Danica Stojiljković

FM šum kao generator slučajnih brojeva

Upoređeni su random generatori implementirani u različitim programskim jezicima sa generatorima brojeva dobijenih iz belog šuma. Testirani su i ocenjeni algoritmi pseudo-random generatora nekih programskih jezika, kao i generator koji slučajne brojeve generiše na osnovu FM šuma. Definisan je kriterijum za ocenjivanje random generatora na osnovu Pirsonovog χ^2 testa. Pretpostavili smo da će uzorkovani brojevi nakon odgovarajuće transformacije davati uniformnu raspodelu. Međutim, u svim slučajevima stvarna raspodela u znatno većoj meri odstupa od očekivane, nego što je to dozvoljeno kada su u pitanju zaista slučajni brojevi. Stoga, na osnovu urađenih testova možemo odbaciti pretpostavku da brojevi dobijeni iz FM šuma daju normalnu raspodelu, tj. pretpostavku da su ti brojevi slučajni.

Uvod

U prirodi postoji mnogo stohastičnih procesa, ali je velik broj onih procesa koji su haotični samo na prvi pogled. Uzorci koji su dobijeni iz prirodnih stohastičnih pojava najčešće imaju Gausovu ili uniformu raspodelu. Normalna (Gausova) raspodela opisuje se Gausovom krivom $g(x) = (1/\sqrt{2\pi})\, e^{-\frac{x^2}{2}}$, dok uniformna raspodela predstavlja raspodelu sa jednakom verovatnoćom pojavljivanja svakog broja.

Možemo iskoristiti prirodne procese za generisanje slučajnih brojeva, ali prvo bi trebalo da odredimo koji su procesi zaista haotični.

Pirsonov χ^2 test

Pirsonov χ^2 test je verovatno najviše korišćeni od svih statističkih testova. Primenjuje se kada želimo da ispitamo u kojoj meri se određena empirijska raspodela slaže sa teorijski očekivanom.

Ovaj test poredi empirijski dobijenu raspodelu događaja sa pretpostavljenom. Ako imamo niz od n slučajnih promenljivih X_i koje se

Danica Stojiljković (1981), Kraljevo, Žarka Zrenjanina 15, učenica 2. razreda Matematičke gimnazije u Beogradu

nalaze u unapred zadatim intervalima A_s, onda njihove frekvencije pojavljivanja u tim intervalima obeležavamo sa f_s. Brojevi f_s su rezultat posmatranja i prikazuju stvarno stanje, dok je np_s matematičko očekivanje. p_s je verovatnoća da se X_i nalazi u odgovarajućem intervalu na osnovu pretpostavljene raspodele. χ^2 test se koristi da bi potvrdili ili odbacili pretpostavljenu raspodelu. Vrednost χ^2_o se definiše kao

$$\chi^2_o = \sum_1^k \frac{(f_s - np_s)^2}{np_s}$$

Po teoremi ova vrednost teži $\chi^2(\nu)$ kada $n \to \infty$ (Merkle, Vasić 1995), gde ν označava broj stepeni slobode, tj. najmanji broj parametara kojim se sistem može potpuno opisati i on je za jedan manji od broja intervala, jer na osnovu poznatih vrednosti f_s za $k-1$ intervala i ukupnog broja događaja možemo vrednost u k-tom intervalu (Kostić 1990).

Pirsonov χ^2 kriterijum za raspodelu slučajnih brojeva kaže da generator možemo smatrati dobrim na nivou poverenja od 95% ako χ^2 leži između kvantila raspodele $\chi^2_{0.975}$ i $\chi^2_{0.025}$ (Knuth 1973). Drugim rečima, χ^2 raspodela nam govori koliku bi vrednost trebalo da ima suma iz prethodne formule, ako su brojevi zaista slučajni. To je zvonasta raspodela i kriterijum kvaliteta generatora mora biti vezan sa rastojanjem χ^2_o od maksimuma raspodele χ^2. Kvantili u ovom slučaju označavaju vrednosti χ^2 za koje je $\int_0^{\chi^2} \chi^2(t)\,dt = 0.025$, odnosno 0.975.

Velike vrednosti statistike χ^2 ukazuju na veliku vrednost razlike između stvarnog i očekivanog, pa zbog toga odbacujemo pretpostavljenu raspodelu. Opet, isuviše male vrednosti χ^2 ukazuju da se očekivane i stvarne vrednosti poklapaju više nego što ova statistika dozvoljava, pa i nju moramo odbaciti.

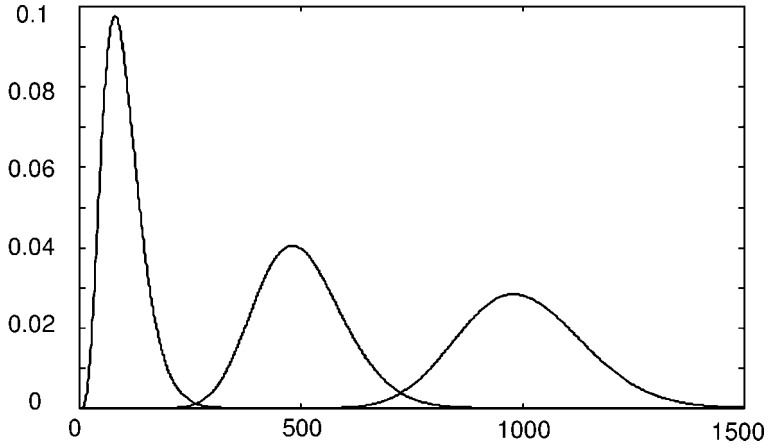
Na osnovu vrednosti p za određeno χ^2_o, podelićemo događaje u kategorije (Knuth 1973):

I – neslučajni ($p < 0.01$ ili $p > 0.99$)

II – sumnjivo slučajni ($0.01 < p < 0.05$ ili $0.95 < p < 0.99$)

III – skoro sumnjivo slučajni ($0.05 < p < .0.10$ ili $0.90 < p < 0.95$)

IV – zadovoljavajuće slučajni ($0.10 < p < 0.90$)

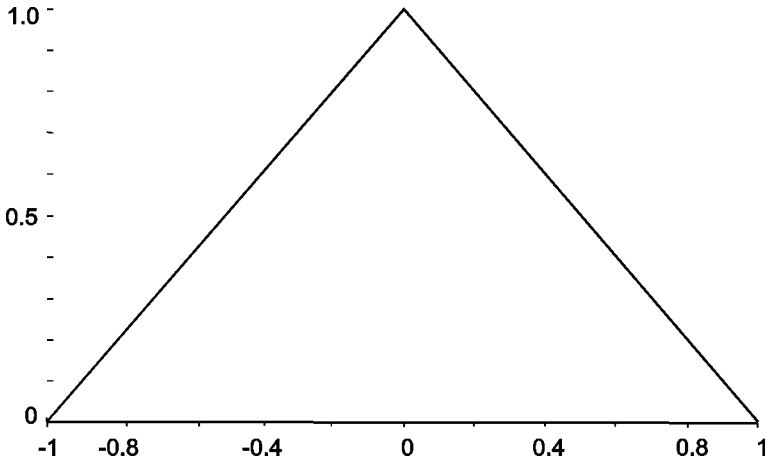


Slika 1.
Krive χ^2 raspodele za različit broj stepeni slobode.

Figure 1.
Curves of χ^2 distribution for different number of degrees of freedom.

Zavisnosti oblika χ^2_o distribucije od broja stepeni slobode govori nam da ona biva sve simetričnija sa porastom brojem stepeni slobode (slika 1). Ako je broj stepeni slobode veći od 5, distribucija je dovoljno simetrična da možemo definisati kriterijum slučajnosti proizvoljne raspodele tako što proveravamo da li se određena vrednost χ^2 nalazi između odgovarajućih kvantila χ^2_o raspodele.

Uz pomoć ovog testa, uspostavljamo pravilo na osnovu koga "ocenjujemo" kvalitet generatora slučajnih brojeva sa uniformnom raspodelom. Najpre definišemo dva pomoćna kriterijuma "saglasnosti". Prvi je kriterijum saglasnosti raspodele samih brojeva sa očekivanom uniformnom raspodelom, dok je drugi kriterijum saglasnosti raspodele razlika uzastopnih brojeva sa očekivanom raspodelom.



Slika 2.
Pretpostavljena raspodela razlika uzastopnih brojeva u slučqajnom nizu.

Figure 2.
The supposed consecutive number difference distribution in a random sequence.

Može se pokazati da bi raspodela razlika prikazana na slici 2 trebalo da zadovoljava jednačinu $f_s = -\frac{n}{k}|X_s| + \frac{n}{k}$. Ako znamo χ^2_o i maksimalne vrednosti χ^2 raspodele za prvi i drugi kriterijum možemo definisati veličine m_1 i m_2, koje predstavljaju njihove odnose i predstavljaju parametar odstupanja raspodela od očekivanih.

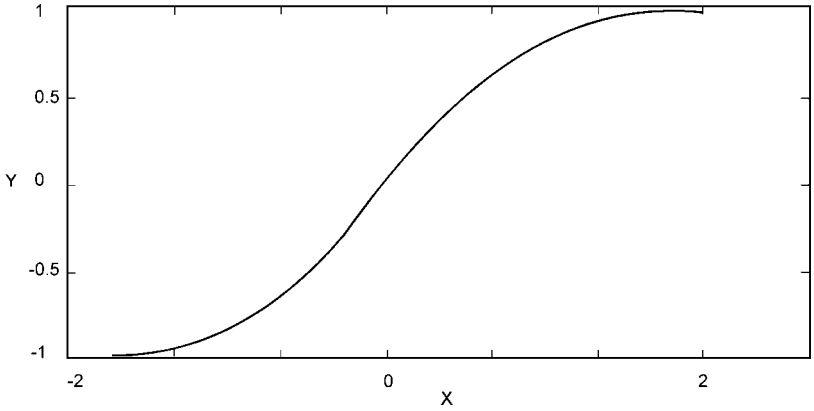
Meru kvaliteta random generatora možemo uvesti definisanjem odgovarajućeg parametra C. S obzirom da su jedine dve veličine koje zaista karakterišu odstupanje dobijenih raspodela od očekivanih – m_1 i m_2, prirodno se nameće izraz za ovaj parametar

$$C = \sqrt{\frac{m_1^2 + m_2^2}{2}}\ .$$

C je definisano na intervalu $[0, 1]$ i trebalo bi da daje objektivnu ocenu slučajnosti niza brojeva sa uniformnom raspodelom U $[0, 1)$. Naravno, slučajnost niza po ova dva kriterijuma ne znači da ne postoji transformacija koja bi pokazala kako je niz prediktabilan ili "ne baš naročito slučajan", ali smo pošli od pretpostavke da su pomenuta dva kriterijuma dovoljna za većinu primena.

Priroda FM šuma

Prilikom slanja signala frekvencije W pomoću ampilitudno modulisanog talasa, pored nosećeg talasa frekvencije f, nastaće i dva bočna talasa frekvencija od $f - \Omega$ i $f + \Omega$. Međutim, polazni signal najčešće ne sadrži samo jednu već čitav spektar frekvencija pa će samim tim pored nosećeg talasa nastati čitav spektar bočnih talasa. Frekvencijski modulisani talas se takođe može predstaviti kao suma sinusnih funkcija i već prilikom prenosa talasa jedne frekvencije javlja se beskonačni spektar bočnih talasa. Za verodostojni prenos signala dovoljan je samo jedan deo ovog spektra ali je čak i taj prekopotrebni deo znatno širi od spektra bočnih talasa koji se javljaju kod amplitudske modulacije. Beli šum frekvencijske modulacije pogodan je za uzorkovanje random brojeva upravo zbog velike širine spektra bočnih talasa. (Opća enciklopedija 1980). Postoje dva osnovna izvora slučajnih grešaka – fluktuacijski i impulsni šum (Kviz 1980). Impulsni šum nastaje usled atmosferskih pražnjena i čovekove delatnosti. Poreklo fluktuacijskog šuma vodi od fizičkih pojava i električnih naboja u strukturi materije. Brojevi dobijeni ovim generatorom imaće raspodelu sličnu normalnoj raspodeli, ali malo "nakrivljenu" na desnu stranu, verovatno, usled nelinearnosti izlaza zvučne kartice, odnosno pojave zasićenja za veće vrednosti napona. Međutim, ako za slučajne vrednosti uzmemo razlike uzastop-



Slika 3.
Zavisnost dobijenih brojeva (Y) koji zadovoljavaju uniformnu raspodelu od eksperimentalnih uzoraka (X).

Figure 3.
The graph of the relation between obtained numbers from uniform distribution (Y) and experimentaly generated samples (X).

nih uzoraka dobićemo raspodelu koja znatno više liči na Gausovu. Nas, međutim, interesuje kakav se generator slučajnih brojeva sa uniformnom raspodelom može dobiti iz ovih uzoraka, pa je potrebno pronaći način da se svakom uzorku X_i iz niza sa normalnom raspodelom pridruži odgovarajući broj Y_i, tako da novodobijeni niz ima uniformnu raspodelu na intervalu $[0, 1)$. Ovo preslikavanje je moguće izvršiti transformacijom

$$Y_i = \int_{-\infty}^{X_i} g(x)\,\mathrm{d}x,$$

gde je $g(x)$ gustina verovatnoće pojavljivanja svakog broja kod normalne raspodele. Grafik zavisnosti dobijenih brojeva Y_i od uzoraka X_i dat je na slici 3. Međutim, kako se integral funkcija $g(x)$ ne može izraziti analitičkim putem koristimo program koji rezultat daje numerički (funkcija Erf u Matlab-u).

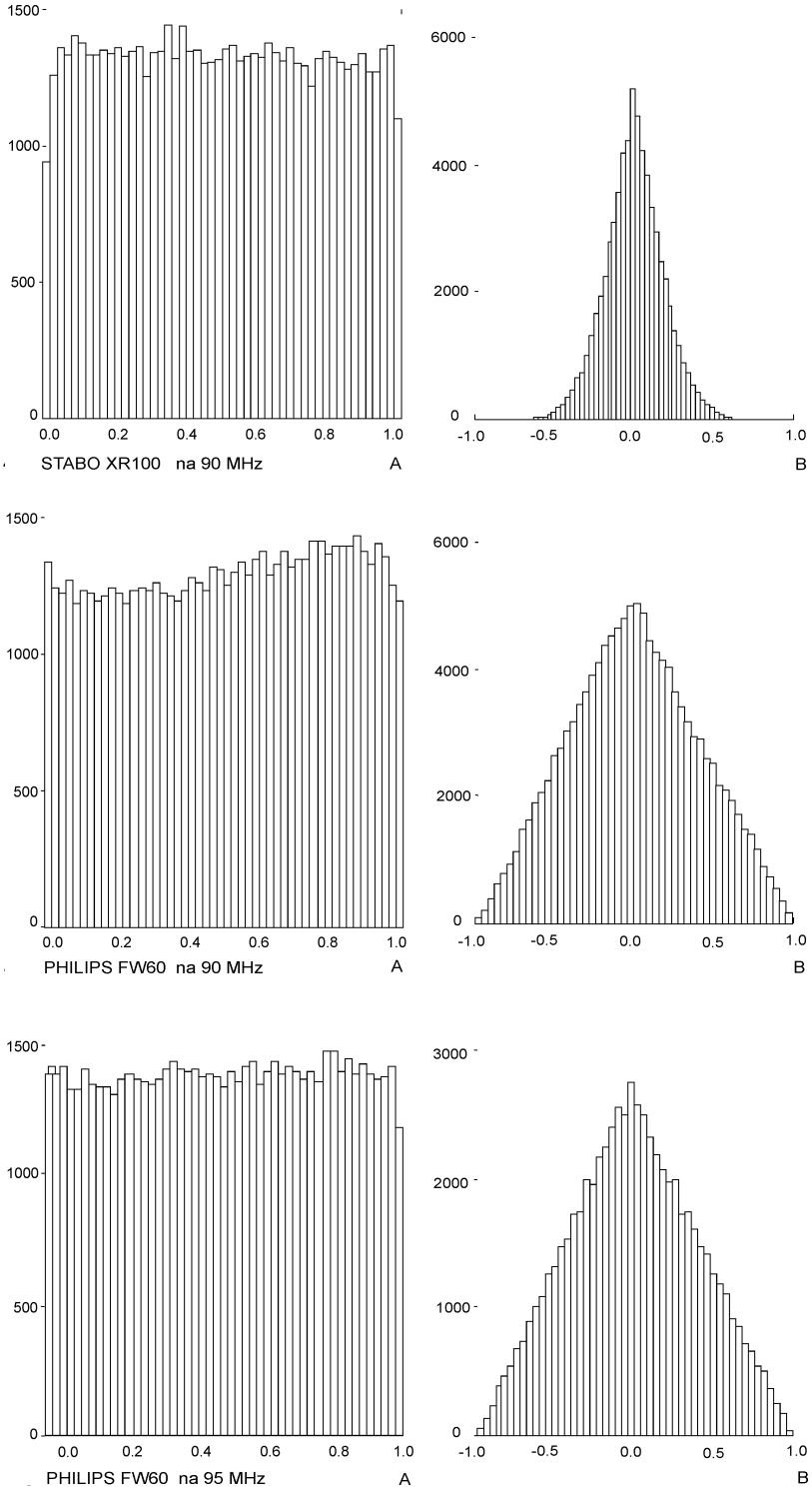
Rezultati generisanja slučajnih brojeva

Pseudo-random generatori

Na opisani način smo testirali i ocenili random generatore koji se isporučuju u standardnim bibliotekama programskih paketa datih u tabeli 1:

Tabela 1. Odstupanja empirijskih raspodela od očekivanih kod razmatanih programskih jezika

Parametar	Generator		
	Turbo Pascal verzija 7.0	Gcc 2.7.2 bibilioteka libc.5.4.13.so	Matlab verzija 5.2.0.3084.



Slika 4.
Raspodele slučajnih brojeva generisanih iz FM šuma (a) i razlika uzastopnih brojeva (b) kod tri analizirana slučaja: 1) STABO XR100 na 90 MHz, 2) PHILIPS FW60 na 90 MHz i 3) PHILIPS FW60 na 95 MHz.

Figure4.
Distributions of the generated random numbers from the FM noise (a) and consecutive number difference distribution (b) at three analyzed cases (1) STABO XR100 at 90 MHz; (2) PHILIPS FW60 at 90 MHz; (3) PHILIPS FW60 at 95 MHz.

raspodele brojeva	χ^2	215.90	44.64	45.39
	m_1	1.6E-21	0.9387	0.9641
raspodele razlika brojeva	χ^2	49.04	54.40	45.17
	m_2	0.9949	0.8219	0.9584
ocena	C	0.70	0.88	0.96

Generisanje slučajnih brojeva na osnovu FM šuma

Testiraćemo random generator koji za slučajne brojeve uzima vrednosti nivoa šuma na određenoj frekvenciji koja se na skali radio prijemnika nalazi između dve FM stanice.

Merenja su vršena na različitim frekvencijama i uz pomoć dva različita uređaja. Na osnovu dobijenih rezultata (slika 4), zaključuje se da rezultati u velikoj meri zavise od uređaja koji su detektovali šum. Na slikama 4.1 i 4.2 su prikazani rezultati dobijeni na istoj frekvenciji (90 MHz) uz pomoć dva različita uređaja.

Merenjem radio-prijemnikom STABO XR100 na drugim frekvencijama (150, 200, 300, 500 MHz i 1 GHz) dobijene su raspodele razlika slične odgovarajućoj raspodeli na frekvenciji od 90 MHz (slika 4.1). Ove raspodele dosta odstupaju od očekivane raspodele. Takođe uočavamo da je broj pojavljivanja brojeva u intervalima (0, 0.05) i (0.95, 1) znatno manji od broja pojavljivanja ostalih brojeva, iz čega se zaključuje da ovaj niz brojeva nije slučajan. Srednja ocena ovih generatora je $1.8 \cdot 10^{-3}$, a najviša ocena je dobijena na frekvenciji od 150 MHz i iznosi $5.6 \cdot 10^{-3}$.

Pri merenju na frekvenciji od 90 MHz radio prijemnikom PHILIPS FW60 primećujemo specifičan oblik raspodele frekvencija šuma kojeg nema kod drugih frekvencija (sl. 4.2), možda zbog blizine neke radio-stanice. Oblici raspodela frekvencija šuma i razlika susednih brojeva merenih na frekvencijama od 95, 100, 105 i 110 MHz su u zadovoljavajućoj korelaciji sa očekivanim raspodelama (4.3). Srednja ocena ovih generatora je 0.017, a najbolja ocena je dobijena na frekvenciji od 95 MHz i iznosi 0.058.

Zaključak

Pretpostavili smo da će uzorkovani brojevi nakon transformisanja davati uniformnu raspodelu. Međutim, u svim slučajevima stvarna raspodela u znatno većoj meri odstupa od očekivane, nego što je to dozvoljeno kada su u pitanju zaista slučajni brojevi. Stoga, na osnovu urađenih testova možemo odbaciti pretpostavku da brojevi dobijeni iz FM šuma daju normalnu raspodelu, tj. pretpostavku da su ti brojevi slučajni.

Literatura

Kviz, B. 1980. *Radiotelemetrija*. Zagreb: Školska knjiga.

Demidovich, M. 1987. *Computational mathematics*. Moscow: Mir

Knuth, D.E. 1994. *The Art of Computer Programming*, vol.2. Addison-Wesley.

Kostić, M.C. 1990. *Metodi statističke analize sa kompjuterskim prilazom*. Beograd: Naučna knjiga.

Danica Stojiljković

The FM Noise as a Random Numbers Generator

There are many stochastic processes in nature, but a large number of them are chaotic only at first glance. The FM noise is often cited as a paradigm of white Gaussian noise. In this paper characteristics of numeri - cal pseudo-random generators and generator obtained by FM noise are analyzed. Generator quality is estimated on the bases of relative disagree - ment between assumed uniform and obtained distribution. We defined an appropriate parameter C in order to introduce the rate of random gener - ator's quality.

We tested and evaluated random generators that are delivered in stan - dard libraries of some program languages in a described way and the re - sults are shown in Table 1.

The frequency-modulated wave could be represented as a sum of si - nusoidal functions and even in case of a signal that contains only one fre - quency, the boundless spectrum of side waves appears. For the credible transmitting of signal only one part of this spectrum is enough and even that necessary part is much wider than the spectrum of side waves which appears at amplitude modulation (Izyumov 1980). White noise of ampli - tude modulation is suitable for sampling random numbers exactly because of the large side wave's spectrum's width. Uniform distribution for the noise is obtained by Ulam's Golden Rule.

Samplings were done at different frequencies using two different ra - dio receivers. According to obtained results (Figure 4) we can draw a con - clusion that the results mainly depend on radio receiver that detects noise. The figures 4.1 and 4.2 shows the results obtained by two different radio receivers at the same frequency of 90MHz.

Sampling by radio receiver STABO XR 100 at other frequencies (150, 200, 300, 500 MHz and 1 GHz) gave the distributions which are similar to the appropriate distribution of numbers sampled at frequency of 90MHz (Fig. 4.1). These distributions deviate from the expected ones in a

high rate. Average evaluation of these generators is $1.8 \cdot 10^{-3}$, and the higher one is obtained at the frequency of 150 MHz and amounts $5.6 \cdot 10^{-3}$.

While sampling at a frequency of 90 MHz by radio receiver PHILIPS FW60, a specific form of noise intensity distribution is noticed that is not noticed at other frequencies (Fig. 4.2). The cause of this phenomenon could be the closeness to some radio station frequency. The forms of noise intensity distribution and distribution of differences obtained at frequencies of 95, 100, 105 and 110 MHz are in sufficient correlation with the expected distributions. Average evaluation is 0.017, and the higher one is obtained at the frequency of 95 MHz and equals 0.058 (Fig. 4.3).

We supposed that after the transformation, the sampled numbers would have uniform distribution. But in all cases the real distribution deviate from the expected one in a much higher rate than this particular statistic allows. Therefore, according to the tests we can reject the assumption that the numbers delivered from the FM noise have Gaussian distribution, i.e. we can reject the hypotesis that those numbers are random.