

# Statističke metode u zaštiti okoliša

---

Zdrilić, Ante

Undergraduate thesis / Završni rad

2015

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Chemical Engineering and Technology / Sveučilište u Zagrebu, Fakultet kemijskog inženjerstva i tehnologije**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:149:697871>

Rights / Prava: [In copyright](#)

Download date / Datum preuzimanja: **2021-10-23**



Repository / Repozitorij:

[Repository of Faculty of Chemical Engineering and Technology University of Zagreb](#)



**SVEUČILIŠTE U ZAGREBU**  
**FAKULTET KEMIJSKOG INŽENJERSTVA I TEHNOLOGIJE**  
**SVEUČILIŠNI PREDDIPLOMSKI STUDIJ**

Ante Zdrilić

**ZAVRŠNI RAD**

Zagreb, rujan 2015.

**SVEUČILIŠTE U ZAGREBU**  
**FAKULTET KEMIJSKOG INŽENJERSTVA I TEHNOLOGIJE**  
**SVEUČILIŠNI PREDDIPLOMSKI STUDIJ**

Ante Zdrilić

**STATISTIČKE METODE U ZAŠTITI OKOLIŠA**

**ZAVRŠNI RAD**

Voditelj rada: prof. dr. sc. Tomislav Bolanča

Članovi ispitnog povjerenstva: prof. dr. sc. Tomislav Bolanča

v.pred.dr.sc. Lidija Furač

izv. prof.dr.sc. Juraj Šipušić

Zagreb, rujan 2015.

Završni rad izrađen je na Fakultetu kemijskog inženjerstva i tehnologije, Sveučilišta u Zagrebu, Zavod za analitičku kemiju, akademske godine 2014./2015.

## **Sažetak**

U današnjemu svijetu važnost točnih, pouzdanih i usporedivih mjernih rezultata od velikog je značaja. Statističke metode pomažu iz skupa podataka izdvojiti informaciju, izračunati bitne činjenice na temelju kojih se može opisati proces i donijeti odluka o njegovom usmjeravanju.

U procesu utvrđivanja značajnih aspekata okoliša i mjera za poboljšanje, jedan od procesnih koraka je vrednovanje aspekata okoliša. Potrebno je odrediti kriterije i metode kojima će se osigurati učinkovito provođenje i nadzor procesa. Neprekidno poboljšanje znači mjeriti, analizirati, nadzirati procese te provoditi postupke u cilju otklanjanja ili smanjenja rizika na prihvatljivu razinu. Ovim aktivnostima skupljaju se brojni kvalitativni ili kvantitativni podaci. Obradom podataka dobiju se informacije na temelju kojih se donose odluke. U obradi podataka primjenjuju se brojne statističke metode.

U radu se ukazuje na mogućnosti programa Excel kao sastavnog dijela Microsoft Office, a koji podržava brojne statističke funkcije i grafičke prikaze rezultata.

***Ključne riječi:*** statističke metode, analiza podataka, prikaz podataka

## **Abstract**

Nowadays, the importance of accurate, reliable and comparable measurement results takes a big place. With statistical methods we can extract information from data collection and calculate the essential facts on the basis of which can be described the process and decision about that process can be made.

In the process of determining significant environmental aspect and measures to improve, one of the process steps is the evaluation of environmental aspects. It is necessary to establish criteria and methods to ensure the effective implementation and monitoring process. Continuous improvement means to measure, analyze, monitor processes and conduct in order to eliminate or reduce the risk to an acceptable level. With these activities we collect a great number of qualitative or quantitative data. Analysis of the data obtain information that are crucial for further decisions. In data processing a variety of statistical methods are applied.

This thesis points to the potential of Excel, as part of Microsoft Office, which supports many statistical functions and graphic results.

**Key words:** statistical methods, data analysis, data display

## SADRŽAJ:

1	UVOD .....	1
2	STATISTIČKI IZAZOVI U ZAŠTITI OKOLIŠA.....	3
3	VALIDACIJA .....	4
3.1	Parametri validacije.....	6
3.1.1	Bias.....	6
3.1.2	Preciznost proračunatih vrijednosti .....	8
3.2	Granica detekcije.....	8
3.3	Granica kvantifikacije .....	10
4	ANALIZA CENZURIRANIH PODATAKA.....	12
5	EKSPERIMENTALNI DIZAJN .....	13
5.1	Principi eksperimentalnog dizajna .....	14
5.1.1	Komparativni dizajn .....	14
5.1.2	Replikacija .....	14
5.1.3	Randomizacija.....	15
5.1.4	Blokiranje.....	15
6	KORELACIJA PODATAKA.....	16
6.1	Kovarijacija.....	16
6.2	Serijska korelacija .....	16
7	PROCJENA RIZIKA.....	17
7.1	Procjena učinka.....	18
7.1.1	Karakterizacija .....	18
7.1.2	Modeliranje .....	18
8	PRIMJENA STATISTIKE U ANALIZI OKOLIŠA .....	19
8.1	Deskriptivna statistika .....	19
8.2	Ispitivanje hipoteze .....	21
8.3	Regresijska analiza .....	22
8.4	Uzorkovanje.....	23
8.5	Simulacije .....	23
8.6	Vremenske serije.....	23
9	ZAKLJUČAK.....	24
10	POPIS KRATICA.....	25
11	LITERATURA .....	26

# 1 UVOD

Statističke metode pomažu iz skupa podataka izdvojiti informaciju, izračunati bitne činjenice na temelju kojih se može opisati proces i donijeti odluka o njegovom usmjeravanju.

U procesima upravljanja zaštitom okoliša najveću primjenu našle su statističke veličine: raspon podataka, medijan i inetrkvartili. Kao grafički prikaz ovih parametara u literaturi se često primjenjuje Box & Whiskers dijagram.<sup>3</sup> Nažalost taj dijagram nije izravno programiran u Excelu, ali ga se može dobiti zaobilaznim putovima.

U procesu utvrđivanja značajnih aspekata okoliša i mjera za poboljšanje, jedan od procesnih koraka je vrednovanje aspekata okoliša. Potrebno je odrediti kriterije i metode kojima će se osigurati učinkovito provođenje i nadzor procesa. Neprekidno poboljšanje znači mjeriti, analizirati, nadzirati procese te provoditi postupke u cilju otklanjanja ili smanjenja rizika na prihvatljivu razinu. Ovim aktivnostima skupljaju se brojni kvalitativni ili kvantitativni podaci. Obradom podataka dobiju se informacije na temelju kojih se donose odluke. U obradi podataka primjenjuju se brojne menadžerske metode, a značajno mjesto zauzimaju statističke metode. U radu se ukazuje na mogućnosti programa Excel kao sastavnog dijela Microsoft Office, a koji podržava brojne statističke funkcije i grafičke prikaze rezultata analiza.

Sustav upravljanja okolišem je upravljački proces kojim se upravlja neželjenim posljedicama nekog proizvodnog ili drugog procesa koji može ugroziti ili već ugrožava okoliš. U odnosu na ekologiju ovdje su procesi jednostavniji i najčešće su dostatne statističke metode koje su programirane u Microsoftovom programskom paketu Excel. Proces i obično ne slijede normalnu raspodjelu, a rasipanja podataka oko srednje vrijednosti mogu biti izrazito velika. Budući skoro svaki menadžer sustava upravljanja okolišem ima na raspolaganju računalo i koristi Microsoft Office programski paket, cilj je članka ukazati na mogućnosti koje pruža Excel u obradi i prikazu podataka za potrebe sustava upravljanja okolišem.<sup>3</sup> Statistika je nauka o principima prikupljanja, organizacije, analize, sažetog prikaza i interpretacije podataka dobivenih promatranjem ili mjerenjem vrijednosti varijable osnovnog skupa, uzorka ili jedinke. Metode statističke analize procesa primjenjuju se i u procesu upravljanja zaštitom okoliša.<sup>3</sup>



Temeljne faze statističkog istraživanja su :

1. Prikupljanje statističkih podataka – prva faza statističkog rada koja ima zadatak upoznati obilježja pojedinačnih elemenata prethodno precizno definiranog statističkog skupa. Metode prikupljanja su brojanje, mjerenje, ocjenjivanje, evidentiranje, anketiranje i sl.
2. Grupiranje statističkih podataka – postupak raščlambe statističkog skupa na određeni broj grupa ili podskupova prema prethodno utvrđenim modalitetima.
3. Tabelarno i grafičko prikazivanje – ima za cilj na jasan, cjelovit i pregledan način prezentirati rezultate provođenja prethodne dvije faze.
4. Statistička analiza i interpretacija dobivenih rezultata - faza u kojoj se, na podatke o jednoj ili više promatranih statističkih masa, primjenjuju odgovarajuće statističke metode da bi se odredile značajke tih pojava.

## 2 STATISTIČKI IZAZOVI U ZAŠTITI OKOLIŠA

Puno je aspekata obuhvaćeno u vezi okolišnih problema, uključujući ekonomske, političke, psihološke, medicinske, znanstvene i tehnološke. U svrhu razumijevanja i rješavanja problema koji se mogu pojaviti u sklopu navedenih aspekata često su uključene određene kvantitativne mjere, osobito prikupljanje i analiza podataka. Efektivno obrađivanje tih kvantitativnih problema je u domeni statistike. Statistika tada može biti shvaćena kao način za produktivno učenje na temelju dobivenih kvantitativnih podataka.<sup>2</sup>

Kada je pojedinac suočen s novim problemom, tada se postavljaju dva temeljna pitanja koja su od velike važnosti, a to su: 'Kako će primjena statistike riješiti taj nastali problem i koje su tehnike najbolje za rješavanje tog problema?'. Mnogo nezavisnih problema nastaje te postoje različite statističke tehnike, krećući se u rasponu od onih tehnika koje uključuju izradu jednostavnih dijagrama do onih tehnika koje uključuju nadograđivanje i evaluaciju ponavljajućih parametara.<sup>1</sup>

Neki problemi mogu biti riješeni ukoliko ih se podvrgne određenoj analitičkoj metodi prema podacima koji su dostupni. Puno je češći slučaj u kojem pojedina analiza mora ići postepeno. Upravo zbog navedenog, statističar mora težiti svestranosti i snalažljivosti baziranima na već prije isprobanim procedurama, no uvijek imajući u obziru kako bi se idući problem na koji nailazi mogao ne uklapati u već prije isprobanu statističku metodu. Korištenje statističkih metoda u zaštiti okoliša zna biti vrlo zahtjevan posao.

Cilj statistike jest učiniti proces samog istraživanja i obrađivanja podataka efikasnim. Analiziranje podataka je dijelom znanost, dijelom zanat, a dijelom umjetnost. Temelji se na iskustvu, vještinama i talentu uz neizostavnu pomoć određenih mjernih instrumenata.

Deskriptivna (opisna) statistika obuhvaća principe i postupke organiziranja, analizu i sažet prikaz prikupljenih podataka. U praćenju odvijanja procesa primjenjuje se ove tehnike: tablični prikazi prikupljenih podataka ,grafički prikazi prikupljenih podataka te statistika sažimanja, tj. računanje određene statističke veličine koja opisuje prikupljeni skup podataka. Zapravo, opisna statistika u cijelosti obuhvaća skup svih promatranih objekata a ima zadaću srediti i sažeti određene rezultate, tako da budu što pregledniji, razumljiviji i pogodniji za interpretaciju, daljnju analizu i primjenu. Sirovi podaci organiziraju se u distribucije frekvencija, prezentiraju se određenim grafikonima te na taj način pripremaju za detaljniju analizu.

Inferencijalna statistika temelji se na parcijalnom (nepotpunom) obuhvatu statističkog skupa ili populacije. Pojam populacije u statistici uključuje skupove kako pojedinca ili subjekata, tako i skupove predmeta ili pak skupove mjerenja.

### **3 VALIDACIJA**

Osnovna zadaća svakog analitičkog laboratorija jest postizanje brzih, točnih i vjerodostojnih rezultata analize. Zbog toga je potrebno opisati analitičke metode u onom obliku i opsegu kako bi se njenom primjenom dobili točni i pouzdani rezultati. Ključni princip za kvalitetu i pouzdanost rezultata je usporedivost laboratorija na široj, međunarodnoj razini. Kako bi bili usporedivi, analitički rezultati moraju biti prijavljeni s izjavom o nesigurnosti mjerenja (*NM*) i mjerne sljedivosti. Najbolji način izbjegavanja problema tijekom uporabe metode jest provođenje validacije analitičke metode. Iako se samom validacijom ne mogu predvidjeti svi problemi koji se mogu javljati tijekom primjene metode, postupci razvoja i validacije metode upućuju na one najčešće. Validacija je postupak kojim dokazujemo da metoda služi svrsi kojoj je namijenjena. Prije svega je potrebno definirati svrhu određene metode. Potom se utvrđuju postupci, tj. planiraju i provode eksperimenti čije rezultate treba prikupiti i prikazati kao dokaze o validnosti metode.<sup>4</sup>

Potrebno je naglasiti kako se isti postupci neće primjenjivati na sve metode; npr-različit je pristup validaciji kvantitativnih i kvalitativnih metoda.

Svakoj se metodi pristupa individualno, procjenjuje se što treba napraviti za dokaz svrhovitosti.

**Tablica 1.** Parametri validacije

Točnost	Izražava podudaranje dobivenih rezultata sa stvarnim. Provodi se najmanje na pet koncentracijskih razina uz najmanje tri ponovljena mjerenja svakog uzork određene koncentracije. Odstupanja od stvarne vrijednosti iskazuju se kao postotak iskorištenja (Recovery, $R$ ): $R = C_{exp} / C_{std} \times 100 \%$
Preciznost	Izražava se kao odstupanje (rasipanje) pojedinog rezultata mjerenja od srednje vrijednosti rezultata serija mjerenja višestrukog uzorkovanja istog homogenog uzorka pod strogo propisanim uvjetima.
Selektivnost	Definirana je kao sposobnost metode da kvantitativno odredi analit u prisustvu drugih komponenata u uzorku (poput onečišćenja, razradnih produkata, pomoćnih sredstava).
Granica detekcije (dokazivanja)	Predstavlja najmanju količinu analita u uzorku koja sa sigurnošću može biti detektirana.
Granica kvantifikacije (određivanja)	Predstavlja najmanju količinu analita koju kvantitativno možemo odrediti u nekom uzorku.
Lineranost i područje mjerenja	Odražava njezinu sposobnost da rezultati analize budu direktno (linearno) proporcionalni količini analita prisutnog u uzorku.
Robusnost (otpornost)	Pokazuje koliko su rezultati mjerenja dobiveni tom metodom neosjetljivi na manje promjene radnih uvjeta mjerenja (kapacitet kompenzacije promjene radnih uvjeta).

### 3.1 Parametri validacije

Iduće mjerenje koje pojedinac izmjeri ili podatak o mjerenju koji dobije je u opasnosti da se nalazi pod eksperimentalnom pogreškom. To je jednostavno činjenica koja se nikako ne smije zanemariti prilikom daljnjeg rada. Pri tome uvelike pomaže statistika koja nastoji otkriti i kvantificirati veličinu eksperimentalnih pogrešaka.

Eksperimentalna pogreška predstavlja zapravo devijaciju promatrane vrijednosti u odnosu na onu stvarnu vrijednost. To je zapravo fluktuacija ili nepodudarnost između ponovljenih mjerenja na identičnim uzorcima. Mjere na uzorcima sa pravom vrijednosti  $\eta$  neće biti identične, iako ljudi koji skupljaju, upravljaju uzorcima te analiziraju uzorke pokušavaju stvoriti što je više moguće identične uvjete. Promatrana vrijednost  $y_i$  će se razlikovati u odnosu na pravu vrijednost u vidu pogreške  $\varepsilon_i$ :

$$y_i = \eta + \varepsilon_i$$

Pogreška može imati sustavnu ili nasumičnu komponentu ili pak obje komponente. Ukoliko je  $\varepsilon_i$  u potpunosti nasumična pogreška, a  $\tau_i$  predstavlja sustavnu pogrešku, tada je:

$$y_i = \eta + (\varepsilon_i + \tau_i)$$

#### 3.1.1 Bias

Sustavne pogreške uzrokuju konstantno odstupanje ili bias od pravih vrijednosti. Mjere su konstantno visoke ili niske zbog neadekvatne opreme (kalibratorskog instrumentarija), nedovoljne pažnje ili potpunih pogrešaka. Jednom otkriven, bias može biti uklonjen pomoću kalibracije i pažljivim provjerama na eksperimentalnoj opremi i tehnici. Bias ne može biti smanjen ukoliko se izvrši još nekoliko dodatnih mjerenja ili ukoliko se uzme prosjek ponovljenih mjerenja.<sup>5</sup> Veličina biasa ne može biti procijenjena ukoliko prava vrijednost nije poznata.

Jednom kad je bias uklonjen, tada su opažanja pogođena eventualnim nasumičnim pogreškama i  $y_i = \eta * \varepsilon_i$ . Promatrani  $\varepsilon_i$  je zbroj svih odstupanja koja se 'ušuljaju' u sami mjerni proces budući da je potrebno puno koraka; počevši od samog prikupljanja uzoraka pa sve do pravilnog rada u laboratoriju. Sveukupni  $\varepsilon_i$  može imati visoku ili nisku vrijednost. Može biti dominantan unutar pojedinih koraka u samom mjernom procesu (npr. tijekom sušenja, vaganja ili ekstrakcije). Pozitivna strana tih pogrešaka jest njihova nasumičnost, a ne sustavnost.

Znak veličine nasumične pogreške nije predvidljiv iz pogreške u sklopu nekog drugog promatranja. Ukoliko potpuno nasumična pogreška  $\varepsilon_i$  iznosi zbroj različitih malih pogrešaka, što je uobičajeni slučaj, tada  $\varepsilon_i$  teži normalnoj distribuciji. Prosječna vrijednost  $\varepsilon_i$  će biti nula tako da će distribucija pogrešaka biti jednaka i u pozitivnom i u negativnom znaku.

Pretpostavimo da je konačni rezultat eksperimenta,  $y$ , određen jednadžbom  $y = a + b$  pri čemu  $a$  i  $b$  predstavljaju izmjerene vrijednosti. . Ukoliko  $a$  i  $b$  imaju sustavnu pogrešku od  $+1$ , jasno je da je sustavna pogreška u  $y$   $+2$ . No, ukoliko  $a$  i  $b$  pojedinačno imaju nasumičnu pogrešku u rasponu između nule i  $\pm 1$ , tada nasumična pogreška u  $y$  ne iznosi  $\pm 2$  rezultat je dobiven iz razloga zato što postoje slučajevi u kojima je nasumična pogreška u  $a$  pozitivna, dok je nasumična pogreška u  $b$  negativna (i obratno). Prosječna nasumična pogreška u pravoj vrijednosti će iznositi nula ukoliko su mjere i izračuni učinjeni mnogo puta. To znači da je očekivana nasumična pogreška u  $y$  jednaka nuli. Varijanca i standardna devijacija pogreške u  $y$  neće iznositi nula, no jednostavna matematička pravila mogu biti korištena za procjenu preciznosti konačnih rezultata ukoliko je preciznost svakog posebnog mjerenja poznata.<sup>7</sup>

Ponavljana mjerenja osiguravaju mogućnost kvantifikacije mjernih pogrešaka i procjenjuju njezinu vrijednost. Učinak nasumičnih pogrešaka može biti reduciran ukoliko se teži uzimanju prosjeka ponovljenih mjerenja. Pogreška koja ostane može biti kvantificirana i može biti iznesena statistička tvrdnja o preciznosti konačnih rezultata. Preciznost je povezana sa razilaženjem u dobivenim rezultatima unutar ponovljenih mjerenja. Precizni rezultati imaju male nasumične pogreške. Razilaženje u rezultatima uzrokovano nasumičnim pogreškama ne može biti eliminirano, ali može biti minimizirano pažljivim tehnikama. Još važnije, može se dobiti prosjek te može biti kvantificirano.<sup>4</sup>

### 3.1.2 Preciznost proračunatih vrijednosti

Inženjeri se koriste jednadžbama kako bi izračunali ponašanje prirodnih i konstruiranih sistema. Ponekad rezultati određenih jednadžbi znaju zavarati i odvesti na krivi trag. Neke od varijabli koje su korištene u jednadžbama predstavljaju mjerenja ili procjene; možda procjene na temelju nekog eksperimenta ili kao rezultata dobivenog iz nekog eksperimenta u priručniku. Neke od konstanti u jednadžbama kao što je recimo  $\pi$  su poznate, ali većina varijabli su zapravo procijenjene vrijednosti. Većinu vremena se ignorira činjenica da je rezultat jednadžbe neprecizan budući da varijable koje su korištene u startu nisu sa sigurnošću precizno utvrđene.

Čineći to, pretpostavljamo da se nesigurnosti varijabli koje koristimo kao input (odnosno, ulazne podatke) u jednadžbama neće pretvoriti u neprihvatljive nesigurnosti krajnjih rezultata (output; izlazni, dobiveni podaci). Nema potrebe za nagađanjem. Ukoliko je preciznost svake izmjerene ili procijenjene vrijednosti poznata, tada se mogu upotrijebiti jednostavna matematička pravila kako bi se dobila preciznost konačnog rezultata. Ova se procedura naziva propagacijom pogrešaka.

## 3.2 Granica detekcije

Metoda granice detekcije (MDL) definira sposobnost mjerne metode da odredi analit u matrici uzoraka, neovisno o njegovom izvoru ili podrijetlu. Obrađujući uzorke npr. razrjeđivanjem, ekstrakcijom, isušivanjem itd. uvodi se varijabla te je neophodno da MDL uzme u obzir tu varijablu.<sup>9</sup>

MDL je metoda koja je često promatrana kao kemijski koncept budući da ovisi od supstance do supstance te se smanjuje ukoliko se analitička metoda poboljša. MDL je statistička metoda u kojoj se procjena vrši na temelju dobivenih/izvađenih podataka. Kao takva, nema nikakvo znanstveno značenje sve dok nije operacijski definirana u terminima mjernih procesa i statističkih metoda za analizu mjerenja. Bez precizne statističke definicije, ne može se odrediti numerička vrijednost za granicu detekcije te ne može ta granica detekcije biti u skladu sa različitim laboratorijima i njihovim određivanjem.<sup>6</sup>

Granica detekcija je definirana kao minimalna koncentracija neke supstance koja može biti izmjerena i zabilježena sa sigurnošću 99% da je koncentracija analita veća od nule i određena je analizom uzorka u dobivenom matriksu koji je sadržavao analit. Neophodno je da u sve korake procesuiranja uzoraka analitičke metode bude uključena domena granice detekcije.

Prije spomenut 'uzorak u dobivenom matriksu' indicira kako MDL može varirati kao funkcija ovisna o samom sastavu uzoraka. Fraza 'sadrži analit' može biti zbunjujuća budući da je procedura dizajnirana za primjenu na širokom rasponu uzoraka, uključujući i praznine reagensa koje ne uključuju sami analit.

Termin granica detekcije je korišten u različitim situacijama kada je bilo potrebno učiniti jasnu razliku. Različite granice detekcije nisu jedinstvene konstante u metodologiji. Ovisе o statističkoj definiciji i o tome kako je procijenjena varijabilnost izmjerenog uzorka niske koncentracije. Također ovisе o iskustvu samog analitičara, kvaliteti kontrolnih procedura korištenih u laboratoriju te mjerama samog matriksa uzoraka. Dakle, dvoje analitičara, u istom laboratoriju i koristeći se istom metodom može dobiti značajne razlike u preciznosti; stoga će se njihove MDL vrijednosti razlikovati. Iz navedenoga slijedi kako objavljene vrijednosti za MDL nemaju primjenu u specifičnim slučajevima, osim za osiguravanje referentnih podataka pomoću kojih bi određeni laboratorij ili analitičar mogli provjeriti posebno izvedeni MDL.<sup>6</sup>

Mjere su sačinjene od minimalno 7 alikvota ( $n \geq 7$ ) pripremljenih otopina koje imaju koncentraciju blizu vrijednosti koja je očekivana granica detekcije. Ovi podaci se koriste za računanje varijanci ponovljenih mjerenja:

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2$$

gdje  $y_i$  ( $i = 1$  do  $n$ ) predstavljaju izmjerene vrijednosti. MDL je:

$$\text{MDL} = s * t_{v,\alpha=0,01}$$

gdje je  $t_{v,\alpha=0,01}$  vrijednost sa 99% sigurnosnom razinom, a standardna devijacija procjenjuje  $\alpha$  za  $v = n - 1$  stupnjeva slobode.  $T$  vrijednost u ovom slučaju 'reže' gornjih 1% od  $t$  distribucije. Za  $n = 7$  i  $v = 6$ ,  $t_{6,0,01} = 3,143$  i procijenjeni MDL je jednak 3.143s.



### 3.3 Granica kvantifikacije

Granica detekcije/ kvantifikacije najmanja je količina analita u uzorku koja se može detektirati / kvantificirati uz odgovarajuću točnost i preciznost. Određivaju se razrijeđivanjem osnovne otopine.

Procjena može biti vizualna, s pomoću omjera signal/ šum ili statistička. Vizualna se procjena može primjeniti i kod neinstrumentalnih i instrumentalnih metoda, uglavnom samo za granicu detekcije, a procjenjuje se najmanji signal koji nedvojbeno može prepoznati.

Statistički se granice detekcije i kvantifikacije mogu odrediti na bazi standardne devijacije signala i nagiba prema jednadžbama:

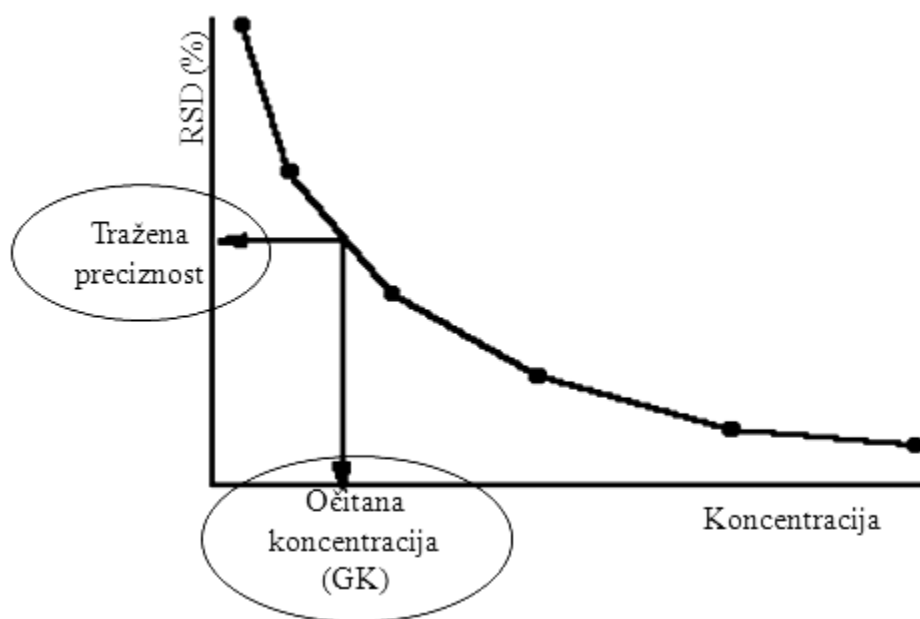
$$GD = \frac{3,3 \cdot \partial}{a}$$

Odnosno  $Kg = 10 \frac{\partial}{a}$  pri čemu je  $a$  nagib, a  $\partial$  standardna devijacija regresijskog pravca.

Ako se zahtjeva da metoda ima zadanu preciznost na granici kvantifikacije, pripremi se više uzoraka poznate koncentracije u području oko moguće granice kvantifikacije, svaki se izmjeri 5 – 6 puta i izračunaju se relativne standardne devijacije za svaku koncentraciju.

Zatim se grafički prikaže odnos RSD-a prema koncentraciji i iz grafa odredi koncentracije na granici kvantifikacije s točno određenom preciznošću kako se vidi na 1. slici. U praksi se obično pripremi uzorak tako određene koncentracije i potvrdi preciznost višekratnim mjerenjem.

Parametar granice kvantifikacije iznimno je važan kod metoda kojima se određuju analiti u tragovima koji i u vrlo niskim koncentracijama mogu štetno djelovati na zdravlje ljudi i okoliš.



**Slika 1.** Određivanje granice kvantifikacije sa zadanom preciznošću

## 4 ANALIZA CENZURIRANIH PODATAKA

Mnogo važnih ekoloških problema se bazira na kemikalijama za koje je očekivano da postoje u vrlo malim koncentracijama ili budu potpuno odsutne. Pod tim uvjetima, skupine podataka mogu uključivati zapažanja zabilježena kao 'nije detektirano' ili 'ispod granice detekcije' (MDL). Za takav podatak se kaže da je cenzuriran.

Cenzurirani podaci su u suštini podaci koji na neki način nedostaju. Vrijednosti koje nedostaju u zapisima podataka nisu rijetkost i ne predstavljaju uvijek veliki problem. Ukoliko bi 50 uzoraka bilo sakupljeno, a od 50 sakupljenih bi nasumično 5 bilo uništeno, mogli bismo provesti analizu iako je preostalo 45 uzoraka. U ovom slučaju, uništeni uzorci bi mogli biti nadomješteni bez značajnijeg utjecaja na sami rezultat ukupne serije uzoraka. Problem sa cenzuriranim podacima jest da oni nisu izabrani nasumično. Oni nedostaju na jednom kraju distribucije te ne možemo ići naprijed ukoliko ti podaci nisu nikad postojali jer bi to utjecalo na pristranost u našim završnim rezultatima.<sup>9</sup>

Neobična osobina koja se odnosi na cenzurirane podatke u vezi kakvoće vode jest ta da cenzurirani podaci ne nedostaju uvijek. Naime, određene brojevnne vrijednosti mogu biti izmjerene, no kemijski analitičar je odredio da je vrijednost ispod granice detekcije (MDL) te je zapisao to u izvještaj kao < MDL umjesto da je zapisao brojevnu vrijednost. Preporuka u praksi je zabilježiti sve vrijednosti skupa sa oznakom o njihovoj preciznosti prilikom mjerenja te prepustiti analitičaru podataka da sam procijeni koju će težinu ti brojevni podaci imati u završnoj interpretaciji podataka.

Nažalost, ne postoji općeprihvaćena shema za zamjenu cenzuriranih opažanja za neke proizvoljne vrijednosti. Zamjenjujući cenzurirana opažanja sa nulom ili 0,5 MDL omogućuje se procjena za varijacije kojima je niski utjecaj na pristranost te za varijacije kojima je visoki utjecaj na pristranost. Pristranost varijacija raste proporcionalno sa vrijednosti MDL.

## 5 EKSPERIMENTALNI DIZAJN

Nije problem navesti nekoliko konačnih eksperimenata u kojima su rezultati inicijalno bili jasni i bez dodatne statističke analize. To se može dogoditi samo u onim slučajevima u kojima postoji odlični eksperimentalni dizajn koji obično uključuje izravnu usporedbu i ponavljanje. Direktna usporedba podrazumijeva isključenje diskutabilnih/nejasnih čimbenika. Ponavljanje podrazumijeva da je vjerodostojnost povećana mogućnošću dokaza kako povoljan rezultat nije samo puka sreća (ukoliko postoji sumnja, rezultati se mogu ponovno dobiti.). S druge strane, postoje eksperimenti čiji su rezultati nejasni i nakon podvrgavanja analize podataka. Takvi eksperimenti su rezultat neučinkovitog eksperimentalnog dizajna.

Statistički eksperimentalni dizajn se odnosi na program rada za upravljanje postavkama od nezavisnih varijabli koje trebaju biti proučene. Druga vrsta eksperimentalnog dizajna se suočava s izgradnjom i radom cijelog eksperimentalnog aparata.<sup>11</sup> Što su teži i skuplji zahvati u pojedinim dizajnim, to taj dizajn više može osigurati efikasnost u traženom području.

Postoji mnogo vrsta eksperimentalnog dizajna. Neke vrste uključuju: upotrebu jednog po jednog faktora, uparene usporedbe, proučavanje faktora na dvije razine, faktore u fragmentima. Učinkovit dizajn daje puno informacija uz malo uloženog posla. Loše izvedeni dizajn kao rezultat daje malo informacija uz puno uloženog posla.

Dizajni koji se odnose na faktore koji se uzimaju u obzir jedan po jedan se smatraju lošim primjerima dizajna te se nastoji statističare usmjeriti na odabiranje dizajna koji uključuju proučavanje faktora na dvije razine te upotreba fragmentiranih faktora kao modele eksperimentalnog dizajna.<sup>11</sup>

Jednostavni eksperimentalni dizajn vodi jednostavnim metodama za analiziranje podataka. Dobar eksperimentalni dizajn se temelji na principima randomizacije, izravne komparacije, replikacije i korištenja kalupa u određenim periodima. Trebao bi osigurati jednostavnost u računanju određenim podacima. Treba težiti i fleksibilnosti, odnosno mogućnosti izvođenja eksperimentalnog dizajna u određenim segmentima i sekvencijama.

Ukoliko je eksperiment pod vodstvom osoba koje su nedovoljno upoznate s korištenjem potrebnih postupaka, tada može biti teško garantirati vjerodostojnost dobivenih rezultata unutar eksperimentalnih uvjeta. Ukoliko je određeni eksperiment izvođen pod posebno kreiranim uvjetima, potrebno je zadirati u te stvorene uvjete što je manje moguće.

U znanstvenom radu, osobito u fazama koje obuhvaćaju pripremne radove za samo istraživanje, od velike je važnosti zadržati fleksibilnost. Inicijalni dio eksperimenta može potvrditi obećavajući značaj cijele cjeline eksperimenta. Stoga nije dobar izbor ukoliko se veliki eksperimenti moraju završiti prije negoli određeni rezultati pojedinih dijelova samog eksperimenta budu prihvaćeni.

## **5.1 Principi eksperimentalnog dizajna**

### **5.1.1 Komparativni dizajn**

Ukoliko dodamo supstancu  $X$  procesu i output (izlazni podaci) se poboljšaju, primamljivo je zaključiti da se napredak dogodio zbog dodavanja supstance  $X$ . No ovo opažanje može u potpunosti biti pogrešno.  $X$  ne mora uopće imati nikakve važnosti u procesu. Njegovo dodavanje u proces je moglo biti slučajno skupa sa promjenom koja se dogodila tijekom procesa u kojeg je dodan. Način na koji će se izbjeći pogrešni zaključci o  $X$  jest provođenje komparativnih eksperimenata. Paralelno pokrenuti procesi, jedni s dodavanjem  $X$ , drugi bez dodavanja  $X$ .

Pasivno promatranje procesa je puno manje pouzdano nego direktna komparacija. Ukoliko želimo znati što će se dogoditi s procesom ako promijenimo nešto, moramo promatrati proces u trenutku kada je faktor aktivno bio promijenjen. Nažalost, postoje i situacije u kojima moramo prihvatiti da se na sistem ne može voljno utjecati. No, osnovni princip jest da bismo trebali, kada god je to moguće raditi dizajnirane i kontrolirane eksperimente. Pod ovim se podrazumijeva kako se trebaju utvrditi specifični eksperimentalni uvjeti ( temperatura, količina dodanoga  $X$ , protok...).

### **5.1.2 Replikacija**

Replikacija (ponavljanje) osigurava internu procjenu nasumičnih eksperimentalnih pogrešaka. Utjecaj pogrešaka na krajnji ishod je procijenjena izračunom standardne pogreške. Ukoliko su svi ostali parametri jednaki, standardna pogreška će se smanjiti kako se brojevi opažanja i replikacija budu povećavali. To znači da preciznost komparacije može biti povećana ukoliko se poveća broj eksperimentalnih izvođenja. Povećana preciznost vodi većoj vjerojatnosti ispravne detekcije malih razlika između postupaka. Ponekad je bolje povećati broj postupaka na način da se povećaju brojevi replikacija, umjesto da se dodaju neke nove postavke u određene postupke.

Ponavljanje postupaka je potrebno kako bi se procijenila nasumična eksperimentalna pogreška. 'Ponavljanje' znači da su postavke od  $X$  iste u dvije ili više ponavljanih postupaka. Podrazumijeva da identične postavke od  $X$  zadržavaju varijacije koje utječu na određene mjere.

### **5.1.3 Randomizacija**

Da bismo potvrdili vrijednost procijenjene eksperimentalne pogreške, oslanjamo se na princip randomizacije (slučajnosti). To vodi procjeni varijance koja nije podložna vanjskim utjecajima i razlikama određenih procesa. Takva procjena označava slobodu od sustavnih utjecaja iz inače drugačijih, nekontroliranih varijacija.

Randomizacija također pomaže u eliminaciji serijski povezanih pogrešaka, zatim u eliminaciji nezapaženih varijabli koje mogu predstavljati pogrešku te eliminaciji nedosljednih podataka.

Postavlja se pitanje mora li se uvijek pristupiti postupku randomizacije? U slučajevima u kojima randomizacija blago utječe na kompliciranost izvođenja postupaka, treba uvijek vršiti randomizaciju. U slučajevima u kojima bi randomizacija proces eksperimenta jako zakomplicirala ili ga učinila nemogućim, ali se može obaviti logična procjena o postojanju varijabli koje bi mogle stvarati problem; tada se vrši eksperiment bez randomizacije. Ukoliko se dogodi situacija u kojoj se smatra da je proces toliko nestabilan, a da se koristi randomizacija bi to onemogućilo sam eksperiment; tada je preporučljivo uopće ne izvoditi eksperiment. Tada je dodatnu pažnju uputnije usmjeriti stabilizaciji samog procesa.<sup>10</sup>

### **5.1.4 Blokiranje**

Upareni *t-test* predstavlja koncept blokiranja. Blokiranje je sredstvo za reduciranje eksperimentalnih pogrešaka. Osnovna ideja jest razdvajanje potpunih eksperimentalnih skupova u podskupove koji su homogeniji što je više moguće. Na ovaj način efekt faktora koji predstavljaju problem u eksperimentu i djeluju na cijeli sustav može biti eliminiran. To vodi puno senzitivnijoj analizi budući da će eksperimentalna pogreška biti procjenjivana u svakom bloku (odnosno, podskupini).<sup>12</sup>

## 6 KORELACIJA PODATAKA

Dvije varijable su izmjerene, a dobiveni podaci sugeriraju da između njih postoji linearna povezanost. Mjera koja kvantificira snagu linearne povezanosti između dvije varijable se naziva koeficijent korelacije. Dakle, koeficijent korelacije pokazuje stupanj linearne zavisnosti između varijabli. Kreće se u rasponu od -1 do +1. Što je koeficijent korelacije bliže 1 ili -1, veća je korelacija između varijabli. Ukoliko su varijable nezavisne, tada je koeficijent korelacije 0. Pretpostavke kod računanja koeficijenta korelacije su: linearna zavisnost između dviju varijabli x i y, kontinuirane slučajne varijable te obje varijable moraju imati normalnu razdiobu.

Mora se pridodati poseban oprez kako se ne bi korelacija zamijenila za uzročnost. Korelacija može, ali ne mora nužno indicirati uzročnosti. Promatrajući da se x povećava isto kada se povećava i y ne znači da promjene u x uzrokuju povećanje u y. Moguće je da se i x i y promijene, ali ako rezultat treće varijable z.<sup>12</sup>

### 6.1 Kovarijacija

Mjera za linearnu ovisnost između dvije varijable x i y je kovarijanca između x i y. Kovarijanca pokazuje koliko se dvije varijable mijenjaju zajedno. To je različito od varijance, koja pokazuje koliko se jedna varijabla mijenja. Kovarijanca postaje više pozitivnom za svaki par vrijednosti koji se razlikuje od njihovih srednjih vrijednosti u istom smjeru; te postaje više negativna za svaki par vrijednosti koji se razlikuje od njihovih srednjih vrijednosti u suprotnim smjerovima.

Način za dobivanje kovarijance između x i y je:

$$\text{COV}(x, y) = \frac{\sum(X_i - \eta_x)(y_i - \eta_y)}{N}$$

### 6.2 Serijska korelacija

Kada su podaci uzeti sekvencijalno, postoji tendencija da se razmatraju oni koji su bliže (vremenski i prostorno) od onih koji su udaljeniji. Temperaturne krivulje, npr. mogu pokazati velike varijacije tijekom cijele godine, dok su temperature uzete svakih sat vremena tijekom jednog dana skoro iste. Pojedine automatizirane nadzorne opreme vrše mjere toliko često sa susjedni dobiveni rezultati gotovo identični. Ta tendencija za susjedna opažanja koja su u

korelaciji jest serijska korelacija ili autokorelacija. Jedna od mjera u serijskoj ovisnosti je autokorelacijski koeficijent koji nije bitno drugačiji u odnosu na prije opisani korelacijski koeficijent.

## **7 PROCJENA RIZIKA**

Proces procjene rizika obuhvaća postupke identificiranja opasnosti i kvantificiranja rizika na ljudsko zdravlje i oštećenje ekosustava.

Procjena rizika na okoliš uključuje:

- procjenu rizika na ljudsko zdravlje,
- procjenu ekološkog i ekotoksikološkog rizika,
- procjenu rizika od specifičnih industrijskih primjena kojima se ispituju pokazatelji promjene u ljudima, biološki uzorcima i ekosustavima.

Rizik, općenito, ovisi o:

- količini kemikalije u različitim medijima okoliša (tlu, vodi, zraku),
- vremenu u kojemu su osoba ili ekološki primatelj bili izloženi onečišćenom okolišu,
- toksičnosti kemikalije.

S ciljem identifikacije i stupnjeva spojeva zabrinjavajućih po okoliš razvijeni su različiti pristupi procjeni rizika radi donošenja odgovarajućih propisa i programa njihova praćenja u okolišu. <sup>14</sup>

Procjena ekološkog rizika obuhvaća sljedeće korake:

- utvrđivanje problema,
- identificiranje opasnosti,
- procjenu ispuštanja u okoliš,
- procjenu izloženosti riziku,
- procjenu posljedica,
- ocjenu rizika.



## 7.1 Procjena učinka

### 7.1.1 Karakterizacija

Sljedeći dio procesa procjene rizika odnosi se na opasnost vezanu uz određeni spoj. S obzirom na njegovu prirodu i uporabu, spoj može imati različite štetne učinke na ekosustav i zdravlje ljudi. Takvi se učinci obično izražavaju tzv. PBT (postojanost, bioakumulacija, toksičnost) pristupom, pri čemu svaki od ta tri izraza označava odgovarajuće mjerljivo sredstvo. Na PBT pristupu temeljen je prvi popis prioriternih organskih tvari, ali i mnoge nedavno predložene sheme prioriternosti onečišćivala. Naknadno je pojam *toksičnost* proširen s obzirom na ogući utjecaj kemikalije na kancerogenost, mutagenost, teratogenost ili poremećaj rada žljezda s unutarnjim lučenjem.<sup>14</sup>

Procjena vrijednosti predviđenih koncentracija ispod kojih se ne očekuju štetni učinci (PNEC), prvenstveno se izvodi iz rezultata laboratorijskih ispitivanja jedne vrste ili iz ispitivanja modela ekosustava. Za određivanje koncentracije pri kojoj nema vidljivih učinaka (NOEC) ili najmanje koncentracije pri kojoj su učinci vidljivi (LOEC) služe dostupni podaci o ekotoksičnosti, koji se uglavnom odnose na vodene organizme.<sup>14</sup>

### 7.1.2 Modeliranje

Kada je izravno mjerenje izloženosti riziku ili biološko praćenje nepredvidivo, pristupa se procjeni rizika modeliranjem. Mnogi modeli omogućuju istovremene procjene izloženosti i rizika, pružajući mogućnost kvantificiranja rizika zbog kemijske izloženosti.

Utvrđene su njihove sljedeće prednosti<sup>14</sup>:

- omogućuju predviđanje potencijalne izloženosti u budućnosti ili scenarije pretpostavljenog ispuštanja,
- omogućuju kombiniranje različitih vrsta onečišćivala i izvora emisija,
- stupanj složenosti modela može se odrediti prema zahtjevima procjene,
- razmatraju izloženost u više smjerova i putova,
- smanjuju potrebu za skupljim programima nadziranja okoliša.

## 8 PRIMJENA STATISTIKE U ANALIZI OKOLIŠA

Statistika je nauka o principima prikupljanja, organizacije, analize, sažetog prikaza i interpretacije podataka dobivenih promatranje ili mjerenjem vrijednosti varijable osnovnog skupa uzoraka ili jedinke.

Metode statističke analize procesa primjenjuju se i u procesu upravljanja zaštitom okoliša.

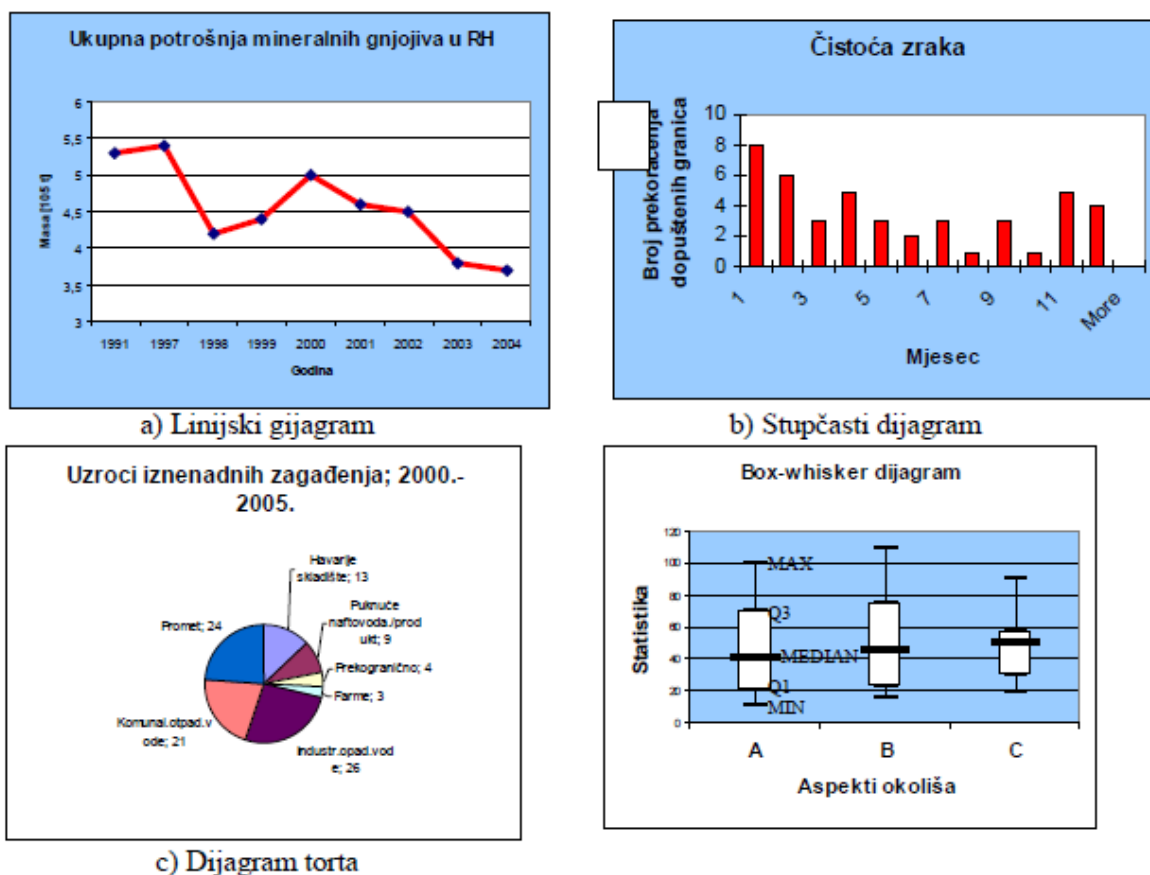
U procesu praćenja odvijanja procesa koriste se ove statističke tehnike:

- tablični prikaz prikupljenih podataka,
- grafički prikazi prikupljenih podataka,
- statistika sažimanja, tj. računanje određene statističke veličine koja opisuje prikupljeni skup podataka.

### 8.1 Deskriptivna statistika

U grafičkom prikazu podataka za potrebe sustava upravljanja zaštitom okoliša primjenjuju se, kao i u sustavu upravljanja kvalitetom: linijski, stupčasti i torta dijagrami, a posebnu je primjenu našao B&W dijagram (Box and Whiskers dijagram).<sup>3</sup>

Primjeri navedenih dijagrama nalaze se na slici 2:



Slika 2. Grafički prikazi podataka

Budući se B&W dijagram rijetko primjenjuje, a veoma je pogodan za prikaz podataka u području zaštite okoliša, ovdje mu je posvećena posebna pozornost. Nažalost ovaj prikaz u Excel-u nije tako jednostavan kao prikazi ostalih dijagrama jer nije programiran već se moramo poslužiti zaobilaznim putovima kako bi ga i u Excel-u nacrtali.

Ovaj grafički prikaz istovremeno prikazuje pet parametara procesa i to: medijan, najveću vrijednost MAX, najmanju vrijednost MIN, prvi kvartil  $Q_1$  i treći kvartil  $Q_3$ .

U zaštiti okoliša većina procesa ne slijedi normalnu raspodjelu. Krivulja funkcije vjerojatnosti najčešće nema simetričan zvonolik oblik, tj. ne slijedi Gaussovu krivulju. Osim toga, rasipanja podataka su obično jako velika. U tom slučaju asimetrična sredina ne opisuje stvarno „težište“ svih podataka. Na rasipanje podataka u procesu ne djeluju samo slučajne veličine već se javljaju i neke koje imaju predvidljiv trend, tako da vrijednosti nisu slučajno raspodijeljene oko neke aritmetičke srednje vrijednosti, već se gomilaju oko neke vrijednosti koja nije na sredini između MAX i MIN vrijednosti. To znači da srednja vrijednost procesa, računata kao aritmetička sredina svih vrijednosti nije jednaka medijanu; tj. vrijednost srednjeg elementa u nizu podataka sortiranih po veličini (rastući ili padajući niz).<sup>3</sup>

Iz B&W dijagrama jasno se vide vrijednosti MAX i MIN, tj. područje u kojem se nalaze svi prikupljeni podaci.

Vrijednost prvog kvartila  $Q_1$  predstavlja granicu područja omeđenu s jedne strane vrijednošću MIN, a s druge s  $Q_1$  unutar kojeg se nalazi 25% svih podataka. U području od MIN do  $Q_3$  nalazi se 75% svih podataka. Između  $Q_3$  i MAX nalazi se preostalih 25% podataka.

Kada se iz nekog skupa podataka žele izdvojiti bitna obilježja procesa pristupa se računanju statistike tog skupa. Statistika se sastoji u računanju srednje tendencije procesa i rasipanja vjerojatnosti parametara procesa oko srednje vrijednosti.

Kao mjera centralne tendencije primjenjuje se:

- aritmetička sredina (u Excel-u je računa funkcija *AVERAGE* ili *MEAN*),
- medijan (u Excel-u ga računa funkcija *MEDIAN*),
- mod (u Excel-u ga računa funkcija *MODE*).

Aritmetička sredina predstavlja težište svih vrijednosti, a računa se kao zbroj svih vrijednosti podijeljeno s brojem vrijednosti (veličinom skupa). Najčešće se ova vrijednost uzima za predstavnika cijelog skupa.

Medijan je vrijednost srednjeg elementa skupa kad se sve vrijednosti sortiraju po veličini. Ako se na primjer skup sastoji od 101 podatka onda je medijan tog skupa jednak vrijednosti

51. elementa skupa sortiranog po veličini. U procesima upravljanja zaštitom okoliša, medijan se u odnosu na aritmetičku srednju vrijednost češće koristi za opis procesa.

Mod je vrijednost s najvećom učestalošću u skupu podataka. Poznato je da svaki proces prati rasipanje vrijednosti pojedinih njegovih parametara zbog neprekidnog djelovanja slučajnih veličina koje imaju utjecaj na taj parametar procesa. Upravo u zaštiti okolišarasipanja su jako velika u odnosu na današnje proizvodne procese.

Potrebno je kontrolirati srednju vrijednost i rasipanje kako se ne bi prekoračile granice dopuštenih vrijednosti tog parametra.

Kvantitativnu mjeru rasipanja parametara procesa opisuju statistike<sup>3</sup>:

- Raspon (U Excel-u  $R=MAX-MIN$ )
- Kvantil (U Excel-u funkcija *PERCENTIVE*)
- Varijanca (U Excel-u funkcija *VAR*)
- Standardno odstupanje (U Excel-u funkcija *STDEV*)

Raspon je razlika između najveće vrijednosti MAX i najmanje vrijednosti MIN u skupu podataka. Na žalost u Excelu-u nije programirana funkcija raspon. Potrebno je prethodno pronaći najveću i najmanju vrijednost u skupu podataka.

Kvantil je skup podataka kojeg sortiramo po rastućim vrijednostima te se kao takav može podijeliti na q jednakih dijelova.

Varijanca predstavlja aritmetičku srednju vrijednost kvadrata odstupanja vrijednosti elementa skupa podataka oko njegove srednje vrijednosti. Što je vrijednost varijance veća to je rasipanje vrijednosti elemenata od srednje vrijednosti, odnosno aritmetičke sredine.

Standardno odstupanje (standardna devijacija) je mjera rasipanja podataka oko srednje vrijednosti skupa, a definirana je kao pozitivna vrijednost kvadratnog korijena varijance.

## **8.2 Ispitivanje hipoteze**

Ekologija je znanost o zaštiti životne sredine, vode, zraka i drugih prirodnih resursa od onečišćenja u cilju skladnog života čovjeka i prirodne sredine. Ekologija brine o zaštiti zraka, vode i drugih prirodnih resursa od onečišćenja i njenih učinaka. Ekologija se bavi relativnim a

ne apsolutnim istinama. Pojave u prirodi se promatraju. Na temelju promatranja stvaraju se teorije koje nakon prosudbe i njene potvrde postaju ljudsko znanje.

Prvo se definira tzv. nulta hipoteza ili istraživačka hipoteza. To je izjava koja se odnosi na uzrok prirodnog fenomena koji se razmatra.

Nakon što je definirana nulta hipoteza definira se alternativna hipoteza. Ako se, recimo, nultom hipotezom definira vrijednost aritmetičke sredine procesa, onda za alternativnu hipotezu postoje tri mogućnosti:

- aritmetička srednja vrijednost prirodnog fenomena nije jednaka onoj iz nulte hipoteze,
- aritmetička srednja vrijednost prirodnog fenomena veća je od one iz nulte hipoteza,
- aritmetička srednja vrijednost prirodnog fenomena manja je od one iz nulte hipoteze.

Nulta hipoteza odbacuje se s određenom razinom povjerenja ako se slučajnim uzorkovanjem prikupi dovoljno čvrstih dokaza koji potvrđuju alternativnu hipotezu. Statistički se smatraju dovoljno čvrsti dokazi ako se, npr. za prvi slučaj definiranja alternativne hipoteze, aritmetička sredina uzorka razlikuje od aritmetičke sredine definirane u nultoj hipotezi za više od  $\pm 2$  standardna odstupanja tog istog uzorka. U tom se slučaju s razinom povjerenja  $P = 95\%$  odbacuje nulta hipoteza.

### **8.3 Regresijska analiza**

U sustavu upravljanja zaštitom okoliša izuzetno je važno poznavati međusobne veze i ovisnosti pojedinih parametara koji utječu na promatrani proces. Tu vezu između dva ili više parametara moguće je istražiti regresijskom analizom. Jakost međusobne veze dva parametra opisuje koeficijent korelacije.<sup>15</sup> Koeficijent korelacije  $r$  može imati vrijednost u intervalu od  $-1$  do  $+1$ . Ako je  $r = +1$ , u raspršenom dijagramu točke definirane parovima vrijednosti oba parametra leže na pravcu koji ima pozitivan koeficijent smjera. Povećanjem vrijednosti jednog parametra povećava se vrijednost drugog parametra, a međusobna veza definirana je jednadžbom pravca. Obrnuto, ako je  $r = -1$ , točke u raspršenom dijagramu isto leže na pravcu, ali on ima negativan koeficijent smjera. Povećanjem vrijednosti jednog parametra smanjuje se vrijednost drugog parametra i obratno. Ako je apsolutna vrijednost koeficijenta korelacije u području  $0 < r \leq 0.65$  kaže se da između parametara postoji slaba statistička veza. Kad je  $r = 0$  kaže se da ne postoji statistička veza između parametara. Za  $0.65 < r \leq 0.85$  kaže se da postoji

srednje jaka statistička veza između promatranih parametara procesa. Konačno za  $0.85 < r \leq 1$  kaže se da postoji jaka korelacijska veza.

U Excelu-u koeficijent korelacije  $r$  računa funkcija *CORREL*.<sup>3</sup>

#### **8.4 Uzorkovanje**

U procesu donošenja odluka potrebno je prikupiti dovoljno pouzdanih činjenica. Do činjenica se dolazi promatranjem procesa, uzorkovanjem i mjerenjima, iz literature i slično. Kod uzimanja uzoraka bitno je izvršiti izbor jedinica na slučajan način, tj. da svaka jedinka ima istu mogućnost izbora. Ukoliko to nije slučaj u rezultat ispitivanja unosi se sustavna pogreška. Drugi važni parametar uzorka je veličina uzorka. Veličina uzorka uvijek je kompromis utrošenih resursa i razine povjerenja dobivenih rezultata.

Kod primjene uzorkovanja u analizi procesa obično se pristupa uzimanjem i dodatnih uzoraka kako bi se dokazala ponovljivost dobivenih rezultata i izbjegle sustavne pogreške.

#### **8.5 Simulacije**

Razvojem informatičke tehnologije i povećanjem ljudskog znanja o prirodnim pojavama i procesima omogućen je matematički opis procesa, a time i računalna simulacija prirodnog fenomena ili procesa. Promjenom parametara u modelu se mogu simulirati različita stanja tijekom procesa i analizirati rezultati njihovih djelovanja.

Razvijeni su mnogi matematički modeli za pojedine prirodne fenomene.

#### **8.6 Vremenske serije**

U procesu odlučivanja bitno je pravovremeno otkriti trendove u procesu kako bi se na vrijeme mogle poduzeti popravne ili preventivne radnje. Upravo analizom vremenskih serija moguće je uočiti trendove i tako predvidjeti buduće ishode ako se proces nastavi istim trendom.

## **9 ZAKLJUČAK**

Statistika je nauka o principima prikupljanja, organizacije, analize, prikaza i interpretacije podataka dobivenih promatranjem ili mjerenjem vrijednosti varijable osnovnog parametra.

Potrebno je neprekidno poboljšavati procese, odnosno mjeriti, analizirati, nadzirati procese te provoditi postupke u cilju otklanjanja ili smanjenja rizika na prihvatljivu razinu. Isto tako bitno je pravovremeno otkriti trendove u procesu kako bi se na vrijeme mogle poduzeti popravne ili preventivne radnje.

## **10 POPIS KRATICA**

MDL – metoda granice detekcije

PNEC – predviđena koncentracije ispod koje se ne očekuju štetni učinci

NOEC – koncentracija pri kojoj nema vidljivih učinaka

LOEC – najmanja koncentracija pri kojoj su učinci vidljivi



## 11 LITERATURA

1. Box, G. E. P. (1974). "Statistics and the Environment," *J. Wash. Academy Sci.*, 64, 52–59.
2. Berthouex, P. M., W. G. Hunter, and L. Pallesen (1981). "Wastewater Treatment: A Review of Statistical Applications," *ENVIRONMETRICS 81—Selected Papers*, pp. 77–99, Philadelphia, SIAM.
3. Antun Benčić, Krešimir Buntak, Tihomir Bbabić, Statistički alati za analizu aspekata okoliša i njihovog utjecaja u sustavu upravljanja okolišem, [http://www.kresimir-buntak.com/Radovi/2012/16\\_Statisticki\\_alati\\_za\\_analizu.pdf](http://www.kresimir-buntak.com/Radovi/2012/16_Statisticki_alati_za_analizu.pdf), 1.8.2015.
4. Kateman, G. and L. Buydens (1993). *Quality Control in Analytical Chemistry*, 2nd ed., New York, John Wiley.
5. Mandel, J. (1964). *The Statistical Analysis of Experimental Data*, New York, Interscience Publishers.
6. American Chemical Society, Committee on Environmental Improvement (1983). "Principles of Environmental Analysis," *Anal. Chem.*, 55(14), 2210–2218.
7. ASTM (1988). *Standard Practice for Determination of Precision and Bias of Applicable Methods of Committee D-19 on Water*, D2777–86, ASTM Standards of Precision and Bias for Various Applications, 3rd ed.
8. Berthouex, P. M. and L. C. Brown (1994). *Statistics for Environmental Engineers*, Boca Raton, FL, Lewis Publishers.
9. Hahn, G. A. and S. S. Shapiro (1967). *Statistical Methods for Engineers*, New York, John Wiley.
10. Box, G. E. P. (1990). "Must We Randomize?," *Qual. Eng.*, 2, 497–502.
11. Box, G. E. P., W. G. Hunter, and J. S. Hunter (1978). *Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building*, New York, Wiley Interscience
12. Chatfield, C. (1983). *Statistics for Technology*, 3rd ed., London, Chapman & Hall
13. Miller, J. C. and J. N. Miller (1984). *Statistics for Analytical Chemistry*, Chichester, England, Ellis Horwood Ltd.
14. Ašperger, D. et al. (2013.) *Analitika okoliša*. Zagreb: HINUS i Fakultet kemijskog inženjerstva i tehnologije, 375-383.
15. Draper, N. R. and H. Smith, (1998). *Applied Regression Analysis*, 3rd ed., New York, John Wiley.

## **ŽIVOTOPIS**

Ante Zdrilić, rođen sam 28. siječnja 1993. godine u Zadru. Završio sam osnovnu školu Poličnik te sam potom nastavio školovanje u Prirodoslovno grafičkoj školi u Zadru. Srednju školu završio sam 2011. godine. Iste godine upisao sam se na Fakultet kemijskog inženjerstva i tehnologije u Zagrebu. Završni rad izradio sam na Zavodu za analitičku kemiju pod vodstvom mentora prof. dr. sc. Tomislava Bolanče.