Univerzitet u Novom Sadu
Prirodno-matematički fakultet
Departman za matematiku i informatiku

Jovana Arsić

Primena Benfordovog zakona u reviziji

Master rad

Mentor

Prof. dr Nataša Spahić

2016, Novi Sad

Predgovor

Master rad pod nazivom "Primena Benfordovog zakona u reviziji" nastao je usled ličnih aspiracija ka naglašavanju zastupljenosti i primeni matematičkih zakonistosti u polju finansija. Smernica pri izradi rada bile su reči istaknutog matematičara Villim Fellera : "U svakoj naučnoj disciplini treba pažljivo razlučiti tri aspekta teorije:

- Formalno logički sadržaj
- Intuitivna pozadina
- Primena."

Cilj rada je ostvarivanje uvida u jedan praktičan, jednostavan i efikasan proces procenjivanja prisustva/odsustva nepravilnosti primenjivog na različitim skupovima podataka. Kao odabrano polje primene u okviru finansija nametnuo se revizijski proces finansijskih izveštaja preduzeća koja posluju na teritoriji Republike Srbije. Dodatni motiv analize finansijskih izveštaja ogleda se u njihovoj dostupnosti javnosti preko baze Agencije za privredne registre, za razliku od većine finansijskih podataka čija je transparentnost ograničena. Rad je namenjen kako studentima, interesentima za polje revizijskih procesa i procesa finansijskog kontrolinga, tako i licima koja u sklopu postojećih informacionih programa koriste analize zasnovane na Benfordovoj zakonitosti, kako bi se bolje upoznali sa matematičkom osnovom problema. Dodatno, rad može poslužiti licima koja vrše kontolne funkcije podataka iz različitih oblasti , a želeli bi da nauče jednostavan metod (primenom MS Office Excel) čiji rezultati mogu poslužiti kao intuicijska osnova procenjivanja istinitosti podataka i smernica za dalja analiziranja.

Sadržaj

Glava 1

Uvod

Poslovni subjekti krajnju evidenciju svojih poslovnih aktivnosti predstavljaju u vidu finsijskih izveštaja. Svrha finansijskog izveštavanja može se podeliti na četiri segmenta:

- Pružanje informacija rukovodstvu subjekta prilikom donošenja daljih odluka o poslovanju.
- Pružanje informacija poreskim institucijama na osnovu kojih se formiraju poreska zaduženja.
- Pružanje informacija kontrolnim institucijama države.
- Pružanje informacija ostalim poslovno pravnim subjektima o stvarnom stanju subjekta.

Finansijki izveštaji bi trebalo da verodostojno predstavljaju finansijsko stanje i poslovne aktivnosti subjekta. On predstavlja izvor informacija za donošenje poslovnih odluka,te je od presudnog je značaja da bude pouzdan, istinit, objektivan i pošten. Upravo takvu, nepristrasnu ocenu, o istinitosti finansijskih izveštaja daju nezavisni revizori. Posmatranjem revizorske profesije kao spone koja stvara poverenje između rukovodstva preduzeća zaduženog za formiranje finansijkih izveštaja i korisnika informacija sadržanih u izveštajima, naznačava se odgovornost i značajnost revizorskih aktivnosti.

Konkurentnost u poslovnim krugovima, dinamičnost poslovanja i složeni sistemi poslovanja koji uključuju baze podataka nametnuli su ozbiljne zadatke revizorima prilikom utvrđivanja tačnosti podataka. Ujedno, informacione tehnologije postale su glavni saveznik u sprovođenju teško uočljivih malverzacija nad podacima. Kako bi ostali korak ispred nekorektnih i malicioznih radnji poslovnih subjekata, revizori su bili primorani na pronalaženje alternativnih načina procenjivanja verodostojnosti finansijskih izveštaja.
Odgovor na pitanje da li određeni skup podataka sadrži nedoslednosti ili nelogičnosti potražili su uz pomoć matematičkog fenomena vezanog za raspodelu cifara na vodećoj poziciji broja, poznatog pod nazivom Benfordov zakon.

Benfordov zakon, zakon vodeće cifre tvrdi da u velikom broju skupova podataka verovatnoća da se na prvoj poziciji broja nađe jedna od devet cifara ne iznosi 1/9, što bi se intuitivno dalo pretpostaviti. Odnosno, tvrdi da će trećina podataka počinjati cifrom 1, dok sa porastom vrednosti cifre verovatnoća njene pojave na početnoj poziciji se smanjuje.
Benfordov zakon dugo je postojao kao uočeno svojstvo skupova iz različitih oblasti, ali nisu se u potpunosti mogli precizirati razlozi njegovog pojavljivanja, kao ni moguće primene.

Kraj 20. veka doneo je ekspanziju u proučavanju Benfordovog zakona, što je dovelo do matematičkog dokazivanja fenomena od strane (Teodor P. Hill, 1996) i proširenja polja primene. Revizorska profesija postala je bogatija za nove metode ispitivanja istinitosti finansijskih izveštaja baziranih na principu opadajućeg redosleda početne cifre. Dodatni doprinos u primeni Benfordovog zakona ostvaren je zahvaljujući američkom naučniku Mark J. Niginiju, koji je svoj rad usmerio ka praktičnoj primeni zakona u polju revizije, uključujući razvoj jednostavnih postupaka za procenjivanje validnosti podataka.

Kroz rad se ostvaruje uvid u istoriju Benfordovog zakona, teorijsku pozadinu i oblasti u kojima se upotreba fenomena vodeće cifre pokazala kao najefikasniji metod otkrivanja nepravilnosti.
U skladu sa temom rada, posebna pažnja posvećena je procesu procenjivanja validnosti finansijskih izveštaja zasnovanom na Benfordovom zakonu. Kao mera istinitosti evidentiranja podataka u finansijskim izveštajima koristi se Financial Statement Divergence Score (FSD).

Celokupan proces revizijske provere istinitosti finansijskih izveštaja uz primenu FSD mere prikazan je kroz uporednu analizu finansijskih izveštaja kompanija "Imlek" i "Apatinska pivara". Dodatno, sprovedene su veštačke izmene na podacima u cilju potpunog razumevanja fenomena, postupaka testiranja kao i izvođenja zaključaka.

Glava 2

Definicije i pojmovi

Značajne cifre broja
Značajne cifre broja su sve njegove cifre počevši od prve nenula cifre. Dodatno, nula se smatra značajnom cifrom ako se nalazi između dve nenula cifre i ako čuva decimalno mesto.

Vodeće cifre broja
Vodeća cifra broja predstavlja prvu nenula cifru sa leve strane broja. U radu se koristi navedena definicija, ako nije drugačije naznačeno. Pojam vodeće cifre može uopštiti na više od jedne vodeće cifre. Vodeća cifra broja 1562 je cifra 1, dok možemo reći da su vodeće cifre i 15, 156, 1562.

Klasična definicija verovatnoće
Neka je $\Omega = \{\omega_1, \omega_2, ..., \omega_n\}$ skup svih jednakoverovatnih elementarnih događaja (svih mogućih ishoda eksperimenta) koji su međusobno nesaglasni, neka je $A = \{\omega_{i_1}, ..., \omega_{i_m}\}$ događaj koji se sastoji od m elementarnih događaja koji imaju osobinu definisanu sa A. Verovatnoća događaja A jednaka je:

$$P(A) = \frac{m}{n}$$

Broj m predstavlja broj povoljnih ishoda koji doprinose realizaciji događaja A, a n je broj svih mogućih ishoda eksperimenta

Definicija verovatnoće (frekvencija dešavanja)
Verovatnoća nekog događaja A, izražava se formulom:

$$P(A) = \lim_{n\to\infty} \frac{N(A)}{N}$$

gde je:
$P(A)$ verovatnoća događaja A

$N(A)$ broj ishoda koji zadovoljavaju uslov definisan sa A
N ukupan broj ishoda eksperimenta
Ovaj pristup podrazumeva ponavljanje eksperimenta dovoljan broj puta kako bi se uočila vrednost kojoj eksperiment teži. Većim brojem eksperimenata, dobijena vrednost će biti bliža stvarnoj vrednosti.

Frekvencija
Apsolutna frekvencija nekog podatka (ili podataka koji zadovoljavaju neku osobinu) predstavlja broj pojavljivanja tog podatka u posmatranom uzorku. Relativna frekvencije predstavlja količnik apsolutne frekvencije i ukupnog broja podataka.

Uslovna verovatnoća i Bejzova (Bayes) teorema Neka je $P(A)$ verovatnoća događaja A i $P(B)$ verovatnoća događaja B, onda je verovatnoća događaja A pod uslovom da se desio događaj B u oznaci $P(A|B)$ jednaka:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{P(A \cap B) * P(A)}{P(A)}}{P(B)} = \frac{P(B|A) * P(A)}{P(B)}$$

Centralna granična teorema
Suma velikog broja nezavisnih slučajnih promenljivih sa identičnom raspodelom verovatnoće, čija su matematička očekivanja i varijansa konačni, teži normalnoj raspodeli verovatnoće.

Logaritamska skala
Logaritamska skala je nelinearna skala koja se koristi pri predstavljanju podataka velikih raspona vrednosti, odnosno podataka sa različitim redovima magnitude.

Logaritamski identiteti primenjeni u radu

- $\log(ab) = \log(a) + \log(b)$
- $\log(a/b) = \log(a) - \log(b)$
- $\log(a^b) = b \log(a)$
- $\log_b(x) = \dfrac{\log_k(x)}{\log_k(b)}$

Naučni zapis broja i signifikant
Naučni zapis broje N:

$$N = \alpha_1 10^n + \alpha_2 10^{n-1} + ... + \alpha_k 10^{n-k+1} + ... + \alpha_m 10^{n-m+1}$$

Normalizovani naučni zapis broja N:

$$S(N) * 10^n$$

Eksponent n je celi broj, a m je realna vrednost koja predstavlja sve značajne cifre broja i naziva se signifikant. Analogno važi za proizvoljnu bazu (umesto baze 10, može se koristiti proizvoljna baza).
Na primer, broj 556.25 u nomralizovanom obliku se može zapisati kao $55625 * 10^{-2}$, ali i kao $5.5625 * 10^2$.
Naučni zapis broja u kome signifikant uzima vrednost između 1.0 i 10 naziva se modifikovani normalizovni oblik. Signifikant iz intervala od 1.0 do 10 naziva se normalizovani signifikant $S(N)$

Red magnitude
Koristeći modifikovani normalizovani zapis broja:

$$N = S(N) * 10_n$$

, redovi magnitude su n stepeni broja 10.
Broj $1200 = 1.2 * 10^3$, dakle red magnitude je 3.

Mantisa
Izvorno, mantisa predstavlja razliku između vrednosti broja i celobrojnog donjeg ograničenja broja Formalano zapisano:

$$\{y\} = y - \lfloor y \rfloor$$

Odnosno, mantisa pozitivnih brojeva je deo broja koji se nalazi iza decimalnog zareza. Kao primer navodimo:

- $\{45.222\} = 45.222 - 45 = 0.222$
- $\{45.222\} = 45.222 - 46 = 0.888$

Uzorak
Uzorak je konačan broj podataka iz neke oblasti na kojima se sprovodi istraživanje.

Deskriptivna statistika
Deskriptivna statistika sadrži metode i procedure za prezentovanje i sumiranje podataka. Predstavlja prvi korak u analizi podataka, a služi za opisivanje prikupljenih podataka. Najčešće korišćene procedure u deskriptivnoj statistici su grafičko i tabelarno prikazivanje podataka i izračunavanje mera centralne tendencije i varijabiliteta.

Glava 3

Benfordov zakon

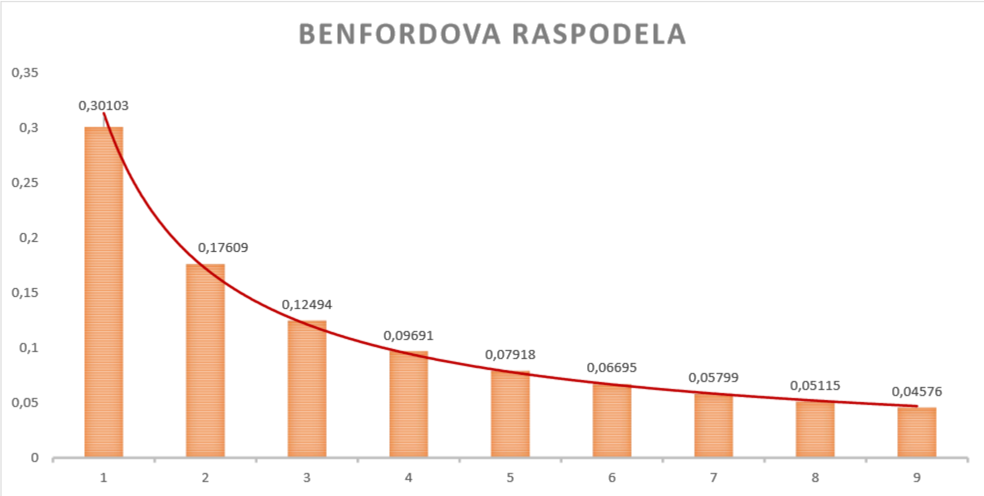
Dekadni brojni sistem podrazumeva zapis vrednosti na način da se na početnoj poziciji može naći jedna od devet cifara iz skupa $\{1, 2, 3, ..., 9\}$. Mogao bi se nametnuti prost zaključak da svaka od cifara ima jednaku verovatnoću pojavljivanja na vodećoj poziciji, odnosno da je raspodela vodeće cifre broja uniformna. Na nivou skupova to bi značilo da 11,11% podataka skupa počinjem svakom od cifara. Međutim, empirijska saznanja su se suprotstavila navedenom razmatranju. Kod značajanog broja skupova dobijenih iz različitih izvora uočene su razlike u količini podataka koji počinju svakom pojedinačnom cifrom. Ova osobina nazvana je zakonom anomalnih brojeva, odnosno fenomenom vodeće cifre.

Fenomen vodeće cifre, poznatiji kao Benfordov zakon predstavlja svojstvo određenih skupova podataka da više od 30 % podataka počinje cifrom jedan, dok se sa povećanjem vrednosti cifre $\{2, 3, ..., 9\}$ verovatnoće njihovog pojavljivanja na vodećim pozicijama smanjuju. Tabelarno prikazano, očekivane verovatnoće, odnosno frekvencije pojave cifara na vodećoj poziciji date su vrednostima:

Benfordov zakon	
Cifra d	*Očekivane verovatnoće pojavljivanja cifre d na vodećoj poziciji*
1	0,30103
2	0,17609
3	0,12494
4	0,09691
5	0,07918
6	0,06695
7	0,05799
8	0,05115
9	0,04576

Tabela 1: Benfordova zakonitost vodećih cifara

Na osnovu navedenih verovatnoća, formira se raspodela učestalosti pojavljivanja cifara na vodećoj poziciji koju nazivamo Benfordovom raspodelom.



Slika 1:Benfordova raspodela

Skupovi podataka koji prate Benfordovu raspodelu nazivaju se Benfordovi skupovi.

3.1 Istorijski pogled

Prva objavljena zapažanja vezana za raspodelu cifri na vodećoj poziciji broja pripadaju Američkom astronomu Simon Newcomb 1881. godine. Newcomb je uvideo da su stranice knjižica logaritama neravnomerno pohabane, odnosno da se pohabanost stranica smanjuje sa porastom njihovog rednog broja. Nametnuo se zaključk da su naučnici daleko više tražili logaritamske vrednosti brojeva koji počinju manjom cifrom. Newcomb se usmerio ka skupovima podataka koji su činili svakodnevnicu čoveka i takve skupove je nazvao "prirodnim"[7], a zaključio je da većina ovih skupova zadovoljava uočenu zakonitost. Pružio je formalan zapis verovatnoće da broj iz skupa koji ima ovo svojstvo počinje cifrom d:

$$P(d) = \log_{10}(d+1) - \log_{10}(d) = \log_{10}\left(\frac{d+1}{d}\right)$$

Njegova zapažanja, nedovoljno matematički i empirijski dokazana, prošla su neopaženo.

3.1.1 Benfoordova istraživanja

Činilo se da će zakon usled nedostatka matematičkog pojašnjenja u poptunosti biti zaboravljen sve dok 1937. godine fizičar Frank Benford nije objavio rezultate svog eksperimenta. Benford je naizgled nezavisno od svog prethodnika primetio istu istrošenost prvih stranica logaritamske knjižice, pa je odlučio da uočeno svojstvo testira na skupovima podataka koji dolaze iz različitih oblasti uključujući geografiju, fiziku i demografiju. Eksperiment je sprovodio nad 20 229 podataka obuhvatajući skupove iz najrazličitijih oblasti. U narednoj tabeli dati su pojedini rezultati Benfordove analize o učestalosti pojave cifara na početnoj poziciji:

Naziv	1	2	3	4	5	6	7	8	9	**Obim skupa**
Dužine reka	31,0	16,4	10,7	11,3	7,2	8,6	5,5	4,2	5,1	335
Naseljenost	33,9	20,4	14,2	8,1	7,2	6,2	4,1	3,7	2,2	3259
Fizičke konstante	41,3	14,4	4,8	8,6	10,6	5,8	1	2,9	10,6	104
Brojevi iz novina	30,0	18,0	12,0	10,0	8,0	6,0	6,0	5,0	5,0	100
Specifična toplota	24,0	18,4	16,2	14,6	10,6	4,1	3,2	4,8	4,1	1389
Pritisak	29,5	18,3	12,8	9,8	8,3	6,4	5,7	4,4	4,7	703
Molekularna težina	26,7	25,2	15,4	10,8	6,7	5,1	4,1	2,8	3,2	1800
Atomska težina	47,2	18,7	5,5	4,4	6,6	4,4	3,3	4,4	5,5	91
$\sqrt{n}$	25,7	20,3	9,7	6,8	6,6	6,8	7,2	8,0	8,9	5000
Narudžbenice	32,4	18,8	10,1	10,1	9,8	5,5	4,7	5,5	3,1	741
n!	25,3	16,0	12,0	10,0	8,5	8,8	6,8	7,1	5,5	900
Natalitet	27,0	18,6	15,7	9,4	6,7	6,5	7,2	4,8	4,1	418
Prosek	30,6	18,5	12,4	9,4	8,0	6,4	5,1	4,9	4,7	14840
Benfordov zakon	30,1	17,6	12,5	9,7	7,9	6,7	5,8	5,1	4,6	

Tabela 2: Benfordov eksperiment[15]

Benford je empirijski uspeo da pokaže značajnost i rasprostranjenost ovog svojstva, a njegov rad pod nazivom " Zakon anomalnih brojeva"[1] priznat je u svetu nauke. Do adekvatnog teorijskog objašnjenja naučnici će doći tek šezdeset godina kasnije.

1994. godine naučnik Boyle, zaključio je da su skupovi brojeva koji zadovoljavju Benfordov zakon dobijeni vršenjem matemtičkih operacija nad brojevima koji dolaze iz različitih izvora, odnosno iz različitih raspodela. Jednostavan primer Boylovog zaključka mogu biti računi narudžbenica, jer se brojevi na ovim računima dobijaju kao proizvod cene i željenje količine, što su dva različita izvora informacija sa zasebnim raspodelama. Ujedno, posmatrajući prethodnu tabelu uočavamo da sumiranje skupova koji su bili predmet Benfordovog istraživanja prati Benfordovu raspodelu, dok pojedinačni skupovi nisu ispoljavali Benfordovo svojstvo.

1995. Theodore P. Hill je pružio matematički dokaz[9] Benfordovog zakona, nakon čega dolazi do potpunog prihvatanja ovog fenomena. Boylova saznanja bila su ključni element Hillovog dokaza. Matematički dokaz se zasniva na činjenici da se podaci koji zadovoljavaju Benfordov zakon formiraju matematičkim operacijama između skupova podataka sa različitim raspodelama, gde pojedinačne raspodele ne moraju zadovoljavati Benfordovu. Hill je to nazvao raspodelom druge generacije ili raspodelom drugog reda.
Od postavke teorijskog dokaza, Benfordov zakon je našao široku primenu u oblastima revizijskih usluga, ekonomije, bankarstva, poreskog poslovanja, kao i informacionih tehnologija.

Poseban doprinos primeni Benfordovog zakonu u računovodstvu i reviziji pri detektovanju finansijskih nedoslednosti pripada Mark Nigriniju, koji je Benfrodov zakon uveo kroz postupak Digitalne analize.4.2. Značajan doprinos u kreiranju mera za ocenjivanje tačnosti finansijskih izveštaja pružila je grupa naučnika okupljena oko Zahna Bozanica[5], profesora Univerziteta Ohajo.

3.2 Benfordovi skupovi

Def 1. *Skupovi koji ispoljavaju fenomen vodeće cifre, odnosno skupovi čiji podaci prate Benfordovu raspodelu nazivaju se Benfordovi skupovi.*

Fenomen vodeće cifre ne predstavlja univerzalnu zakonitost svih skupova podataka i nije prirodno svojstvo koje karakteriše sve oblasti čovekovog postojanja. Skup brojeva koji su nastali slučajnim uzorkom i ne predstavljaju numeričku interpretaciju stvarnih procesa sa sigurnoću neće pratiti Benfordovu raspodelu. Takođe, skupovi podataka koji ne dozvoljavaju sirok opseg cifara na vodećoj poziciji, odnosno blisko su vezani za jednu početnu cifru neće ispoljavati Benfordovo svojstvo. Ovakvi primeri upućuju na postojanje određenih ograničenja u pogledu osobina skupova koji prate zakonitost prve cifre. Analiza podataka zasnovana na Benfordovom zakonu koristiti se isključivo pri testiranju skupa podataka od koga se očekuje praćenje Benfordove raspodele. Odnosno, odstupanja skupova za koje se očekuje praćenje raspodele, signaliziraju moguće nepravilnosti, nelegalne radnje ili korekcije stvarnih podataka.

Nametnuo se zaključak o postojanju zajedničkih karakteristika skupova od kojih se očekuje ispoljavanje Benfordovog svojstva. Sumirajući empirijske zaključke, potrebni uslovi da bi skup bio Benfordov su:

- Skup sadrži dovoljan broj elemenata da omogući raspodelu cifara[1].
- Skup je raspodeljen preko nekoliko redova magnitude.
- Skup podataka mora biti homogen.[2]
- Skup podataka ne čine potpuno slučajni brojevi. [3]
- Podaci ne smeju biti u potpunoj suprotnosti sa slučajnim brojevima, odnosno ne smeju sadržati ograničenja vrednosti, niti biti večtačkog porekla formirani nekim modelom šifrovanja.

Naredna tabela obezbediće jednostavan pregled podataka za koje se očekuje praćenje Benfordove raspodele, odnosno skupova za koje ima smisla razmatrati odstupanje od zakonitosti kao signal neispravnosti, kao i onih gde se primena ne smatra mogućom i korisnom.

[1] Testirani skup sadrži minimalno 100 elemenata

[2] Jedinice koje čine skup su istovrsne.

[3] U slučajnom broju svaka cifra ima jednaku verovatnoću pojavljivanja na bilo kojoj poziciji broja.

Benfordova analiza primenjiva na:
Skupovima brojeva koji su nastali kombinacijom dve ili više raspodela. Računi uplatnica i narudžbenica su nastali proizvodom cene i količine.
Transakcionim podacim kao što su uplate, isplate i refundiranja.
Skupovima podataka kod kojih je spoljoštenost raspodele pozitivna, a sredina je veća od mediane. Većina računovodvenih skupova zadovoljava ova dva uslova.

Benfordova analiza nije primenjiva na:
Skupovima podataka koji su šifrovani, kao što su poštanski brojevi, brojevi telefona.
Brojevi uslovljeni čovekovim razmišljanjima, odnosno svi dodeljeni brojevi.
Skupovi podataka koji su, da bi bili zabeleženi, morali biti donja / gornja granica u određenom trenutku.

Tabela 3: Indikatori i kontraindikatori Benfordovih skupova

Analiza upotrebom Benfordovog zakona ne smatra se korisnom u slučajevima kada tesitirani skup sadrži veliki broj dupliranih podataka. Prilikom proveravanja skupa transakcija preduzeća koje uključuju određene vrste fiksno zadatih naknada, čija je učestalost pojavljivanja velika, potrebna je dodatna predostrožnost i eventualno prilagođavanje skupa podataka testiranju uklanjanjem dupliranih vrednosti.

Dodatno, problem predstavljaju i nepostojeći podaci. Ukoliko je prevara izvedena na način da ne postoji pisana evidencija o nelegalnim aktivnostima kao što su krađe i mita, Benfordov zakon u njihovom otrkivanju je potpuno nemoćan. Pri primeni Benfordovog zakona moglo bi se formulisati pravilo:

$$nepostojeći\ podatak = nepostojanje\ malverzacija.$$

Ukoliko razmatramo prevaru u kojoj je pojedinac na jednom od podataka napravio promene, recimo napisao neosnovan ček, upotreba Benfordovog zakona neće biti korisna pri detekciji pojedinačne prevare. Benfordov zakon se koristi za uočavanje ponovljenih manipulacija, kao i manipulacija koje su uzročno-posledično povezane, koje mogu napraviti značajna odstupanje od zakonitosti.

3.2.1 Oblasti primene Benfordovog zakona

Smatrajući Benfordov zakon lepom osobinom bez posebnog uporišta nisu se posebno razmatrala polja primene. Krajem dvadesetog veka razvoj informacionih tehnologija, pokrenuće i usmeriti primenu Benfordovog zakona u dva smera:

- Modelovanje
  Naučnici su ustanovili da ukoliko podaci prate Benfordovu raspodelu, predviđanja budućih vrednosti ovih podataka na osnovu nekog matematičkog modela takođe će pratiti raspodelu. Utvrđeno je da neki od poznatih berzanskih indeksa prate Benfordovu raspodelu, pa se ova osobina može iskoristiti za utvrđivanje kvaliteta predikcionih modela.

- Otkrivanje nepravilnosti
  U osnovi primene Benfordove zakonitosti pri otkrivanju mogućih nepravilnosti, malverzacija i prevara nalazi se prost zaključak da ukoliko je na skupu podataka za koji se očekuje praćenje Benfordove raspodele izvršna bilo kakva namerna promena vodeće cifre neće zadovoljavati Benfordovu raspodelu. Primenom različitih testova, utvrđujemo praćenje / odstupanje Benfordovih skupova, odnosno neosnovanost / osnovanost sumnji o postojanju nepravilnosti. Benfordovom analizom su se do sada ispitivali skupovi novčanih transakcija preduzeća, finansijski izveštaji, sociološke promene, ali i makroekonomski podaci[4].

Forenzička analiza je nauka koja se bavi otkrivanjem, prikupljanjem i analizom manipulacija nad finansijkim podacima u formi koja bi mogla poslužiti kao validan dokaz pri sudskom postupku protiv izvršilaca nelegalnih radnji. Jedan od segmenata sprovođenja Forenzičke analize nad podacima, uključuje testiranja primenom Benfordovog zakona odnosno sprovođenje Digitalne Benfordove analize, koja se smatra validnim dokazom nepravilnosti u većini zemalja sveta.
Pojedini programi, kao što su ACL, Kirix Strate u sklopu svojih funkcija sadrže proveru zadovoljavanja Benfordove raspodele .

Svetska javnost se upoznala sa Benfordovim zakonom kroz slučaj nepravilnosti makroekonomskih izveštaja Grčke vlade prilikom pristupanja Evropskoj Uniji. Finansijski izveštaji koji su dostavljeni EU u trenutku pristupanja tretirali su se kao verodostojan prikaz finansijskog stanja u Grčkoj. Naknadne analize primenom Benfordovog zakona pokazale su veliku devijaciju izveštaja u odnosu na Benfordov zakon[17] i na taj način sugerisale manipulatorne aktivnosti nad podacima. Ekonomski problemi i krize sa kojima se Grčka susrela u godinama nakon prijema, potvrdili su osnovanost rezultata Benfordove analize.

2009. godine održani su izbori za predsednika Irana, na kojima je prema prvim rezultatima pobeda pripala dotadašnjem predsedniku Mahmoud Ahmadinejadu velikom većinom od 62% u odnosu na protivkandidata Mir-Hossein Mousavija. Usledile su demonstracije u svim gradovima Irana, kao i Iranaca širom sveta verujući u neregularnost i nameštenost izbora. Svetski analitičari su sprovodili opsežne numeričke analize ispravnosti glasanja, uključujući analiziranje broja glasača na biralištima primenom Benfordovog zakona. Rezultati su ukazali na velika odstupanja, odnosno jasnu neregularnost izbora[26].

Dodatno, sprovode se brojna istraživanja o primeni Benfordovog zakona na polju informacionih tehnnologija. Grupa švajcarskih naučnika u cilju povećanja efikasnosti bavi se testiranjem hard diska na kome su podaci grupisani primenom Benfordovog principa u odnosu na veličine pohranjenih datoteka u bajtovima.

3.3 Matematička pojašnjenja Benfordovog zakona

Def 2. ***(Benfordov zakon vodeće cifre)*** *Skup brojeva zadovoljava Benfordov zakon vodeće cifre ako je verovatnoća pojavljivanja cifre d,* $d \in \{1, 2, 3, ..., 9\}$ *na prvoj poziciji jednaka:*

$$P(d) = \log_{10}(d+1) - \log_{10}(d) = \log_{10}\left(\frac{d+1}{d}\right) = \log_{10}\left(1 + \frac{1}{d}\right)$$

Naznačimo, verovatnoće vodećih cifara zadovoljavaju jednakost:

$$P(1) + P(2) + ... + P(9) = 1$$

Odnosno,

$$\log_{10}\left(\frac{2}{1}\right) + \log_{10}\left(\frac{3}{2}\right) + ... + \log_{10}\left(\frac{10}{9}\right) = 1$$

Navedena definicija ima nedostatak u domenu primene. Naime, vrednost sa leve strane, odnosno verovatnoće pojavljivanja su racionalni brojevi, dok izraz sa desne strane jednakosti predstavlja iracionalnu vrednost. Prethodnu definiciju možemo korigovati. Matematički, problem bi se mogao rešiti posmatranjem beskonačnih skupova[3], dok u praksi to nije moguće, jer su skupovi podataka iz realnog sveta konačni. Dakle, kako bi se definicija mogla primenjivati, potrebno je izvršiti neznatne promene.

Def 3. ***(Benfordov zakon vodeće cifre(Upotrebna definicija))*** *Skup brojeva zadovoljava Benfordov zakon vodeće cifre ako je verovatnoća pojavljivanja cifre d na prvoj poziciji približno jednaka:*

$$P(d) = \log_{10}\left(\frac{d+1}{d}\right)$$

Pojam "približno" može uneti dodatne nejasnoće. Za početak, smatraćemo da skup podataka "približno" prati Benfordovu raspodelu ukoliko imamo približno grafičko poklapanje raspodela.

U nastavku, razmotrićemo uticaj koji mantisa9 logaritma broja ima na vodeću cifru.

3.3.1 Određivanje vodeće cifre broja

Mantisa određuje vodeću cifru broja. U nastavku navodimo algoritam određivanja vodeće cifre.

Algoritam[5]

- Izračunati log [4] broja
- Označiti mantisu (deo broja sa desne strane decimalnog zareza)
- Podići 10 na stepen vrednosti mantise: $10^{mantisa}$
- Vodeća cifra dobijenog broja je prva cifra početnog broja.

Primer 1. *Ukoliko na broju 542.325 primenimo prethodni algoritam*

- $\log 542.325 = 2.734$
- *Mantisa=0.734*
- $10^{0.734} = 5.42$
- *Vodeća cifra dobijenog broja je 5.*

Algoritam nam pokazuje na koji način mantisa logaritamske vrednosti broja (u oznaci $M(\log(N))$) određuje cifru na vodećoj poziciji. Formalno, broj N počinje cifrom d, $d \in \{1, 2, 3, ..., 9\}$ ako i samo ako

$$M(\log(N)) \in [\log(d), \log(d+1))$$

Odnosno, broj N počinje cifrom:

- $d = 1 \Leftrightarrow$ mantisa $\log(N) \in [\log(1), \log(2)) = [0, 0.301)$
- $d = 2 \Leftrightarrow$ mantisa $\log(N) \in [\log(2), \log(3)) = [0.301, 0.477)$

$$\vdots$$

- $d = 9 \Leftrightarrow$ mantisa $\log(N) \in [\log(10), \log(9)) = [1, 0.954)$

Verovatnoće pojavljivanja cifre d kao vodeće cifre definisane Benfordovim zakonom su dužine intervala kojima pripadaju mantise logaritama brojeva sa početnom cifrom d.

Iz pokazanog, nameće se formulacija zakona: *"Verovatnoće pojavljivanja brojeva na vodećoj poziciji definisane su mantisama njihovih logaritama koje su jednako verovatne"*

[4] U nastavku rada, oznaka $\log(N)$ predstavlja logaritam sa bazom 10. Logaritamske vrednosti drugih baza biće posebno naglašene

Sada se podsetimo normalizovanog naučnog zapisa broja N:

$$S(N) * 10^n$$

gde $S(N)$ nazivamo signifikantom8 i n je red magnitude9 , a u slučaju modifikovanog normalizovanog zapisa vrednost $S(N) \in [1, 10)$.
Kada bi umesto vodeće cifre posmatrali ceo signifikant, umesto postavljenog pitanja da li je na vodećoj poziciji cifra d mogli bismo postaviti pitanje: Da li se posmatrani signifikant nalazi između *d* i *d+1*?
Primenom definisanog dolazimo do nove formulacije Benfordovog zakona.

Def 4. ***(Benfordov zakon(I))*** *Skup brojeva zadovoljava Benfordov zakon vodeće cifre ako je verovatnoća pripadanja signifikanta intervalu* $[1, s)$ *jednaka* $\log(s)$.

Prethodna definicija implicira Benfordovo svojstvo, jer verovatnoća da se na vodećoj poiziciji nađe cifra d jednaka je verovatnoći da signifikant broja pripada intervalu $[d, d + 1)$. Ako bismo interval $[d, d + 1)$ zapisali kao $[1, d + 1) \,/\, [1, d)$, jasno je da je verovatnoća pripadanja signifikanta intervalu jednaka verovatnoći da se na vodećoj poziciji nađe cifra d i iznosi:

$$\log\,(d + 1) - \log\,(d)$$

Jasno je da se navedene definicije odnose na vodeću cifru u smislu prve cifre u zapisu broja, pa se postavlja pitanje o ponašanju cifara na ostalim vodećim pozicijama.
U tebeli se navode vorovatnoće pojavljivanja cifara na prve četiri vodeće pozicije. Oznaka D_i predstavlja i-tu poziciju broja.[7]

Očekivane verovatnoće pojavljivanja cifara				
Cifra d	D_1	D_2	D_3	D_4
0		0,11968	0,10178	0,10018
1	0,30103	0,11389	0,10138	0,10014
2	0,17609	0,19882	,010097	0,10010
3	0,12494	0,10433	0,10057	0,10006
4	0,09691	0,10031	0,10018	0,10002
5	0,07918	0,09668	0,09979	0,09998
6	0,06695	0,09337	0,09940	0,09994
7	0,05799	0,09035	0,09902	0,09990
8	0,05115	0,08757	0,09864	0,09986
9	0,04576	0,08500	0,09827	0,09982

Tabela 4: Verovatnoće pojavljivanja cifara na vodećim pozicijama

Primetimo, udaljavanje od početne pozicije broja praćeno je promenom učestalosti pojavljivanja cifara. Zapravo, povećanjem i, učestalost pojave cifre na D_i poziciji teži ravnomernoj raspodeli. Dovoljno je posmatranje do D_4, da uočimo približno jednake verovatnoće pojave svih deset cifara.

Formalno zapisano, ako sa D_i označimo i-tu poziciju broja, a sa d_i cifru na i-toj poziciji, verovatnoća da se na drugoj poziciji broja D_2 pojavi cifra d_2 jednaka je:

$$P(D_2 = d_2) = \sum_{d_1=1}^{9} log_1(1 + (1/d_1 d_2)), \quad d_i \in \{1, 2, 3, ..., 9\}$$

Dakle, verovatnoća da se na prvoj poziciji broja nađe cifra d_1 i na drugoj poziciji cifra d_2 jednaka je :

$$P(D_1 = d_1, D_2 = d_2) = \log(1 + (1/d_1 d_2)), \quad d_i \in \{1, 2, 3, ..., 9\}$$

Odnosno, verovatnoća da će se na poziciji D_2 naći cifra d_2, ako se na poziciji D_1 nalazi cifra d_1 jednaka je:

$$P(D_1 = d_1 | D_2 = d_2) = \log(1 + (1/d_1 d_2)) / \log(1 + (1/d_1)), \quad d_i \in \{1, 2, 3, ..., 9\}$$

Generalizaciju Benfordovog zakona možemo usmeriti i u odnosu na bazu logaritma. Ukoliko zakon važi na skupu podataka upotrebom baze 10 logaritma, prirodno je očekivati da će se zakon održati u odnosu na proizvoljnu bazu. Ova osobina naziva se bazna invarijantnost.
Umesto baze 10 možemo koristiti proizvolju bazu B, usled čega verovatnoća vodeće cifre postaje:

$$P_B(d) = \log_B \left(\frac{d+1}{d} \right) = \log_B(d+1) - \log_B(d)$$

.
Sumiranjem definicije 4 i osobine bazne invarijantnosti formira se najopštija formulacija Benfordovog zakona.

Def 5. ***(Benfordov zakon(I) baze B)*** *Skup podataka zadovoljava Benfordov zakon (I) baze B ako je verovatnoća pripadanja signifikanta intervalu* $[1, s)$ *u bazi B jednaka* $\log_B(s)$.

3.3.2 Pojašnjenja rasprostranjenosti Benfordovog zakona

Empirijski je pokazano da veliki broj skupova iz različitih oblasti nauke, pojava u prirodi, finansija zadovoljava Benfordov zakon. U nastavku pažnja se usmerava na matematičke razloge rasprostranjenosti Benfordove raspodele kroz :

- Geometrijsko pojašnjenje
- Pojašnjenje skalarne invarijantnosti
- Pojašnjenje upotrebom Centralne granične teoreme
- Pojašnjenje zasnovano na funkciji gustine

3.3.2.1 Geometrijsko pojašnjenje

Ideja geometrijskog pojašnjenja zasniva se na procesu sa konstantnom stopom rasta, usled čega će veći procenat posmatranih podataka počinjati manjim ciframa. Posmatrajmo kretanje broja stanovnika malog mesta u Srbiji[15]. Pratimo podatke o rastu broja stanovnika od trenutka kada je mesto imalo 1000 stanovnika. Godišnji prirast od 10% nam govori da ć u periodu od 7.3 godine broj stanovnika se udvostručiti, odnosno dostići vrednost 1999. Za ponovno udvostručavanje sa 2000 na 3999 biće potrebno 7.3 godine. Ove godine biće raspoređne između brojeva koji počinju cifrom 2 i brojeva koji počinju cifrom 3. Period rasta broja stanovnika od 4000 do 7999 takođe će trajati 7.3 godine, a kao početne cifre biće obuhvaćene 4,5,6,7. Kada broj stanovnika dosegne 10000, krug vezan za pojavljivanje cifara na vodećoj poziciji broja ponovo započinje. Zapravo, posmatrajući kretanje broja stanovnika tokom više decenija prosečno vreme koje broj stanovnika provede na određenoj početnoj cifri teži Benfordovoj raspodeli. Ukoliko formalno zapišemo proces rasta populacije sa 1000 stanovnika na 2000 stanovnika uz godišnji prirast od 10% u n-tom periodu:

$$1000 * (1.01)^n = 2000$$

$$n = \frac{\log 2}{\log 1.01}$$

Uopšteno, vreme provedeno sa brojem stanovnika koji cifru d ima kao vodeću, jednako je:

$$\frac{\dfrac{\log \dfrac{d+1}{d}}{log 1.01}}{\dfrac{\log 10}{log 1.01}} = \frac{\log \dfrac{d+1}{d}}{log 10}$$

što predstavlja očekivane verovatnoće definisane Benfordovim zakonom.

Veliki broj pojava u prirodi kao što je razmnožavanje bakterija, ali i matematičkih zakonistosti kao što je Fibonačijev niz[3], uslovljeni su geometrijskim rastom i prate Benfordov zakon.

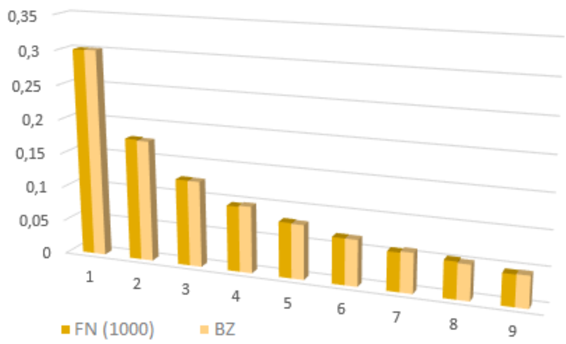
Primer 2. *Fibonačijev niz brojeva nastaje na osnovu pravila da zbir prethodna dva člana u nizu daje naredni član niza.*

$$F_0 = 0$$

$$F_1 = 1$$

$$F_n = F_{n-1} + F_{n-2} \quad n \geq 2$$

Ako formirano prvih 1000 članova niza, dobijamo:



Slika 2: Raspodela vodeće cifre Fibonačijevog niza

3.3.2.2 Pojašnjenje skalarne invarijantnosti

Posmatrajući skupove za koje se empirijski pokazalo da su Benfordovi skupovi uočava se da dolaze iz različitih životnih oblasti, odnosno da merne jedinice u kojima se podaci izražavaju mogu biti različite. U osnovi, različite zemlje, pojedinci i sistemi koriste različite skale za izražavanje podataka iz istog skupa. Skup finansijskih podataka poseduje najveći broj interpretacija u različitim valutama. Jasno je da Benfordov skup podataka treba da zadrži Benfordovo svojstvo nezavisno od brojnog sistema i mernih jedinica u kojima će se podaci predstaviti. Zamislimo da smo prinuđeni na prevođenje podataka iz dekadnog brojnog sistema koji priznaje devet cifara na vodećoj poziciji u brojni sistem u kome je cifra sedam najveća moguća vrednost. Promena brojnog sistema će uticati na promenu većine vodećih cifara, ali uprkos izmenama Benfordovo svojstvo će se održati. Dodatno, raspodela prve značajne cifre bi trebala ostati nepromenjena kada se svi brojevi množe sa konstantnim faktorom, što je u osnovi pretvaranja mernih jedinica. Opisana osobina naziva se sklarana invarijantnost.

Dokazivanje skalarne invarijantnosti zahteva uvrštavanje gradiva iz više oblasti[8] uključujući i σ-algebre, što nije predmet ovog rada, stoga ćemo se intuitivno osloniti na ovu osobinu.

Pretpostavimo da imamo skup podataka koji zadovoljava Benfordov zakon (4). Odnosno, verovatnoća pripadanja signifikanta intervalu:

$$S(N) \in [a, b] \subset [1, 10) \quad jednakaje \quad \log\left(\frac{b}{a}\right)$$

Množenje elemenata skupa konstantnom vrednosti $\sqrt{3}$, interval $[1, 10)$ prevodi u interval $\left[\sqrt{3}, 10\sqrt{3}\right)$. Sada, za broj N koji na vodećoj poziciji ima cifru 1 intervali pripadanja signifikanta su $\left[\sqrt{3}, 2\right)$ i $\left[10, 10\sqrt{3}\right)$. Upotrebom definicije 4 , verovatnoće pripadanja signifikanata intrevala su $\log\left(\frac{2/\sqrt{3}}{1}\right)$ i $\log\left(\frac{10/\sqrt{3}}{10}\right)$. Sumirajući ove dve verovatnoće dobijamo konačnu verovatnoću pojavljivanja cifre 1 na vodećoj poziciji koja je jednaka očekivanoj verovatnoći Benfordovog zakona:

$$\log\left(\frac{2/\sqrt{3}}{1}\right) + \log\left(\frac{10/\sqrt{3}}{10}\right) = \log(2)$$

3.3.2.3 Pojašnjenje upotrebom Centralne granične teoreme

Kako bismo pojasnili uticaj Centralne granične teoreme, potrebne su nam određene transformacije nad podacima.

Ako $y \in \Re_0^+$, *y moduo 1* u oznaci $ymod1$ predstavlja ostatak iza celog dela broja, odnosno

$$\{y\} = y - \lfloor y \rfloor$$

Ako $y \in \Re^-$, onda je ostatak iza celog dela broja:

$$\{y\} = 1 - (-ymod1)$$

Odnosno $ymod1$ je jedinstven broj iz intervala $[0, 1)$, a $(y - (ymod1))$ je celi broj.

Primer 3.
Neka je y=2.41, onda je $2.41mod1 = 0.41$
Neka je y=-2.41, onda je $-2.41mod1 = 1 - 0.41 = 0.59$

Primenjujemo ideju da podatke ne posmatramo u dobijenom obliku, nego da uvedemo transforamciju:

$$x \mapsto \log(x)mod1 \tag{3.1}$$

Ukoliko se prisetimo naučnog zapisa brojeva i značajnosti signifikanta, dobijamo da su za skup podataka $\{x_1, x_2, ..., x_k\}$, podskupovi brojeva koji počinju cifrom d definisani prema Benfordovim zakonom:

$$\{x_j : S(x_j) \in [d, d+1)\}$$

Primenom logaritmovanja dobijamo:

$$\{x_j : \log S(x_j) \in [\log d, \log d + 1)\}$$

Dva pozitivna broja imaju istu vodeću cifru, ako i samo ako im siginifikanti pripadaju istim intervalima.

Primenom transformacije 3.1 na naučni zapis broja $x = S(x)10^n$ dobija se:

$$\log(x)mod1 = \log S(x)$$

Dakle, ovim postupkom dolazimo do logaritamske vrednosti signifikanta, koja prema 4 predstavlja željene verovatnoće pojave cifre d kao vodeće. Prethodno pojašnjenje predstavlja formalan zapis algoritma.17

Veliki broj skupova podataka, uglavnom je nastao kao kombinacija različitih raspodela, odnosno operacijama sa veličinanama koje pripadaju različitim oblastima i izražene su u različitim mernim jedinicama . Recimo, računi za struju nastali su proizvodom utrošene energije i cene. Posmatrajmo slučajnu promenljivu X, koja je nastala kombinacijom n različitih slučajnih promenljivih.

$$X = X_1X_2...X_n$$

Prethodno razmatranje nam je pokazalo da raspravu u vezi praćenja Benfordove raspodele možemo poistoveti sa proučavanjem raspodele $log(X)mod1$. Dakle,

$$\log X = \log(X_1X_2...X_n) = \log X_1 + \log X_2 + ... + \log X_n$$

Centrala granična teorema8 tvrdi da će navedena suma logaritama biti normalno raspodeljena, za dovoljno veliko n.

Nas u osnovi zanima

$$(\log X)\ mod\ 1 = (\log X_1 + \log X_2 + ... + \log X_n)mod1$$

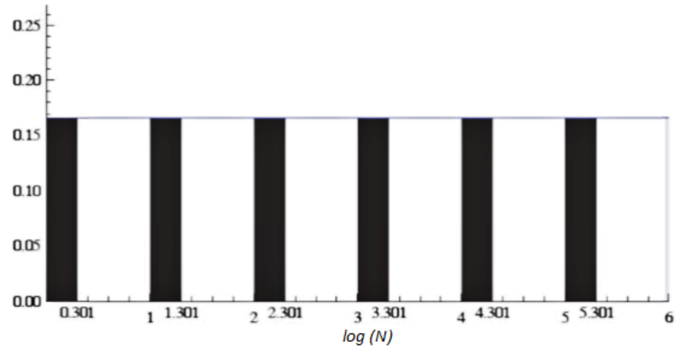
Pojednostavljeno, neka je Z normalno raspoređena sa sredinom μ i velikom varijansom σ. Ako razmotrimo funkciju gustine nove slučajne promenljive $Zmod1$ dobijamo da je približno uniformno raspoređena na intervalu $[0, 1)$. Navedeno znači da je verovatnoća pripadanja $Z \in [\log(d), \log(d + 1))$ jednaka $\log(d + 1) - \log(d)$, što je u skladu sa Benfordovom zakonitosti.
Iako ne deluje prirodno da tražimo logaritamsku vrednost, pa *moduo 1*, nakon primene takvog postupka skup podataka ima uniformnu raspodelu, ako je početni skup podataka bio Benfordov.[9] Na ovo možemo gledati i u suprotnom smeru, kako postoji prirodna tranformacija koja zahteva Benfrodov skup podataka, dok kao izlaznu vrednost dobijamo skup uniformno raspoređenih brojeva. Navedeni primer predstavlja množenje slučajnih promenljivih, dok se isti zaključak može izvesti i primenom drugih matematičkih operacija, kao što su sabiranje, stepenovanje. Dakle, u pozadini svega nalazi se Centralna granična teorema, samo je neophodno da svoj način gledanja prilagodimo kako bismo otkrili njenje efekte.

3.3.2.4 Pojašnjenje zasnovano na funkciji gustine

U nastavku se objašnjavaju razlozi učestalosti pojave Benfordovih skupova u različitim oblastima upotrebom funkcije gustine. Objašnjenje se zasniva na zapažanju da je površina ispod funkcije gustine jednaka verovatnoći da broj izvučen iz ove rapodele bude u jednom od osenčenih intrevala. Razmatraju se funkcije gustine i površine ispod funkcije gustine uniformne raspodele, normalne raspodele i generičke raspodele. Posmatraćemo raspodelu prvih cifara na logaritamskoj skali8 sa osnovom 10, umesto linearne skale.

Funkcija gustine uniformne raspodele

Posmatramo unifomnu raspodelu u intervalu od 0 do 6 logaritamske skale.



Slika 3: Uniformna raspodela vrednosti $\log(N)$

Funkcija gustine jednaka je:

$$\varphi(\log(N)) = \frac{1}{6}$$

Uočimo da svaki broj koji je između n i $n + 0.301$ na logaritamskoj skali, počinje cifrom 1 gledano u odnosu na linearnu skalu, jer mantisa se nalazi u intervalu $[0, 0.301)$. Kako bi se izračunala verovatnoća da broj na lineranoj skali počinje 1, moramo naći površine ispod funkcije gustine u intevalima n i $n+0.301$. U osnovi, osenčene površine ispod krive možemo dobiti računanjem integrala funkcije gustine u granicama n i $n + \log 2$, pa je verovatnoća da broj na vodećoj poziciji ima cifru 1 jednaka:

$$\sum_{n=0}^{5} \int_{n}^{n+\log 2} \frac{1}{6} dN = \frac{1}{6} * ((0.301 - 0) + (1.301 - 1) + ... + (5.301 - 5) = 0.301$$

Uopšteno, skup podataka koji prati uniformnu raspodelu u odnosu na na logaritamsku skalu, pratiće Benforodov zakon u odnosu na lineranu skalu, jer je verovatnoća da se na vodećoj poziciji pojavi cifra d jednaka $log(d + 1) - log(d)$, odnosno:

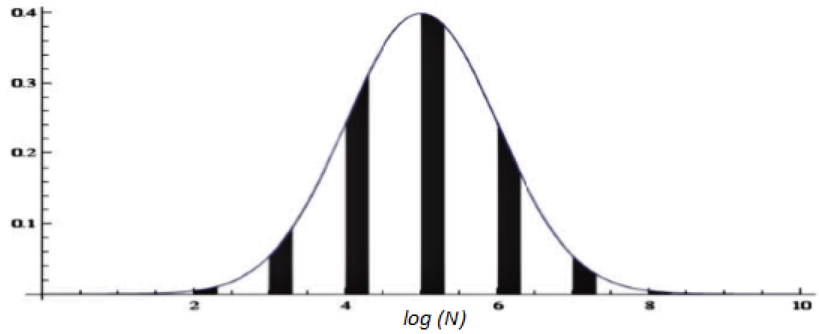
$$\sum_{n=0}^{5} \int_{n+\log(d)}^{n+\log(d+1)} \frac{1}{6} dN = \log(d + 1) - \log(d)$$

Funkcija gustine normalne raspodele

U ispitivanju skupova podataka se najčešće susrećemo sa normalnom i log-normalnom raspodelom, što se može razumeti kao posledica CGT, ali i prirodna osobina različitih procesa. Dodatno, potrebno je da se raspodela prostire u opsegu većem od nekoliko celih podioka tj. redova magnitude logaritamske skale, što znači da se prostire npr. od 1 do 10000 vrednosti linearne skale i praćenje Benfordove raspodele biće osigurano.

Posmatrajmo normalnu raspodelu u odnosu na logaritamsku skalu, što je zapravo log-normalna raspodela u odnosu na lineranu skalu. Neka je data funkcija gustine normalne raspodele sa očelivanjem $\mu = 5$ i odstupanjem $\sigma = 1$:

$$\varphi(\log(N)) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x-5)^2}{2}\right)$$



Slika 4: Normalna raspodela vrednosti $\log(N)$

Kao i u slučaju uniformne raspodele, osenčeni delovi predstavljaju intervale $[n, n+0.301]$, a površine svih osenčenih delova ispod funkcije gustine normalne raspodele daju verovatnoću da broj na vodećoj poziciji ima cifru 1, u odnosu na lineranu skalu. Formalno zapisano, verovatnoća da broj počinje cifrom 1:

$$\sum_{n=-\infty}^{\infty} \int_{n+\log 1}^{n+\log 2} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x-5)^2}{2}\right) \cong \log 2 - \log 1$$

Uopšteno, možemo nači učestalost pojavljivanja za svaku cifru d na vodećoj poziciji brojeva iz normalne raspodele:

$$\sum_{n=-\infty}^{\infty} \int_{n+\log d}^{n+\log(d+1)} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x-5)^2}{2} dN\right)) \cong \log(d+1) - \log(d)$$

Generičke raspodele

Generalno, za proizvoljnu raspodelu, verovatnoća cifre d kao vodeće može da se dobiti posmatranjem površine ispod funkcije gustine definisane tom raspodelom. Formalno zapisano:

$$\sum_{n=-\infty}^{\infty} \int_{n+\log(d)}^{n+\log(d+1)} \varphi\left(\log(N)\right) dN$$

Ukoliko je površina ispod funkcije gustine, definisana navedenom formulom jednaka $\log(d+1) - \log(d)$ onda će i verovatnoće da broj uzet iz ove raspodele počinje cifrom d zadovoljavati Benfordov zakon. Drugačije rečeno, ako imamo glatku i simetričnu funkciju gustine u odnosu na logaritamsku skalu, opsega većeg od nekoliko redova magnitude imaćemo praćenje Benfordovog zakona. Posmatrajući funkcije gustine koje su simetrične i glatke u odnosu na logaritamsku skalu , ali čija visina po segmentima nije jednaka, nameće se potreban uslov većeg reda magnitude. Dakle, kako bi suma oblasti po segmentima koji određuju vodeću cifru d dostigla vrednosti verovatnoća predviđenih Benfordovim zakonom potreban je dovoljno veliki red magnitude, da bi očekivana sredina visine fukcije gustine po segmentima bila jednaka za svaku cifru. Prostor ispod funkcije gustine u intervalima $n + \log(d)$ i $n + \log(d + 1)$ jednaki su verovatnoćama da se na vodećoj poziciji broja nađe cifra d predviđenih Benfordovim zakonom.

Kako većina raspodela zadovoljava osobinu glatkoće i simetričnosti funkcije gustine u odnosu na logaritamsku skalu kao posledica varijacija CTG, velika učestalost pojave Benfordovog zakona ne bi trebala da bude začuđujuća. Za raspodele koje generalno prate Benfordov zakon, odstupanje znači da su se pojavili određeni vidovi grešaka koje prave raspodelu manje simetričnom ili glatkom. Ovo zapažanje postalo je ključno za primenu Benfordovog zakona u otkrivanju nepravilnosti.

Glava 4

Primena Benfordovog zakona u reviziji

4.1 Motivi i oblasti primene

Pouzdanost i autentičnost finansijskog izveštavanja predstavlja osnov za donošenje svih odluka vezanih za poslovanje subjekta čije izveštaje razmatramo. Finansijski izveštaji su osnovna svedočanstva o imovinskom, finansijskom položaju i poslovnoj uspešnosti subjekta. Dosadašnja istraživanja na polju finansija pokazala su da kvalitet računovodstvenih podataka bitno utiče na efikasnu alokaciju resursa na tržištu, finansijski i privredni razvoj subjekta koji je predmet posmatranja, sklapanje ugovora o zaduživanju kao i na obim zapošljavanja. Osim korisnika finansijskih izveštaja kao što su vlasnici kompanija, rukovodioci, zaposleni i poslovni partneri, verodostojni finansijski podaci osnova su za izmirenje finansijskih obaveza prema državi od strane preduzeća[1].

Poreske i pravosudne institucije zadužene za kontrolu finansijskih kriminalnih radnji nemaju dovoljne kapacitete za borbu sa sve naprednijim i teško uočljivim malverzacijama. Glavnu ulogu u ispitivanju istinitosti preuzima revizorska profesija, u čijim okvirima se izdvaja posebna oblast pod nazivom forenzička revizija, specijalizovana za pitanja detekcije malverzacija, odnosno ocenjivanja zakonitosti, stručnog evidentiranja i izveštavanja na način da dokazi o nelegalnosti podataka mogu predstavljati validne argumente pri sudskom procesu.

Revizorska profesija, zahvaljujući velikom broju interesenata za finansijske izveštaje bila je prinuđena da osmisli neke nove i jednostavnije metode otkrivanja mogućih nepravilnosti računovodsvenih podataka. Dodatni problem predstavlja i razvoj međunarodne trgovine koji uključuje različite standarde finansijke evidencije. Poteškoće različitog vođenja računovodstvene evidencije između zemalja donekle su se prevazišle stvaranjem jedinstvenih međunarodnih računovodstvenih standarda MRS[2]. Međutim, sloboda samostalne interpretacije standarda ostavila

[1] U nastavku teksta govoićemo o preduzećima kao poslovnim subjektima čiji finansijski izveštaji podležu reviziji

[2] IFRS-međunarodni standardni finansijskog izveštavanja

je prostor subjektima za ostvarivanje skrivenih poslovnih ciljeva. Stvorila se potreba za metodom koja bi proveravala ispravnost podataka nezavisno od načina evidentiranja transakcija.

Jedan od osnovnih koraka u sprovođenju revizije predstavlja procenjivanje verovatnoća da navedeni podaci iz finansijskih izveštaja pojedinačno predstavljaju stvarne finansijske ocene poslovnih aktivnosti. Prilikom procenjivanja validnosti svake od pozicija u finansijskim izveštajima revizor se susreće sa rizikom procesa revizije koji obuhvata tri vrste rizika: inherentni, kontrolni i detekcioni.

- Inherentni rizik ili rizik pojave značajnih grešaka u finansisjkim izveštajima.
- Kontrloni rizik ili rizik da se značajne greške neće detektovati i otkloniti sistemom interne kontrole.
- Detekcioni rizik ili rizik neoktivanja, odnosno rizik da revizor neće svojim metodama otkriti značajne greške u računovodsvenim podacima.

Ukupan nivo rizičnosti dobija se proizvodom navedena tri rizika. Kako bi smanjio početnu rizičnost revizor skuplja što je moguće širi uzorak , kao i dovoljan broj informacija o samom poslovanju i već postojećim postupcima kontrole. U okvirima ovog procesa nudi mu se i alat koji se oslanja na Benfordovo svojstvo. Jednostavno analiziranje podataka upotrebom Benfordovog zakona, nametnulo se kao još jedna dodatna prednost, koja omogućava revizorima da pored intuicijskog i iskustvenog segmenta procenjivanja, dobiju i neke numeričke naznake o ispravnosti podataka koje procenjuju.

Analiza povezana sa Benfordovim zakonom, koja podrazumeva oslanjanje na konkretne vrednosti vodećih cifara i precizno definisan statistički aparat, oslobođena je svake subjektivnosti procenjivanja. Benfordov zakon se nametnuo kao efikasno rešenje za relativno brzo ispitivanje mogućih grešaka u računovodstvenim podacima. Odnosno, razmatranjem učestalosti pojave cifara, revizori pokušavaju da uoče informacije o istinitosti numeričkog evidentiranja transakcija, bez upuštanja u analiziranje poslovnih aktivnosti koje ti podaci predstavljaju. Jedna od osnovnih prednosti ispitivanja baziranih na Benfordovom zakonu jeste poželjnost većeg obima uzorka, odnosno testnog skupa. Veći broj podataka daje bolje rezultate analize koja se oslanja na Benfordov zakon[2]. Pozovemo li se na osnovne karakteristike skupova koji zadovoljavaju Benfordov zakon, prvenstveno na raspodelu drugog reda, uvidećemo da mnoštvo skupova podataka čije se sumirane vrednosti nalaze u finansijskim izveštajima prate Benfordovu raspodelu. Eventulana odstupanja od raspodele signal su za moguće nelegalne izmene podataka.

Pri sprovođenju stvarnog revizijskog procesa, omogućen je uvid u sve transakcije, pa je analiza podataka primenom Benfordove zakonitosti lako izvodljiva. Naše mogućnosti u pristupanju podacima su ograničene, odnosno možemo pristupiti isključivo završnim ocenama u vidu podataka iz finansijskih izveštaja. Podaci koji se nalaze u finansijskim izveštajima takođe zadovoljavaju Benfordovu raspodelu.[6]

Računovodstveni podaci iz finansijskih izveštaja u osnovi predstavljaju kombinacije ocena finansijskih realizacija podloga poslovanja, odnosno dobijaju se kombinovanjem raspodela različitih novčanih tokova koji čine poslovanje, npr. tokova novca od prodaje, zarada zaposlenih. Realizacije prihoda, isplate dobavljačima, isplate zaposlenima, plaćene takse uglavnom su nastale kao kombinacije podataka iz različitih rapsodela (*npr.*raspodele količine i raspodele cena), pa se očekivanje praćenja Benfordovog zakona čini opravdano.

Potrebno naglasiti da nema ograničenja u pogledu materijalnosti[3] poslovnog subjekta čiji se finansijski izveštaji procenjuju, odnosno mogu se ispitivati istovetno podaci velikih kompanija, kao i malih preduzeća. Takođe, moguće je raditi testiranja u različitim vremenskim razdobljina, nezavisno od obima i novih poslovnih aktivnosti koje preduzeće obavlja.

Sa psihloškog stanovišta čini se da bi podaci mogli biti izmanipulisani na način da pri promenama održavamo veću zastupljenost vodećih cifara manjih vrednosti, no ovakav pristup zbog povezanosti pozicija u FI prilično je otežan. Takođe, ukoliko bi se pokušale izvesti promene u finansijskim izveštajima koje bi trebale da kamufliraju već formirane malverzacije, troškovi tih promena bi bili previše veliki, pa je to dovoljan razlog za odvraćanje od bilo kakvog rada nad podacima.

[3] Udeo preduzeća na tržištu iz oblasti poslovanja subjekta

4.2 Digitalna Benfordova analiza

Postupak detektovanja nedoslednosti, malverzacija i prevara upotrebom Benfordovog zakona naziva se Digitalna Benfordova analiza.[14]
Digitalnu analizu mogu sprovoditi sva preduzeća, kompanije i finansijke institucije koje imaju računovodstvenu evidenciju podatka u cilju samokontrolisanja. Postupak je posebno namenjen institucijama koje su zadužene za poboljšanje, praćenje i kontrolisanje njihovog rada kao što su revizorske kuće, centralne banke i poreske institucije. Osnovi koraci digitalne analize su:

1. Odabir uzorka
2. Sprovođenje deskriptivne9 statistike
3. Prilagođavanje uzorka
4. Računanje raspodele koju formiraju elementi uzorka
5. Testiranja koja pružaju odgovor na pitanje: Da li raspodela uzorka prati Benfordovu raspodelu?

Prilikom odabira uzorka, potrebno je da se vodi računa o veličini uzorka. Odnosno, da skup podataka koji ispitujemo bude dovoljnog obima13.
Benfordov zakon se primenjuje isključivo na pozitivne vrednostima brojeva. Primenom apsolutne vrednosti na sve elemente uzorka, prilagodićemo podatke za dalja testiranja.

4.3 Testovi praćenja Benfordove raspodele

Merenje kvaliteta finansijkih izveštaja u čijoj osnovi se nalazi Benfordov zakonon, moguće je izvršiti različitim statističkim testovima[21]. Prilikom testiranja suprostavljaju se dve hipoteze:

$$H_0 \text{ (Podaci prate Benfordovu raspodelu)}$$

$$\updownarrow$$

$$H_1 \text{(Podaci odstupaju od Benfordove raspodele)}$$

U zavisnosti od rezultata testa, ili prihvatamo nultu hipotezu H_0, ili je odbacujemo u korist alternativne hipoteze H_1. Postupak testiranja odvija se kroz sledeće korake:

- Računanje test statistike uzorka.
- Izbor nivoa značajnosti, na osnovu koga se određuje kritična p-vrednost.
- Ukoliko je vrednost statistike uzorka veća od kritične vrednosti, onda se odbacuje nulta hipoteza za izabrani nivo značajnosti. U suprotonom, nulta hipoteza se prihvata.

Za potvrđivanje jedne od hipoteza, koriste se statistički testovi:

- Z-test
- χ^2 test
- Kolmogorov-Smirnov (KS) test
- MAD (*eng.*Mean absolute deviation) test

Z-test

Pri testiranju Z-test[22] koristi Z-statistiku oblika:

$$Z = \frac{(O_i - C_i)}{s_i}$$

gde je,

$$s_i = \sqrt{\left[\frac{O_i(1 - C_i)}{n}\right]}$$

C_i- frekvencija pojavljivanja cifre i utvrđena ispitivanjem uzorka
O_i- teorijska frekvencija pojavljivanja cifre i koja se očkuje Benfordovim zakonom n- obim uzorka, odnosno testiranog skupa. U zavisnosti od odabranog nivoa poverenja, Z-test koristi jedinstvene kritične vrednosti. Na primer,ako je pri 95% nivou poverenja apsolutna vrednost Z-statistike veća od 2.77, doće do odbijanja nulte hipoteze u koristi alternativne hipoteze. Nedostatak Z-testa se ogleda ispitivanju frekvencije pojavljivanja pojedinačne cifre i poređenja sa očekivanom frekvencijom prema Benfordovom zakonu za tu cifru, a ne da li je pojavljivanje svih cifara u skladu sa očekivanim.

χ^2 test

χ^2 test prevazilazi posmatranje pojedinačnih cifara. χ^2 test je vrlo praktičan test koji može poslužiti onda kad želimo utvrditi da li frekvencije dobijene na osnovu uzorka odstupaju od frekvencija koje bismo očekivali pod određenom pretpostavkom. Konkretno, test posmatra frekvencije dobijene na osnovu uzorka, pokušavajući utvrditi da li je učestalost pojavljivanja svih cifara na vodećoj poziciji u skladu sa Benfordovim zakonom. χ^2 test koristi χ^2-statistiku oblika:

$$\chi^2 = \sum_{i=1}^{9} \frac{(C_i - O_i)^2}{O_i}$$

C_i- frekvencija pojavljivanje cifre i utvrđena ispitivanjem podataka
O_i- teorijska frekvencija pojavljivanja cifre i koja se očekuje Benfordovim zakonom.

Dobijenu vednost χ^2-statistike sa 8 stepeni slobode upoređujemo sa kritičnom p-vrednosti u zavisnosti od željenog stepena poverenja.

Metod χ^2 testa u osnovi je veoma sličan KS testu,rezultati im se zasnivaju na poređenju χ^2-statistike koja se dobija na osnovu podataka i kritične p-vrednosti. Nedostaci su im povezani sa zavisnošću zaključivanja od odbima skupa, odnosno ovaj test se ne smatra korisnim kod testiranja uzoraka malog obima.

Kolmogorov - Smirnov test

Kolmogorov-Smirnov test se zasniva na formiranju Kolmogorov-Smirnov statistike koja predstavlja maksimalnu vrednost odstupanja od Benfordovog zakona, na način da sumira razlike između frekvencija cifara od 1 do 9 koje se dobijaju testiranjem podataka uzorka i frekvencija koje se teorijski očekuju. Matematički zapisano: C_i-frekvencija pojavljivanje cifre i utvrđena ispitivanjem podataka

$$KS = \max \{ |C_1 - O_1|, |(C_1 + C_2) - (O_1 + O_2)|, ..., |(C_1 + C_2 + ... + C_9) - (O_1 + O_2 + ... + O_9)| \}$$

O_i-teorijska frekvencija pojavljivanja cifre i koja se očkuje Benfordovim zakonom

Primer 4. *Primer se oslanja na podatke iz finansijskih izveštaja kompanije "Imlek" u periodu od 2012. do 2015. godine.*

$$KS = \max \{ |0,305 - 0,301|, ..., |(0,305 + ... + 0,044) - (0,301 + ... + 0,046)| \} = 0.04541$$

Prilikom KS testiranja, dobijenu vrednost KS statistike poredimo da kritičnom vrednošću. Kritična vrednost ili p-vrednost sa 95% nivoom poverenja iznosi:

$$p = 1.36/\sqrt{T}$$

Gdje T predstavlja broj elemenata skupa čije se praćnje Benfordove raspodele ispituje. Nedostaci KS testa uočvaju se prilikom ispitivanja skupova velikog obima. Porastom vrednosti T, dolazi do smanjenjenja p-vrednosti, što ukazuje da bi raspodela skupa morala skoro savršno odgovarati Benfordovoj raspodeli, kako ne bismo odbacili nultu hipotezu.

MAD test

MAD (*eng.* Mean absolute deviation) test zasniva se na MAD statistici. MAD statistiku definišmo kao sumu apsolutnih vrednosti razlika između frekvencije pojavljivanja svake od cifara utvrđne na podacima i očekivane frekvencije pojavljivanja prema Benfordovom zakonu. Suma se dodatno deli sa brojem cifara koje su razmatrane na vodećoj poziciji. Matematički zapisano:

$$MAD = \left(\sum_{i=1}^{D} |C_i - O_i| \right) / D$$

C_i-frekvencija pojavljivanje cifre i utvrđena ispitivanjem podataka
O_i-teorijska frekvencija pojavljivanja cifre i koja se očkuje Benfordovim zakonom
D - broj vodećih cifara uzetih u razmatranje.

Primer 5. *Primer se oslanja na finansijske izveštaje kompanije "Imlek" (2012-2015). Razmatraćemo odstupanja svih devet cifara.*

$$MAD = \left(\sum_{i=1}^{9} |C_i - O_i|\right)/9 = 0.00115$$

MAD test predstavlja logično ispitivanje odstupanja svih vodećih cifara od Benfordovog zakona i ne uključuje nikakve kritične vrednosti. Činjenica da MAD statistika ne zavisi od veličine skupa koji se testira čini je korisnim alatom za ispitivanje skupova podataka različitog obima. Posebno, MAD statistika je našla primenu kod ispitivanja istinitosti finansijskih izveštaja, poređenja izveštaja tokom vremena i između preduzeća. Dodatno, MAD statistika se koristi pri sprovođenju testiranja na podacima koji predstavljaju konkretne transakcije, gde se količina podataka koja se ispituje razlikuje od preduzeća do preduzeća.

4.4 FSD - mera istinitosti finansijskih izveštaja

Kako bi se mogao oceniti stepen tačnosti podataka iz finansijskim izveštajima primenom Benfordovog zakona, kreirana je posebna mera FSD (*eng.* Financial statement deviation) koja se oslanja na upotrebu MAD ili KS testa, u zavisnosti od obima testnog skupa. Uzimajući u obzir nezavisnot MAD testa od veličine testnog skupa, u daljim razamtranjima koristimo FSD konstruisan uz pomoć MAD statistike.

FSD meri odstupanje raspodele dobijene na osnovu testiranja uzorka i očekivane raspodele Benfordovog zakona. Formalno zapisano:

$$FSD = \frac{1}{9} * \sum_{d=1}^{9} | \left[\sum_{n=-\infty}^{\infty} \int_{n+\log d}^{n+\log(d+1)} PDF(\log N)dN - (\log(d+1) - \log(d)) \right] |$$

Primetimo, kod raspodela čije su funkcije gustine glatke i simetrične, verovatnoća da broj kao vodeću ima cifru d jednaka je $\log(d+1) - \log(d)23$, odnosno za svaku pojedinačnu cifru nema odstupanja od Benfordove raspodele, pa je srednje apsolutno odstupanje jednako nuli, odnosno vrednost FSD je jedanaka nuli.

Osobine FSD-a

Osobine FSD-a posmatramo kroz prednosti upotrebe ove mere pri procenjivanju finansijskih izveštaja:

- FSD ne ostvaruje povezanost sa karakteristikama poslovanja subjekta na čije finansijske podatke se primenjuje. Odnosno, bilo kakve promene poslovnih aktivnosti i modela rada preduzeća, koje su korektno zabeležene i dokumentovane, neće uticati na FSD vrednost.
- Ne zahteva vremenske serije podataka kako bi se formirala FSD vrednost.
- FSD ne zahteva nikakva predviđanja vezana za budućnost. Svako predviđanje budućih kretanja u osnovi se zasniva na procenama trenutnog stanja i istorijskih podataka, pa ukoliko trenutni podaci sadrže nepravilnosti, za očekivati je da će i predviđanja biti nekorektna.
- FSD je u potpunosti nezavisna od mernih jedinica oblasti na koju se primenjuje.

  Navedene prednosti FSD-a nastale su kao rezultat prevazilaženja poblema sa kojima su se susretali postojeći Modeli procene istinitosti podataka[19] u finansijskim izveštajima koji se zanivaju na principima obračunskog računovodstva[4]. Većina modela je bila blisko povezana sa poslovnim karakteristiama subjekta i zasniva se na podeli troškova na neizbežne i dodatne. Jedan od najpoznatijih je Beneishov model[20] koji se zaniva na korišćenju Beneish M količnika i 8 promenljivih. Osam promenljivih se bazira na podacima

[4] Prihodi se priznaju kada su zarađeni, a rashodi kada se pojave, bez obzira kada se novac naplaćuje ili plaća.

iz finansijskih izveštaja, nakon čega se računa M-skor i poredi sa vrednošču -2,22. Ukoliko je vrednost M-skora manja, podaci su istiniti, dok vrednost veća pd navedene ukazuje na mogućnost nepravilnosti U osnovi, FSD greške ne posmatra kao odstupanja od postojećih Modela procene istinitosti. Naime, koeficijeneti modela se zasnivaju na analizama istorisjkih podataka i formiraju koeficijente modela koji se tokom vremena ne menjaju. Neke bitne promene u poslovnim aktivnostima mogu uticati na promene koeficijenata modela, što može uticati na procenjivanje ispravnosti finansijkih izveštaja. Dakle ovaj problem može kreirati grešku merenja. Moguća je pristrasnost zaključivanja zbog bliskog uticaja načina poslovanja subjekta. Posmatranje raspodele cifara i njeno odstupanje od očekivanih vrednosti u potpunosti eliminište ove probleme.

Detaljnije, primenu FSD i vrednovanje razmatramo kroz primer Digitalne Benfordove analize kompanija "Imlek" i "Apatinska pivara" čija su poslovanja zastupljena i na domaćem i na stranom tržištu.

4.5 Zašto podaci iz finansijskih izveštaja prate Benfordov zakon?

Posmatrajući finansijki izveštaj preduzeća, jasno je da podaci sadržani u izveštaju predstavljaju ocene realizovanih transakcija i stanja subjekta. Dakle, mi nismo u mogućnosti da sprovedemo testiranja na skupu stvarnih transakcija, jer kompanje i preduzeća ne dozvoljavaju pristup takvoj vrsti podataka, osim u slučaju sprovođenja revizije preduzeća. Kako bi Digitalna Benfordova analiza finansijskih izveštaja bila opravdana, pokazaćemo sledeće:

- Pod određenim pretpostavkama računovodstveni podaci iz finansijskih izveštaja zadovoljavaju poseban numerički model.
- Podaci iz finansijskih izveštaja oslobođeni nepravilnosti zadovoljavaju Benfordov zakon.
- Finansijski izveštaji određenih vrsta preduzeća zadovoljavaju Benfordov zakon, uprkos postojećim malverzacijama.
- Numeričke karakteristike tipova grešaka koje uzrokuju odstupanje od Benfordove raspodele.
- Nepravilnosti koja uzrokuju odstupanje od raspodele impliciraju porast vrednosti FSD-a.

Analiza raspodele podataka u finansijskim izveštajima

Zamislimo da se bavimo izradom jednogodišnjeg projekta. Početni vektor sredstava jednak je $X = \{X_1, X_2, ..., X_k\}$ i sastoji se iz slučajnih, različitih tokova(*npr.* X_1- odliv gotovine na račune dobavljača) čiji je vreme realiazacije dve godine, i konstruisan je kao vektor pozitvnih vrednosti.
Pretpostavimo:

- Svaki tok, odnosno element vektora X_i ima log-normalnu raspodelu sa sredinom μ_i i standardnom devijacijom σ_i, $i \in \{1, 2, ..., K\}$. Odnosno,
$$\log(X_i) : N(\mu_i, \sigma_i)$$
- Greška koja nastaje pri procesu ocenjivanja toka, odnosno pri formiranju podatka iz finansijskog izveštaja nezavisna je od samog toka.
- Raspodela greške je log-normalna.

Na kraju godine, pristupa se zakonski obaveznoj izradi finansijkog izveštaja. Dakle, izveštaj u osnovi predstavlja vektor ocena Y svakog$i$-tog toka, odnosno Y_i drugačije nazivamo pozicijama u finansijskim izveštajima .

$$Y = \{Y_1, Y_2, ..., Y_k\}$$

Elementi vektora Y_i u osnovi su:

$$Y_i = X_i * Z_i$$

Z_i predstavlja grešku ocenjivanja X_i u trenutku sprovođena izrade FI. Vektor grešaka ocenjivanja je:

$$Z = \{Z_1, Z_2, ..., Z_k\}$$

Ukoliko je $Z_i = 1$, ne postoje grešeke pri ocenjivanju. Veća vrednost od jedinice ukazuje na precenjenost stvarnog toka X_i, dok manja vrednost od jedinice ukazuje ne podcenjenost. Primetimo,

$$\log(Y_i) = \log(X_i) + \log(Z_i)$$

za vrednost $Z_i = 1$, dobijamo tačnu reprezentaciju $\log(Y_i) = \log(X_i)$. Nametanjem uslova *log*-normalne raspodele na $Z_i : logN(\mu_{i_z}, \sigma_{i_z})$, dobijamo

$$\log(X_i) + \log(Z_i) : N(\mu_i + \mu_{i_z}, \sigma_i + \sigma_{i_z})$$

$$\log(Y_i) : N(\mu_{yi}, \sigma_{yi})$$

Normalna raspodela u odnosu na logaritamsku skalu osigurava da početni skup podataka zadovoljava Benfordovu raspodelu, dakle nametanjem navedenih pretpostavki osiguravamo praćenje Benfordovog zakona.

Ukoliko bismo u jednom trenutku odlučili napraviti presek finansijskog stanja, raspodela vektora X nastala kombinacijom raspodela X_i formalno bi se mogla zapisati:

$$\varphi(X) = \sum_{i=1}^{K} \alpha_i \varphi(X_i)$$

gde je α_i težinski koeficijent pojedinačne raspodele u kombinovanoj raspodeli. Primetimo, raspodela X neće biti normalna raspodela, nezavisno od jednakih srednjih vrednosti. Primenom funkcije gustine normalne raspodele X_i na logaritamskoj skali dobijamo:

$$\varphi(\log(X)) = \sum_{i=1}^{K} \frac{1}{K} \left(\frac{1}{\sigma_i \sqrt{2\pi}} e^{\dfrac{(x-\mu_i)}{2\sigma_i^2}} \right)$$

Dolazi i do promena u formulacije FSD-a : Hill je dokazao da kombinovana ras-

$$FSD = \frac{1}{9} * \sum_{d=1}^{9} | \left[\left(\sum_{n=-\infty}^{\infty} \int_{n+\log d}^{n+\log(d+1)} \sum_{i=1}^{K} \frac{1}{K} \left(\frac{1}{\sigma_i \sqrt{2\pi}} e^{\dfrac{(x-\mu_i)}{2\sigma_i^2}} \right) dX_i \right) - (\log(d+1) - \log(d)) \right] |$$

podela zadovoljava Benfordov zakon, uz nametanje potrebnih dodatnih uslova.[9]

Zanimljiva činjenica je da ukoliko je kombinovana raspodela nastala kombinacijom normalnih raspodela čije su srednje vrednosti manje od dve pojedinačne standardne devijacije, simetričnost i izgled normalne raspodele će se održati. Ovo je osnovni razlog za postojanje malverzacija koje mogu proći neotkriveno, jer se neće narušiti normalnost raspodele u odnosu na logaritamsku skalu, a time ni Benfordovo svojstvo.

4.6 Odstupanje finansijskih izveštaja od Benfordovog zakona

Nakon što smo utvrdili da podaci iz finansijskih izveštaja prate Benfordov zakon, razmotrićemo kako određeni vidovi grešaka utiču na održavanje Benfrdovog svojstva izveštaja . Dakle, greške ćemo podeliti na one koje dovode do odstupanja finansijskih izveštaja od Benfordovog zakona i one koje ne utiču na normalnost raspodele podataka iz izveštaja u odnosu na logaritamsku skalu, gde Benfordovo svojstvo ostaje očuvano.

Greške koje ne impliciraju odstupanje od Benfrdovog zakona

Nakon što smo utvrdili razloge zbog kojih se od podataka u finansijskim izveštajima očekuje praćenje Benfordove raspodele, pretpostavke koje moraju biti zadovoljene i mere koje se primenjuju u procenjivanju, možemo razmotriti osobine greška koje mogu uticati na odstupanja podataka od Benfordove raspodele. Slučajevi u kojima neće doći do odstupnja od Benfordove raspodele:

- Ukoliko su greške nastale na način da se identična greška ponavlja prilikom ocenjivanja svakog od tokova. Odnosno da se greška ponavlja prilikom formiranja svakog Y_i. Ovakvi vidovi grešaka će dovesti do promene sredine raspodele greške $\mu_{zi}+c$ za određenu fiksnu vrednost za koju će se promeniti i odstupanje, pa će novo odstupanje biti jednako $\sigma_{zi}+c$. Što praktično znači da će se raspodele za određeni broj podeoka pomeriti u stranu, ali normalnost raspodele u odnosu na logaritmasku skalu će se zadržati

  $$\log(X_i)+\log(Z_i) : N(\mu_i+\mu_{i_z+c}, \sigma_i+\sigma_{i_z}+c)$$

  Dakle, nepravilnosti koje se pojavljuju kod svih tokova, na identičan način i stvaraju jednake razlike neće se detektovati kao prevare. Izvodi se zaključak da greške nastale kao posledica grešaka u formirnaju ocenjivača neće uzrokovati odstupanja, što se smatra jednom od prednosti.

- Ukoliko vektor grešaka utiče samo na odstupanje raspodele , ali ne i na sredinu raspodele, rapsodela podataka iz finansijskih izveštaja se neće razlikovati mnogo od normalne raspodele bez uticaja na odstupanje, odnosno zadržaće se osobine normalnosti, ali uz nešto širi ili uži opseg, što implicira da na praćenje Benforodovog zakona. Vrednost FSD skora ostaće male, bliske nuli.

Greške koje impliciraju odstupanje od Benfordovog zakona

Promatrajući greške čija su standardna odstupanja konstantna, ali sredine različite, Benfordovo svojstvo podataka biće narušeno. Razlikujemo tri slučaja formiranja ovakvih nepravilnosti.

- Posmtrajmo slučaj kada greška predstavlja precenjivanje ili potcenjivanje neke of pozicija u finansijkom iveštaju Y_i. Konkretno, grešku koja je nastala kao uvećanje sredine $\mu_{zi} + c$. Dakle, promena u sredimi greške će uticati na vrednost sredine Y_i, dok će standardno odstupanje ostati nepromenjeno. Na ovaj način narušava se oblik normalne raspodele u odnosu na logaritmsku skalu, ujedno dolazi do narušavanja Benfordovog svojstva pozicija iz finansijskih izveštaja, praćeno rastom FSD-a.

- Situacija u praksi najčešće nalaže blisku povezanost grešaka u finansijskim izveštajima. Kroz primer, ukoliko se utiče na izmenu podataka povezanih sa prihodima, nužno će doći i do promena podataka povezanih sa poreskim zaduženjima. Povezanost podataka u finansijskim izveštajima implicira da pojava greške povezane sa jednim od ocenjivača toka Y_i usloviće promenu sredine drugog Y_i, što će uticati na normalnost raspodele dva i-ta toka i na taj način narušiti zadovoljavanje Benfordove raspodele od strane kombinovane raspodele. Vrednost FSD-a će svojim povećeanjem ispratiti nastale promene.

- Moguće je da su greške ocenjivanja povezane sa vrednostima tog toka. Odnosno, sredina greške može zavisiti od sredine raspodele toka. Formalno:

  $$\mu_{zi} = \psi(\mu_i)$$

  Tokovi, čije su sredine veće usloviće veća odstupanja sredine ocenjivača i na taj način narušiti normalnost raspodele toka.

Kao zaključak prethodnog razamatranja nameće se da odstupanja koja su blisko povezana za sredine raspodele pri čemu ne dolazi do promena u standardnim odstupanjima dovešće do narušavanja pretpostavke o normalnosti raspodele, odnosno o praćenu Benfordove zakonitosti od strane finansijskih izveštaja. Nasuprot tome, jednake promene sredina i odstupanja ili promene odstupanja, pri konstantnim sredinama neće narušavati normalnost raspodele pojedinačnih tokova, pa ni Benfordovo svojstvo kombinovane raspodele. U situacijama u kojima Benfordov zakon ostaje očuvan, FSD skor ima niske vrednosti bliske nuli Dakle, porastom grešaka prilikom ocenjivanaj dolazi do povećanja vrednosti FSD-a.

4.7 Efikasnost Digitalne Benfordove analize

Rezultati Digitalne Benfordove analize (DBA) mogu poslužiti isključivo kao indikator mogućih nepravilnosti, ali negativan rezultat praćenja raspodele podataka iz finansijskih izveštaja može biti uslovljen greškama u zaokruživanju, odnosno nekim procesom prilikom formiranja vrednosti iz finansijskih izveštaja koje su u osnovi ocene realnih tokova. Ukoliko bi se greška sprovedena na stvarnim podacima transakcija dovoljno puta ponovila, mogla bi se preslikati na vrednosti iz finansijskih izveštaja i na taj način dovesti do odstupanja.

Dakle, Digitalna Benfordova analiza može poslužiti kao jednostavan analitički alat, čiji rezultati mogu usmeriti postupak formiranja revizorske strategije. Pozitivno praćenje zakona ne može poslužiti kao krajnji dokaz ravizijskog postupka, nego je potrebno sprovesti dalje revizorske analize, ali nam govori sa kolikim stepenom obazrivosti pristupamo daljim postupcima. Suprotno tome, negativno praćenje zakona, daje jasne naznake mogućih nepravilnosti i utiče na revizorsku svest pri pristupanju detaljnijim ispitivanjima.

Uzimajući u obzir značajnost gubitaka u slučaju neotrkivanja prevara, potrebno je detaljno razmotriti i ograničenja prilikom zaključivanja na osnovu dobijenih rezultata. Osnovna ograničenja sa kojima se susreće postupak DBA su:

- Pažljivo tumačenje rezultata statističkih testova.
- DBA se može primenjivati samo na Benfordove skupove podataka.
- Određeni tipovi grešaka ne mogu biti otkriveni postupkom DBA.

Potpunim razumevanjem Benfordovih skupova, pažljivim sprovođenjem DBA i poznavanjem ograničenja ove analize, moguće je efikasno usmeriti dalji revizijski proces.

Osnovno pitanje koje se nameće jeste sposobnost DBA da uspešno detektuje postojanje prevare. Pri započinjanju revizijskog procesa upotrebom DBA, dalje revizijske odluke se susreću sa dve vrste rizičnosti:

- Nepoznato je koliko će rezultati DBA biti tačni pri upotrebi na stvarnim podacima.
- Nepoznata je sklonost uzorka koji se ispituje (*npr.* finansijskih izveštaja konkretnog preduzeća) ka mogućim prevarama i ova sklonost meri se verovatnoćom prevare posmatranog uzorka.

Najednostavniji način utvrđivanja ispravnosti rezultata DBA ogleda se u poređenju dobijenih rezultata pri saznanju da postoje nepravilnosti i rezultata uzoraka oslobođenih nepravilnosti. No, praktično ne bismo ni provodili postupak DBA da sa sigurnošću možemo tvrditi koji to podaci sadrže nepravilnosti, a koji ne sadrže. Ni generalnu sklonost podataka određenog poslovnog subjekta ka prevarama nije moguće jednostavno otkriti, a osnovni razlog predstavlja odbijanje subjekata da pruže uvid u postojeće informacije o nepravilnostima, i u slučajevima kada se utvrdi da su podaci iz finansijskih izveštaja istiniti. Možemo zakjučiti da o postojanju prevara govorimo na nivou ličnih procena revizora i nivou špekulacija.

Kako bismo razmotrili koliko u ovakvim okolnostima DBA može biti korisna, poslužićemo se formulom uslovne verovatnoće. Ukoliko je DBA signalizirala postojanje prevare, verovatnoća da usled postojećih neizvesnosti prevara zaista postoji jednaka je:

$$P(N|D) = \frac{P(D|N) * P(D)}{P(N) * P(D|N) + P(\overline{N}) * P(D|\overline{N})} = \frac{P(D|N) * P(D)}{P(D)}$$

Navedene događaji čije verovatnoće računamo su:
N- prisustvo prevare;
$\overline{N}$- odsustvo prevare;
D- detektovnje prevare primenom DBA.

Navedeni formalni zapis nam govori da je verovatnoća prisutnosti prevare ukoliko je prevara detektovana DBA, odnosno verovatnoća da DBA daje tačne signale neispravnosti podataka, jednaka količniku verovatnoće da DBA detektuje prevaru ukoliko ona stvarno postoji pomnoženim sa verovatnoćom da podaci sadrže prevaru i verovatnoćom detektovanja (signaliziranja) prevare. Verovatnoća detektovanja prevare u ozbaci $P(D)$ predstavlja sumu procentulanih vrednosti u kojima je DBA korekto detektovala prevaru (DBA detektovao prevaru, prevara prisutna) i slučajeva u kojima je DBA dao pogešne rezultate (DBA detektovao prevaru,a prevara nije postojala). Sumirano, verovatnoća postojanja prevara, usled njenog detektovanja od strane DBA, odnosno verovatnoća korisnosti DBA jednaka je količniku tačnih detekcija prevare i ukupnih detekcija.

Pri procesu računanja, verovatnoća postojanja prevare $P(N)$ igra ključnu ulogu, a njena stvarna vrednost je nepoznata usled:

- Preduzeća ne žele da otkrivaju slabosti vezane za sam proces transakcija, slabosti prilikom ocenjivanja, kao ni da su eventualno sami bili žrtve određenih prevara.
- Informacije supervizora o prevarama formirane su ikljičivo na prevarama koje su sami otkrili ili prevarama koje su javno objavljene.
- Nije moguće utvrditi opseg prevara.

Dakle, verovatnoća postojanja prevare ostavlja se na određivanje revizorima, koji dodatno pažnju moraju obratiti na privredna okruženja podložnija malverzacijama i prevarama.

Proces procenjivanja efikasnosti DBA razmatramo kroz primer:

Primer 6. *Pretpostavimo sledeće:*
Verovatnoća postojanja nepravilnosti jednaka $P(N) = 0.04$ *Kako smo prethodno pojasnili ova vrednost se zasniva na saznanjima, iskustvu revizora i okolnostima transakcija sa kojima su povezani podaci.*
Verovatnoća da DBA uspešno detektuje nepravilnosti je $P(D|N) = 0.80$ *Praktično, 80% vremena DBA će uspešno signalizirati prevaru koju čini 4% transakcija uzorka.*

$$P(N|D) = \frac{P(D|N) * P(D)}{P(N) * P(D|N) + P(\overline{N}) * P(D|\overline{N})} = \frac{0.8 * 0.04}{(0.04 * 0.8) + (0.96 * 0.2)} = 0.1429$$

Dakle, uslovna verovatnoća da će prevara biti otkrivena ukoliko je DBA detektovala nepavilnosti jednaka je 0.1429, što znači da su nam šanse za otkrivanjem pravera približno 14% . Odnosno, da će 80%vremena DBA uspešno detektovati prevaru zastupljenu u 4% transakcija skupa na kome se sprovodi analiza. Ova uslovna verovatnoća jednaka je količniku broja uspešnih i ukupnih detekcija DBA, pa se koristi kao mera korisnosti DBA.

Glava 5

Uporedna Digitalna Benfordova analiza finansijkih izveštaja kompanija "Imlek" i "Apatinska pivara"

Postupak Digitalne Benfordove analize sprovodimo nad podacima iz finansijskih izveštaja kompanija "Imlek" i "Apatinska pivara". Izveštaji su peuzeti iz baze Agencije za privredne registre[27].

Razlozi uzimanja upravo ovih komapanija ogledaju se u njihovom poslovanju koje uključuje poslovanje u zemlji i inostranstvu, ali i prepoznatljivoti brendova kompanija. Broj pozicija iz finansijskih izveštaja u Srbiji nije najpogodniji za jednogodišnja testitanja, pa smo prinuđeni na razmatranje finansijskih izveštaja u rasponu od 2012. do 2015. godine. Okolnosti razmatranja finansijskih izveštaja koji uključuju period duži od jedne godine mogu biti posebno korisne. Na ovaj način teret mogućih malverzacija sprovedenih u godinama pre trenutka sprovođenja poslednje revizije ostaje očuvan, što predstavlja dodatnu sigurnost revizorima u rezultate analize, jer detektuje opštiju sklonost preduzeća ka iznošenju neistinitih tvrdnji. Vrednosti FSD-a u odnosu na MAD testove i KS testove se svodi na računanje prostih MAD i KS statistika.

Iz izveštaja isključujemo vrednosti nule i na taj način prilagođavamo skupove. Finansijski izveštaji se sastoje iz pozitivnih vrednosti, pa taj vid prilagođavanja skupa ne moramo sprovoditi.

5.1 Prate li finansijski izveštaji kompanija "Imlek" i "Apatinska pivara" Benfordov zakon?

Pre nego razmotrimo praćenje Benfordove raspodele, poželjno je sprovesti deskriptivnu analizu finansijskih izveštaja.

Analaiza finansijskih izveštaja kompanije "Imlek"

Kao rezultat razmatranja frekvencija pojavljivanja cifara na vodećim pozicijama finansijskih izveštaja kompanije "Imlek" dobili smo:

Cifra	203	Frekvencije (Imlek)	Frekvencije (BZ)	MAD	HI	KS
1	62	0,305418719	0,301029996	0,00438872	0,9999999997633	0,004388724
2	29	0,142857143	0,176091259	0,03323412	0,9999999997633	0,028845393
3	22	0,108374384	0,124938737	0,01656435	0,9999999997633	0,045409745
4	23	0,113300493	0,096910013	0,01639048	0,9999999997633	0,029019265
5	16	0,078817734	0,079181246	0,00036351	0,9999999997633	0,029382777
6	15	0,073891626	0,06694679	0,00694484	0,9999999997633	0,022437941
7	14	0,068965517	0,057991947	0,01097357	0,9999999997633	0,011464371
8	13	0,064039409	0,051152522	0,01288689	0,9999999997633	0,001422515
9	9	0,044334975	0,045757491	0,00142252	0,9999999997633	3,46945E-17
				0,01146322	**0,9999999997633**	**0,045409745**

Tabela 5: Deskriptivna statistika:Finansijski izveštaji kompanije "Imlek" (2012-2015)

Sprovodi se Digitalna Benfordova analiza upotrebom tri testa:

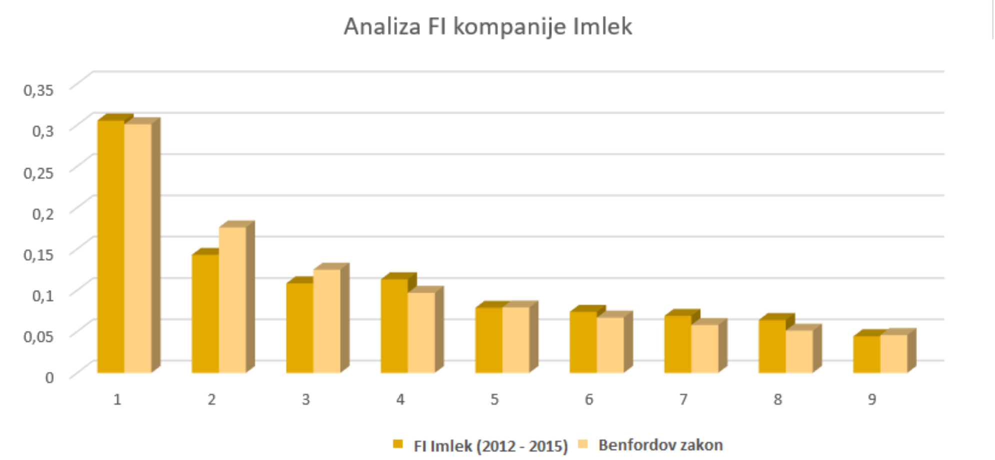
- MAD test: Rezultati MAD testa ukazuju da je vrednost MAD statistike niska 0,0115 , što implicira održavanje hipoteze o praćenju Benfordove raspodele. Ujedno, ova vrednost se uzima kao vrednost FSD-a. U skladu sa izloženom teorijom, niska vrednost FSD-a upućje na isinitost prikazanih podataka.

- χ^2 test: Kritična p-vrednost χ^2 za 8 stepeni slobode iznosi 2,753 sa 95% nivoom poverenja , a vrednost test statistike jednaka je jako blizu 1, što ukazuje na prihvatanje nulte hipoteze o praćenju Benfordove raspodele.

- KS test: Vrednost KS statistike iznosi 0,04541, dok je kritična vrednost jednaka 0,09545, pa je i KS test potvrdio održavanje Benfordove zakonitosti, odnosno prihvatanje nulte hipoteze o odsustvu nepravilnosti.

Najjednostavniji i najbrži način pružanja indikacija o prisustvu/odsustvu nepravilnosti predstavlja izrada grafika raspodela. Izradom grafičke reprezentacije raspodele učestalosti vodećih cifara testiranih podataka uočavamo da su raspodele izuzetno bliske, odnosno da su odstupanja od Benfodove zakonitosti vizuelno mala.

5.1 Prate li finansijski izveštaji kompanija "Imlek" i "Apatinska pivara" Benfordov zakon?



Slika 5: Finansijski izveštaji kompanije "Imlek" (2012-2015) u poređenju sa Benfordovim zakonom

Analiza finansijskih izveštaja kompanije "Apatinska pivara"

Analogno, na podacima iz finansijskih izveštaja kompanje "Apatinska pivara" sprovode se testiranja učestalosti pojavljivanja cifara na vodećoj poziciji. Rezultati statističkih testova:

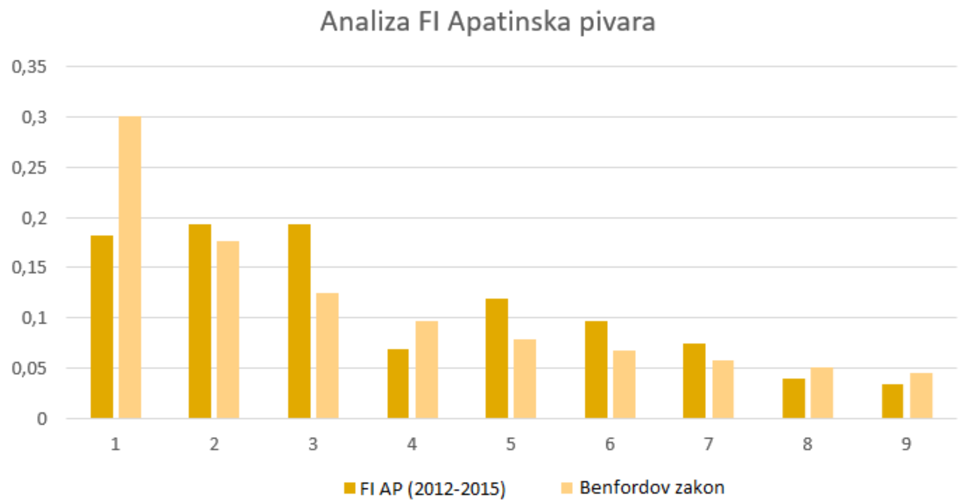
Cifra	176	Frekvencije (AP)	Frekvencije (BZ)	MAD	HI	KS
1	32	0,181818182	0,301029996	0,119211814	0,999999107	0,1192118138457990000
2	34	0,193181818	0,176091259	0,017090559	0,999999107	0,1021212547196620000
3	34	0,193181818	0,124938737	0,068243082	0,999999107	0,0338781731461442000
4	12	0,068181818	0,096910013	0,028728195	0,999999107	0,0626063679723824000
5	21	0,119318182	0,079181246	0,040136936	0,999999107	0,0224694322018255000
6	17	0,096590909	0,06694679	0,029644119	0,999999107	0,0071746872584704500
7	13	0,073863636	0,057991947	0,015871689	0,999999107	0,0230463766444202000
8	7	0,039772727	0,051152522	0,011379795	0,999999107	0,0116665814697662000
9	6	0,034090909	0,045757491	0,011666581	0,999999107	0,0000000000000002776
	176			**0,037996975**	**0,999999107**	**0,119211814**

Tabela 6: Deskriptivna statistika: Finansijski izveštaji kompanije "Apatinska pivara" (2012-2015)

- MAD test: Rezultati MAD testa ukazuju da vrednost MAD statistike iznosi 0,038, što je veća vrednost u odnosu na kompaniju "Imlek", dakle naznačava dodatno nepoverenje u istinitost podataka.
- χ^2 test: Kritična p-vrednost sa 95% nivoom poverenja i osam stepeni slobode iznosi 2,753, a vrednost test statistike približno je jednaka 1, što ukazuje na prihvatanje nulte hipoteze o praćenju Benfordove raspodele.
- KS test: Vrednost KS statistike iznosi 0,119, dok je kritična vrednost jednaka 0,102, pa je KS test odbio nultu hipotezu o praćenu Benfordovog zakona u korist hipoteze o odstupanju od raspodele, odnosno KS je podatke iz FI označio kao podatke sa mogućim nepravilnostima.

Dakle, rezultati testova nisu dali jedinstven zaključak.

5.1 Prate li finansijski izveštaji kompanija "Imlek" i "Apatinska pivara" Benfordov zakon?



Slika 6: Finansijski izveštaji kompanije "Apatinska pivara" (2012-2015) u poređenju sa Benfordovim zakonom

Grafička interpretacija , ukazuje da podaci iz finansijskih izveštaja imaju sličnosti, ali ne prate u potpunosti Benfordov zakon.
Posebno odstupanje se uočava kod frekventnosti cifre 1. Tumačenje grafičke interpretacije, dalo je ideju za upotrebu Z-testa, odnosno testa koji ispituje pojedinačno slaganje učestalosti pojave cifara na vodećim pozicijama. Sa 95% nivoom

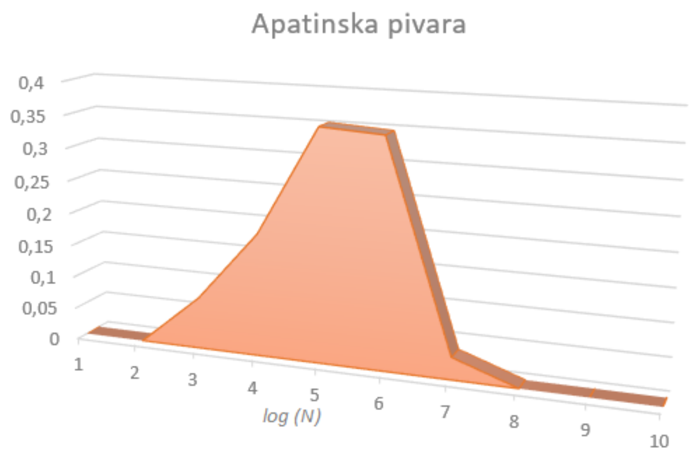
Cifre	F(AP)	BZ	Z
1	0,1818182	0,30103	3,18673413
2	0,1931818	0,1760913	0,60152803
3	0,1931818	0,1249387	2,8515341
4	0,0681818	0,09691	1,26827886
5	0,1193182	0,0791812	2,01641475
6	0,0965909	0,0669468	1,59914477
7	0,0738636	0,0579919	0,90856955
8	0,0397727	0,0511525	0,68119309
9	0,0340909	0,0457575	0,73620833

Tabela 7: Z-test: Finansijski izveštaji kompanije "Apatinska pivara" (2012-2015)

poverenja, uočavamo da cifre 1 i 3 značajno odstupaju od Benfordove zakonitosti, gde Z-statistika prelazi kritičnu vrednost od 2,77. U slučajevima kada jedan od testova, ukaze na odstupanje od Benfordove raspodele poželjno je razmotriti osobine testova i moguće razloge odstupanja koji su najćešće povezani sa obimom testnog skupa.

Kako je obim skupa od 203 elementa pogodan za upotrebu KS testa, koji ne gubi svoju efikasnost, rezultate ovoga testa smatraćemo relevantnim. Odnosno, Digitlana Benfordova analiza nam je pružila neslaganje sa Benfordovim zakonom i naznake za prisustvo nepravilnosti. Potrebno je podsetiti da DBA ne daje popunu sigurnost zaključivanja, nego može poslužiti samo kao indikator.

Razmotrimo kroz prizmu prethodno pojašnjenog teorijskog stanovišta da ukoliko je skup podataka normalno raspodeljen u odnosu na logaritamsku skalu zadovoljavaće Benfordrovo svojstvo. U primeru kompanije "Apatinska pivara" na grafiku podataka u odnosu na logaritamsku skalu možemo uočiti da postoje određene promene koje narušavaju normalnost raspodele za šta se može pretpostaviti da je posledica postojećih grešaka koje su dovele do odstupanja od zakonitosti.



Slika 7: Finansijski izveštaji kompanije "Apatinska pivara" (2012-2015) u odnosu na logaritamsku skalu

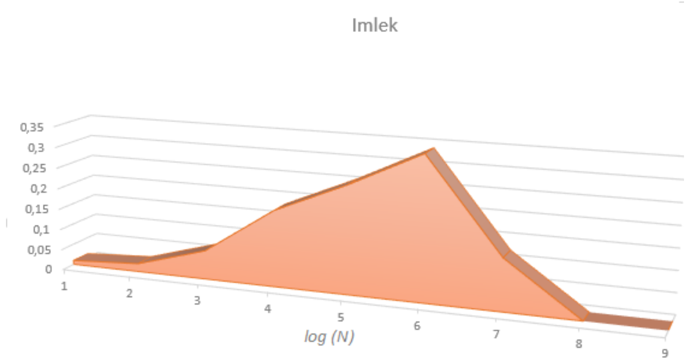
5.2 Manipulacija nad podacima iz finansijskih izveštaja komapanije "Imlek"

Prilikom upotrebe Benfordovg zakona, svesnost o njegovoj relativnoj tačnosti uvek mora postojati. Verovatnoće tačne detekcije DBA su osnova svih daljih razmatranja i zaključivanja.

Kako bismo stekli bolji uvid u povezanost količine izmenjnih podataka i stepena efikasnosti Digitalne Benfordove analize u detekciji nepravilnosti, sproveli smo namerne manipulacije vodećih cifara u finansijskim izveštajima za koje su prethodno izvedeni statistički testovi pokazali praćenje Benfordove raspodele, odnosno odsustvo malverzacija.

Posledice različitog obima manipulacija pratićemo kroz vrednosti FSD upotrebom MAD statistike, a ujedno ćemo određivati efikasnost DBA u daljem revizijskom procesu u zavisnosti od procenata promena koje smo napravili. Potrebno je naglasiti da će se promene na početnim pozicija sprovoditi po principu promena za jednu cifru na gore, odnosno jednu cifru na dole, kako bi se poboljšale performase preduzeća. Prvenstveno izmene će se sprovoditi na podacima koji su vrednosno bliski većim, odnosno manjim početnim ciframa.

Pre sprovođenja manipulacija, izgled raspodele kompanije "Imlek" u odnosu na logaritamsku skalu je sledeći:

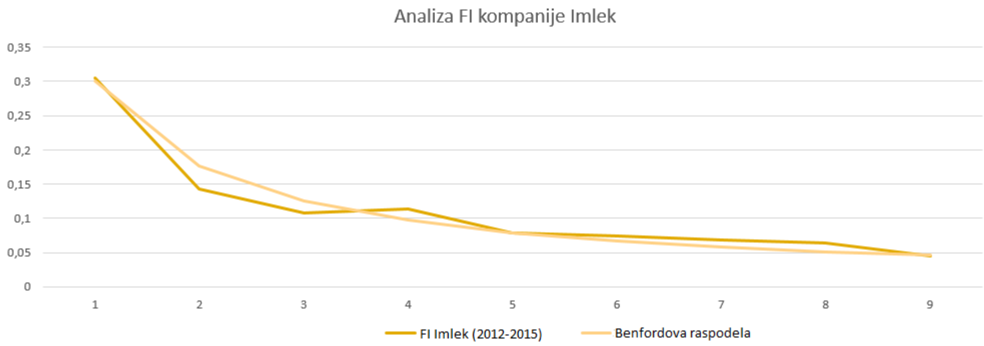


Slika 8: Finansijski izveštaji kompanije "Imlek" (2012-2015) u odnosu na logaritamsku skalu

Raspodela ima oblik blizak normalnosti, što je dodatna potvrda ispravnosti sprovedenih testova.

Podsetimo se početna vrednost FSD-a iznosi 0,0115. FSD će nam poslužiti kao glavna mera spovođenja manipulacija, porastom broja manipulacija, očekuje se njegov rast.

Slaganje sa Benfordovim zakonom je prikazano i na slici(9).



Slika 9: Finansijski izveštaji kompanije "Imlek" (2012-2015) u odnosu na Benfordov zakon

Sprovođenje manipulacija u 5% obimu

Manipulacija nad 5% podataka finansijskih izveštaja kompanije "Imlek" iz perioda (2012-2015), nije dovela do značajnih promena u rezultatima testova.

Cifra	203	Frekvencije (Imlek)	Frekvencije (BZ)	MAD	HI	KS
1	58	0,285714286	0,301029996	0,01531571	0,9999999998029	0,01531571
2	30	0,147783251	0,176091259	0,02830801	0,9999999998029	0,043623718
3	24	0,118226601	0,124938737	0,00671214	0,9999999998029	0,050335853
4	23	0,113300493	0,096910013	0,01639048	0,9999999998029	0,033945374
5	15	0,073891626	0,079181246	0,00528962	0,9999999998029	0,039234994
6	17	0,083743842	0,06694679	0,01679705	0,9999999998029	0,022437941
7	14	0,068965517	0,057991947	0,01097357	0,9999999998029	0,011464371
8	12	0,0591133	0,051152522	0,00796078	0,9999999998029	0,003503593
9	10	0,049261084	0,045757491	0,00350359	0,9999999998029	3,46945E-17
				0,01236122	**0,9999999998029**	**0,050335853**

Tabela 8: Finansijski izveštaji kompanije "Imlek" (2012-2015) sa izmenom 5% vodećih cifara

Vrednost FSD-a, odnosno MAD statistike povećala se na 0,012, te ovaj rast prati pojavu nepravilnosti. Testovi i dalje rezultiraju praćenje Benfordove raspodele, pošto je vrednost statistika manja od kritičnih vrednosti, pa nulta hipoteza o praćenju raspodele se održava.

Sprovođenje manipulacija u 10% obimu

Ukoliko namerno izazovemo malverzacije na desetini vodećih cifara, vrednost FSD(MAD) se dodatno povećala na 0,0173, dok testovi još uvek podržavaju nultu hipotezu praćenja raspodele. Konkretno, KS statistika je i dalje manja od kritične vrednosti 0,095.

Cifra	203	Frekvencije (Imlek)	Frekvencije (BZ)	MAD	HI	KS
1	52	0,256157635	0,301029996	0,04487236	0,9999999967581	0,04487236
2	32	0,157635468	0,176091259	0,01845579	0,9999999967581	0,063328151
3	23	0,113300493	0,124938737	0,01163824	0,9999999967581	0,074966395
4	24	0,118226601	0,096910013	0,02131659	0,9999999967581	0,053649807
5	18	0,088669951	0,079181246	0,0094887	0,9999999967581	0,044161103
6	13	0,064039409	0,06694679	0,00290738	0,9999999967581	0,047068483
7	16	0,078817734	0,057991947	0,02082579	0,9999999967581	0,026242696
8	15	0,073891626	0,051152522	0,0227391	0,9999999967581	0,003503593
9	10	0,049261084	0,045757491	0,00350359	0,9999999967581	3,46945E-17
				0,01730528	**0,9999999967581**	**0,074966395**

Tabela 9: Finansijski izveštaji kompanije "Imlek" (2012-2015) sa izmenom 10% vodećih cifara

Sprovođenje manipulacija u 15% obimu

Pomenom 15% vodećih cifara testiranog skupa doći će do značajnih promena u razultatima Digitalne Benfordove analize.

Cifra	203	Frekvencije (Imlek)	Frekvencije (BZ)	MAD	HI	KS
1	48	0,236453202	0,301029996	0,06457679	0,9999999686896	0,064576794
2	31	0,15270936	0,176091259	0,0233819	0,9999999686896	0,087958693
3	23	0,113300493	0,124938737	0,01163824	0,9999999686896	0,099596937
4	24	0,118226601	0,096910013	0,02131659	0,9999999686896	0,078280349
5	19	0,093596059	0,079181246	0,01441481	0,9999999686896	0,063865536
6	13	0,064039409	0,06694679	0,00290738	0,9999999686896	0,066772917
7	17	0,083743842	0,057991947	0,0257519	0,9999999686896	0,041021021
8	17	0,083743842	0,051152522	0,03259132	0,9999999686896	0,008429702
9	11	0,054187192	0,045757491	0,0084297	0,9999999686896	3,46945E-17
				0,02277874	**0,9999999686896**	**0,099596937**

Tabela 10: Finansijski izveštaji kompanije "Imlek" (2012-2015) sa izmenom 15% vodećih cifara

Vrednsost FSD-a 0,022 naznačava skok u nepravilnosti, dok KS statistika iznosi 0,099 i vrednost je prerasla kritičnu vrednost od 0,095, odnosno ukazala je na odstupanje od Benfordovog svojstva. Dok χ^2 test i dalje ne uvećava vrednost svoje statistike, odnosno ne detektuje odstupanja. Razlog sprovođenja oba testa, kao i razlika u rezultatima može se objasniti posledicom obima skupa koji se ispituje. Naprosto, obim od 203 podatka ne narušava pitanje KS testa koji je osetljiv na velike skupove, dok χ^2 test nije najbolje primenjiv pri testiranju skupova ovog obima, pa njegove značajne promene nije ni trebalo očekivati.

Ustanovili smo da je detektovanje nepravilnosti primenom DBA ostvareno tek pri 15% obimu malverzacija. Postavlja se pitanje efikasnosti DBA, odnosno tačne detekcije nepravilnosti koja bi bila potvrđena daljim revizijskim procesom. Ukoliko pretpostavimo da je verovatnoća detekcije nepravilnosti upotrebom DBA jednaka 0,75.

Efikasnosti DBA pri uspešnom signaliziranju nepravilnosti u obimu koji smo sami izazavali (5%,10%,15%) jednaka je:

P(N)	*Efikasnost DBA*
0,05	0,109375
0,1	0,197183
0,15	0,296230

Tabela 11: Efikasnost DBA u zavisnosti od verovatnoće greške

Dakle, dobili smo još jednu potvrdu da značajna količina izmenjenih podataka utiče na samu efikasnost testa odnososno veće obim nepravilnosti, povećavaju šansu uspešnog detektovanja primenom Benfordove zakonitosti.

Takođe, vrednosti FSD-a pratiće procentualni rast sprovedenih manipulacija:

P(N)	*FSD*
0	0,0115
0,05	0,012
0,1	0,017
0,15	0,022

Tabela 12: Vrednosti FSD-a u zavisnosti od obima izmenjenih podataka

5.3 Efikasnost DBA u slučaju kompanije "Apatinska pivara"

Postojeća saznanja o ispravnosti poslovanja komanije "Apatinska pivara" su javno dostupna kroz revizorske izveštaje. Godišnji revizorski izveštaji u periodu od 2012. do 2015. ukazuju na pozitivna revizorska mišljenja, dakle revizijski proces nije ustanovio značajne materijalne greške, niti postojanje bilo kakvih nepravilnosti. Sprovođenje Digitalne Benfordove analize, ukazalo se na značajna odstupanje, posebno u zastupljenosti cifre 1. Pokušaćemo da na temeljima određenih pretpostavki razmotrimo efikasnost DBA u pogledu otkrivanja stvarnh nepravilnosi u kompaniji "Apatinska pivara".

Pretpostavke:

- Verovatnoće postojanja prevare 0,05 (odnosno 5%obim nepravilnosti), 0,10 (odnosno 10%obim nepravilnosti) i 0,15 (odnosno 15%obim nepravilnosti).
- Verovatnoća uspešnog detektovanja nepravilnsti upotrebom DBA jednaka je 0,80.

Primenom pretpostavki, efikasnost DBA pri 5% obimu nepravilnosti jednaka je:

$$P(N|D) = \frac{P(D|N) * P(D)}{P(N) * P(D|N) + P(\overline{N}) * P(D|\overline{N})} = \frac{0,8 * 0,05}{(0,05 * 0,8) + (0,95 * 0,2)} = 0,173913043$$

Uspešnost DBA u detekciji pevara u 10% obimu jednaka je:

$$P(N|D) = \frac{P(D|N) * P(D)}{P(N) * P(D|N) + P(\overline{N}) * P(D|\overline{N})} = \frac{0,8 * 0,1}{(0,1 * 0,8) + (0,9 * 0,2)} = 0,307692308$$

Uspešnost DBA u detekciji prevare 15% obima jednaka je: Dakle, kako je DBA

$$P(N|D) = \frac{P(D|N) * P(D)}{P(N) * P(D|N) + P(\overline{N}) * P(D|\overline{N})} = \frac{0,8 * 0,15}{(0,15 * 0,8) + (0,85 * 0,2)} = 0,413793103$$

detektovala nepravilnosti, pod pretpostavkom da je prevara u obimu od 5% verovatnoća da će kasniji revizijski koraci otkriti nepravilnosti jednaka je 0,17. Pretpostavka da je verovatnoća postojanja nepravilnosti jednaka 0,1 uputiće nas na 30% efikasnost DBA. Dok pretpostavka o obimu nepravilnosti u iznosu od 15% jednaka je 40%. Ove procentualne vrednosti ukazuju na ograničeno poverenje u rezultate dobijene DBA.

5.4 Dodatne analize u slučaju kompanije "Apatinska pivara"

Upitna istinitost podataka iz finansijskih izveštaja primenom DBA, podstakla nas je da sprovedemo dodatni eksperiment. Ukoliko bi smo početni skup od 176 elemenata koji predstavljaju pozicije iz finansijskih izveštaja u rasponu od 2012. do 2015. ponovo testirali primenom DBA, ali tako što ćemo izvršiti podelu početnog skupa. Poznato je da skup koji se testira vezano za Benfordovu zakonitost ne bi smeo da sadrži manje od 100 elemenata, pa smo proveli DBA nad prve 102 vrednosti (I skup) i poslednje 102 vrednosti (II skup)[1]. Spovođenjem podele testnog skupa pokušava se utvrditi koji deo podataka sugeriše na neistinitost evidentiranja.

Digitalna Benfordova analiza I skupa

Sprvođenjem DBA na I skupu, dobijeni su sledeći rezultati:

Cifra	102	Frekvencije (AP)	BZ	MAD	KS
1	15	0,147058824	0,30103	0,1539712	0,1539711721345690
2	21	0,205882353	0,17609126	0,0297911	0,1241800782490740
3	21	0,205882353	0,12493874	0,0809436	0,0432364619161976
4	5	0,049019608	0,09691001	0,0478904	0,0911268670811169
5	10	0,098039216	0,07918125	0,018858	0,0722688974424671
6	12	0,117647059	0,06694679	0,0507003	0,0215686282495508
7	10	0,098039216	0,05799195	0,0400473	0,0184786404590369
8	7	0,068627451	0,05115252	0,0174749	0,0359535689920477
9	1	0,009803922	0,04575749	0,0359536	0,0000000000000001
				0,0528478	**0,153971172**

Tabela 13: DBA I dela finansijskih izveštaja kompanije "Apatinska pivara" (2012-2015)

Kritična p-vrednost KS testa iznosi 0,135, dok je vrednost KS-statistike 0,154 što ukazuje na odbacivanje nulte hipoteze u korist alternativne hipoteze koja zagovara odsustvo praćenja Benfordovog zakona. Uočavamo da je vrednost FSD(MAD) u iznosu od 0,052 prilično visoka.

[1]Presek ova dva skupa je neprazan

Digitalna Benfordova analiza II skupa

Sprovođenjem DBA na II skupu, dobijaju se sledeći rezultati: Vrednost KS-

Cifra	102	Frekvencje (AP)	BZ	MAD	KS
1	24	0,235294118	0,30103	0,0657359	0,06573587801692240
2	16	0,156862745	0,1760913	0,0192285	0,08496439197456440
3	20	0,196078431	0,1249387	0,0711397	0,01382469721031530
4	9	0,088235294	0,09691	0,0086747	0,02249941610072460
5	14	0,137254902	0,0791812	0,0580737	0,03557423981243480
6	7	0,068627451	0,0669468	0,0016807	0,03725490116221370
7	5	0,049019608	0,0579919	0,0089723	0,02828256202766430
8	1	0,009803922	0,0511525	0,0413486	0,01306603885108960
9	6	0,058823529	0,0457575	0,013066	0,00000000000000003
				0,0319911	**0,084964392**

Tabela 14: DBA II dela finansijskih izveštaja kompanije "Apatinska pivara" (2012-2015)

statistike od 0,085 za drugi deo podataka iz finansijskih izveštaja manja je u odnosu na kritičnu vrednost koja i za ovaj skup obima 102 elementa iznosi 0,135. Kolmogorov-Smirnov test je potvrdio praćenje Benfordovog zakona, odnosno ukazuje na odsustvo nepravilnosti na II skupu. Vrednost FSD(MAD) je manja u odnosu na vrednost I skupa i iznosi 0,032.

Konačno, rezultati Digitalne Benfordove analize ukazali su na neistinitosti podatka iz FI (2012-2015) komapanije "Apatinska pivara". Dodatna analiza ukazala je da problem postoji kod podataka iz prve polovine posmatranog perioda. Ponavljamo, rezultate dobijene DBA ne možemo uzeti konačan dokaz neistinitosti, već kao značajan signal opreznosti pri daljim ispitivanjima.

Glava 6

Zaključak

Master radom pod nazivom "Primena Benfordovog zakona u reviziji" na najjednostavniji način obuhvaćena su sva dosadašnja saznanja vezana za teorijsku intuicijsku osnovu Benfordovog svojstva, oblasti upotrebe i primenu fenomena vodeće cifre. Postupak smo sproveli kroz proces Digitalne Benfordove analize koja uključuje i statističke postupke deskriptivne analize i upotrebu različitih statističkih testova. Celokupan proces DBA primenjen je u sklopu početnih koraka revizijskog procesa finansijskih izveštaja preduzeća koja posluju na teritoriji Republike Srbije, konkretno kompanija "Imlek" i "Apatinska pivara". Dodatno, naznačene su prednosti upotrebe FSD-a pri izražavanju pouzdanosti podataka iz finansijskih izveštaja, merenjem odstupanja stvarnih frekvencija pojave cifara ustanovljene na uzorku i očekivanih frekvencija ustanovljenih Benforvodim zakonom.

Rad postaje jednostavno sredstvo za vođenje početnih postupaka analiza istinitosti različitih vrsta podataka upotrebom Benfordovog zakona. Analiza primenom zakonistosti može poslužiti kao indikator za strožija ili fleksibilnija dalja razmatranja, ali nikako kao krajnji dokaz odsustva/postojanja nepravilnosti bez sprovođenja daljih provera. Opreznost se nalaže zbog relativne efikasnosti Digitalne Benfordove analize, zbog ograničenja na vidove skupova na kojima se ova analiza može primenjivati, ali i zbog osobina pojedinih grešaka koje mogu da prođu nezapaženo.

Praktična primena DBA kompanija "Imlek" i "Apatinska pivara" ukazala je na ograničenja pri upotrebi različitih statističkih testova, kao i na uticaj obima grešaka u otkrivanju nepravilnosti i ocenjivanju efikasnosti DBA. Logično, veći obim grešaka doprineo je većoj efikasnost procesa DBA u detekciji nepravilnosti.
Cilj rada je bio pomoć kako ljudima koji sprovode sve vidove kontrola, tako i upućivanje lica da pri stupanju u bilo kakve poslovne aktivnosti sa poslovnim subjektima mogu tražiti i ovakve vidove analiza, kao potvrde ispravnosti podataka. Jednostavnu upotrebu Benfordovog zakona, na prvi pogled neobičanog matematičkog fenomena, primenjivog u raznim oblastima života poželjno je savladati, jer Non scholae, sed vitae discimus[1].

[1] Ne za školu, za život učimo

Literatura

[1] Frank Benford, *The Law of Anomalous numbers*, Schenectady, Njujork, 1994.

[2] Teodor P. Hill, *The First Digit Phenomenon*, American Scientist, 1998.

[3] Steven J. Miller, Arno Berger, Teodor P. Hill, *Theory and Applications of Benford's Law*, Princeton University , 2010, pp. 5-19

[4] Jesus Gonzales-Garcia, Gonzalo Pastor, *Benford's Law and Macroeconomics Date Quality*, International Monetary Fund, 2009.

[5] Dan Amiram, Zahn Bozanic, Ethan Rouen, *Detecting Financial Statement Irregularities: Evidence from the Distributional Properties of Financial Statement Numbers* , Columbia University, 2015.

[6] Dan Amiram, Zahn Bozanic, Ethan Rouen, *Detecting Financial Statement Irregularities: Evidence from the Distributional Properties of Financial Statement Numbers* , Columbia University, 2013.

[7] Cindy Durtschi, William Hillison, Carl Pacini , *The Effective Use of Benford's Law to Assist in Detecting Fraud in Accounting Data* , Edwards, Inc, 2004.

[8] Daniel I. A. Choen, *An Explanation of the First Digit Phenomenon* , Journal of Combinational Theory, 1975.

[9] Adrian Jamain , *Benford's Law* , Imperial Collage of London, Ensimag, 2001.

[10] Marko Milojević, Ivica Terzić, Vojisalav Marjanović , *Primena Benfordovog zakona u otkrivanju anomalija u finansijkim izveštajima - slučaj velikih preduzeća u Srbiji* , Univerzitet Singidunum, Beoggrad, 2014.

[11] Jon Walthoe , *Looking out for number one* , Brington, 2011.

[12] Alex Ely Kossovsky , *Benford's Law: Theory, the General Law of Relative Quantities, and Forensic Detection Applications* , World Scientific Publishing, 2014.

[13] Miljenko Huzak , *Benfordov zakon* , Sveučilište Zagreb, 2001.

[14] Mark Nigrini , *Forensic Analytics: Methods and Techniques for Forensic Accounting Investigations* , Hoboken, 2011.

[15] Omer Jukić, Nedžad Muhurdarević , *Benfordov zakon* , BiH, 2003.

[16] Miljenko Huzak , *Benfordov zakon* , Sveučilište Zagreb, 2001

[17] Stefan Engel , *Fact and Fiction in EU-Governmental Economic Date* , Catholic University of Eichstätt-Ingolstadt, German Economic Review, 2011

[18] P.Dechow, W. Ge, C. Larson, R. Sloan, *Predicting Material Accounting Misstatements*, Contemporary Accounting Research, 2011

[19] Tianran Chen , *Analysis on Accrual-Based Models in Detecting Earnings Management* , Lingan Journal of Banking, 2010

[20] Jae Jun , *How to Detect Earnings Manipulation Using the Beneish M Score* , 2011

[21] Zagorka Lozanov-Crvenković , *Statistika* ,Novi Sad, 2012

[22] Daniel P. Pike, *Testing for the Benford Property*, Rochester Institute of Tecnology , pp. 15-17

[23] A. Diekmann,J. Ben, *Benford's Law and Fraud Detection: Facts and Legends*, German Economic Review, 2010

[24] K. Todter, *Benford's Law as an Indicator of Fraud in Economics*, German Economic Review, 2009.

[25] M. Andrić, B. Krsmanović, D. Jakšić, *Revizija - teorija i praksa*, Ekonomski fakultet Subotica, 2009.

[26] Results of the Iranian presidental election 2009, 2011, dostupno na `http://en.wikipedia.org`

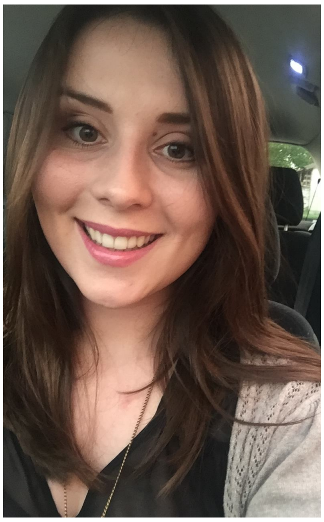
[27] Agencija za privredne registre `http://www.apr.gov.rs`

Spisak tabela i slika

Spisak tabela

Spisak slika

Biografija



Jovana Arsić rođena je 12. aprila 1993. godine u Ilijašu, Bosna i Hercegovina. Od 1996. nastanjena na prostoru opštine Brčko. Gimnaziju „Vaso Pelagić“ završila 2011. godine u Brčkom. Iste godine postaje student Prirodno-matematičkog fakulteta u Novom Sadu, smer Primenjena matematika, modul Matematika finansija.

Osnovne akademske studije završila je 2014. godine i obrazovanje nastavlja upisom na Master studije Primenjene matematike na Prirodno-matematičkom fakultetu.

Nosilac je diploma učenika generacije, volonter Centra za socijalni rad i učesnik festivala nauke. U periodu od juna do septembra 2016. godine kao polaznik Letnje prakse u Centru za superviziju informacionih sistema Narodne banke Srbije bila je zadužena za izradu Modela procene rizičnosti informacionih sistema finansijskih institucija.

UNIVERZITET U NOVOM SADU PRIRODNO - MATEMATIČKI FAKULTET KLJUČNA DOKUMENTACIJSKA INFORMACIJA

Redni broj:
RBR

Identifikacioni broj:
IBR

Tip dokumentacije: *monografska dokumentacija*
TD

Tip zapisa: *tekstualni Stampani materijal*
TZ

Vrsta rada: *master rad*
VR

Autor: *Jovana Arsić*
AU

Mentor: *vanredni profesor dr Nataša Spahić*
MN

Naslov rada: *Primena Benfordovog zakona u reviziji*
MR

Jezik publikacije: *srpski (latinica)*
JP

Jezik izvoda: *s/e*
JI

Zemlja publikovanja: *Republika Srbija*
ZP

Uže geografsko područje: *Vojvodina*
UGP

Godina: *2016*
GO

Izdavač: *autorski reprint*
IZ

Mesto i adresa: *Novi Sad, Trg Dositeja Obradovića 4*
MA

Fizički opis rada: *6 poglavlja, 60 strana, 27 lit. citata, 14 tabela, 9 slika*
OFO

Naučna oblast: *matematika*
NO

Naučna disciplina: *primenjena matematika*
ND

Ključne reči: *revizija, Benfordov zakon, Digitlana Benfordova analiza, statistički testovi, FSD*
PO

UDK

Čuva se: *u biblioteci Departmana za matematiku i informatiku, Prirodno-matematičkog fakulteta, u Novom Sadu*
ČU

Važna napomena:
VN

Izvod: *Cilj master rada je da predstavi primenu Benfordovog zakona u reviziji. Benforov zakon se zasniva na razmatranju da kod određenih skupova podataka veći broj podataka počinje manjim ciframa. Odstupanja učestalosti pojavljivanja cifara na vodećoj poziciji testiranog skupa podataka u odnosu na učestalosti pojavljivanja cifara predviđenih Benfordovim zakonom ukazuju na moguće prevare. Svrha rada je da posluži kao vodič za primenu Benfordovog zakona prilikom proveravanja istinitosti finansijkih izveštaja upotrebom statitstičkih testova i FSD-a.*
IZ

Datum prihvatanja teme od strane NN veća: *04. maj 2016.*
DP

Datum odbrane:
DO

Članovi komisije:
KO
Predsednik: *dr Nataša Krejić, redovni profesor Prirodno-matematičkog fakulteta u Novom Sadu*
Mentor: *dr Nataša Spahić, vanredni profesor Prirodno-matematičkog fakulteta u Novom Sadu*
Član: *dr Sanja Rapajić, vanredni profesor Prirodno-matematičkog fakulteta u Novom Sadu*

UNIVERSITY OF NOVI SAD FACULTY OF SCIENCE KEY WORDS DOCUMENTATION

Accession number:
ANO

Identification umber:
INO

Document type: *monograph type*
DT

Type of record: *printed text*
TR

Contents Code: *master thesis*
CC

Author: *Jovana Arsić*
AU

Mentor: *associate professor dr Nataša Spahić*
MN

Title: *Application of the Benford's law in audit*
XI

Language of text: *serbian (latin)*
LT

Language of abstract: *s/e*
LA

Country of publication: *Republic of Serbia*
CP

Locality of publication: *Vojvodina*
LP

Publication year: *2016*
PY

Publisher: *author's reprint*
PU

Publ. place: *Novi Sad, Trg Dositeja Obradovića 4*
PP

Physical description: *6 chapters, 60 pages, 27 references, 14 tables, 9 pictures*
PD

Scientific field: *mathematics*
SF

Scientific discipline: *applied mathematics*
SD

Key words: *audit, Benford's law, Digital Benford's analysis, statistical tests, FSD*
UC

Holding data: *Department of Mathematics and Informatics Library, Faculty of Science, Novi Sad*
HD

Note:
N

Abstract: *The purpose of the paper is to present the application of the Benford's law in audit. Benford's law is based on observation that in certain date sets more numbers have smaller digits in the leading position. The deviation of empirical date sets digit frequency from theoretical Benford's frequency could be signal of the fraud. This paper attempts to provide guidance for application Benford's law when testing accuracy of the financial statements using statistical tests and FSD score.*
AB

Accepted by the Scientific Board on: *04 May 2016*
ASB

Defended:
DE

Thesis defend board:
DB
President: *dr Nataša Krejić, full professor*
Mentor: *dr Nataša Spahić, associate professor*
Member: *dr Sanja Rapajić, associate professor*