка презентација за влијанието на коефициентите

4.11.2. Пример 2:

Се проучува процесот на дестилација на суровини во апаратура со периодично дејство. Целта на проучувањето е да се одреди зависноста на излезните продукти (во % во однос на внесената суровина во постројката) од брзината на загревање како и од крајната температурата на суровината. Сознанијата за процесот дозволуваат да се воспостават технолошко разумни граници кои можат да ги имаат факторите: крајна температура од 300 °C до 400 °C; брзина на загревање од 4 °C/min до 8 °C/min.

Кодирање на факторите е неопходно за да се конвертираат натуралните фактори (температурата и брзината на загревање) во без димензионални големини за да би можеле да составиме стандардна ортогонална план-матрица на експериментот. За конвертирање на натуралните променливи (варијабли) во кодирани x_i ја пополнуваме табелата за кодирање на променливите на две нивоа, табела 4.6. Како почетно ниво на факторите, обично се избира центарот на интервалот во кој се планира да се изврши експериментот:

Нулто ниво за:

$$x_1 = \frac{300 + 400}{2} = 350^o C$$

$$x_2 = \frac{4 + 8}{2} = 6^o C / \min$$

Табела 4.6. Кодирање на факторите

Интервал на варирање и нивоа на факторите	**Крајна температура °C**	**Брзина на загревање °C/min**
Нулто ниво	350	6
Интервал на варирање	50	2
Долно ниво, $x_i = -1$	300	4
Горно ниво, $x_i = +1$	400	8
Кодирана ознака	x_1	x_2

Во индустриски услови почетното ниво се совпаѓа со вредноста на факторите при нормалниот технолошки режим.

Матрицата за експериментот е дадена во табелата 4.7.

Табела 4.7. Матрица за сефакторен експеримент од типот 2^2

Опит (проба)	x_1	x_2
1	- 1	- 1
2	+ 1	- 1
3	- 1	+ 1
4	+ 1	+ 1

Изборот на интервалот на варирање може да биде сложена задача. Често, при оптимизација на процесите, во почетокот, пригодно е да се опишат со линеарна равенка (како што тоа ние го правиме) и затоа интервалот на варирање треба да биде доволно мал за да се добие линеарна зависност но сепак доволно голем за да не се добие погрешен заклучок за незначително влијание на кој било од факторите.

Релацијата меѓу кодираните и натуралните единици за нивно конвертирање како што ни е познато е:

$$x = \frac{A - A_0}{Интервал}$$

Каде A е натуралната вредност на факторот; A_0 е вредноста на факторот на нулто ниво; „*Интервал*“ е интервалот на варирање на факторот

Табела 4.8. Услови и резултати од пробите

Опит (проба)	x_1	x_2	x_1x_2	y_{j1}	y_{j2}	$\bar{y}_j$ (*одзив*)
1(A)	- 1	- 1	+ 1	27,0	28,0	27,5
2(B)	+ 1	- 1	- 1	15,9	17,1	16,5
3(C)	- 1	+ 1	- 1	22,1	22,9	22,5
4(D)	+ 1	+ 1	+ 1	13,4	13,6	13,5

Во табелата 4.8 се презентирани резултатите од тестовите. Како што може да се види, секој тест е вршен два пати y_{j1} и y_{j2} а во последната колона е презентирана нивната средна вредност, која ја пресметуваме по долната релација и која ни го претставува одзивот за дадените услови на тестот.

$$\bar{y}_j = \frac{y_{j1} + y_{j2}}{2}$$

Табелата 2.7.2.3 претставува проширена матрица на експериментот, затоа што во неа е воведена уште една колона x_1x_2, за интеракцијата меѓу двата фактори. Како и во претходниот пример ги пресметуваме коефициентите на равенката. Добиените вредности се:

$$a_0 = 20\ ;\ \ a_1 = -5\ ;\ \ a_2 = -2\ ;\ \ a_{12} = 0{,}5\ ;$$

Според тоа равенката на одзивот ќе биде:

$$y = 20 - 5x_1 - 2x_2 + 0{,}5x_1x_2$$

Во Excel 2010 табелата од претходниот пример, во колоната D, внесете ги новите вредности за одзивот за автоматски да ги добиете вредностите на коефициентите во колоната F.

4.12. Бенефит од користење на софтвер

Како што видовме од претходните два примери со планираниот експеримент можно е да одговориме на многу прашања, дури и за оние за кои не сме правела проба. Меѓутоа овие калкулации, кои ги применивме, се репетитивни, тривијални и макотрпни и не ни даваат многу додатни информации за самиот процес. Нашите напори можеме да ги искористиме за многу попродуктивни задачи отколку на прости калкулации. Затоа, препорачливо е да се користи софтвер за планирање на експериментите. Сите калкулации кои ги правевме (и многу други) мануелно, всушност, можат лесно и многу побрзо да ги прават компјутерите. Компјутерите, исто така, цртаат графикони и ги појаснуваат и истакнуваат карактеристичните принципи на проучувањето. Тие ја мултиплицираат снагата на експериментаторот за резонирање и вршат калкулации кои се невозможно мануелно да се извршат. Всушност, штом бројот на факторите и нивоа само малку се зголеми, калкулациите стануваат претерано комплицирани. На пример, се вршат експерименти во кои учествуваат и повеќе од 10 фактори. Во такви случаи, тешкотијата произлегува не од теоријата, туку од имплементацијата и контролата на сите нивоа на факторите во текот на експериментот. Ако, на пример, имаме 11 фактори тогаш 110 нивоа треба внимателно да се контролираат за секоја проба и замислете колку напор е потребен да се проучи секој резултат. Со 12 фактори се јавуваат 132 нивоа. При вакви обемни проучувања мала грешка на било кое од овие нивоа може да го компромитира проучувањето. Сето тоа може лесно да се реши со користење на компјутерска програма за планирање на експерименти.

На примерот 2 (точка 4.11 од претходното поглавје) ќе покажеме како поставениот проблем можеме да го решиме со помош на компјутерска програма следејќи ги горните чекори. При тоа програмата ни овозможува додатна контрола врз експериментот и филтрација на податоците.

Програмата која е презентирана во продолжение ги следи следниве чекори при реализација на експериментот:

- Кодирање на факторите
- Составување на план матрица на експериментот
- Рандомизација на опитите
- Реализација на планот од експериментот
- Проверка на репродуктивноста на опитите
- Проценка на значењето на коефициентите на регресија
- Проверка на адекватноста на линеарниот модел

Ова не е процедура по која работи програмата, па ние ја користиме овде за да ја опишеме програмата, туку обратно, тоа е процедура која се следи при реализација на планирањето на експерименти, а програмата е напишана според таа процедура. Значи, користејќи ја програмата, ќе ја проследиме комплетната практична постапка за планирање и анализа на експерименти.

4.12.1. Кодирање на факторите

Кодирање на факторите, како што веќе е речено, е неопходно за да се конвертираат натуралните фактори (температурата и брзината на загревање) во без димензионални величини, за да би можеле да составиме стандардна ортогонална план-матрица на експериментот. За конвертирање на натуралните променливи (варијабли) во кодирани x_i ја пополнуваме табелата за кодирање на променливите на две нивоа, табела 4.9. Како референца за почетно ниво на факторите обично се избира центарот на интервалот во кој се планира да се изврши експериментот. Како што рековме во индустриски услови почетното ниво се совпаѓа со вредноста на факторите при нормалниот технолошки режим.

Табела 4.9. Кодирање на факторите

Интервал на варирање и нивоа на факторите	***Крајна температура °C***	***Брзина на загревање °C/min***
Почетно ниво	350	6
Интервал на варирање	50	2
Долно ниво, $x_i = -1$	300	4
Горно ниво, $x_i = +1$	400	8
Кодирана ознака	x_1	x_2

За изборот на интервалот на варирање ситуацијата е малку посложена. Често, особено при оптималност на процесите, во почетокот, пригодно е да се опишат со линеарна равенка (како што тоа го правиме) и затоа интервалот на варирање треба да биде доволно мал за да се добие линеарно равенство, но истовремено доволно голем за да не се добијат погрешни заклучоци дека влијанието на некој фактор е незначајно

Релацијата за конверзија меѓу кодот и натуралните единици, како што ни е познато, е:

$$x = \frac{A - A_0}{Интервал}$$

Каде A е натуралната вредност на факторот (се означува и со големо X); A_0 е вредноста на факторот на нулто (основно) ниво (се означува и со X_0), и обично се зема дека претставува средина меѓу горното и долното ниво:

$$A_0 = \frac{300 + 400}{2} = 350\ {}^oC$$

„$Интервал$“ е интервалот на варирање на тој фактор. Тој претставува една половина од разликата меѓу горното и долното ниво во инженериски единици. Во нашиов пример интервалот е:

$$Интервал = \frac{400-300}{2} = 50\ ^{o}C$$

4.12.2. Составување на план - матрица на експериментот

Составувањето на план- матрицата на експериментот се имплементира на следниов начин. За:

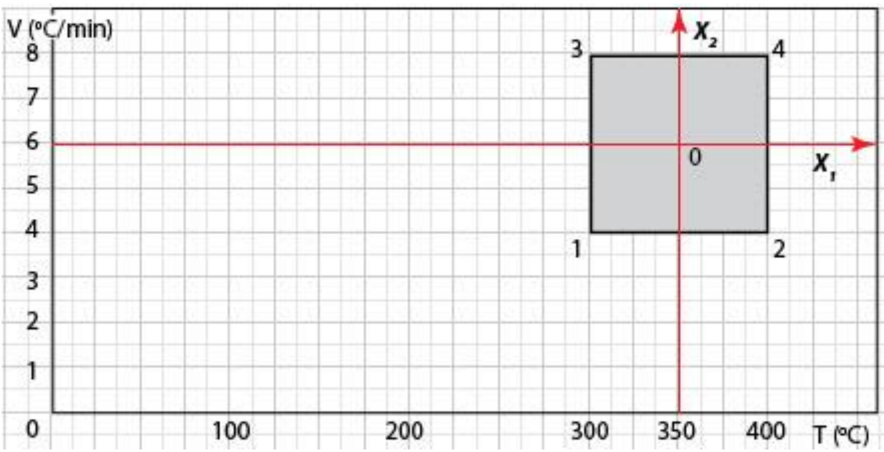
- x_1 - нивоата се менуваат на секој опит (т.е. серија на опити)
- x_2 - нивоата се менуваат на секои два опити
- x_3 - нивоата се менуваат на секои четири опити, итн.

План матрицата за разгледуваниот пример е прикажана во табелата 4.10 и на сликата 4.20.

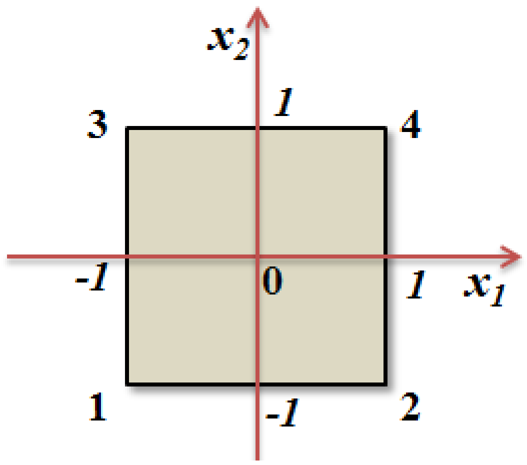
Табела 4.10. Матрица за сефакторен експеримент од типот 2^2

Опит (проба)	x_1	x_2
1	- 1	- 1
2	+ 1	- 1
3	- 1	+ 1
4	+ 1	+ 1

Планот на експериментот презентиран во табелата 2.8.2.1 се нарекува ортогонален план од прв ред. Основна предност на таквиот план е во независната (поединечната или јасно разграничената) проценка на коефициентите на регресија.



(a)



(б)

Слика 4.20. Распоред на опитите (1,2,3,4) во натурални (а) и кодирани променливи (б)

4.12.3. Рандомизација на опитите

Нормално дека на процесот на дестилација покрај x_1 и x_2 вршат влијание и мноштво други фактори кои можат апсолутно да му бидат непознати на експериментаторот. За донекаде да се компензираат систематските грешки на експериментот, се користи метода која се нарекува *рандомизација*. Рандомизацијата се состои во тоа што опитите се вршат по случаен редослед кој се одредува со помош на табели со случајни броеви, прилог 2. На пример, ако треба да рандомизираме 6 опити означени со I, II, ...VI. Соодветно на нив ќе поставиме кои било кои шест последователни броеви земени од која било колона или ред од табелата во прилог 2. Ако при тоа се најдат исти броеви треба да се отстранат. На пример, можат да бидат добиени следниве двојки:

I	II	III	IV	V	VI
60	12	05	15	34	30

Сега ги подредуваме случајните броеви по растечки или паѓачки редослед и ќе го добиеме редоследот на извршување на опитите:

III, II, IV, VI, V, I или **I, V, VI, IV, II, III**

Во нашиов конкретен пример добиен е следниов редослед на извршување на опитите:

Редослед	1	2	3	4	5	6	7	8
Опит	**2**	**3**	**1**	**2**	**4**	**1**	**3**	**4**

Имаме вкупно четири серии на опити и за секоја серија (комбинација на нивоа) вршиме по две испитувања (вкупно осум опити), затоа во вториот ред секоја комбинација на опити ни се јавува по два пати или стручно кажано имаме две ***реплики*** за секоја проба.

Паралелните опити (репликите) се предвидуваат за проценка на репродуктивноста на експериментот и за извршување на статистичка проценка.

4.12.4. Реализација на планот на експериментот

Реализацијата на планот на експериментот е претставена во табелата 4.11. Табелата претставува проширена матрица на експериментот, затоа што во неа е воведена уште една колона x_1x_2, за проценка на интеракцијата меѓу двата фактори.

Табела 4.11. Услови и резултати од пробите

Опит (проба)	x_1	x_2	x_1x_2	y_{j1}	y_{j2}	$\bar{y}_j$ *(одзив)*
1(A)	- 1	- 1	+ 1	27,0	28,0	27,5
2(B)	+ 1	- 1	- 1	15,9	17,1	16,5
3(C)	- 1	+ 1	- 1	22,1	22,9	22,5
4(D)	+ 1	+ 1	+ 1	13,4	13,6	13,5

За секоја серија на проби, т.е. за секоја комбинација на факторите и нивоата вршиме две мерења (две реплики или две паралелни проби) при што добиваме два одзиви, y_{j1} и y_{j2}. Ако го разгледаме пр. првиот опит, двата фактори се на долното ниво (300 °C и 4 °C/min) и при овие услови се извршени две мерења при што се добиени вредности за % на излезните продукти од 27 % и 28 %. Која од овие вредност да ја земеме како валиден одзив при овие услови, дали 27 % или 28 %? Која била вредност, од овие две, да ја земеме ќе направиме грешка. Помала грешката ќе биде ако ја земеме нивната средната вредност т.е. 27.5 %. Токму затоа како валиден одзив ја земаме средната вредност кај сите серии на проби. Неа ја пресметуваме според релацијата:

$$\bar{y}_j = \frac{y_{j1} + y_{j2}}{2}$$

За првата серија добиваме:

$$\bar{y}_1 = \frac{27 + 28}{2} = 27{,}5$$

истото тоа го правиме и за другите серии при што за втората, третата и четвртата серија добивме вредности: 16,5; 22,5; 13,5; соодветно. Во конкретниов пример имаме по две паралелни мерења (проби) за секоја серија на мерења, но и ако се повеќе (препорачливо е да се 5), секогаш ја пресметуваме средната вредност.

4.12.5. Проверка на репродуктивноста на опитите

За проверка на репродуктивноста на опитите ја пресметуваме варијантата (стандардната девијација степенувана на квадрат) за секоја серија на паралелни проби според релацијата:

$$s_j^2 = \frac{1}{k-1}\sum_{i=1}^{k}(y_{ji} - \bar{y}_j)^2$$

Каде што:

S^2_j - варијантата која го карактеризира растурањето (расејувањето) на резултатите од опитите на ***j***-та серија на опити

$i = 1, 2, 3\ldots k$ - е бројот на паралелни опити (реплики)

На пример S^2_j за првата серија ќе биде:

$$s_1^2 = \frac{1}{2-1}\sum_{1}^{2}(27-27{,}5)^2 + (28-27{,}5)^2 = \sum_{1}^{2}(-0{,}5)^2 + 0{,}5^2 = 0{,}5$$

На истиот начин пресметуваме:

$$s_2^2 = 0,72;\;\; s_3^2 = 0,32;\;\; s_4^2 = 0,02$$

за проверка на репродуктивноста на опитите (изводливоста или апликативноста на експериментот) се наоѓа односот на најголемата варијанта и сумата на сите варијанси:

$$G_P = \frac{\max s_j^2}{\sum_{i=1}^{N} s_j^2}$$

N – е бројот на серии т.е. различни експерименти (во нашиов пример, 4)

Оваа величина се нарекува пресметано значење на критериумот на **Кохрен**. Табеларните вредности на ***критериумот на Кохрен*** (*Cochran*), G_T, се дадени на крајот од ова поглавје во прилог 1. Тие одговараат на ниво на доверба (*confidence level*) од $P = 0,95$ со која се применува хипотезата за репродуктивност на опитите.

Треба да се напомне дека величината

$$p = 1 - P$$

се нарекува ***ниво на значајност*** *(significance level)* и во случајов изнесува 0,05.

За наоѓање на табеларната вредност на Кохреновиот критериум неопходно е да се знае вкупниот број на девијации ***N*** (или серии на опити) и бројот на степени на слобода, f, сврзани со секоја од нив, при што:

$$f = N - 1$$

Во некои табели за Кохреновиот критериум може да биде дадена вредноста на f а во други на N. Кога користите таква табела задолжително проверете која вредност е презентирана.

Ако е исполнет условот: $G_P \leq G_T$, опитот се смета за репродуктивен а варијансите за ***еднородни***.

При тоа ***варијансата на репродуктивноста*** или ***грешката на опитите*** се пресметува по равенката за варијансата (стандардната девијација е квадратен корен од варијансата) :

$$s_y^2 = \frac{1}{N}\sum_{j=1}^{N} s_j^2$$

Во нашиов пример варијансата е:

$$s_y^2 = \frac{0{,}50 + 0{,}72 + 0{,}32 + 0{,}02}{4} = 0{,}39$$

Да провериме дали нашиот експеримент е репродуктивен. Од добиените резултати можеме да видиме дека најголемата девијација изнесува 0,72; со замена во равенката за G_P добиваме:

$$Gp = \frac{0{,}72}{0{,}5 + 0{,}72 + 0{,}32 + 0{,}02} = 0{,}4615$$

Табеларната вредност т.е. максимално дозволената вредност за Кохреновиот критериум ја наоѓаме од прилогот 1; за $N = 4$ и $k = 2$, Кохреновиот критериум е $G_T = 0{,}9065$. Бидејќи е исполнет условот $G_P \leq G_T$ нашиот експеримент е репродуктивен и можеме да го продолжиме.

Ако се покаже дека опитите не се репродуктивни треба да се обидеме тие да се направат репродуктивни со идентификување и отстранување за причините за нестабилноста на експериментот, а исто така и со користење на поточни методи и средства за мерење. И конечно, ако на никој начин не е можно да се постигне репродуктивност, тогаш дадените математички методи за планирање на тој експеримент не смеат да се применат.

Коефициентите на регресија ги пресметуваме според претходно презентираните релации:

$$a_1 = \frac{1}{4}[-27{,}5 + 16{,}5 - 22{,}5 + 13{,}5] = -5$$

$$a_2 = \frac{1}{4}[-27{,}5 - 16{,}5 + 22{,}5 + 13{,}5] = -2$$

$$a_{12} = \frac{1}{4}[27{,}5 - 16{,}5 - 22{,}5 + 13{,}5] = 0{,}5$$

$$a_0 = 16{,}5 - 2 + 0{,}5 + 5 = 20$$

4.12.6. Проценка на значењето на коефициентите на регресија

Некои коефициенти можат да се покажат занемарливо мали – незначајни. За да се потврди дали тие се значајни или не, пред се, треба да се процени варијансата при која тие се определуваат.

$$s_a = \sqrt{\frac{s_y^2}{N}}$$

Треба да се напомне дека кај сефакторниот план на експерименти сите коефициенти се определуваат со еднаква погрешност т.е. точност. Се зема дека коефициентите на регресија се значајни ако е исполнет условот:

$$|a| \geq s_a t$$

t – табеларна вредност на Студентовиот (*Student*) критериум, прилог 4.

$$s_a = \sqrt{\frac{0{,}39}{4}} = 0{,}3122$$

претходно пресметавме дека S^2_y = 0,39.

Сега можеме да ја одредиме вредноста за Студентовиот критериум од прилогот 4 за P = 0,95 и *f* = 4, и наоѓаме дека изнесува *t* = 2,78 (со четири децимали вредноста е 2.7764)

Вредноста на степените на слобода ја одредуваме од релацијата:

$$f = N(k-1) = 4(2-1) = 4$$

За нашиот експеримент :

$$|a| \geq s_a t \geq 0{,}3122 \cdot 2{,}7764 \geq 0{,}8668$$

што значи, секој коефициент по апсолутна вредност треба да биде поголем од или еднаков на 0,8668. Тоа ќе го провериме за секој коефициент:

$$|a_0| = 20 > 0{,}8668$$
$$|a_1| = 5 > 0{,}8668$$
$$|a_2| = 2 > 0{,}8668$$
$$|a_{12}| = 0{,}5 < 0{,}8668$$

Како што можеме да видиме коефициентот a_{12} не го задоволува критериумот и се отфрла како незначаен.

Според тоа регресивната равенка со кодирани променливи ќе биде:

$$y = 20 - 5x_1 - 2x_2$$

4.12.7. Проверка на адекватноста на линеарниот модел

За добиената регресивна равенка треба да се провери нејзината адекватност т.е. нејзината способност доволно добро да ја опишува површината на одзивот. Таа проверка се врши со помош на критериумот на Фишер (Fisher), видете прилог 3. Равенката е адекватна ако е исполнет условот:

$$F_P \leq F_T$$

$$F_P = \frac{s^2_{ad}}{s^2_y}$$

$$s_{ad}^2 = \frac{\sum_{j=1}^{N} (\bar{y}_j - y_j)^2}{N-k-1}$$

Во оваа равенка y_j ги означува вредностите на одзивот пресметани според регресивната равенка. На пример, за првата серија, ако ги замениме вредностите за кодираните фактори ќе добиеме:

$$y_1 = 20 - 5x_1 - 2x_2 = 20 - 5(-1) - 2(-1) = 27$$
$$y_2 = 20 - 5x_1 - 2x_2 = 20 - 5(+1) - 2(-1) = 17$$
$$y_3 = 20 - 5x_1 - 2x_2 = 20 - 5(-1) - 2(+1) = 23$$
$$y_4 = 20 - 5x_1 - 2x_2 = 20 - 5(+1) - 2(+1) = 13$$

Сите пресметки се претставени во табелата 4.12.

Табела 4.12. Пресметки за одредување на исполнувањето на Фишеровиот критериум

Опит	$\bar{y}_j$	y_j	$\bar{y}_j - y_j$	$(\bar{y}_j - y_j)^2$
1	27,5	27,0	0,5	0,25
2	16,5	17,0	0,5	0,25
3	22,5	23,0	0,5	0,25
4	13,5	13,0	0,5	0,25

$$\sum_{j=1}^{4} (\bar{y}_j - y_j)^2 = 0{,}25 + 0{,}25 + 0{,}25 + 0{,}25 = 1$$

Бројот на степените на слобода е:

$$N - k - 1 = 4 - 2 - 1 = 1$$
$$s_{ad}^2 = 1$$

Претходно пресметавме дека:

$$s_y^2 = 0{,}39$$

$$F_P = \frac{1}{0{,}39} = 2{,}5641$$

Табеларната вредност на Фишеровиот критериум ја наоѓаме од прилогот 3 за P = 0,95 , f = 1 и N=4, наоѓаме дека Фишеровиот критериум изнесува 7,7086, и бидејќи:

$$F_P \leq F_T$$
$$2{,}5641 < 7{,}7086$$

заклучуваме дека моделот е адекватен.

4.12.8. Одредување на моделот во натурални единици

Ако во равенката

$$y = 20 - 5x_1 - 2x_2$$

замениме:

$$x_1 = \frac{X_1 - 350}{50} \qquad x_2 = \frac{X_2 - 6}{2}$$

ќе ја добиеме:

$$y = 50 - 5\left(\frac{X_1 - 350}{50}\right) - 2\left(\frac{X_2 - 6}{2}\right)$$

$$y = 61 - 0{,}1X_1 - X_2$$

равенката во инженериски единици. Пригодно е равенката во инженериски единици да ја пишуваме со инженериски ознаки. Така на пример ако факторот X_1 (температурата) го означиме со *t* а факторот X_2 (брзината) со *v* равенката ќе ја добиеме во поприфатлива форма:

$$Q = 61 - 0{,}1t - v$$

4.12.9. Проверка на моделот

Користејќи ја равенката со кодирани варијабли ќе пресметаме колкав е приносот на дестилацијата ако е температурата на дестилација 380 ^oC и брзината на загревање – 5,4 ^oC/min.

Прво треба вредноста на овие натурални факторите да ја конвертираме во кодирани:

$$x_1 = \frac{380 - 350}{50} = 0{,}6 \qquad x_2 = \frac{5{,}4 - 6}{2} = -0{,}3$$

Сега можеме да замениме во равенката со кодирани варијабли:

$$y = 20 - 5 \cdot 0{,}6 - 2(-0{,}3) = 17{,}6$$

Доколку моделот е во ред, истата оваа вредност треба да ја добиеме и во натурални варијабли:

$$Q = 61 - 0{,}1 \cdot 380 - 5{,}4 = 17{,}6$$

5. ПРЕЗЕНТАЦИЈА И РАБОТА СО СОФТВЕРОТ FFED 7.5

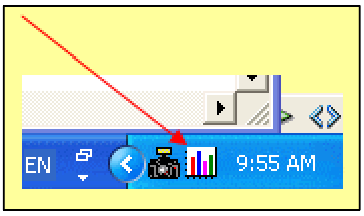
Од сè она што е досега презентирано можеме да се сложиме дека мануелното планирање, изведување, презентирање на резултатите и нивна статистичка обработка е долготрајна и макотрпна постапка. При тоа можноста за грешка во некој резултат е навистина голема и во таков случај потребно е целата процедура да се повтори. Замислете колку многу пресметки треба да направиме кога регресивната равенка од кодирани ќе ја конвертираме во инженериски единици на пример при добивањето на равенката 4.7. Целата постапка е уште подолготрајна и по макотрпна кога работиме со три фактори (видете следно поглавје). Сето тоа ни одзема многу ментална енергија и не дефокусира од главните цели на експериментот – учење на процесот/системот и донесување валидни статистички заклучоци. Иако, со користењето на Microsoft Excel можеме значително да си ја олесниме работата, но и тука постапката е долготрајна и подложна на субјективни грешки. Microsoft Excel е олеснителна алатка, но сепак не е замена за вистинската ствар.

Она што долготрајно и макотрпно мануелно го правевме досега, со користење на софтвер за планирање на експерименти можеме да го завршиме за неколку десетина секунди. Користењето на соодветен софтвер силно се препорачува во наставата по предметот „Дизајн и анализа на експерименти“, затоа што енормно се скратува времето за пресметување, можноста за грешка во калкулациите се елиминира, се заштедува ментална енергија и се овозможува експериментаторот да се фокусира на стекнување знаења за системот или процесот.

Во наставата, предавањата и вежбите, по предметот „Дизајн и анализа на експерименти“, студентите ќе го имаат на располагање некомерцијалниот софтверскиот пакет FFED 7.5, развиен од авторите, кој е прилагоден на процедурите за дизајн и анализа на експерименти презентирани во овој учебник. Со оглед дека студентите ќе го изучуваат и полагаат овој предмет користејќи го овој софтвер, подетално ќе ги презентираме неговите можности и постапката за користење. Можностите и постапката за работа со програмата ќе ги презентираме користејќи го нашиот претходен пример.

5.1. Инсталација на програмата

Зборуваме за инсталација на програмата затоа што таа ќе им биде слободно дадена на студентите од Технолошко-техничкиот факултет и на Универзитетот (затоа што предметот „*Дизајн и анализа на експерименти*" е и универзитетски предмет) да ја инсталираат на домашните компјутери. За оние кои не се во можност да ја користат во домашни услови, програмата ќе биде инсталирана на компјутерите во компјутерската лабораторија на факултетот. Програмата работи во оперативниот систем Windows 9x/XP или понова. Инсталацијата се почнува со активирање на инсталациониот фајл со кој се активира инсталациониот Wizard кој ве води во текот на инсталацијата. По дифолт програмата се инсталира во фолдерот C:\PFE. Ако е инсталацијата успешна по неговото стартување се покажува неговата икона во системската лента на оперативниот систем во десниот долен агол, слика 5.1.

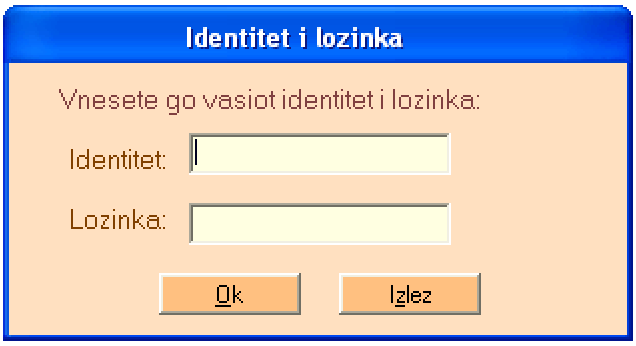


Слика 5.1. Икона на активиранаиата програма

5.2 Стартување на програмата

Програмата ја стартуваме на вообичаен начин како секој друга егзекутивна програма во оперативниот систем Windows. Со двојно кликање со глушецот на иконата од програмата на десктопот; со кликање на копчето *Start* ⇒ *Programs* ⇒ *PFE*; со двојно кликање на егзекутивната датотека *pfe 7.50.exe* во фолдерот каде што програмата е инсталирана.

За да влеземе во програмата потребно е да се логираме со идентитет и лозинка, слика 5.2. Секој студент ќе добие свој идентитет и лозинка за да би можел да ја користи програмата.

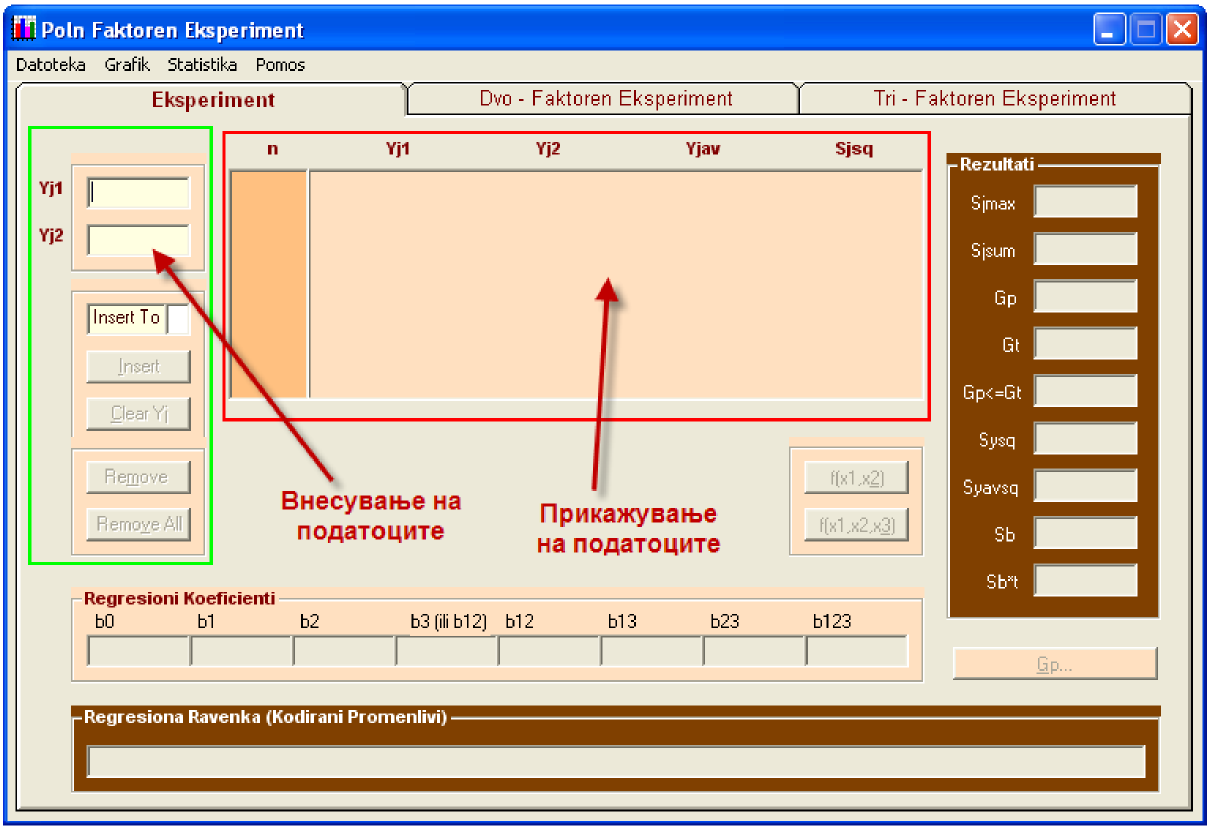


Слика 5.2. Логирање во програмата

По внесувањето на идентитетот и лозинката кликаме на копчето ОК за да влеземе во програмата. Алтернативно, го држиме копчето *Alt* од тастатурата притиснато и истовремено притискаме на типката О од тастатурата (која е подвлечена на називот од копчето). Програмата е овозможена комплетно да работи само со тастатурата, без користење на глушецот. Секое копче го активираме со држење на копчето *Alt* притиснато и истовремено ја притискаме соодветната типка (буква) на тастатурата која е подвлечена во називот на копчето.

5.3. Опис на интерфејсот

По успешното логирање ни се прикажува интерфејсот на програмата, слика 5.3. Интерфејсот е поделен, според функцијата, на неколку целини, групирани во рамки. Зелената рамка служи за внесување на податоците, црвената за нивно прикажување. Во рамката ***Rezultati*** се прикажуваат резултатите од статистичката обработка на податоците, а ни служи и за проверка и прикажување на Кохреновиот коефициент.



Слика 5.3. Интерфејс на програмот PPFE 7.5

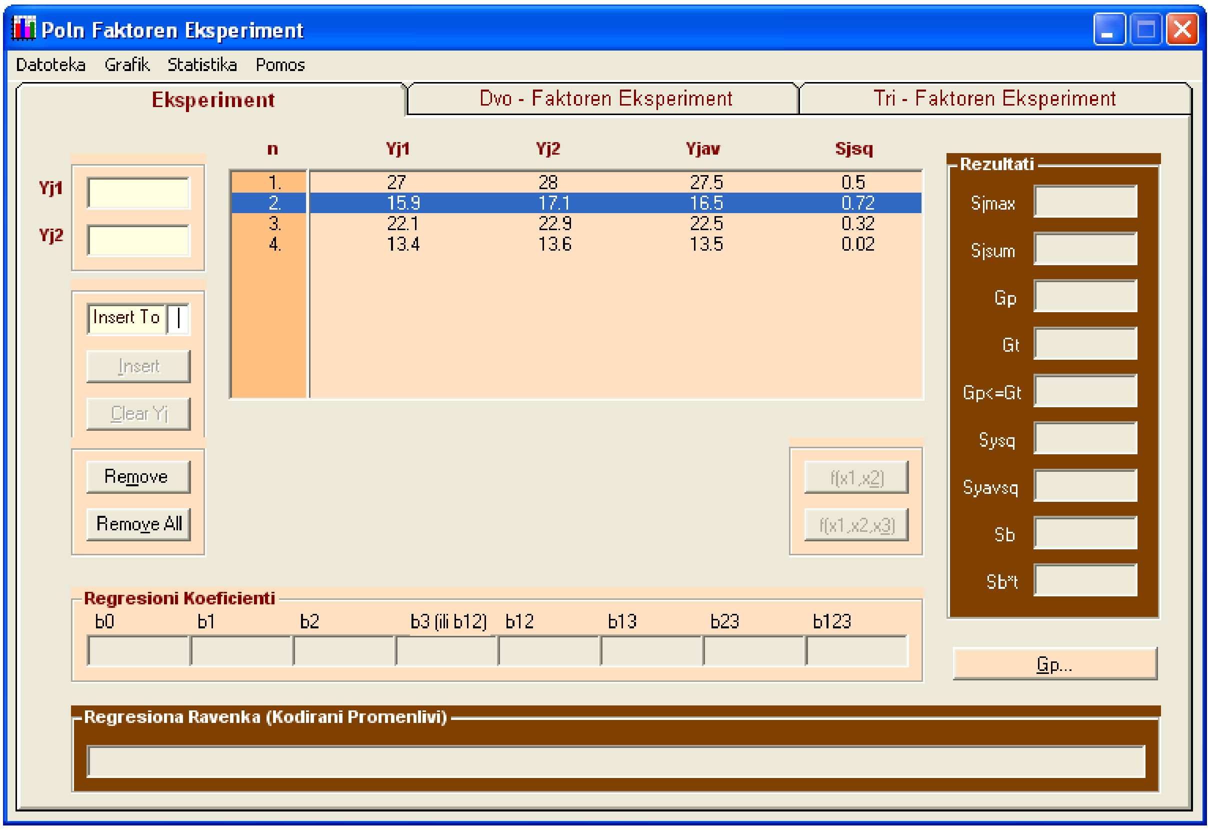
Во рамката ***Regresioni Koeficienti*** ни се прикажуваат регресионите коефициенти од равенката со кодирани варијабли додека во рамката ***Regresiona Ravenka (Kodirani Promenlivi)*** ни се прикажува регресивната

равенка со кодирани варијабли. Копчињата f(x1,x2) и f(x1,x2,x3) ни служат за пресметување и презентирање на регресионите равенки со две и три варијабли соодветно.

Во горниот дел од интерфејсот табовите ***Dvo-Faktoren Eksperiment*** и ***Tri-Faktoren Eksperiment*** ни служат за продолжување на процедурата со експерименти со два и три фактори. Описот на интерфејсот по активирање на овие табови ќе го продолжиме како што ќе продолжиме со актуелниот експеримент кој следи. Табот ***Eksperiment*** е дифолт по стартувањето на програмата, што е логично, затоа можноста за внесување на податоци е изводлива само кога е тој активиран. Бидејќи на почетокот немаме внесено никакви податоци сите копчиња за манипулирање на податоците се неактивни.

5.4. Внесување на податоците од експериментот

По стартувањето на програмата фокусот е во полето Y_{j1} и можеме да почнеме со внесување на податокот, со притискање на тастерот ENTER од тастатурата (по конвенција во овој текст, секаде каде што е употребен терминот „тастер“ се мисли на копче од тастатурата и се користат големи букви за негово означување; терминот „копче“ се однесува на копче од програмата, и го означува со болд фонт) курсорот се преместува во полето Y_{j2} во кое го внесуваме вториот податок, повторно притискаме на ENTER, фокусот оди на копчето **Insert** (кое дури откако ќе се внесат податоци во двете полиња станува активно) и со уште едно притискање на ENTER внесените податоците се прикажуваат во листата (црвен рамка); податоците од полињата Y_{j1} и Y_{j2} се бришат и курсорот повторно оди во полето Y_{j1} за внесување на следниот податок. Ова го повторуваме додека не ги внесеме сите податоци (четири двојки за експеримент со два фактори и осум за експеримент со три фактори).



Слика 5.4. Внесување на податоците од експериментот

По внесување на податок во едното или двете полиња Y_j, копчето **Clear** $\mathbf{Y_j}$ станува активно. Со него можеме да ги избришеме податоците доколку сме погрешиле при внесувањето. Ова е наједноставната и нај брзата постапка (само со тастатурата) за внесување на податоците. Се разбира, внесувањето на податоците можеме да извршиме и со користење на глушецот и/или TAB тастерот за навигација и фокусирање на полињата/копчињата. Исто така бришењето на податоците може да се врши на вообичаениот начин, со селектирање на полето и притискање на тастерот DELETE или со кликање со десниот тастер од глушецот и со избор на DELETE од контекстуалното мени што ќе се појави. Ако погрешно избришеме податок, кликаме со десното копче во полето и од контекстуалното мени избираме UNDO за да го вратиме назад.

Ако забележиме грешка во податоците откако се внесени во листата, неа можеме да ја коригираме. При помала грешка, со глувчето го селектираме редот со погрешните податоци, слика 5.4, и со притискање на копчето **Remove** (кое станува активно, само ако има селектирано податоци во листата) ги отстрануваме. Ако сакаме да ги отстраниме сите податоци од листата го користиме копчето **Remove All**.

Понатамошната презентација на програмата и начинот на користење ќе ја продолжиме користејќи ја програмата за одредување на моделот на процесот од примерот 2 од претходната глава. За да би можеле да извршиме статистичка обработка на резултатите програмата овозможува да се работи со две реплики, полето Y_{j1} и Y_{j2}. Една двојка од реплики ни претставува еден тест (n) во листата.

Програмата овозможува максимална заштита од правење на грешка при неговото користење.

- Во дадена фаза од работата активни се само оние копчиња кои е логично да ги притиснеме за да продолжиме со работа, останатите се неактивни.
- Во дадена фаза од работа, при селектирани податоци, се активираат само копчињата наменети за манипулација со тие податоци (пр. **Remove**)
- Во полињата Y_{j1} и Y_{j2} можат да се внесат *само нумерички вредности и децимална точка* (стрингови т.е. букви не е можно)

На веќе погоре опишаниот начин ги внесуваме податоците од експериментот. Внесените податоци мора да бидат строго според проширената матрица, табела 4.8. Ако не е така, тогаш планираниот експеримент нема смисла. Доколку се погреши во редоследот на внесување на податоците програмата дава можност за корекција. Имено, го селектираме редот со податоци кои се внесени со погрешен редослед, во полето “Insert To” го внесуваме бројот на редот во кој што сакаме податоците да се најдат, и по кликање на копчето “Insert” (кое е неактивно ако нема податок во “Insert To” полето), податоците се сместуваат на саканата позиција, а другите (по дадениот ред) соодветно се поместуваат за едно место. Доколку пак се погреши во вредноста на внесени во листата тогаш редот се селектира и со кликање на копчето “Remove” се отстранува. Со оваа можност за инсертирање на податоци на дадена позиција и/или отстранување, во случај на грешка, избегнуваме одново да ги внесуваме сите податоци. Кај експеримент со два фактори можеби не е толку битна потребата од користење на оваа функција, но кај експерименти со тир фактори, каде што работиме со 16 податоци (8 x 2), секако дека е битна.

На пример, да претпоставиме дека сме внесле погрешни во вториот ред од сликата, овој ред го селектираме, слика 5.4, и со **Remove** го бришеме. Сега редовите кои се под него се поместуваат за едно место нагоре, третиот ред станува втор, а четвртиот – трет. Сега кога ќе ги внесеме точните податоци од вториот ред, по дифолт одат на последната позиција т.е. на четвртиот ред. Овој редослед не е во согласност со матрицата на експериментот. Затоа, е многу корисно користењето на копчето Insert To. По внесувањето на податоците во полињата Y_{j1} и Y_{j2} , во полето **Insert To** пишуваме „2“ со што му „кажуваме“ на програмата да ги внесе податоците во вториот ред (иако таму веќе има податоци). По притискање ENTER програмата ги внесува податоците во второт ред, а претходниот трет и четврт ред ги поместува на старите позиции во склад со матрицата на експериментот.

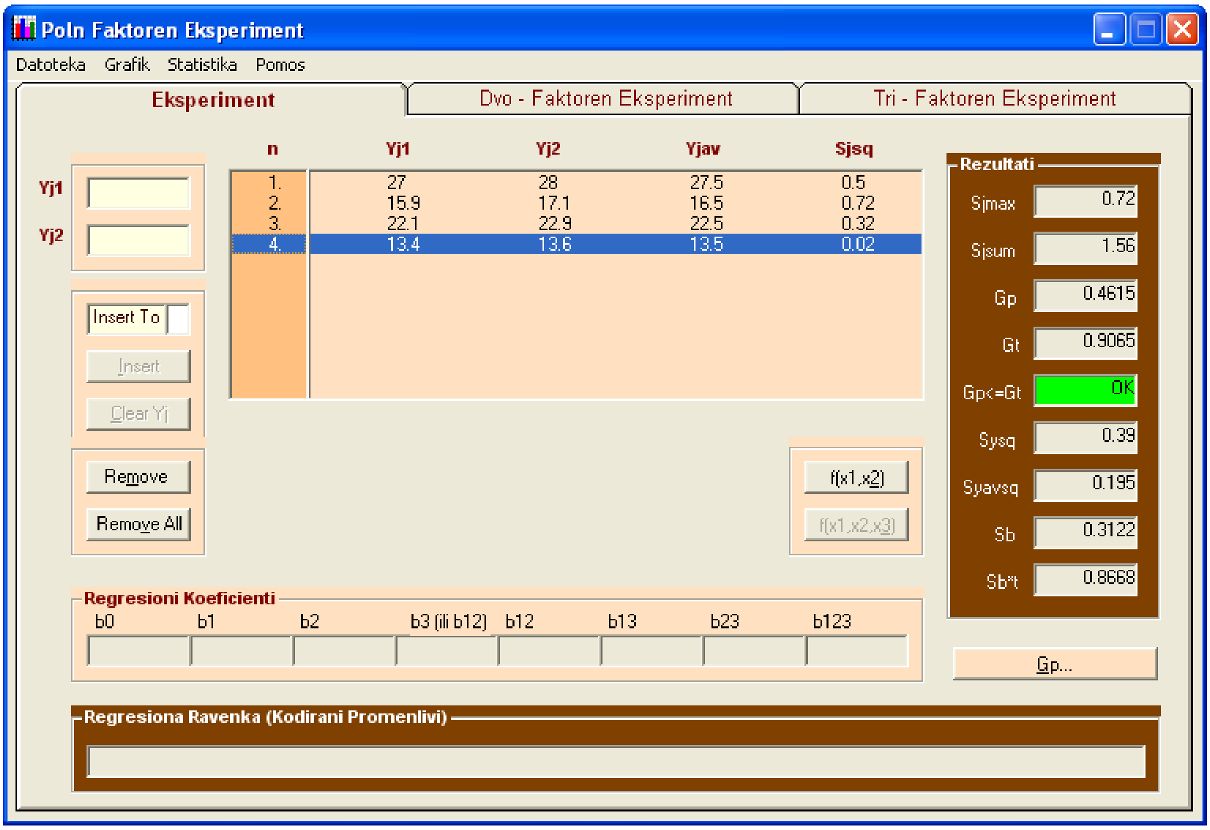
При внесувањето на податоците, цело време имаме увид во претходно внесените податоци кои стојат прикажани во листата (колона Y_{ji} и Y_{j2}). Покрај тоа, за секоја двојка на реплики програмата ни ја покажува нивната средна вредност (колона Y_{jav}) и варијансата (колона S_{jsq}) како и бројот на проби (колона n).

Програмата функционира само ако го имаме комплетниот број на опити, 4 за експеримент со два фактори и 8 за експеримент со три фактори.

Проверка на Кохреновиот критериум

Се додека не внесеме минимум четири податоци (за експеримент со два фактори, односно осум (за експеримент со три фактори) во листата, програмата не овозможува никаква обработка на податоците освем нивна корекција.

Откако ќе ја внесеме четвртата проба (n = 4) дури тогаш копчето “**G_p...**” станува активно. Со копчето “**G_p...**” проверуваме дали е задоволен Кохреновиот критериум. Ако е задоволен, полето G_p добива зелена позадина и во него се појавува знакот “ОК“, слика 5.5. Покрај тоа, сите полиња под и над ова поле (во рамката **Rezultati**) се пополнуваат со релевантните податоци кои програмата ги пресметала. Тоа се сите податоци кои мануелно ги пресметувавме во примерот 2 од претходната глава и чие значење може да се види од лабелите покрај соодветните полиња.



Слика 5.5. Проверка на Кохреновиот критериум

Програмата не дозволува менување на овие податоци. Тие се генерираат исклучиво од податоците во листата. Само ако таму извршиме измена можат овие податоци да се променат.

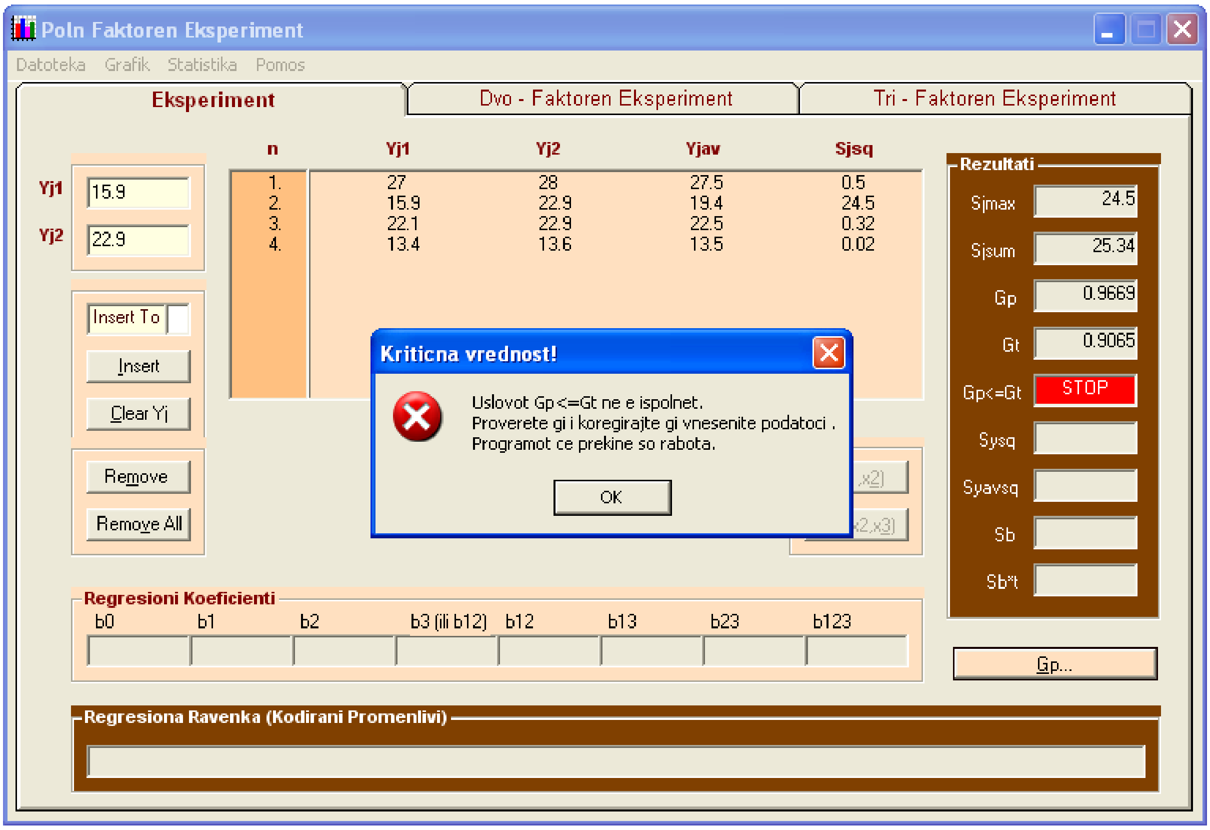
Копчето “f(x1,x2,x3)” е и понатаму неактивно, тоа се активира само ако има осум податоци во листата – за експеримент со три фактори.

Бидејќи извршивме проверка на Кохреновиот критериум и тој задоволува, дури сега копчето “f(x1,x2)” станува активно. во спротивно ова копче не е активно.

Меѓутоа, ако на пример податокот од вториот ред „17.1“ го замениме со произволен податок кој не е добиен од експериментот пример со „22.9“ и ако го провериме Кохреновиот критериум тој нема да биде задоволен. Во ваков случај програмата не опоменува

„*Условот $G_T \leq G_T$ не е исполнет. Проверете ги и коригирај те ги внесените податоци. Програмата ќе прекине со работа*“,

како што е прикажано на слика 5.6.



Слика 5.6. Кохреновиот критериум не е исполнет

Во ваков случај потребно е внимателно да се проверат податоците за да се открие кој е податокот што доведува до неисполнување на критериумот. Во конкретен пример, ако ја погледнеме вредноста на варијансата во петтата колона ќе забележиме значителна разлика на нејзината вредност во вториот ред во однос на другите вредности. Тоа ни е индикација дека сме внесле погрешна вредност за резултатот од експериментот или дека сме извршиле погрешно мерење на одзивот. Потребно е податоците во овој ред да ги коригираме и повторно да извршиме проверка на Кохреновиот критериум.

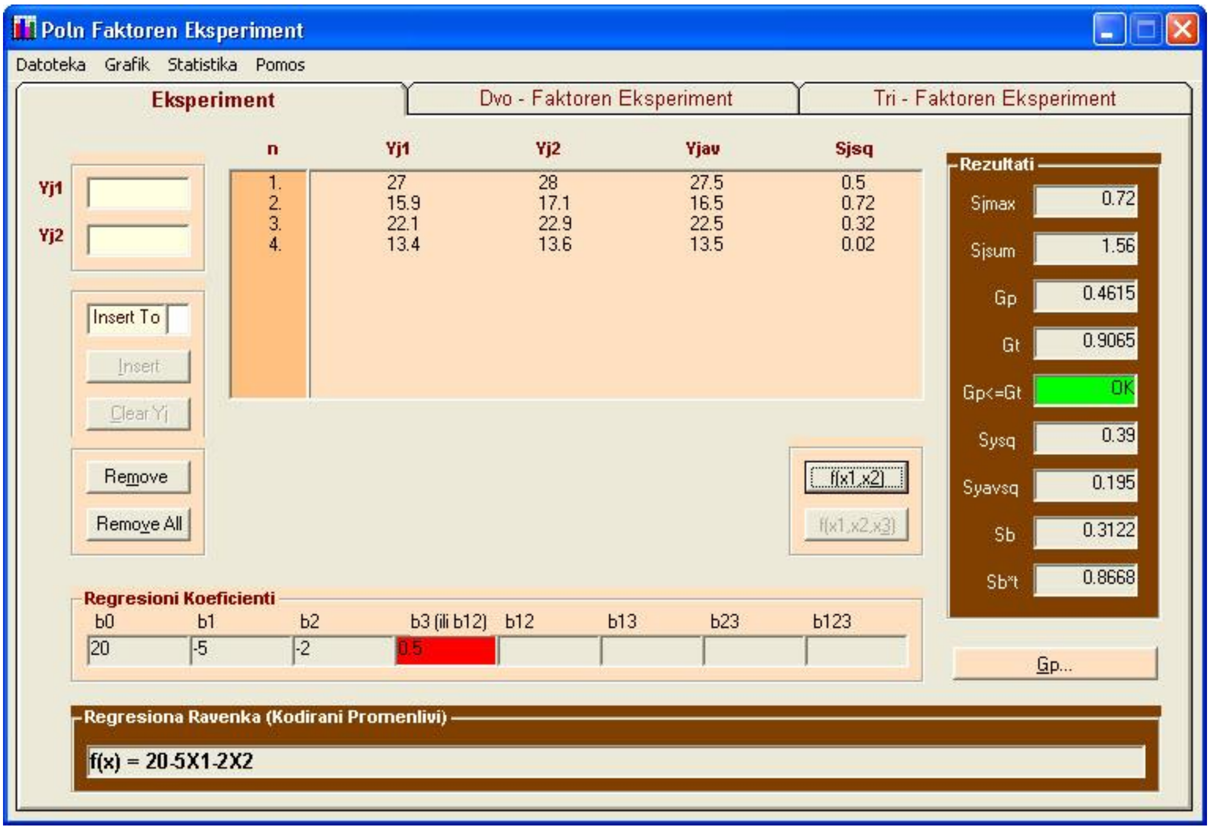
Пресметување на регресивната равенка со кодирани варијабли

Доколку е задоволен Кохреновиот критериум со активирање на копчето **f(x1,x2)** програмата ги пресметува регресионите коефициенти (означени со b_0, b_1, b_2,...) и ни ги презентира во рамката **Regresioni Koeficient** како и регресивната равенка која ни ја презентира во рамката **Regresiona Ravenka (Kodirani Promenlivi**

Доколку некој од коефициентите не го задоволува критериумот

$$|a| \geq s_a t \geq 0{,}8668$$

тој е означен со црвена позадина, и како незначаен не е вклучен во регресивната равенка, слика 5.7.



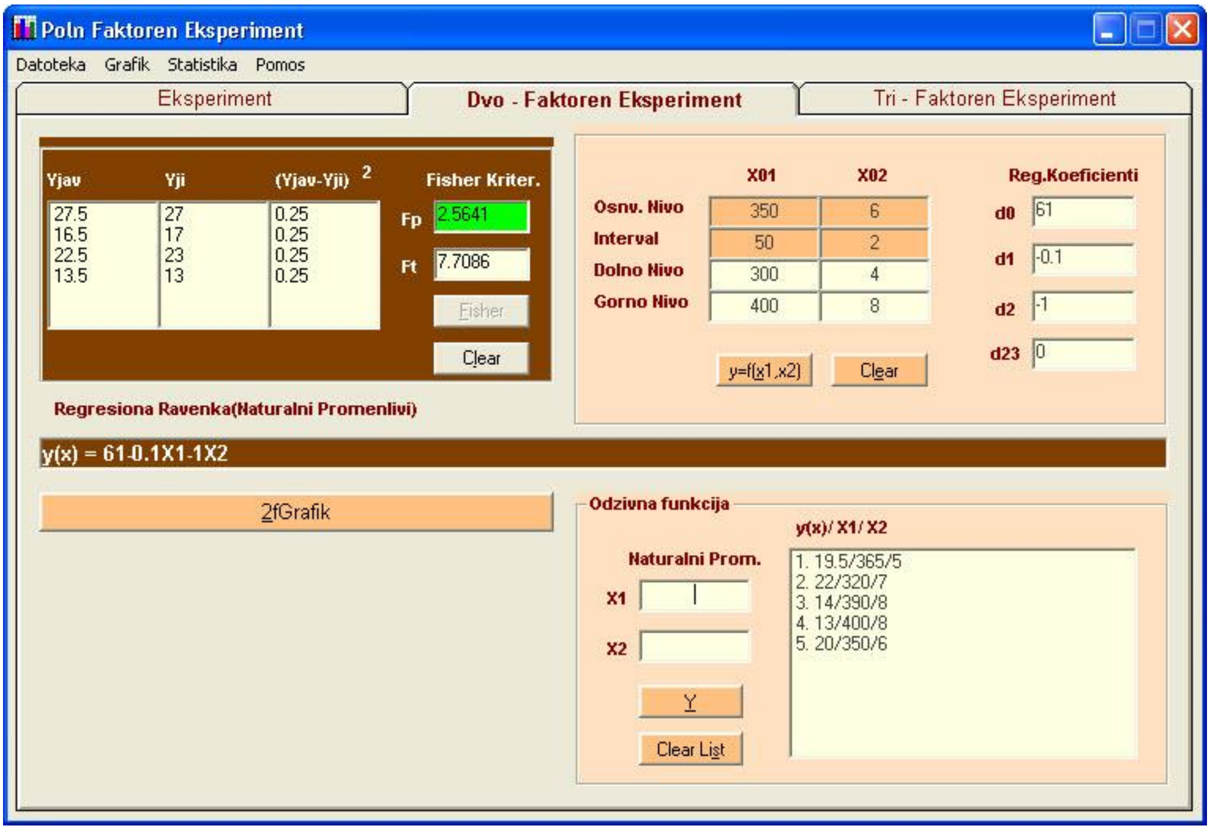
Слика 5.7. Пресметување на коефициентите и на регресивната равенка со кодирани варијабли

Пресметување на Фишеровиот критериум

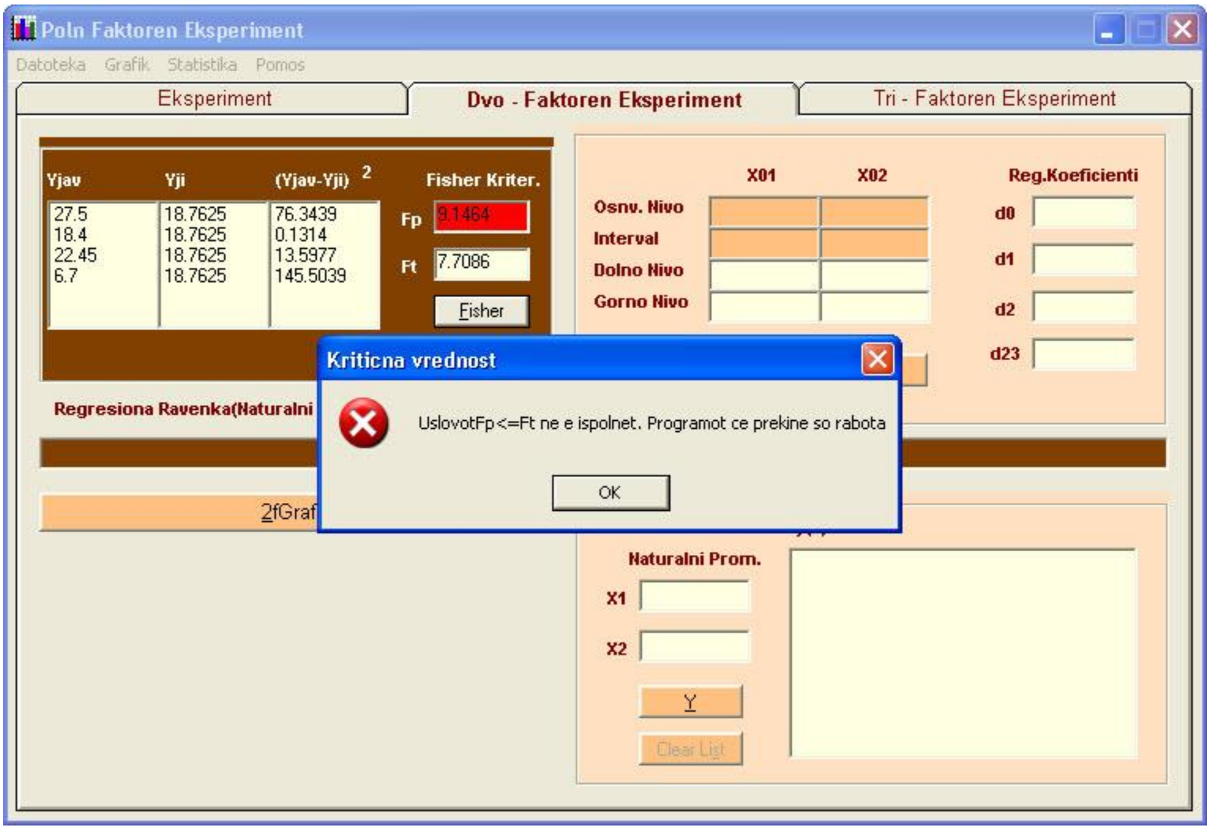
Откако ја определивме равенката со кодирани променливи понатамошниот тек на работа зависи од тоа дали се работи за експеримент со два или со три фактори. Со оглед дека работиме со експеримент со два фактори, во горниот дел од интерфејсот притискаме на табот **Dvo-Faktoren Eksperiment“** Овој таб ни презентира нов интерфејс , слика 5.8, кој се состои од неколку рамки во кои според функцијата се групирани команди за внесување на податоци или извршување одредени команди.

Во кафеавата рамка се презентираат статистичките податоци, според кои се одредува Фишеровиот критериум и копче за негова проверка. Во рамката во десната горна четвртина од интерфејсот се сместени командите за пресметување на регресионите коефициенти во инженериски единици. Во рамката **Odzivna funkcija** се сместени командите за одредување на вредноста на одзивот во инженериски единици за дадени вредности на факторите. Со копчињата **f(x1,x2)** и **2fGrafik** ја одредуваме функцијата на одзивот во инженериски единици и цртаме графикон, соодветно.

Кога ќе ни се презентира новиот интерфејс активно е само копчето **Fisher** со кое ја одредуваме адекватноста на моделот. По кликањето врз него, појавата на зелената позадина во полето **F_p** значи дека е задоволен Фишеровиот критериум. Покрај пресметаната вредност, тука ни е дадена и табличката вредност на Фишеровиот критериум во полето F_T.



Слика 5.8. Проверка на адекватноста на моделот и генерирање на равенка со натурални фактори



Слика 5.9. Неисполнување на Фишеровиот критериум

Доколку критериумот не е задоволен, полето F_T добива црвена позадина и програмата опоменува со пораката:

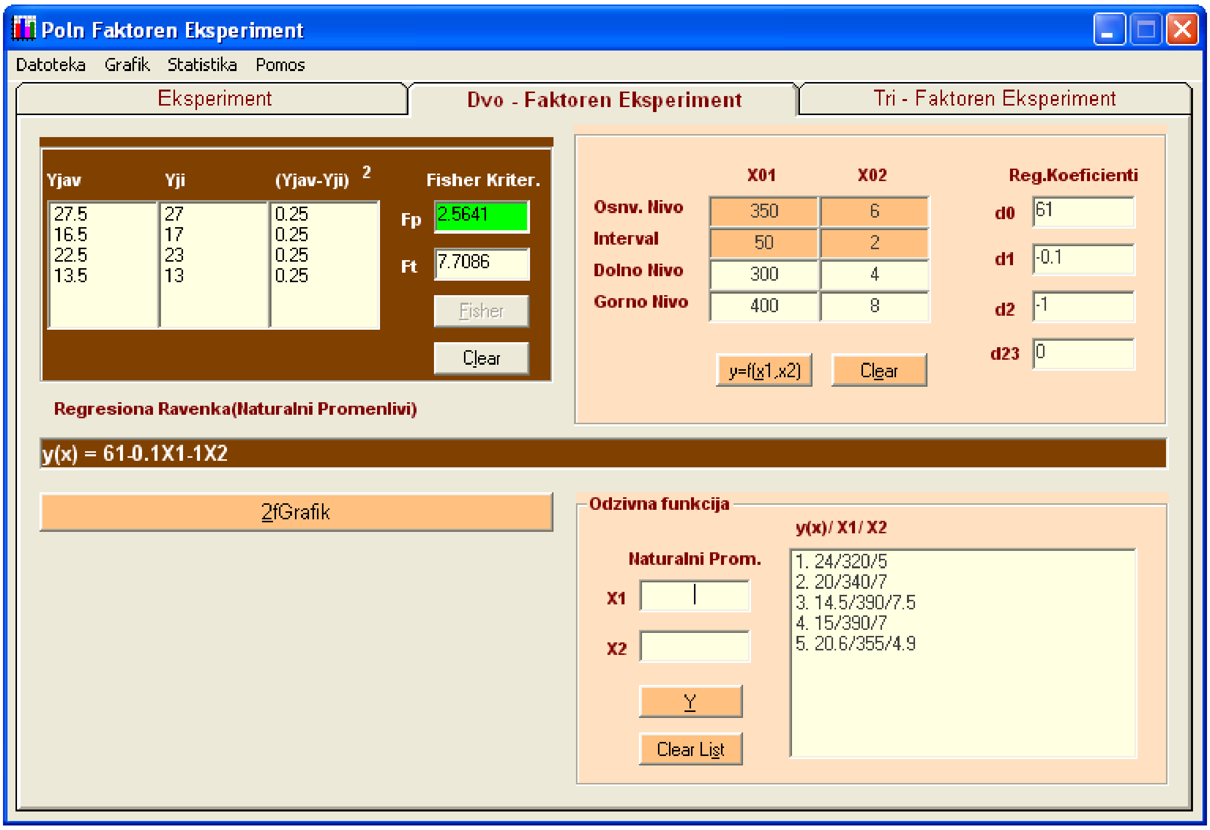
„*Условот $F_p \leq F_T$ не е исполнет. Програмата ќе прекине со работа*“

како што е прикажано на сликата 5.9.

Во тој случај треба да се проанализираат податоците кои се релевантни за Фишеровиот критериум и треба да се обиде причината да се идентификува и коригира или отстрани.

Одредување на регресивната равенка во инженериски варијабли

За да се одредат коефициентите и равенката во инженериски единици потребно е да се внесат податоци за основното ниво и интервалот на варирање за секоја варијабла. Полињата за внесување на овие податоци се активни само ако Фишеровиот критериум е задоволен. Ако ги внесеме податоците за основното (нулто) ниво (350 °C) и брзината (6 °C/min) програмата сама го пресметува долното и горното ниво на факторите во инженериски единици.



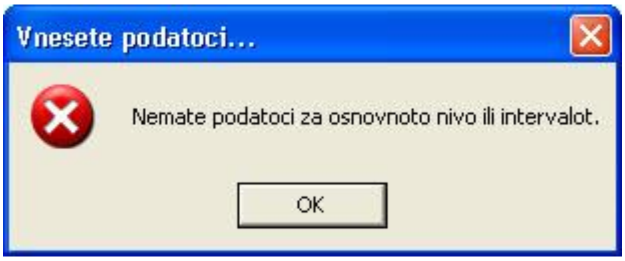
Слика 5.10. Одредување на моделот во инженерски единици

Ако сите полиња за основното ниво, интервалот, долно и горно ниво се исполнети со валидни податоци фокусот оди кон копчето **f(x1,x2)** со чие кликање ни се презентираат регресионите коефициентите и равенката во инженериски варијабли, слика 5.10.

Доколку во некое од полињата нема податок програмата не е опоменува со пораката:

Немате податоци за основното ниво или интервалот

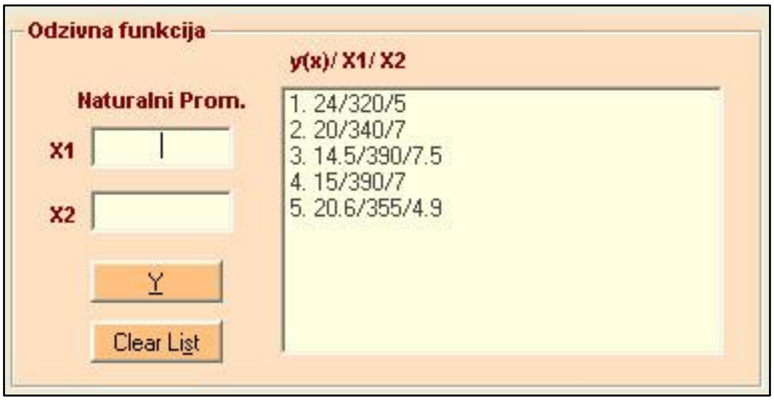
како што е прикажано на сликата 5.11.



Слика 5.11. Недоволно податоци за пресметка на коефициентите на регресија и регресивната равенка

Пресметување на одзивот

Сега, кога веќе ја имаме регресивната равенка во инженериски варијабли, лесно можеме да го пресметаме одзивот т.е. приносот при дестилацијата за различни комбинации на варијаблите во нивниот предвиден опсег. Тоа пресметување го правиме во рамката **Odzivna funkcija**, слика 5.12.



Слика 5.12. Пресметување на одзивот

Во полињата X1 и X2 ги внесуваме вредностите за факторите, температура и брзината на загревање соодветно и во рамката **y(x)/X1/X2** програмата ни го презентира резултатот во формат:

Вредност на одзивот/Вредност на температурата/Вредност на брзината

Со оваа опција на програмата можеме да го пресметаме одзивот за која било комбинација на варијаблите во веќе проучуваниот домен, без да заменуваме вредности во регресивната равенка и мануелно да го пресметуваме одзивот.

Од сликата 5.12 можеме да видиме дека одзивот е:

1. 24 %, ако е температурата на дестилација 320 ^oC а брзината на загревање 5 $^oC/min$
2. 20 %, ако е температурата на дестилација 340 ^oC а брзината на загревање 7 $^oC/min$
3. 14,5 %, ако е температурата на дестилација 390 ^oC а брзината на загревање 7,5 $^oC/min$, итн.

Како што веќе претходно е споменато, одзивот е валиден доколку факторите се во рамките на проучуваниот домен. Доколку факторите се вон доменот планираниот експеримент не може да ни даде веродостоен одзив. Доколку еден од факторите или двата се вон доменот, програмот не опоменува со пораката:

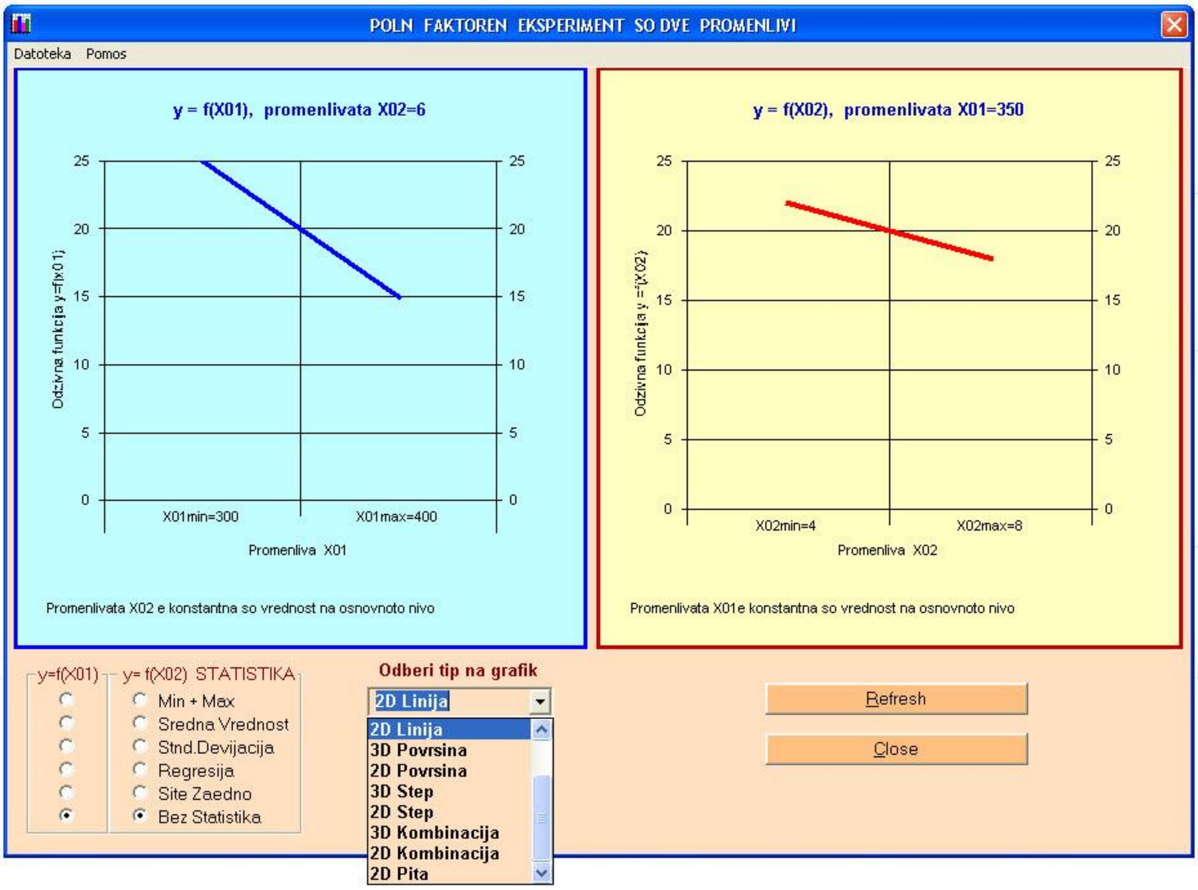
Променливата X1или X2 е вон опсегот!

како што е прикажано на сликата 5.13.

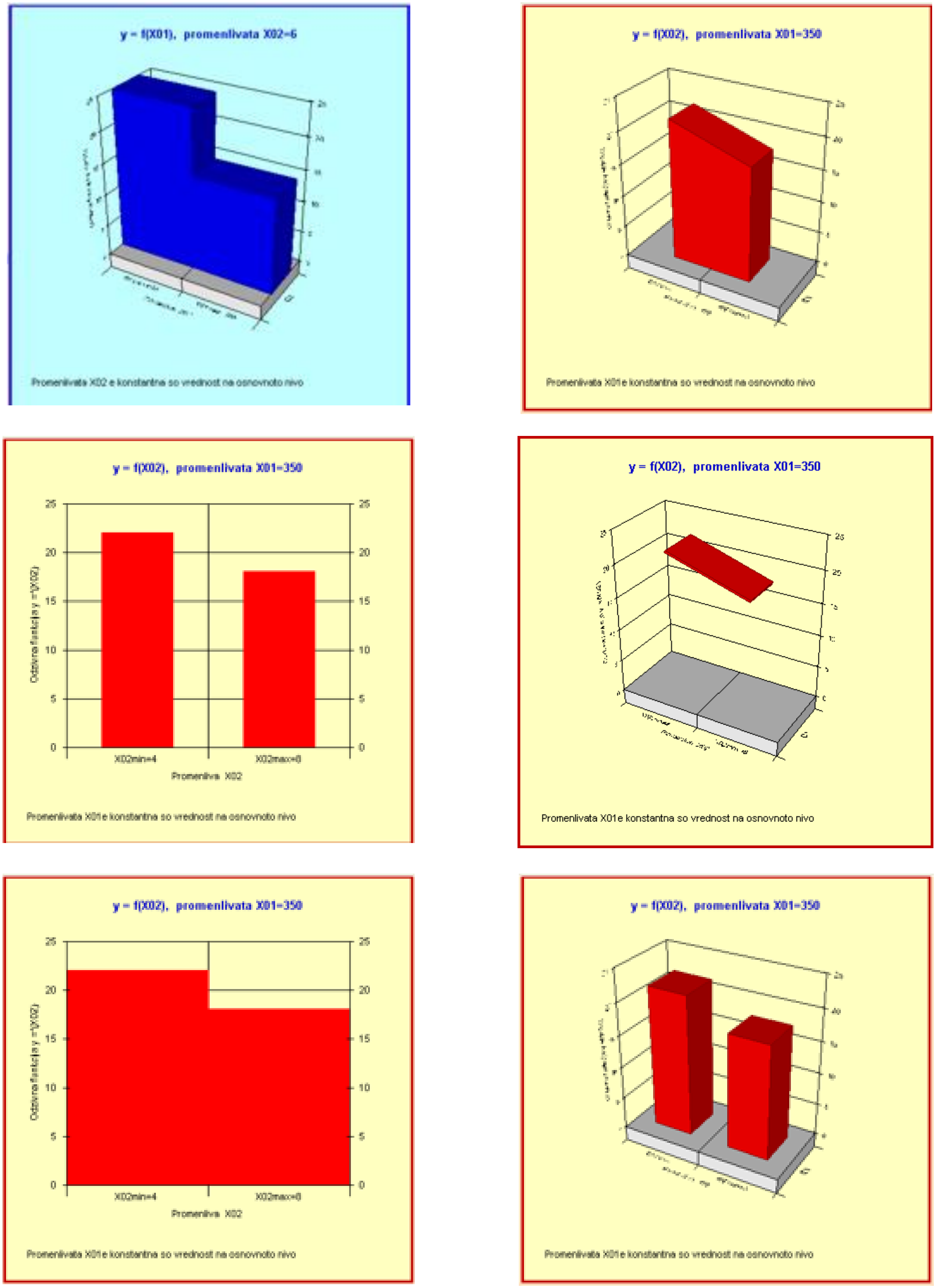
Слика 5.13. Променливи надвор од предвидениот домен програмата не прифаќа

Графичка презентација на резултатите

Најдобра претстава за влијанието на факторите, како што знаеме, е ако го претставиме нивниот ефект графички. Програмата дава и таква можност со избор на повеќе типови на 2Д и 3Д графици. Тоа е овозможено, по презентацијата на регресивната равенка во инженериски променливи, со кликање на копчето **2fGrafik**, слика 5.10. Со тоа ни се отвора нов интерфејс со дифолт 2Д линиски график, слика 5.14.



Слика 5.14. Графичка презентација на ефектот од факторите



Слика 5.15. Некои типови на графици на FFED 7.5

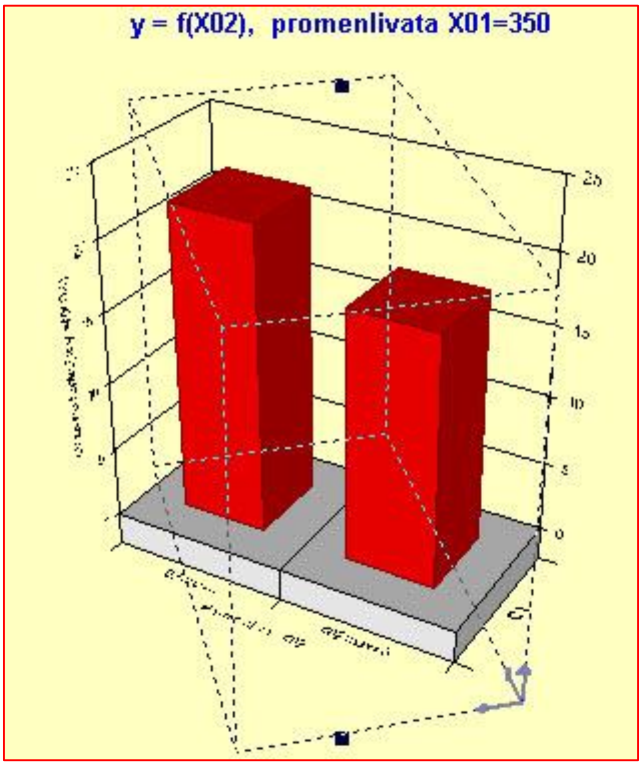
Покрај графички приказ во вид на линија, од комбо-боксот **Odberi tip na grafik** можеме да избереме уште многу други типови на графички приказ, како што е прикажано во табелата 5.1.

Табела 5.1. Типови на графици што можат да се креираат FFED 7.5

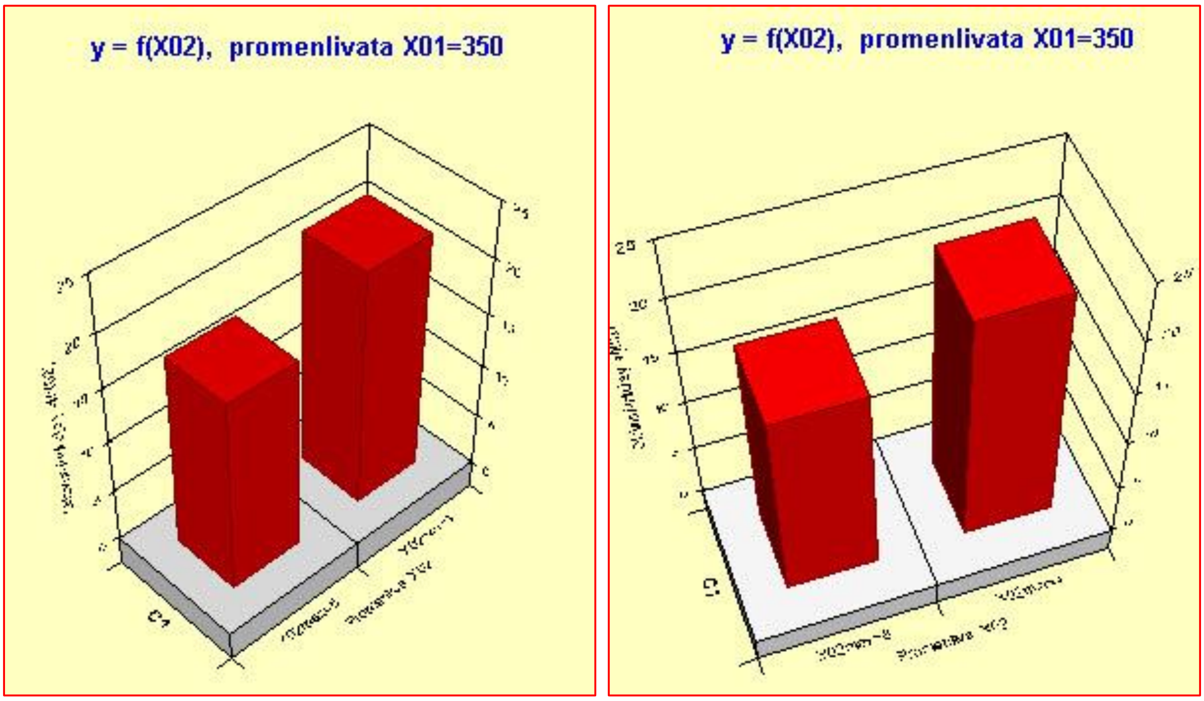
1	2Д линија	7	3Д линија
2	2Д столб	8	3Д столб
3	2Д Површина	9	3Д Површина
4	2Д Степ	10	3Д Степ
5	2Д Пита	11	3Д Комбинација
6	2Д Комбинација		

Некои типови од овие графици се прикажани на слика 5.15.

Доколку избереме 3Д графички приказ (на сите називот им почнува со 3D) тогаш него можеме да го ротираме, односно да му го промениме аголот на гледање. Тоа се прави на тој начин што се држи притиснат тастерот Ctrl од тастатурата и со левото копче од глушецот го менуваме аголот на гледање на графичкиот приказ во три оски. При тоа имаме визуелни водилки во вид на испрекинати линии, слика 5.16, кои ни ја покажуваат моменталната позиција. Откако ќе се одлучиме за саканата позиција го отпуштаме тастерот од маусот и графикот ја зазема саканата позиција.



Слика 5.16. Ротирање на графички приказ во простор со три димензии



Слика 5.17. Ротација на 3Д графички приказ

На сликата 5.17 прикажани се две различни позиции на графичкиот приказ од слика 5.16..

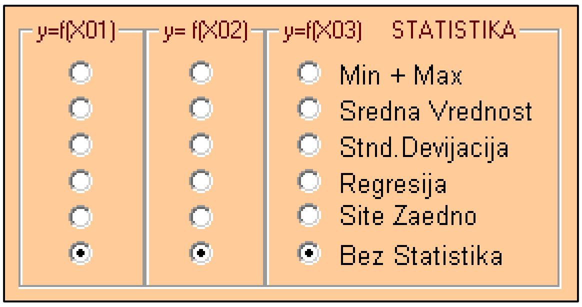
Со копчето **Refresh** е овозможено повторно вчитување на графичкиот приказ. Тоа е практично ако, откако еднаш сме го презентирале приказот, сакаме да направиме измени во податоците, тогаш, со притискање на ова копче го ажурираме графичкиот приказ со новите податоци.

Со копчето **Close** го затвораме графичкиот интерфејс и се враќаме на главната програма.

Од менито **Datoteka,** во графичкиот интерфејс, со избор на командата **Snimi**, можеме секој графички приказ поединечно да го снимиме како засебен графички фајл. Поддржани формати се: gif, jpg, bmp, wmf и ico. Исто така, од ова мени, со избор на командата **Pecati** можеме секој графички приказ поединечно да го отпечатиме.

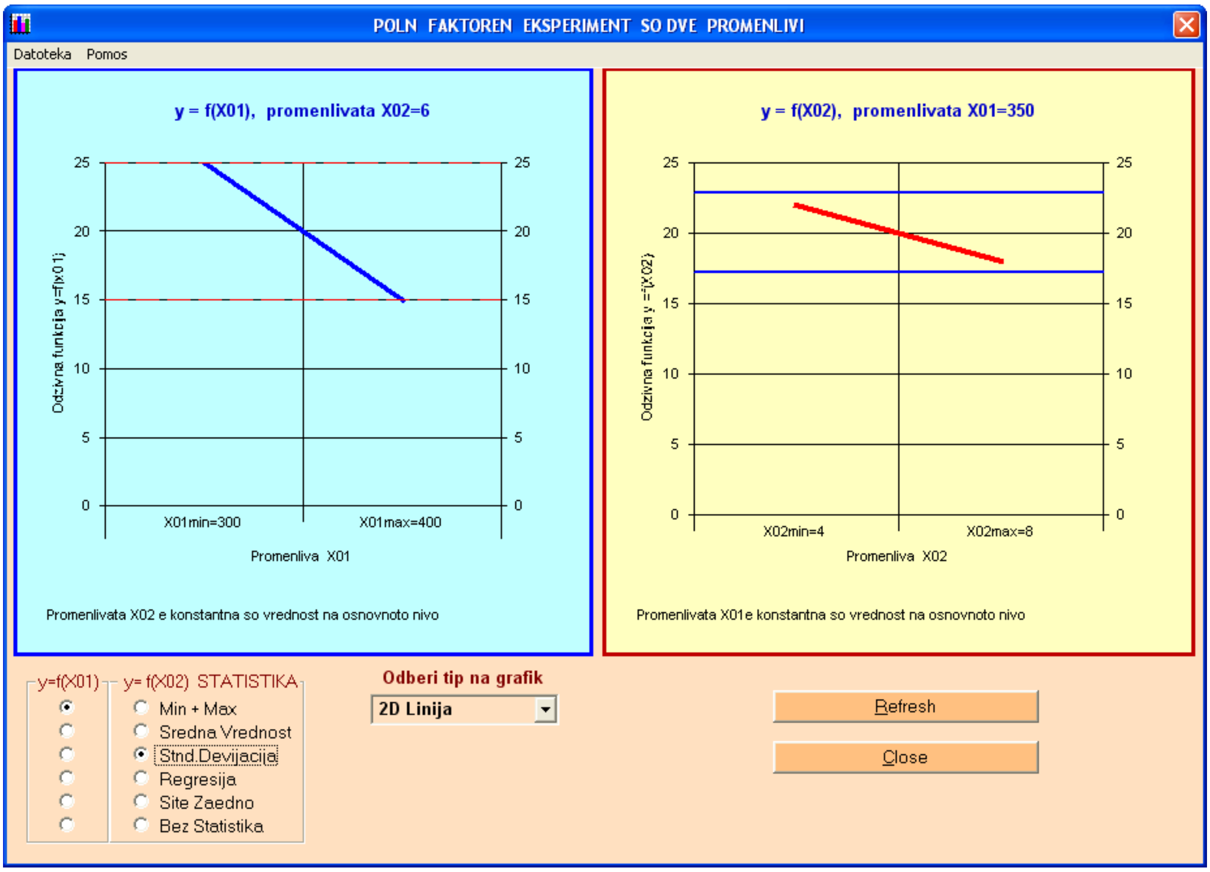
Визуелна статистика

Програмата поединечна овозможува 2Д линија график (кој е дифолт) графичка презентација на основните статистички податоци, како што се: минимална и максимална вредност, средна вредност, стандардна девијација, насока на регресија и презентација на сите заедно. Тоа е овозможено преку рамката **STATISTIKA** од интефејсот за график, слика 5.18.



Слика 5.18. Избор на статистичка карактеристика

Визуелната статистика е овозможена за секој фактор посебно и независно од другиот со селектирање на кругот, слика 5.18 пред соодветната статистичка карактеристика. На пример, на сликата 5.19 можеме паралелно да ги видиме минималната и максималната вредност на одзивот во зависност од ефектот на факторот x1 додека факторот x2 е на основно ниво (синиот графички приказ), како и стандардната девијација на одзивот во зависност од ефектот на факторот x2 кога факторот x1 е на нулта – основно ниво (црвен графички приказ). По желба, со селектирање на кругот **Bez statistika** визуелната статистика можеме да ја исклучиме од графичиот приказ.



Слика 5.19. Паралелно прикажување на различни статистички карактеристики на факторите

Креирање на извештај

За да се креира извештај експериментот треба да е напреднат барем до фаза да ја има одредено регресивната равенка со кодирани променливи. Доколку покушаме да направиме извештај без оваа равенка програмата опоменува тоа да се стори за да креира извештај. Можеме да креираме два типови на извештаи:

1. Извештај како Microsoft Word документ
2. Текстуален извештај

За да креираме извештај од менито **Datoteka** избираме **Kreiraj izvestaj kako…** и од подменито **Word Dokument** или **Tekst (txt) Datototeka** избираме меѓу едниот и другиот тип на извештаи. За да креираме Word документ извештај потребно е во компјутерот на кој се дизајнира експериментот да има инсталирано Microsoft Word. Извештајот се креира во посебен формулар, кој се инсталира со програмата и кој FFED 7.5 го пополнува со сите релевантни податоци од експериментот. Примерок од извештајот е приложен во прилог 5. Предноста на овој извештај е што можеме да ги користиме сите можности на Word за да го едитираме и зачуваме извештајот. Секој извештај може да се зачува како посебен фајл.

И текстуалниот извештај ги содржи сите релевантни податоци од експериментот. Предноста на текст извештајот е што можеме да креираме извештаи од повеќе експерименти во еден ист фајл. Недостаток е неможноста за современо форматирање на извештајот. Во прилог 6 даден е примерок на ваков извештај за актуелниов експеримент.

Преглед на менито

Некои од содржините на менито веќе спомнавме, да ги прегледаме оние кои не ги сретнавме во досегашново излагање, другите кои не се овде спомнати објаснети се во менито **Pomos**.

Paralelen Test – оваа команда овозможува отворање повеќе од една инстанца на програмот, всушност можат по желба да се отворат безброј инстанци. Тоа е пригодна опција што овозможува пр. на едната инстанца да го водиме експериментот со едни параметри а на другата со променети и да го споредуваме ефектот онлајн.

Otpecati Ekran – оваа команда овозможува печатење на интерфејсот со сите податоци кои се моментално прикажани

Otpecati Regresiona Ravenka – оваа команда овозможува печатење само на регрсионата равенка и/или нејзино снимање во pdf формат.

Statistika – ова е мал, независен помошен програм кој овозможува на група од податоци да се пресмета средната вредност и стандардната девијација

Pomos – ова е фајл за помош во кој детално, со илустрации (html формат) е објаснета работата со програмот

§6. ТРИФАКТОРЕН ПЛАН НА ЕКСПЕРИМЕНТ

Во §4 видовме колку многу информации можеме да собереме со едноставен дизајн на експериментот со два фактори (4 точки). Ако само со два фактори добивме толку многу информации, лесно е да се претпостави дека мулти-факторниот план ќе биде уште побогат со корисни информации. Во ова поглавје ќе се запознаеме со планирање на експерименти кај кои одзивот зависи од три фактори. Принципот на се факторниот план на експерименти е ист без разлика за колку фактори се работи.

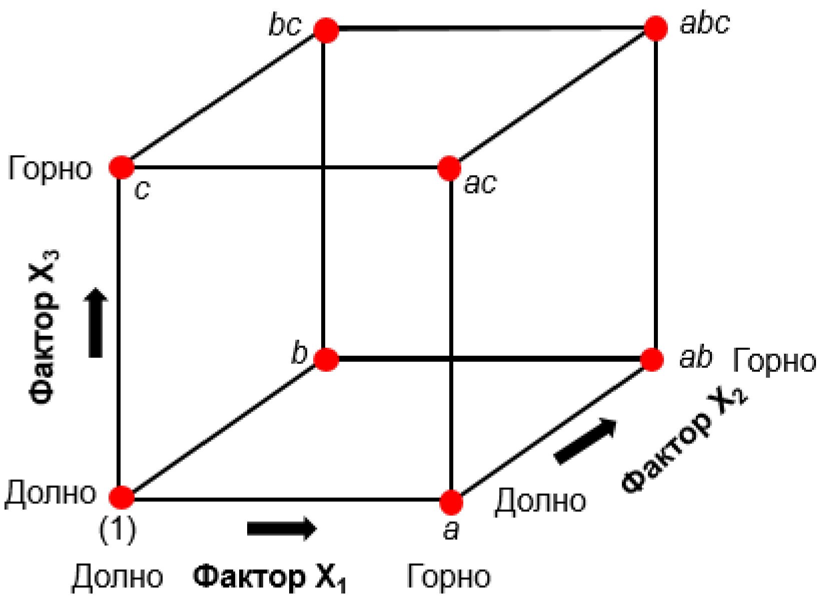
Да претпоставиме дека имаме процес кој зависи од три фактори, x_1, x_2, x_3 и дека секој фактор има две нивоа. Ваквиот дизајн е наречен **2^3 факторен дизајн.**

Минималниот број на експерименти кои треба да се направат е:

$$N = 2^k = 2^3 = 8$$

Што значи, за да го имплементираме овој дизајн треба да експериментираме со минимум 8 комбинации на фактори и нивоа. Ако секоја комбинација ја реплицираме два пати, заради статистичка обработка на податоците, тогаш минималниот број на тестови е 16.

Осумте комбинации на тестовите можеме геометриски да ги претставиме во вид на коцка, слика 6.1. Често пати, по конвенција, тестот кога сите фактори се на долното ниво го означуваме со „(1)“. Стрелките ја покажуваат насоката на пораст на факторите.



Слика 6.1. Геометриска претстава на дизајнот на експериментот

Ако ги кодираме долното и горното ниво осумте тестови можеме да ги претставиме со матрицата на табела 6.1.

Табела 6.1. Матрица на дизајнот на експерименти од типот 2^3

	Фактор		
Тест	x_1	x_2	x_3
1	-1	-1	-1
2	+1	-1	-1
3	-1	+1	-1
4	+1	+1	-1
5	-1	-1	+1
6	+1	-1	+1
7	-1	+1	+1
8	+1	+1	+1

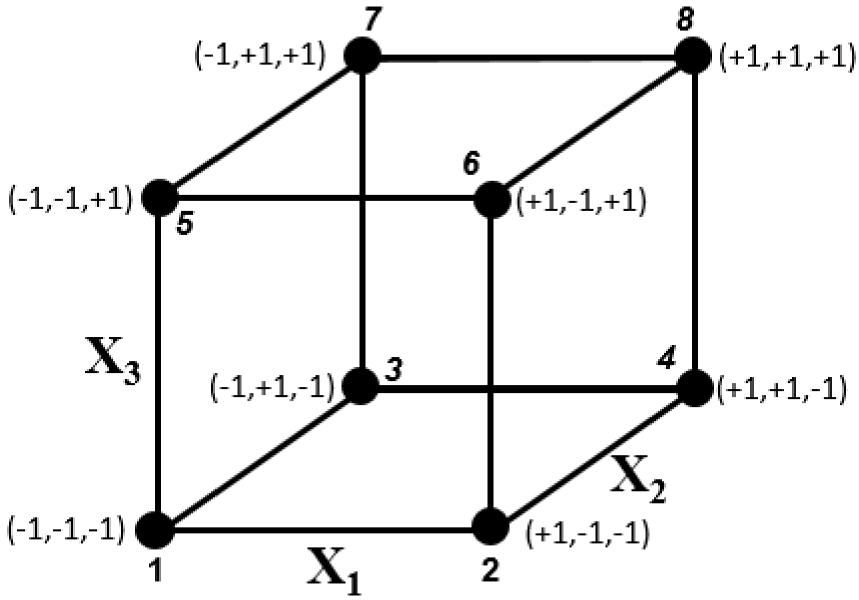
Ако во матрицата го внесеме редоследот на тестовите, ќе добиеме друга проширена матрица од која може да се види која комбинација од факторите одговара на соодветниот тест, табела 6.2.

Табела 6.2. Проширена матрица на дизајнот на експериментот

Опит	Редослед на опитите	Фактор		
		x_1	x_2	x_3
1	(1)	-1	-1	-1
2	***a***	+1	-1	-1
3	***b***	-1	+1	-1
4	***ab***	+1	+1	-1
5	***c***	-1	-1	+1
6	***ac***	+1	-1	+1
7	***bc***	-1	+1	+1
8	***abc***	+1	+1	+1

Постојат седум степени на слобода меѓу осумте комбинации кај 2^3 дизајнот. Три степени на слобода се поврзани со главните ефекти на x_1, x_2 и x_3. Четири степени на слобода се поврзани со интеракциите: по еден за x_1x_2, x_1x_3 и x_2x_3 и еден за $x_1x_2x_3$.

Сега можеме и геометриски да го претставиме редоследот (не рандомизиран) на тестовите според дизајнот на експериментот, слика 6.2.



Слика 6.2. Редослед на тестовите според дизајнот на експериментот

Ако ја внесеме и интеракцијата во матрицата ќе добиеме матрица за дизајнот, со внесени интеракции меѓу факторите, табела 6.3.

Табела 6.3.Матрица на главните ефекти и нивните интеракции

Опит	x_1	x_2	x_3	x_1 x_2	x_1 x_3	x_2 x_3	x_1 x_2 x_3
1	-1	-1	-1	+1	+1	+1	-1
2	+1	-1	-1	-1	-1	+1	+1
3	-1	+1	-1	-1	+1	-1	+1
4	+1	+1	-1	+1	-1	-1	-1
5	-1	-1	+1	+1	-1	-1	+1
6	+1	-1	+1	-1	+1	-1	-1
7	-1	+1	+1	-1	-1	+1	-1
8	+1	+1	+1	+1	+1	+1	+1

Да ги одредиме главните ефекти.

Главен ефект на x_1

Ефектот на x_1 кога x_2 и x_3 се на долното ниво е :

$$\frac{a-(1)}{n}$$

n е бројот на реплики (најчесто n = 2)

Ефектот на x_1 кога x_2 е на горното ниво а x_3 е на долното ниво е:

$$\frac{ab-b}{n}$$

Ефектот на x_1 кога x_2 е на долното ниво а x_3 е на горното ниво, е:

$$\frac{ac-c}{n}$$

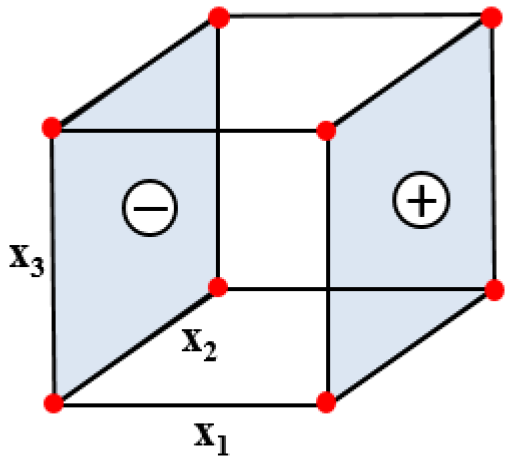
И конечно, ефектот на x_1 кога x_2 е на горното ниво и x_3 е на горното ниво, е:

$$\frac{abc-bc}{n}$$

Така, средниот (просечниот) ефект на x_1 е просечната вредност од овие четири ефекти:

$$x_1 = \frac{1}{4n}[a-(1)+ab-b+ac-c+abc-bc] \tag{6.1}$$

Истата оваа равенка ќе ја добиеме како разлика меѓу четирите тестови на десната страна од коцката, слика 6.3, кога x_1 е на горното ниво, и четирите тестови на левата страна од коцката, кога x_1 е на долното ниво. Т.е. ефектот



Слика 6.3. Геометриска презентација на разликата во ефектот на факторот x_1

на x_1 е средната вредност од четирите тестови кога x_1 е на горното ниво ($\bar{y}_{+x1}$) минус средната вредност од четирите тестови кога x_1 е на долното ниво ($\bar{y}_{-x1}$), или:

$$x_1 = \bar{y}_{+X_1} - \bar{y}_{-X_1} = \frac{a + ab + ac + abc}{4n} - \frac{(1) + b + c + bc}{4n}$$

Ако се прераспредели оваа равенка, ќе се добие:

$$x_1 = \frac{1}{4n}[a - (1) + ab - b + ac - c + abc - bc]$$

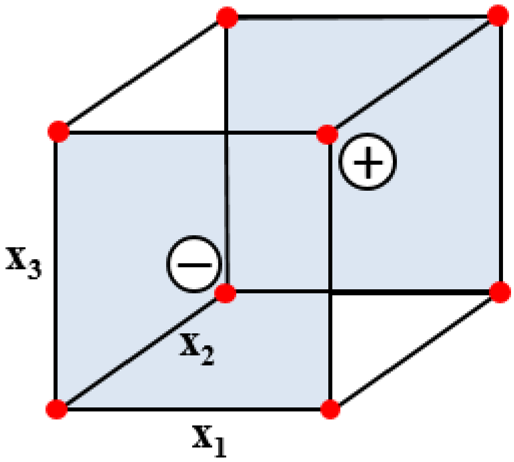
што е идентична равенка на 6.1.

Главен ефект на x_2

На сличен начин како за x_1 така и за x_2 главниот ефект е:

$$x_2 = \bar{y}_{+X_2} - \bar{y}_{-X_2} = \frac{1}{4n}[b + ab + bc + abc - (1) - a - c - ac] \qquad (6.2)$$

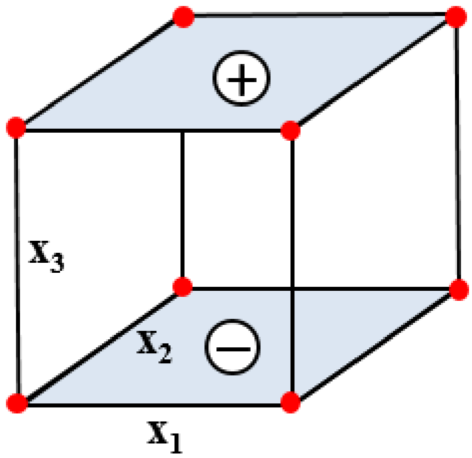
Геометриската разлика на ефектот на факторот е претставена на сликата 6.4.



Слика 6.4. Геометриска презентација на разликата во ефектот на факторот x_2

Главен ефект на x_3

Ефектот на x_3 е разликата во средните вредности меѓу четирите тестови на горната површина и четирите тестови на долната, слика 6.5.



Слика 6.5. Геометриска презентација на разликата во ефектот на факторот x_3

$$x_3 = \bar{y}_{+X_3} - \bar{y}_{-X_3} = \frac{1}{4n}[c + ac + bc + abc - (1) - a - b - ab] \tag{6.3}$$

Ефект на интеракцијата x_1x_2

Интеракцијата меѓу два фактори може лесно да се пресмета. Мерка за интеракцијата x_1x_2 е разликата меѓу просечниот ефект на x_1 на двете нивоа од x_2. По конвенција, половината од оваа разлика се нарекува x_1x_2 интеракција. Симболично тоа го претставуваме вака:

x_2	Просечен x_1 ефект
Горно (+)	$\frac{[(abc-bc)+(ab-b)]}{2n}$
Долно (-)	$\frac{\{[(ac-c)+[a-(1)]\}}{2n}$
Разлика	$\frac{[abc-bc+ab-b-ac+c-a+(1)]}{2n}$

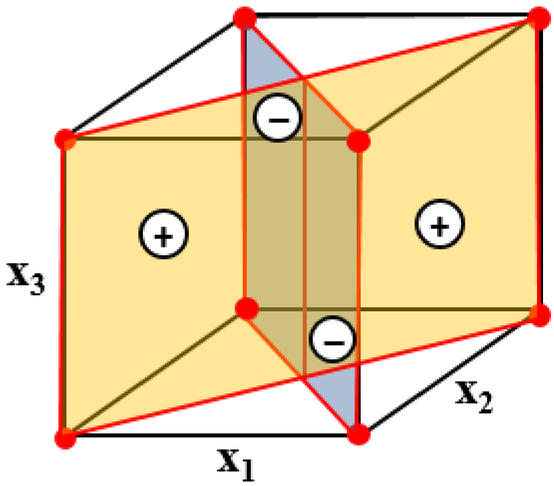
Бидејќи интеракцијата е една половина од оваа разлика пишуваме:

$$x_1x_2 = \frac{[abc-bc+ab-b-ac+c-a+(1)]}{4n} \tag{6.4}$$

Равенката 3.4 можеме да ја запишеме и на следниов начин:

$$x_1x_2 = \frac{abc+ab+c+(1)}{4n} - \frac{bc+b+ac+a}{4n}$$

Во оваа форма лесно може да се види дека интеракцијата x_1x_2 претставува разлика во средните вредности на тестовите меѓу двете дијагонални рамнини на коцката, слика 6.6.

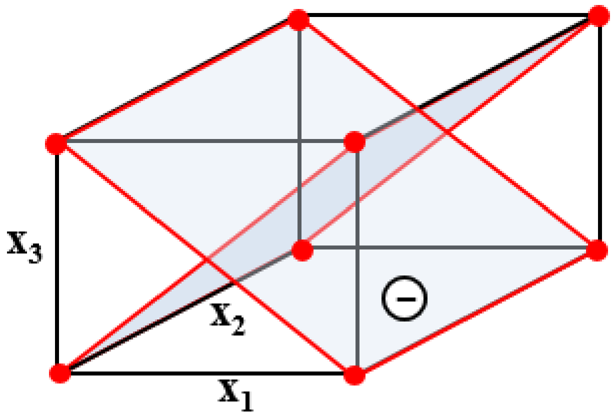


Слика 6.6. Ефект на интеракцијата x_1x_2

На сличен начин ги пресметуваме и останатите интеракции:

Ефект на интеракцијата x_1x_3

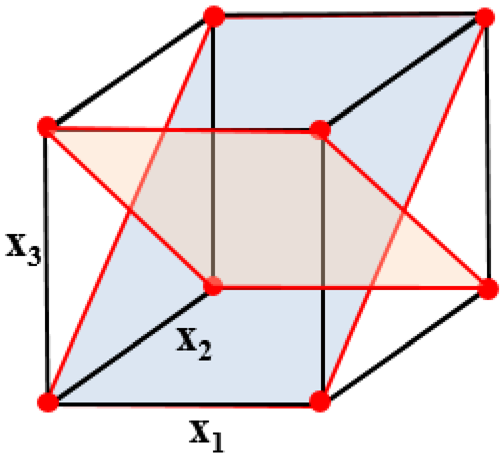
$$x_1x_3 = \frac{1}{4n}[(1)-a+b-ab-c+ac-bc+abc] \tag{6.5}$$



Слика 6.7. Ефект на интеракцијата x_1x_3

Ефект на интеракцијата x_2x_3

$$x_2x_3 = \frac{1}{4n}[(1) + a - b - ab - c - ac + bc + abc] \qquad (6.6)$$



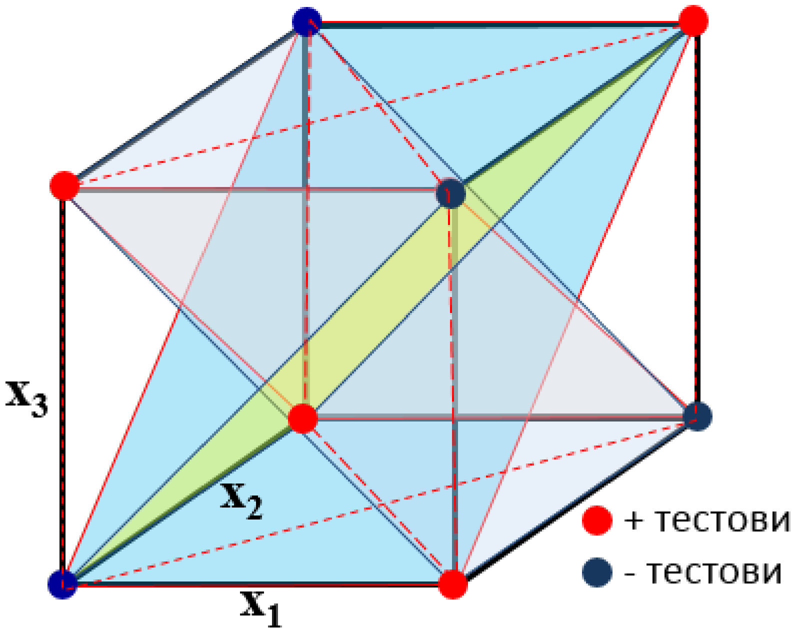
Слика 6.8. Ефект на интеракцијата x_2x_3

Ефект на интеракцијата $x_1x_2x_3$

Интеракцијата $x_1x_2x_3$ е дефинирана како средна разлика меѓу x_1x_2 интеракцијата при две различни нивоа на x_3 . Па според тоа:

$$x_1x_2x_3 = \frac{1}{4n}\{[abc - bc] - [ac - c] - [ab - b] + [a - (1)]\}$$

$$x_1x_2x_3 = \frac{1}{4n}[abc - bc - ac + c - ab + b + a - (1)] \qquad (6.7)$$



Слика 6.9. Интеракција меѓу три фактори

На следниот пример ќе ја покажеме имплементацијата на се факторниот дизајн на експерименти со три променливи.

Пример 6.1.

Во хемиска пилот постројка се одвива процес чии принос зависи од три фактори: температурата, концентрацијата и типот на катализаторот. Матрицата на експериментот е дадена во табелата 6.4.

Табела 6.4. Матрица на експериментот

Експе-римент #	Темпе-ратура x_1	Концен трација x_2	Катали затор x_3	Принос		$\bar{y}$
				y_{i1}	y_{i2}	
1	-1	-1	-1	59	61	60
2	+1	-1	-1	74	70	72
3	-1	+1	-1	50	58	54
4	+1	+1	-1	69	67	68
5	-1	-1	+1	50	54	52
6	+1	-1	+1	81	85	83
7	-1	+1	+1	46	44	45
8	+1	+1	+1	79	81	80
- 1 Ниво	160°C	10 %	A			
+ 1 Ниво	180°C	40 %	B			

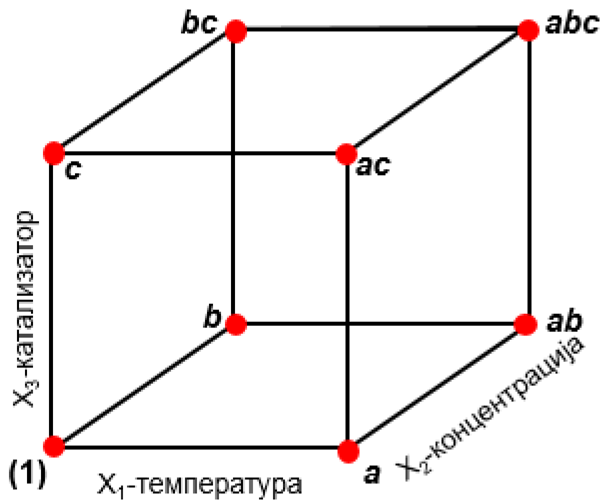
При стандардниот редослед

- Во првата колона варијаблата се менува (-1, +1, -1....) на секоја 2^0 вредност.
- Во втората колона варијаблата се менува (-1, -1, +1, +1...) на секоја 2^1 вредност.
- Во третата колона варијаблата се менува (-1,-1,-1,-1,+1,+1,...) на секоја 2^2 вредност.
- n-тата варијабла се менува на секоја 2^{n-1} вредност.

Геометриски овој експеримент ќе го претставиме во вид на коцка, слика 6.10, означувајќи ги променливите во 3Д просторот. Според дефиницијата за сефакторниот експеримент ако има ***a*** нивоа на факторот x_1, ***b*** нивоа на факторот x_2 и ***c*** нивоа на факторот x_3 сите комбинации ***abc*** се тестираат. Кога факторите ќе се подредат во експерименталната матрица често се нарекуваат вкрстени (*crossed*) фактори.

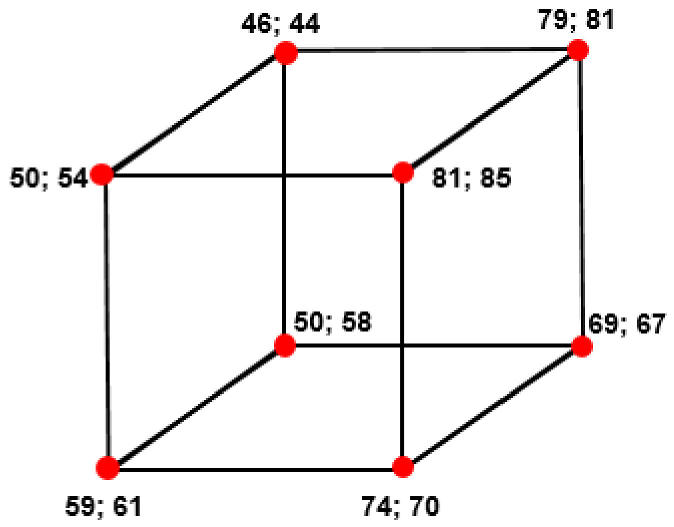
Претпоставениот модел од прв ред за процес со три фактори е:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 + \beta_{123} x_1 x_2 x_3 + \varepsilon$$



Слика 6.10. Ортогоналност на променливите

На сликата 6.10 претставена е ортогоналноста на трите фактори на процесот со назнака за нивната позиција во геометриската презентација. За 2^3 сефакторен експеримент потребни се 8 тестови (2^3 = 8, или abc = 2 x 2 x 2), секој со различна комбинација на факторите и нивоата. Со оглед дека за секоја комбинација се направени 2 теста (2 реплики, y_{i1} и y_{i2}), вкупно се извршени 16 тестови. Резултатите од секој тест (со соодветното ниво на факторите) се претставени на сликата 6.11.



Слика 6.11. Геометриска презентација на одзивот за секој од 16-те тестови

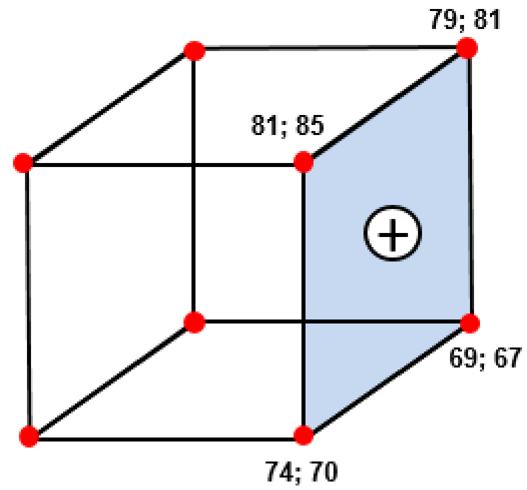
Според дефиницијата, ефектот на некој фактор го дефинираме како промена на одзивот (y) во зависност од промената на нивото на факторот. Тоа е **главен ефект** бидејќи се однесува на **основните фактори** кои се од интерес за експериментот.

Пресметување и интерпретација на главниот ефект X_1

Пресметување на главниот ефект x_1 се врши на тој начин што се пресметува просечната вредност на одзивот од сите тестови кога е x_1 на горното ниво и од неа се одзема просечната вредност на одзивот од сите тестови кога е x_1 на долното ниво.

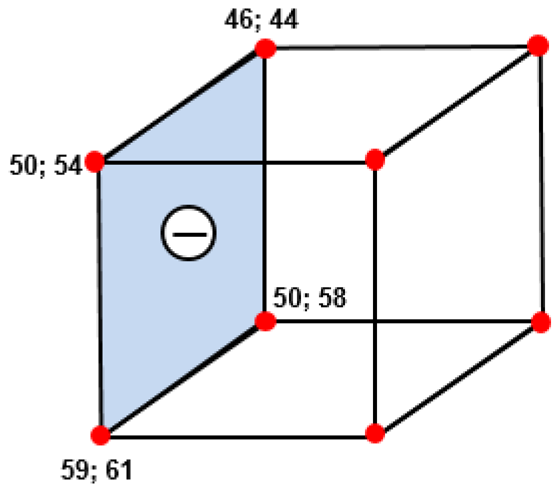
Просечниот одзив кога е x_1 на горно ниво е:

$$\bar{y}_{+X_1} = \frac{74+70+69+67+81+85+79+81}{8} = 75{,}75$$



Просечниот одзив на x_1 кога е на долното ниво е:

$$\bar{y}_{-X_1} = \frac{59+61+50+58+50+54+46+44}{8} = 52{,}75$$



$$x_1 = \bar{y}_{+X_1} - \bar{y}_{-X_1} = 75{,}75 - 52{,}75 = 23{,}0$$

Што значи дека ефектот на x_1 е 23.

Истиот резултат ќе го добиеме доколку директно работиме со средните вредности на одзивот. Ако во равенката 3.1 ги замениме овие вредности (последната колона од табела 3.1.1) ќе добиеме

$$x_1 = \frac{1}{4n}[a-(1)+ab-b+ac-c+abc-bc]$$

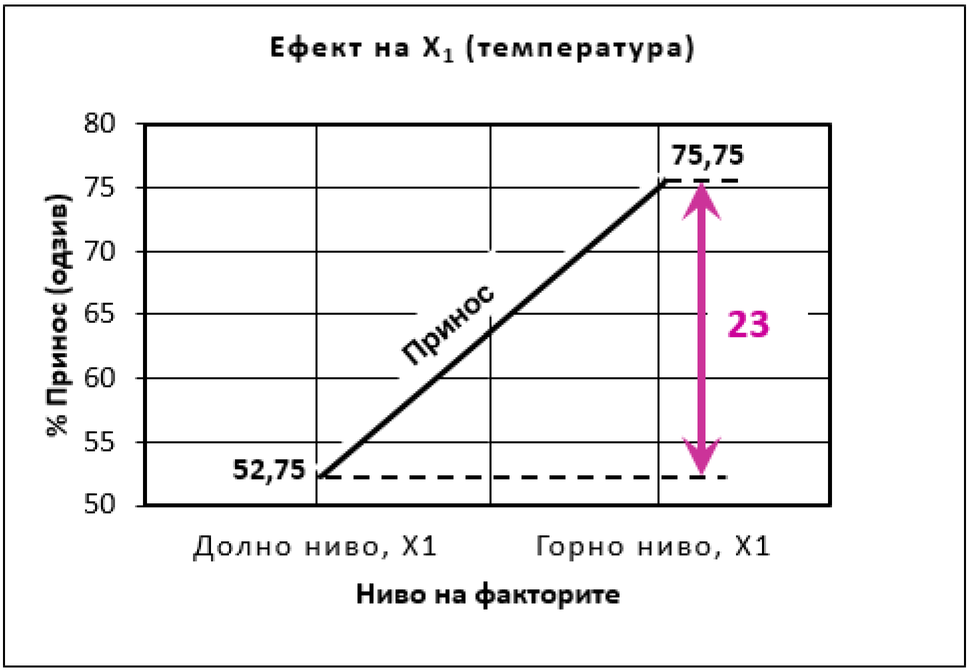
$$x_1 = \frac{1}{4 \cdot 1}(72-60+68-54+83-52+80-45) = 23{,}0$$

во овој случај, бидејќи работиме само со една вредност на одзивот, n=1.
Бидејќи е поедноставна постапката ако работиме со средна вредност, во продолжение, за пресметување на другите главни ефекти ќе работиме со средните вредности.

Сумарно ефектот на x_1, користејќи ќе ги средните вредности, ќе го претставиме во табела:

	Горно ниво, X_1	**Долно ниво, X_1**	**Ефект на X_1**
	72	54	
	68	60	
	83	52	
	80	45	
Средна вредност	**75,75**	**52,75**	**23**

Интерпретација на главниот ефект X_1



Пресметување и интерпретација на главниот ефект X2

Пресметување на главниот ефект x_2 се врши на тој начин што се пресметува просечната вредност на одзивот од сите тестови кога е x_2 на горното ниво и од неа се одзема просечната вредност на одзивот од сите тестови кога е x_2 на долното ниво. Таа математичка операција скратено е претставена со равенката:

$$x_2 = \bar{y}_{+X_2} - \bar{y}_{-X_2} = \frac{1}{4n}[b + ab + bc + abc - (1) - a - c - ac]$$

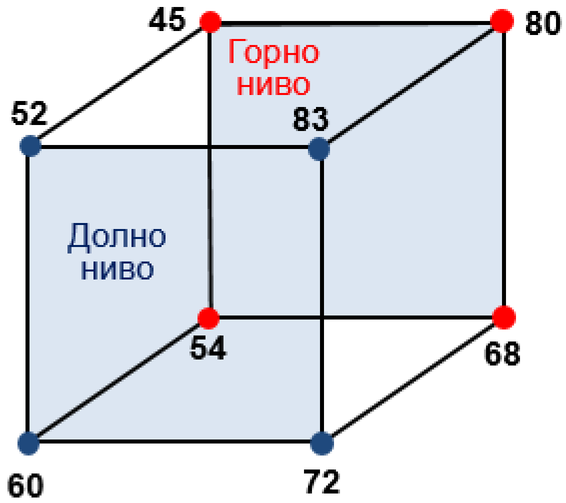
Во оваа равенка наместо со секоја реплика поединечно, заради поедноставување и скратување на постапката, ќе работиме со нивните средни вредности.

Ефектот на факторот x_2 е:

$$x_2 = \frac{1}{4 \cdot 1}(54 + 68 + 45 + 80 - 60 - 72 - 52 - 83) = -5$$

Бидејќи работиме само со една вредност на одзивот n = 1.

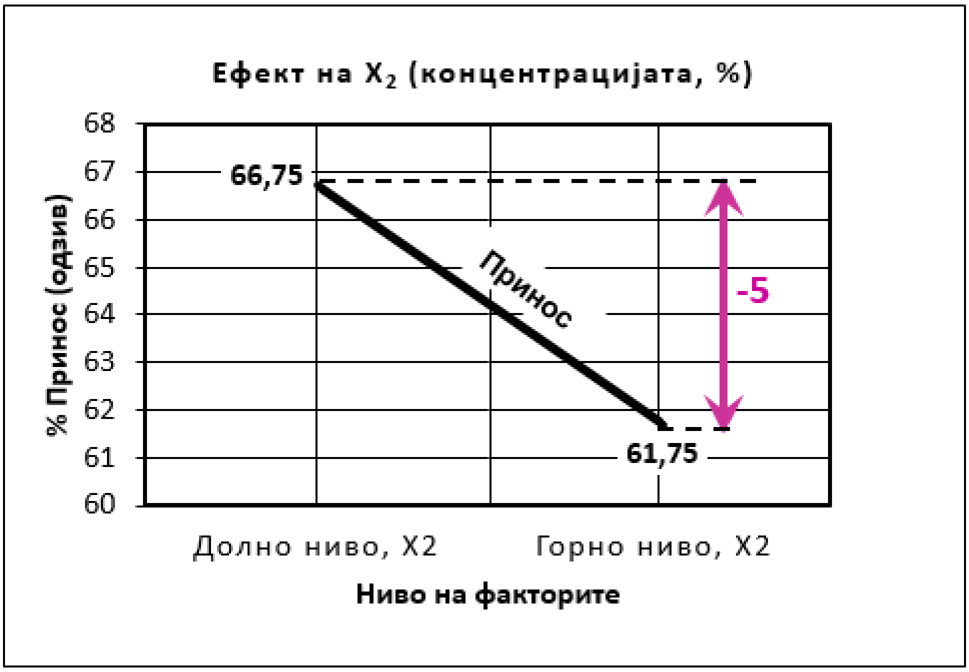
Графички овој ефект ќе го претставиме на следниов начин:



Сумарно ефектот на x_2, користејќи ги средните вредности, ќе го претставиме во табела:

	Горно ниво, X_2	Долно ниво, X_2	Ефект на X_2
	54	60	
	68	72	
	45	52	
	80	83	
Средна вредност	**61,75**	**66,75**	**- 5**

Интерпретација на главниот ефект X_2



Пресметување и интерпретација на главниот ефект X_3

Пресметување на главниот ефект x_2 се врши на тој начин што се пресметува просечната вредност на одзивот од сите тестови кога е x_2 на горното ниво и од неа се одзема просечната вредност на одзивот од сите тестови кога е x_2 на долното ниво. Таа математичка операција скратено е претставена со равенката:

$$x_3 = \bar{y}_{+X_3} - \bar{y}_{-X_3} = \frac{1}{4n}[c + ac + bc + abc - (1) - a - b - ab]$$

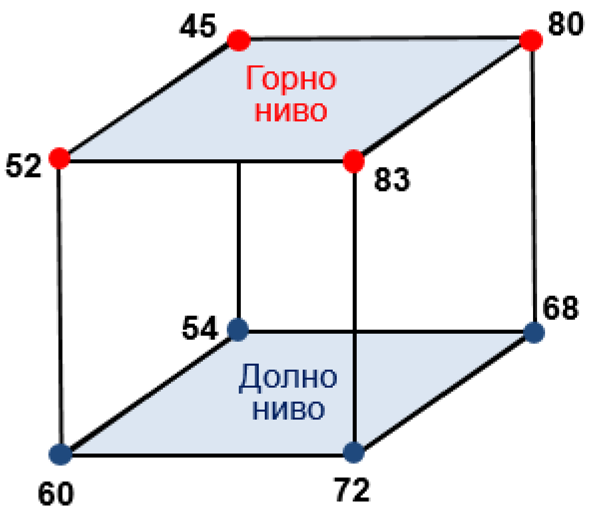
И овој ефект ќе го пресметаме од средните одзиви на репликите (одзивот).

Ефектот на факторот x_3 е:

$$x_3 = \frac{1}{4 \cdot 1}(52 + 83 + 45 + 80 - 60 - 72 - 54 - 68) = 1{,}5$$

Бидејќи работиме само со една вредност на одзивот n = 1.

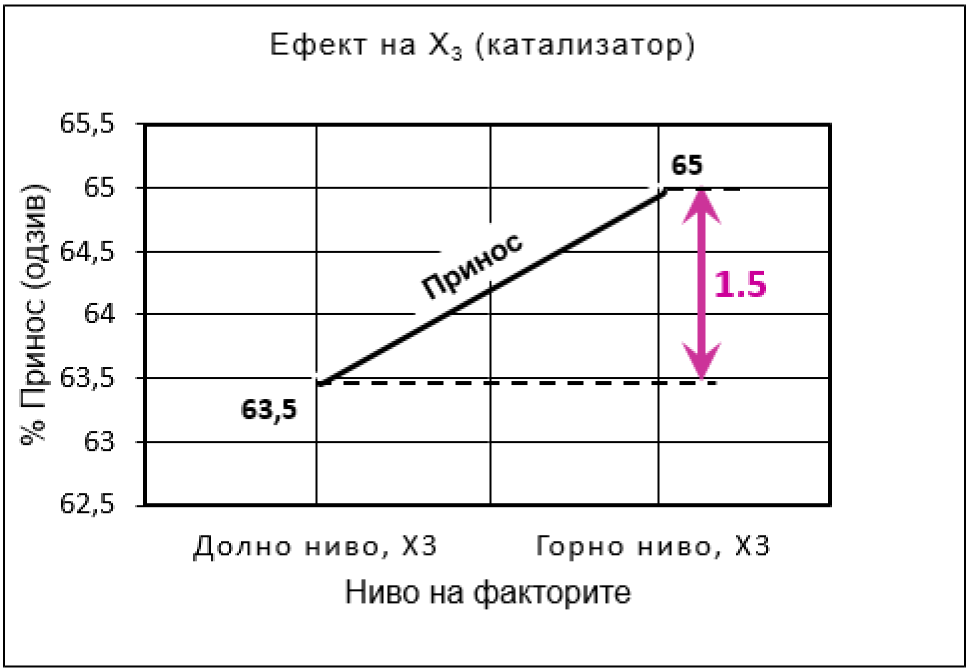
Графички овој ефект ќе го претставиме на следниов начин:



Сумарно ефектот на x_3, користејќи ќе ги средните вредности, ќе го претставиме во табела:

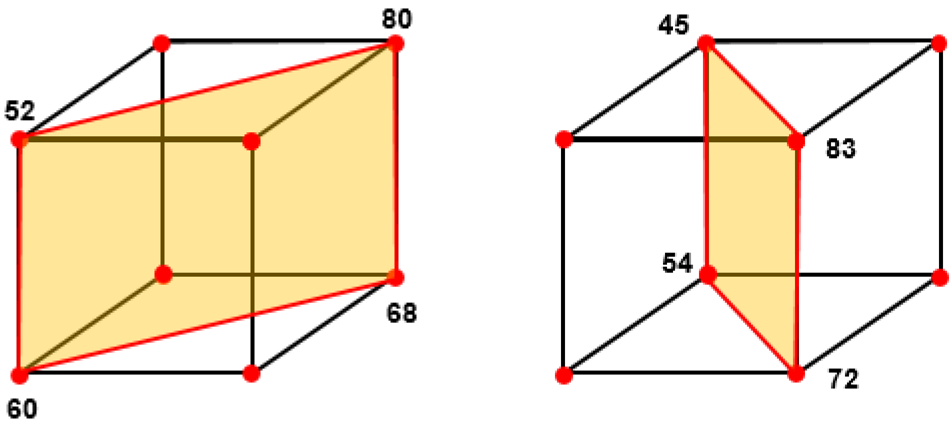
	Горно ниво, X_3	Долно ниво, X_3	Ефект на X_3
	52	60	
	83	72	
	45	54	
	80	68	
Средна вредност	**65**	**63,5**	**1,5**

Интерпретација на главниот ефект X_3



Пресметување и интерпретација на интеракцијата X_1X_2

Ефектот на интеракцијата го пресметуваме како просечната вредност на сите позитивни вредности на производот $x_1 \cdot x_2$ минус сите негативни вредности на истиот производ, како што е претставено на сликата.



Позитивни вредности X_1X_2 Негативни вредности X_1X_2

Таа математичка операција скратено е претставена со равенката:

$$x_1x_2 = \frac{[abc - bc + ab - b - ac + c - a + (1)]}{4n}$$

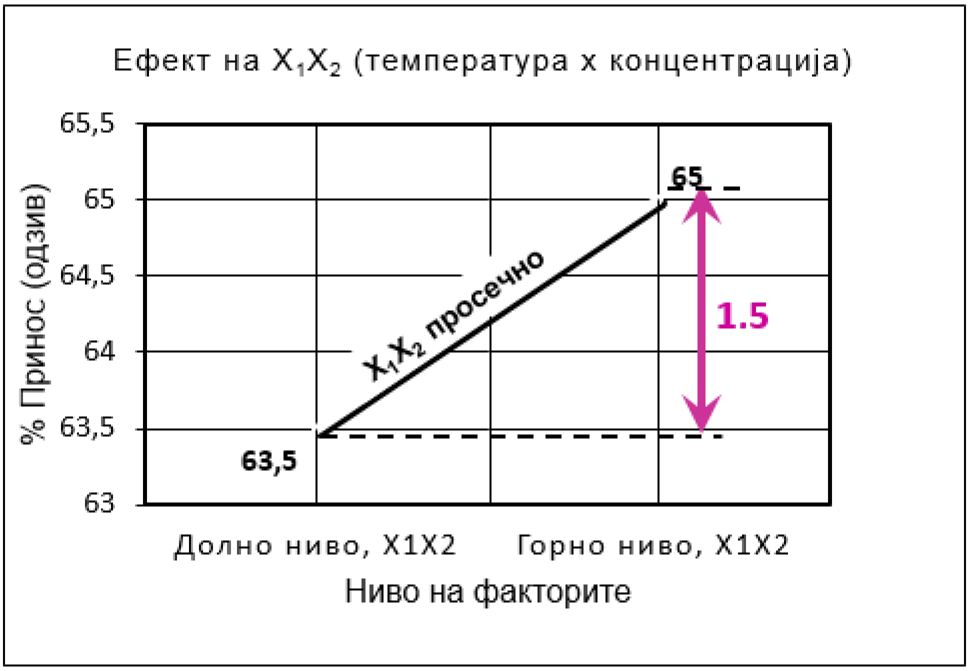
Ефектот на факторот x_1x_2 е:

$$x_1 x_2 = \frac{(80-45+68-54-83+52-72+60)}{4\cdot 1} = 1{,}5$$

Сумарно ефектот на интеракцијата $x_1 \cdot x_2$, користејќи ги средните вредности, ќе го претставиме во табела:

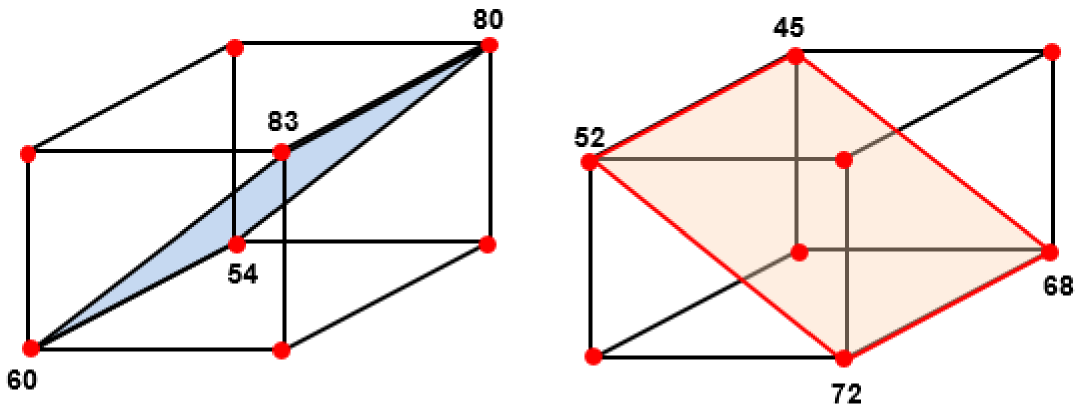
Позитивни вредности на $x_1 \cdot x_2$		Негативни вредности на $x_1 \cdot x_2$		Ефект на $x_1 \cdot x_2$
- -	60	+ -	72	
+ +	68	- +	54	
- -	52	+ -	83	
+ +	80	- +	45	
Средно	**65**		**63,5**	**1,5**

Интерпретација на интерактивниот ефект $x_1 \cdot x_2$`



Пресметување и интерпретација на интеракцијата X_1X_3

Ефектот на интеракцијата го пресметуваме како просечната вредност на сите позитивни вредности на производот $x_1 \cdot x_3$ минус сите негативни вредности на истиот производ, како што е претставено на сликата.



Позитивни вредности X_1X_3 Негативни вредности X_1X_3

Таа математичка операција скратено е претставена со равенката:

$$x_1x_3 = \frac{1}{4n}[(1) - a + b - ab - c + ac - bc + abc]$$

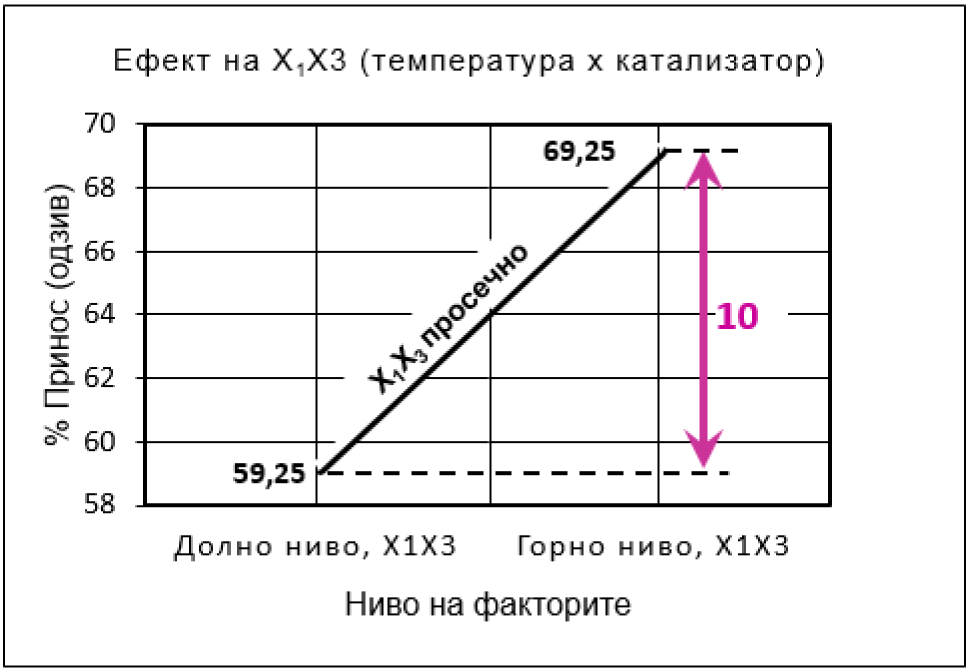
Ефектот на факторот x_1x_3 е:

$$x_1x_3 = \frac{1}{4}(60 - 72 + 54 - 68 - 52 + 83 - 45 + 80) = 10$$

Сумарно ефектот на интеракцијата $x_1 \cdot x_3$, користејќи ги средните вредности, ќе го претставиме во следнава табела:

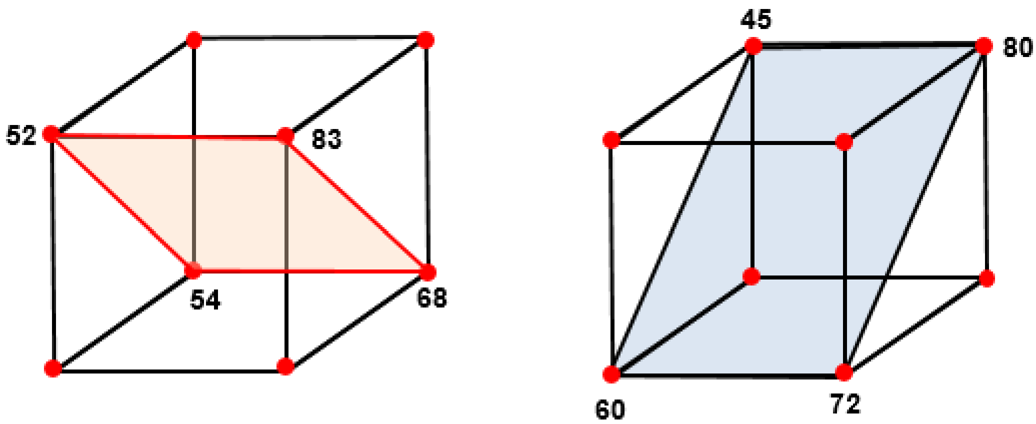
Позитивни вредности на $x_1 \cdot x_3$		**Негативни вредности на $x_1 \cdot x_3$**		**Ефект на $x_1 \cdot x_3$**
- -	60	+ -	72	
+ +	54	- +	45	
- -	83	+ -	68	
+ +	80	- +	52	
Средно	**69,25**		**59,25**	**10**

Интерпретација на интерактивниот ефект $x_1 \cdot x_3$`



Пресметување и интерпретација на интеракцијата X_2X_3

Ефектот на интеракцијата го пресметуваме како просечната вредност на сите позитивни вредности на производот $x_2 \cdot x_3$ минус сите негативни вредности на истиот производ, како што е претставено на сликата.



Позитивни вредности X_2X_3 Негативни вредности X_2X_3

Таа математичка операција скратено е претставена со равенката:

$$x_2x_3 = \frac{1}{4n}[(1) + a - b - ab - c - ac + bc + abc]$$

Ефектот на факторот x_2x_3 е:

$$x_2x_3 = \frac{1}{4}(60+72-54-68-52-83+45+80) = 0$$

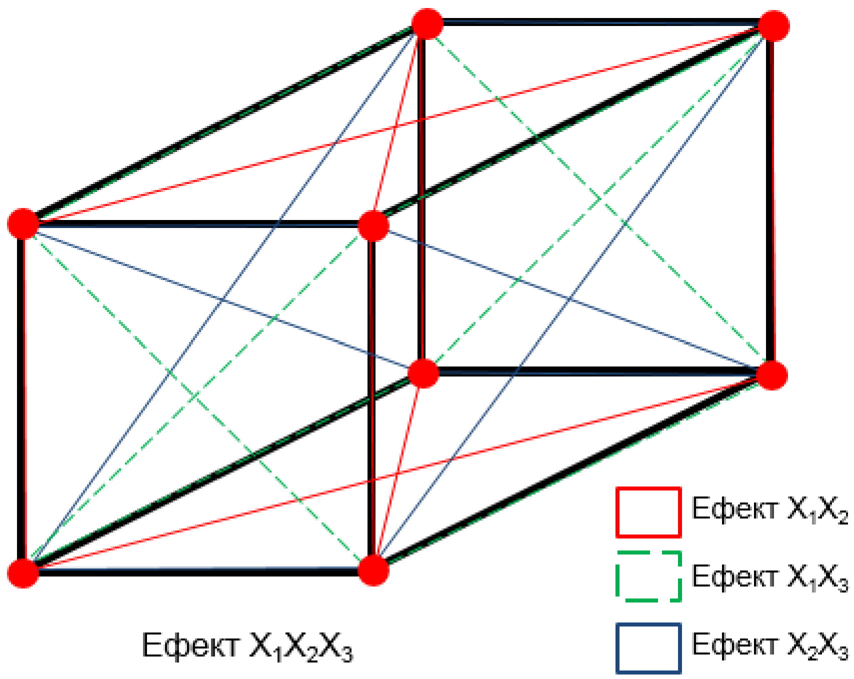
Сумарно ефектот на интеракцијата $x_2 \cdot x_3$, користејќи ги средните вредности, ќе го претставиме во следнава табела:

Позитивни вредности на $x_2 \cdot x_3$		**Негативни вредности на $x_2 \cdot x_3$**		**Ефект на $x_2 \cdot x_3$**
- -	60	+ -	54	
+ +	45	- +	83	
- -	72	+ -	68	
+ +	80	- +	52	
Средно	**64,25**		**64,25**	**0**

Интерпретација на интерактивниот ефект $x_2 \cdot x_3$.

Пресметување и интерпретација на интеракцијата $X_1X_2X_3$

Ефектот на интеракцијата го пресметуваме како просечната вредност на сите позитивни вредности на производот $x_1x_2 \cdot x_3$ минус сите негативни вредности на тој производ. Односно, интеракцијата претставува разлика на сите вредности на производот $x_1x_2x_3$ кога x_3 е на горно ниво, минус сите вредности на производот $x_1x_2 \cdot x_3$ кога x_3 е на долно ниво (види табела), , како што е претставено на сликата. На сликата, заради по прегледност, ефектите се претставени со разнобојни рамки.



Таа математичка операција скратено е претставена со равенката:

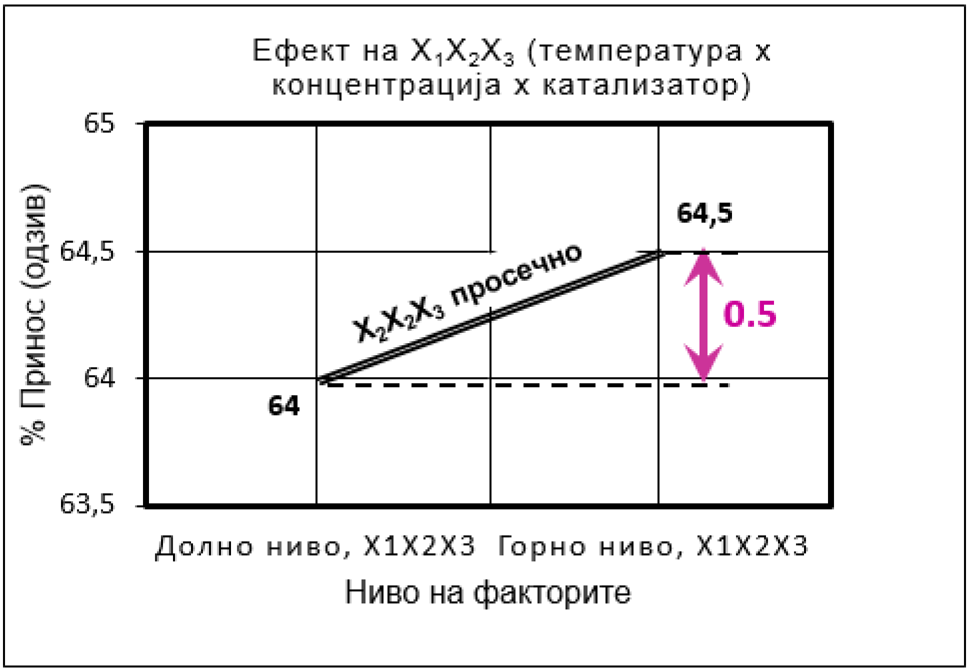
$$x_1x_2x_3 = \frac{1}{4n}[abc - bc - ac + c - ab + b + a - (1)]$$

$$x_1x_2x_3 = \frac{1}{4}(80 - 45 - 83 + 52 - 68 + 54 + 72 - 60) = 0{,}5$$

Сумарно ефектот на интеракцијата $x_1x_2 \cdot x_3$, користејќи ги средните вредности, ќе го претставиме во следнава табела:

Позитивни вредности на $x_1x_2x_3$		Негативни вредности на $x_1x_2x_3$		Ефект на $x_1x_2x_3$
+ - -	72	- - -	60	
- + -	54	+ + -	68	
+ + +	80	- ++	45	
- - +	52	+ - +	83	
Средно	**64,5**		**64**	**0,5**

Интерпретација на интерактивниот ефект $x_1x_2 \cdot x_3$`



Според дизајнот на експериментот ова е последниот ефект кој може да има влијание врз одзивот. Во следната табела, ќе ги претставиме сумарно сите ефекти:

Ефект	Вредност
X_1	23
X_2	-5
X_3	1,5
X_1X_2	1,5
X_1X_3	10
X_2X_3	0
$X_1X_2X_3$	0,5

Прашањето е сега кој од нив е значаен?

Следниот чекор, со оглед дека експериментот е реплициран, е пресметувањето на варијансата според следнава релација:

$$s_j^2 = \frac{1}{k-1}\sum_{i=1}^{k}(y_{ji} - \bar{y}_j)^2$$

За првиот експеримент одзивите се 59 и 61 а нивната средна вредност 60, па варијансата ќе биде:

$$s_1^2 = \frac{1}{2-1}\sum_1^2 (59-60)^2 + (61-60)^2 = \sum_1^2 (-1)^2 + 1^2 = 2$$

На сличен начин ја пресметуваме варијансата и за останатите тестови:

$$s_2^2 = \frac{1}{2-1}\sum_1^2 (74-72)^2 + (70-72)^2 = 8$$

$$s_3^2 = \frac{1}{2-1}\sum_1^2 (58-54)^2 + (50-54)^2 = 32$$

Резултатите од останатите пресметки на варијансата се дадени во табелата:

Тест	y_{j1}	y_{j2}	$\bar{y}$	S_{yj}^2
1	59	61	60	2
2	74	70	72	8
3	50	58	54	32
4	69	67	68	2
5	50	54	52	8
6	81	85	83	8
7	46	44	45	2
8	79	81	80	2

Средната вредност на варијансата е:

$$S_y^2 = \frac{1}{N}\sum_{j=1}^{N} S_y^2$$

$$S_y^2 = \frac{1}{8}\sum_{j=1}^{N} 2+8+32+2+8+8+2+2 = 8$$

Сега можеме да го пресметаме Кохреновиот критериум

$$G_P = \frac{\max s_j^2}{\sum_{i=1}^{N} s_j^2}$$

Најголемата варијанса изнесува 32, па ако замениме во равенката ќе добиеме:

$$G_P = \frac{32}{2+8+32+2+8+8+2+2} = 0{,}5$$

Од табелата, прилог 1, наоѓаме дека Кохреновит критериум изнесува $G_T = 0{,}68$. Бидејќи е исполнет условот $G_P \leq G_T$ нашиот експеримент е репродуктивен и можеме да продолжиме со обработка на резултатите.

Некои коефициенти можат да се покажат занемарливо мали – незначајни. За да се утврди дали тие се значајни или не, пред се, треба да се процени варијансата при која тие се определуваат.

$$s_\beta = \sqrt{\frac{s_y^2}{N}}$$

$$S_\beta = \sqrt{\frac{8}{8}} = 1$$

Треба да се напомене дека кај сефакторниот план на експерименти сите коефициенти се определуваат со еднаква грешка т.е. точност. Се зема дека коефициентите на регресија се значајни ако е исполнет условот:

$$|\beta| \geq S_\beta \cdot t$$

t – табеларна вредност на Студентовиот (Student) критериум од прилог 4.

Сега можеме да ја одредиме вредноста на Студентовиот критериум од прилог 4 за P = 0,95 и f = 8, и наоѓаме дека изнесува t = 2,31

Вредноста на степените на слобода ја одредуваме од релацијата:

$$f = N(k-1) = 8(2-1) = 8$$

За нашиот експеримент :

$$|\beta| \geq s_\beta t \geq 1 \cdot 2{,}31 \geq 2{,}31$$

што значи, секој коефициент по апсолутна вредност треба да биде поголем од или еднаков на 2,31, но претходно да ги пресметаме вредностите на секој коефициент од равенката за модел од прв ред со три фактори:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 + \beta_{123} x_1 x_2 x_3 + \varepsilon$$

Коефициентот β_0 претставува средна вредност од сите одзиви:

$$\beta_0 = \frac{60+72+54+68+52+83+45+80}{8} = 64{,}25$$

Останатите коефициенти ги пресметуваме како една половина од вредноста на соодветниот ефект и ќе ги претставиме во следнава табела:

Ефект	**Вредност**	**Коефициент**	**Половина од вредноста**
X_1	23	β_1	11,5
X_2	-5	β_2	-2,5
X_3	1,5	β_3	0,75
X_1X_2	1,5	β_{12}	0,75
X_1X_3	10	β_{13}	5
X_2X_3	0	β_{23}	0
$X_1X_2X_3$	0,5	β_{123}	0,25

Значајни коефициенти се само оние чија апсолутна вредност го задоволува горниот услов, и тоа ќе го провериме за секој коефициент:

$$|\beta_0| = 64{,}25 > 2{,}31$$

$$|\beta_1| = 11{,}5 > 2{,}31$$

$$|\beta_2| = 2{,}5 > 2{,}31$$

$$|\beta_3| = 0{,}75 < 2{,}31$$

$$|\beta_{12}| = 0.75 < 2{,}31$$

$$|\beta_{13}| = 5 > 2{,}31$$

$$|\beta_{23}| = 0 < 2{,}31$$

$$|\beta_{123}| = 0.25 < 2{,}31$$

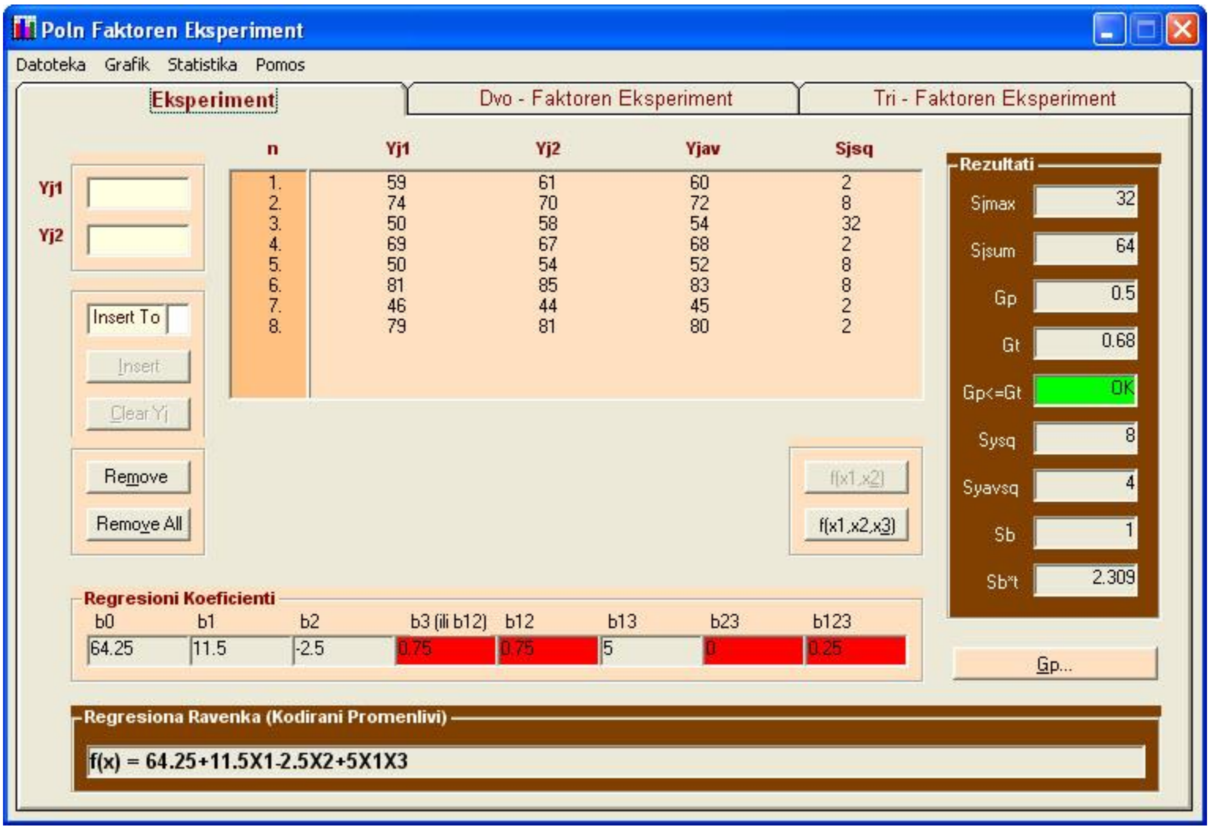
Како што можеме да видиме коефициентите и не го задоволуваат условот што значи дека се безначајни.

Според тоа регресивната равенка со кодирани променливи ќе биде:

$$y = 64{,}25 + 11{,}5x_1 - 2.5x_2 + 5x_1x_3$$

6.1. Проверка на моделот со FFED 7.5

И кај овој експеримент ќе го употребиме софтверскиот пакет FFED 7.5 за да ја провериме мануелно спроведената постапка. На сликата 6.12 можеме да видиме кои податоци ни ги дава софтверот до фазата на одредување на регресивната равенка со кодирани променливи.



Слика 6.12. Одредување на регресивната равенка со кодирани варијабли

Како што можеме да видиме (иако, нотацијата во софтверот е различна), софтверот ни ги дава точно истите вредности на релевантни податоци кои до сега ги пресметавме и точно истата регресивна равенка со кодирани променливи. Можеме да заклучиме дека софтверот ни го потврди моделот со кодирани варијабли.

Сета макотрпна постапка која до сега ја поминавме софтверот ни ја овозможува само за неколку десетина секунди.

За добиената регресивна равенка треба да се провери адекватноста т.е. нејзината способност доволно добро да ја опишува површината на одзивот. Таа проверка се врши со помош на критериумот на Фишер (Fisher), видете прилог 3. Равенката е адекватна ако е исполнет условот:

$$F_P \leq F_T$$

$$F_P = \frac{s_{ad}^2}{s_y^2}$$

$$s_{ad}^2 = \frac{\sum_{j=1}^{N} (\bar{y}_j - y_j)^2}{N-k-1}$$

Во оваа равенка y_j ги означува вредностите на одзивот пресметани според регресивната равенка. На пример, за првата серија, ако ги замениме вредностите за кодираните фактори ќе добиеме:

$$y_1 = 64{,}25 + 11{,}5(-1) - 2.5(-1) + 5(-1)(-1) = 60{,}25$$
$$y_2 = 64{,}25 + 11{,}5(+1) - 2.5(-1) + 5(+1)(-1) = 73{,}25$$
$$y_3 = 64{,}25 + 11{,}5(-1) - 2.5(+1) + 5(-1)(-1) = 55{,}25$$
$$y_4 = 64{,}25 + 11{,}5(+1) - 2.5(+1) + 5(+1)(-1) = 68{,}25$$
$$y_5 = 64{,}25 + 11{,}5(-1) - 2.5(-1) + 5(-1)(+1) = 50{,}25$$
$$y_6 = 64{,}25 + 11{,}5(+1) - 2.5(-1) + 5(+1)(+1) = 83{,}25$$
$$y_7 = 64{,}25 + 11{,}5(-1) - 2.5(+1) + 5(-1)(+1) = 45{,}25$$
$$y_8 = 64{,}25 + 11{,}5(+1) - 2.5(+1) + 5(+1)(+1) = 78{,}25$$

Ова се вредностите на одзивот што ги добивме од моделот со замена на соодветните вредности на варијаблите , +1 и -1, т.е. со нивните вредности кога се на долно и горно ниво, соодветно. Во следнава табела напоредно ќе ги презентираме вредностите добиени од експериментот $\bar{y}_j$ со вредностите добиени од моделот, $\boldsymbol{y_j}$.

Експеримент	$\bar{y}_j$	y_j	$\bar{y}_j - y_j$	$(\bar{y}_j - y_j)^2$
1	60	60,25	-0,25	0,0625
2	72	73,25	-1,25	1,5625
3	54	55,25	-1,25	1,5625
4	68	68,25	-0,25	0,0625
5	52	50,25	1,75	3,0625
6	83	83,25	-0,25	0,0625
7	45	45,25	-0,25	0,0625
8	80	78,25	-1,75	3,0625

Сега ја бараме сумата на сите вредностите во последната колона:

$$\sum_{j=1}^{8}(\bar{y}_j - y_j)^2 = 0.0625 + 1.5625 + 1.5625 + 0.0625 + 3.0625 + 0.0625 +$$

$$+ 0.0625 + 3.0625 = 9,5$$

Бројот на степените на слобода е:

$$N - k - 1 = 8 - 2 - 1 = 5$$

од тука следи дека:

$$s_{ad}^2 = \frac{9,5}{5} = 1,9$$

Претходно пресметавме дека:

$$s_y^2 = 8$$

$$F_P = \frac{1,9}{8} = 0,2375$$

Табеларната вредност на Фишеровиот критериум ја наоѓаме од прилогот 3 за P = 0,95 , f = 5 и N=8, наоѓаме дека Фишеровиот критериум изнесува 3,69, и бидејќи е задоволен условот

$$F_P \le F_T$$

заклучуваме дека моделот е адекватен.

Со тоа заклучуваме дека моделот на експериментот

$$y = 64,25 + 11,5x_1 - 2.5x_2 + 5x_1x_3$$

е сосема добро го објаснува процесот..

Одредување на моделот во инженериски единици

Во практика, по практично е ако место со кодирани, работиме со инженериски варијабли. За да го добиеме одзивот во вистинските единици потребно е да извршиме конвертирање на кодираните варијабли во инженериски или натурални варијабли. Во релацијата за претворање на варијаблите од еден вид во друг ќе

$$x = \frac{X - X_0}{\Delta X}$$

X – инжeнериска варијабла (фактор)

X_0 – средното ниво на инженериската варијабла

ΔX – интервалот на промена на X

x – кодирана варијабла

За секој од трите фактори ќе извршиме соодветна замена на кодираните фактори (мало x) во инженериски (големо X) според следниве релации:

$$x_1 = \frac{X_1 - 170}{10} \qquad x_2 = \frac{X_2 - 25}{15} \qquad x_3 = \frac{X_3 - 1{,}5}{0{,}5}$$

Факторот x_1 е температурата која има вредности од 160 °C до 180 °C. Нултото ниво на овој фактор е 170 °C додека интервалот на варирање е 10 °C (170 ± 10 °C). Соодветно за факторот x_2 (концентрацијата) кој има вредности од 10% до 40 %. Нултото ниво е 25 %, додека интервалот, 15 % (25 ± 15%).

Со оглед дека факторот x_3 е квалитативна (катализатор) променлива, може да има само две вредности, катализатор A – кој го означуваме со „1“ и катализаторот B, кој го означуваме со „2“. Тоа го правиме за да можеме да го вклучиме во равенката. За овој фактор земаме (фиктивно) нулта вредност 1,5 и интервал 0,5 (1,5 ± 0,5).

Кај овој фактор не може да има меѓувредности, меѓу 1 и 2 бидејќи по својата природа претставува Буленов фактор Со квалитативни променливи не се вршат математички операции па затоа ја правиме оваа супституција

Сега во равенката со кодирани фактори можеме да извршиме замена како би добиле равенка во инженериски единици:

$$y = 64{,}25 + 11{,}5(\frac{X_1 - 170}{10}) - 2.5x(\frac{X_2 - 25}{15}) + 5(\frac{X_1 - 170}{10})(\frac{X_3 - 1{,}5}{0{,}5})$$

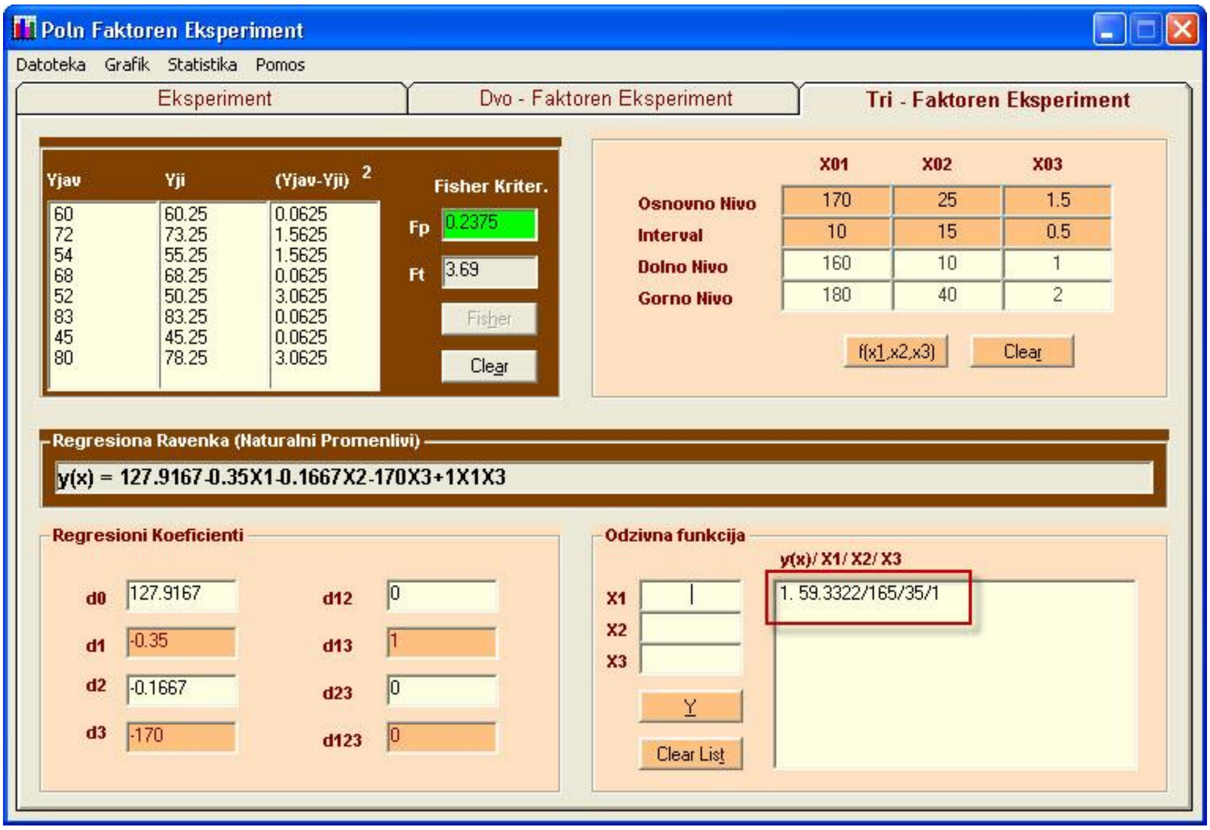
$$y = 127{,}92 - 0.35X_1 - 0{,}17X_2 - 170X_3 + X_1X_3$$

Ако во оваа равенка ги замениме кои било вредности, во опсегот што е предмет на испитување, за факторите во инженериски единици ќе го добиеме одзивот во %. На пример да землеме дека температурата е 165 °C, концентрацијата – 35 % и ако работиме со катализаторот A (=1), со замена на овие вредности ќе добиеме принос во %.

$$y = 127{,}92 - 0.35(165) - 0{,}17(35) - 170(1) + (165)(1) = 59{,}3\%$$

При избраното ниво на фактор приносот на процесот е 59,3 %.

Од сликата 6.13 можеме да видиме дека со програмата за планирање на фактор експериментите ја добивме точно, истата регресиона равенка со инженериски променливи. За да го провериме одзивот при ниво на променливите од горниот пример во рамката **Odzivna funkcija** и ги зададовме



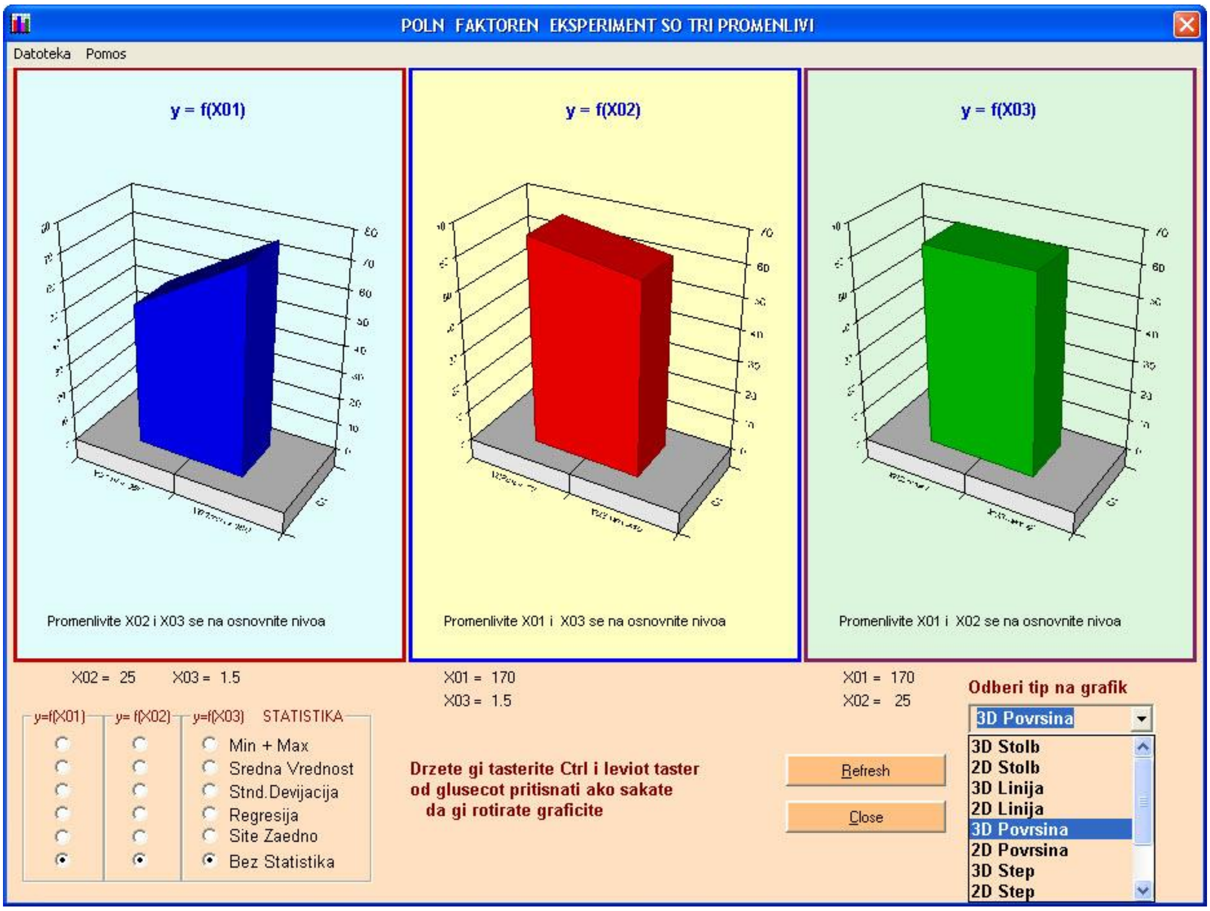
Слика 6.13. Презентирање на регресивната равенка во инженериски единици

на програмата истите вредности на променливите и ја добивме точно истата вредност на одзивот, 59,3 % (на сликата 6.13 во црвена рамка).

Работата на програмата со три фактори незначително се разликува од онаа со два фактори. Можеме да забележиме дека во интерфејсот на експериментот со три фактори му недостасува копчето за креирање на графички приказ. Причината за тоа е многу проста, едноставно на интерфејсот, физички нема место за тоа копче, со оглед дека сега се прикажани многу повеќе податоци отколку при експеримент со два фактори.

Можноста за креирање на графички приказ е сместена во главното мени **Grafik** од каде ја избираме оваа опција за да креираме графички приказ, слика 6.14.

Сè она што зборуваме кај експериментот со два фактори (поединечно печатење и снимање на графичките прикази, визуелна статистика, креирање на извештај, итн.) е апликативно и за експериментот со три фактори.



Слика 6.14. 3Д графички приказ за ефектот на променливите: температура, концентрација и катализатор.

7. ЗАДАЧИ ЗА САМОСТОЈНО РЕШАВАЊЕ

Планирањето на експерименти, како што се уверивме, претставува подолга постапка која вклучува многу аритметика. Мала, тривијална грешка може да не доведе во ќор-сокак и неможност постапката да ја истераме до крај. Од причини за да можете да ја контролирате вашата работа во секоја фаза на планираниот експеримент, како водич кон планирањето ќе биде презентирана слика од интефејсот на FFED 7.5 со поставката на експериментот и решението.

Пробајте целосно самостојно да ги решите задачите (без да ја гледате сликата) и само ако дојдете во ситуација да не можете самостојно да дојдете до решението, погледајте ја сликата за да видите каде сте згрешиле. По корекцијата на грешката, повторно продолжете самостојно до крајното решение.

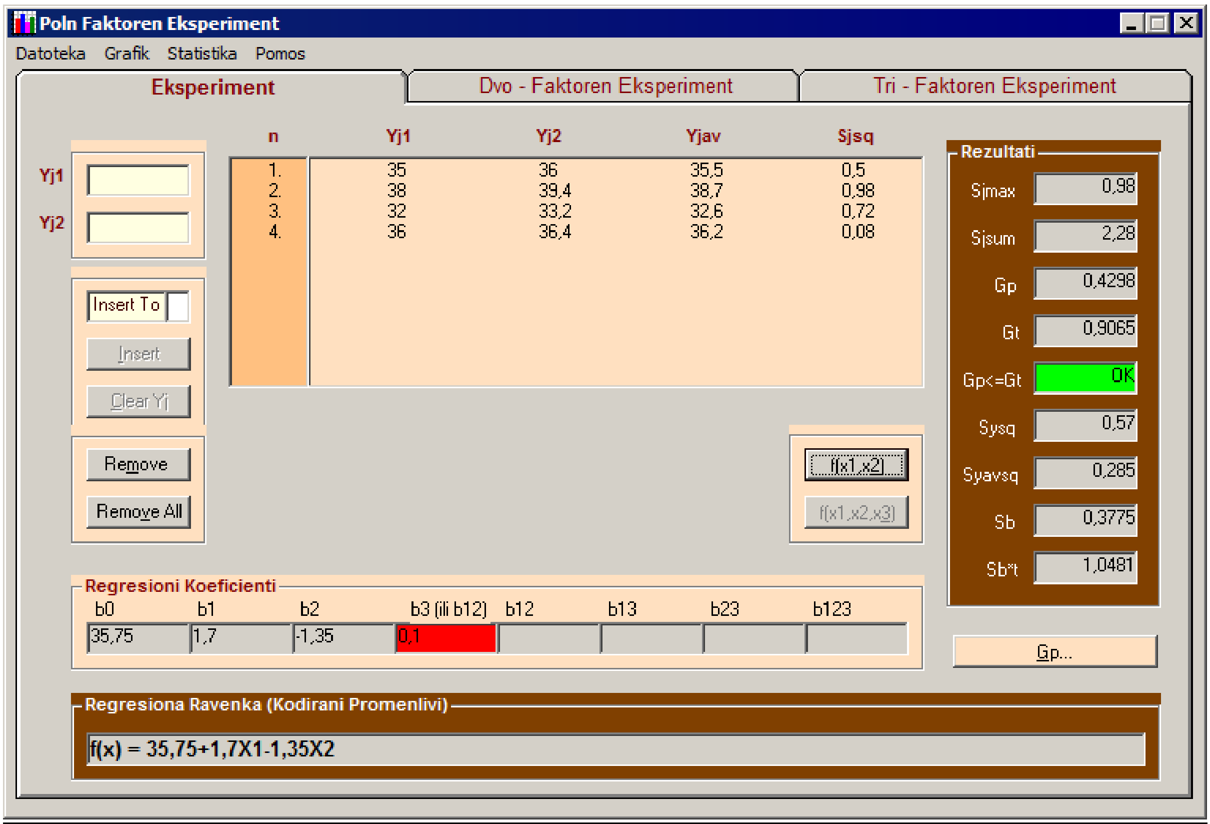
7.1. Да го разгледаме хемискиот процес во кој приносот на продуктите на реакцијата y (%) зависи од температурата на реакционата смеса x_1 (^oC) и концентрацијата на реагентите x_2 (%). Со помош на сефакторен експеримент треба да се најде математички опис на тој процес ако температурата варира меѓу 45 ^oC и 55 ^oC, а концентрацијата меѓу 24 % и 26 %.

Матрицата на планирање на експериментот и резултатите од сефакторниот експеримент се дадени во табелата,

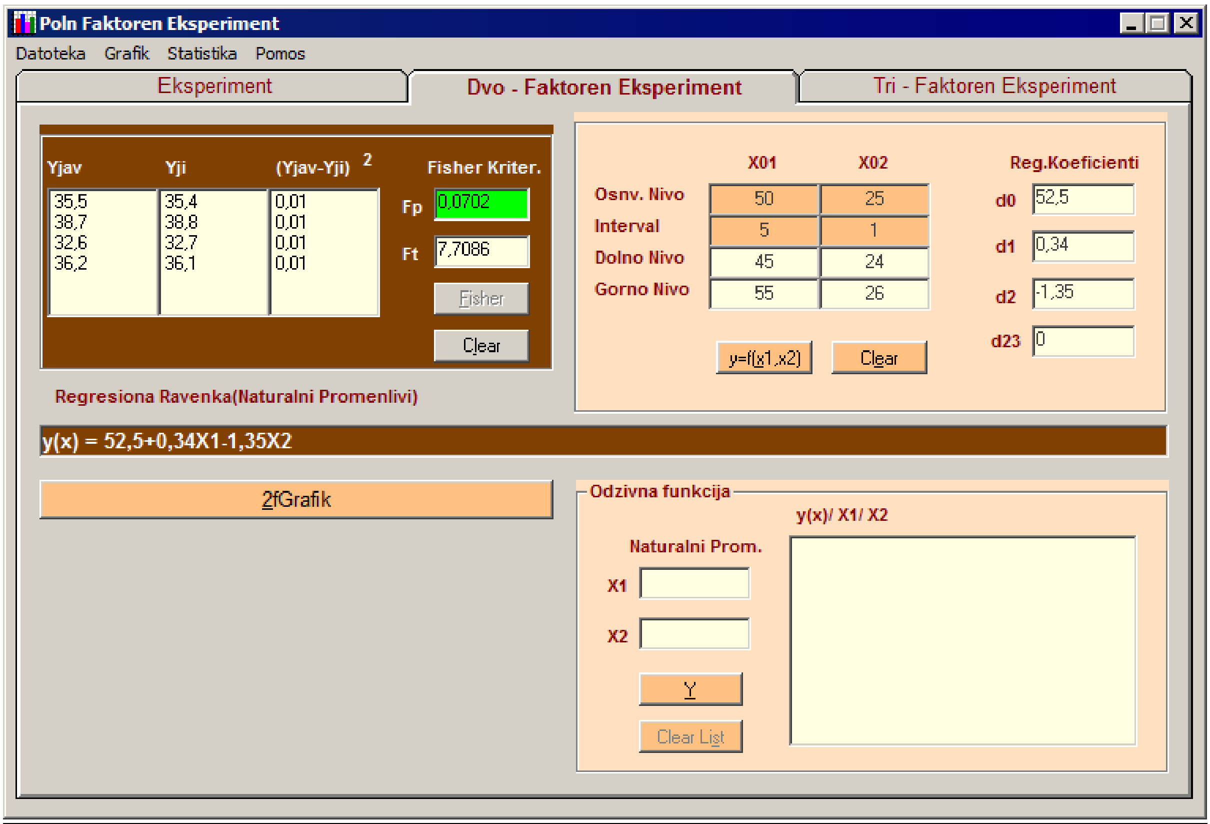
Бр. на пробата	**x_1 ^oC**	**x_2 %**	**y_{j1}**	**y_{j2}**	$\bar{y}$
1	45	24	35	36	35,5
2	55	24	38	39,4	38,7
3	45	26	32	33,2	32,6
4	55	26	36	36,4	36,2

<u>Водич кон решението и решение:</u>

- Равенка со кодирани варијабли



- Равенка со натурални варијабли

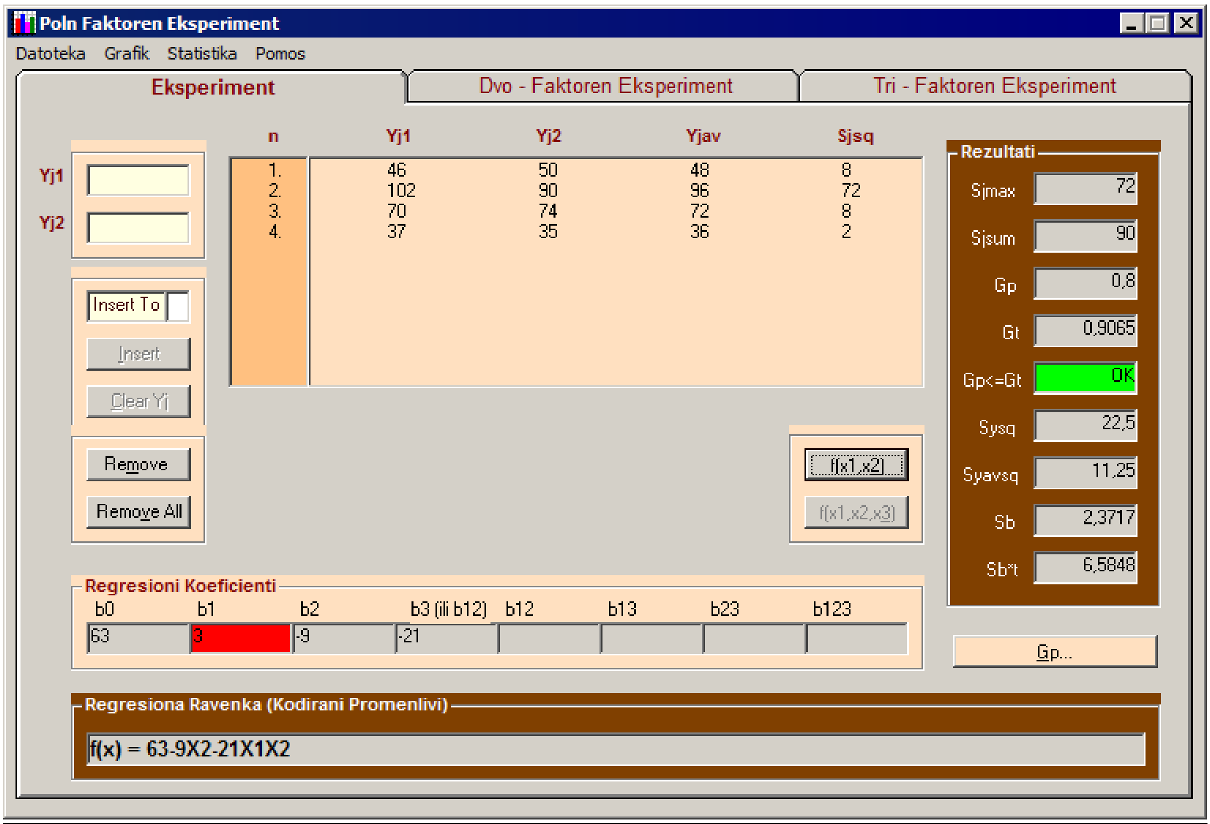


7.2. Продуктивноста на еден процес зависи од факторите А и В. План - матрицата со резултатите е дадена во табелата,

Бр. на пробата	А (X1)	В (X2)	y_{j1}	y_{j2}	$\bar{y}$
1	-1	-1	46	50	48
2	+1	-1	102	90	96
3	-1	+1	70	74	72
4	+1	+1	37	35	36

Да се определи равенката со кодирани варијабли за процесот.

<u>Водич кон решението и решение</u>



7.3. Дизјниран е експеримент за проучување на ефектот на протокот на кислород и времето врз дебелината на оксидната облога нанесена на метална подлога во текот на сува оксидација. Секој фактор е проучуван на две нивоа, види табела:

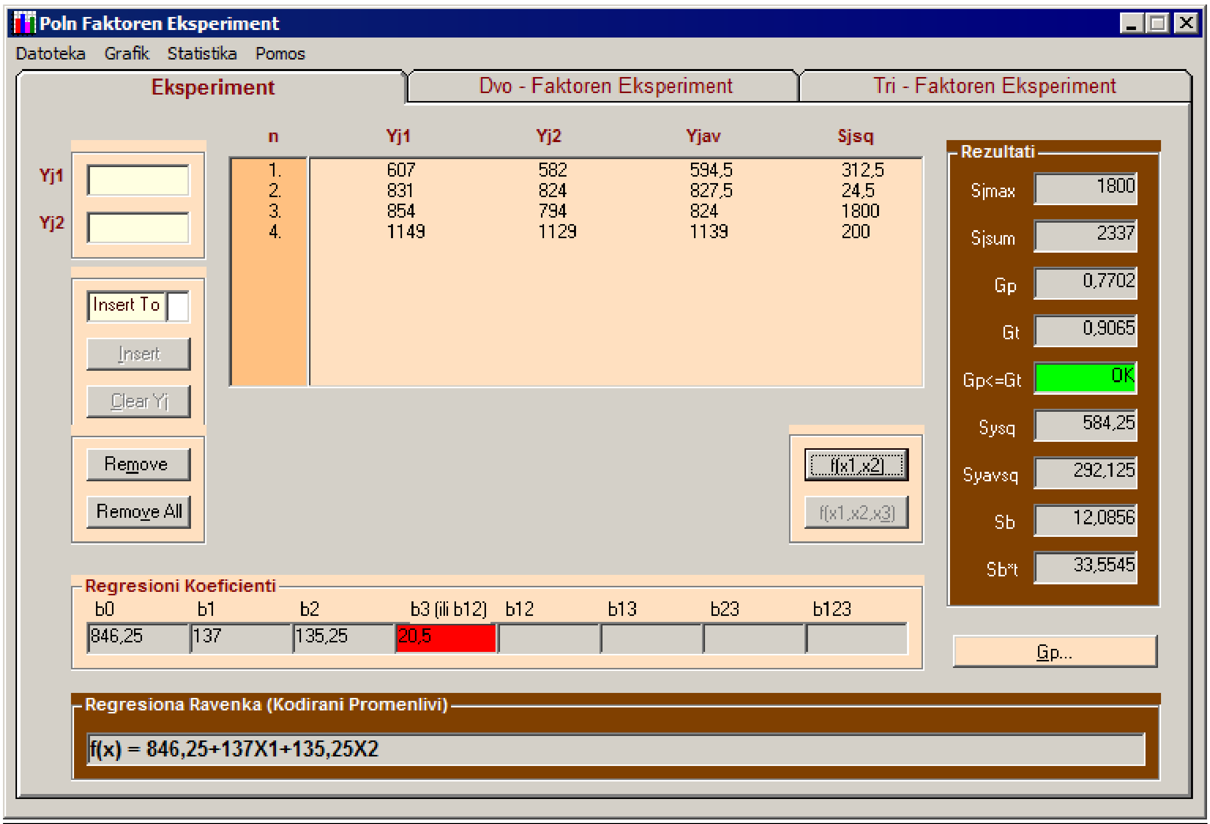
Ниво	Фактор А Проток на O_2 (*l/min*)	Фактор В Време на оксидација (*min*)
Долно (-)	1,5	21
Горно (+)	9	40

Правени се по две реплики, а дебелината на облогота е мерена во ангстреми. Податоците се дадени во оваа табела:

Бр. на пробата	A X1	B X2	y_{j1}	y_{j2}
1	-1	-1	607	582
2	+1	-1	831	824
3	-1	+1	854	794
4	+1	+1	1149	1129

Да се одреди моделот за предвидување на дебелината од оксидот.

<u>Водич кон решението и решение</u>



7.4. Приносот на еден хемиски процес зависи од температурата , притисокот и времето на реакција. Процесните параметри се дадени во оваа табела:

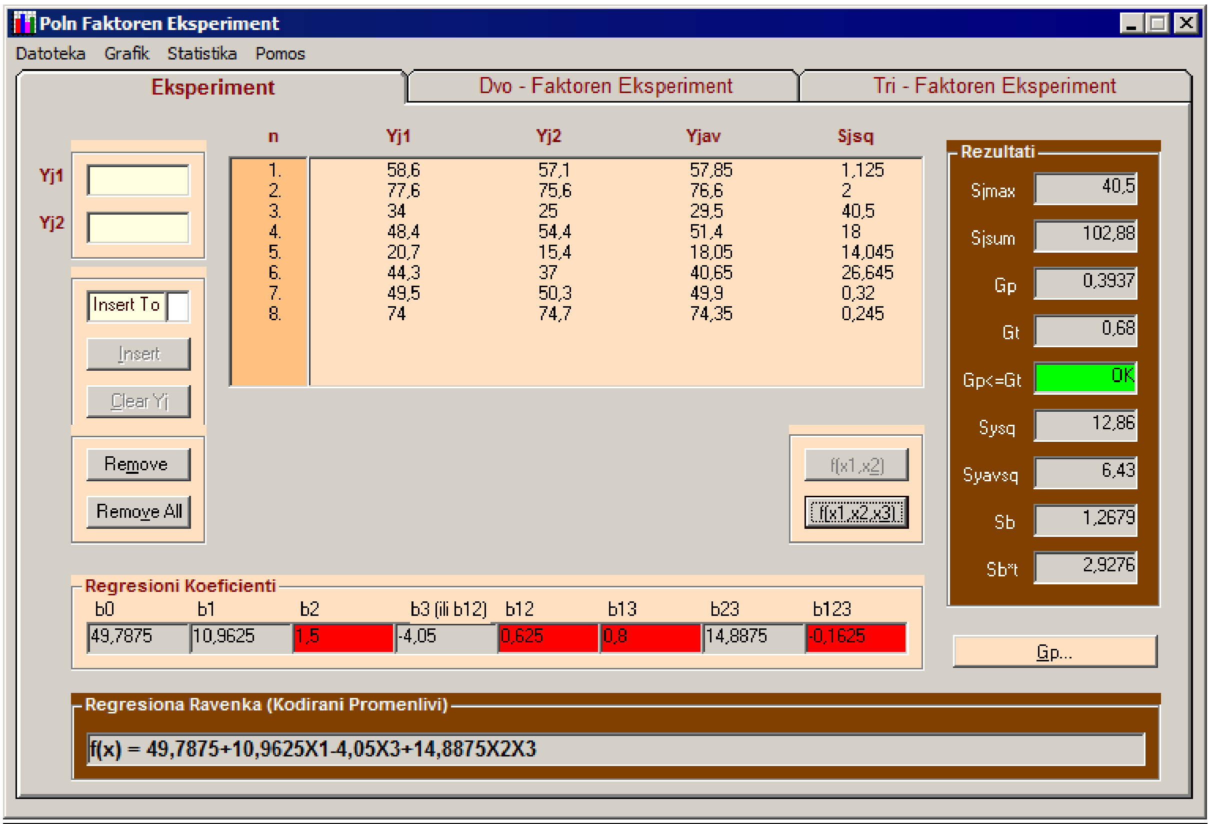
Параметар	Ознака	Долно ниво	Горно ниво
Температура	X_1	80 ºC	120 ºC
Притисок	X_2	5 Pa	7 Pa
Реакционо време	X_3	5 min	15 min

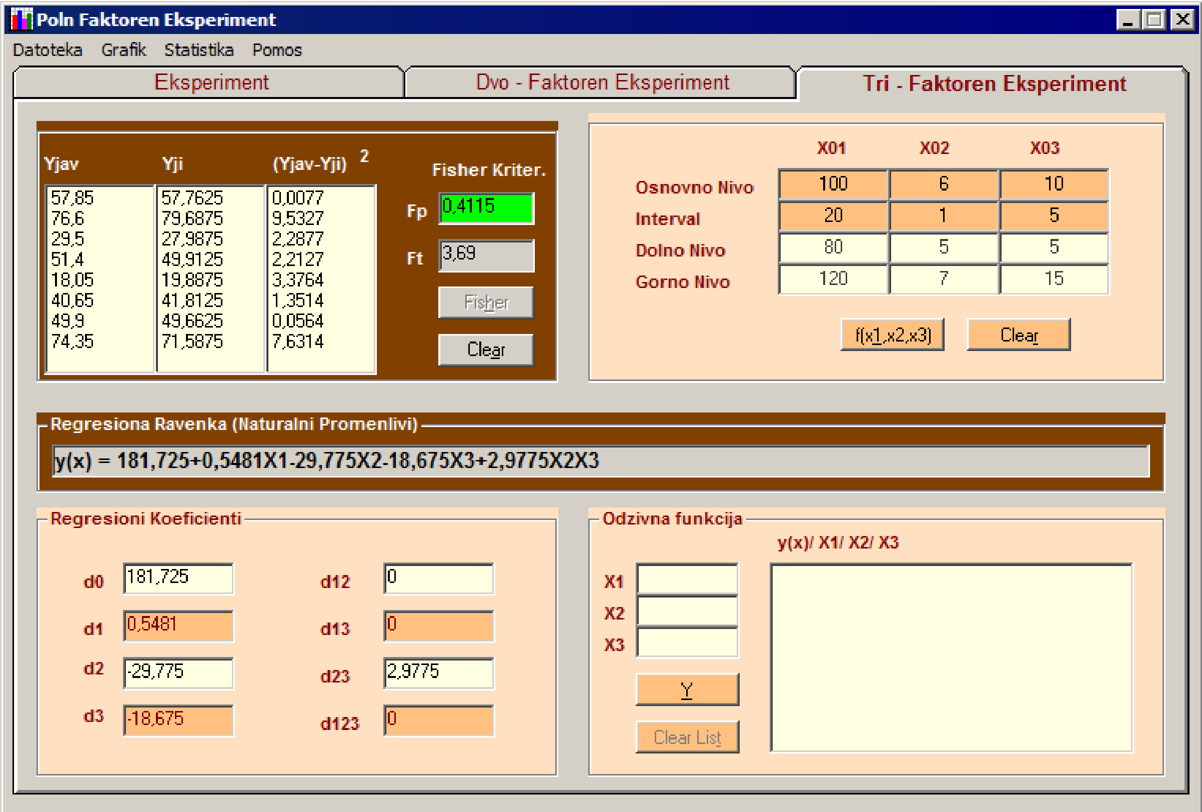
План - матрицата на експериментот и резултатите се:

Проба	X_1	X_2	X_3	Принос 1	Принос 2
1	80	50	5	58,6	57,1
2	120	50	5	77,6	75,6
3	80	70	5	34,0	25,0
4	120	70	5	48,4	54,4
5	80	50	15	20,7	15,4
6	120	50	15	44,3	37,0
7	80	70	15	49,5	50,3
8	120	70	15	74,0	74,7

Да се определи моделот на процесот (равенките) во кодирани и инженериски варијабли.

Водич кон решението и решение:



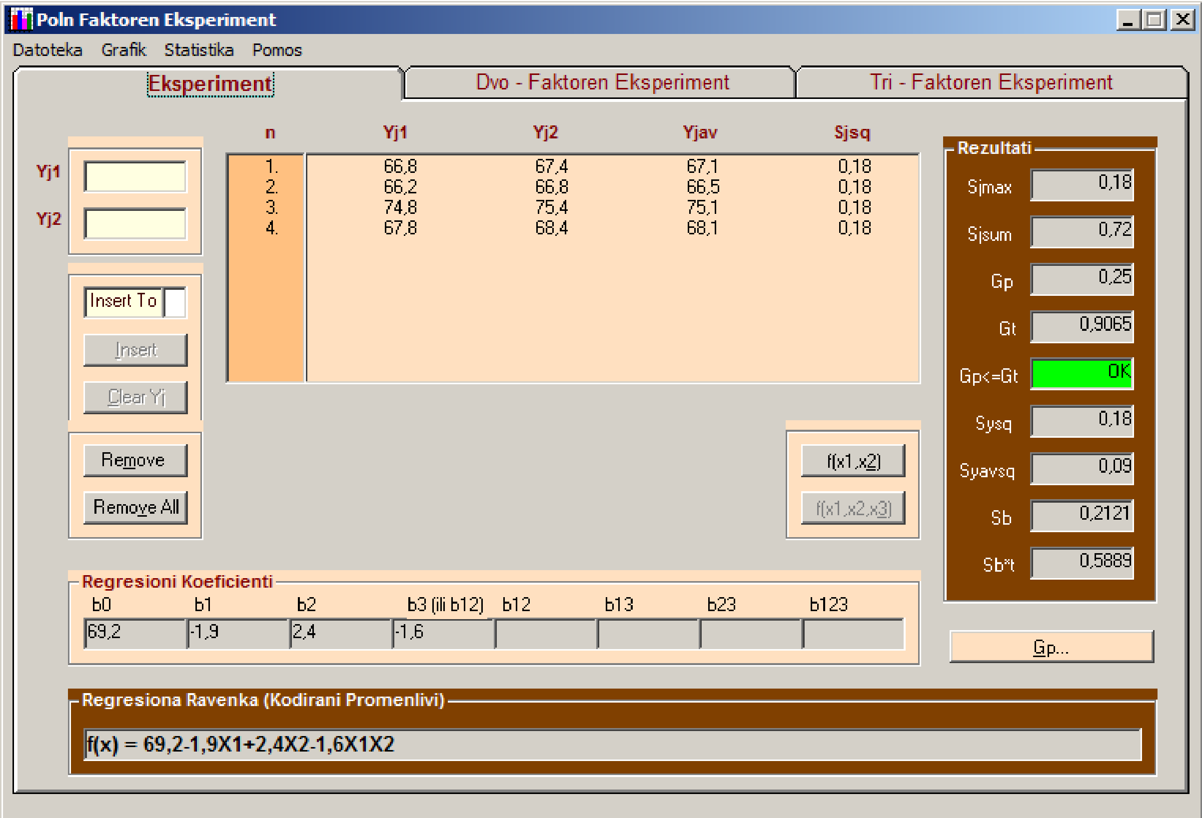


7.5. Одзивите од еден планиран експеримент се дадени во табелата:

Бр. на пробата	X1	X2	y_{j1}	y_{j2}
1	-1	-1	66,8	67,4
2	+1	-1	66,2	66,8
3	-1	+1	74,8	75,4
4	+1	+1	67,8	68,4

Да се определи регресивната равенка со кодирани варијабли.

<u>Водич кон решението и решение:</u>


Poln Faktoren Eksperiment
Datoteka Grafik Statistika Pomos
Eksperiment
Dvo - Faktoren Eksperiment
Tri - Faktoren Eksperiment
Yj1
Yj2
Insert To
Insert
Clear Yj
Remove
Remove All
n Yj1 Yj2 Yjav Sjsq
1. 66,8 67,4 67,1 0,18
2. 66,2 66,8 66,5 0,18
3. 74,8 75,4 75,1 0,18
4. 67,8 68,4 68,1 0,18
f(x1,x2)
f(x1,x2,x3)
Rezultati
Sjmax 0,18
Sjsum 0,72
Gp 0,25
Gt 0,9065
Gp<=Gt OK
Sysq 0,18
Syavsq 0,09
Sb 0,2121
Sb*t 0,5889
Gp...
Regresioni Koeficienti
b0 b1 b2 b3 (ili b12) b12 b13 b23 b123
69,2 -1,9 2,4 -1,6
Regresiona Ravenka (Kodirani Promenlivi)
f(x) = 69,2-1,9X1+2,4X2-1,6X1X2

Литература:

[1] G.Ruxton, N.Colegrave, Experimental Design for the Life Science, Oxford University Press, 2013
[2] J.Goupy, L.Creighton, Introduction to Design of Experiments, SAS Publishing, 2007.
[3] A.Dean, D.Voss, Design and Analysis of Experiments, Springer-Verlag New York, Inc. 1999
[4] D. Montgomery, Design and Analysis of Experiments, 5th Edition, John Wiley & Sons, Inc, 2001.
[5] Montgomery, Design and Analysis of Experiments – Solution Manual for the 5th Edition, John Wiley & Sons, Inc, 2001.
[6] A. Wheeler, A. Ganji, Introduction to Engineering Experimentation, 3td Edition, Prentice Hall, 2010
[7] G.Quinn, M. Koeugh, Experimental Design and Data Analysis for Biologists, Cambridge University Press, 2002
[8] Thomas Ryan, Modern Experimental Design, Wiley Interscience John Wiley & Sons, Inc, 2007
[8] Н.Спирин, В.Лавров, Методи планированија и обработки резултатов инженерного експеримента, ГОУ ВПО УГТУ – УПИ, Екатеринбург, 2004
[9] В.Риков, В. Иткин, Математическаја статистика и планирование експерименатa, РГУ нефти и газа им. И.М. Губкина, Москва, 2008
[10] И. Агајинц, А. Орлов, Планирование експеримента и анализ данних, МИТХТ, Ломоносов, Москва, 1998
[11] С.А. Семонов, Планирование експеримента в химии и хемическој технологии, Москва, 2005 (www.mitht.ru/e-library)
[12] American National Standard ASTM E 178-00, Standard Practice for Dealing with Outlying Observations
[13] S. Deming, S. Morgan, Experimental Design: a chemometric approach, Elsevier, Amsterdam, 1987
[14] Джонсон, Ф.Лион, Статистика и планирование експеримента в технике и науке, МИР, Москва, 1980
[15] С.Н. Саутин, Планирование експеримента в химии и химическој технологии, издатељство „Химија“, 1975
[16] Ч. Хикс, Основние принципи планированија експеримента, издатељство „МИР“, Москва, 1967
[17] Г.И. Крсовскии, Г.Ф. Филаретов, Планирование експеримента, издатељство НГУ „Ленин“, Минск, 1982
[18] Ј.П. Адлер, Введение в планирование експеримента, издатељство, „Металургија“, Москва, 1969.
[19] К. Хартман и др, Планирование експеримента в иследовании технологических процесов, издатељство „МИР“, Москва, 1977

[20] Jiny Antony, Design of Experiments for Engineers and Scientists, Elsevier Science and Technology Books, 2003
[21] F.Sani, J.Todman, Experimental Design and Statistics for Psychology, Blackwell Publishing, Oxford, 2006
[22] A.Clewer, D.Scarisbrick, Practical Statistics and Experimental Design for Plant and Crop Science, John Wiley & Sons, Inc, 2009
[23] T.Lorenzen, V. Anderson, Design of Experiments a No-name Approach, CRC Press, 1993
[24] Gary Oehlert, A First Course in Design and Analysis of Experiments, Gary W. Oehlert, 2010
[25] Е.И. Короткова, Планирование и организација екперимента, Издатељство Томского Политехнического Университета, 2010
[26] J.Kiefer, Optimum Designs in Regression Problems, Annals of Mathematical Statistics, Vol.30,pp.271-294

Прилог 1*. Критични вредности за горно 0.05 ниво на значајност или P = 0,95*

Кохренов критериум

(Во редот кој е врамен со црвена рамка е вредноста на **к,**

N *е број на серии на паралелни опити*

N	2	3	4	5	6	7	8	9	10	11	17	37	145	∞
2	.9985	.9750	.9392	.9057	.8772	.8534	.8332	.8159	.8010	.7880	.7341	.6602	.5813	.5000
3	.9669	.8709	.7977	.7457	.7071	.6771	.6530	.6333	.6167	.6025	.5466	.4748	.4031	.3333
4	.9065	.7679	.6841	.6287	.5895	.5598	.5365	.5175	.5017	.4884	.4366	.3720	.3093	.2500
5	.8412	.6838	.5981	.5441	.5065	.4783	.4564	.4387	.4241	.4118	.3645	.3066	.2513	.2000
6	.7808	.6161	.5321	.4803	.4447	.4184	.3980	.3817	.3682	.3568	.3135	.2612	.2119	.1667
7	.7271	.5612	.4800	.4307	.3974	.3726	.3535	.3384	.3259	.3154	.2756	.2278	.1833	.1429
8	.6798	.5157	.4377	.3910	.3595	.3362	.3185	.3043	.2926	.2862	.2462	.2022	.1616	.1250
9	.6385	.4775	.4027	.3584	.3286	.3067	.2901	.2768	.2659	.2568	.2226	.1820	.1446	.1111
10	.6020	.4450	.3733	.3311	.3029	.2823	.2666	.2541	.2439	.2353	.2032	.1655	.1308	.1000
12	.5410	.3924	.3264	.2880	.2624	.2439	.2299	.2187	.2098	.2020	.1737	.1403	.1100	.0833
15	.4709	.3346	.2758	.2419	.2195	.2034	.1911	.1815	.1736	.1671	.1429	.1144	.0889	.0667
20	.3894	.2705	.2205	.1921	.1735	.1602	.1501	.1422	.1357	.1303	.1108	.0879	.0675	.0500
24	.3434	.2354	.1907	.1656	.1493	.1374	.1286	.1216	.1160	.1113	.0942	.0743	.0567	.0417
30	.2929	.1980	.1593	.1377	.1237	.1137	.1061	.1002	.0958	.0921	.0771	.0604	.0457	.0333
40	.2370	.1576	.1259	.1082	.0968	.0887	.0827	.0780	.0745	.0713	.0595	.0462	.0347	.0250
60	.1737	.1131	.0895	.0765	.0682	.0623	.0583	.0552	.0520	.0497	.0411	.0316	.0234	.0167
120	.0998	.0632	.0495	.0419	.0371	.0337	.0312	.0292	.0279	.0266	.0218	.0165	.0120	.0083
∞	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000

Прилог 2. Случајни броеви

Row number	Column number											
	1	2	3	4	5	6	7	8	9	10	11	12
1	20	288	215	98	44	202	323	226	162	262	248	276
2	331	171	204	131	185	39	183	338	209	352	285	333
3	250	142	85	145	348	55	327	137	289	135	146	282
4	259	226	225	215	219	306	267	181	29	328	247	280
5	143	103	64	332	199	215	159	145	84	87	135	250
6	234	261	211	152	181	198	59	179	99	29	106	342
7	360	61	177	310	265	150	23	8	83	122	303	112
8	86	103	313	117	254	284	17	347	296	196	153	251
9	120	191	52	93	297	209	290	251	155	210	47	357
10	6	105	155	77	191	147	356	245	262	36	268	32
11	37	2	259	115	3	24	105	115	130	246	256	61
12	357	180	52	190	311	20	48	348	12	322	223	355
13	326	227	40	93	217	100	336	262	272	83	179	48
14	336	322	174	335	94	358	303	337	253	253	42	163
15	240	213	217	8	184	72	219	136	211	64	154	217
16	289	327	291	216	112	106	147	181	33	311	7	111
17	140	32	297	319	342	15	30	53	54	55	10	139
18	125	81	128	59	121	236	307	264	328	120	247	127
19	351	21	132	172	148	91	93	334	74	126	223	62
20	80	91	98	280	41	146	219	271	248	246	185	179
21	198	255	26	201	330	134	100	19	41	50	192	349
22	80	225	321	260	161	330	260	333	195	145	42	169
23	66	54	140	179	179	46	311	87	340	341	287	98
24	19	185	137	346	163	224	13	72	208	48	124	288
25	281	321	20	16	5	48	70	91	122	305	240	75
26	126	61	317	336	188	184	23	31	174	280	115	88
27	355	267	117	253	186	72	1	221	345	86	89	355
28	321	327	358	78	320	261	187	325	100	91	179	184
29	201	341	29	14	341	63	253	249	226	170	311	127
30	308	233	229	344	215	341	318	334	144	326	70	148
31	195	42	88	337	122	97	84	215	298	168	300	312
32	22	115	136	100	110	73	219	290	133	70	23	55
33	48	190	121	352	200	249	135	55	336	17	263	269
34	175	225	341	307	283	280	359	348	143	230	102	283
35	133	17	293	356	45	151	49	69	334	340	151	251
36	188	43	324	343	104	302	168	92	97	43	321	268
37	133	103	334	241	207	12	208	168	16	153	121	268
38	351	360	102	5	85	226	246	16	155	323	113	260
39	80	354	191	267	197	169	172	348	271	231	40	164
40	299	311	165	41	211	64	281	91	95	186	283	299

Select a row from 1 to 40 and a column from 1 to 12. Use the number at the intersection of the selected row and column.

Прилог 3. Фишеров критериум

Табела за F-статистика, P = 0,05

df2\df1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.76	8.74	8.73	8.71	8.70	8.69	8.68	8.67	8.67	8.66
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.94	5.91	5.89	5.87	5.86	5.84	5.83	5.82	5.81	5.80
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.70	4.68	4.66	4.64	4.62	4.60	4.59	4.58	4.57	4.56
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.03	4.00	3.98	3.96	3.94	3.92	3.91	3.90	3.88	3.87
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.60	3.57	3.55	3.53	3.51	3.49	3.48	3.47	3.46	3.44
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.31	3.28	3.26	3.24	3.22	3.20	3.19	3.17	3.16	3.15
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.10	3.07	3.05	3.03	3.01	2.99	2.97	2.96	2.95	2.94
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.94	2.91	2.89	2.86	2.85	2.83	2.81	2.80	2.79	2.77
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.82	2.79	2.76	2.74	2.72	2.70	2.69	2.67	2.66	2.65
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.72	2.69	2.66	2.64	2.62	2.60	2.58	2.57	2.56	2.54
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.63	2.60	2.58	2.55	2.53	2.51	2.50	2.48	2.47	2.46
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.57	2.53	2.51	2.48	2.46	2.44	2.43	2.41	2.40	2.39
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.51	2.48	2.45	2.42	2.40	2.38	2.37	2.35	2.34	2.33
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.46	2.42	2.40	2.37	2.35	2.33	2.32	2.30	2.29	2.28
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.41	2.38	2.35	2.33	2.31	2.29	2.27	2.26	2.24	2.23
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.37	2.34	2.31	2.29	2.27	2.25	2.23	2.22	2.20	2.19
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.34	2.31	2.28	2.26	2.23	2.21	2.20	2.18	2.17	2.16
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.31	2.28	2.25	2.23	2.20	2.18	2.17	2.15	2.14	2.12
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.26	2.23	2.20	2.17	2.15	2.13	2.11	2.10	2.08	2.07
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.22	2.18	2.15	2.13	2.11	2.09	2.07	2.05	2.04	2.03
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.18	2.15	2.12	2.09	2.07	2.05	2.03	2.02	2.00	1.99
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.15	2.12	2.09	2.06	2.04	2.02	2.00	1.99	1.97	1.96
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.13	2.09	2.06	2.04	2.01	1.99	1.98	1.96	1.95	1.93

Прилог 4. Студентов критериум

Студентов критериум	
P=0,95	
f	**t**
1	12,71
2	4,30
3	3,18
4	2,78
5	2,57
6	2,45
7	2,36
8	2,31
9	2,26
10	2,23
11	2,20
12	2,18
13	2,16
14	2,14
15	2,13
16	2,12
17	2,11
18	2,10
19	2,09
20	2,09

Прилог 5. Извештај од експериментот

Извештај од експериментот			Бр: 01	5/20/2014
n	**Yj1**	**Yj2**	**Yav**	**Sjsq**
1	27	28	27.5	0.5
2	15.9	17.1	16.5	0.72
3	22.1	22.9	22.5	0.32
4	13.4	13.6	13.5	0.02

Карактеристики	***Резултати***
Sjmax	0.72
Sjsum	1.56
Gp	0.4615
Gt	0.9065
Gp<=Gt	OK
Sysq	0.39
Syavsq	0.195
Sb	0.3122
Sb*t	0.8668
b0	20
b1	-5
b2	-2
b12	0.5

Регресиона равенка со кодирани променливи		
f(x) = 20-5X1-2X2		
Fp	2.5641	
Ft	7.7086	
	X01	**X02**
Осно. ниво	350	6
Интервал	50	2
Долно ниво	300	4
Горно ниво	400	8
do	61	
d1	-0.1	
d2	-1	
d12	0	
Регресиона равенка со натурални променливи		
y(x) = 61-0.1X1-1X2		
y(x) / X1 / X2		
1. 24/320/5 2. 20/340/7 3. 14.5/390/7.5 4. 15/390/7 5. 20.6/355/4.9		

Прилог 6. Текстуален извештај од експериментот

```
IZVESTAJ OD POLN FAKTOREN EKSPERIMENT
Testot e izvrsen na 30.03.2014 16:44:03
```

n	Yj1	Yj2	Yav	Sjsq
1	27	28	27.5	0.5
2	15.9	17.1	16.5	0.72
3	22.1	22.9	22.5	0.32
4	13.4	13.6	13.5	0.02

```
Sjmax=0.72
Sjsum=1.56
Gp=0.4615
Gt=0.9065
Gp<=Gt=OK
Sysq=0.39
Syavsq=0.195
Sb=0.3122
Sb*t=0.8668


Regresioni Koeficienti(Kodirani Promenlivi) :


b0=20
b1=-5
b2=-2
b3=0.5


Regresiona Ravenka(Kodirani Promenlivi)
f(x) = 20-5X1-2X2


Fiserov Koeficient
Fp=2.5641
Ft=7.7086


Naturalni Promenlivi
X01
Osnovno nivo=350
Interval=50
Lower Level=300
Upper Level=400


X02
Basic Level=6
```

```
Interval=2
Lower Level2=4
Upper Level=8


Coefficients of Regression (Natural Variables):


d0=61
d1=-0.1
d2=-1
d3=0


Regresiona Ravenka (Naturalni Promenlivi)
y(x) = 61-0.1X1-1X2


Vrednosti na odzivnata funkcija, y(x).
   y(x)/x1/x2
1. 24/320/5
2. 20/340/7
3. 14.5/390/7.5
4. 15/390/7
5. 20.6/355/4.9
```