Statistikë e aplikuar

UNIVERSITETI I EJL
ЈИЕ УНИВЕРЗИТЕТ
SEE UNIVERSITY

Shpërndarjet e zgjedhjeve

Faton Berisha

Kapitulli 6

Shpërndarjet e zgjedhjeve

Shpërndarjet e zgjedhjeve

6.1 Shpërndarja e zgjedhjeve e mesatares së mostrës

6.2 Shpërndarja e zgjedhjeve e proporcionit të mostrës

Shpërndarja e zgjedhjeve e mesatares së mostrës

Shpërndarja e zgjedhjeve e mesatares së mostrës $\bar{x}$ është shpërndarja e probabilitetit e popullimit të mesatareve të të gjitha mostrave të mundshme për t'u zgjedhur të madhësisë *n* nga një popullim i madhësisë *N.*

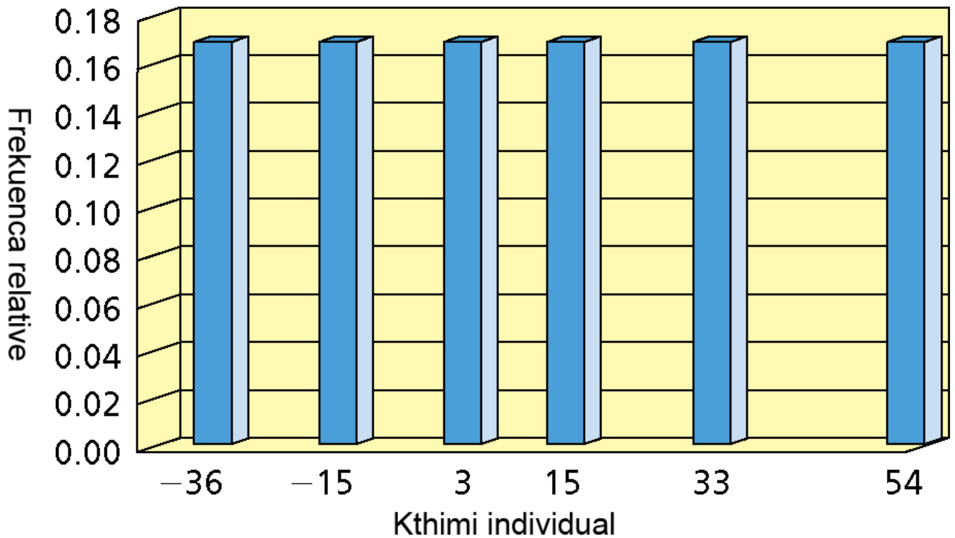
Shembull: Zgjedhja e % të kthimeve

- Popullimi i përqindjes së kthimeve nga 6 letra me vlerë
  - Të rradhitura, vlerat e % së kthimeve:

    -36%, -15%, 3%, 15%, 33%, 54%
    - Shenojmë secilën letër me vlerë me A, B, C, …, F në rradhitje të % së kthimeve
    - Mesatarja e kthimeve është 9% me devijim standard 29.65%
  - Secila nga këto letra me vlerë ka gjasë të njëjtë të zgjedhjes sikur secila tjetër
    - Shpërndarje uniforme me $N = 6$
    - Secila letër me vlerë ka probabilitet të zgjedhet $1/6 = 0.1667$

Shembull: Zgjedhja e % të kthimeve. (Vazhdim)

Letra me vlerë	*% e kthimeve*	*Frekuenca*	*Frekuenca relative*
A	10	1	1/6
B	20	1	1/6
C	30	1	1/6
D	40	1	1/6
E	50	1	1/6
F	60	1	1/6
Gjithsej		**6**	**1**



Shembull: Zgjedhja e % të kthimeve. (Vazhdim)

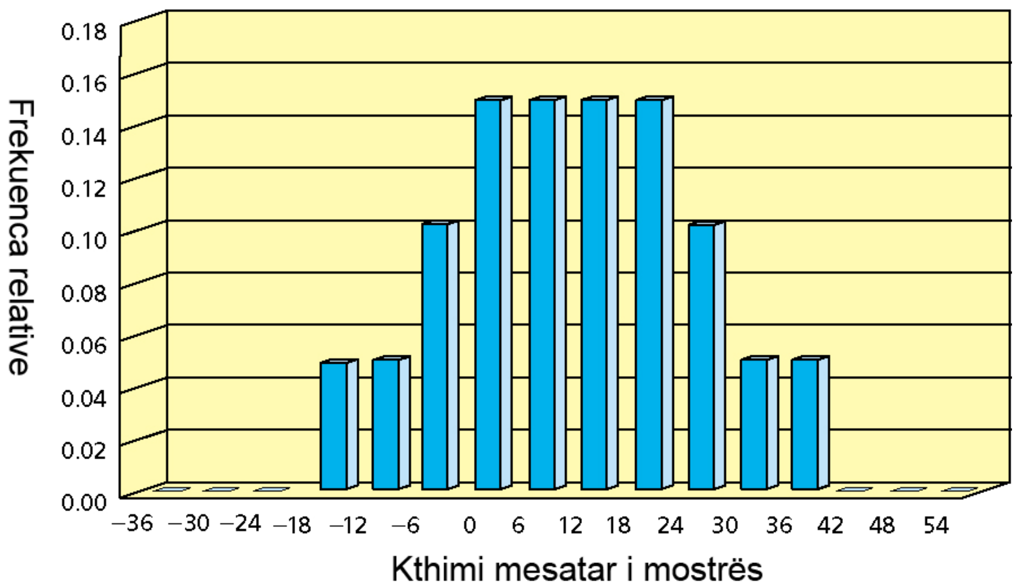
- Tani, zgjedhim të gjitha mostrat e mundura të madhësisë $n = 3$ nga ky popullim i letrave i madhësisë $N = 6$
- Si të zgjedhet një mostër?
  - Zgjedhja e rastësishme
  - Zgjedhja pa kthim
  - Zgjedhja pa marrë parasysh renditjen

Shembull: Zgjedhja e % të kthimeve. (Vazhdim)

- Rezultati: Ekzistojnë 20 mostra të madhësisë n = 3.
- Llogarisim mesataren e mostrës të secilës mostër

Shembull: Zgjedhja e % të kthimeve. (Vazhdim)

	n = 3 Kthimet			*Mesatarja*
Mostra	*në mostrën*			*e mostrës*
1	-36	-15	3	-16.00
2	-36	-15	15	-12.00
3	-36	-15	33	-6.00
4	-36	-15	54	1.00
5	-36	3	15	-6.00
6	-36	3	33	0.00
7	-36	3	54	7.00
8	-36	15	33	4.00
9	-36	15	54	11.00
10	-36	33	54	17.00
11	-15	3	15	1.00
12	-15	3	33	7.00
13	-15	3	54	14.00
14	-15	15	33	11.00
15	-15	15	54	18.00
16	-15	33	54	24.00
17	3	15	33	17.00
18	3	15	54	24.00
19	3	33	54	30.00
20	15	33	54	34.00



Observime

- Edhepse popullimi i kthimeve të *N* = 6 letrave me vlerë ka shpërndarje uniforme,...
- … histogrami i 20 kthimeve mesatare të mostrave:
  1. Duket të jetë i përqendruar rreth kthimit mesatar të mostrës prej 35%
  2. Duket të jetë në formë këmbane dhe më pak i përhapur sesa histogrami i kthimeve individuale.

Përfundime të përgjithshme

- Në qlftë se popullimi i njësive individuale është normal, atëherë popullimi i mesatareve të të gjitha mostrave është poashtu normal.
- Madje edhe në qoftë se popullimi i njësive individuale nuk është normal, ekzistojnë rrethana kur popullimi i mesatareve të mostrave është normal (Teorema qendrore limite).

Përfundime të përgjithshme. (Vazhdim)

- Mesatarje e të gjitha mesatareve të mostrave të mundshme është e barabartë me mesataren e popullimit
  - D.m.th., $\mu = \mu_{\bar{x}}$
- Devijimi standard $\sigma_{\bar{x}}$ i të gjitha mesatareve të mostrave është më i vogël sesa devijimi standard i popullimit.
  - D.m.th., $\sigma_{\bar{x}} < \sigma$
    - Secila mesatare mostre gjen mesin aritmetik të masave të larta dhe të ulëta, prandaj është më afër μ sesa shumë nga masat e popullimit individual.

Rregulla empirike

- Vlen rregulla empirike për shpërndarjen e zgjedhjeve të mesatareve të mostrave
  - 68.26% të të gjitha mesatareve të mostrave të mundshme janë brenda (plus ose minus) onjë devijimi standard $\sigma_{\bar{x}}$ nga μ.
  - 95.44% të të gjitha vlerave të mundshme të observuara të $\bar{x}$ janë brenda dy $\sigma_{\bar{x}}$ nga μ
  - 99.73% të të gjitha vlerave të mundshme të observuara të $\bar{x}$ janë brenda tri $\sigma_{\bar{x}}$ nga μ

Vetitë e shpërndarjes së zgjedhjeve të mesatareve të mostrave

- Në qoftë se popullimi nga i cili zgjedhen mostrat është normal, atëherë e tillë është edhe shpërndarja e mesatares së mostrës, $\bar{x}$
- Mesatarja $\sigma_{\bar{x}}$ e shpërndarjes së zgjedhjeve x̄ është

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

- D.m.th., mesatarja e të mesatareve të të gjitha mostrave të mundshme është e barabartë me mesataren e popullimit.

Vetitë e shpërndarjes së zgjedhjeve të mesatareve të mostrave. (Vazhdim)

- Varianca $\sigma^2_{\bar{x}}$ e shpërndarjes së zgjedhjeve $\bar{x}$ është

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}$$

  - D.m.th., varianca e shpërndarjes së zgjedhjeve të $\bar{x}$ është
    - në proporcion të drejtë me variancën e popullimit
    - në proporcion të zhdrejtë me madhësinë e mostrës

Vetitë e shpërndarjes së zgjedhjeve të mesatareve të mostrave. (Vazhdim)

- Devijimi standard $\sigma_{\bar{x}}$ e shpërndarjes së zgjedhjeve të $\bar{x}$ është

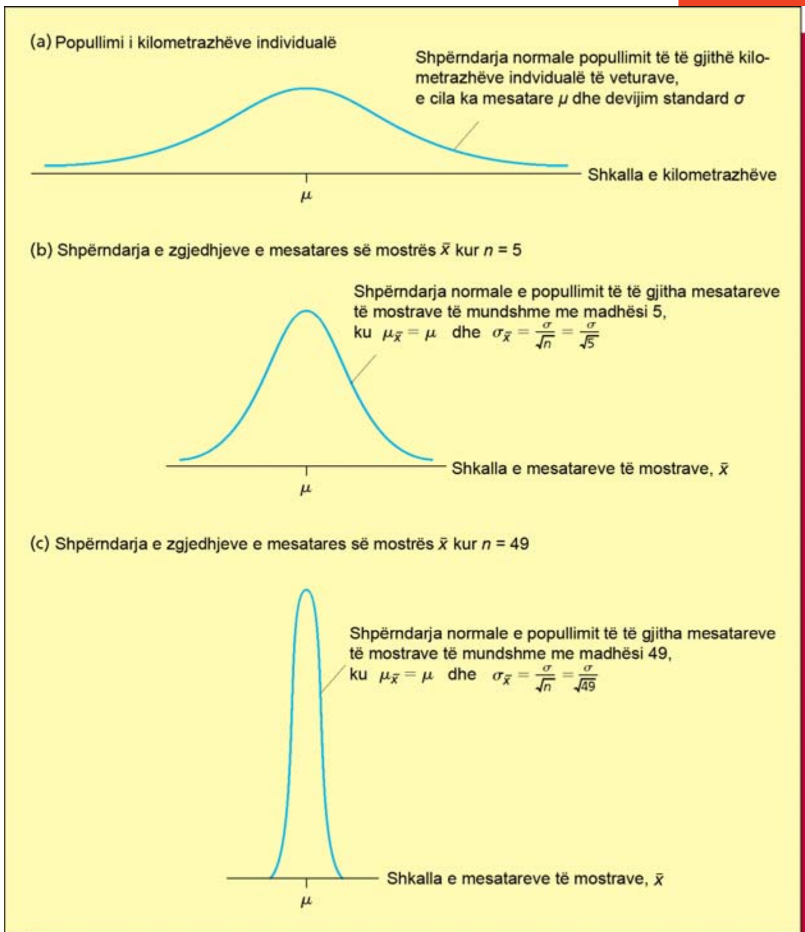
$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

  – D.m.th., varianca e shpërndarjes së zgjedhjeve të $\bar{x}$ është
    - në proporcion të drejtë me devijimin standard të popullimit
    - në proporcion të zhdrejtë me rrënjën katrore të madhësisë së mostrës

Vërejtje

- Formulat për $\sigma^2_{\bar{x}}$ dhe $\sigma_{\bar{x}}$ janë të sakta në qoftë se popullimi nga i cili zgjedhet mostra është i pafundmë.
- Formulat janë përafërsisht të sakta në qoftë se popullimi nga i cili zgjedhet mostra është i fundmë por *N* është shumë më i madh (së paku 20 herë) sesa *n* (*N*/*n* ≥ 20)
  - $\bar{x}$ është vlerësimi pikësor i μ, dhe sa më e madhe të jetë madhësia e mostrës *n*, aq më i saktë është vlerësimi
  - Me rritjen e *n*, $\sigma_{\bar{x}}$ zvogëlohet meqë $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$
    - Përmëtepër, sa më i madh të jetë *n*, aq më reprezentative është mostra e popullimit.
      - Prandaj, për të zvogëluar $\sigma_{\bar{x}}$, merrni mostra të mësha!

Efekti i madhësisë së mostrës



Shembull 6.2. Rasti i kilometrazhit të veturave

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{5}} = \frac{\sigma}{2.2361}$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{49}} = \frac{\sigma}{7}$$

Pra, mesataret e ndryshme të mostrave të mundshme do të jenë më afër të grumbulluara rreth μ në qlftë se n = 49 sesa në qoftë se n =5.

Të konkluduarit nga shpërndarja e zgjedhjeve

- Rikujtoni shembullin e kilometrazhit nga kap. 2, $\bar{x}$ = 31.5531 mpg për një mostër madhësie n=49.
  - Me s = 0.7992
- A ofron kjo një dëshmi statistikore se mesatarja e popullimit μ është më e madhe se 31 mpg?
  - D.m.th., a ofron mesatarja e mostrës dëshmi se μ është së paku 31 mpg?
- Llogaritet prbabiliteti i observimit të një mesatare mostre e cila është më e madhe ose baraz me 31.5531 mpg në qoftë se μ = 31 mpg
  - Gjejmë $P(\bar{x} \geq 31.5531$ në qoftë se $\mu = 31)$

Të konkluduarit nga shpërndarja e zgjedhjeve. (Vazhdim)

- Shfrytëzojmë s si vlerësim pikësor për σ:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{s}{\sqrt{n}} = \frac{0.7992}{\sqrt{49}} = 0.1143$$

$$\begin{aligned} P(\bar{x} \geq 31.5531 \text{ në qoftë se } \mu = 31) &= P\left(z \geq \frac{31.5531 - \mu_{\bar{x}}}{\sigma_{\bar{x}}}\right) \\ &= P\left(z \geq \frac{31.5531 - 31}{0.1143}\right) \\ &= P(z \geq 4.84) \end{aligned}$$

- Por z = 4.84 është jashtë tabelës standarde normale.
- Vlera më e madhe e z në tabelën është 3.09, e cila ka syprinë të bishtit të djathtë 0.001.

Të konkluduarit nga shpërndarja e zgjedhjeve. (Vazhdim)

- $z = 4.84 > 3.09$, prandaj $P(z \geq 4.84) < 0.001$
- D.m.th., në qoftë se $\mu = 31$ mpg, atëherë më pak se 1 në 1,000 nga të gjitha mostrat e mundshme kanë mesataren së paku aq të madhe sa e observuara.
- Imponohen dy shpjegimet vijuese:
  - Në qoftë se μ është aktualisht 31 mpg, atëherë zgjedhje tepër fatkeqe e mostrës

    OSE
  - Jo fatkeqe, por μ nuk është 31 mpg, por është në fakt më e madhe
- Vështirë për të besuar se do të ndodhte rast aq i pagjasë, prandaj konkludojmë se ekziston dëshmi e fortë se μ është më e madhe se 31 mpg

Teorema qendrore limite

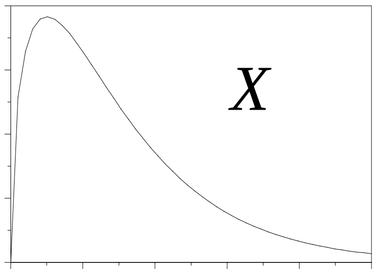
- Shqyrtojmë zgjedhje mostrash nga një popullim me shpërndarje jo normale.
- Akoma kemi: $\mu_{\bar{x}} = \mu$ dhe $\sigma_{\bar{x}} = \sigma / \sqrt{n}$
  - Saktësisht i saktë në qoftë se popullimi i pafundmë
  - Përafërsisht i saktë në qoftë se madhësia *N* e popullimit e fundme por shumë më e madhe se madhësia *n* e mostrës.
    - Sidomos në qoftë se $N / n \geq 20$
- Por në qoftë se popullimi është jo normal, çfarë është forma e shpërndarjes së zgjedhjeve të mesatares së mostrës?
  - A është normale, siç do të ishte sikur popullim të ishte normal?
  - Po, shpërndarja e zgjedhjeve është përafërsisht normale në qoftë se mostra është mjaft e madhe, madje edhe për popullime jo normale.
    - Sipas “Teoremës qendrore limite”

Teorema qendrore limite. (Vazhdim)

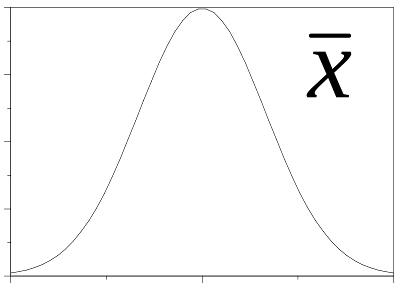
- Pa marrë parasysh çfarë është shpërndarja e probabilitetit e popullimit, në qoftë se madhësia *n* e mostrës është mjaft e madhe, atëherë popullimi i të gjitha mesatareve të mostrave të mundshme është ***përafërsisht*** normal me mesatare $\mu_{\bar{x}} = \mu$ dhe devijim standard $\sigma_{\bar{x}} = \sigma / \sqrt{n}$
- Përmëtepër, sa më e madhe të jetë madhësia *n* e mostrës aq më afër për të qenë normale është shpërndarja e zgjedhjeve të mesatares së mostrës.
  - Me fjalë tjera, sa më e madhe *n*, aq më i mirë përafrimi

Teorema qendrore limite. (Vazhdim)

Mostër e çfarëdoshme (x_1, x_2, …, x_n)



për $n \to$ të madh



Shpërndarja e popullimit

(μ, σ)

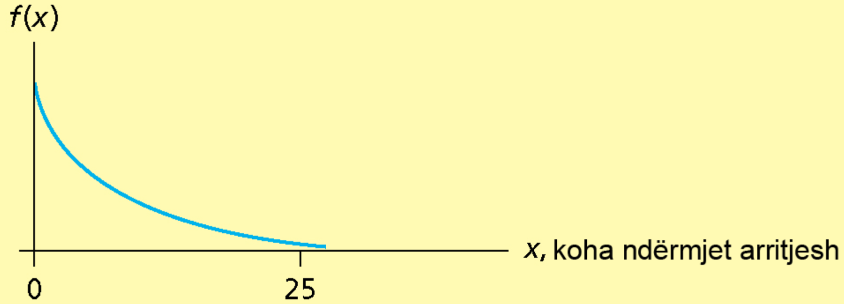
(e shtrembëruar djathtas)

Shpërndarja e zgjedhjeve e mesatarës së mostrës

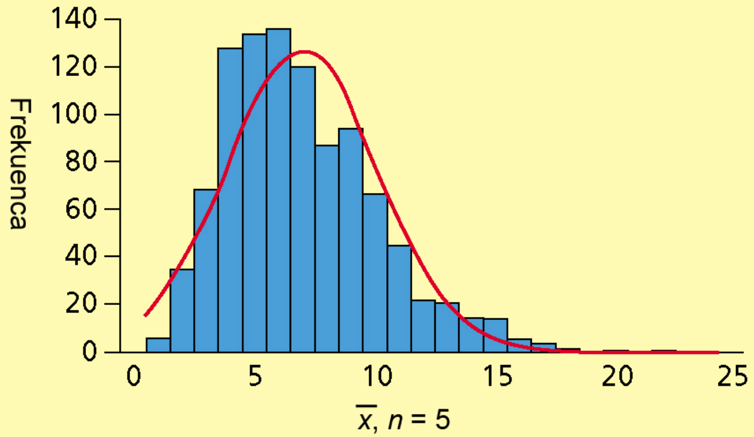
$$\left(\mu_{\bar{x}} = \mu, \sigma_{\bar{x}} = \sigma/\sqrt{n}\right)$$

(afërsisht normale)

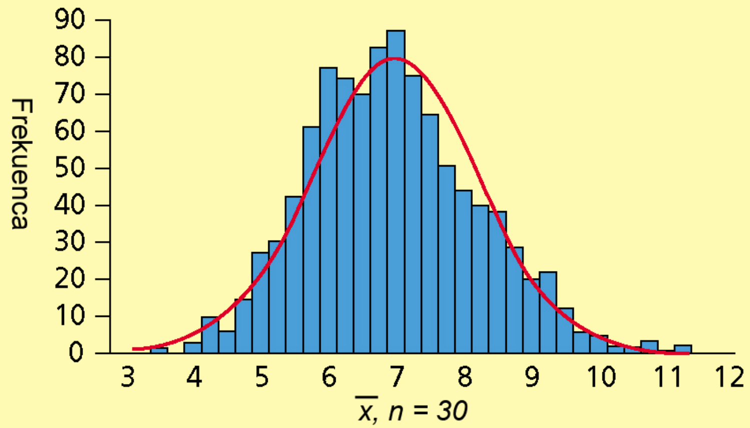
Shembull: Simulimi i teoremës qendrore limite



(a) Shpërrndarja eksponenciale e cila përshkruan kohët ndërmjet arritjesh në dhomën e emergjencës

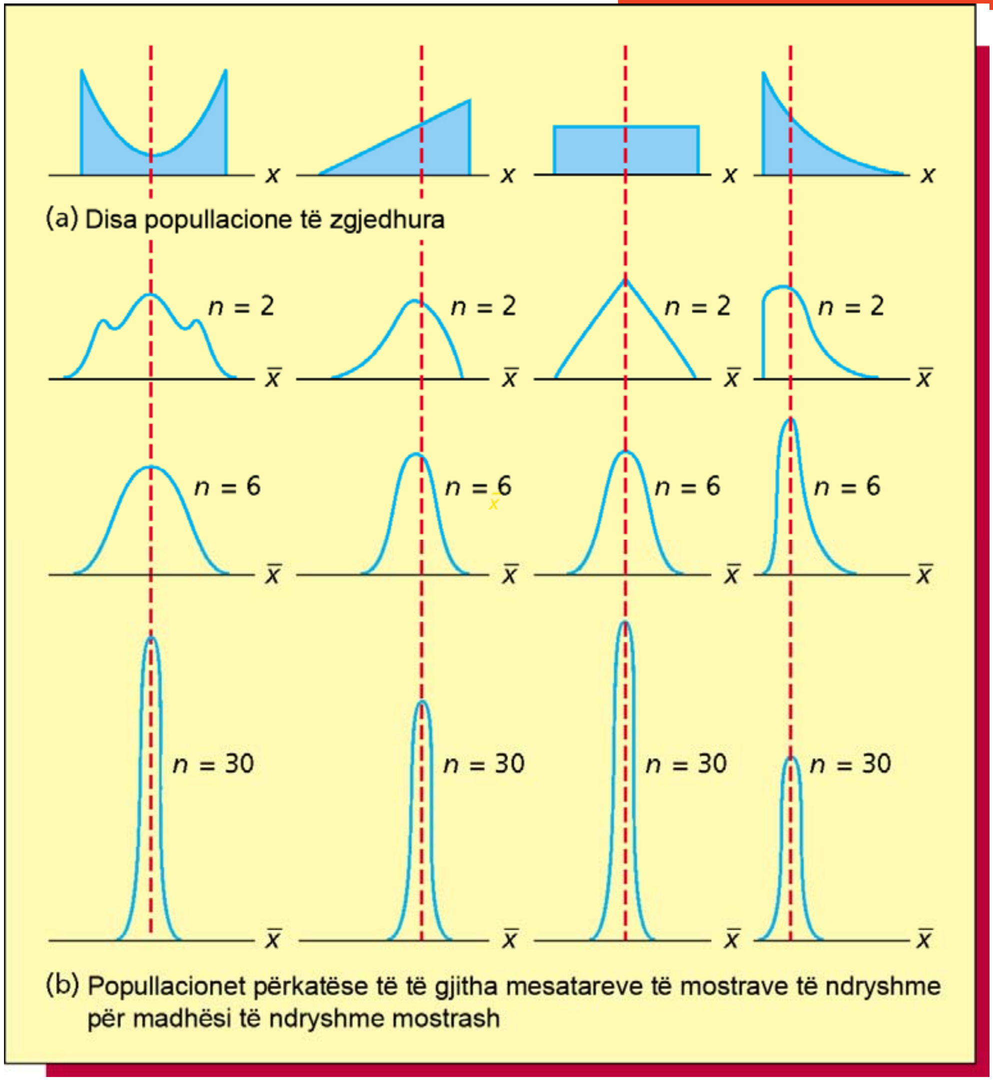


(b) Histogram i 1000 mesatareve mostrash mbështetur në mostra të madhësisë 5



(c) Histogram i 1000 mesatareve mostrash mbështetur në mostra të madhësisë 30

Shembull: Efekti i madhësisë së mostrës



Sa më e madhe madhësia e mostrës, aq më afërsisht normale është shpërndarja e popullimit të të gjitha mesatareve të mostrave të mundshme.

Poashtu, me rritjen e madhësisë së mostrës, përhapja e shpërndarjes së zgjedhjeve zvogëlohet.

Shpërndarja e zgjedhjeve e proporcionit të mostrës

Shpërndarja e probabilitetit të proporcioneve të të gjitha mostrave të mundshme është *shpërndarje e zgjedhjeve e proporcionit të mostrës* $\hat{p}$

Në qoftë se një mostër e rastësishme e madhësisë n merret nga popullimi, atëherë shpërndarja e zgjedhjeve e $\hat{p}$ është

- përafërsisht normale, në qoftë se n është i madh
- ka mesataren $\mu_{\hat{p}} = p$
- ka devijimin standard $\sigma_{\hat{p}} = \sqrt{\dfrac{p(1-p)}{n}}$

ku p është proporcioni i popullimit dhe $\hat{p}$ është proporcioni i një mostre të zgjedhur