LABORATORIJSKA VEŽBA IZ VEROVATNOĆE I STATISTIKE

Anja Vujačić 0307/21
Projekat iz predmeta Verovatnoća i statistika (13S082VS)
Elektrotehnički fakultet, Univerzitet u Beogradu

Abstract

Glavni cilj ove laboratorijske vežbe je realizacija simulacija n Bernulijevih upita sa datom verovatnoćom uspeha p, pri čemu će se simulacija sa istim zadatim parametrima ponoviti N puta. Analiza rezultata data je na samom kraju.

UVOD I POSTAVKA PROBLEMA

Postavka problema data je primerom košarkaša koji ubacuje loptu u koš 10 ($n = 10$) puta sa verovatnoćom uspeha ($p = 0.7$). Program simulira 10 ovakvih gadjanja i ponavlja simulaciju 100 puta. Na izlazu programa za svako $k \in 0,1,...,10$ odredjena je relativna frekvencija k pogodaka. Takodje su izvresene dodatne analize i poredjenja rezultata sa očekivanom verovatnoćom k pogodaka.

BERNULIJEVI UPITI

Bernoullijev upit predstavlja statistički eksperiment ili slučajni proces koji ima samo dva moguća rezultata: uspjeh ili neuspeh. Naziv je potekao od Švajcarskog matematičara Jakoba Bernoullija, koji je prvi formalizirao ovu vrstu slučajnih eksperimenata.

Verovatnoća uspeha se obično označava sa "p", dok se verovatnoća neuspeha obično označava sa "q" (gdje je q = 1 - p). Važno je napomenuti da svaki pojedinačni eksperiment u Bernoullijevom upitu ne zavisi o prethodnim eksperimentima i zato se naziva nezavisnim Bernoullijevim eksperimentom.

Ovi upiti imaju važnu ulogu u statistici i verovatnoći jer služe kao osnova za razne statističke raspodele. Dobar primer je binomna raspodela koja prikazuje broj uspeha u nizu nezavisnih Bernoullijevih eksperimenata.

BINOMNA RASPODELA

Binomna raspodela je statistička raspodela koja se koristi za prikaz broja uspeha u seriji nezavisnih Bernulijevih eksperimenata. Kao što je prethodno rečeno, u Bernulijevim eksperimentima postoje samo dva moguća rezultata: uspeh ili neuspeh. Binomna raspodela se primenjuje za analizu broja uspeh koji će se dogoditi u određenom broju ponavljanja eksperimenta.

Binomna raspodela ima dva parametra: n i p.

- n predstavlja ukupan broj eksperimenata ili pokušaja
- p predstavlja verovatnoću uspeha u svakom pojedinačnom eksperimentu

Verovatnoća da se tačno k uspeha dogodi u n nezavisnih Bernulijevih eksperimenata može se izračunati koristeći binomnu formulu:

ALGORITAM

Funkcija BernoulisTrials

Algorithm 1 Funkcija BernoulisTrials(n, p, N)

$f = arrayofzerosofsizen + 1$
for $s \in (1, N)$ **do**
  $countSuccessfulTrials = 0$
  for $trial \in (1, n)$: **do**
    if $rndNumber\ between\ 0\ and\ 1 < p$: **then**
      $countSuccessfulTrials += 1$
    end if
  end for
  $f[countSuccessfulTrials] += 1$
end for
$return\ f/N$

Navedene instrukcije čine funkciju BernullisTrials koja posmatra n Bernulijevih upita.

- n puta generiše nasumičan broj iz segmenta 0,1.
- Vrednost ovog broja uporedjuje se brojem p.
- Ukoliko je broj manji od p, upit smatramo uspešnim.

Povratnu vrednost funkcije čini vektor dužine n+1 koji sadrži relativne frekvencije broja uspešnih upita prilikom simulacije (od 0 do n).

Algorithm 2 Funkcija BinomialDistribution(n, p):

$f = array\ of\ size\ (n+1)\ of\ type\ float$
for $x\ in\ f\ converted\ to\ integer$ **do**
  $f[x] = \binom{n}{k} \cdot (p^x) \cdot (1-p)^{(n-x)}$
end for
$return\ f$

Navedene instrukcije čine funkciju BinomialDistribution koja izračunava tačne verovatnoće mogućih vrednosti broja uspešnih upita po formuli:

$$P(X = k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{(n-k)} \tag{1}$$

REZULTATI I DISKUSIJA

Na slikama ispod prikazani su histogrami relativne funkcije dobijene simulacijom Bernulijevih upita (svetlija nijansa) kao i tačne verovatnoće (tamnija nijansa). Kod simulacije dat jeu prilogu, na progrmaskom jeziku Python. Simulacija je radjena u Google Colabu. Radi bolje analize eksperimenta, simulacija je pokrenuta nad 100 (slika levo) i nad 1000 iteracija (slika desno).

Table 1: N=100

Broj uspesnih upita	Relativna frekvencija	Tacna verovatnoca	Apsolutna greska
0	0.00000	0.00001	0.00001
1	0.00000	0.00014	0.00014
2	0.00000	0.00145	0.00145
3	0.00000	0.00900	0.00900
4	0.01000	0.03676	0.02676
5	0.15000	0.10292	0.04708
6	0.24000	0.20012	0.03988
7	0.28000	0.26683	0.01317
8	0.22000	0.23347	0.01347
9	0.10000	0.12106	0.02106
10	0.00000	0.02825	0.02825

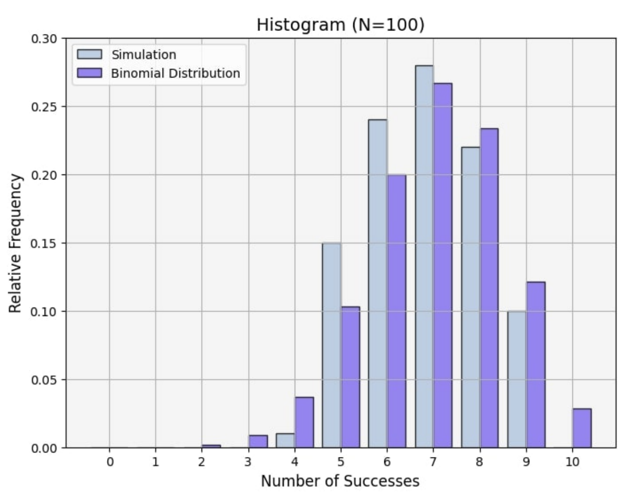

Figure 1

Sudeći po datim graficima i tabelama, jasno je da se sa povećanjem broja iteracija smanjuje greška, što je i očekivano za binomnu raspodelu. Kako bismo utvrdili prethodno tvrdjenje, program je pokrenut još 2 puta. Primećujemo da su rezultati oba pokretanja različiti ali je ponašanje isto kao i u prethodnom primeru, odnosno sa povećanjem broja iteracija smanjuje se greška.

Table 2: N=100

Broj uspesnih upita	Relativna frekvencija	Tacna verovatnoca	Apsolutna greska
0	0.00000	0.00001	0.00001
1	0.00000	0.00014	0.00014
2	0.00100	0.00145	0.00045
3	0.01600	0.00900	0.00700
4	0.04200	0.03676	0.00524
5	0.08100	0.10292	0.02192
6	0.23000	0.20012	0.02988
7	0.25600	0.26683	0.00153
8	0.23500	0.23347	0.1083
9	0.11200	0.12106	0.00906
10	0.02700	0.02825	0.00125

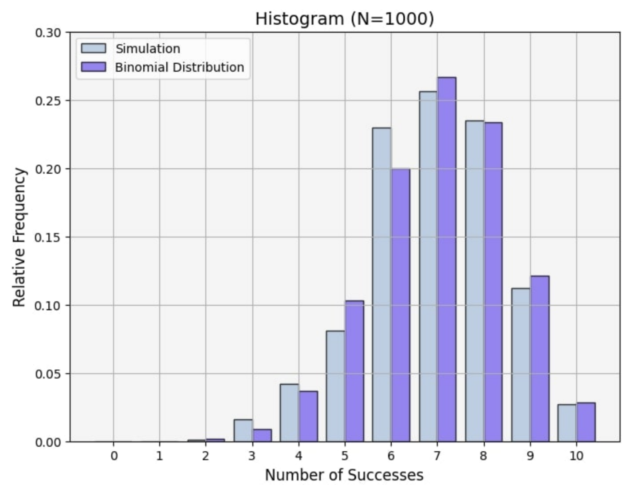

Figure 2

REFERENCES

[1] Merkle, M., 2016. *Verovatnoća i statistika za inženjere i studente tehnike.* Akademska Misao 2016.

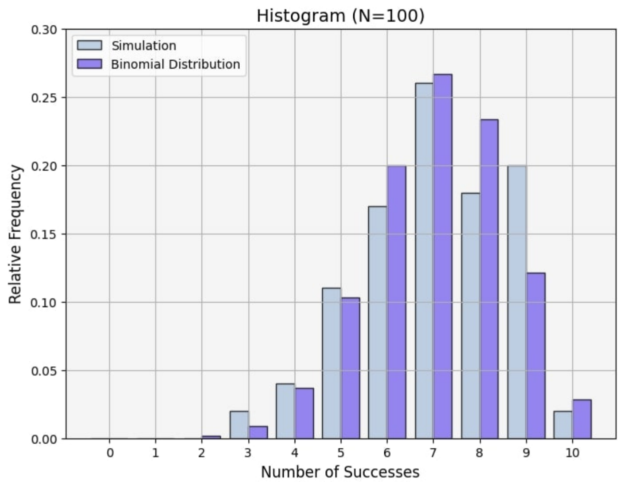

Figure 3

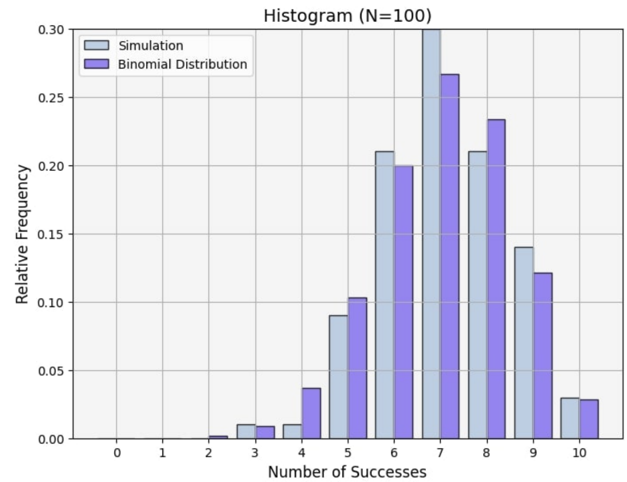

Figure 5

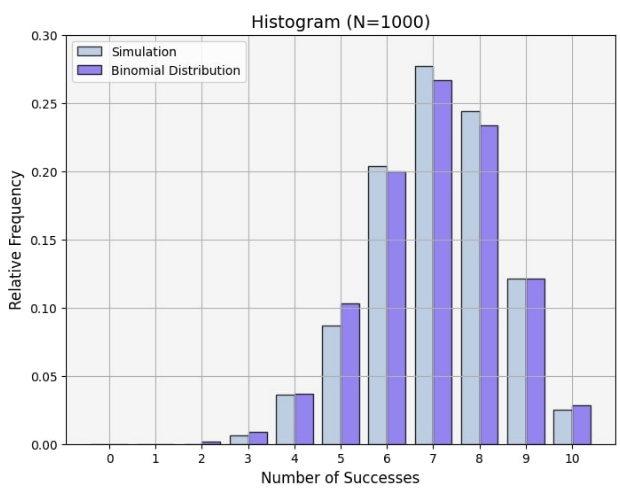

Figure 4

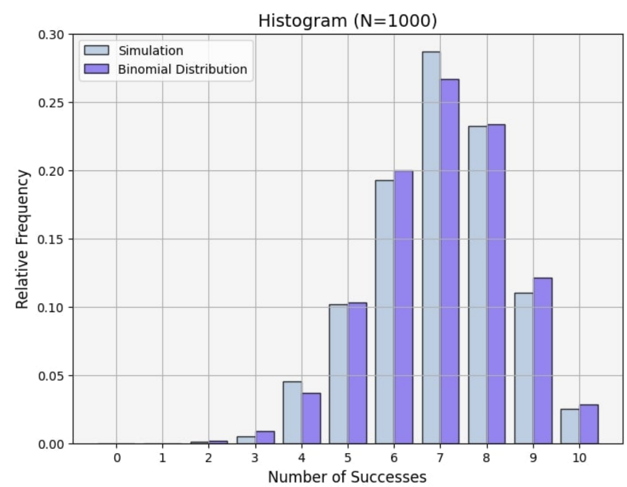

Figure 6