Uporaba statističkih metoda u vojnom sustavu

Toljan, Marko

Undergraduate thesis / Završni rad

2019

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Mechanical Engineering and Naval Architecture / Sveučilište u Zagrebu, Fakultet strojarstva i brodogradnje**

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:235:243204

Rights / Prava: Attribution-NonCommercial 4.0 International/Imenovanje-Nekomercijalno 4.0 međunarodna

Download date / Datum preuzimanja: **2024-04-24**

Repository / Repozitorij:

Repository of Faculty of Mechanical Engineering and Naval Architecture University of Zagreb

SVEUČILIŠTE U ZAGREBU
FAKULTET STROJARSTVA I BRODOGRADNJE

ZAVRŠNI RAD

Marko Toljan

Zagreb, 2019. godina.

SVEUČILIŠTE U ZAGREBU
FAKULTET STROJARSTVA I BRODOGRADNJE

ZAVRŠNI RAD

Mentor: Student:

Doc. dr. sc. Hrvoje Cajner Marko Toljan

Zagreb, 2019. godina.

Izjavljujem da sam ovaj rad izradio samostalno koristeći znanja stečena tijekom studija i navedenu literaturu.

Zahvaljujem mentoru doc. dr. sc. Hrvoju Cajneru, dipl. ing. na dodjeljivanju zadatka, ukazanoj pomoći i savjetima, a najviše na strpljenju.

Zahvaljujem i obitelji, kolegama i prijateljima te djevojci na svim oblicima pomoći i podrške koje su mi pružili.

Marko Toljan

SVEUČILIŠTE U ZAGREBU
FAKULTET STROJARSTVA I BRODOGRADNJE
Središnje povjerenstvo za završne i diplomske ispite
Povjerenstvo za završne ispite studija strojarstva za smjerove:
proizvodno inženjerstvo, računalno inženjerstvo, industrijsko inženjerstvo i menadžment, inženjerstvo materijala i mehatronika i robotika

Sveučilište u Zagrebu Fakultet strojarstva i brodogradnje	
Datum	Prilog
Klasa:	
Ur.broj:	

ZAVRŠNI ZADATAK

Student: **Marko Toljan** Mat. br.: 0035201294

Naslov rada na hrvatskom jeziku: **Uporaba statističkih metoda u vojnom sustavu**

Naslov rada na engleskom jeziku: **The use of statistical methods in military system**

Opis zadatka:

Realni sustavi konstantno su podložni promjenama kako zbog vanjskih tako i zbog unutrašnjih čimbenika. Generirani podaci iz sustava daju sliku o stanju sustava i njegovom stupnju varijabilnosti. Kako bi sustav bio kvalitetno analiziran potrebno je koristiti reprezentativne podatke i adekvatne metode analize. Stoga je uporaba statističkih metoda vrlo česta u svim aspektima društva, od društvenih istraživanja, inženjerstva pa sve do vojnih sustava. Metode primjerene sustavima iz područja inženjerstva vrlo lako je moguće aplicirati i na područje vojnog sustava odnosno vojnih operacija.

U radu je potrebno:

1. Klasificirati statističke metode pogodne za rješavanje problema u vojnom sustavu.
2. Teorijski razraditi najzastupljenije metode analize podataka.
3. Dati presjek uporabe statističkih metoda u ratnim operacijama.
4. Na temelju realnih primjera demonstrirati neke od metoda koristeći se gotovim programskim rješenjima.

Zadatak zadan:
29. studenog 2018.

Rok predaje rada:
1. rok: 22. veljače 2019.
2. rok (izvanredni): 28. lipnja 2019.
3. rok: 20. rujna 2019.

Predviđeni datumi obrane:
1. rok: 25.2. - 1.3. 2019.
2. rok (izvanredni): 2.7. 2019.
3. rok: 23.9. - 27.9. 2019.

Zadatak zadao:

Dr. sc. Hrvoje Cajner, doc.

Predsjednik Povjerenstva:

Prof. dr. sc. Branko Bauer

SADRŽAJ

SADRŽAJ I

POPIS SLIKA III

POPIS TABLICA IV

POPIS OZNAKA V

SUMMARY VIII

1. UVOD 1

2. PREGLED KORIŠTENJA STATISTIČKIH METODA U VOJNIM SUSTAVIMA 2

2.1. Analiza učinkovitosti opreme i metoda 2

2.1.1. Raspored borbenog napora 2

2.1.1.1. Plan broja tenkova za invaziju [3] 3

2.1.1.2. Plan patrola za suzbijanje aktivnosti podmornica [4] 3

2.1.2. Ocjena adekvatnosti i efikasnosti korištenja opreme 5

2.1.2.1. Unaprjeđenje bombi s dubinskom detonacijom 5

2.1.3. Aktivnosti podrške vojnim operacijama 7

2.1.3.1. Vođenje statistike u medicinskim ustanovama 7

2.1.3.2. Ispitivanje utjecaja eksplozija [6] 8

2.1.3.3. Optimalna veličina konvoja 9

2.2. Simulacijske metode 11

2.2.1. Simulacija ratnih operacija [10] 11

2.2.2. Planiranje ljudskih potencijala 12

2.3. Probabilističke metode 12

2.3.1. Vjerojatnost uništavanja mete [4] 12

2.3.2. Vjerojatnost preživljavanja pogotka [12] 14

3. PRIMJERI NAJČEŠĆE KORIŠTENIH METODA 16

3.1. Hi-kvadrat test 16

3.1.1. Hi-kvadrat test za dva nezavisna uzorka [14] 17
3.1.2. Analiza učinkovitosti protuzračne opreme na trgovačkim brodovima [4] 17
3.1.2.1. Uspješnost neprijateljskih napada – broj pogodaka.....18
3.1.2.2. Uspješnost neprijateljskih napada – broj potapanja.....20
3.2. Fisherov egzaktni test [15] 22
3.2.1. Analiza učinkovitosti obrane od torpeda [4].....23
3.3. Regresijska analiza.....25
3.3.1. Višestruka regresijska analiza 25
3.3.2. Logistička regresija.....26
3.3.3. Primjer - Mjera borbene uspješnosti [17] 26
3.3.3.1. Regresija.....26
3.3.3.2. Mjera borbenog uspjeha.....27
3.3.3.3. Predviđanje vjerojatnosti ishoda.....29
3.4. Markovljev lanac 30
3.4.1. Apsorpcijski Markovljev lanac.....31
3.4.2. Primjer - Model dinamike vojnog osoblja [21].....32
3.4.2.1. Opis podataka i podloga za model.....32
3.4.2.2. Model Markovljeva lanca za činove 36
3.4.2.3. Model Markovljevog lanca za sva stanja 38
4. ZAKLJUČAK 42
LITERATURA.....43
PRILOG.....45

POPIS SLIKA

Slika 1. Vjerojatnost potapanja podmornice s obzirom na dubinu 6
Slika 2. Kružni dijagram Florence Nightingale [9].....8
Slika 3. Udio uništenih meta.....14
Slika 4. Slučajna raspodjela pogodaka 15
Slika 5. Stvarna raspodjela pogodaka.....15
Slika 6. Raspodjela vjerojatnosti ishoda za Grupu 1.....28
Slika 7. Usporedba predviđene i stvarne vjerojatnosti ishoda 30
Slika 8. Grafički prikaz Markovljeva lanca s tri stanja 31
Slika 9. Tranzicijski dijagram za sva stanja.....33
Slika 10. Vjerojatnosti nastavka službe unutar trenutnog čina 34
Slika 11. Vjerojatnost napredovanja s obzirom na trajanje trenutnog čina 35
Slika 12. Tranzicijski dijagram za razrede.....36

POPIS TABLICA

Tablica 1. Učinak patrolnih letova 4
Tablica 2. Učinak letenja ovisno o udaljenosti od obale 4
Tablica 3. Podaci o napadima podmornica na konvoje 10
Tablica 4. Bombardiranje velike površine 13
Tablica 5. Raspored pogodaka na zrakoplovima 14
Tablica 6. Šteta pretrpljena pri bombardiranju 18
Tablica 7. Metoda formiranja tablice kontingencije 19
Tablica 8. Tablica kontingencije - opažene frekvencije pogodaka 19
Tablica 9. Tablica kontingencije - očekivane frekvencije pogodaka 19
Tablica 10. Tablica kontingencije - opažene frekvencije potapanja 20
Tablica 11. Tablica kontingencije - očekivane frekvencije pogodaka 21
Tablica 12. Računska tablica za hi-kvadrat 21
Tablica 13. Tablica kontingencije - opažene frekvencije 24
Tablica 14. Nezavisne varijable 26
Tablica 15. Oznake varijabli 27
Tablica 16. Regresijski koeficijenti 27
Tablica 17. Matrica nesigurnosti 28
Tablica 18. Točnost predviđanja s y_0 29
Tablica 19. Rezultati logističke regresije 29
Tablica 20. Prisilno napuštanje po stanjima 35
Tablica 21. Iznosi indeksa optimalne odluka 40

POPIS OZNAKA

Oznaka	Mjerna jedinica	Opis
n_t	-	predviđen broj jedinica
m_t	-	najveći opažen broj jedinice
k_t	-	broj opaženih jedinica
A	m^2	površina pod nadzorom
T_L	h	ukupno vrijeme letenja
N_p	-	broj opaženih podmornica
n_p	-	broj pratećih brodova
N_K	-	veličina konvoja
L	m^2	smrtonosna površina
P_d	-	vjerojatnost uništenja mete pogotkom u točku
P_D	-	vjerojatnost uništenja mete
P_D'	-	vjerojatnost uništenja mete s više bombi
χ^2	-	hi-kvadrat parametar
f_o	-	opažene frekvencije
f_t	-	očekivane (teoretske) frekvencije
H_0	-	nulta hipoteza
H_1	-	alternativna hipoteza
π_1, π_2	-	frekvencije u populacijama
α	-	granica značajnosti
N	-	veličina populacije
K	-	broj varijabli traženih svojstava u populaciji
n	-	veličina uzorka
y	-	zavisna varijabla linearne regresije
$x_1, x_2, \ldots$	-	nezavisne varijable linearne regresije
ε	-	slučajna promjenjiva varijabla linearne regresije
$\beta_0, \beta_1, \ldots$	-	nepoznati parametri linearne regresije
$\mathbb{R}$	-	skup realnih brojeva
CASA	-	broj žrtava napadača
CASD	-	broj žrtava napadnute strane
STRA	-	veličina napadačevih postrojbi
STRD	-	veličina postrojbi napadnute strane

KMDA	km	udaljenost za koju je napadač prešao preko obrambene linije
WOFA1	km	širina napadačevog fronta na početku bitke
ACHA	-	uspjeh napadačevih misija
ACHD	-	uspjeh misija napadnute strane
R^2	-	koeficijent determinacije
MCS	-	mjera borbenog uspjeha
y_L	-	MCS ispod kojeg je napadač uvjerljivo u prednosti
y_R	-	MCS iznad kojeg je napadač uvjerljivo u prednosti
s_0, s_1, ...	-	diskretne varijable
s_0, s_1, ...	-	stanja diskretnih varijabli
E1, E2,...	-	činovi
TIG	godina	broj godina provedenih u činu
VL	-	dobrovoljno napuštanje čina
IVL	-	nedobrovoljno napuštanje čina
RCP	-	kontrolna točka zadržavanja
θ_{ij}	-	vjerojatnost prelaska iz stanja *i* u stanje *j*
ω_t	-	funkcija predviđene kompenzacije

SAŽETAK

Tema rada je uporaba statističkih metoda u vojnim sustavima. U vojnim sustavima zbog velike važnosti i visokih ulaganja vlada stalna potreba za olakšavanjem ili unaprjeđivanjem aktivnosti. Veliki razmjeri i kompleksnost sustava taj zadatak otežavaju i zato je potrebno pronaći metode za dobivanje reprezentativnih podataka i metode kojima će se ti podaci obraditi tako da se dobije što točniji zaključak.

U radu je iznesena klasifikacija područja najčešće primjene statističkih metoda temeljena na zastupljenosti u dostupnoj literaturi, područja su ukratko opisana i ilustrirana primjerima, a neke od metoda su teorijski razrađene i primjeri njihovog korištenja u praksi su detaljno objašnjeni.

Ključne riječi: analiza vojnog sustava, operacijska istraživanja

SUMMARY

The topic of the paper is the use of statistical methods in military systems. In military systems, due to the great importance and high investment, there is a constant need to facilitate or enhance activities. The large scale and complexity of the system make the task difficult, and it is therefore necessary to find methods for obtaining representative data and the methods for processing that data in order to draw the most accurate conclusion.

The paper presents a classification of the areas where statistical methods are most frequently used based on the representation in the available literature, the areas are briefly described and illustrated with examples, and some of the methods are theoretically elaborated and examples of their use in practice are explained in detail.

Key words: military system analysis, operations research

1. UVOD

Vojska je jedna od temeljnih institucija svake države, prosječno u svijetu izravno zapošljava 0.8% radne snage [1], a neizravno, kroz proizvodnju vojne opreme i izgradnju infrastrukture, višestruko više. Kod tako velikih sustava je posebno važno da su što bolje isplanirani, organizirani i vođeni jer pogreške u bilo kojem segmentu mogu imati neprocjenjive dalekosežne posljedice. Isto vrijedi i za operacije u vojnim sukobima. Kako bi se što objektivnije mogle donositi tako važne odluke, potrebno je sustave promatrati kroz veličine koje se mogu iskazati brojčano. U radu je prikazano kako se za to najčešće koriste vjerojatnosti. Nekad je ta vjerojatnost dovoljno izražena da se na temelju nje donese odluka, a nekad ona služi samo kao pomoć osobi zaduženoj za donošenje odluke.

Pregledom literature (knjiga, priručnika, znanstvenih radova…) je stečen uvid u područja najčešće primjene statističkih metoda pri donošenju odluka, a to su *analiza učinkovitosti* koja služi za pomoć pri odabiru opreme i metoda, *simulacijske metode* kojima se predviđa ponašanje sustava u budućnosti te *probabilističke metode* koje obuhvaćaju uglavnom probleme preciznosti oružja. Najveći iskorak u području analize vojnih sustava se dogodio za vrijeme Drugog svjetskog rata, kad su znanstvenici iz raznih grana postavili temelje danas široko raširene discipline operacijskih istraživanja, pa je tako većina korištene literature vezana uz to razdoblje.

Najviše pozornosti je posvećeno prvim dvama područjima jer se metode korištene kod njih ipak češće pronalazi primjena u inženjerskoj struci dok su probabilističke metode više matematičke prirode i primjenjuju se samo u nekim specifičnim slučajevima.

2. PREGLED KORIŠTENJA STATISTIČKIH METODA U VOJNIM SUSTAVIMA

Prostora za primjenu neke statističke metode postoji u svakoj aktivnosti u vojnom sustavu, a u ovom poglavlju je prikazan pokušaj klasifikacije područja primjene na neke najosnovnije grupe.

- analiza učinkovitosti opreme i metoda
    - raspored borbenog napora
    - ocjena adekvatnosti i efikasnosti korištenja opreme
    - aktivnosti podrške vojnim operacijama
- simulacijske metode
    - simulacije ratnih operacija
    - planiranje ljudskih potencijala
- probabilističke metode
    - vjerojatnost pogađanja mete
    - vjerojatnost preživljavanja pogotka

Uz klasifikaciju su ukratko prikazani primjeri koji ilustriraju navedeno područje.

2.1. Analiza učinkovitosti opreme i metoda

U vojsci, kao konzervativnom i strogo hijerarhijski uređenom sustavu, izražena je pojava praćenja uvriježenih procedura bez preispitivanja njihove ispravnosti što može dovesti do nepotrebnih troškova. Zanimljiva je anegdota koju prepričava jedan od pionira operacijskih istraživanja Solandt [2], koji je pri analizi procedure korištenja određenog modela topa utvrdio da jedan član jedinice samo stajao sa strane. Kad je provjerio koja je njegova funkcija, otkrio je da je on zadužen za držanje konja. Naravno, konji su prestali vući topove 20 godina prije toga, ali taj zadatak nije izbačen iz procedura prilikom modernizacije. Iz tog razloga treba kritički pristupiti i analizirati svaku aktivnost, koliko god da se banalna činila.

2.1.1. Raspored borbenog napora

Uvjeti u kojima se odvijaju vojne operacije ne dozvoljavaju preveliku slobodu pri raspodjeli resursa, nepotrebna dodjela resursa jednom području može imati dalekosežne posljedice na onom gdje resursa zbog te odluke fali. Zato treba kvantitativnim metodama opravdati donošenje odluke kako bi se eliminiralo subjektivne faktore, iskustvo, predviđanje, itd.

2.1.1.1. Plan broja tenkova za invaziju [3]

Prilikom planiranja aktivnosti vezanih uz invaziju Normandije, Savezničke snage su se susrele s nepoznanicom koja je mogla utjecati na ishod operacije. Razvoj njemačkog tenka *Panther* je predstavljao opasnost za inače superiorne Američke *Sherman* tenkove pa je bilo potrebno znati sa koliko točno tenkova Njemačka raspolaže. Do tad su bila zarobljena samo dva tenka i pomoću njih je bilo potrebno odrediti ukupni proizvodni kapacitet. Već ranije su provedene analize serijskih brojeva na raznim dijelovima, gumama, šasijama, reduktorima, itd. pa su bili poznati neki od algoritama dodjeljivanja. Dijelovi tenka koji su se pokazali najpogodnijima za korištenje u analizi su kotači gusjenica jer ih je na svakom tenku 48.

$$n_t \approx m_t - 1 + \frac{m_t}{k_t} \tag{1}$$

Ako imamo niz serijskih brojeva $\sum_{i=1}^{n} n + 1$, prikazani izraz predstavlja procjenu najvećeg broja n_t, gdje je m_t najveći viđen broj, a k_t broj viđenih brojeva. Korištenjem te logike, na temelju serijskih brojeva s kotača gusjenica je procijenjen broj kalupa s kojim njemački proizvođači raspolažu, a iz toga je, u konzultacijama s britanskim proizvođačima, procijenjen njemački proizvodni kapacitet. Procjena broja tenkova *Panther* je iznosila 270 komada, a nakon rata je pregledom podatka iz proizvodnje otkriven stvarni broj od 276 komada.

Dobiveni podatak je upozorio na veći broj tenkova od očekivanog i Saveznici su mogli shodno tome prilagoditi aktivnosti. Analiza nije doprinijela samo operaciji invazije, već su njom otkrivene razne karakteristike njemačke proizvodnje koje su iskorištene za taktičko bombardiranje postrojenja.

2.1.1.2. Plan patrola za suzbijanje aktivnosti podmornica [4]

Britanske snage su od samog početka 2. Svjetskog rata velik udio resursa ulagali u borbu protiv njemačkih podmornica koje su predstavljale veliku opasnost sjevernoatlantskim trgovačkim putevima, britanskoj slamci spasa. Ta borba je bila pod povećalom jer je britanska vlast bila vrlo naklona ideji bombardiranja njemačkih radničkih gradova kako bi se oslabila proizvodnja, a za ostvarivanje te ideje je bilo potrebno što više raspoloživih zrakoplova. Zato je postojala potreba maksimalnog iskorištenja flote. Tablica 1. prikazuje podatke o tri mjeseca patrolnih letova.

Tablica 1. Učinak patrolnih letova

	Područje A	Područje B	Područje C
Površina (tisuće kvadratnih milja), A	300	600	900
Ukupno vrijeme letenja, (tisuće sati), T_L	69	22	16
T_L/A, (sati po kvadratnoj milji)	0,23	0,037	0,018
Broj opaženih podmornica, N_P	116	51	24
N_P/T_L, (opaženih po tisuću sati letenja)	1,68	2,32	1,5

Na prvi pogled je vidljivo da su patrole nad područjem A znatno manje učinkovite, sat letenja u tom području daje jednako ili manje susreta s podmornicama iako se pokriva dvostruko ili trostruko manja površina. Niti pokušaji slanja iskusnijih eskadrila nisu poboljšali učinkovitost u području A tako da je donesena odluka o dodatnoj analizi rasporeda letenja.

Prva pretpostavka koju je trebalo ispitati je bila da u tom prostoru nije jednaka gustoća podmornica kao u ostalim područjima. Pregledom podataka prikazanih (Tablica 2.) je utvrđeno da je u području A puno veći udio letova prolazio uz obalu gdje je aktivnost podmornica manja.

Tablica 2. Učinak letenja ovisno o udaljenosti od obale

Udaljenost od obale, (milje)	0-60	60-120	120-180	180-240
Ukupno vrijeme letenja, (tisuće sati), T_L	15,5	3,7	0,6	0,17
Broj opaženih podmornica, N_P	21	11	5	2
N_P/T_L, (opaženih po tisuću sati letenja)	1,36	2,97	8,33	11,77

Jasno je da letenje uz samu obalu nije učinkovit oblik patroliranja pa je potrebno preraspodijeliti jedinice tako da više pokrivaju vanjski rub područja. Ako bi se 10% jedinica iz zone najbliže obali prebacilo na dvije vanjske zone, gustoća letenja u prvoj zoni bi se smanjila neznatno s obzirom na N_P/T_L, a u vanjskim zonama bi se višestruko povećala i gustoća, i broj opaženih podmornica. U stvarnosti, kad je uvedena ravnomjernija raspodjela letenja, podmornice su

prilagodile obrasce kretanja (vrijeme i doba dana pod vodom) tako da je N_P/T_L za vanjske zone pao na 4-5, ali to je i dalje rezultiralo povećanjem ukupne učinkovitosti.

Ovaj primjer dobro ilustrira kako stanje na terenu nije sasvim jasno dok se ne provede detaljnija analiza iako se rezultati retrospektivno čine očitima. Zapovjednici u početku vidjeli da zrakoplovi nailaze na puno više podmornica bliže obali, a kako nisu odmah imali raspoložive podatke slične izloženima, zaključak je bio kako je najbliže obali najviše podmornica.

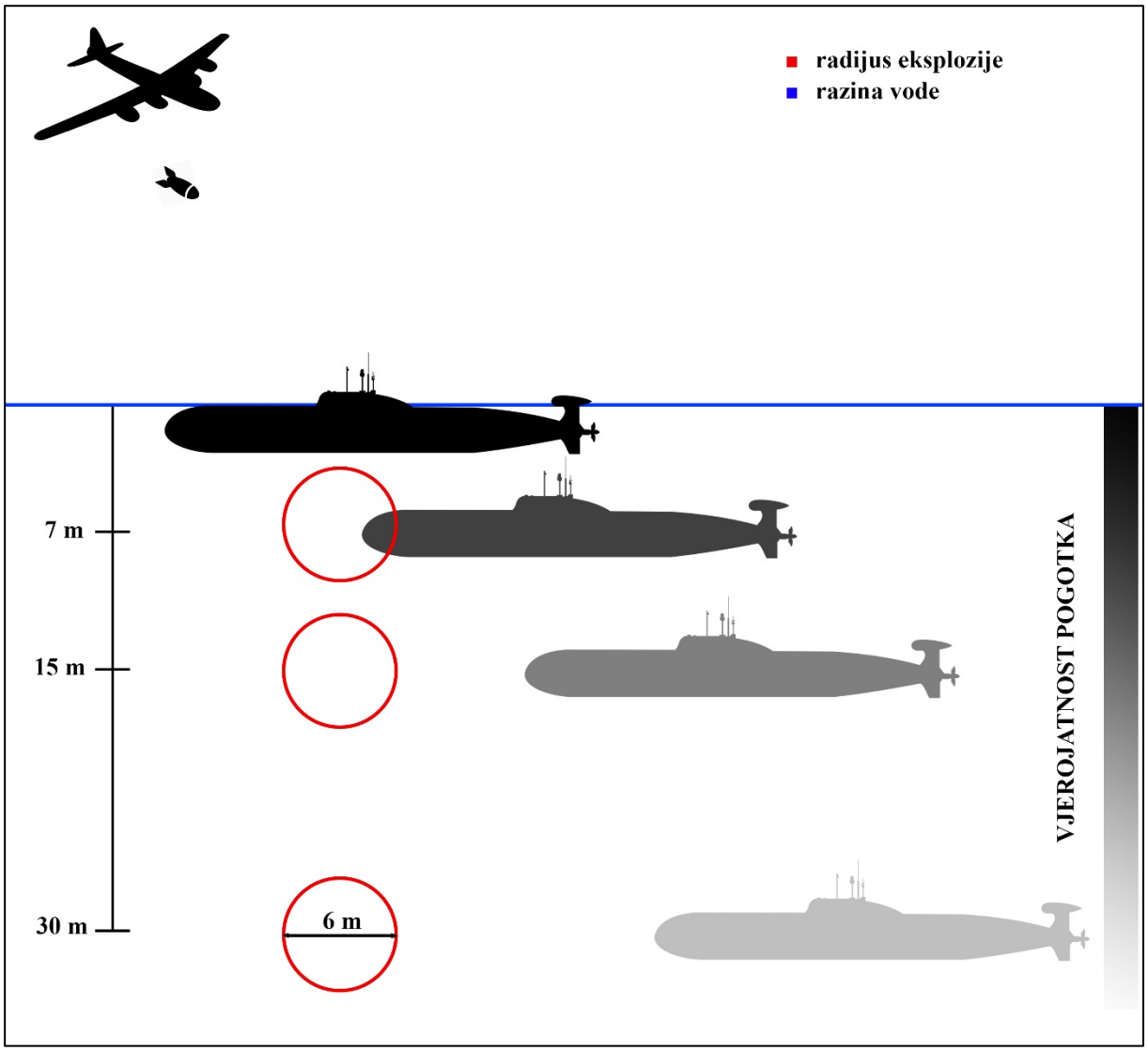
2.1.2. Ocjena adekvatnosti i efikasnosti korištenja opreme

U prethodnom primjeru su zrakoplovi koji su vrlo upotrebljivi za traženje podmornica korišteni na krivi način, ali su možda su i češći primjeri korištenja opreme za zadatak za koji uopće nije namijenjena. Tako Blackett [5] navodi primjer korištenja noćnih bombardera za traženje podmornica i rezultate ispod očekivanja. Analizirane su metode treninga posada, razmotrene opcije boljih dalekozora, istražen utjecaj kuta letenja u odnosu na sunčeve zrake, ali pitanje koje je dovelo do konačnog rješenja je bilo „Koje su boje zrakoplovi?“. Noćni bombarderi su obojenu anti-reflektivnom crnom bojom kako bi se smanjio utjecaj neprijateljskih svjetala, ali naknadno provedeni eksperimenti su utvrdili da se svijetlo obojeni avioni opaženi na 20% manjoj udaljenosti. To je podmornicama ostavilo manje vremena za zaron i rezultati su se poboljšali. Nisu sva poboljšanja bila tako jednostavna, ali obično je potreba za poboljšanjima otkrivana iz jednostavne analize učinkovitosti. Dva primjera jednostavne analize su prikazana u poglavljima 3.1 i 3.2.

2.1.2.1. Unaprjeđenje bombi s dubinskom detonacijom

U početku su zrakoplovi podmornice napadali s običnim bombama koje bi eksplodirale pri dodiru s površinom vode. Naravno, zbog ograničenog radijusa eksplozije, učinkovitost je bila niska jer su podmornice uspijevale zaroniti na sigurnu dubinu. Iz tog razloga je započelo naoružavanje zrakoplova bombama koje bi eksplodirale kad bi potonule na određenu dubinu. Problem je bio taj što su te bombe razvijene za izbacivanje s brodova pa je zato dubina aktiviranja postavljena na 30 metara kako eksplozija ne bi zahvatila brodove. Brodovi su također bili opremljeni sonarima pa su bombe ispuštali dovoljno precizno. [6] Zrakoplovi nisu mogli znati gdje se podmornica nalazi nakon zarona i bombe su ispuštali nad mjestom zarona tako da su rijetko pogađali metu zbog radijusa eksplozije od samo 6 metara. Ta vrsta bombi je proglašena neučinkovitom i počeo je razvoj nove, znatno kompleksnije i skuplje bombe koja bi eksplodirala kad bi se približila podmornici, neovisno na kojoj se to dubini dogodilo. [5]

Međutim, tim zadužen za to je uvidio očite probleme postojeće metode napada i predložio da se smanji dubina eksplozije kako bi se smanjio učinak pomicanja podmornice u horizontalnoj ravnini. Predložena je dubina aktivacije od 15 metara, ali analiza uspješnih pogodaka je pokazala da je u 40% slučajeva podmornica bila na površini, a u još 10% je barem neki dio podmornice bio iznad površine pa je aktivacija na dubini od 15 metara potpuno neučinkovita u 50% slučajeva, a niti u ostalih 50% nije potpuno učinkovita jer se neke podmornice ipak uspiju udaljiti dovoljno da ih eksplozija ne zahvati. Slika 1. ilustrira vjerojatnost potapanja podmornice.



Slika 1. Vjerojatnost potapanja podmornice s obzirom na dubinu

Analiza vjerojatnosti uspjeha potapanja kao funkcije dubine zarona podmornice je pokazala da je optimalna dubina detonacije na 7 metara čak triput učinkovitija u pogađanju mete nego na 15 metara. Ta dubina je postala standardna praksa te su piloti upućeni da uopće ne pokušavaju

pogoditi podmornice koje su zaronile prije više od pola minute jer toliko vremena treba da izađu van područja učinkovitosti bombi. [4]

2.1.3. Aktivnosti podrške vojnim operacijama

Statističke metode korištene u ratnom razdoblju služe, u konačnici, većem vojnom uspjehu, ali primjena nije nužno vezana uz vojne operacije, već uz logističku podršku.

2.1.3.1. Vođenje statistike u medicinskim ustanovama

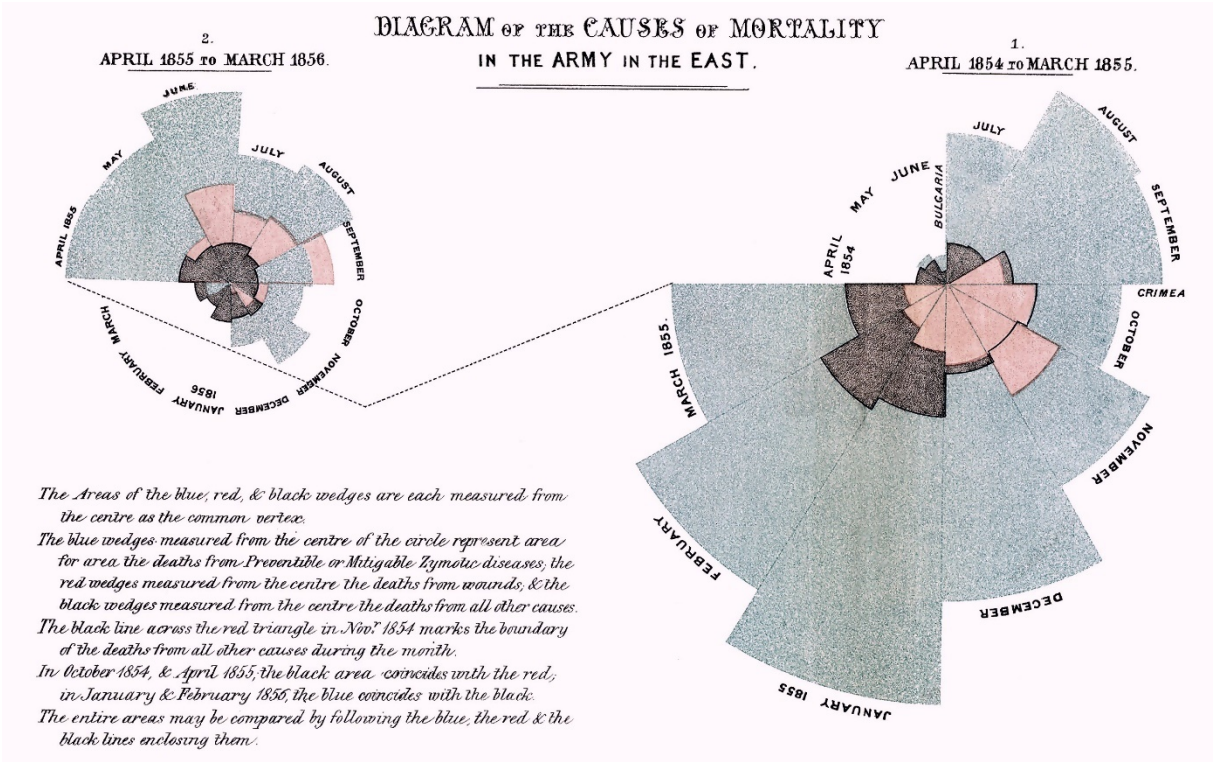
Florence Nightingale je poznata kao utemeljiteljica modernog sestrinstva i reformatorica zdravstva, upamćena je kao „Dama s lampom" koja obilazi ranjenike, ali njezine ideje ne bi niti postojale, niti bi bile tako široko prihvaćene da nije vodila detaljnu evidenciju, i još važnije, osmislila način kako te podatke prikazati nekom tko nije upućen u temu.

Sa svojim djelovanjem je počela u Krimskom ratu. Primijetila je da medicinsko osoblje ne zna koliko ranjenika u bolnicu ulazi i koliko iz nje izlazi pa je počela voditi evidenciju. Njena pretpostavka je da toliko velik broj smrti u bolnici uzrokuje loša hrana i nedostatak lijekova, ali nakon što je u bolnici obavljeno čišćenje ventilacijski, kanalizacijskih i vodovodnih instalacija, stopa smrtnosti je pala sa 47% na 2% [7] bez da se išta promijenilo po pitanju hrane i lijekova. Takvo i slična zapažanja su ju navela da postavi higijenu kao jedan od osnovnih uvjeta za medicinske ustanove.

Nakon povratka u Britaniju, sastavila je izvještaj od 850 stranica [8] u kojem je pokazala kako je glavni uzrok smrti vojnika bio povezan uz bolesti i infekcije, a ne posljedice ranjavanja, suprotno onom što su tvrdili vojni zapovjednici. Kako bi demonstrirala napredak koji je njezina bolnica ostvarila, usporedila je stopu smrtnosti u njoj ss najboljim londonskim bolnicama i pokazala kako više vojnika umire u Londonu nego na bojišnici.

Krimski rat je bio prvi rat popraćen novinarskim izvještajima, javnost je bila više upućena i zainteresirana za zbivanja, a Florence Nightingale je zbog toga mogla lakše obznaniti svoja otkrića. Način na koji je prilagodila podatke za javnost je također revolucionaran. Koristila je tisak u boji, grafičke prikaze omjera umjesto suhoparnog nabrajanja brojeva, a prva je popularizirala oblik kružnog dijagrama (*pie chart)* sličnog današnjem histogramu u polarnim

koordinatama (*circular histogram, polar area diagram)* (Slika 2.).



Slika 2. Kružni dijagram Florence Nightingale [9]

2.1.3.2. Ispitivanje utjecaja eksplozija [6]

Pod stalnom prijetnjom njemačkih bombardiranja je bilo potrebno istražiti kako se određene građevinske strukture ponašaju nakon eksplozije. Iako je primarni fokus bio razvoj što otpornijih industrijskih postrojenja, rezultati do kojih je došao zaduženi odjel su doveli i do povećanja sigurnosti stambenih prostora. Podaci dobiveni u ispitivanjima su onda iskorišteni za predviđanje rezultata neprijateljskih napada i planiranje potrebnih popratnih aktivnosti, broja raspoloživog medicinskog osoblja, vatrogasaca itd. Podaci su korišteni i u drugom smjeru, za predviđanje uspjeha bombardiranja strateških ciljeva. Analize su bile toliko precizne da kad je grad Coventry bombardiran, posljedice su bile gotovo jednake onima koje je neka ranija analiza predvidjela.

Istražen je i utjecaj eksplozija na ljudsko tijelo. Ranije se vjerovalo kako ljudsko tijelo može preživjeti pretlak do 35 kPa, ali istraživanje provedeno nad životinjama je pokazalo da čak i kod pretlaka od 3500 kPa postoji vjerojatnost preživljavanja od 50%. Kad su obrađeni podaci prikupljeni s terena to i dokazali, došlo je do velikih promjena u planiranju gubitaka ljudstva i potrebnih kapaciteta bolnica.

2.1.3.3. Optimalna veličina konvoja

Konvojima nije posvećivana velika pažnja jer se smatralo da su resursi potrebniji u vojnim operacijama, ali nakon što su njemačke snage počele koristiti metodu „čopora vukova“ i nanositi sve veću štetu konvojima, nastala je potreba za promišljanjem isplativosti metoda u uporabi.

Prvi razmotreni problem je bila raspodjela resursa između izgradnje trgovačkih brodova i brodova za zaštitu konvoja. Da bi se više isplatilo usmjeriti resurse na brodove za zaštitu, jedan takav brod bi morao „spasiti“ jedan trgovački brod u konvoju kojeg prati. Podaci su pokazali da konvoji s 9 brodova u pratnji u prosjeku gube 25% manje brodova nego oni sa 6. Iz ukupnog broja i prosječne veličine konvoja te prosječnog broja pratnje je dobiven iznos od 2-3 „spašena“ broda po pratećem brodu. To jasno dovodi do zaključka da je isplativije proizvesti jedan brod za pratnju, međutim, zbog nemogućnosti brzog prilagođavanja brodogradilišta drugom tipu brodova, ta opcija je odbačena. [5]

Sljedeći je istražen učinak brzine kretanja konvoja i zračnog nadzora. Podaci su pokazali da su konvoji s brzinom od 9 čvorova prosječno gubili upola manje brodova nego sporiji konvoji s brzinom od 7 čvorova. To je lako objašnjeno brzinom podmornice pod vodom koja je manja od 9 čvorova pa su podmornice morale biti na površini kako bi sustigle konvoje, ali to nije bilo izvedivo zbog zračne pratnje. No ni to poboljšanje nije izvedivo jer se konvoj kreće brzinom najsporijeg člana, a kako su konvoje sačinjavali brodovi iz brojnih država, nije bilo moguće pronaći rješenje ubrzanja konvoja. S druge strane, pojačanje zračnog nadzora je bilo i izvedivo, i učinkovito poboljšanje. Saveznici su raspolagali velikim brojem aviona dugog dometa, ali ih je većina bila korištena za bombardiranje Njemačke što se ispostavilo vrlo neučinkovitom upotrebom. Naime, jedan avion tog tipa bi prosječno za vrijeme službe izbacio oko 100 tona bombi, uništio uglavnom civilne objekte i nanio beznačajnu štetu industriji. Takav avion korišten za pratnju konvoja bi „spasio“ barem 6 trgovačkih brodova. Prijedlog je prihvaćen i dio aviona je prenamijenjen za pratnju konvoja, uz predviđene rezultate. [5]

Kao sljedeće područje za analizu je odabrana sama organizacija konvoja, broj trgovačkih i pratećih brodova. Konvoji su se u pravilu sastavljali od situacije do situacije, postojalo je pravilo za određivanje broja pratećih broja ovisno o veličini konvoja i ograničenje maksimalnog broja brodova, a sve ostalo je bilo pitanje slučajnosti. Analiza je pokazala da su oba ograničenja pogrešna. Broj pratećih brodova je određivan po izrazu $n_{pr} = 3 + \frac{N_k}{10}$ što je trebalo osigurati jednaku sigurnost konvoja svih veličina. Teško je pronaći smisao tog izraza jer, iako govori da

za veće konvoje treba biti veća pratnja, račun za određene veličine nije dosljedan s tom tvrdnjom. Ako se sastave tri konvoja po 20 brodova, za svakog će biti raspoređeno 5 brodova u pratnju, tj. 15 ukupno. Ako se svi brodovi pošalju kao jedan konvoj, u pratnju će ići samo 9 brodova. Jasno je da je izraz proizvoljan, a ne temeljen na logičkom zaključku.

Maksimalni dozvoljeni broj brodova u konvoju postavljen na 60 je odbačen nakon pregleda podataka o potopljenim brodovima. Tablica 3. prikazuje dio tih podataka. [4]

Tablica 3. Podaci o napadima podmornica na konvoje

Veličina konvoja	15-24	25-34	35-44	45-54
Broj napada podmornica	8	11	13	7
Prosječan broj brodova potopljenih po napadu	5	6	6	5
Prosječna veličina pratnje	7	7	6	7
Prosječan broj podmornica po napadu	7	5	6	5

Vidljivo je da je prosječan broj potopljenih brodova neovisan o veličini konvoja što znači da je kod većih konvoja manji udio potopljenih. Kako bi se Admiralitetu dokazalo da rezultati nisu slučajni, a čemu su oni bili skloni vjerovati jer se rezultati kose s odlukom koju su oni donijeli, vjerojatnost potapanja broda je razložena na tri dijela. Prvi je vjerojatnost da će konvoj biti opažen, a ona je gotovo jednaka za svaku veličinu jer su i mali konvoji vrlo lako uočljivi. Drugi je vjerojatnost da će podmornica probiti kroz pratnju. Iz podataka je zaključeno da ta vjerojatnost ovisi samo o gustoći pratnje, broju pratećih brodova po kilometru opsega konvoja. Treći je vjerojatnost da će brod biti potopljen nakon što podmornica probije kroz pratnju, a ta vjerojatnost je padala s porastom veličine konvoja jer jednak broj podmornica prosječno potapa jednak broj (a ne udio) brodova. Taj dokaz je bio dovoljan da se donese odluka o povećanju veličine konvoja. Procijenjeno je da bi gubitak brodova u konvojima bio 20%, odnosno 200 brodova manji da je odluka donesena godinu dana ranije. [5]

2.2. Simulacijske metode

Simulacijske metode se koriste kako bi se predvidjelo ponašanje sustava s obzirom na određene okolnosti. U poglavlju 3.3 je prikazan primjer zaključivanja na temelju regresije iz poznatih podataka, a u poglavlju 3.4 je razrađen primjer korištenja tzv. *Markovljevog svojstva* niza stanja za predviđanje budućih stanja.

2.2.1. Simulacija ratnih operacija [10]

Procjenu planiranog oružja i taktika pogodno je izvršiti u stvarni uvjetima korištenja. Međutim, u stvarnosti postoje brojni faktori čiji je utjecaj teško procijeniti i kvantificirati pa nije jednostavno odrediti uzročno posljedične odnose.

Prvi korak je zato odvajanje fizičkih komponenti s definiranim performansama od subjektivnih razmišljanja, planova, namjera i odluka zapovjednika. Dva načina odvajanja ljudskog faktora su korištenje *standardiziranih odluka* i *minimax* metoda teorije igara. U klasičnoj mehanici, poznavanje stanja sustava omogućuje točno predviđanje prelaska u novo stanje, ali u realnim sustavima je puno češći slučaj postojanja određene vjerojatnosti da će od velikog broja sustava istog stanja jednak broj završiti u nekom drugom određenom stanju. Određivanje te vjerojatnosti za odluke zapovjednika je moguće sakupljanjem i analizom velikog broja podatka o ranije donesenim odlukama. To je metoda standardiziranih odluka. Logika minimax metode je da svaki zapovjednik povlači potez s kojim je siguran da će pretrpjeti najmanju moguću štetu pod pretpostavkom da će mu protivnik htjeti nanijeti najveću štetu, ali bez da zna koji će potez povući protivnik. Korištenjem jedne od metoda je zadovoljena pretpostavka da je utjecaj ljudskog faktora eliminiran.

Kad je postignuto da se mogu proučavati samo veličine koje se mogu kvantificirati, moguće je izraziti uvjetnu vjerojatnost promjene stanja $(x, t) \rightarrow (x', t')$:

$$a(t, x; t', x') = P(S\ u\ stanju\ x'u\ t' | S\ u\ stanju\ x\ u\ t) \tag{2}$$

zvanu tranzicijska vjerojatnost, gdje:

- S … promatrani sustav
- X … set svih mogućih stanja
- x, x' … stanja
- t, t' … točke u vremenu

Kad bi sve tranzicijske vjerojatnosti *a* bile poznate, bilo bi moguće predvidjeti vjerojatnosti svih mogućih ishoda iz svih početnih stanja. Problem definiranja sustava je upravo traženje tranzicijskih vjerojatnosti. Elementarne tranzicijske vjerojatnosti, odnosno stope tranzicije su vjerojatnosti dobivene iz kvantifikacije osnovnih elemenata sustava.

Glavna pretpostavka u postupku određivanja tranzicijskih vjerojatnosti je da niz stanja posjeduje Markovljevo svojstvo, a više o tome govori poglavlje 3.4. Simulacijski model može biti *deterministički*, gdje iz svako stanja slijedi samo jedno, točno određeno stanje, ili *stohastički*, gdje se sljedeće stanje odabire slučajno s unaprijed određenom vjerojatnošću. Druga vrsta je zapravo osnova široko raširenog Monte Carlo simulacijskog modela.

2.2.2. Planiranje ljudskih potencijala

Planiranje ljudskih potencijala (engl. manpower planning) je važna aktivnost u vojnom sustavu zbog samih proporcija i nemogućnosti naglih promjena. Osnovni problem se može definirati kao utvrđivanje broja osoblja i njihovih vještina kako bi se što bolje zadovoljile kratkoročne ili dugoročne operativne potrebe. [11] Za svakog pojedinca je moguće opisati stanje s parametrima poput staža, čina, dobi i kvalifikacije te se, uz utvrđene tranzicijske vjerojatnosti, može predvidjeti sastav jedinica nakon određenog vremena. Tema je detaljnije razrađena u poglavlju 3.4.2

2.3. Probabilističke metode

Probabilističke metode su uglavnom matematički izrazi koji povezuju neke geometrijske karakteristike poput oblika i površine s parametrima oružja kako bi se dobile vjerojatnosti ishoda nakon pogodaka.

2.3.1. Vjerojatnost uništavanja mete [4]

Broj potrebnih metaka, projektila, bombi i sl. za uništavanje mete ovisi o dvije veličine, destruktivnoj moći, tj. vjerojatnosti da će pogodak uništiti metu i preciznosti, tj. vjerojatnosti da će meta biti pogođena. Dodatno, ako se koristi više komada oružja, raspored ispaljivanja odnosno pucanja ima utjecaj na ukupno razaranje.

Najjednostavniji slučaj je onaj gdje je prvi pogodak dovoljan za uništavanje mete. Takvi slučajevi su rijetkost pa se u stvarnosti obično mora definirati vjerojatnost uništenja mete s obzirom na dio koji je pogođen. Tako se dolazi do izraza:

$$L = \iint P_d(x,y)dxdy \tag{3}$$

gdje L označava „smrtonosnu“ površinu, a $P_d(x,y)$ vjerojatnost uništenja mete pogotkom u točku (x,y).

Kod nasumičnog bombardiranja velikih površina, ta „smrtonosna“ površina je dovoljan parametar za određivanje ukupnog razaranja. Vjerojatnost da će bomba pasti u neku određenu točku je jednostavno $\frac{dxdy}{A}$, gdje je A ukupna površina zahvaćena napadom, a iz toga slijedi vjerojatnost razaranja određene mete:

$$P_D = \iint P_d(x,y)\frac{dxdy}{A}, \tag{4}$$

odnosno:

$$P_D = \frac{L}{A}. \tag{5}$$

Uz pretpostavku da prethodni pogodak ne mijenja vjerojatnost uspjeha idućeg te da je L mala veličina u odnosu na A, može se iz izraza slaganje neovisnih vjerojatnosti dobiti izraz za dobivanje vjerojatnosti razaranja s n bačenih bombi:

$$P_D{}' = 1 - Q_D' = 1 - (1 - P_D)^n = 1 - \left(1 - \frac{L}{A}\right)^n = 1 - e^{-n\frac{L}{A}} \tag{6}$$

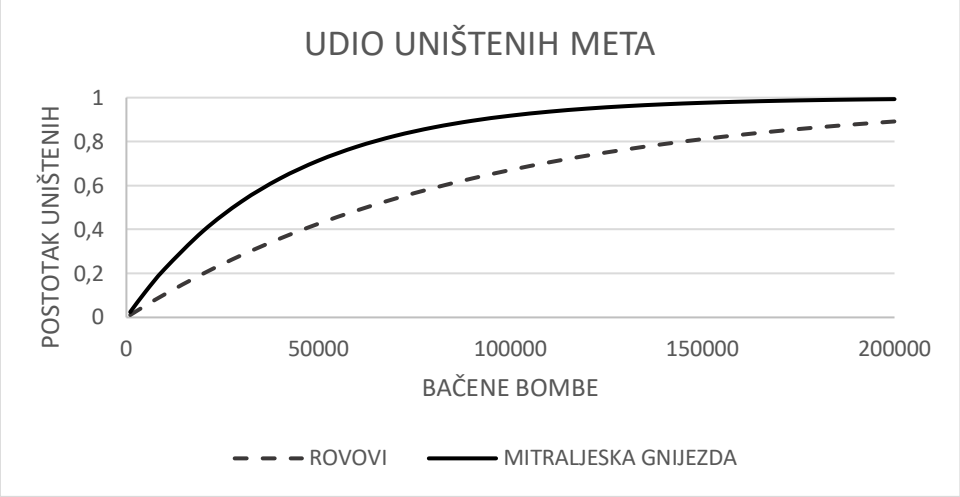
Na površini s većim brojem meta se vrijednost dobivena gornjim izrazom može smatrati i udjelom uništenih meta.

Tablica 4. prikazuje stanje na površini od 3,4 kvadratnih kilometara napadnutoj s 1000 bombi određenog tipa.

Tablica 4. Bombardiranje velike površine

	Broj	L, m^2	L/A
Mitraljeska gnijezda	100	37	1/90000
Vojnici u rovovima	100	83	1/40000

Slika 3. je prikaz dobivenih vrijednosti iz izraza *(6)* uz korištenje koeficijenata iz gornje tablice.



Slika 3. Udio uništenih meta

Izloženi izrazi su najjednostavniji oblik računanja vjerojatnosti pogotka i uništenja mete, ali i za kompleksnije račune se koristi ista logika. Slika 3. pokazuje kako i ovakav jednostavan račun daje realistične vrijednosti jer je poznato da je nasumično bombardiranje velikih površina jako neučinkovit napad, a korištenje 200 000 bombi za uništavanje 90 od 100 mitraljeskih gnijezda na to ukazuje.

2.3.2. Vjerojatnost preživljavanja pogotka [12]

Promatrajući rupe od metaka na zrakoplovima koji su se iz operacija vraćali u bazu, zapažena je sljedeća raspodjela:

Tablica 5. Raspored pogodaka na zrakoplovima

Dio	Udio u površini, %	Broj zapaženih pogodaka
Motori	26,9	19
Trup	34,6	39
Sustav za gorivo	15,4	18
Sve ostalo	23,1	26

Od strane zapovjedništva je bilo primijećeno gdje je najveći broj pogodaka te je donesen zaključak kako baš te dijelove, iz razloga što su očito češće na meti pogodaka, treba ojačati.

Greška u tom zaključku je što je gotovo nemoguće da je neprijatelj ciljao baš u te dijelove, već je kod pogodaka s tako velikih udaljenosti za očekivati slučajnu raspodjelu. Abraham Wald je u skladu s tim zaključio suprotno, budući da je očekivana slučajna raspodjela rupa, onaj dio aviona koji ima rupe po motorima se nije vratio u bazu, tj. pogoci u motor su fatalni. Slika 4. i Slika 5. ilustriraju situaciju.

Slika 4. Slučajna raspodjela pogodaka

Slika 5. Stvarna raspodjela pogodaka

Naravno, bez pregleda zrakoplova koji se nisu vratili, to je i dalje samo pretpostavka jer se ne može utvrditi da je na njima zaista takva raspodjela kakvu je Wald predvidio, već je on to morao dokazati kako bi se prijedlog pojačavanja oklopa zrakoplova prihvatio i primijenio na čitavu flotu i buduće proizvedene zrakoplove. To je i dokazao kroz 8 memoranduma [13] čiji je sadržaj previše kompleksan i ne ulazi u opseg rada pa neće biti iznesen.

3. PRIMJERI NAJČEŠĆE KORIŠTENIH METODA

Ovo poglavlje sadrži detaljnije razrađene neke od metoda s kojima se najviše može susresti u obrađenoj literaturi. Prikazana je teorijska podloga te primjeri sa stvarnim podacima i prikaz dobivanja zaključka na temelju njih.

3.1. Hi-kvadrat test

Hi-kvadrat test je metoda testiranja razlike skupa podataka od neke očekivane vrijednosti koja se koristi u slučaju da su podaci izraženi čestinom (frekvencijom) pojavljivanja ili ako im distribucija značajno odstupa od normalne. [14]

Formula za hi-kvadrat jednaka je u gotovo svim slučajevima i glasi:

$$\chi^2 = \Sigma \frac{(f_o - f_t)^2}{f_t} \tag{7}$$

gdje su f_o opažene, a f_t teoretske frekvencije, tj. frekvencije očekivane pod nekom hipotezom H_0.

Vrijednost χ^2 raste s porastom razlike između opaženih i očekivanih frekvencija. Što je χ^2 bliži nuli, to je vjerojatnije da nultu hipotezu treba prihvatiti. Parametar χ^2 prati hi-kvadrat razdiobu s *k* stupnjeva slobode, a očitava se iz tablica.

Ako je izračunati parametar χ^2 manji od vrijednosti χ^2 očitane iz tablice hi-kvadrat distribucije za *k* stupnjeva slobode i granicu značajnosti α, prihvaća se nultu hipotezu H_o koja glasi:

$$H_0: \pi_1 = \pi_2, \tag{8}$$

odnosno zaključuje se da ne postoji razlika u opaženim frekvencijama u populacijama iz kojih dolaze uzorci. Ako je izračunata vjerojatnost veća od χ^2 iz tablice, odbija se nulta hipoteza H_o i prihvaća alternativna hipoteza H_1,

$$H_1: \pi_1 \neq \pi_2, \tag{9}$$

odnosno zaključuje se da postoji razlika između opaženih frekvencija u populacijama. [15]

Hi-kvadrat test se upotrebljava kad imamo:

- frekvenciju jednog uzorka i želimo utvrditi odstupaju li vrijednosti od očekivanih,

- frekvencije dvaju ili više nezavisnih uzoraka i želimo utvrditi razlikuju li se opažene vrijednosti, odnosno pripadaju li istoj populaciji,
- frekvencije dvaju dihotomnih zavisnih uzoraka i želimo utvrditi postoji li razlika među njima, odnosno je li došlo do promjene. [14]

3.1.1. Hi-kvadrat test za dva nezavisna uzorka [14]

Kod ovog oblika hi-kvadrat testa, prvo se sastavlja tzv. tablica kontingencije ($r \times s$) kod koje svaka os predstavlja jednu varijablu, a u nju se unose frekvencije pojavljivanja svake kombinacije varijabli. Nakon toga je potrebno popuniti tablicu očekivanih frekvencija, a taj postupak će biti objašnjen kasnije u primjeru. Broj stupnjeva slobode se računa prema izrazu:

$$k = (r - 1) \times (s - 1). \tag{10}$$

Postoje ograničenja po pitanju minimalnih iznosa očekivanih frekvencija. Za tablice 2 × 2, test se smije koristiti uvijek ako je ukupan zbroj frekvencija N veći od 40, a za 40 > N > 20, niti jedna frekvencija ne smije biti manja od 5. Za tablice s više od jednog stupnja slobode, hi-kvadrat se smije koristiti ako je manje od 20% frekvencija manje od 5 i nijedna nije manja od 1. Ako uvjeti nisu postignuti, potrebno je spojiti susjedna polja (samo ako to logički ima smisla, ne mogu se spajati oprečne tvrdnje) kako bi se povećala očekivana frekvencija. [14]

3.1.2. Analiza učinkovitosti protuzračne opreme na trgovačkim brodovima [4]

Na samom početku Drugog svjetskog rata, britanski Admiralitet je donio odluku o naoružavanju 5500 trgovačkih brodova kako bi ih se što uspješnije obranilo od neprijateljskih zračnih napada. Nije puno vremena moralo proći da se pokaže kako neiskusna i nevješta posada opremljena malim brojem komada oružja vrlo teško može pogoditi i oboriti napadača (stopa uspjeha oko 4%). S obzirom na skupoću projekta i potrebu za protuzračnom obranom na mnogo drugih mjesta, oporba je postajala glasnija, ali kako je oružje na brodu davalo posadi barem privid sigurnosti, došlo je do potrebe za analizom podataka kako bi se konačno dobio odgovor na pitanje treba li nastaviti s projektom ili odustati od njega. Tablica 6. u nastavku prikazuje prikupljene podatke.

Nulta hipoteza glasi:

$$H_0: \pi_1 = \pi_2, \tag{11}$$

i govori da ne postoji razlika među populacijama koje koriste i onima koje ne koriste protuzračnu opremu, a alternativna hipoteza govori da protiv populaciji koja je koristila protuzračnu opremu neprijatelj ima manje uspjeha i glasi:

$$H_1: \pi_1 > \pi_2. \tag{12}$$

Tablica 6. Šteta pretrpljena pri bombardiranju

	Obrana pucala	Obrana nije pucala
Broj bačenih bombi	632	304
Broj pogodaka	50	39
Postotak pogodaka	8	13
Broj napadnutih brodova	155	71
Broj potopljenih	16	18
Postotak potopljenih	10	25

Sam pogled na tablicu jasno pokazuje da su neprijateljski napadi manje uspješni u slučajevima kad obrana broda uzvraća, ali kako bi se provjerilo je li razlika između dva slučaja značajna, izračunati će se hi-kvadrat parametar, prvo za pogotke, a onda i za potapanja.

3.1.2.1. Uspješnost neprijateljskih napada – broj pogodaka

Prvo se popunjava tablica opaženih frekvencija, odnosno prepisuju prva dva retka (Tablica 6.) i upisuju zbrojevi svih redaka i stupaca kao što je prikazano (Tablica 7.).

Tablica 7. Metoda formiranja tablice kontingencije

	Obrana pucala	Obrana nije pucala	Ukupno
Broj bačenih bombi	a	b	a + b
Broj pogodaka	c	d	c + d
Ukupno	a + c	b + d	a + b + c + d = N

Tablica 8. Tablica kontingencije - opažene frekvencije pogodaka

	Obrana pucala	Obrana nije pucala	Ukupno
Broj bačenih bombi	632	304	936
Broj pogodaka	50	39	89
Ukupno	682	343	1025

Zatim se izrađuje tablica očekivanih frekvencija, i to tako da se u polje unosi umnožak zbroja stupca i zbroja retka kojem pripada podijeljen ukupnim brojem frekvencija.

$$a_{tij} = \frac{\sum_{i=1}^{s} a_{oij} \cdot \sum_{j=1}^{r} a_{oij}}{\sum_{i=1}^{s} \sum_{i=1}^{r} a_{oij}} \tag{13}$$

Tablica 9. Tablica kontingencije - očekivane frekvencije pogodaka

	Obrana pucala	Obrana nije pucala	Ukupno
Broj bačenih bombi	622,8	313,2	936
Broj pogodaka	59,2	29,8	89
Ukupno	682	343	1025

Za dva (ili više) uzorka, hi-kvadrat parametar se računa po izrazu:

$$\chi^2 = \sum_{i=1}^{s}\sum_{j=1}^{r}\frac{\left(f_{oij} - f_{tij}\right)^2}{f_{tij}}. \quad (14)$$

Računski hi-kvadrat iznosi:

$$\chi^2 = 4{,}676$$

Prema izrazu *(10)*, $k = 1$, a očitana odgovarajuća vrijednost iz tablice A (u privitku) za $k = 1$ i granicu značajnosti $\alpha = 0{,}05$ iznosi:

$$\chi_t^2 = 3{,}48.$$

$$\chi^2 \geq \chi_t^2 \rightarrow H_0 \; se \; odbacuje$$

Ne može se ustvrditi da nema značajne razlika između dvaju uzoraka jer je vjerojatnost slučajnog pojavljivanja toliko visokog χ^2 manja od granice značajnosti 5%. Za promatrani slučaj to dovodi do zaključka da i sam čin pucanja na protivničke zrakoplove, iako ih se vrlo rijetko pogađa, utječe na njihovu preciznost i sposobnost pogađanja mete.

3.1.2.2. Uspješnost neprijateljskih napada – broj potapanja

Kako bi se dodatno provjerila učinkovitost obrambenog sustava, postupak će se ponoviti i za frekvencije potopljenih brojeva.

Tablica 10. Tablica kontingencije - opažene frekvencije potapanja

	Obrana pucala	Obrana nije pucala	Ukupno
Broj bačenih bombi	155	71	226
Broj pogodaka	16	18	34
Ukupno	171	89	260

Tablica 11. Tablica kontingencije - očekivane frekvencije pogodaka

	Obrana pucala	Obrana nije pucala	Ukupno
Broj bačenih bombi	147,8	78,2	226
Broj pogodaka	23,2	10,8	34
Ukupno	171	89	260

I u ovom, i u prethodnom primjeru se može primijetiti da je zbroj opaženih i očekivanih frekvencija jednak, a to je važan uvjet za ispravno provođenje testa.

Za matrice 2 × 2, odnosno mali broj polja, način računanja prikazan u prethodnom primjeru nije pretjerano zahtjevan, ali s porastom polja, postoji veća mogućnost računske pogreške. Zato će u nastavku biti prikazan tablični račun hi-kvadrat parametra.

Tablica 12. Računska tablica za hi-kvadrat

f_o	f_t	$f_o - f_t$	$\frac{(f_o - f_t)^2}{f_t}$
155	147,8	6,4	0,277
71	78,2	-6,4	0,524
16	23,2	-6,4	1,766
18	10,8	6,4	3,793
		$\chi^2 = \sum \frac{(f_o - f_t)^2}{f_t} =$	6,360

Suma zadnjeg stupca je upravo iznos hi-kvadrat parametra. Tablični hi-kvadrat je jednak jer je i broj stupnjeva slobode je jednak te:

$$\chi^2 \geq \chi_t^2 \rightarrow H_0 \; se \; odbacuje.$$

Vjerojatnost značajnosti razlike dvaju uzoraka je još veća, gotovo 99%. Osim jednakog zaključka kao i u prošlom primjeru, da obrambeni sustav djeluje na sposobnost pogađanja broda, vidi se da su i napadi koji uspiju pogoditi metu manje učinkoviti, a to svakako potvrđuje učinkovitost cijelog projekta.

Zaključak analize je da trgovački brodovi koji koriste protuzračnu opremu imaju znatno veću vjerojatnost preživljavanja nego oni koji ju ne koriste iako ne uspijevaju izravno oštetiti napadača te da s projektom ugradnje opreme treba nastaviti.

3.2. Fisherov egzaktni test [15]

Fisherov egzaktni test je neparametarska metoda koja se koristi za analizu diskretnih varijabli, najčešće kod uzoraka premalih za propisno provođenje hi-kvadrat testa, ali je valjan i za sve veličine uzoraka. Kod velikih uzoraka se ne koristi i zbog opsežnog matematičkog računa.

Test počiva na hipergeometrijskoj raspodjeli vjerojatnosti dobivanja točno određenog seta opaženih frekvencija iz tablice kontingencije. Hipergeometrijska raspodjela je raspodjela vjerojatnosti koja opisuje vjerojatnost *k* odabira varijabli određenih svojstava iz *n* pokušaja, iz populacije veličine *N* s točno *K* varijabli traženih svojstava, a opisuje ju funkcija distribucije vjerojatnosti:

$$p(k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}. \tag{15}$$

Važno je pripaziti da se ne izračuna vjerojatnost samo onog stanja zapisanog u tablici kontingencije, već i vjerojatnost svih još ekstremnijih slučajeva zajedno.

Ako je izračunata vjerojatnost manja od pola odabrane granice značajnosti, prihvaća se nulta hipoteza H_o,

$$H_0: \pi_1 = \pi_2, \tag{16}$$

odnosno omjer opaženih frekvencija u populacijama iz kojih dolaze uzorci nije različit. Ako je izračunata vjerojatnost veća od $\alpha/2$, odbija se nulta hipoteza H_o i prihvaća alternativna hipoteza H_1,

$$H_1: \pi_1 \neq \pi_2, \tag{17}$$

odnosno zaključuje se da postoji razlika između opaženih frekvencija u populacijama. Moguće je nultu hipotezu postaviti i usmjereno tako da se testira jesu li vrijednosti u jednoj populaciji veće ili manje nego u drugoj, samo što se onda dobiveni *p* uspoređuje s α.

3.2.1. Analiza učinkovitosti obrane od torpeda [4]

Slično kao u prošlom primjeru s protuzračnom obranom, postojali su pokušaji zaštite trgovačkih brodova od njemačkih podmornica (tzv. *U-Boat*) i njihovih napada torpedom. Jedan od prvih pokušaja je bilo opremanje brodova mrežama razapetim na određenoj udaljenosti od trupa. Stupanj zaštite se može procijeniti kao uvjetna vjerojatnost zaustavljanja torpeda. Mreže su bile u stanju zaustaviti oko 85% električnih torpeda (ET) i oko 20% torpeda pogonjenih zrakom (ZT), a s obzirom njemačke podmornice su bile opremljene 60% s torpedima prvog tipa, a 40% drugog pa slijedi:

$$p' = P(zET) \cdot P(pET) + P(zZT) \cdot P(pZT) = 0{,}59 \tag{18}$$

gdje p' predstavlja vjerojatnost da će mreža zaustaviti torpedo koji pogodi unutar njezine površine, *P(zET)* je vjerojatnost zaustavljanja električnog torpeda, *P(zZT)* vjerojatnost zaustavljanja torpeda pogonjenog zrakom, *P(pET)* vjerojatnost da će podmornica ispaliti električni torpedo, a *P(pZT)* vjerojatnost da će ispaliti torpedo pogonjen zrakom. Kad se još u obzir uzme da mreže pokrivaju oko 3/4 površine trupa, konačno se dobiva:

$$p = p' \cdot P(tm) = 0{,}44 \tag{19}$$

gdje *p* predstavlja vjerojatnost da će mreža zaustaviti torpedo ispaljen unutar površine broda, a *P(tm)* vjerojatnost da će torpedo biti ispaljen unutar površine mreže.

Na prvi pogled se čini kako je vjerojatnost obrane od torpeda korištenjem mreže dovoljno visoka kako bi opravdala visoke troškove ugradnje i održavanja, ali i troškove nastale većom potrošnjom goriva i dužim trajanjem puta. Ta tvrdnja će biti provjerena Fisherovim testom.

Prikupljeni podaci o rezultatima napada na brodove s ugrađenim mrežama prikazuje Tablica 13.:

Tablica 13. Tablica kontingencije - opažene frekvencije

	Mreže korištene	Mreže nisu korištene	Ukupno
Broj oštećenih brodova	7	12	19
Broj neoštećenih brodova	3	0	3
Ukupno	10	12	22

Želimo provjeriti koliko je vjerojatno da je veći omjer preživjelih brodova rezultat slučajnosti, a ne stvarnog utjecaja mreže na sigurnost. Nulta hipoteza glasi:

$$H_0: \pi_1 = \pi_2, \tag{20}$$

i govori da ne postoji razlika među populacijama, a alternativna hipoteza govori da je u populaciji koja je koristila mrežu više preživjelih i glasi:

$$H_1: \pi_1 > \pi_2. \tag{21}$$

Vjerojatnost da iz skupa oštećenih brojeva 7 budu oni kod kojih su mreže korištene iznosi: $\binom{19}{7}$, od svih neoštećenih, 3 budu on kod kojih su mreže korištene iznosi $\binom{3}{3}$, a da se od svih *n* slučajeva odaberu samo oni kod kojih su mreže korištene iznosi $\binom{22}{10}$. Vidljivo je da se može se zapisati opći izraz za računanje vjerojatnosti dobivanja seta frekvencija za 2 × 2 tablicu kontingencija, i to korištenjem označavanja prikazanog (Tablica 7.) u prošlom primjeru:

$$p = \frac{\binom{a+c}{a}\binom{b+d}{b}}{\binom{N}{a+b}}. \tag{22}$$

Za pojednostavljivanje računa, u izrazu *(22)* se može binomne koeficijente prikazati pomoću faktorijela korištenjem izraza *(23)*.

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!} \tag{23}$$

Tako dobiveni izraz glasi:

$$p = \frac{(a+b)!\,(c+d)!\,(a+c)!\,(b+d)!}{a!\,b!\,c!\,d!\,N!}, \tag{24}$$

a njime izračunata vjerojatnost:

$$p = 0{,}0779 \geq 0{,}05 \rightarrow H_0\ se\ odbacuje$$

dovodi do zaključka da se odbacuje nulta hipoteza i prihvaća alternativna, odnosno postoji značajan učinak mreža na vjerojatnost preživljavanja. Ipak, unatoč rezultatima analize, zaštitne mreže su odbačene kao metoda obrane zbog različitih ekonomskih razloga.

3.3. Regresijska analiza

Regresijska analiza je skup statističkih postupaka za utvrđivanje odnosa među varijablama, tj. za utvrđivanje načina promjene zavisne varijable s obzirom na promjenu nezavisne te postoji li uopće veza i kolika je njena jakost. [16] Postoje linearna i nelinearna regresija, ovisno o matematičkom modelu zavisnosti. Kod linearne, zavisna varijabla ovisi o sumi nezavisnih varijabli.

Opća jednadžba linearne regresije [16]:

$$y = f(x_1, x_2, \ldots, x_i) + \varepsilon \tag{25}$$

gdje je varijabla y zavisna varijabla, varijabla čiju vrijednost predviđamo, a varijable x_1, x_2, ... , x_i nezavisne (prediktori). $f\ (x_1, x_2, \ldots, x_i)$ izražava zavisnost, a ε je slučajna promjenjiva varijabla, greška predviđanja.

3.3.1. Višestruka regresijska analiza

Slučaj kod kojeg se utvrđuje veza između samo dviju varijabli se naziva jednostruka regresija, a kod postojanja više nezavisnih varijabli se koristi višestruka regresija.

Opća jednadžba višestruke linearne regresije [16]:

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_i x_i + \varepsilon \tag{26}$$

gdje su β_0, β_1, ... , β_i nepoznati parametri (težinski faktori) koji se određuju pomoću metode najmanjih kvadrata.

3.3.2. Logistička regresija

Dok se kod linearne regresije predviđa vrijednost kontinuirane varijable, kod logističke je zavisna varijabla diskretna (kategorička), poprima samo određene vrijednosti. U nastavku će biti objašnjena samo binarna logistička regresija, kod koje zavisna varijabla poprima samo dvije vrijednosti (npr. da/ne, muško/žensko, poraz/pobjeda).

Logistička funkcija izražena je kao:

$$p\colon \mathbb{R} \rightarrow (0,1), \quad p(x) = \frac{1}{1+e^{-x}} \tag{27}$$

gdje je *p(x)* vjerojatnost ishoda.

3.3.3. Primjer - Mjera borbene uspješnosti [17]

N. K. Jaiswal i B. S. Nagabhushana su proveli analizu baze podataka CDB90 [18] koja sadrži detalje 660 bitaka vođenih između 1600. i 1973. s ciljem dobivanja veze između određenih karakteristika (nezavisnih varijabli) sukobljenih strana i njihovog uspjeha (zavisnih varijabli) u bitkama te predviđanja koja je sukobljena strana u prednosti u određenom trenutku.

3.3.3.1. Regresija

Karakteristike koje su uzete u obzir i način njihovog računanja:

Tablica 14. Nezavisne varijable

OZNAKA	KARAKTERISTIKA	RAČUN
x_1	udio žrtava napadača	CASA / STRA
x_2	udio žrtava napadnute strane	CASD / STRD
x_3	taktička učinkovitost	KMDA / WOFA1
x_4	relativna borbena snaga	(STRA - CASA) / (STRD - CASD)
x_5	stopa ubijanja napadača	CASD / STRA
x_6	stopa ubijanja napadnute strane	CASA / STRD
x_7	omjer razmjene žrtava	CASA / CASD

Kao zavisna varijabla *y* je uzeta razlika između uspjeha misija obaju strana na ljestvici 1-10, ACHA - ACHD.

Tablica 15. Oznake varijabli

CASA / CASD	broj žrtava napadača / napadnute strane
STRA / STRD	veličina napadačevih postrojbi / postrojbi napadnute strane
KMDA	udaljenost za koju je napadač prešao preko obrambene linije
WOFA1	širina napadačevog fronta na početku bitke
ACHA / ACHD	uspjeh napadačevih misija / misija napadnute strane

Cilj regresije je dobivanje koeficijenata $\beta_0, \ldots, \beta_7$ tako da izraz:

$$y = \beta_0 + \sum_{i=1}^{7} \beta_i x_i \tag{28}$$

što bolje opisuje skup podataka.

Tablica 16. prikazuje koeficijente dobivene pomoću programskog paketa SPSS. Prije regresije, uklonjene su bitke kod kojih nisu poznate sve potrebne karakteristike te su od preostalih izdvojene dvije grupe, *Grupa 1* kod koje nije bilo utvrđene obrane i *Grupa 2* kod koje je bio prisutan neki oblik utvrđivanja. To je učinjeno iz razloga što utvrđivanje znatno mijenja karakteristike borbe i utjecaj nezavisnih varijabli na zavisnu.

Tablica 16. Regresijski koeficijenti

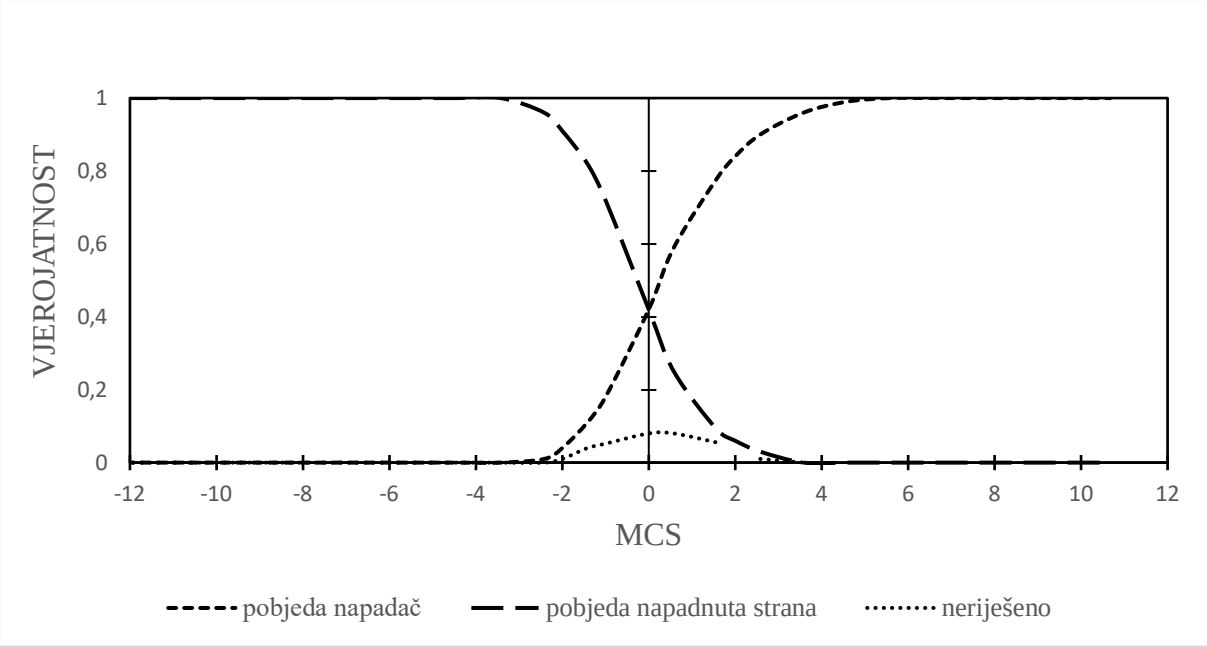
		REGRESIJSKI KOEFICIJENTI								R^2
GRUPA	BROJ BITAKA	β_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7	
1	452	-0,43	-3,39	10,73	0,63	0,00	0,91	-4,00	-0,05	0,49
2	89	-0,22	-1,68	2,50	0,79	0,17	3,46	-2,40	-0,15	0,58

R^2 se naziva koeficijentom determinacije i predstavlja mjeru u kojoj je promjena zavisne varijable objašnjena promjenom nezavisne. [19] Na primjeru Grupe 2 to znači da jednadžba regresije pokriva 58% varijance vrijednosti zavisne varijable.

3.3.3.2. Mjera borbenog uspjeha

Provedena analiza je dala koeficijente koji uvršteni u izraz *(27)* daju jednadžbu regresije. Uvrštavanjem iznosa zavisnih varijabli u tu jednadžbu se dobivaju iznosi zavisne varijable *y*, u nastavku definirane kao mjera borbenog uspjeha (MCS).

Vjerojatnost ishoda dobije se izračunavanjem omjera broja određenog ishoda i ukupnog broja bitaka koje zadovoljavaju kriterij grupe. Podjelom MCS na intervale te računanjem vjerojatnosti za svaki interval zasebno može se grafički prikazati raspodjela vjerojatnosti.



Slika 6. Raspodjela vjerojatnosti ishoda za Grupu 1

Iz raspodjele vjerojatnosti, vidljivo je da u rasponu MCS od -1,95 (y_L) do 3,91 (y_R) i jedna i druga strana imaju vjerojatnost za pobjedu i taj raspon se naziva područjem nesigurnosti. MCS veći od y_R ukazuje na pobjedu napadača, a manji od y_L na pobjedu napadnute strane. Također je vidljivo da su sve bitke iz baze podataka obuhvaćene intervalom [-12,11;10,68]. Središnja granica y_0 označava MCS za kojeg nijedna strana nema prednost, desno od nje je u prednosti napadač, a lijevo napadnuta strana. Za Grupu 1, ta granica je 0,42. Matrica nesigurnosti uspoređuje stvarne ishode borbi s ishodima predviđenima pomoću granice y_0:

Tablica 17. Matrica nesigurnosti

GRUPA 1		STVARNI ISHOD	
		NAPADAČ	NAPADNUTA STRANA
PREDVIĐENI ISHOD	NAPADAČ	245	51
	NAPADNUTA STRANA	8	123

Točnost korištenja y_0 za predviđanje ishoda je definirana kao omjer zbroja točnih predviđanja i ukupnog broja događaja i prikazana je, kao i granice područja nesigurnosti i krajnje granice MCS (Tablica 18.):

Tablica 18. Točnost predviđanja s y_0

GRUPA	y_0	TOČNOST	GRANICE PODRUČJA NESIGURNOSTI		GRANICE MCS ZA SVE BITKE	
			y_L	y_R	min	max
1	0,42	81,46	-1,95	3,91	-12,11	10,86
2	0,23	86,50	-0,07	1,80	-7,43	8,86

U ratnim simulacijama i igrama, uvrštavanjem vrijednosti zavisnih varijabli u izraz *(27)* za određeno vrijeme *t* se može dobiti iznos MCS za to vrijeme te ga usporediti s vrijednostima y_0, y_L, y_R kako bi se odlučilo koja strana je u prednosti u odabranom trenutku.

3.3.3.3. Predviđanje vjerojatnosti ishoda

Kako bi se dobila vjerojatnost pobjede određene strane (ili neriješenog rezultata), provedena je logistička regresija s vjerojatnošću ishoda kao zavisnom varijablom *p*, a središnjim vrijednostima intervala MCS kao nezavisnom varijablom *m*.

Logistička funkcija korištena za predviđanje vjerojatnosti pobjede napadača:

$$p(w_1) = \frac{e^{k_1+k_2 m}}{1+e^{k_1+k_2 m}} \tag{29}$$

gdje su k_1 i k_2 težinski faktori.

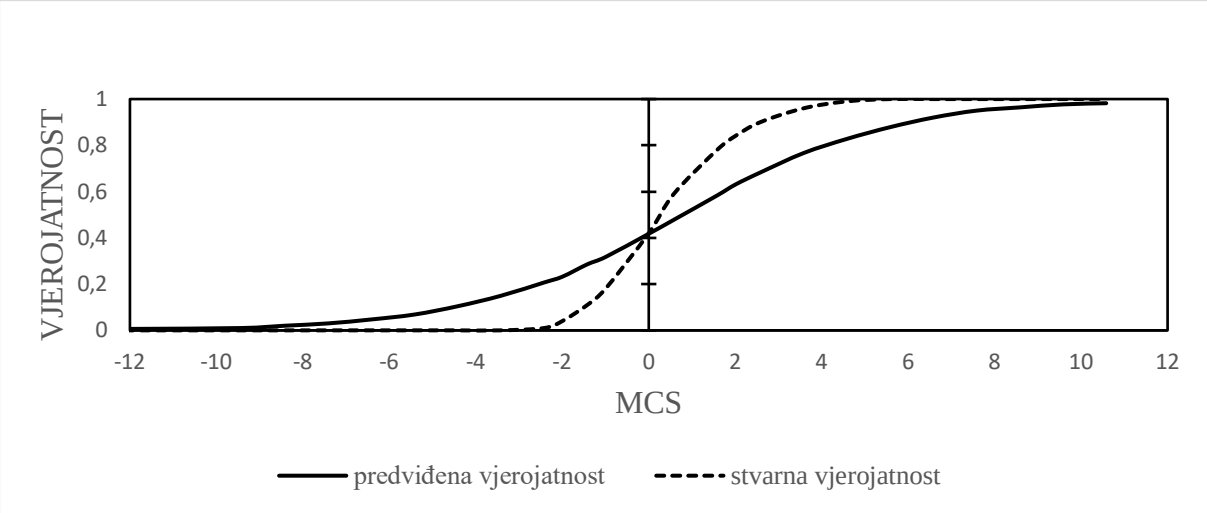
Tablica 19. prikazuje rezultate regresijske analize za Grupu 1 i s pobjedom napadača kao zavisnom varijablom.

Tablica 19. Rezultati logističke regresije

GRUPA	k_1	k_2	χ^2
1	-0,33	0,42	1.11

Kod logističke regresije ne postoji mjera odstupanja poput R^2 kod linearne, ali zato se može koristiti vrijednost χ^2 , dobivena Hosmer-Lemeshow testom, koja za ovaj slučaj ukazuje na zadovoljavajuću prikladnost modela.

Za predviđanje pobjede napadača, vrijednosti zavisnih varijabli x_i se uvrštavaju u izraz *(27)*, a dobiveni *y* se uvrštava u izraz *(28)*. Slika 7. prikazuje predviđenu raspodjelu vjerojatnosti pobjede napadača te stvarnu raspodjelu iz povijesnih rezultata:



Slika 7. Usporedba predviđene i stvarne vjerojatnosti ishoda

Vidljivo je odstupanje stvarnih ishoda od predviđenih, pogotovo u području neizvjesnosti, ali model ipak donekle prati stvarnu raspodjelu pa se može pretpostaviti da bi obradom novijih podataka došlo do još veće točnosti.

3.4. Markovljev lanac

Markovljev lanac je stohastički proces koji zadovoljava Markovljevo svojstvo, „svojstvo zaboravljivosti“. To znači da vjerojatnost svakog događaja ovisi samo o stanju dobiveno u prethodnom događaju, a ne o stanjima u kojima je bio u događajima prije toga. [20]

Korištenjem uvjetnih vjerojatnosti i zavisnih varijabli, definicija je sljedeća. Stohastički lanac, niz diskretnih varijabli X_0, X_1, ..., za čija sva stanja s_0, s_1, ..., $s_n \in S$ vrijedi:

$$P(X_n = s_n | X_{n-1} = s_{n-1}, \dots, X_0 = s_0) = P(X_n = s_n | X_{n-1} = s_{n-1}), \qquad (30)$$

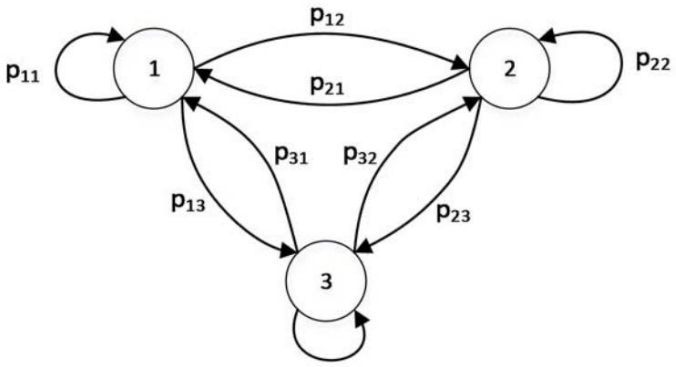
naziva se Markovljevim lancem. [20]

Vjerojatnosti prelaska u novo stanje se nazivaju tranzicijskim vjerojatnostima, a prikazuju se tranzicijskom matricom. Prikazana je tranzicijska matrica P_M za lanac s tri stanja:

$$P_M = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix}, \tag{31}$$

pri čemu je

$$\sum_{j=1}^{3} p_{ij} = 1, i = 1,2,3. \tag{32}$$



Slika 8. Grafički prikaz Markovljeva lanca s tri stanja

3.4.1. Apsorpcijski Markovljev lanac

Apsorpcijski Markovljev lanac je posebna vrsta Markovljevog lanca kod kojeg postoji jedno stanje kojeg je nemoguće napustiti i kojeg svako stanje može doseći nakon određenog broja koraka, tj. uz pozitivnu vjerojatnost.

Stanja koja je moguće napustiti se zovu tranzijentna, a stanja koja nije moguće napustiti se zovu apsorpcijska i za njih vrijedi:

$$P(X_t = s|X_{t-1}) = 1. \tag{33}$$

Ako je u lancu *t* tranzijentnih stanja i *r* apsorpcijskih, vrijedi ovakav zapis kanonske forme tranzicijske matrice:

$$P_M = \begin{bmatrix} Q & R \\ 0 & I \end{bmatrix}, \tag{34}$$

gdje je Q tranzijentna matrica dimenzija $t \times t$, R apsorpcijska matrica dimenzija $t \times r$, 0 nul-matrica dimenzija $r \times t$, a I jedinična matrica dimenzija $r \times r$.

Za apsorpcijski Markovljev lanac, matrica $I - Q$ ima inverz N_M, a

$$N_M = I + Q + Q^2 + \cdots \tag{35}$$

i naziva se temeljnom matricom. Sadržaj matrice N_M, $n_{ij,}$ govori koliko puta će događaj završiti u j ako započne u i.

Umnožak temeljne matrice i jedinične matrice I daje matricu T kod koje t_{ij} pokazuje koliko vremena će proteći do apsorpcije za svako početno stanje i.

Umnožak temeljne matrice i apsorpcijske matrice R daje matricu B koje b_{ij} pokazuje kolika je vjerojatnost da će lanac koji krene iz stanja i biti apsorbiran u stanje j.

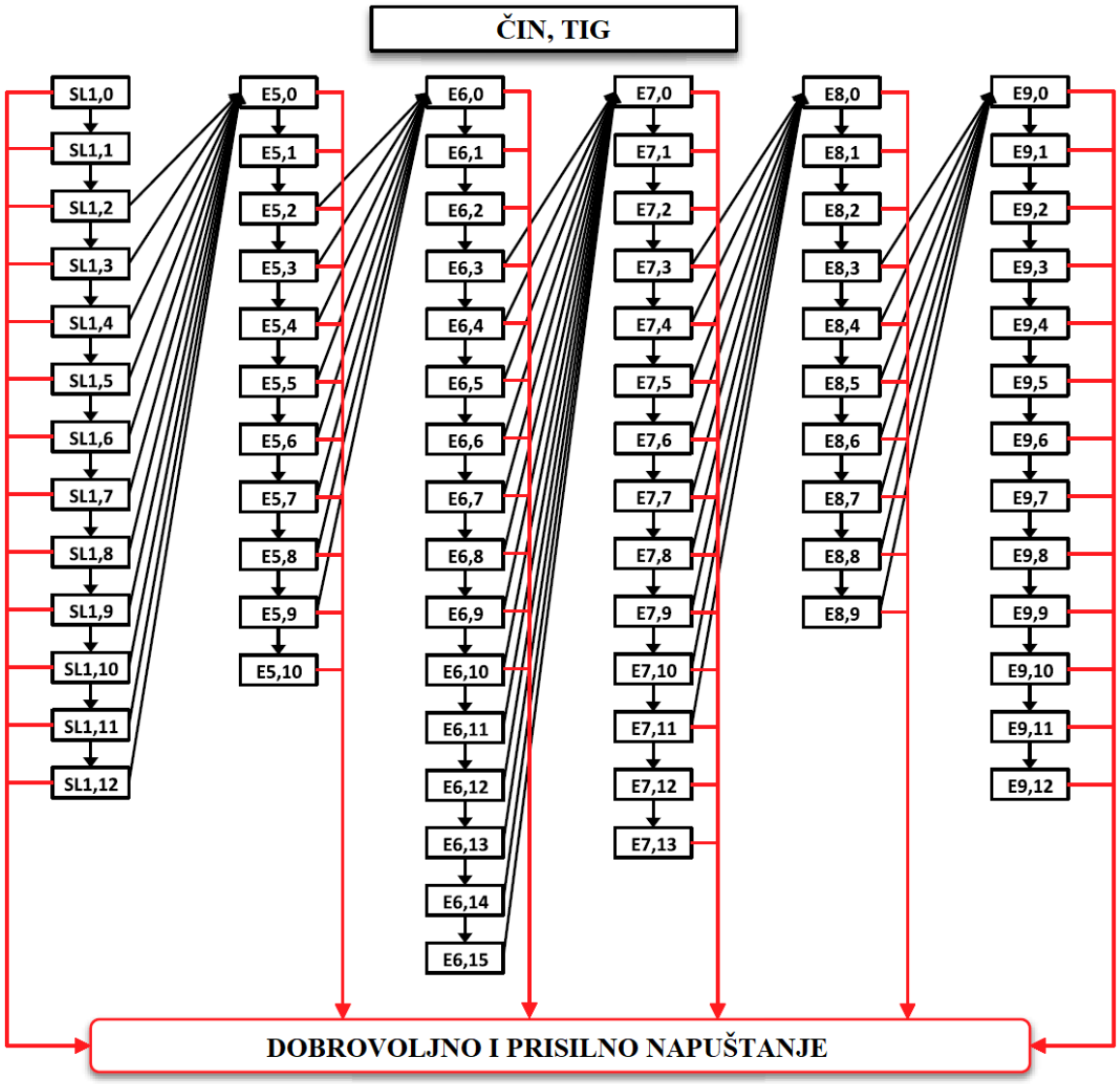
3.4.2. Primjer - Model dinamike vojnog osoblja [21]

Američka vojska je kao velik mehanizam podložna i osjetljiva na razne ekonomske, društvene i ostale utjecaje. Pod tim utjecajima, mijenja se dostupan i potreban broj ljudi u službi. Kako bi se ta dva broja zadržala u ravnoteži, tj. kako ne bi došlo do manjka vojnog osoblja, potrebno je znati točan broj ljudi koji će napustiti službu u narednom periodu kako bi ih se moglo nadomjestiti. Također, kako bi se smanjio potreban broj novaka, poželjno je znati u kojim grupama ima najviše odlazaka kako bi se ih se određenim, specifičnim mjerama moglo potaknuti na ostanak. Takvo predviđanje je svakako moguće pomoću ranije obrađene regresijske analize, ali zbog velikog broja ulaznih varijabli, puno je veći opseg informacija kojima treba raspolagati, a zbog toga što su mnoge varijable vezane uz osobne karakteristike, postoji mogućnost diskriminacije. M. Zais i D. Zhang predlažu opisivanje napredovanja karijere vojnog osoblja pomoću Markovljeva lanca čime nestaje potreba za poznavanjem osobnih informacija.

3.4.2.1. Opis podataka i podloga za model

Analizom je obuhvaćeno 9 činova, od vojnika (E1) do vodnika (E9). Kategorije E1-E4 pripadaju istoj razini vještine i posjeduju slične obrasce ponašanja te su objedinjeni u jednoj kategoriji, SL1. Trenutna stanja su funkcija čina i broja godina provedenih u činu (time-in-grade, TIG). Definirana su i dva dodatna stanja koja obuhvaćaju dobrovoljno (VL) i prisilno (IVL) napuštanje, a pretpostavljeno je da pojedinci koji na taj način napuste vojsku nemaju

vjerojatnost ponovnog uvrštavanja pa se iz ta dva stanja ne može napredovati ni u kojem smjeru. Za svako stanje je definiran minimalan TIG da bi se ispunio uvjet za napredovanjem. Također je definiran i maksimalan TIG za svaki čin nakon kojeg slijedi kontrolna točka zadržavanja (retention control point, RCP). Ukupno postoji 79 stanja, a Slika 9. prikazuje sve moguće smjerove napredovanja, tranzicijski dijagram.



Slika 9. Tranzicijski dijagram za sva stanja

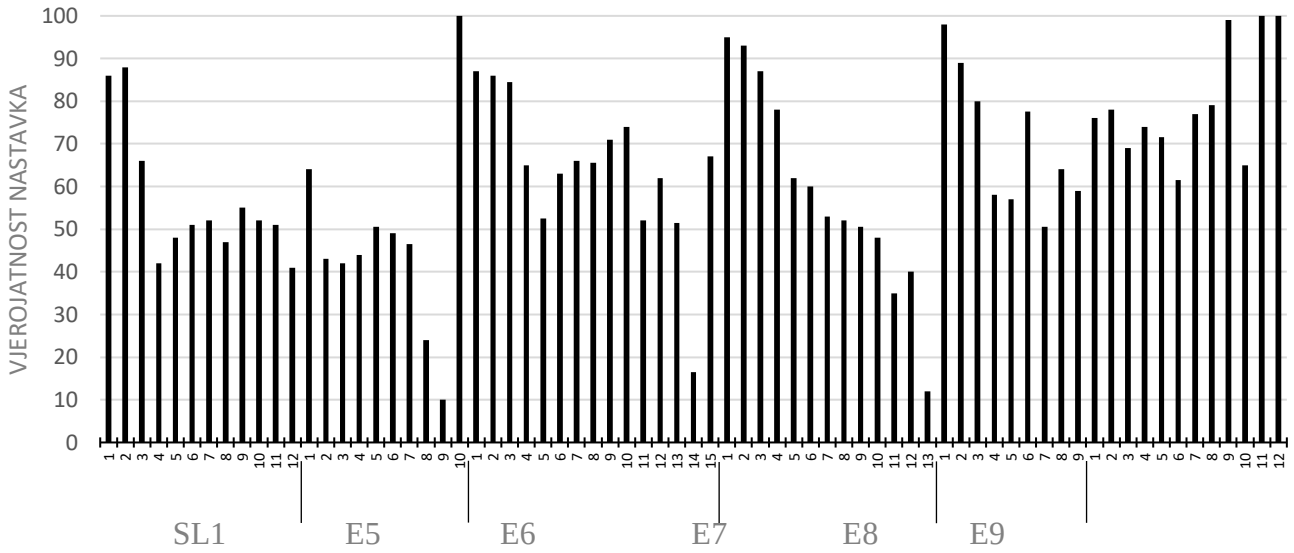
Četiri su vrste tranzicija: ostanak u istom stanju, napredovanje u sljedeće stanje, dobrovoljno napuštanje i prisilno napuštanje.

Prvi korak je izrada matrice broja tranzicija $C = [c_{ij}]$ dimenzija $m \times m$, gdje m označava ukupan broj stanja, a koji iznosi 79. c_{ij} predstavlja ukupan broj tranzicija iz stanja i u stanje j za promatrano razdoblje. S obzirom na dimenzije matrice, njezin prikaz je izostavljen.

Druga matrica, također dimenzija $m \times m$, je $M = [\theta_{ij}]$, gdje je θ_{ij} vjerojatnost prelaska iz stanja i u stanje j, a računa se kao:

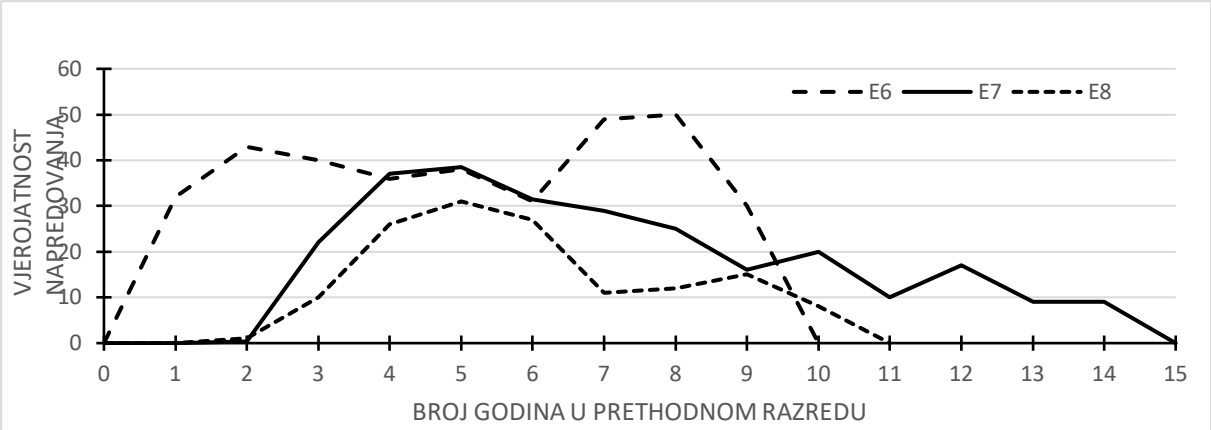
$$\theta_{ij} = \frac{c_{ij}}{\sum_{j=1}^{m} c_{ij}}. \tag{36}$$

Matrica M također nije prikazana u cijelosti zbog veličine, a grafički prikaz njenog sadržaja dan je u nastavku.



Slika 10. Vjerojatnosti nastavka službe unutar trenutnog čina

Slika 10. prikazuje vjerojatnosti nastavka službe unutar istog čina, tj. vjerojatnost da se neće dogoditi napredovanje ili otpuštanje. Pojedinac u devetoj godini u činu E5 ima vjerojatnost 10% da će i sljedeće godine imati čin E5 što govori da će, ako toliko vremena ima isti čin, najvjerojatnije zaslužiti promaknuće, a ako ga ne dobije, u sljedećoj ga sigurno neće dobiti jer je vjerojatnost nastavka unutar istog čina za sljedeće stanje 100% što znači da je vjerojatnost promaknuća 0%. Pojedinac koji u devetoj godini u činu E5 ne dobije promaknuće će sigurno biti otpušten iz vojske jer će doći do RCP.



Slika 11. Vjerojatnost napredovanja s obzirom na trajanje trenutnog čina

Iz Slika 11. je vidljivo kako se na temelju broja godina provedenih u trenutnom činu može predvidjeti vjerojatnost odlaska iz njega. Npr., pojedinci u E5 imaju podjednaku vjerojatnost napredovanja u čin E6 kroz cijelo vrijeme provedeno u činu E5, osim ako je to trajanje vrlo blizu RCP dok je za promociju iz čina E6 u E7 ipak potrebno određeno vrijeme provedeno u činu jer je promocija u prve dvije godine gotovo nemoguća.

Tablica 20. Prisilno napuštanje po stanjima

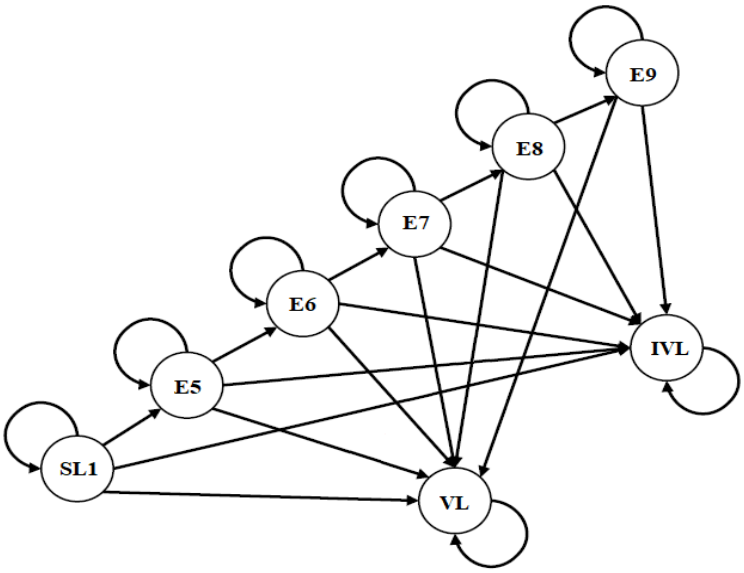
TIG	SL4	E5	E6	E7	E8	E9
0	27.6%	4.7%	2.4%	2.3%	0.9%	0.0%
1	11.6%	7.2%	3.6%	2.2%	0.8%	0.4%
2	10.0%	10.2%	4.0%	2.5%	0.9%	1.4%
3	14.7%	14.2%	4.0%	1.8%	0.7%	1.3%
4	18.0%	12.9%	4.9%	1.2%	1.0%	6.1%
5	20.2%	10.9%	4.0%	0.9%	0.8%	0.0%
6	18.0%	19.4%	3.5%	1.4%	2.3%	0.0%
7	20.9%	8.4%	2.5%	2.6%	0.0%	0.0%
8	22.5%	20.1%	0.7%	2.2%	0.0%	0.0%
9	21.1%	0.0%	0.9%	1.3%	RCP	0%
10	33.1%	RCP	0.7%	7.6%	–	0%
11	39.5%	–	2.3%	5.0%	–	RCP
12	RCP	–	2.9%	0.0%	–	–
13	–	–	12.5%	RCP	–	–
14	–	–	0.0%	–	–	–
15	–	–	RCP	–	–	–

Tablica 20. prikazuje stopu prisilnog napuštanja, a s obzirom na to da glavni dio tih slučajeva čine smrtna stradanja, jasno je zašto je stopa najviša kod dvaju najnižih činova.

3.4.2.2. Model Markovljeva lanca za činove

Iako su do sada pojedinci bili raspoređeni po stanjima s obzirom na čin i na vrijeme provedeno u njemu, za ovaj model će biti korištena podjela samo po činovima jer detaljna zapažanja koja nam omogućuje prethodna podjela nisu ključna za ovaj model.

Uz šest činova (SL1, E5, E6, E7, E8, E9), kao dva dodatna razreda, iako u stvarnosti ne postoje, će biti razmatrana stanja dobrovoljnog i prisilnog napuštanja.



Slika 12. Tranzicijski dijagram za razrede

Iz matrice M se može dobiti matrica P_M koja sadrži vrijednosti tranzicija između pojedinih razreda tako što se tranzicije između stanja unutar činova promatraju kao tranzicija nazad u isti razred, tj. zadržavanje u njemu, dok se samo promocije i napuštanja smatraju tranzicijama između razreda.

$$P_M = \begin{bmatrix} 0{,}6691 & 0{,}1588 & 0 & 0 & 0 & 0 & 0{,}0646 & 0{,}1075 \\ 0 & 0{,}5335 & 0{,}2542 & 0 & 0 & 0 & 0{,}1495 & 0{,}0627 \\ 0 & 0 & 0{,}6992 & 0{,}1110 & 0 & 0 & 0{,}1408 & 0{,}0490 \\ 0 & 0 & 0 & 0{,}8100 & 0{,}0800 & 0 & 0{,}1000 & 0{,}0100 \\ 0 & 0 & 0 & 0 & 0{,}8150 & 0{,}0499 & 0{,}1282 & 0{,}0069 \\ 0 & 0 & 0 & 0 & 0 & 0{,}8380 & 0{,}1554 & 0{,}0066 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (37)$$

Prikladnost matrice se provjerava korištenjem hi-kvadrat testom po izrazu:

$$\chi^2 = \sum_{i=1}^{8} \frac{(O_i' - E_i)^2}{E_i}. \quad (38)$$

O je vektor stvarnih početnih vrijednosti (brojeva članova razreda), a O' je vektor vrijednosti na kraju razdoblja. Vektor procijenjenih vrijednosti se dobiva iz $E = O^T P^3$. Dobivena vrijednost χ^2 je 6,76 što je manje od χ^2 za $p=0.05$ i 7 stupnjeva slobode te je matrica prikladna.

Razlog iz kojeg su svi prethodni koraci provedeni nije donošenje zaključaka na temelju podataka u matrici P_M (iako i oni imaju svoju primjenu), već transformacija matrice P_M u matricu N_M, temeljnu matricu iz koje se onda mogu dobiti informacije upotrebljive u dugoročnom planiranju.

$$N_M = \begin{bmatrix} 3{,}0221 & 1{,}0287 & 0{,}8694 & 0{,}5079 & 0{,}2196 & 0{,}0677 \\ 0 & 2{,}1436 & 1{,}8115 & 1{,}0583 & 0{,}4577 & 0{,}1410 \\ 0 & 0 & 3{,}3245 & 1{,}9422 & 0{,}8399 & 0{,}2587 \\ 0 & 0 & 0 & 5{,}2632 & 2{,}2760 & 0{,}7011 \\ 0 & 0 & 0 & 0 & 5{,}4054 & 1{,}6650 \\ 0 & 0 & 0 & 0 & 0 & 6{,}1728 \end{bmatrix} \quad (39)$$

n_{ij} govori koliko puta će lanac završiti u *j* ako krene iz *i*, odnosno to je očekivani broj tranzicija iz početnog stanja *i* u konačno stanje *j* prije apsorpcije.

Množenjem matrice N_M i apsorpcijske matrice *R*, dobiva se matrica vjerojatnosti prelaska u jedno od apsorpcijskih stanja iz bilo kojeg od tranzijentnih stanja:

$$N_M R = \begin{bmatrix} 0{,}5609 & 0{,}4390 \\ 0{,}7619 & 0{,}2378 \\ 0{,}8102 & 0{,}1898 \\ 0{,}9270 & 0{,}0730 \\ 0{,}9517 & 0{,}0483 \\ 0{,}9593 & 0{,}0407 \end{bmatrix}. \quad (40)$$

Vidi se da razlika vjerojatnosti postanka dobrovoljnog i prisilnog gubitka izrazito divergira s porastom čina što je i očekivano (Tablica 20.).

Matrica *T* sastavljena od suma redova temeljne matrice govori koliko perioda će proći u jednom stanju do tranzicije u apsorpcijsko.

$$T = \begin{bmatrix} 5{,}7153 \\ 5{,}6121 \\ 6{,}3652 \\ 8{,}2402 \\ 7{,}0704 \\ 6{,}1728 \end{bmatrix}. \quad (41)$$

Rast vremena do apsorpcije raste do čina E7 jer i poticaji za ostankom do mirovine rastu (viša mirovina, manje rizika, manje naporne obveze, teži pronalazak posla u civilnom sektoru koji pruža jednake prednosti...), a onda ponovo pada jer se i vrijeme do mirovine skraćuje, a i poticaji za ostankom stagniraju.

3.4.2.3. Model Markovljevog lanca za sva stanja

Uz pretpostavku da pojedinac usporedi očekivanu plaću i mirovinu s mogućim uvjetima u civilnom sektoru, ovaj model procjenjuje odluku pojedinca o ostanku ili odlasku u bilo kojem trenutku. Model bilježi dinamiku razvoja karijere, uključujući plaću i mogućnosti za napredak, što omogućuje oblikovanje poticaja za ciljane grupe kako bi ih se navelo na ostanak. U nastavku će biti iznesen zamršeni matematički postupak dobivanja funkcije optimalne odluke čije detaljno razjašnjavanje nije u domeni teme tako da će biti ukratko navedeni glavni principi.

Pretpostavka je da se pojedinac nakon napuštanja vojske ne može vratiti u službu pa je vjerojatnost prelaska iz stanja *i* u stanje *j*:

$$q_{ij} = \frac{\theta_{ij}}{1 - \sum_{j \neq VL} \theta_{ij}}. \tag{42}$$

Iz određenih izraza za modeliranje plaće i mirovine vojnog djelatnika te pogodnosti za duži radni staž:

$$\omega_t = \begin{cases} 0, & t < 20 \\ 0.5 + 0.025(t - 20), & t \geq 20 \end{cases}. \tag{43}$$

se dolazi do dviju funkcija predviđene kompenzacije, jedne za ostanak u službi, jedne za odlazak:

$$m_{i,t} + c_{i,t} + \beta \sum_{j=1}^{S} P_{i,j} V_{t+1}(j), \tag{44}$$

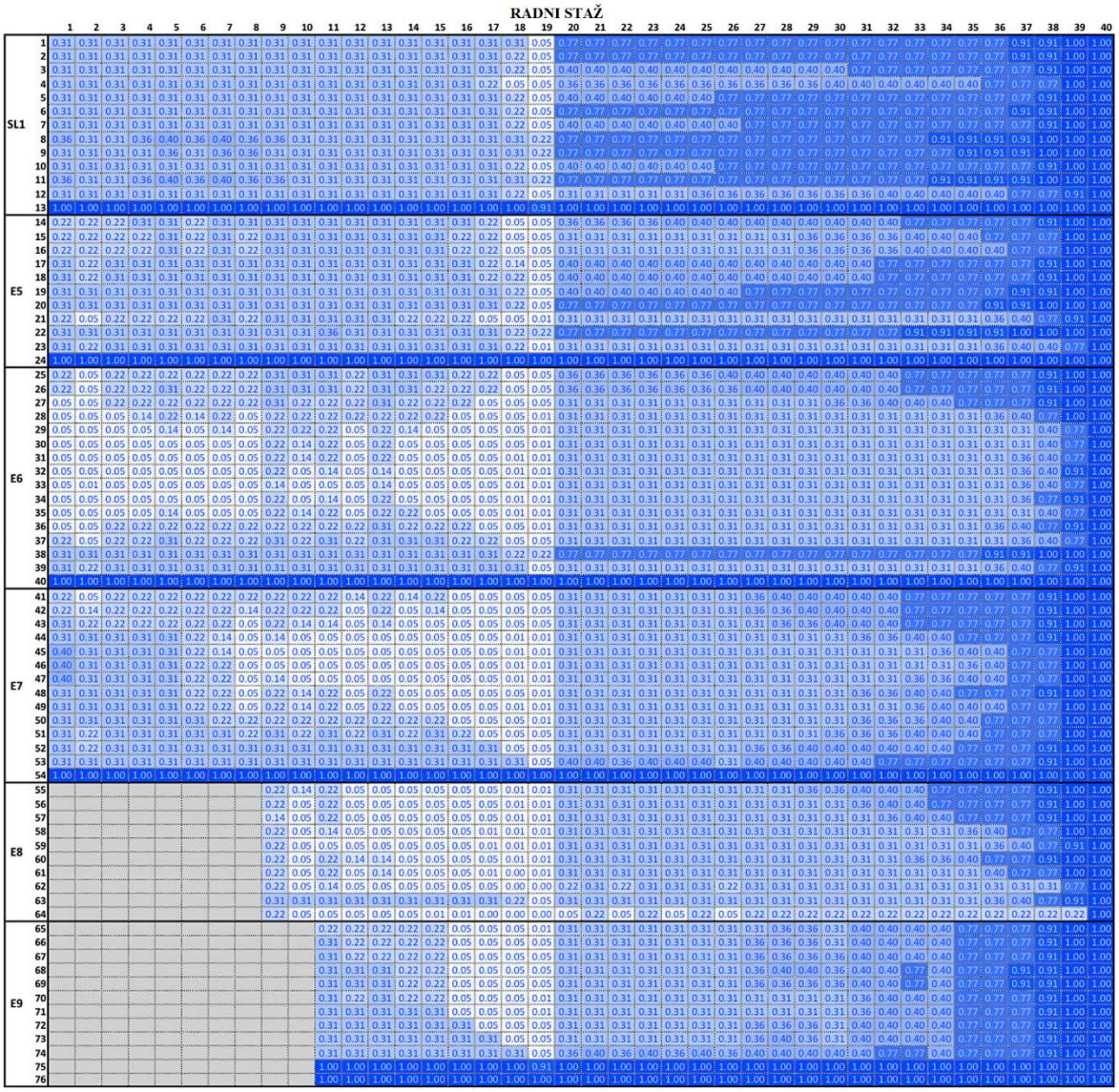
$$w_t \widehat{m}_{i,t} \sum_{k=t}^{\tau-\eta} \beta^{k-t} + \delta_x m_{i,t} \sum_{k=t}^{T-1} \beta^{k-t} (1 + \xi_{k-t}) \tag{45}$$

te se, uz uzimanje u obzir vjerojatnosti za postajanje prisilnog gubitka iz njih dolazi do funkcije optimalne vrijednosti:

$$\begin{aligned} V_t(i) = (1 - q_{i,\mathrm{IL}}) E \Bigg[\max \Bigg\{ & m_{i,t} + c_{i,t} + \beta \sum_{j \neq \mathrm{IL}} \frac{q_{ij} V_{t+1}(j)}{(1 - q_{i,\mathrm{IL}})}, \\ & w_t \widehat{m}_{i,t} \sum_{k=t}^{\tau-\eta} \beta^{k-t} + \delta_x m_{i,t} \sum_{k=t}^{T-1} \beta^{k-t} (1 + \xi_{k-t}) \Bigg\} \Bigg] \\ & + q_{i,\mathrm{IL}} \left(w_t \widehat{m}_{i,t} \sum_{k=t}^{\tau-\eta} \beta^{k-t} + \delta_x m_{i,t} \sum_{k=t}^{T-1} \beta^{k-t} (1 + \xi_{k-t}) \right), \end{aligned} \tag{46}$$

iz koje se uvrštavanjem iznosa životne dobi, dobi pri pristupanju vojnoj službi, radnog vijeka u civilnom sektoru te parametara koji utječu na promjenu plaće dobiva indeks optimalne odluke, odnosno vjerojatnost da je odlazak bolja opcija.

Tablica 21. Iznosi indeksa optimalne odluka



Tamnije obojena polja ukazuju na viši indeks, a iz toga je jasno vidljivo kako su pogodnosti dobivene prekoračenjem staža od 20 godina vrlo utjecajne u donošenju odluke o odlasku, i po naglom porastu vjerojatnosti za odlaskom nakon te granice, i po smanjenju vjerojatnosti kako se granica bliži. Horizontalna tamna područja označavaju RCP gdje zapravo ne postoji izbor već je napuštanje jedini mogući ishod.

Tablica 21. ukazuje na grupe kod kojih je moguće ne toliko opsežnim ulaganjima povisiti stopu ostanaka u službi što je od velike koristi u situacijama kad se primijeti opasnost od nastajanja manjkova, npr. u slučaju nižeg broja unovačenih pojedinaca od planiranog. Ako bi se općenite metode poticanja primijenilo na cijeli skup ili npr. na određeni čin, djelovalo bi se i na velik broj pojedinaca kod kojih taj poticaj ne bi doveo do smanjenja indeksa koje bi uzrokovalo

promjenu odluke i kod kojih bi se indeks dodatno smanjio iako je već otprije na sigurnoj razini. Ako se jednaka sredstva usmjere samo na usku grupu, npr. visoke činove koji su tek napunili 20 godina staža, moguće je dovesti do značajnijeg sniženja indeksa i tako djelovanjem na manji broj pojedinaca zapravo dovesti do većeg broja ostanaka.

4. ZAKLJUČAK

Pregledom literature je stečena slika o raširenosti korištenja statističkih metoda u vojnim sustavima te je na temelju nje napravljena klasifikacija područja korištenja.

Najzastupljenijim područjem se pokazalo ocjenjivanje prikladnosti korištenja određene taktike, metode ili opreme. Zbog vremena u kojem su se ti primjeri odvijali se stječe dojam da danas fokus više nije na tim područjima, ali predmeti današnjih istraživanja i njihovi rezultati još dugo godina neće biti javni. Bez obzira na zastarjelost tadašnjih otkrića, iz svih prikazanih primjera je jasan značaj logičkog i kritičkog pristupa analizi, nečeg na što su i autori literature izričito stavljali naglasak. Danas je računski dio metoda neusporedivo lakši zahvaljujući primjeni računala i nije nikakav problem izvući razne parametre iz seta podataka, ali dio sa zaključivanjem je jednak kao i u 2. svj.ratu.

Područje koje moderna saznanja mogu itekako unaprijediti je simuliranje ishoda s obzirom na promjenjive okolnosti. Razvoj neuronskih mreža i strojnog učenja daje novu dimenziju dvjema metodama obrađenima u radu, regresijskoj analizi i Markovljevim lancima. Mogućnost učenja iz raspoloživih podataka i prošlih događaja znači da simulacije mogu postati još točnije bez proširivanja kompleksnosti van kapaciteta čovjeka.

LITERATURA

[1] The World Bank, »Armed forces personnel (% of total labor force)« https://data.worldbank.org/indicator/MS.MIL.TOTL.TF.ZS. [Pokušaj pristupa 17.9.2019.].

[2] O. Solandt, Observation, Experiment and Measurement in Operations Research, 1955.

[3] R. Ruggles, An Empirical Approach to Economic Intelligence in WW2, 1947.

[4] P. M. Morse i G. E. Kimball, Methods of Operations Research, 1946.

[5] P. Blackett, Studies of War, Nuclear and Conventional, 1962.

[6] J. F. McCloskey, British Operationa Research in WW2, 1987.

[7] D. Neuhauser, »Florence Nightingale gets no respect: as a statistician« 2003.

[8] »Florence Nightingale: Measuring Hospital Care Outcomes« *Joint Commission on Accreditation of Healthcare Organizations,* 1999.

[9] *Diagram of the causes of mortality in the Army in the East,* Bethesda, MD : U.S. National Library of Medicine, National Institutes of Health, Health & Human Services.

[10] B. O. Koopman, A Study of the Logical Basis of Combat Simulation, 1969.

[11] S. I. Gass, Military manpower planning models, 1990.

[12] M. Mangel, »Abraham Wald's work on aircraft survivability« 1984.

[13] A. Wald, »A method of estimating plane vunerability based on damage of survivors« 1943.

[14] B. Petz, Petzova statistika, 2012.

[15] D. J. Sheskin, Handbook of Parametric and Nonparametric Statistical Procedures, 2000.

[16] S. I. Gass, Encyclopedia of Operations Research and Management Science, 2013.

[17] N. Jaiswal i B. Nagabushana, »Measure of Combat Success Through Historical Data« 1995.

[18] »CAA Database of Battles, Version 1990« https://github.com/jrnold/CDB90. [Pokušaj pristupa 18.2.2019.].

[19] »Materijali s predavanja i vježbi, Prehrambeno-tehnološki fakultet Osijek« http://www.mathos.unios.hr/ptfstatistika/zakljucivanje2sv.pdf.

[20] F. S. Hiller, Introduction to Operations Research, 2010.

[21] M. Zais i D. Zhang, »A Markov chain model of military personnel dynamics« 2015.

Napomena:

Za obradu podataka korištena je probna verzija programskog paketa *IBM SPSS Statistics 23.*

PRILOG

I. CD-R disk

II. Statističke tablice

Tablica A. Hi-kvadrat distribucija [14]

Stup. slob.	P												
	0.995	0.990	0.975	0.950	0.900	0.750	0.500	0.250	0.100	0.050	0.025	0.010	0.005
1	0.00004	0.00016	0.0010	0.0039	0.016	0.102	0.455	1.32	2.71	3.84	5.02	6.63	7.83
2	0.0100	0.0201	0.0506	0.103	0.211	0.575	1.39	2.77	4.61	5.99	7.38	9.21	10.6
3	0.0717	0.115	0.216	0.352	0.584	1.21	2.37	4.11	6.25	7.81	9.35	11.3	12.8
4	0.207	0.297	0.484	0.711	1.06	1.92	3.36	5.39	7.78	9.49	11.1	13.3	14.9
5	0.412	0.554	0.831	1.15	1.61	2.67	4.35	6.63	9.24	11.1	12.8	15.1	16.7
6	0.676	0.872	1.24	1.64	2.20	3.45	5.35	7.84	10.6	12.6	14.4	16.8	18.5
7	0.989	1.24	1.69	2.17	2.83	4.25	6.35	9.04	12.0	14.1	16.0	18.5	20.3
8	1.34	1.65	2.18	2.73	3.49	5.07	7.34	10.2	13.4	15.5	17.5	20.1	22.0
9	1.73	2.09	2.70	3.33	4.17	5.90	8.34	11.4	14.7	16.9	19.0	21.7	23.6
10	2.16	2.56	3.25	3.94	4.87	6.74	9.34	12.5	16.0	18.3	20.5	23.2	25.2
11	2.60	3.05	3.82	4.57	5.58	7.58	10.3	13.7	17.3	19.7	21.9	24.7	26.8
12	3.07	3.57	4.40	5.23	6.30	8.44	11.3	14.8	18.5	21.0	23.3	26.2	28.3
13	3.57	4.11	5.01	5.89	7.04	9.30	12.3	16.0	19.8	22.4	24.7	27.7	29.8
14	4.07	4.66	5.63	6.57	7.79	10.2	13.3	17.1	21.1	23.7	26.1	29.1	31.3
15	4.60	5.23	6.26	7.26	8.55	11.0	14.3	18.2	22.3	25.0	27.5	30.6	32.8
16	5.14	5.81	6.91	7.96	9.31	11.9	15.3	19.4	23.5	26.3	28.8	32.0	34.3
17	5.70	6.41	7.56	8.67	10.1	12.8	16.3	20.5	24.8	27.6	30.2	33.4	35.7
18	6.26	7.01	8.23	9.39	10.9	13.7	17.3	21.6	26.0	28.9	31.5	34.8	37.2
19	6.84	7.63	8.91	10.1	11.7	14.6	18.3	22.7	27.2	30.1	32.9	36.2	38.6
20	7.43	8.26	9.59	10.9	12.4	15.5	19.3	23.8	28.4	31.4	34.2	37.6	40.0
21	8.03	8.90	10.3	11.6	13.2	16.3	20.3	24.9	29.6	32.7	35.5	38.9	41.4
22	8.64	9.54	11.0	12.3	14.0	17.2	21.3	26.0	30.8	33.9	36.8	40.3	42.8
23	9.26	10.2	11.7	13.1	14.8	18.1	22.3	27.1	32.0	35.2	38.1	41.6	44.2
24	9.89	10.9	12.4	13.8	15.7	19.0	23.3	28.4	33.2	36.4	39.4	43.0	45.6
25	10.5	11.5	13.1	14.6	16.5	19.9	24.3	29.3	34.4	37.7	40.6	44.3	46.5
26	11.2	12.2	13.8	15.4	17.8	20.8	25.3	30.4	35.6	38.9	41.9	45.6	48.3
27	11.8	12.9	14.6	16.2	18.1	21.7	26.3	31.5	36.7	40.1	43.2	47.0	49.6
28	12.5	13.6	15.3	16.9	18.9	22.7	27.3	32.6	37.9	41.3	44.5	48.3	51.0
29	13.1	14.3	16.0	17.7	19.8	23.6	28.3	33.7	39.1	42.6	45.7	49.6	52.3
30	13.8	15.0	16.8	18.5	20.6	24.5	29.3	34.8	40.3	43.8	47.0	50.9	53.7