

MC studija za ispitivanje utjecaja distribucije na snagu t testa

Projektni zadatak iz kolegija Računarska statistika

Belcar Eva, Detić Eleonora

8. lipnja 2022.

Opis zadatka

Generiranje
podataka

Snaga t testa

Algoritam 1

Algoritam 2

Rezultati

Fiksna veličina uzorka

Fiksna distribucija

PROC POWER

LAW dataset

- ▶ MC eksperiment za ispitivanje utjecaja distribucije na jednostranu hipotezu t testa

$$H_0 : \mu \geq \mu_0$$

$$H_1 : \mu < \mu_0$$

- ▶ $\mu_0 = 500, 520, 540, 560, 580, 600, 620, 640, 660, 680, 700$
- ▶ veličina uzorka: 10, 15, 20, 50, 100
- ▶ distribucija: Normalna, Laplace, Gamma, Weibul, Uniformna
- ▶ $\mu = 600$
- ▶ $\sigma = 42$
- ▶ broj replikacija: 500
- ▶ alpha: 0.01, 0.05

Generiranje podataka

Zadatak 3

Belcar Eva, Detić
Eleonora

Opis zadatka

Generiranje
podataka

Snaga t testa

Algoritam 1

Algoritam 2

Rezultati

Fiksna veličina uzorka

Fiksna distribucija

PROC POWER

LAW dataset

- ▶ traženje parametra distribucija

$$X \sim N(600, 42^2)$$

$$X \sim \text{Laplace}(600, \frac{42}{\sqrt{2}})$$

$$X \sim 42\sqrt{2} \cdot \text{Gamma}(\frac{1}{2}, 1) + 600 - 21 \cdot \sqrt{2}$$

$$X \sim \frac{21}{\sqrt{5}} \cdot \text{Weibull}(\frac{1}{2}, 1) + 600 - \frac{42}{\sqrt{5}}$$

$$X \sim U(600 - 21 \cdot \sqrt{12}, 600 + 21 \cdot \sqrt{12})$$

- ▶ RAND

Algoritam 1

Zadatak 3

Belcar Eva, Detić
Eleonora

Opis zadatka

Generiranje
podataka

Snaga t testa

Algoritam 1

Algoritam 2

Rezultati

Fiksna veličina uzorka

Fiksna distribucija

PROC POWER

LAW dataset

1. simuliraj 500 uzorka veličine n iz distribucije D
2. za svaki uzorak izračunaj vrijednost testne statistiku t i spremi tu vrijednost
3. za pripadni α izračunaj proporciju uzorka za koje je odbačena H_0 , tj. one uzorke kod kojih je testna statistika t upala u kritično područje
4. izračunata proporcija predstavlja snagu t testa

Algoritam 2

Zadatak 3

Belcar Eva, Detić
Eleonora

Opis zadatka

Generiranje
podataka

Snaga t testa

Algoritam 1

Algoritam 2

Rezultati

Fiksna veličina uzorka

Fiksna distribucija

PROC POWER

LAW dataset

1. simuliraj 500 uzorka veličine n iz distribucije D
2. za svaki uzorak provedi t test s jednostranom alternativnom hipotezom $H_1 : \mu < \mu_0$
3. za svaki provedeni test spremi p vrijednost
4. za pripadni α izračunaj proporciju uzorka za koje je odbačena H_0 , tj. one uzorke kod kojih je $p < \alpha$
5. izračunata proporcija predstavlja snagu t testa

Rezultati za fiksne veličine uzorka

Zadatak 3

Belcar Eva, Detić
Eleonora

Opis zadatka

Generiranje
podataka

Snaga t testa

Algoritam 1

Algoritam 2

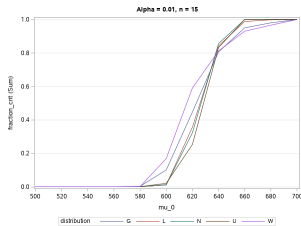
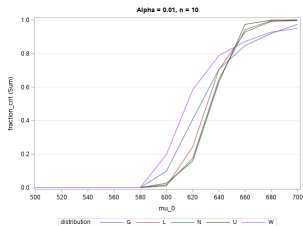
Rezultati

Fiksna veličina uzorka

Fiksna distribucija

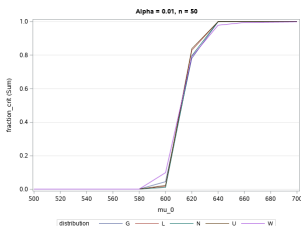
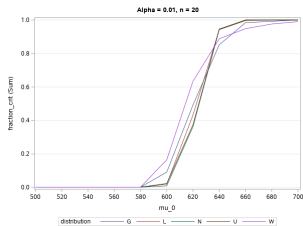
PROC POWER

LAW dataset



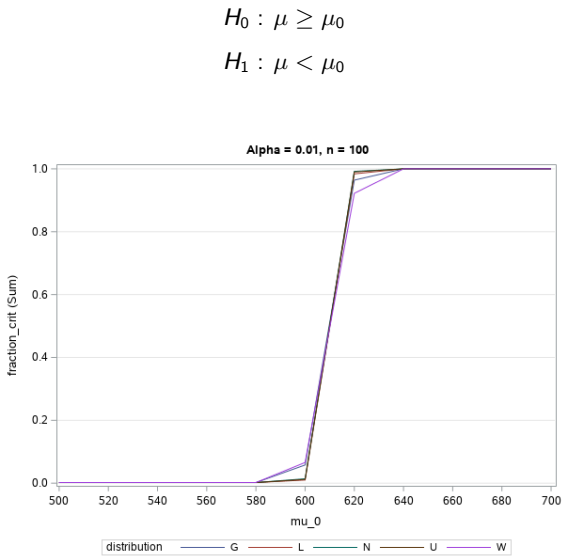
N = 10

N = 15



N = 20

N = 50



N = 100

Rezultati za fiksne distribucije

Zadatak 3

Belcar Eva, Detić
Eleonora

Opis zadatka

Generiranje
podataka

Snaga t testa

Algoritam 1

Algoritam 2

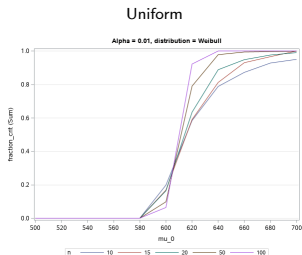
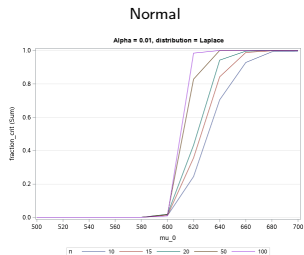
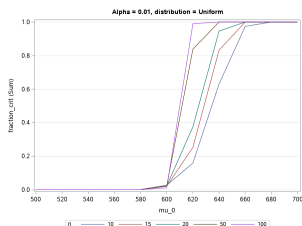
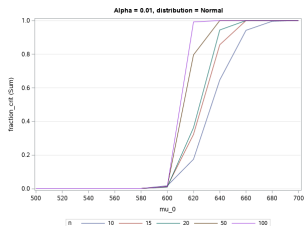
Rezultati

Fiksna veličina uzorka

Fiksna distribucija

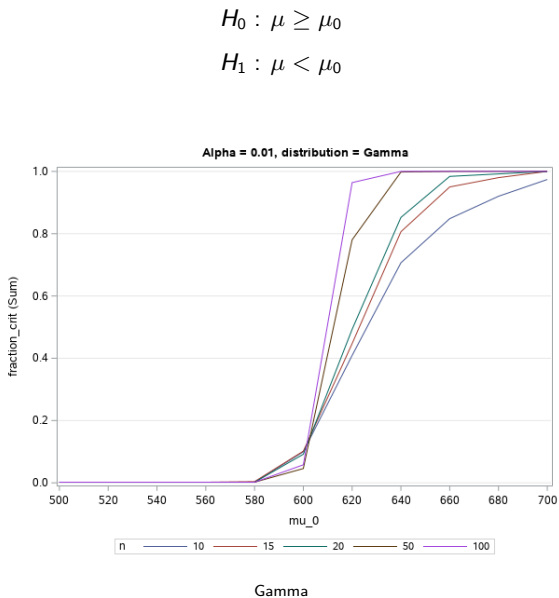
PROC POWER

LAW dataset



Laplace

Weibull



PROC POWER procedura

Zadatak 3

Belcar Eva, Detić
Eleonora

Opis zadatka

Generiranje
podataka

Snaga t testa

Algoritam 1

Algoritam 2

Rezultati

Fiksna veličina uzorka

Fiksna distribucija

PROC POWER

LAW dataset

- ▶ samo za normalnu distribuciju

```
proc power;  
  onesamplemeans test = t sides = L  
  nullmean= 580  
  mean = 600  
  stddev = 42  
  ntotal = 10 to 100 by 5  
  alpha = 0.01 0.05  
  power = .;  
run;
```

- $\alpha = 0.05$, $\mu_0 = 620$, $\mu_1 = 640$

n	MC eksperiment	proc POWER	MC eksperiment	proc POWER
10	0.453	0.400	0.872	0.871
15	0.611	0.544	0.968	0.968
20	0.657	0.658	0.992	0.993
50	0.944	0.953	1.000	> 0.999
100	1.000	0.999	1.000	> 0.999

- zadovoljavajuće podudaranje
- analogni zaključak i za $\alpha = 0.01$

Primjena na stvarnim podacima

Zadatak 3

Belcar Eva, Detić
Eleonora

- ▶ LAW dataset, varijabla LSAT
- ▶ ispitivanje normalosti

Opis zadatka

Generiranje
podataka

Snaga t testa

Algoritam 1

Algoritam 2

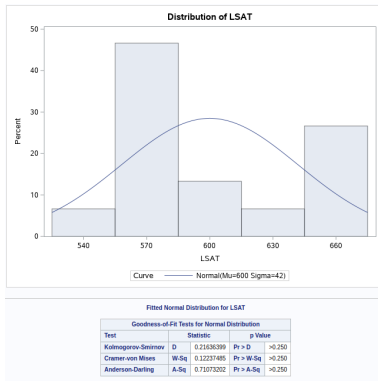
Rezultati

Fiksna veličina uzorka

Fiksna distribucija

PROC POWER

LAW dataset



- ▶ ne odbacujemo H_0

Analysis Variable : LSAT				
N	Mean	Std Dev	Minimum	Maximum
15	600.2666667	41.7945086	545.0000000	666.0000000

- ▶ ispitujemo sljedeću jednostranu hipotezu

$$H_0 : \mu \geq 580$$

$$H_1 : \mu < 580$$

- ▶ p vrijednost = 0.9593
- ▶ Kolika je veličina uzorka potrebna da bi se H_0 odbacila na razini statističke značajnosti $\alpha = 0.01$, a kolika na razini $\alpha = 0.05$ uz snagu testa od otprilike 0.80 (ili više)?

Opis zadatka

Generiranje
podataka

Snaga t testa

Algoritam 1

Algoritam 2

Rezultati

Fiksna veličina uzorka

Fiksna distribucija

PROC POWER

LAW dataset

Opis zadatka

Generiranje
podataka

Snaga t testa

Algoritam 1

Algoritam 2

Rezultati

Fiksna veličina uzorka

Fiksna distribucija

PROC POWER

LAW dataset

Computed Power		
Index	N Total	Power
1	10	<.001
2	15	<.001
3	20	<.001
4	25	<.001
5	30	<.001
6	35	<.001
7	40	<.001
8	45	<.001
9	50	<.001
10	55	<.001
11	60	<.001
12	65	<.001
13	70	<.001
14	75	<.001
15	80	<.001
16	85	<.001
17	90	<.001
18	95	<.001
19	100	<.001

$$\alpha = 0.01$$

Computed Power		
Index	N Total	Power
1	10	0.001
2	15	<.001
3	20	<.001
4	25	<.001
5	30	<.001
6	35	<.001
7	40	<.001
8	45	<.001
9	50	<.001
10	55	<.001
11	60	<.001
12	65	<.001
13	70	<.001
14	75	<.001
15	80	<.001
16	85	<.001
17	90	<.001
18	95	<.001
19	100	<.001

$$\alpha = 0.05$$

Promjena hipoteze

Zadatak 3

Belcar Eva, Detić
Eleonora

$$H_0 : \mu \geq 620$$

$$H_1 : \mu < 620$$

Opis zadatka

Generiranje
podataka

Snaga t testa

Algoritam 1

Algoritam 2

Rezultati

Fiksna veličina uzorka

Fiksna distribucija

PROC POWER

LAW dataset

Computed Power		
Index	N Total	Power
1	40	0.719
2	41	0.732
3	42	0.744
4	43	0.756
5	44	0.768
6	45	0.779
7	46	0.790
8	47	0.800
9	48	0.810
10	49	0.819
11	50	0.828

$\alpha = 0.01$

Computed Power		
Index	N Total	Power
1	20	0.658
2	21	0.678
3	22	0.697
4	23	0.715
5	24	0.732
6	25	0.748
7	26	0.763
8	27	0.778
9	28	0.791
10	29	0.804
11	30	0.816

$\alpha = 0.05$

Hvala na pažnji.