UCG
University of Montenegro

Centralna granična teorema

Vježbe VI

04.04.2018.
Nemanja Batrićević

Uzorak i populacija

- Do sad smo radili sa pojedinačnim vrijednostima, određivali njihovu poziciju u ukupnoj populaciji **(z-skor)** i računali vjerovatnoće da izvučemo takvu vrijednost **(z-tabela)**

- **Problem**: većina istraživačkih poduhvata zahtijeva rad sa većim uzorcima, ne pojedinačnim vrijednostima

- **Cilj:** primijeniti istu logiku i pojmove prilikom rada sa većim uzorcima - *pretvaranje aritmetičke sredine uzorka u z-skor*
- **Suština** ostaje ista: objasniti đe se konkretni uzorak nalazi u distribuciji svih ostalih potencijalnih uzoraka
  - Baš kao i pojedinačne vrijednosti, ako uzorak ima z-skor blizu "0" to znači da imamo tipičan uzorak. Ukoliko ima vrijednosti veće od 2 ili manje od -2, to znači da je uzorak ekstreman

Uzorak i populacija

- Uzorci nude **nekompletnu** sliku populacije
- **Greška uzorkovanja** – *prirodna diskrepanca između karakteristika uzorka i populacije*
- Uzorci **variraju**, sadrže različite pojedince, imaju različite aritmetičke sredine, itd.
- U većini slučajeva moguće je izvući na hiljade potencijalnih uzoraka iz jedne populacije

- ***Kako onda možemo znati išta o populaciji na osnovu uzorka?***

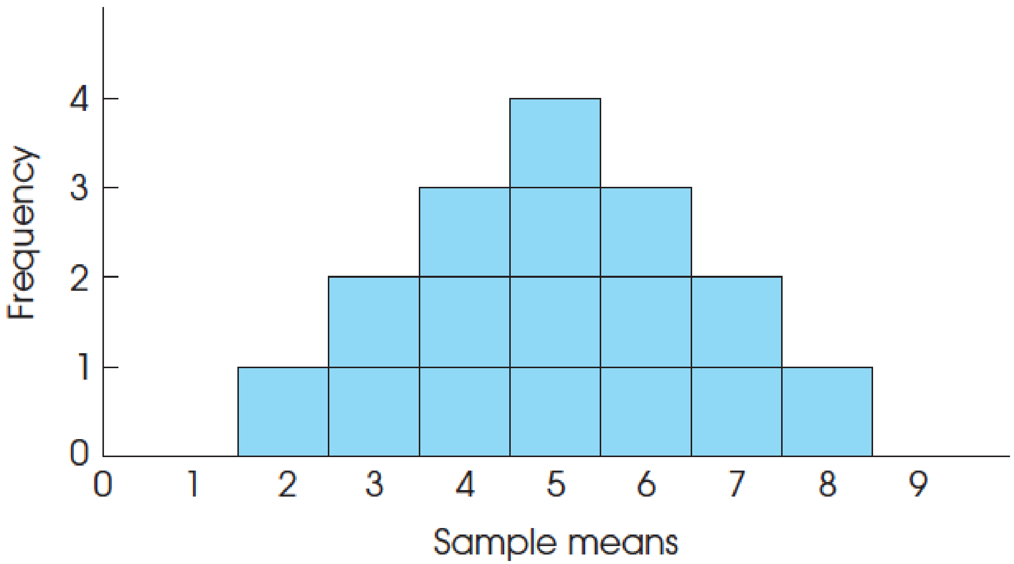
Distribucija sredina uzoraka

- Veliki broj izvučenih uzoraka kreira “šablon” koji omogućava da predvidimo karakteristike uzorka – **distribucija sredina uzoraka**
- **Definicija**: *Distribucija sredina uzoraka je kolekcija uzoračkih sredina svih mogućih nasumičnih uzoraka veličine (n).*
- VAŽNO: **distribucija svih mogućih uzoraka!**
- Sredina uzorka bi se trebale ”gomilati” oko aritmetičke sredina u populaciji
- “Gomila” prati normalnu distribuciju = rijetki su uzorci sa netipičnom strurkturom
- Što je uzorak veći to se odstupanje od aritmetičke sredine smanjuje

Primjer:

FIGURE 7.2

The distribution of sample means for $n = 2$. The distribution shows the 16 sample means from Table 7.1.



Centralna granična teorema

- U realnim situacijama uzorak se **ne sastoji** od 2 već od stotina/hiljada vrijednosti

- Broj mogućih uzoraka je izuzetno velik, pa je **nemoguće** zapravo izvući svaki od potencijalnih uzoraka

- Srećom, moguće je da odredimo kako izgleda distribucija sredina uzoraka **bez** da smo primorani da izvlačimo na stotine uzoraka

- **Centralna granična teorema** – nudi precizan opis distribucije koji bi dobili da smo izvukli svaki mogući uzorak, izračunali sredinu uzorka, itd.

Centralna granična teorema

- **Definicija:** *Za svaku populaciju sa aritmetičkom sredinom* **μ** *i standardnom devijacijom* **σ**, *distribucija sredina uzoraka veličine "n" imaće aritmetičku sredinu M i standardnu devijaciju* **σ**/korijen(n), i približavače se normalnoj distribucij kako se "n" približava beskonačnosti.

- **Upotrebna vrijednost: 1.)** važi za svaku populaciju, nezavisno od oblika, prosjeka, standardne devijacije; **2.)** dostribucija sredina uzoraka postaje normalna već nakon 30 opservacija
- Distribucija je normalna ako je jedan od dva uslova ispunjen: 1.) populacija je normalno distribuirana 2.) veličina uzorka je veća od 30
- Karakteristike populacije su najčešće nepoznate, pa se oslanjamo na **2.)**

Ključne formule

- Svaki uzorak će imati svoju aritmetičku sredinu
- Mjera varijabilnosti aritmetičkih sredina iz uzorka u uzorak je izražena **standardnom greškom aritmetičke sredine**
- Ako populacija ima normalnu distribuciju (raspored) sa aritmetičkom sredinom μ i standadnom devijacijom σ, onda distribucija aritmetičkih sredina uzoraka je takođe normalna sa:

$$\mu_{\bar{X}} = \mu$$

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$$

$$Z = \frac{(\bar{X} - \mu_{\bar{X}})}{\sigma_{\bar{X}}} = \frac{(\bar{X} - \mu)}{\frac{\sigma}{\sqrt{n}}}$$

$\bar{X}$ = aritmetička sredina uzorka
μ = aritmetička sredina populacije
σ = standardna devijacija populacije
n = veličina uzorka

Primjeri

- *Za populaciju sa aritmetičkom sredinom $\mu = 70$ i standardnom devijacijom $\sigma = 20$, koliku razliku očekujete, u prosjeku, između aritmetičke sredine uzorka i populacija, za svaku od sljedećih veličina uzorka:*

A. n = 4
B. n = 16
C. n = 25

Primjeri

- *Za populaciju sa aritmetičkom sredinom $\mu = 70$ i standardnom devijacijom $\sigma = 20$, koliku razliku očekujete, u prosjeku, između aritmetičke sredine uzorka i populacija, za svaku od sljedećih veličina uzorka:*

A. n = 4 10
B. n = 16 5
C. n = 25 4

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

Primjeri

- *Za populaciju sa aritmetičkom sredinom $\mu = 80$ i standardnom devijacijom $\sigma = 12$, pronađi z-skor koji odgovara sljedećim uzorcima, i odredi da li je uzorak "reprezentativan":*

A. M = 83 sa veličinom uzorka n = 4
B. M = 83 sa veličinom uzorka n = 16
C. M = 83 sa veličinom uzorka n = 36

Primjeri

- *Za populaciju sa aritmetičkom sredinom $\mu = 80$ i standardnom devijacijom $\sigma = 12$, pronađi z-skor koji odgovara sljedećim uzorcima, i odredi da li je uzorak "reprezentativan":*

A. M = 83 sa veličinom uzorka n = 4 $\sigma_x = 6$ sa **z = 0.50**
B. M = 83 sa veličinom uzorka n = 16 $\sigma_x = 3$ sa **z = 1.00**
C. M = 83 sa veličinom uzorka n = 36 $\sigma_x = 2$ sa **z = 1.50**

$$\boxed{Z = \frac{(\bar{X} - \mu_{\bar{X}})}{\sigma_{\bar{X}}} = \frac{(\bar{X} - \mu)}{\frac{\sigma}{\sqrt{n}}}}$$

Primjeri

- *Populacija IQ skorova ima normalnu distribuciju (raspored) sa aritmetičkom sredinom $\mu = 100$ i standardnom devijacijom $\sigma = 15$. Koja je vjerovatnoća da ćemo izvući uzorak sa aritmetičkom sredinom većom od 97:*

A. Za uzorak sa veličinom n = 9

B. Za uzorak sa veličinom n = 25

Primjeri

- *Populacija IQ skorova ima normalnu distribuciju (raspored) sa aritmetičkom sredinom $\mu = 100$ i standardnom devijacijom $\sigma = 15$. Koja je vjerovatnoća da ćemo izvući uzorak sa aritmetičkom sredinom većom od 97:*

A. Za uzorak sa veličinom n = 9 $\sigma_x = 5$; $z = -0.60$

B. Za uzorak sa veličinom n = 25 $\sigma_x = 3$; $z = -1.00$

Primjeri

- *Populacija IQ skorova ima normalnu distribuciju (raspored) sa aritmetičkom sredinom $\mu = 100$ i standardnom devijacijom $\sigma = 15$. Koja je vjerovatnoća da ćemo izvući uzorak sa aritmetičkom sredinom većom od 103:*

A. Za uzorak sa veličinom n = 9 $\sigma_{\bar{x}} = 5$; $z = 0.60$

B. Za uzorak sa veličinom n = 25 $\sigma_{\bar{x}} = 3$; $z = 1.00$

$$Z = \frac{(\bar{X} - \mu_{\bar{X}})}{\sigma_{\bar{X}}} = \frac{(\bar{X} - \mu)}{\frac{\sigma}{\sqrt{n}}}$$

(A) z	(B) Proportion in Body	(C) Proportion in Tail	(D) Proportion Between Mean and z
0.58	.7190	.2810	.2190
0.59	.7224	.2776	.2224
0.60	.7257	.2743	.2257
0.61	.7291	.2709	.2291
0.62	.7324	.2676	.2324
0.63	.7357	.2643	.2357
1.00	.8413	.1587	.3413
1.01	.8438	.1562	.3438
1.02	.8461	.1539	.3461