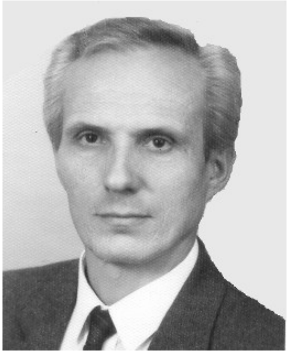
Gligorije Perović

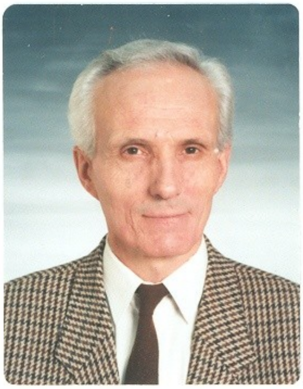
Analiza Modela i Odluka

(AMO)



Srpska geodetska trilogija

3. Teorija grešaka merenja
4. Analiza modela i odluka
5. Geodetske mreže – teorija i primene



UNIVERZITET U BEOGRADU
Građevinski fakultet, Beograd

GLIGORIJE PEROVIĆ

ANALIZA MODELA I ODLUKA

(MONOGRAFIJA 4)

Sa 40 tablica i 58 slika

AGM knjiga
Beograd 2017

ANALIZA MODELA I ODLUKA

Autor: Gligorije Perović

Recenzenti: Jovan Mališić, Matematički fakultet – Univerzitet u Beogradu
Marko Pejić, Građevinski fakultet – Univerzitet u Beogradu
Milan Kilibarda, Građevinski fakultet – Univerzitet u Beogradu

Izdavači: AGM KNJIGA DOO: www.agmknjiga.co.rs;
Građevinski fakultet: www.grf.bg.ac.rs

Glavni i odgovorni urednik:
Slavica Sarić Ahmić

Priprema: Gligorije Perović

Štampa: Donat graf
Tiraž: 250

Rešenjem br. 22/93-2 od 28. 11. 2017. Nastavno-naučno veće Građevinskog fakulteta Univerziteta u Beogradu odobrilo je štampanje ove monografije.

ISBN: 978-86-86363-85-5 (AGMK)

CIP - Каталогизација у публикацији - Народна библиотека Србије, Београд

528.06:519.2
528.06:519.87

ПЕРОВИЋ, Глигорије, 1940-
Analiza modela i odluka : (monografija 4) : sa 40 tablica i 58 slika / Gligorije Perović. - Beograd : AGM knjiga : Građevinski fakultet, 2017 (Beograd : Donat graf). - XII, 247 str. : ilustr. ; 24 cm

Tiraž 250. - Srpsko - rusko - engleski rečnik: str. 237-247. - Bibliografija: str. 200-206. - Registri.

ISBN 978-86-86363-85-5 (AGMK)

a) Геодезија - Статистички методи b) Геодезија - Математички модели
COBISS.SR-ID 247846924

Svi modeli su pogrešni,
ali neki su korisni i ilustrativni.

G. E. P. Box, 1979.

Matematički modeli su pogodno sredstvo
za opisivanje realnih sistema (*pojava, procesa, objekata*).

G. Perović, 1997.

Nikakvo naučno istraživanje nije konačno, ono samo po sebi predstavlja
najverovatniji zaključak, koji autor može izvući iz njemu raspoloživih podataka.
Potpuniji skup podataka ili savremenija analiza, eksperiment
i opažanje dovode do novih formula i teorija.
*U tome je **suština razvoja nauke**.*

(*K. Pirson, 1899*)

Teorija vodi
merenje odlučuje

(*G. Perović, 1978*)

OSNOVNI PRINCIPI ANALIZE MODELA I ODLUKA

1. *Istraživanje tipa procesa* (*objekta, sistema, pojave*)
2. *Uključenje svih mogućih regresora*
3. *Optimizacija eksperimenta*
4. *Definisanje i izbor matematičkog modela*
5. *Ocena čiste greške*
6. *Ocena greške modela*
7. *Istraživanje optimalnog sastava regresora*
8. *Prognoza, planiranje, dijagnostika, ...*

(*G. Perović, 1986*)

PREDGOVOR

Pre svega ovde se ističe da je za razumevanje i praćenje materije monografije Analiza modelâ i odlukâ *neophodno i dovoljno znanje teorije verovatnoće i matematičke statistike iz standardnih kurseva*.

Kako je nastala Analiza modelâ i odlukâ, i, uopšte, kako je nastala Srpska geodetska trilogija.

U svom dugogodišnjem nastavnom radu, uvodeći stalno nova znanja u nastavne predmete, Teoriju grešaka i Račun izravnanja, ja sam 1985. godine u nastavu na Geodetskom odseku Građevinskog fakulteta Univerziteta u Beogradu uveo nastavni predmet Teorija modelâ i odlukâ, koji sam već sledeće 1986. godine preimenovao u *Analizu modelâ i odlukâ*. U ovome sam ogromnu pomoć dobio od velikana Srpske geodezije profesora doktora Vladete Milovanovića diplomiranog inženjera geodezije.

U svetu je odavno postojala želja da se pojave – procesi – sistemi – opisuju matematičkim i/ili matematičko-statističkim modelima, nadalje *modelima*, kao pogodnim sredstvom za analizu. Tih 80-ih godina prethodnog milenijuma razvoj računara i mernih tehnologija bio je intenzivan. Tako su se merenja, računanja i analize modelâ mogle mnogo lakše, brže i kvalitetnije izvoditi nego do tada. Tih godina su se iz oblasti spomenutih modelâ pojavile fundamentalne knjige Ajvazjana (Айвазян С. А., Енюков И. С., Мешалкин Л. Д.: Прикладная статистика: *Основы моделирования и первичная обработка данных*. Финансы и статистика, Москва, 1983, и Айвазян С. А., Енюков И. С., Мешалкин Л. Д.: Прикладная статистика: ***Исследование зависимостей***, Справочное издание, под редакцией С. А. Айвазяна. Финансы и статистика, Москва, 1985; (engl.: S. A. Aivazyan, I. S. Yenykov, L. D. Meshalkin: APPLIED STATISTICS – STUDY OF RELATION-SHIPS, Reference edition. Edited by prof. S. A. Aivazyan. Finansy i statistika, Moscow, 1985), koje su značajno doprinele naučnom utemeljenju i razvoju modelâ i odlukâ.

Utemeljenje nekih fundamenata ove analize nalazi se u mojoj knjizi Metod najmanjih kvadrata i Least Squares (Perović 2005a, 2005b, glave: 4, 5, 23, 24, 25, 26, 28, 29, 30, 31, 33, 34, 35, 36, 37 i 38).

„***Zlatno doba***" *statističke obrade podataka* počinje sa pojavom elektronskih uređaja za merenje i računanje od kada se ona primenjuje u svim oblastima ljudskog delovanja. U geodeziji, *nauci o merenjima*, ona je ne samo nezaobilazna već je i fundamentalna, jer se skoro isključivo njenim poznavanjem *razdvaja profil inženjera od profila tehničara*.

PRIZNANJA

Recenzent dr ***Jovan Mališić***, diplomirani matematičar, profesor Matematičkog fakulteta u Beogradu, detaljno je pregledao tekst i pružio ogromnu pomoć u tačnom definisanju pojmova i fundamentalnih postavki u rešavanju konkretnih zadataka, što je suštinski poboljšalo izloženu materiju. Dr Mileva Samardžić Petrović, diplomirani inženjer geodezije, docent na Katedri za geodeziju i geoinformatiku Građevinskog fakulteta u Beogradu, pomogla je oko definisanja vrsta promenljivih i vrsta modela. Dr Vesna Jevremović, diplomirani matematičar, profesor Matematičkog fakulteta u Beogradu, pomogla je oko srpsko – rusko – engleskog rečnika.

Veliku finansijsku i moralnu podršku za izdavanje i štampanje ove monografije dali su:

- Inženjer geodezije ***Marko Božić***, direktor firme MEIXNER iz Zagreba, i
- ***Ivica Walz***, vlasnik firme ALHIDADA DOO iz Petrijevaca.

DARODAVCI

Štampanje ove monografije pomogli su:

1. MEIXNER, Zagreb, **glavni darodavac**;
2. ALHIDADA DOO, Petrijevci, **glavni darodavac**;
3. NE-POK, Ivanić Grad;
4. GEODETSKA RADNJA MARINKOVIĆ, Veliko Gradište;
5. GEOINŽENJERING, Novi Sad;
6. GEOZETA doo, Podgorica;
7. GEOS doo, Zenica;
8. ZONING, Herceg Novi;

Analiza modelâ i odlukâ je deo matematičke statistike koji u suštini predstavlja ***dijagnostiku sistêma*** (procesâ, pojavâ) zasnovanu na merenjima. Ukoliko su sve relevantne informacije, u vezi posmatranog sistema i samog eksperimenta, prikupljene onda će ta dijagnostika biti ispravna i potpuna.

Naučni razlozi koji su me ka ovom cilju stalno vodili jesu Gausov princip da „*Nauka treba da bude prijateljica prakse, ali ne i njena robinja, ona treba da je daruje, a ne da joj služi*”, koji stoji i u mome opredeljenju, i fundamentalni značaj statističke obrade podataka; a osim toga i shvatanje da je *neznanje najveći čovekov neprijatelj*.

U monografiji Analiza modelâ i odlukâ ustanovljene su ***mere i kriterijumi kvaliteta modêla***, kao najvažniji njen deo.

Iz ove Analize danas se svojim razvojem izdvaja oblast sa nazivom ***Geostatistika***.

Srpsku geodetsku trilogiju čine monografije: Teorija grešaka merenja (monografija 3), Analiza modelâ i odlukâ (monografija 4) i Geodetske mreže – teorija i primene (monografija 5). U njima su dati neophodni fundamenti za analizu merenja, a teorije i metode su propraćene primerima iz geodetske prakse. Osim toga dao sam i *nova rešenja za neke geodetske probleme* koji do sada nisu bili rešeni.

Opšta je tendencija u svetu stvaranja *stručnjaka SOFTWARE-aša*, gotovih „preciznih metoda merenja” i „automatskih softverskih analiza” – za šta su potrebni niži nivoi znanja od inženjerskih pa ih takvi stručnjaci rado koriste što samo po sebi dovodi do pogrešnih analiza i odluka, uz istovremeno stvaranje sve manjeg broja ***inženjera mislilaca***[1)], tj. takvih stručnjaka koji znaju ***zašto – kako – koliko***; *zašto* se nešto radi, *kako* se to (naučno i stručno) radi i *koliko* to istovremeno *najmanje košta*.

Stoga je *ova trilogija namenjena očuvanju inženjera mislilaca*.

Prema tome, monografija Analiza modelâ i odlukâ namenjena je svima koji se bave obradom i analizom merenja – *inženjerima*, *studentima*, *naučnicima*.

Osim toga dao sam i SRPSKO–RUSKO–ENGLESKI rečnik korišćenih pojmova, kao važnu materiju savremenim istraživačima.

U Beogradu

24. avgusta 2017.

Gligorije Perović

Univerzitet u Beogradu
Građevinski fakultet, Beograd

[1)] Inače, tvorci novog načina učenja datog u „Bolonjskoj deklaraciji” su osim sniženja nivoa znanja koja na fakultetima treba osvojiti snizili i sama zvanja; tako su inženjera preimenovali u bachelor-a, a diplomiranog inženjera – izraz koji se koristi već 2 000 godina – u master-a.

IZVOD IZ RECENZIJE

(a) Sadržaj rukopisa

Predloženi rukopis „Analiza modelâ i odlukâ” predstavlja nastavak serije monografija u oblasti geodetskih nauka autora Gligorija Perovića, redovnog profesora Građevinskog fakulteta u Beogradu, sada u penziji. Radi se o monografiji 4. (prethodne četiri su: 1 Metod najmanjih kvadrata, Beograd, 2005; 2 Least Squares, Beograd, 2005; 3 Teorija grešaka merenja, Beograd, 2015; Precizna geodetska merenja, 2017).

Osnovne teme ovog rukopisa su formiranje, *analiza*, *diskusija adekvatnosti* i *mogućnosti poboljšanja izabranih matematičkih modela*, posebno stohastičkih modela, kao i *donošenje odgovarajućih statističkih* (i praktičnih) *odluka*, uz primene, pre svega, u geodetskim disciplinama.

Rukopis je podeljen na 16 glava, koje čine obim od 199 stranica kucanog teksta, sa većim brojem tabela, slika i primera primene tretirane teorije u geodetskim merenjima. Na kraju rukopisa se daju literatura, tablice moći testova, izvod iz singularnih izravnanja, registar autora pominjanih u tekstu i registar pojmova. …

… autor polazi od premise da je planiranje eksperimenata jedna od najvažnijih faza u istraživanju zavisnosti. Pri tome, u planiranju eksperimenata imamo dve grupe oprečnih zahteva: „*sto više i bolje*” i „*sto manje i jeftinije*”. Kompromis za koji se istraživač odlučuje dovodi do dobrog opredeljenja za određene klase eksperimenata (u datim uslovima) …

Glava 15 odnosi se na izučavanje topografsko-katastarskih planova, sa naglaskom na izučavanje transformacija grafičkih nosilaca informacija. …

(b) Analiza rukopisa

Ceo tekst monografije napisan je na vrlo visokom stručnom i naučnom nivou. Za njegovo detaljnije čitanje korisnik mora biti naoružan pouzdanim znanjima iz raznih oblasti geodezije i matematičke statistike (teorija matrica, metoda najmanjih kvadrata; teorija statističkih raspodela, testiranje statističkih hipoteza, regresiona analiza, kovarijaciona analiza,...). Da bi se tekst više približio njegovom čitaocu autor monografije na više mesta daje praktična uputstva za istraživanja i donošenja odgovarajućih odluka, bez velikog korišćenja komplikovanih matematičko-statističkih formula. Osim toga, poslednje glave monografije odnose se na primene tretiranje teorije u konkretnim geodezijskim zadacima. Tu će čitalac naći dobre ideje za proširivanja i preciziranja svojih znanja, a sve radi adekvatnije primene u praksi.

Ova monografija (kao i prethodne tri) je rezultat višegodišnjeg stručnog i naučnog, a takođe i pedagoškog rada njenog autora. Tu je autor pokazao da suvereno vlada velikim brojem naučnih disciplina koje se odnose na izbore stohastičkih modela i donošenje odgovarajućih odluka. Komplikovanu matematičku stranu u samom rukopisu autor nastoji da svojim odmerenim pristupom i komentarima što više približi čitaocu. Ipak, čitalac-istraživač mora biti naoružan strpljenjem, upornošću i trudom za razumevanje teksta, a takođe i prethodnim znanjima iz geodezije i matematičke statistike.

U monografiji se *na velikom broju mesta nalaze i autorovi teorijski i praktični rezultati*. Na tim mestima se daju razne nove metode autora Gligorija Perovića u pitanjima izbora statističkih modela, što dovodi do popravke, poboljšanja i preciziranja u dosadašnjim izborima tih statističkih modela. *To predstavlja esencijalno dobru stranu ove monografije*.

Sugestije i predloge koje smo kao recenzenti davali na ponuđeni tekst autor je savesno i detaljno razmatrao i uneo u konačan tekst rukopisa,

(c) Zaključak i predlog

Ukupno uzevši, predloženi tekst monografije predstavlja jedno ***fundamentalno***, ***kapitalno delo***. Nijedna monografija takvog tipa, formata, obima i sadržine (koliko je nama kao recenzentima poznato) kod nas do sada nije napisana. Ova monografija značajno popunjava tu prazninu. Tu će svoja interesovanja naći kako odlični studenti (pre svega, studenti geodezije), inženjeri, specijalisti u toj i srodnim naučnim i stručnim oblastima, tako i naučni radnici raznih profila koji rade na svom daljem naučnom usavršavanju.

Iz navedenih razloga, sa izuzetnim zadovoljstvom predlažemo da se rukopis monografije „Analiza modelâ i odlukâ” autora Gligorija Perovića prihvati za štampu i objavi na srpskom jeziku. Bilo bi dobro razmotriti i mogućnosti njenog prevođenja na neki od svetskih jezika.

10. oktobar 2017. godine
Beograd

Recenzenti rukopisa

Dr Jovan Mališić,
redovni profesor Matematičkog fakulteta,

Dr Marko Pejić,
docent Građevinskog fakulteta,

Dr Milan Kilibarda,
docent Građevinskog fakulteta.

S A D R Ž A J

G l a v a 1

O STATISTIČKOM ISTRAŽIVANJU ZAVISNOSTI

1.1. Pojave i promenljive veličine i njihova unutrašnja povezanost

Stalna je čovekova težnja da se za bilo koji zakon prirode ili društvenog razvitka može opisati karakter ili struktura međusobnih veza koje postoje među izučavanim *pojavama* ili pokazateljima – *promenljivim veličinama*.

Međusobna povezanost pojava je jedno od temeljnih načela naučnog istraživanja. Nauka, između ostalog, ima zadatak da što bolje utvrdi takvu povezanost pojava u datom području istraživanja.

Tako je ***prvi cilj*** naučnog istraživanja *da se na osnovu jedne pojave predvide promene i dešavanja u drugoj pojavi ili drugim pojavama koje su sa prvom povezane*.

Drugi cilj svakog naučnog istraživanja je *razumevanje pojava i događaja na nekom području*.

Povezanost pojava ima bar dvojaki karakter. Može biti *uzročno-posledična* ili *korelativna*. Ako je jedna pojava ili događaj uzrok nastanka neke druge pojave, događaja ili promene onda je to ***uzročno-posledična povezanost*** pojava. Druga pojava je, dakle, posledica prve; obratno ne važi.

Stoga je jedan od zadataka naučnog istraživanja utvrđivanje uzročno-posledične povezanosti dveju ili više pojava u određenom području. Pri tome može biti situacija da postoji istovremeno delovanje više uzroka i da se javlja više posledica.

Pri *postojanju* ***korelativne povezanosti pojava*** promene u jednoj i u drugoj pojavi mogu se javljati paralelno tako da jedne nisu uzrok drugima.

Na osnovu korelativnih promena ne možemo suditi o uzročno-posledičnim odnosima, ali nam proučavanje korelativnih promena omogućava bolje razumevanje događaja koje istražujemo i bolje predviđanje. Na osnovu merenja određenih promena možemo predvideti promene koje će se dogoditi u nekoj drugoj pojavi, koja je u korelativnoj vezi sa posmatranom. Stoga proučavanje korelativnih promena i povezanosti ima jednaku važnost kao i proučavanje uzročno-posledičnih odnosa.

Podaci dobijeni merenjem promena na pojavama su predmet analiziranja i proučavanja i nazivaju se ***promenljivim veličinama***.

Promenljive, prema tome, odražavaju promene u pojavama koje se proučavaju, a njihova međusobna povezanost meri se *koeficijentima korelacije* ili ***koeficijentima asocijacije*** između promenljivih.

Cilj je, dakle, kako na osnovu rezultata statističkih opažanja na događajima ili pokazateljima otkriti i opisati njihove međusobne veze. Naime, to je ***problem statističkog istraživanja zavisnosti*** i on se pojavljuje kao glavni pri rešavanju tipičnih zadataka prakse takvih kao što su:

- *normiranje,*
- *prognoza,*
- *planiranje,*
- *dijagnostika,*
- *ocena teško dostupnih za opažanje i merenje karakteristika analiziranog sistema,*
- *ocena efektivnosti funkcionisanaja ili kvaliteta objekta,*
- *regulisanje parametara procesa ili sistema* i sl.

Pri naučnom proučavanju pojava teži se obuhvatiti što veći broj pojava, pogotovo kada su one višestruko povezane. Takav pristup dovodi do većeg broja promenljivih, a posledično i do većeg broja koeficijenata korelacije među promenljivim, što rezultira podacima *koji su sami po sebi materijal za dalja proučavanja, odnosno podacima koji se ne mogu prihvatiti kao konačni*.

Zbog toga se u nauci nastoji *objasniti što veći broj pojava ili događaja na osnovu što manjeg broja promenljivih*.

1.2. Vrste promenljivih i vrste modela

U glavama 4 i 5 monografije METOD NAJMANJIH KVADRATA (Perović 2005a i 2005b) detaljno su prikazane vrste promenljivih i vrste modela. Stoga će se ta materija ovde izneti kratko i dati definicija matematičkog modela. Analiza modelâ i odlukâ se intenzivno tazvija od 80-tih godina prošlog veka, čemu su znatno doprinele i fundamentalne knjige Ajvazjana (1983 i 1985). Do danas je ona prodrla u skoro sve pore društvenog razvitka (Heuvelink 2007), a naročito se razvija oblast *geostatistike*, iz koje na Katedri za geodeziju i geoinformatiku Građevinskog fakulteta u Beogradu imamo disertacije (Kilibarda 2013, Samardžić – Petrović 2014 i Pejović 2016).

Karakter ili strukturu međusobnih odnosa među ***pojavama*** – *sistemima* – *procesima* – *objektima* – *obeležjima*, (*promenljivim veličinama* ili kratko ***promenljivim***), možemo opisati ***matematičkim modelom*** – videti glavu 3, a ovde ćemo reći da je to izraz gde se *matematičkim jezikom opisuje ponašanje i/ili osobina posmatranog fizičkog sistema*. Tako, na primer, zakon Njutna u mehanici i sva, na njemu zasnovana, klasična mehanika predstavlja skup modela mehaničkih pojava. Maksvelova jednačina u fizici predstavlja matematički model elektrodinamičkih pojava itd.

U mnogim slučajevima matematički model se može formirati čisto matematičkim putem na osnovu poznatih zakona mehanike, fizike i drugih disciplina koje se koriste količinskim odnosima. Na primer, rastojanje između dveju tačaka može se izraziti matematički, koristeći zakone geometrije; *geometrijska figura može se jednoznačno opisati* (*definisati*) *neophodnim i dovoljnim brojem njenih nezavisnih elemenata* i sl.

Međutim, postoje i takve pojave za koje je principijelno nemoguće konstruisati model čisto matematičkim putem. Kao primer može poslužiti čovek ili kolektiv ljudi koji vrše određene funkcije, zavod, grana privrede, ekonomika i slično. U geodeziji, na primer, pri merenju koordinata GPS tehnologijom *nije moguće čisto matematičkim putem konstruisati model kojim bi se uzimali u obzir uticaji kvazistacionarnih blokova refrakcije*. Ali, savremeni naučno-tehnički progres zahteva stvaranje matematičkih modela i za takve sisteme.

Ako odnosi među promenljivim omogućuju da se prema podacima jednih veličina jednoznačno odrede vrednosti drugih onda se model koji ih opisuje naziva ***determinestičkim.*** A model koji prepoznaje slučajnu prirodu ulaznih veličina je ***stohastički***. U njemu je ulaz po prirodi slučajan – isto kao i ulazi koji su slučajne veličine.

Da bi se mogli formulisati modeli zavisnosti, usvojeno je da se opis funkcionisanja izučavanog realnog fizičkog procesa opisuje skupom sledećih promenljivih:

$\mathbf{x}$ – vektorom „***ulaznih***” ***promenljivih***, koje opisuju *uslove funkcionisanja* ispitivanog sistema (deo njih, po pravilu, dâ se regulisati ili se njima može delimično upravljati); u odgovarajućim matematičkim modelima nazivaju se ***nezavisnim, faktorima*** - *argumentima, egzogenim, prediktornim*, ili prosto **prediktorima,** tj. predskazivačima, ili *objašnjavajućim promenljivim*;

$\mathbf{y}$ – vektorom „***izlaznih***” ***promenljivih***, koje karakterišu ponašanje ili efektivnost funkcionisanja ispitivanog sistema; u matematičkim modelima nazivaju se ***zavisnim****, odzivnim, endogenim, rezultujućim* ili *opisivim* **promenljivim** (noviji izraz: *kovarijetne* promenljive; engl.: covariate), a najčešće su to ***rezultati merenja***; i

$\boldsymbol{\varepsilon}$ – vektorom ***latentnih*** (tj. skrivenih, koje se ne mogu neposredno meriti) *slučajnih rezidualnih komponenti,* koje odražavaju uticaj na odgovarajuće $\mathbf{y}$, ne uzetih u obzir „ulaznih” faktora, a takođe i ***slučajnih grešaka merenja*** (sl. 1.2.1). To određuje *stohastički* (*statistički*) model.

Sl. 1.1.1. *Opšta šema međusobnog dejstva promenjivih pri statističkom istraživanju zavisnosti*

.

Napomena 1.2.1. *Promenljive su ovako klasifikovane sa tačke gledišta njihovog korišćenja u regresionim modelima. Inače, detaljna klasifikacija promenljivih sa* *raznih aspekata, može se naći u Australian Bureau of Statistics*: http://www.abs.gov.au/websitedbs/a3121120.nsf/home/statistical+language+-+what+are+variables. ▲

Merenja (podaci) se mogu urediti na razne načine. Ali, najčešće se koristi sledeći poredak prikazivanja merenja (promenljivih).

		Pojedino	**merenje**					|	**Očekivanje**
Promenljiva *j*:	$\mathbf{y}_j$:	**1.**	**2.**	$\cdots$	***i***	$\cdots$	$\boldsymbol{n}_j$	|	$\mathsf{M}[\mathbf{y}_j]$
Promenljiva 1:	$\mathbf{y}_1$:	y_{11}	y_{12}	$\cdots$	y_{1i}	$\cdots$	y_{1n_1}	|	η_1
Promenljiva 2:	$\mathbf{y}_2$:	y_{21}	y_{22}	$\cdots$	y_{2i}	$\cdots$	y_{2n_2}	|	η_2
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	|	$\cdots$
Promenljiva *j*:	$\mathbf{y}_j$:	y_{j1}	y_{j2}	$\cdots$	$\boxed{y_{ji}}$	$\cdots$	y_{jn_j}	|	η_j
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	|	$\cdots$
Promenljiva *k*:	$\mathbf{y}_k$:	y_{k1}	y_{k2}	$\cdots$	y_{ki}	$\cdots$	y_{kn_k}	|	η_k

Matematička očekivanja su: $\mathsf{M}[y_{1i}] = \eta_1$, $\cdots$, $\mathsf{M}[y_{ji}] = \eta_j$, $\cdots$, $\mathsf{M}[y_{ki}] = \eta_k$, $i = 1, 2, \cdots, n_j$.

1.3. Zadaci i ciljevi statističkog istraživanja zavisnosti

Opšti zadatak statističkog istraživanja zavisnosti može se formulisati na sledeći način:

Na osnovu rezultata merenja veličina X *i* Y *ispitivanih promenjivih na objektima* (*sistemima, procesima, pojavama*) *analiziranog skupa treba konstruisati takvu funkciju* (*vektorsku*):

$$\mathbf{y} = \mathbf{f}(\mathbf{x}), \tag{1.3.1}$$

gde je $\mathbf{Y} = \mathbf{y} + \boldsymbol{\varepsilon}$ – vektor rezultata merenja, $\mathbf{x}$ vektor konkretnih vrednosti, *pomoću koje bi se na najbolji* (*u određenom smislu*) *način određivala vrednost rezultujućih* (*prognostičkih*) *promenljivih* $\mathbf{y}$ *po zadatim vrednostima objašnjavajućih* (*prediktornih*) *promenljivih* $\mathbf{x}$, pri čemu je $\boldsymbol{\varepsilon}$ normalno raspoređeno sa $\boldsymbol{\varepsilon} \sim N[\mathbf{0}, \mathbf{K}_{\boldsymbol{\varepsilon}}]$.

Ali, ova formulacija je dosta uopštena i zahteva odgovore na detalje, a posebno na:

- **a)** kakva je ***struktura modela*** tražene zavisnosti između $\mathbf{y}$ i $\mathbf{x}$, zapisana pomoću $\mathbf{y}, \mathbf{x}, \mathbf{f}(\mathbf{x})$ i $\boldsymbol{\varepsilon}$,
- **b)** ***izbor kriterijuma aproksimacije*** $\mathbf{y}$ pomoću $\mathbf{f}(\mathbf{x})$, i
- **c)** *radi kog **praktičnog cilja*** se sve ovo sprovodi.

Razjašnjenje ovih i sličnih pitanja dobrim delom će se naći u primerima 4.1.1, 4.1.2 i 4.2.6 u MNK (Perović, 2005a i 2005b).

Pri istraživanju zavisnosti među promenljivim pojavljuju se sledeći **osnovni ciljevi** (Ajvazjan i dr. 1985, Perović, 2005a i 2005b):

1°. polazeći od konkretnih praktičnih ciljeva, *kako **odrediti smisao** u kome se shvata istraživana zavisnost*;

2°. *da li postoji **bilo kakva veza*** među istraživanim promenljivim, kakva je struktura tih veza (u slučaju više promenljivih), *i kako izmeriti stepen te veze*;

3°. *kakva je **opšta struktura** odgovarajućeg matematičkog modela* (opšti matematički oblik tražene veze između $\boldsymbol{\eta} = \mathsf{M}[\mathbf{Y}]$ i $\mathbf{x}$);

4°. polazeći od usvojene opšte strukture modela, ***kako izvesti** neophodna **računanja*** (ovde se neretko javljaju slabo uslovljeni sistemi linearnih jednačina gde je *kondicion* matrice normalnih jednačina reda 10^{10} (videti primer 4.2.2 poglavlja 4.2), pa čak i viši) *radi dobijanja konkretnog oblika zavisnosti* $\boldsymbol{\eta}$ od $\mathbf{x}$;

5°. *kako **oceniti stepen tačnosti** naših zaključaka*;

6°. *kako rešiti sve probleme u situacijama kada među prediktornim promenljivim **mogu biti i nekoličinske***, i

7°. na kraju, *ako se skup polaznih statističkih informacija dobija u uslovima aktivnog eksperimenta, kako **odrediti vrednosti prediktornih promenljivih** i kako **rasporediti zadati** (planirani) **broj opažanja**, da bi se dobio najveći stepen tačnosti naših statističkih zaključaka.*

Jedno od bitnih pitanja je: *kakav je **konačni praktični cilj** statističkog istraživanja zavisnosti*? S tim pitanjem treba da počne svako statističko istraživanje zavisnosti jer od odgovora na njega suštinski zavisi: *plan istraživanja*, *izbor opšte strukture matematičkog modela*, *interpretacija dobijenih statističkih karakteristika i zaključaka* itd. Konačni praktični ciljevi statističkog istraživanja zavisnosti u osnovi mogu biti triju tipova (Ajvazjan i dr. 1985, Perović, 2005a, 2005b):

1) ***Ustanovljavanje*** same činjenice ***postojanja*** (*ili odsustva*) ***statistički značajne veze*** između Y i X, i *istraživanje strukture te veze*;

2) ***Prognoza*** (***predikcija***, *uspostavljanje*) nepoznatih vrednosti, pojedinih - $Y(x) = (Y \mid X = x)$ ili očekivanih - $(\eta(x) = \mathsf{M}(Y \mid X = x))$ vrednosti istraživanog rezultujućeg pokazatelja po zadatim vrednostima prediktornih promenljivih X, i

3) ***Otkrivanje uzročnih veza*** između opisnih promenljivih X i odzivnih pokazatelja Y i delimično upravljanje vrednostima Y putem regulisanja veličina opisnih promenljivih X, (takva postavka zadatka pretenduje da pronikne u sami mehanizam transformisanja „ulaznih” promenljivih X i ε u rezultujuće pokazatelje Y, vidi sl. 1.2.1).

Pri izučavanju veze između analiziranih količinskih pokazatelja treba ustanoviti kom ***tipu zavisnosti*** *ona pripada*, u kom slučaju se podrazumeva da na izbor analitičkog vida funkcije $\eta = f(x,\theta)$ (θ – konstantni parametar) bitno utiče ***priroda analiziranih promenljivih*** (X,y).

O ovim zadacima i ciljevima statističkog istraživanja zavisnosti detaljnija objašnjenja mogu se naći kod Ajvazjana i dr. (1985), Ermakova i dr. 1983 i kod Perovića (2005a i 2005b), a ovde će konkretni zadaci i ciljevi biti detaljnije objašnjeni kod odgovarajućih oblasti; npr. cilj **4º** – „*kako izvesti neophodna računanja*” objašnjen je u primeru 4.2.2, poglavlja 4.2.

1.4. Strukturni prikaz statističkog istraživanja zavisnosti

Ceo proces statističkog istraživanja zavisnosti pogodno je razložiti na *osnovne etape*. Te etape su u daljem opisane saglasno hronologiji njihove realizaije, međutim neke od njih, koje su u hronološkom redu, mogu biti međusobno povezane u iteracionom procesu, jer rezultati realizacije kasnijih etapa mogu sadržati zaključak o neophodnosti ponovnog „prolaza” (sa uzimanjem u obzir novih informacija dobijenih na prethodnim etapama) kroz već pređene etape (v. na primer, šemu uzajamne povezanosti etapa 3, 4, 5 i 6 na sl. 1.4.1). Izložena šema u osnovnom je prilagođena za istraživanje zavisnosti među količinskim promenljivim, međutim s minimalnim (i očiglednim) modifikacijama ona se može primeniti i pri statističkoj analizi veza među nekoličinskim i raznorodnim promenljivim.

Kao osnovne možemo izdvojiti sledeće etape (*Ajvazjan S. A. i dr.*, 1985):

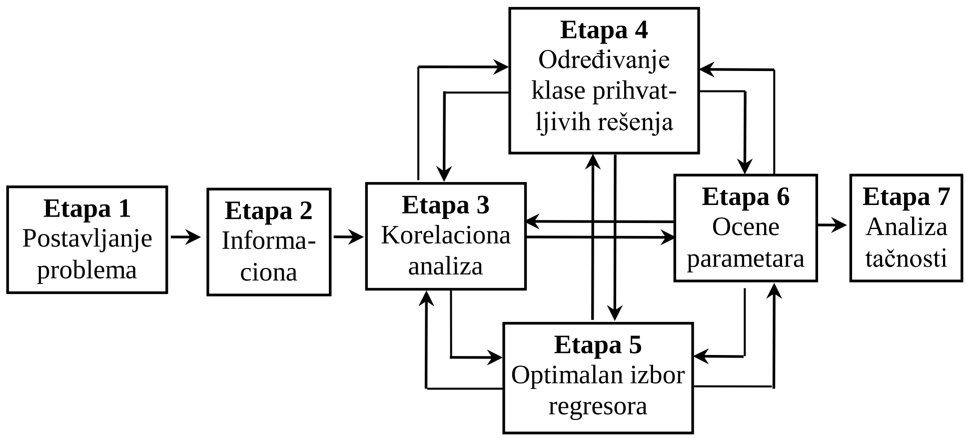
Etapa 1. Postavka problema. Pre svega, istraživač je dužan da odredi:

1) *elementarni objekt*;
2) *skup „ulaznih” i „izlaznih” pokazatelja* ($x^{(1)}, x^{(2)}, \cdots, x^{(p)}; y^{(1)}, y^{(2)}, \ldots, y^{(m)}$), koji se registruju na svakom statistički istraživanom objektu, s jasnim definisanjem načina njihovog merenja;
3) *konačne praktične uslove* istraživanja, tip istraživanih zavisnosti i željenu formu statističkih zaključaka;
4) *skup* elementarnih *objekata istraživanja* na koji se mogu primeniti statističke zavisnosti otkrivene u analizi;
5) *ukupno vreme i troškove rada*.

U rešavanju svih nabrojanih pitanja prve etape istraživanja glavnu ulogu, nesporno, ima istraživač, tj. specijalista dotične oblasti, za koju se planira izvođenje tog istraživanja;

Etapa 2. Informaciona. Sastoji se u prikupljanju neophodnih statističkih informacija (v. poglavlje 1.2). Pri tom su moguće dve principijelno različite situacije:

1) istraživač ima *mogućnost prethodnog planiranja istraživanja*, i
2) istraživač *dobija polazne podatke* takve kakvi su prikupljeni bez njegovog učešća (*uslovi pasivnog eksperimenta*);



Sl. 1.4.1. *Šema hronološki-iteracionih povezanosti osnovnih etapa statističkog istraživanja zavisnosti..*

Etapa 3. Korelaciona analiza. Ova etapa treba da odgovori na pitanja, da li postoji bilo kakva veza među istraživanim promenljivim, kakva je struktura tih veza i kako izmeriti stepen te veze. (O koeficijentu korelacije r slučajnih veličina X i Y - parni koeficijent korelacije dat je detaljan prikaz kod Ajvazjana i dr. 1985, Halda 1957 i Perovića 2017, a koeficijent determinacije R^2 – kod Ajvazjana i dr. 1985, Halda 1957 i ovde u glavi 8).

Etapa 4. Određivanje klase prihvatljivih rešenja. Glavni cilj istraživača u toj etapi je ***određivanje*** *opšteg oblika, tj.* ***strukture tražene veze*** između Y i X, ili, drugim rečima, opis klase funkcija F, u okviru koje će on izvoditi dalje traženje konkretnog oblika zavisnosti koja ga interesuje. Taj opis se najčešće daje u formi neke parametarske familije $f(X,Y)$, pa se stoga ova etapa naziva i ***etapom parametrizacije modela***.

Treba primetiti da, osim što je ključna, u određenom smislu presudna karika u celom procesu statističkog istraživanja zavisnosti, ta etapa je istovremeno i u najnepovoljnijem položaju u poređenju sa drugim etapama (s pozicije (postojanja) prisutnosti strogih i završenih matematičkih preporuka po njenoj realizaciji). Prema tome, za njenu realizaciju je neophodan zajednički rad specijaliste odgovarajuće predmetne oblasti (ekonomike, tehnike, medicine itd.) i matematičara – statističara, usmeren na što dublje prodiranje (shvatanje) u „fizički mehanizam" istraživane veze.

Sledeće dve etape, 5-ta i 6-ta, realizuju se, po svojoj suštini, paralelno;

Etapa 5. Optimalan izbor regresora. U regresionoj analizi može se pojaviti multikolinearnost, tj. uske (bliske) statističke veze među predskazujućim promenljivim $x^{(1)}, x^{(2)}, \ldots, x^{(p)}$, zbog čega se, posebno, javlja slaba uslovljenost njihove korelacione matrice, što zadaje poteškoće i stvara nepogodnosti pri statističkom istraživanju zavisnosti.

Za optimalan izbor regresora treba se koristiti rezultatima glave 8 – Istraživanje optimalnog sastava regresora, pri čemu autor preporučuje metodu PERIREG M, (zbog visoke moći kriterijuma i malog broja istraživanih kombinacija regresora);

Etapa 6. Izračunavanje ocena nepoznatih parametara *koji ulaze u ispitivanu jednačinu statističke veze.*

U ovoj etapi se određuje najbolja aproksimacija $f_{\mathrm{T}}(X)$ funkcije $f_{\mathrm{T}}(X)$, koja se pojavljuje kao *rešenje zadatka optimizacije*;

Etapa 7. Analiza tačnosti dobijenih jednačina veze. Istraživač je dužan dati izveštaj o stepenu saglasnosti dobijene aproksimacije $f_{\mathrm{T}}(X)$ nepoznate teorijske funkcije $f_{\mathrm{T}}(X)$, (pri čemu je $f_{\mathrm{T}}(X)$ nazvana ***empirijskom funkcijom regresije***), sa istinitom zavisnosti $f_{\mathrm{T}}(X)$. Pri tom se greške opisa nepoznate istinite funkcije $f_{\mathrm{T}}(X)$ pomoću $f_{\mathrm{T}}(X)$ u opštem slučaju sastoje od dve komponente: a) ***greške** aproksimacije - **modela***, i b) ***greške merenja***. Veličina prve zavisi od uspeha u realizaciji etape 4, tj. od pravilnosti izbora klase prihvatljivih rešenja *F*. Posebno, ako je klasa *F* izabrana na taj način, da uključuje u sebe i nepoznatu funkciju $f_{\mathrm{T}}(X)$, onda je greška aproksimacije nula. Ali, ipak, i u tom slučaju ostaje slučajna komponenta (izazvana slučajnom greškom merenja).

1.5. Princip optimalnog planiranja u regresionim eksperimentima

Detaljniji prikaz optimalnog planiranja eksperimenata biće dat u glavama 14 i 8, dok će ovde biti reči samo o *principima optimalnog planiranja u regresionim eksperimentima*.

Neka se posmatrani *zadatak planiranja eksperimenta opisuje modelom linearne regresije* (o regresionim i ostalim zavisnostima videti glavu 2):

$$Y_{ij} = f(\mathbf{x}_i)^{\mathrm{T}}\boldsymbol{\theta} + \varepsilon_{ij}, \quad i = 1, \cdots, k, \quad j = 1, \cdots, n_i; \quad \sum_1^k n_i = N \tag{1.5.1}$$

(u oznakama (2.2.4) i (2.2.9a) $f(\mathbf{x})$ je umesto matrice $\mathbf{A}$), gde su $\mathbf{x}_i$ – *kontrolisane promenljive*, koje ne zavise od rezultata opažanja, $\boldsymbol{\theta}$ – *vektor nepoznatih parametara*,

Y_{ij} – *opažanja*, a ε_{ij} – *greške* opažanja sa očekivanjem nula, $\mathsf{M}[\varepsilon_{ij}]=0$, i disperzijom $\mathsf{D}[\varepsilon_{ij}]=\sigma_i^2$. Saglasno (2.2.13) i (2.2.13'), ovde za (1.5.1) i $f(\mathbf{x}_i)^{\mathrm{T}}\boldsymbol{\theta}$ imamo nazive:

$$Y_{ij} = f(\mathbf{x}_i)^{\mathrm{T}}\boldsymbol{\theta}+\varepsilon_{ij} \qquad \text{– \textbf{\textit{jednačina}} \textit{regresije}} \tag{1.5.2}$$

$$f(\mathbf{x}_i)^{\mathrm{T}}\boldsymbol{\theta} \qquad \text{– \textbf{\textit{funkcija}} \textit{regresije.}} \tag{1.5.2′}$$

Zadatak (1.5.1) mora posedovati sledeća *suštinski značajna svojstva*:

(1) Svojstvo 1: Funkcije $f(\mathbf{x})$ su iz poznate klase.

(2) Svojstvo 2: Postoji mogućnost kontrole x, tj. postoji mogućnost da se ono bira iz skupa X po nahođenju istraživača – eksperimentatora. Skup X se često naziva ***oblast dejstva***, a promenljive x - *kontrolisanim promenljivim*.

(3) Svojstvo 3: Greške imaju ***aditivnu strukturu***, sa matematičkim očekivanjem nula, disperzijom $\mathsf{D}[\varepsilon_{ij}]=\sigma_i^2=\sigma^2(x_i)$, *postoje njihovi drugi momenti* i one su *nekorelisane*: $\mathsf{M}[\boldsymbol{\varepsilon}_{ij}\boldsymbol{\varepsilon}_{i'j'}^{\mathrm{T}}]=\mathbf{0}$ pri $i\neq i'$ i $j\neq j'$.

(4) Svojstvo 4: Postoji mogućnost ocene parametara $\boldsymbol{\theta}$ ili neke linearne funkcije od njih, $\mathbf{h}^{\mathrm{T}}\boldsymbol{\theta}$.

Ocenom po *metodi najmanjih kvadrata* (MNK) naziva se ocena

$$\hat{\boldsymbol{\theta}} = agr\min_{\boldsymbol{\theta}\in\boldsymbol{\Theta}}\sum_i\sum_j\sigma^{-2}(x_i)[Y_{ij}-\eta(x_i,\boldsymbol{\theta})]^2, \tag{1.5.3}$$

gde je $\boldsymbol{\Theta}\in R^u$. Pri tome se pojavljuju dve vrste grešaka: *sistematske greške – greške modela* (detaljno prikazane u poglavljima 4, 5, 6 i 7) i *slučajne greške* (greške određivanja parametara modela).

A sada ćemo dati eksplicitne izraze za ocene (1.5.3) i za ocene funkcije $\mathbf{h}^{\mathrm{T}}\boldsymbol{\theta}$.

Iz regresione teorije sledi da u svojstvu ocena parametara možemo koristiti NLN (najbolje linearne nepomerene) ocene, odnosno MNK ocene [Perović, 2005a i 2005b: formule (6.3.2) glave 6, formule (23.4.13) glave 23 i formule (24.1.7) glave 24 – za $\mathbf{A}$ sa potpunim rangom kolona, (gde $f(\mathbf{x})$ odgovara matrici $\mathbf{A}$, tj. $f(\mathbf{x})\equiv\mathbf{A}$), a ovde date sa (1.5.3.a); formule (14.2.4) glave 14 i formule (14.7.5d) glave 14 – za $\mathbf{A}$ (t. j. za $f(\mathbf{x})$) sa nepotpunoim rangom kolona, a ovde date sa (D.30)]. U ovim formulama slobodni član je $\mathbf{f}=\mathbf{L}_0-\mathbf{l}$, u našem modelu (1.5.2) $\mathbf{l}$ je jednako $\mathbf{Y}$, ($\mathbf{l}\equiv\mathbf{Y}$), pa pošto je (1.5.2) linearna funkcija po $\mathbf{Y}$ to se za približnu vrednost može uzeti $\mathbf{L}_0=\mathbf{0}$ i tako postaje $\mathbf{f}=-\mathbf{l}=-\mathbf{Y}$. Tako zaredom dobijamo:

$$\hat{\boldsymbol{\theta}}=\mathbf{N}^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{PY} \qquad \text{- za regularno } \mathbf{N}, \tag{1.5.3a}$$

$$\tilde{\boldsymbol{\theta}}=\mathbf{N}^{-}\mathbf{A}^{\mathrm{T}}\mathbf{PY} \qquad \text{- za singularno } \mathbf{N} \text{ (videti (D.30),} \tag{1.5.3b}$$

$$\hat{\boldsymbol{\theta}}=\mathbf{N}^{+}\mathbf{A}^{\mathrm{T}}\mathbf{PY} \qquad \text{- za singularno } \mathbf{N} \text{ (videti (D.66),} \tag{1.5.3c}$$

gde je $\mathbf{N}^-$ uopštena inverzija od $\mathbf{N}$, $\mathbf{N}^+$ pseudoinverzija od $\mathbf{N}$, $\mathbf{N}=\mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{A}$, $\mathbf{Y}$ je vektor rezultata merenja: $\mathbf{Y}^{\mathrm{T}}=[Y_{11}\cdots Y_{1n_1}\cdots\ \cdots Y_{k1}\cdots Y_{k\,n_k}]$, a $\mathbf{P}$ matrica težina definisana sa (2.2.7'), $\mathbf{P}_{N\times N}=diag\{\overbrace{P_1,\cdots,P_1}^{n_1\,puta},\ \overbrace{P_2,\cdots,P_2}^{n_2\,puta},\ \cdots,\ \overbrace{P_k,\cdots,P_k}^{n_k\,puta}\}$, pri čemu je $P_{i\,j}=P_i=\sigma^{-2}(x_i)$, $\forall\, j\in(1,2,\cdots,n_i)$, $i=1,2,\cdots,k$.

U singularnom slučaju, prema poglavlju 14.2.3. – NLNO I OCENLJIVE FUNKCIJE monografije Metod najmanjih kvadrata (Perović, 2005a i 2005b) za ***ocenljive funkcije*** $\mathbf{h}^{\mathrm{T}}\boldsymbol{\theta}=\mathbf{c}^{\mathrm{T}}\mathbf{Y}$, za koje mora važiti (14.2.10) i (14.2.13) poglavlja 14.2.3. – NLNO I OCENLJIVE FUNKCIJE, postoji *linearna nepomerena ocena sa najmanjom disperzijom*. Ta ocena je data sa (D.38) i (D.70) (videti i (14.2.14) poglavlja 14.2.3 – Perović, 2005a i 2005b), tako da ovde glasi:

$$\mathbf{h}^{\mathrm{T}}\boldsymbol{\theta}=\mathbf{h}^{\mathrm{T}}\tilde{\boldsymbol{\theta}}=\mathbf{h}^{\mathrm{T}}\hat{\boldsymbol{\theta}}=\mathbf{c}^{\mathrm{T}}\hat{\mathbf{Y}},\quad (\mathbf{l}\equiv\mathbf{Y}),\tag{1.5.4}$$

t.j. dobija se sa bilo kojim rešenjem normalnih jednačina, (1.5.3b) ili (1.5.3c) – za singularan slučaj, pri čemu je $\hat{\mathbf{Y}}=\mathbf{Y}+\hat{\mathbf{v}}$ MNK ocena vektora merenih veličina $\mathbf{L}$, odnosno $\mathbf{Y}$, ($\mathbf{l}\equiv\mathbf{Y}$).

Prema (1.5.3), parametar $\boldsymbol{\theta}$ se, dakle, ocenjuje pomoću statistike

$$\hat{\boldsymbol{\theta}}(x_{i1},x_{i2},\cdots,x_{ik};Y_{i1},Y_{i2},\cdots,Y_{n_i}).$$

Pri tome se, po pravilu, pojavljuje zavisnost greške $\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}$ određivanja prave vrednosti parametra $\boldsymbol{\theta}$ od izbora tačaka $x_{i1},x_{i2},\cdots,x_{ik}$ u kojima se meri funkcija y. To daje mogućnost konstruisanja ***kriterijuma kvaliteta eksperimenta*** (obično neke norme greške parametara $\hat{\boldsymbol{\theta}}$) i *planiranja eksperimenta* – ako eksperimentator može birati tačke $x_{i1},x_{i2},\cdots,x_{ik}$ – kako je rečeno ispred u svojstvu (2).

Skup tačaka u kojima se meri, $\{x_1,x_2,\ldots,x_k\}$, sa skupom njihovih mera $\{p_1,p_2,\ldots,p_k\}$, $p_i=n_i/N$, t.j. skup:

$$\xi_N=\{x_1,x_2,\ldots,x_k;\,p_1,p_2,\ldots,p_k\};\quad p_i=\frac{n_i}{N},\tag{1.5.5}$$

nazivamo ***planom eksperimenta***. Tačke x_i nazivamo ***osnovnim*** *tačkama plana*, a veličine p_i – ***merama*** tih tačaka (ako su opažanja iste preciznosti, onda one predstavljaju *težine* opažanja). Prema učinjenim pretpostavkama kovarijaciona matrica NLN ocena potpuno je određena planom eksperimenta:

$$\mathbf{K}[\hat{\boldsymbol{\theta}}_N]=N^{-1}\mathbf{M}^{-1}(\xi_N)=N^{-1}\mathbf{K}(\xi_N),\tag{1.5.6}$$

gde je $\mathbf{M}(\xi_N)$ – (*normirana*) *informaciona matrica*, $\mathbf{K}(\xi_N)=\mathbf{M}^{-1}(\xi_N)$. Sa uvedenom pretpostavkom o nekorelisanosti grešaka $\varepsilon_{i\,j}$ (svojstvo (3) ispred)

$$\mathbf{M}(\xi_N) = \sum_{i=1}^{k} n_i P_i \mathbf{K}(x_i)\,, \tag{1.5.7}$$

gde je $\mathbf{K}(x_i)$ *prirast informacione matrice*, proizveden ponašanjem (uticajem) jednog opažanja u tački x_i. U slučaju skalarnog odziva, kada je $f(\mathbf{x}) = [f(x_1), \cdots, f(x_k)]$ matrica $\mathbf{K}(\mathbf{x})$ ima jedinični rang i jednaka je $\sigma^{-2}(x) f(\mathbf{x}) f^{\mathrm{T}}(\mathbf{x})$. A u opštem slučaju kada je vektor merenja $\mathbf{Y}$ dimenzija $N \times 1$ tada je:

$$\mathbf{K}(\mathbf{x}) = f(\mathbf{x})^{\mathrm{T}} \mathbf{K}_{\boldsymbol{\varepsilon}}^{-1}(\mathbf{x}) f(\mathbf{x}), \quad \text{sa} \quad \mathbf{K}_{\boldsymbol{\varepsilon}}(\mathbf{x})_{N\times N} = \mathsf{M}[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{\mathrm{T}}]_{N\times N}\,. \tag{1.5.8}$$

Pošto matrica $\mathbf{M}(\xi_N)$, a samim tim i matrica $\mathbf{M}^{-1}(\xi_N)$, *ne zavisi od nepoznatih parametara* $\boldsymbol{\theta}$, onda ima smisla govoriti o ***apriornom istraživanju planova*** koji *minimiziraju neku zadatu funkciju* R od matrice $N\mathbf{M}(\xi_N)$ (*Ermakov i dr., 1983*) :

$$\xi_N^* = \arg\inf_{\xi_N} R[N\mathbf{M}(\xi_N)]\,. \tag{1.5.9}$$

Dakle, *variranjem plana* (1.5.5), *pri čemu ključnu ulogu igra skup osnovnih tačaka plana* $\{x_1, x_2, \ldots, x_k\}$, *dolazimo do optimalnog plana eksperimenta*.

Dâti zadatak o ekstremumu nazivamo ***zadatkom optimalnog planiranja eksperimenta***. Rešenje tog zadatka nazivamo ***optimalnim planom***, a funkciju R – ***kriterijumom optimalnosti***, ili *funkcijom kriterijuma*. Napomenimo da smo u razmatranje uzeli funkciju $R[N\mathbf{M}(\xi_N)] > 0$ od matrice $\mathbf{M}(\xi_N)$, a ne od matrice $\mathbf{K}(\xi_N)$, jer je pogodnije, pošto obuhvata slučajeve kada je $r(\mathbf{M}(\xi_N)) < u$, (kao što je, na primer, slučaj pri ocenjivanju linearnih kombinacija parametara $\boldsymbol{\theta}$); pri tome je u – broj nepoznatih parametara $\boldsymbol{\theta}$.

Neka je u broj parametara $\boldsymbol{\theta}$ onda zavisno od ranga matrice $\mathbf{M}(\xi_N)$ imamo:

$r(\mathbf{M}(\xi_N)) < u$ - ***singularni*** plan, i

$r(\mathbf{M}(\xi_N)) = u$ - ***regularni*** plan.

Na optimalno rešenje ξ_N^* iz (1.5.9), preko funkcije kriterijuma $R[N\mathbf{M}(\xi_N)$ utiče i neadekvatnost modela (definicija neadekvatnosti modela u poglavlju 4.6), o čemu se detaljnije govori u poglavlju 14.2. – Optimalni planovi u regresionim eksperimentima i u poglavlju 14.5. tačka B5 – Moć kriterijuma.

U vezi sa planiranjem pogodno je istaći sledeće tipove eksperimenata: *ekstremalni eksperiment*, *eksperiment razlikovanja*, *eksperiment izostavljanja* i *imitacioni eksperiment*. Oni su opisani u poglavlju 14.3.

A o vrstama kriterijuma biće reči u poglavlju 14.2. – Optimalni planovi u regresionim eksperimentima.