

Sveučilište Josipa Jurja Strossmayera u Osijeku
Odjel za matematiku
Diplomski studij - Financijska matematika i statistika

Istraživanje o Nobelovoj nagradi

Seminarski rad

Student: Lara Juzbašić
Kolegij: Statistika

Osijek, 2020.

Sadržaj

1	Uvod	1
1.1	Nobelova nagrada	1
1.2	Opis varijabli	1
2	Analiza podataka	2
2.1	Spol dobitnika	2
2.1.1	Spol i godina dodjele nagrade	3
2.2	Kategorija	3
2.2.1	Kategorija i spol	5
2.3	Dob dobitnika	6
2.3.1	Dob dobitnika i godina dodjele nagrade	6
2.3.2	Dob dobitnika i kategorija	7
2.4	Godišnji broj dobitnika	7
3	Zaključak	9
	Literatura	10

1 Uvod

U ovom seminarskom radu statističkom analizom ispitani su podaci o dobitnicima Nobelove nagrade. Podaci korišteni u ovom istraživanju nalaze se u bazi "nobel_final.csv". Statističke analize rađene su u programskom jeziku R.

1.1 Nobelova nagrada

Nobelova nagrada jedna je od najznačajnijih svjetskih nagrada za postignuća u područjima znanosti i ljudskog djelovanja. Osnivač fonda Nobelove nagrade bio je Alfred Nobel, švedski kemičar, izumitelj i industrijalac. Od 1901., nagrade se dodjeljuju svake godine na dan njegove smrti.

Pretpostavke koje želimo provjeriti su sljedeće:

- Proporcija žena među dobitnicima manja je od 10%.
- Proporcija žena među dobitnicima se povećala u posljednjih 50 godina.
- Zastupljenost kategorija je različita među nagradama koje su osvojile žene i nagradama koje su osvojili muškarci.
- Prosječna dob dobitnika se s godinama povećava.

1.2 Opis varijabli

Baza sadrži 14 varijabli i 923 jedinice koje predstavljaju dobitnike Nobelove nagrade od 1901. do 2019. godine. Opis varijabli korištenih u analizi nalazi se u Tablici 1.

Ime varijable	Tip varijable	Opis varijable
year	Godina u kojoj je nagrada dodijeljena	Numerička diskretna
gender	Spol dobitnika nagrade ("male" ili "female")	Kvalitativna nominalna
category	Kategorija u kojoj je dodijeljena nagrada	Kvalitativna nominalna
age_get_prize	Dob dobitnika prilikom dodjele nagrade	Numerička neprekidna

Tablica 1: Opis varijabli

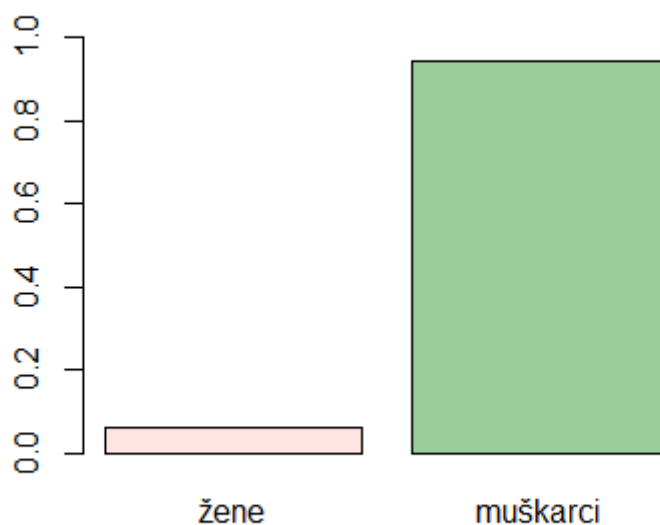
2 Analiza podataka

2.1 Spol dobitnika

Zbog slabe zastupljenosti žena u znanstvenim područjima u ranijim godinama dodjele nagrade, možemo pretpostaviti da je proporcija žena među dobitnicima manja od proporcije muškaraca. U Tablici 2 nalaze se frekvencije dobitnika po spolu te na Slici 1 stupčasti dijagram relativnih frekvencija.

spol	frekvencija
muški	869
ženski	54

Tablica 2: Tablica frekvencija dobitnika po spolu

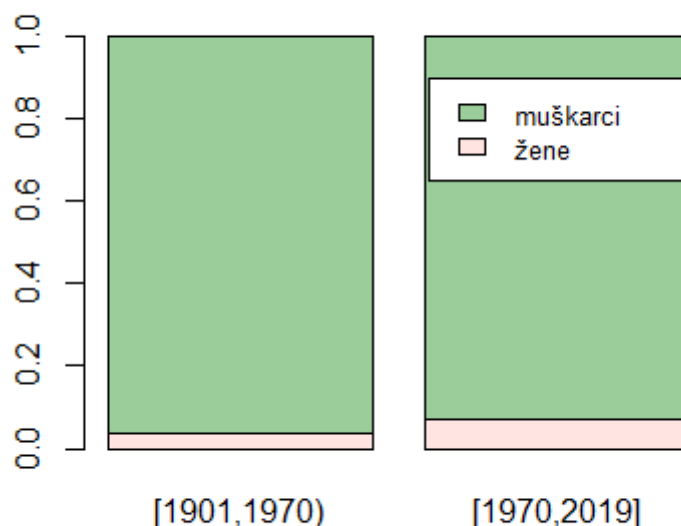


Slika 1: Stupčasti dijagram relativnih frekvencija dobitnika po spolu

Od ukupnog broja dodijeljenih nagrada, tek 54 pripalo je ženama. Primjećujemo da je proporcija žena manja od 10% pa provjerimo je li statistički značajno manja od 10%. Provodeći binomni test dobivamo vrijednost p manju od 0.05 ($p=4.473e-06$), stoga na razini značajnosti 0.05 možemo tvrditi da je proporcija nagrada koje su osvojile žene statistički značajno manja od 10%.

2.1.1 Spol i godina dodjele nagrade

Kako je broj znanstvenica u posljednjih 50 godina porastao, očekujemo veći udio žena među dobitnicima u posljednjih 50 godina nego što je bio prije. Test o jednakosti proporcija za dva nezavisna uzorka potvrđuje našu slutnju ($p=0.01963$), odnosno možemo tvrditi da je proporcija žena veća među dobitnicima kojima je nagrada dodijeljena u razdoblju od 1970. do 2019. godine nego među onima kojima je dodijeljena od 1901. do 1970. godine. Međutim, kao što vidimo na Slici 2, porast proporcije je vrlo malen.



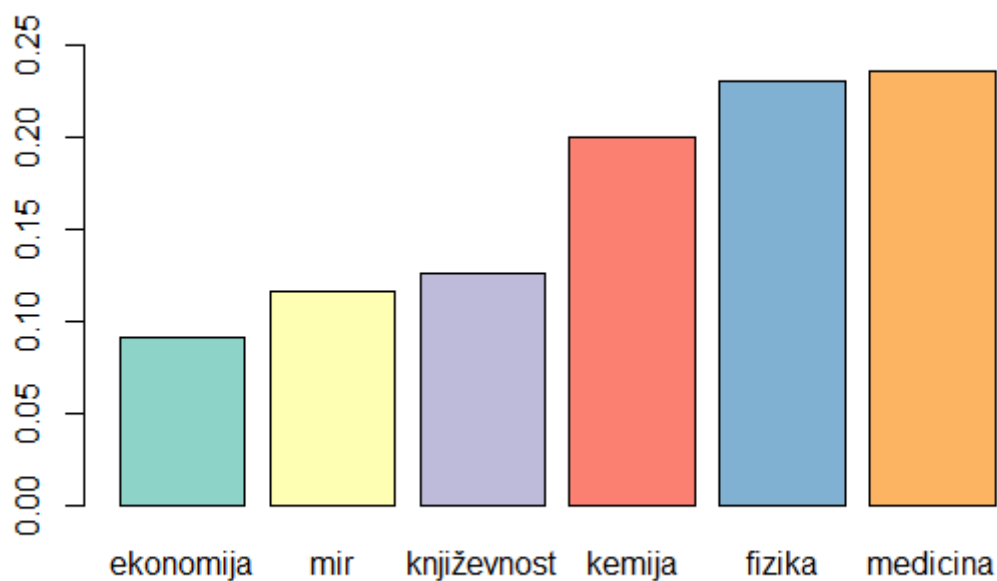
Slika 2: Stupčasti dijagram relativnih frekvencija dobitnika po spolu prije i nakon 1970. godine

2.2 Kategorija

Nobelova nagrada se dodjeljuje za kemiju, fiziku, medicinu, književnost i mir, a od 1969. godine dodjeljuje se i za ekonomiju. U pojedinoj kategoriji mogu biti dodijeljene najviše tri nagrade, a i ne mora biti dodijeljena niti jedna. Promatrajući frekvencije i relativne frekvencije u Tablici 3 te pripadni stupčasti dijagram na Slici 3, možemo uočiti da je među svim dobitnicima najveća proporcija dobitnika nagrade za medicinu, a najmanja za ekonomiju.

kategorija	frekvencija	%
kemija	184	19,9
ekonomija	84	9,1
književnost	116	12,6
medicina	219	23,7
mir	107	11,6
fizika	213	23,1

Tablica 3: Tablica frekvencija i relativnih frekvencija dobitnika po kategorijama

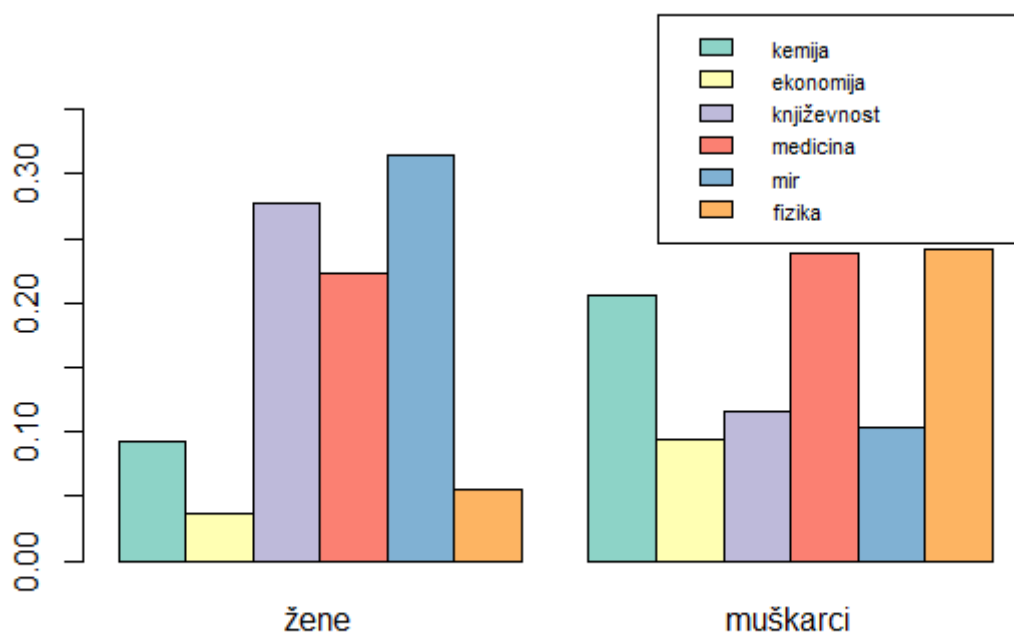


Slika 3: Stupčasti dijagram relativnih frekvencija dobitnika po kategorijama

Provjerimo je li statistički značajno veća proporcija dobitnika iz znanstvenih disciplina (kemije, fizike ili medicine), nego onih koji su osvojili nagradu iz ostalih disciplina. U tu svrhu koristimo binomni test te dobivamo izrazito malu vrijednost p ($p < 2.2e - 16$) pa možemo tvrditi da je proporcija dobitnika iz znanstvenih disciplina statistički značajno veća od proporcije onih koji su osvojili nagradu u ostalim disciplinama.

2.2.1 Kategorija i spol

Zbog različite zastupljenosti žena i muškaraca u pojedinim područjima, možemo naslutiti da kategorije nisu jednako zastupljene među nagradama koje su osvojile žene i nagradama koje su osvojili muškarci. Provedeći χ^2 -test o homogenosti na razini značajnosti 0.05 možemo potvrditi našu slutnju ($p=3.82e-08$). Iz stupčastog dijagrama na Slici 4 možemo vidjeti da je među nagradama osvojenih od strane žena najzastupljenija nagrada za književnost, dok su među nagradama koje su osvojili muškarci najzastupljenije nagrade za fiziku i medicinu.



Slika 4: Stupčasti dijagram relativnih frekvencija kategorija prema spolu

2.3 Dob dobitnika

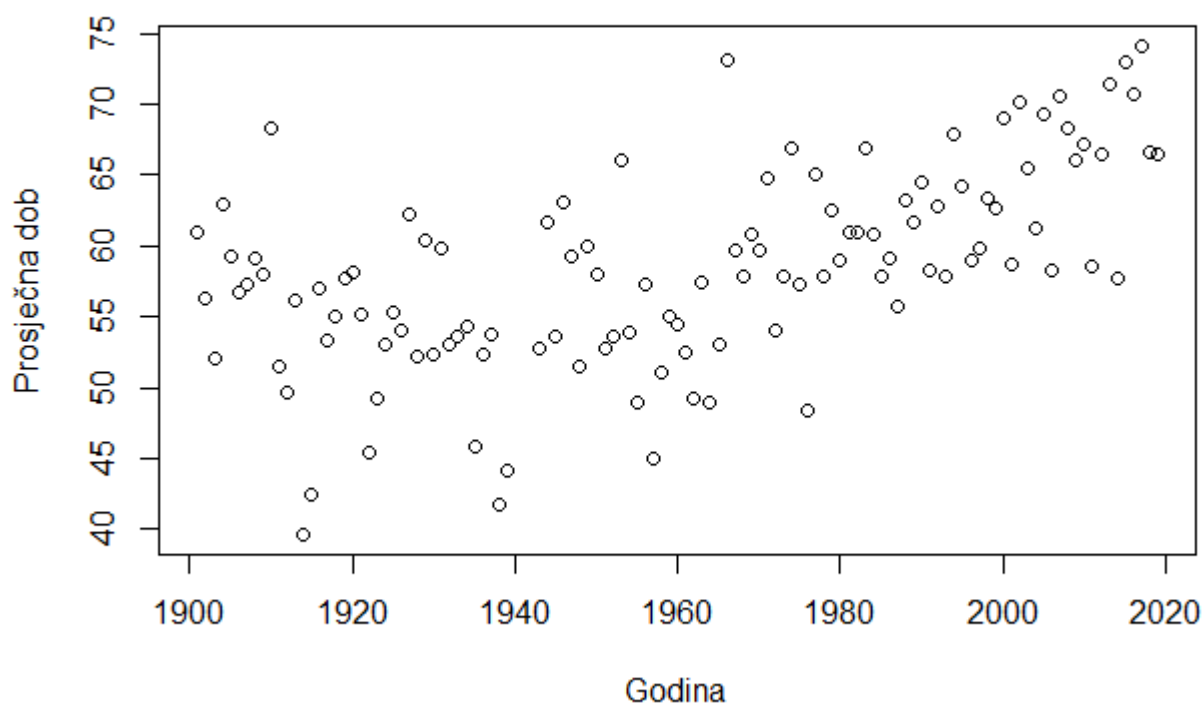
Prosječna dob dobitnika Nobelove nagrade je 59.82 godine. Najstariji dobitnik osvojio je nagradu s 97 godina, dok je najmlađi imao samo 17 godina.

minimum	donji kvartil	medijan	gornji kvartil	maksimum	prosjek	st. dev.
17	51	60	69	97	59.82	12.57

Tablica 4: Deskriptivna statistika dobi dobitnika Nobelove nagrade

2.3.1 Dob dobitnika i godina dodjele nagrade

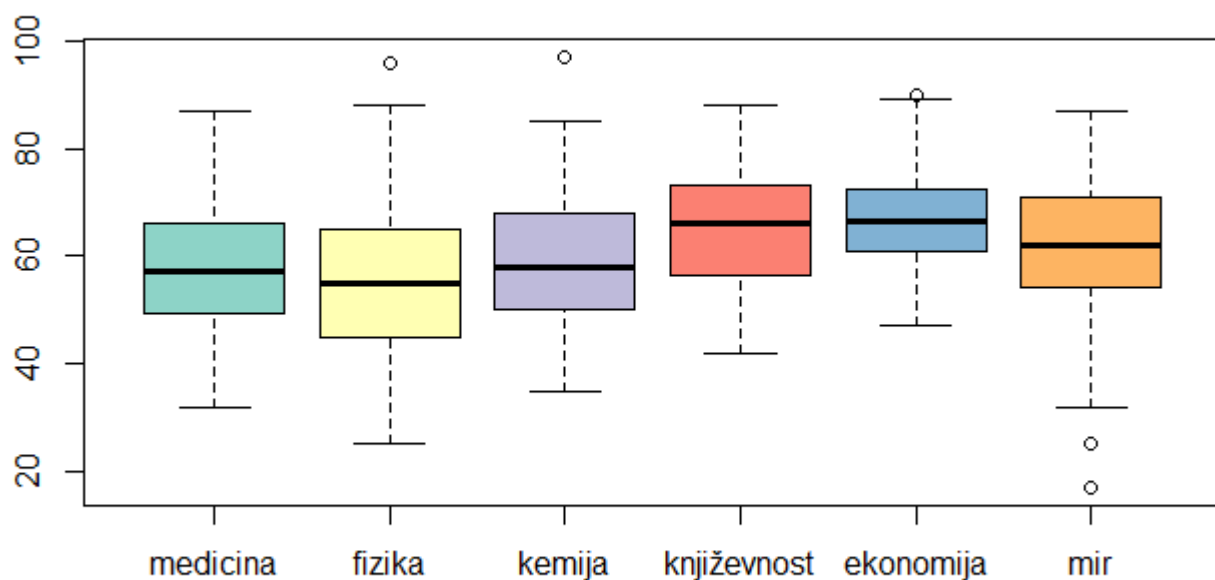
Sve veći broj kandidata za Nobelovu nagradu uzrokuje sve duže vrijeme čekanja od postignuća do dobivanja nagrade. Iz tog razloga možemo pretpostaviti da prosječna dob dobitnika s godinama raste. Korištenjem testa korelacije (Spearman metoda) potvrđujemo našu slutnju ($p=3.242e-12$), tj. možemo tvrditi da postoji rastuća monotona veza između godine dodjele nagrade i prosječne dobi dobitnika.



Slika 5: Odnos godine dodjele nagrade i prosječne dobi dobitnika

2.3.2 Dob dobitnika i kategorija

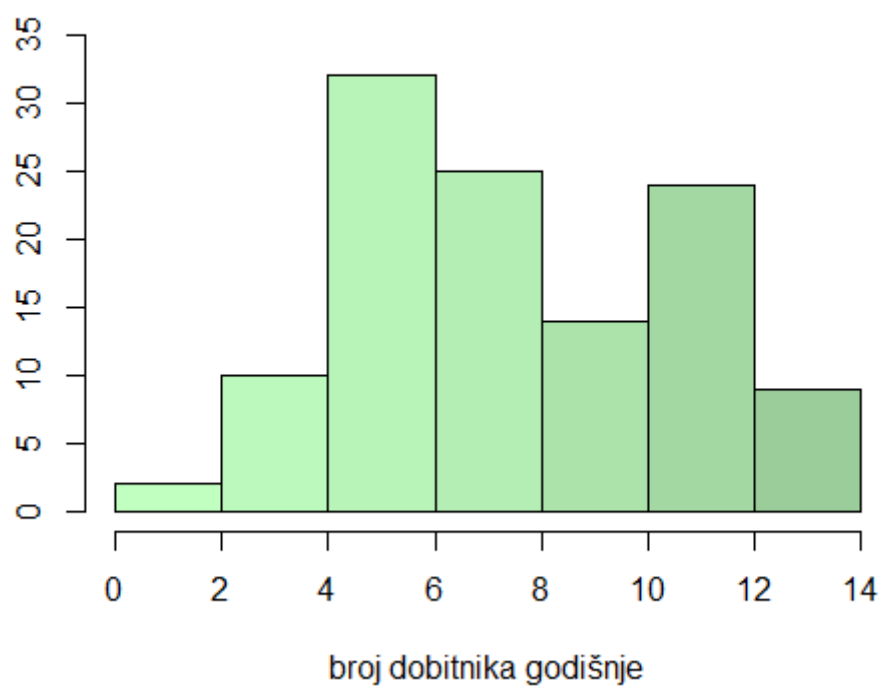
Promatrajući kutijaste dijagrame na Slici 6, primjećujemo razlike u dobi dobitnika iz različitih kategorija. Provjerimo je li ta razlika statistički značajna. Provodeći jednosmjerni ANOVA test potvrđujemo našu slutnju, tj. na razini značajnosti 0.05 možemo tvrditi da očekivana dob dobitnika nije jednaka u svim kategorijama ($p=4.56e-13$). Nadalje, provođenjem HSD Tuckey testa, uočavamo da su očekivane dobi dobitnika nagrade za ekonomiju i književnost statistički značajno veće od očekivanih dobi dobitnika nagrade za kemiju, medicinu i fiziku, na razini značajnosti 0.05 (p-vrijednosti: 0.0000166, 0.0008231, 0.0000016, 0.00, 0.0000955, 0.00).



Slika 6: Kutijasti dijagram dobi dobitnika prema kategorijama

2.4 Godišnji broj dobitnika

Želimo provjeriti koliki je očekivani godišnji broj dobitnika Nobelove nagrade. Procijenimo li očekivani broj dobitnika aritmetičkom sredinom, dobivamo 7.96 dobitnika, stoga provjerimo je li očekivani godišnji broj dobitnika različit od 8. Korištenjem Welchovog t-testa dobivamo vrijednost p veću od 0.05 ($p=0.8773$), tj. nema dokaza da je očekivani godišnji broj dobitnika različit od 8.



Slika 7: Histogram godišnjeg broja dobitnika Nobelove nagrade

3 Zaključak

Analizom podataka iz baze "nobel_final.csv", na razini značajnosti 0.05, donešeni su sljedeći zaključci o dobitnicima Nobelove nagrade:

- Proporcija žena među dobitnicima manja je od 10%.
- Proporcija žena među dobitnicima se povećala u posljednjih 50 godina.
- Proporcija dobitnika nagrade iz znanstvenih disciplina veća je od proporcije onih koji su osvojili nagradu u ostalim disciplinama.
- Zastupljenost kategorija je različita među nagradama koje su osvojile žene i nagradama koje su osvojili muškarci.
- Prosječna dob dobitnika se s godinama povećava.
- Očekivana dob dobitnika nagrade za ekonomiju i književnost je veća od očekivane dobi dobitnika nagrade za kemiju, medicinu i fiziku.
- Očekivani godišnji broj dobitnika je 8.

Literatura

- [1] <https://github.com/BahramJannesar/NobelPrize/tree/master/dataset>
- [2] https://hr.wikipedia.org/wiki/Nobelova_nagrada
- [3] <https://www.nobelprize.org>