მენეჯერული სტატისტიკა

თავი 7
შერჩევა და შერჩევის განაწილება

თავის მიზნები

თავის შესწავლის შემდეგ თქვენ შეძლებთ:

- აღწეროთ მარტივი შემთხვევითი შერჩევა და ახსნათ რატომაა შერჩევა მნიშვნელოვანი
- ახსნათ განსხვავება აღწერით და დასკვნით სტატისტიკებს შორის
- ახსნათ შერჩევითი განაწილების ცნება
- განსაზღვროთ საშუალო და სტანდარტული გადახრა შერჩევის საშუალოს შერჩევითი $\bar{X}$ განაწილებისთვის
- ახსნათ ცენტრალური ზღვრის თეორემა და მისი მნიშვნელოვნება
- განსაზღვროთ საშუალო და სტანდარტული გადახრა შერჩევითი პროპორციის p შერჩევითი განაწილებისთვის
- ახსნათ შერჩევის ვარიაციის შერჩევითი განაწილება

ბიზნეს სტატისტიკის ინსტრუმენტები

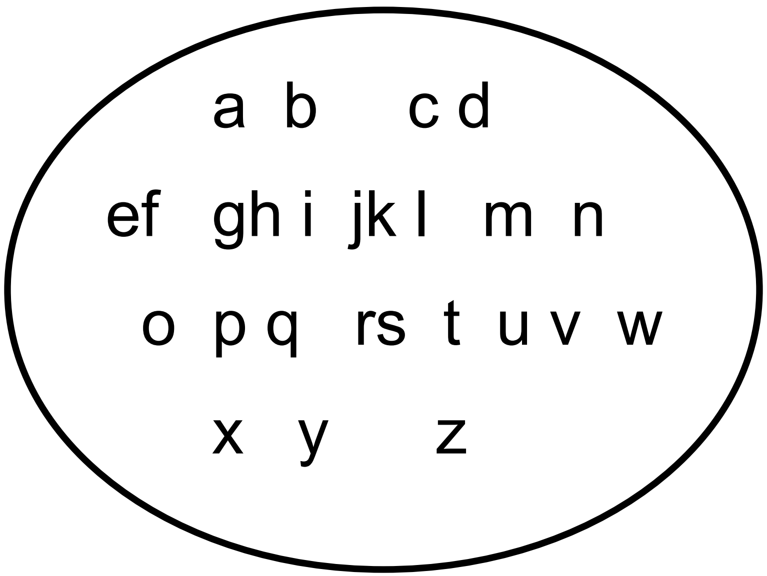
- **აღწერითი სტატისტიკა**
  - მონაცემების შეგროვება, წარდგენა და აღწერა
- **დასკვნითი სტატისტიკა**
  - დასკვნების გაკეთება და/ან შერჩევის მონაცემებზე დაყრდნობით პოპულაციის შესახებ გადაწყვეტილების მიღება

შერჩევა და პოპულაცია

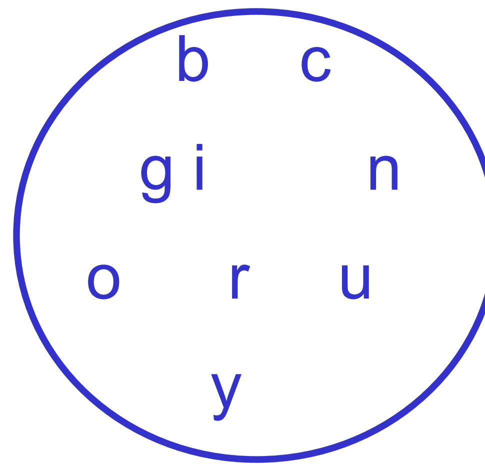
- **პოპულაცია (Population)** არის შესასწავლი ობიექტების (ერთეულების) ყველა შესაძლო მნიშვნელობათა ერთობლიობა
  - მაგალითები: ყველა შესაძლო ამომრჩეველი, რომელიც მივა არჩევნებზე, ყველა წარმოებული ნაწილი, ნოემბრის თვის მთლიანი გაყიდვები.
- **შერჩევა (Sample)** არის პოპულაციის გარკვეული ნაწილი (ქვესიმრავლე)
  - მაგალითები: შემთხვევით შერჩეული 1000 ამომრჩეველი, ტესტირებისთვის შერჩეული რამოდენიმე ნაწილი, აუდიტისთვის შერჩეული რამოდენიმე ქვითარი.

პოპულაცია vs. შერჩევა

პოპულაცია



შერჩევა

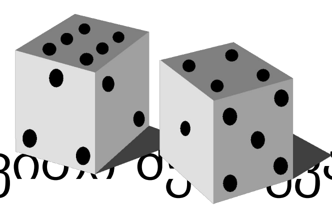


რატომ შერჩევა?

- საჭიროებს ნაკლებ დროს ვიდრე პოპულაციის აღწერა
- საჭიროებს ნაკლებ ხარჯებს ვიდრე პოპულაციის აღწერა
- შესაძლებელია მიღებულ იქნას მაღალი სიზუსტის სტატისტიკური შედეგები შერჩევაზე დაყრდნობით

მარტივი შემთხვევითი შერჩევა

- პოპულაციის ყველა ობიექტს აქვს შერჩევაში მოხვედრის თანაბარი შანსი
- ობიექტების შერჩევა ხდება ერთმანეთისგან დამოუკიდებლად
- შერჩევს განხორციელება შეიძლება შემთხვევითი ცხრილების ან კომპიუტერული პროგრამის გამოყენებით

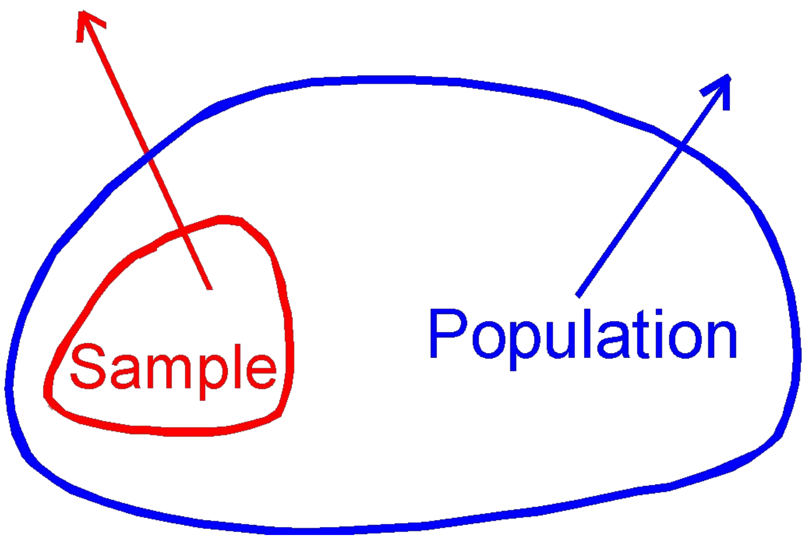


- მარტივი შემთხვევითი შერჩევა საუკეთესოა სხვა შერჩევის მეთოდებთან შედარებით

დასკვნითი სტატისტიკა

- შერჩევის შედეგებზე დაყრდნობით, აკეთებს განაცხადს პოპულაციის შესახებ შერჩევის სტატისტიკა პოპულაციის პარამეტრი

  (ცნობილი) დასკვნა (უცნობი, მაგრამ შესაძლოა შეფასებულ იქნას შერჩევაზე დაყრდნობით)



დასკვნითი სტატისტიკა

დასკვნის გაკეთება და/ან შერჩევის მონაცემებზე დაყრდნობით პოპულაციის შესახებ გადაწყვეტილების მიღება

- **შეფასება**
  - მაგ, შერჩევის საშუალო წონაზე დაყრდნობით, პოპულაციის საშუალო წონის განსაზღვრა.
- **ჰიპოთეზის შემოწმება**
  - მაგ, მტკიცების შემოწმება, რომ მოსახლეობის საშუალო სიმაღლე არის 170 სმ.

შერჩევის განაწილება

- **შერჩევის განაწილება** არის პოპულაციიდან მიღებული გარკვეული ზომის შერჩევის სტატისტიკის ყველა შესაძლო მნიშვნელობის განაწილება.

თავის მიმოხილვა

შერჩევითი განაწილება

- შერჩევის საშუალოს შერჩევითი განაწილება
- შერჩევის პროპორციის შერჩევითი განაწილება
- შერჩევის ვარიაციის შერჩევითი განაწილება

შერჩევის საშუალოს შერჩევითი განაწილება

- შერჩევითი განაწილება
  - შერჩევის საშუალოს შერჩევითი განაწილება
  - შერჩევის პროპორციის შერჩევითი განაწილება
  - შერჩევის ვარიაციის შერჩევითი განაწილება

შერჩევითი განაწილების განვითარება

- **დავუშვათ მოცემული გვაქვს პოპულაცია…**
- პოპულაციის ზომა N=4
- შემთხვევითი ცვლადი, X, არის ინდივიდების ასაკი
- X-ის მნიშვნელობები: 18, 20, 22, 24 (წელი)

შერჩევითი განაწილების განვითარება

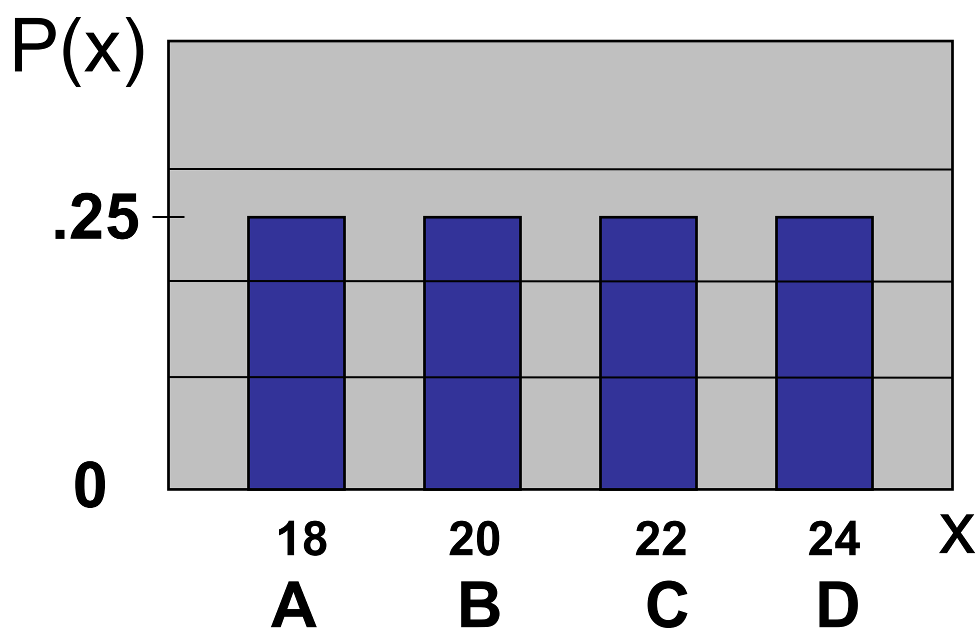
(გაგრძელება)

პოპულაციის განაწილების შემაჯამებელი მაჩვენებლები:

$$\mu = \frac{\sum X_i}{N}$$

$$= \frac{18+20+22+24}{4} = 21$$

$$\sigma = \sqrt{\frac{\sum (X_i - \mu)^2}{N}} = 2.236$$



თანაბარი განაწილება

შერჩევითი განაწილების განვითარება

(გაგრძელება)

ახლა განვიხილოთ შერჩევა, რომლის ზომაა n = 2

1st Obs	2nd Observation 18	20	22	24
18	18,18	18,20	18,22	18,24
20	20,18	20,20	20,22	20,24
22	22,18	22,20	22,22	22,24
24	24,18	24,20	24,22	24,24

16 შესაძლო შერჩევა (შერჩევა ჩანაცვლებით)

16 შერჩევის საშუალო

1st Obs	2nd Observation 18	20	22	24
18	18	19	20	21
20	19	20	21	22
22	20	21	22	23
24	21	22	23	24

შერჩევითი განაწილების განვითარება

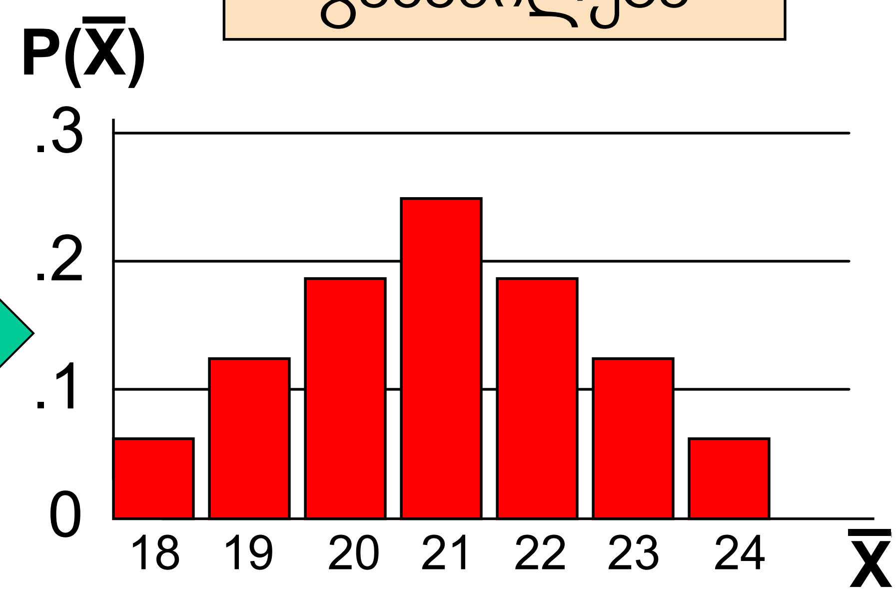
(გაგრძელება)

შერჩევითი საშუალოს შერჩევითი განაწილება

16 შერჩევითი საშუალო

1st Obs	2nd Observation			
	18	20	22	24
18	18	19	20	21
20	19	20	21	22
22	20	21	22	23
24	21	22	23	24

შერჩევითი საშულოს განაწილება



(არა ერთგვაროვანი)

შერჩევითი განაწილების განვითარება
(გაგრძელება)

შერჩევითი განაწილების შემაჯამებელი საზომები:

$$E(\bar{X}) = \frac{\sum \bar{X}_i}{N} = \frac{18 + 19 + 21 + \cdots + 24}{16} = 21 = \mu$$

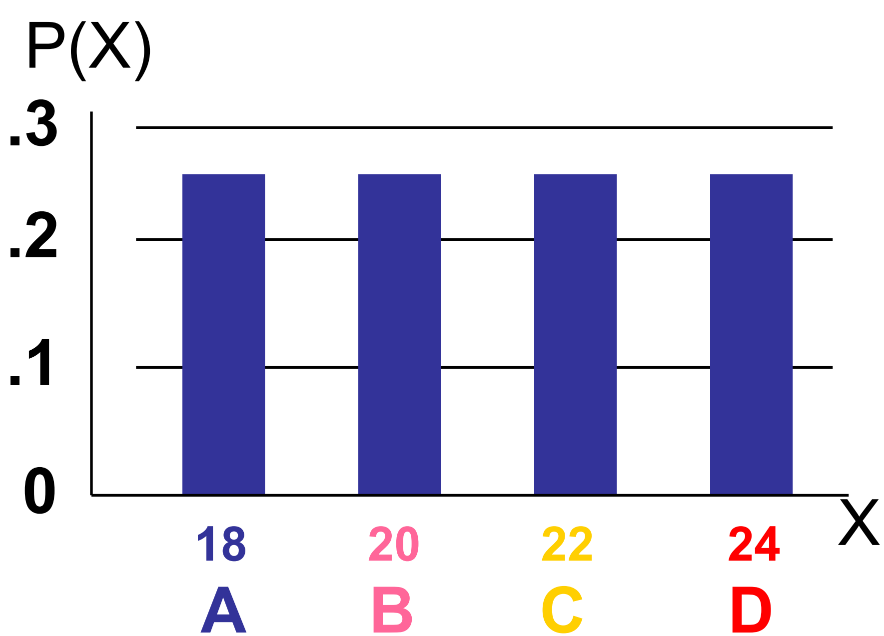
$$\sigma_{\bar{X}} = \sqrt{\frac{\sum (\bar{X}_i - \mu)^2}{N}}$$

$$= \sqrt{\frac{(18 - 21)^2 + (19 - 21)^2 + \cdots + (24 - 21)^2}{16}} = 1.58$$

პოპულაციის შედარება თავის შერჩევის განაწილებასთან

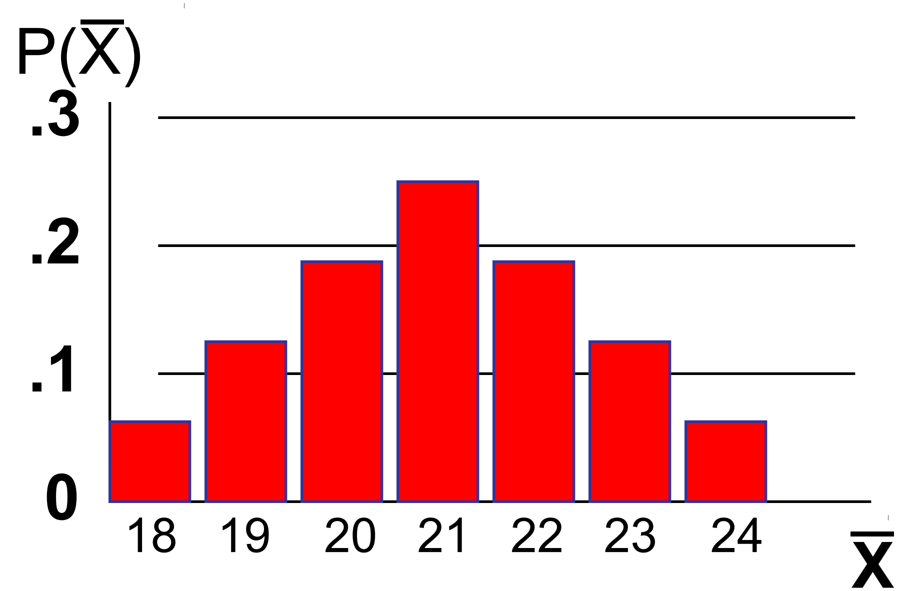
პოპულაცია N = 4

$\mu = 21 \qquad \sigma = 2.236$



შერჩევის საშუალოს განაწილება
n = 2

$\mu_{\bar{X}} = 21 \qquad \sigma_{\bar{X}} = 1.58$



შერჩევის საშუალოს მოსალოდნელი მნიშვნელობა

- დავუშვათ $X_1, X_2, \ldots X_n$ არის პოპულაციიდან მიღებული შემთხვევითი შერჩევა
- შერჩევის საშუალო განისაზღვრება, როგორც:

$$\boxed{\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i}$$

შერჩევის სტანდარტული შეცდომა

- ერთიდაიგივე პოპულაციის, ერთი და იგივე ზომის სხვადასხვა შერჩევა მოგვცემს სხვადასხვა შედეგებს:
- საშუალოს ცვალებადობის საზომი შერჩევიდან შერჩევამდე განსაზღვრულია შერჩევის სტანდარტული შეცდომით:

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

- ყურადღება მიაქციეთ, რომ სტანდარტული შეცდომა მცირდება, როდესაც შერჩევის ზომა იზრდება

თუ პოპულაცია ნორმალურია

- თუ მოპულაცია ნორმალურია საშუალო μ-თი და სტანდარტული გადახრა σ-ით, შერჩევითი განაწილება იქნება ასევე ნორმაულრად განაწილებული $\bar{X}$

$$\mu_{\bar{X}} \text{ და } = \mu$$

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

Z-მნიშვნელობა საშუალოს შერჩევითი განაწილებისათვის

- Z-მნიშვნელობა $\bar{X}$-ის შერჩევითი განაწილებისათვის არის:

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

სადაც: $\bar{X}$ = შერჩევის საშუალო

μ = პოპულაციის საშუალო

σ = პოპულაციის სტანდარტული გადახრა

n = შერჩევის ზომა

პოპულაციის სასრული კორელაცია

- გამოიყენეთ პოპულაციის სასრული კორელაცია თუ:
  - პოპულაციის წევრი არ შეიძლება ერთზე მეტჯერ შედიოდეს შერჩევაში (შერჩევა დაგანაცვლებადობის გარეშე), და შერჩევა დიდი ზომისაა პოპულაციიდან გამომდინარე

    (n არის N-ის 5%-ზე მეტი)
- მაშინ

$$\mathrm{Var}(\overline{X}) = \frac{\sigma^2}{n}\frac{N-n}{N-1}$$

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$

პოპულაციის სასრული კორელაცია

- თუ შერჩევის ზომა n არ არის პატარა პოპულაციის ზომა N-თან შედარებით, მაშინ გამოიყენეთ

$$\boxed{Z = \frac{(\bar{X} - \mu)}{\dfrac{\sigma}{\sqrt{n}}\sqrt{\dfrac{N-n}{N-1}}}}$$

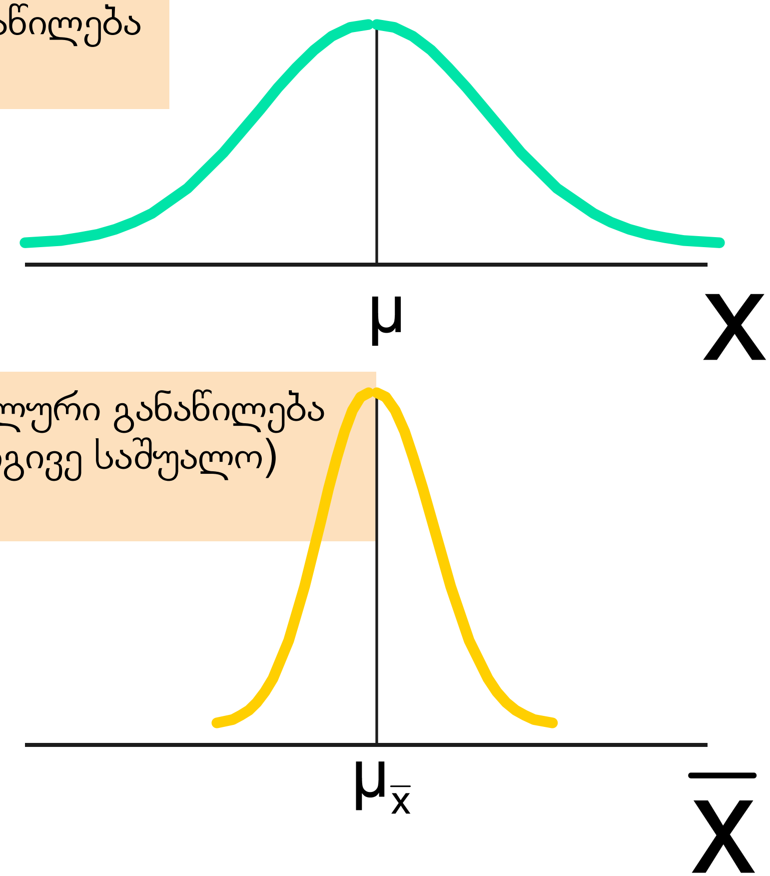
შერჩევითი განაწილების თვისებები

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

(ანუ $\bar{X}$ არის მიუკერძოებელი)

პოპულაციის ნორმალური განაწილება

შერჩევის ნორმალური განაწილება (აქვთ ერთი და იგივე საშუალო)

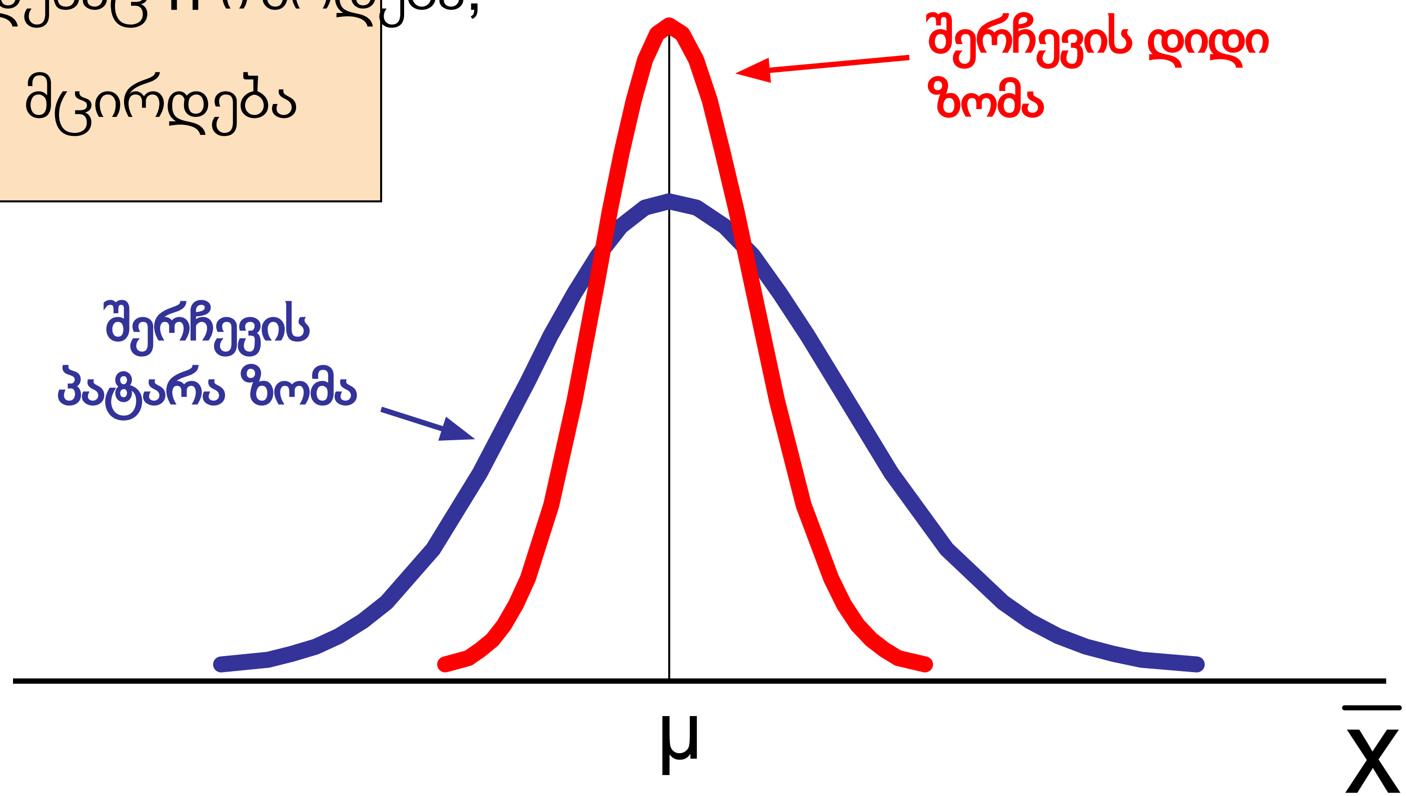


შერჩევითი განაწილების თვისებები
(გაგრძელება)

- შერჩევისთვის გადანაცვლებადობის შემთხვევაში:

როდესაც n იზრდება,

$\sigma_{\bar{x}}$ მცირდება



თუ პოპულაცია **არაა** ნორმალურად განაწილებულია

- ჩვენ შეგვიძლია გამოვიყენოთ ცენტრალური ზღვრის თეორემა:
  - მაშინაც კი თუ პოპულაცია არაა ნორმალური,
  - …პოპულაციიდნ მიღებული შერჩევის საშუალო იქნება მიახლოებით ნორმალური, თუ შერჩევის ზომა იქნება საკმარისად დიდი.

შერჩევითი განაწილების თვისებები:

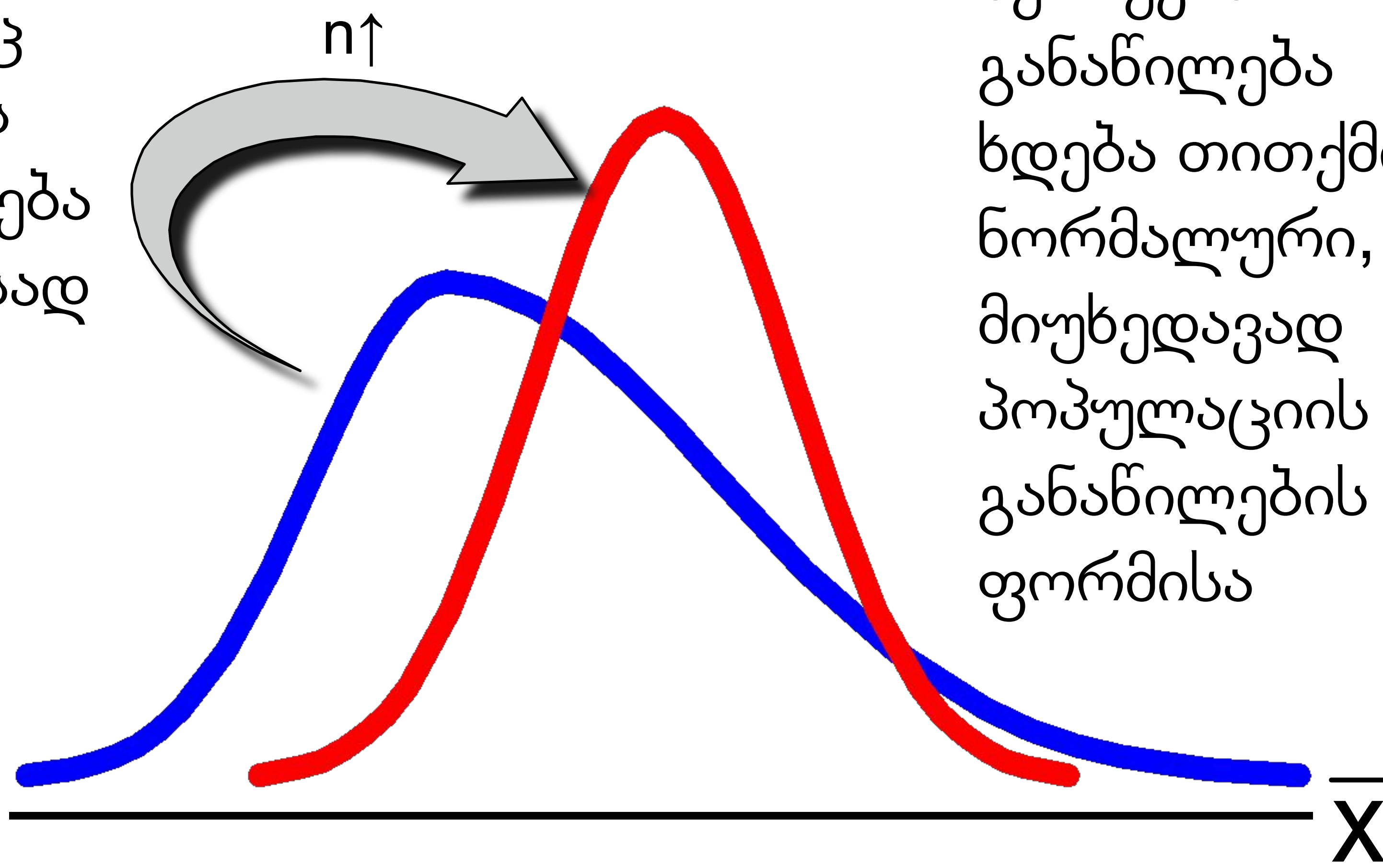
$$\boxed{\mu_{\bar{x}} = \mu}$$

და

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

ცენტრალური ზღვრის თეორემა

როდესაც შერჩევის ზომა ხდება საკმარისად დიდი...



შერჩევითი განაწილება ხდება თითქმის ნორმალური, მიუხედავად პოპულაციის განაწილების ფორმისა

თუ პოპულაცია **არაა** ნორმალურად განაწილებული

(გაგრძელება)

შერჩევითი განაწილების თვისებები:

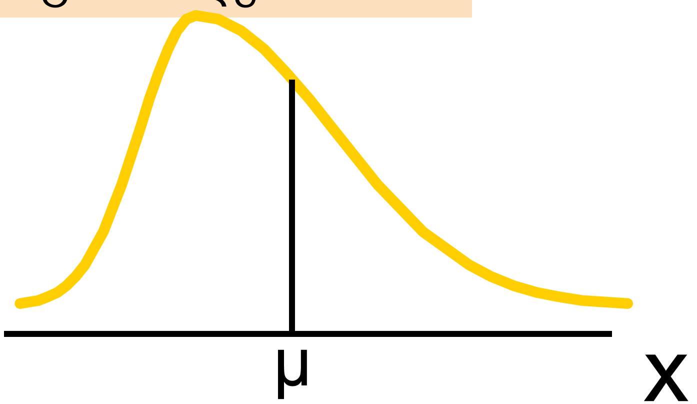
ცენტრალური ტენდენცია

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

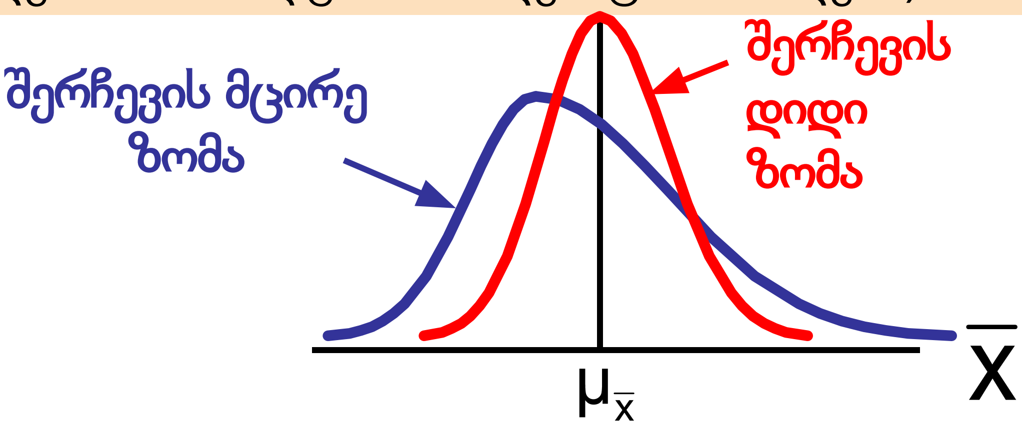
ვარიაცია

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

პოპულაციის განაწილება



შერჩევითი განაწილება
(ხდება ნორმალური, როდესაც n იზრდება)



რამდენია საკმაოდ დიდი?

- უმრავლესობა განაწილებისათვის, $n > 25$ მოგვცემს შეერჩევით განაწილებას, რომელიც ახლოსაა ნორმალურთან
- ნორმალურად განაწილებული პოპულაციისათვის, შერჩევითი განაწილების საშუალო ყოველთვის ნორმალურია.

მაგალითი

- დავუშვათ პოპულაციის საშუალოა $\mu = 8$ და სტანდარტული გადახრაა $\sigma = 3$. დავუშვათ შემთხვევითი შერჩევის ზომაა $n = 36$.

- რა არის ალბათობა იმისა, რომ შერჩევის საშუალო მოთავსებული იქნება 7.8-სა და 8.2-ს შორის?

მაგალითი

(გაგრძელება)

ამოხსნა:

- მაშინაც კი თუ პოპულაცია ნომრალურად არაა განაწილებული, ცენტრალური ზღვრის თეორეა შეიძლება იქნას გამოყენებული(n > 25)
- … ანუ $\bar{X}$ -ის შერჩევითი განაწილება იქნება მიახლოებით ნორმალური
- … საშუალო $\mu_{\bar{x}}$ = 8
- …და სტანდარტული გადახრა

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{3}{\sqrt{36}} = 0.5$$

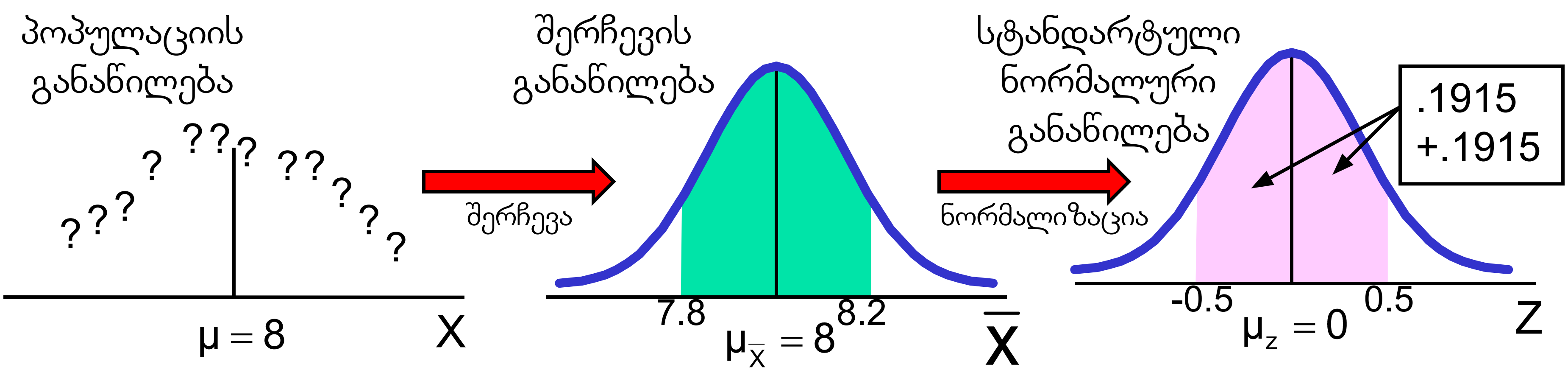
მაგალითი

(გაგრძელება)

ამოხსნა (გაგრძელება):

$$P(7.8 < \mu_{\bar{X}} < 8.2) = P\left(\frac{7.8-8}{3/\sqrt{36}} < \frac{\mu_{\bar{X}}-\mu}{\sigma/\sqrt{n}} < \frac{8.2-8}{3/\sqrt{36}}\right)$$

$$= P(-0.5 < Z < 0.5) = \boxed{0.3830}$$



მისაღები ინტერვალი

- მიზანი: პოპულაციის საშუალოსა და ვაციაციაზე დაყდრნობით, იმ დიაპაზონის განსაზღვრა, რომელშიც მაღალია შერჩევის საშუალოს მოხვედრა $\bar{}$
  - ცენტრალური ზღვრის თეორემის თანახმად, ჩვენ ვიცით, რომ $\bar{X}$-ის განაწილება მიახლოებით ნორმალურია, თუ n არის საკმარისად დიდი, რომ საშუალოა μ და სტანდარტული გადახრა $\sigma_{\bar{X}}$
  - დავუშვათ $z_{\alpha/2}$ არის z-მნიშვნელობა, რომელიც ნორმალური განაწილების ზედა კუდში ტოვებს ფართობს $\alpha/2$ (ე.ი. ინტერვალი -$z_{\alpha/2}$-დან $z_{\alpha/2}$-მდე ახლოსაა ალბათობა $1 - \alpha$-სთან)
  - მაშინ

$$\boxed{\mu \pm z_{\alpha/2}\sigma_{\bar{X}}}$$

არის ინტერვალი, რომელიც მოიცავს $\bar{X}$-ს ალბათობით $1 - \alpha$