სტატისტიკური ფუნქციები მონაცემთა ანალიზისათვის

რიცხვით მონაცემთა სხვადასხვა სტატისტიკური მახასიათებლების განსაზღვრა ნებისმიერ მონაცემთა ანალიზისას მნიშვნელოვანი ნაწილია. MATLAB-ში მრავლადაა მონაცემთა სტატისტიკური დამუშავებისათვის საჭირო ფუნქციები.

დასაწყისისთვის განვიხილოთ რამდენიმე მარტივი ფუნქცია.

მაქსიმუმი და მინიმუმი

max (x)	თუ x ვექტორია, ეს ფუნქცია გვაძლევს მის ელემენტებს შორის უდიდეს მნიშვნელობას, თუ მატრიცაა, მაშინ – სტრიქონ- ვექტორს, რომლის ელემენტებიც წარმოადგენენ თითოეული სვეტის უდიდეს მნიშვნელობებს
[y,I] = max (x)	y წარმოადგენს x მატრიცის ყოველი სვეტის უდიდესი მნიშვნელობებისგან შემდგარ ვექტორს, ხოლო I ვექტორში მოცემულია თითოეულ სვეტში უდიდესი მნიშვნელობის შესაბამისი სტრიქონი ანუ ინდექსი. თუ რომელიმე სვეტში ორი ერთმანეთის ტოლი მაქსიმუმია, მოიცემა პირველი მნიშვნელობის ინდექსი
min (x)	თუ x ვექტორია, ეს ფუნქცია გვაძლევს მის ელემენტებს შორის უმცირეს მნიშვნელობას, თუ მატრიცაა, მაშინ – სტრიქონ- ვექტორს, რომლის ელემენტებიც წარმოადგენენ თითოეული სვეტის უმცირეს მნიშვნელობებს
[y, I] = min (x)	y წარმოადგენს x მატრიცის ყოველი სვეტის უმცირესი მნიშვნელობებისგან შემდგარ ვექტორს, ხოლო I ვექტორში მოცემულია თითოეულ სვეტში უმცირესი მნიშვნელობის შესაბამისი სტრიქონი ანუ ინდექსი. თუ რომელიმე სვეტში ორი ერთმანეთის ტოლი მინიმუმია, მოიცემა პირველი მნიშვნელობის ინდექსი

საშუალო არითმეტიკული და მედიანა

x ვექტორის μ საშუალო არითმეტიკული გამოითვლება შემდეგნაირად:

$$\mu = \frac{\sum_{k=1}^{N} x_k}{N}$$

სადაც

$$\sum_{k=1}^{N} x_k = x_1 + x_2 + x_3 + \cdots + x_n$$

x ვექტორის მედიანა არის ზრდის მიხედვით დალაგებული მისი მნიშვნელობების შუა მონაცემი. თუ ზრდის მიხედვით დალაგებული x ვექტორი შეიცავს ელემენტთა კენტ რაოდენობას N-ს, მედიანა იქნება იმ ელემენტის ტოლი, რომლის ინდექსია **ceil**(N/2) (თუ N=5, მედიანა იქნება მესამე ელემენტის ტოლი), თუ N ლუწია, მაშინ მედიანა იქნება შუა ორი ელემენტის საშუალო არითმეტიკული მნიშვნელობის ტოლი. მაგალითად, თუ x=[2 3 5 7 8] მედიანა იქნება 5, ხოლო თუ x=[2 3 5 7 8 10] მედიანა 6-ის ტოლია.

MATLAB-ში საშუალო არითმეტიკული და მედიანა შემდეგი ფუნქციებით გამოითვლება: **mean** და **median**.

mean (x)	თუ x ვექტორია, ეს ფუნქცია გამოითვლის x ელემენტების საშუალო მნიშვნელობას. თუ x მატრიცაა, ეს ფუნქცია გვაძლევს ვექტორს, რომლის ელემენტებია შესაბამისი სვეტების ელემენტთა საშუალო მნიშვნელობები
median (x)	თუ x ვექტორია, ეს ფუნქცია გამოითვლის x ელემენტების მედიანას. თუ x მატრიცაა, ეს ფუნქცია გვაძლევს ვექტორს, რომლის ელემენტებია შესაბამისი სვეტების ელემენტთა მედიანები

ჯამი და ნამრავლი

MATLAB აქვს ფუნქციები, რომლებიც გამოითვლის მატრიცის ელემენტების ჩვეულებრივსა და კუმულაციურ ჯამსა და ნამრავლს.

sum (x)	თუ x ვექტორია, ეს ფუნქცია გამოითვლის მისი ელემენტების ჯამს. თუ x მატრიცაა, ის გვაძლევს ვექტორს, რომლის ელემენტები წარმოადგენს x ვექტორის თითოეული სვეტის ელემენტთა ჯამს
prod (x)	თუ x ვექტორია, ეს ფუნქცია გამოითვლის მისი ელემენტების ნამრავლს. თუ x მატრიცაა, ის გვაძლევს ვექტორს, რომლის ელემენტები წარმოადგენს x ვექტორის თითოეული სვეტის ელემენტთა ნამრავლს
cumsum (x)	თუ x ვექტორია, ეს ფუნქცია გვაძლევს x ვექტორის კუმულაციურ ჯამს, ანუ იგივე სიგრძის ვექტორს, რომლის k-ური ელემენტი წარმოადგენს x-ის პირველი k ელემენტის ჯამს, თუ x მატრიცაა, ეს ფუნქცია იძლევა იგივე ზომის მატრიცას, რომლის თითოეული სვეტი წაროადგენს x მატრიცის შესაბამისი სვეტის კუმულაციურ ჯამს
cumprod (x)	თუ x ვექტორია, ეს ფუნქცია გვაძლევს x ვექტორის კუმულაციურ ნამრავლს, ანუ იგივე სიგრძის ვექტორს, რომლის k-ური ელემენტი წარმოადგენს x-ის პირველი k ელემენტის ნამრავლს, თუ x მატრიცაა, ეს ფუნქცია იძლევა იგივე ზომის მატრიცას, რომლის თითოეული სვეტი წაროადგენს x მატრიცის შესაბამისი სვეტის კუმულაციურ ნამრავლს
sort (x)	თუ x ვექტორია, დაალაგებს ელემენტებს ზრდის მიხედვით. თუ x მატრიცაა, ეს ფუნქცია თითოეულ სვეტს დაალაგებს ზრდის მიხედვით
[y, i]=sort(x)	ასეთი ფორმით ეს ფუნქცია y ვექტორის სახით მოგვცემს ზრდის მიხდვით დალაგებულ x ვექტორის მნიშვნელობებს, ხოლო i ვექტორში მოცემული იქნება შესაბამისი ელემენტების ინდექსები საწყისს x ვექტორში.

savarjiSo

მოცემულია მატრიცები:

w = [0 3 –2 7];
x = [3 –1 5 7];

$$y = \begin{bmatrix} 1 & 3 & 7 \\ 2 & 8 & 4 \\ 6 & -1 & -2 \end{bmatrix}$$

გამოითვალეთ შემდეგი სიდიდეები და შეამოწმეთ MATLAB-ის საშუალებით

1. max (w)
2. min (y)
3. min (w,x)
4. [z, i] = max (y)
5. mean (y)
6. median (w)
7. cumprod (y)
8. sum (x)
9. sort (2*w + x)
10. sort (y)

დისპერსია და სტანდარტული გადახრა

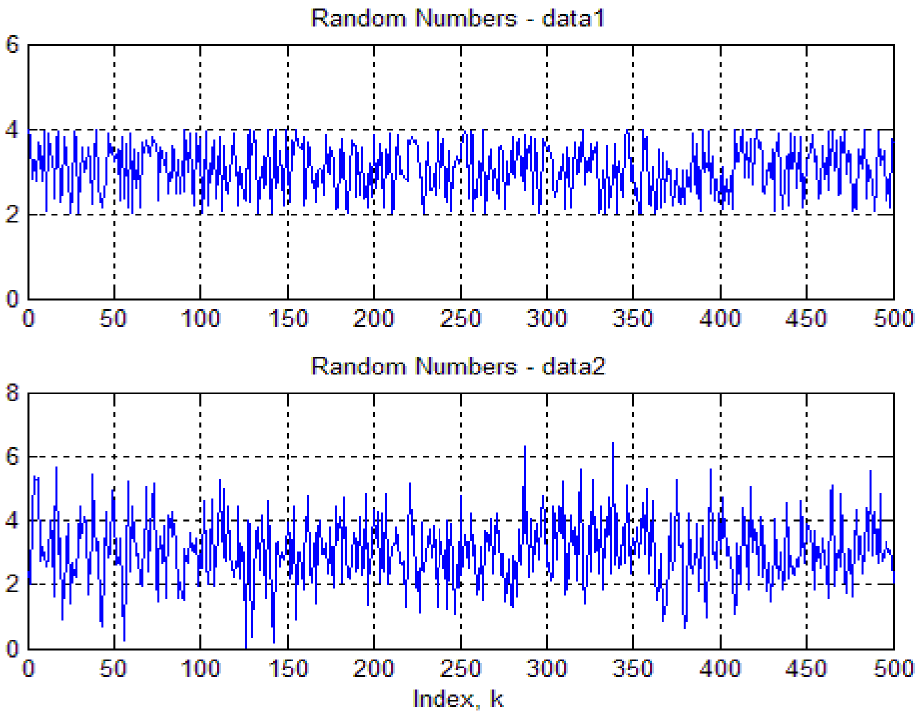
რიცხვით მონაცემთა ერთ-ერთი მნიშვნელოვანი სტატისტიკური მახასიათებელია დისპერსია. ვიდრე განვიხილავდეთ დისპერსიის მათემატიკურ შინაარსს, განვიხილოთ ასეთი მაგალითი: ვთქვათ გვაქვს მონაცემთა ორი ჯგუფი data1 და data2 (ნახ. 1). როგორც ვხედავთ ორივე მონაცემთა სიმრავლის საშუალო არითმეტიკული მნიშვნელობაა 3.0, მიუხედავად იმისა, რომ ამ მონაცემებს განსხვავებული მახასიათებლები გააჩნიათ. data2 მონაცემები უფრო მკვეთრად გადაიხრება საშუალო მნიშვნელობიდან, ვიდრე data1–sa.

x ვექტორის მათემატიკურად დისპერსია σ^2 გამოითვლება შემდეგი ფორმულით:

$$\sigma^2 = \frac{\sum_{k=1}^{N}(x_k - \mu)^2}{N-1}$$

თუ კარგად დავაკვირდებით გამოსახულება საკმაოდ მარტივია. ის წარმოადგენს x ვექტორის ყოველი ელემენტის საშუალო არითმეტიკულიდან გადახრის მნიშვნელობათა კვადრატების ჯამს გაყოფილს x ვექტორის მონაცემთა N რაოდენობაზე ერთით ნაკლებ სიდიდეზე და ამდენად წარმოადგენს მონაცემთა საშუალო სიდიდიდან გადახრის ზომას. სწორედ ამიტომ კვადრატული ფესვის მნიშვნელობას დისპერსიიდან, უწოდებენ საშუალო კვადრატულ ან სტანდარტულ გადახრას:

$$\sigma = \sqrt{\sigma^2}$$



ნახ. 1. data1 და data2 მონაცემთა გრაფიკები

MATLAB-ში სტანდარტული გადახრა გამოითვლება **std** ფუნქციით. დისპერსიის გამოსათვლელად კი საკმარისია მისი კვადრატში აყვანა.

std (x) თუ x ვექტორია, ეს ფუნქცია გამოითვლის მის მნიშვნელობათა სტანდარტულ გადახრას. თუ x მატრიცაა, მოგვცემს სტრიქონ ვექტორს, რომლის ყოველი ელემენტი x-ის შესაბამისი სვეტის სტანდარტული გადახრას წარმოადგენს.

სავარჯიშო

განსაზღვრეთ შემდეგ ფუნქციათა მნიშვნელობები, თუ მოცემულია მატრიცები:

w = [0 3 –2 7];

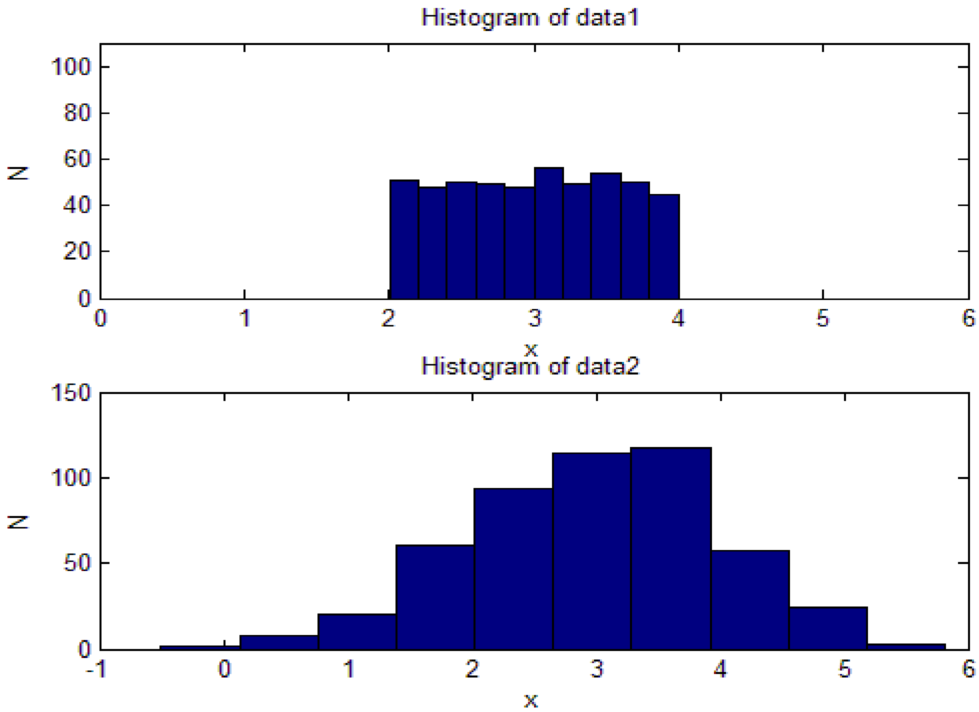
x = [3 –1 5 7];

$$y = \begin{bmatrix} 1 & 3 & 7 \\ 2 & 8 & 4 \\ 6 & -1 & -2 \end{bmatrix}$$

1. std (w)
2. std (x)^2
3. std (y(:2))
4. std (y)

ჰისტოგრამა

ჰისტოგრამა სპეციალური გრაფიკული საშუალებაა სტატისტიკური მონაცემების ვიზუალიზაციისათვის. ის გვიჩვენებს მონაცემთა განაწილების სურათს. MATLAB–ში **hist** ფუნქცია გამოითვლის იმ სიდიდეთა რაოდენობას, რომლებიც ხვდებიან მონაცემთა უმცირეს და უდიდეს მნიშვნელობას შორის 10 თანაბარ ნაწილად დაყოფილ შუალედში. მაგალითად, თუ ავაგებთ data1 და data2 მონაცემების ჰისტოგრამას, მივიღებთ შემდეგ გრაფიკს (ნახ. 2).

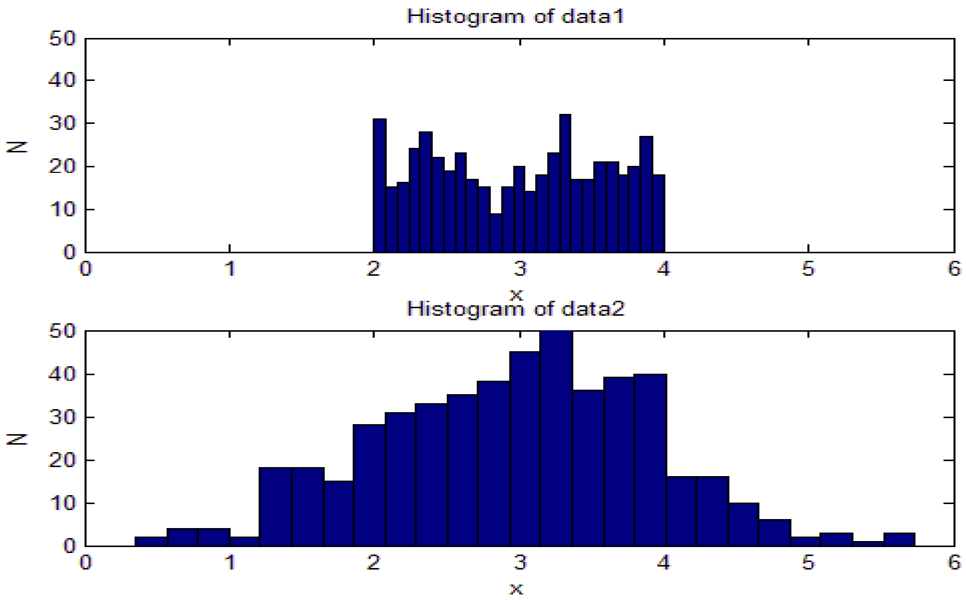


ნახ. 2 ჰისტოგრამა 10 სვეტით

ჰისტოგრამა გვაძლევს არა მარტო მნიშვნელობათა არეს, არამედ ინფორმაციას იმის თაობაზე, თუ როგორ არიან მონაცემები განაწილებული. მაგალითად data1 მონაცემები უჩვენებს რომ მონაცემები თითქმის თანაბრადაა განაწილებული უმცირეს (2) და უდიდეს (4) სიდიდეს შორის (ასეთი ტიპის განაწილებას თანაბარი ეწოდება). data2 მონაცემები არ არის განაწილებული თანაბრად, მონაცემთა უმრავლესობა კონცენტრირებულია საშუალო მნიშვნელობის მახლობლად (ასეთი ტიპის განაწილებას გაუსისებური, ან ნორმალური განაწილება ეწოდება).

MATLAB ბრძანება ჰისტოგრამის ასაგებად ამგვარია: **hist (x)** სადაც x ვექტორია იმ მონაცემებით, რომლის ჰისტოგრამაც უნდა ავაგოთ. ეს ბრძანება პირველ რიგში დაალაგებს x ვექტორის ელემენტებს ზრდის მიხედვით, შემდეგ ინტერვალს x-ის მინიმალურ და მაქსიმალურ მნიშვნელობებს შორის გაყოფს 10 ტოლ ნაწილად და დაითვლის თითოეულ ნაწილში რამდენი მონაცემი მოხვდა.

თუ მოცემული გვაქვს მონაცემები მატრიცის სახით და გვსურს ავაგოთ მისი მეორე სვეტის მნიშვნელობათა ჰისტოგრამა, MATLAB-ს უნდა მივცეთ შემდეგი სახის ბრძანება: hist (data (: , 2)). **hist** ბრძანება საშუალებას იძლევა შევარჩიოთ ჰისტოგრამის სვეტების რაოდენობაც. თუ გვსურს გავზარდოთ ჰისტოგრამის გარჩევა ისე, რომ იგი შეიცავდეს 25 სვეტს ნაცვლად ათისა, მივმართავთ შემდეგ ბრძანებას: hist (x, 25). შესაბამისი ჰისტოგრამა ნაჩვენებია ნახ. 3-ზე.



ნახ. 3. ჰისტოგრამა 25 სვეტით

ინფორმაცია, რომლის მიხედვითაც ჰისტოგრამა იგება, შეიძლება [n, x] ვექტორთა წყვილის სახით ჩაიწეროს:

[n, x] = hist (data1);

[n, x] = hist (data1,25);

ამ ბრძანებების შესრულებისას ჰისტოგრამის აგების მაგივრად იქმნება ორი ვექტორი – n და x. n ვექტორი შეიცავს თითოეულ სვეტში მოხვედრილი მონაცემების რაოდენობას, x ვექტორი კი - თითოეული სვეტის შუაწერტილის მნიშვნელობას. მეორე ბრძანება ისეთივე ნაირია რაც პირველი, მაგრამ 25 სვეტისათვის.

შემთხვევითი რიცხვები

შემთხვევითი რიცხვები არ გამოითვლება ფორმულით. მათი მნიშვნელობა შემთხვევაზეა დამოკიდებული. წინასწარ შეუძლებელია იმის განსაზღვრა, თუ რა მნიშვნელობას მიიღებს შემთხვევითი სიდიდე. მის შესასწავლად პირველ რიგში უნდა ვიცოდეთ ის რიცხვითი მნიშვნელობები, რომელთა მიღებაც მას შეუძლია და რა სიხშირით ღებულობს იგი ამა თუ იმ შესაძლო მნიშვნელობას. ამას კი განსაზღვრავს მისი განაწილების კანონი, რომელიც მოიცემა განაწილების ფუნქციით.

დასაწყისში გავეცნობით თანაბრი და ნორმალური განაწილების ფუნქციის მქონე შემთხვევით რიცხვებს.

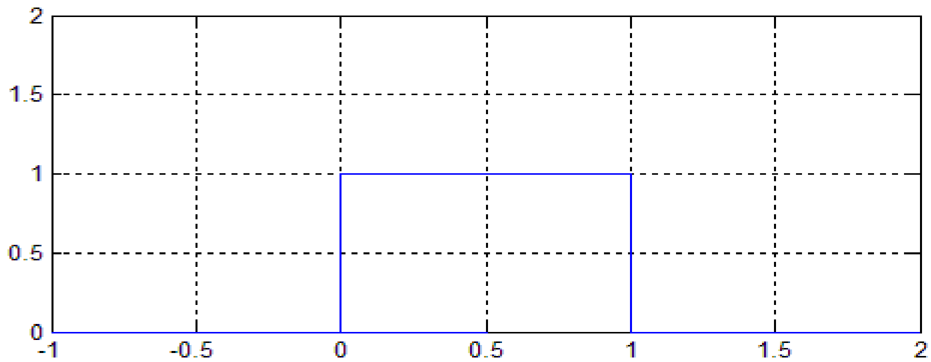
ფუნქციები შემთხვევითი რიცხვებისათვის

ფუნქცია **rand** MATLAB-ში ქმნის შემთხვევით რიცხვებს ინტერვალში [0, 1], ასე მიღებულ შემთხვევით სიდიდეთა განაწილება თანაბარია.

rand (n)	ეს ფუნქცია გვაძლევს n სტრიქონიან და n სვეტიან მატრიცას, რომლის ელემენტებიც 0 და 1 შორის მოთავსებული თანაბრად განაწილებული შემთხვევითი რიცხვებია.
rand (m,n)	გვაძლევს m სტრიქონიან n სვეტიან მატრიცას, რომლის ელემენტებიც 0 და 1 შორის მოთავსებული თანაბრად განაწილებული შემთხვევითი რიცხვებია.
randn (n)	ეს ფუნქცია გვაძლევს n სტრიქონიან, n სვეტიან მატრიცას, რომლის ელემენტებიც ემორჩილებიან ნულოვანი საშუალოსა და ერთეულოვანი დისპერსიის მქონე ნორმალურ განაწილებას
randn (m,n)	ეს ფუნქცია გვაძლევს m სტრიქონიან, n სვეტიან მატრიცას, რომლის ელემენტებიც ემორჩილებიან ნულოვანი საშუალოსა და ერთეულოვანი დისპერსიის მქონე ნორმალურ განაწილებას

შემთხვევით სიდიდეთა თანაბარი განაწილება

შემთხვევითი სიდიდეები ხასიათდება სიმკვრივის ფუნქციით, რომელიც განაწილების ფუნქციის წარმოებულს წარმოადგენს. ეს ფუნქცია ჰგავს ჰისტოგრამას, გვიჩვენებს შემთხვევით სიდიდეთა ინტერვალს და განსაზღვრავს ცალკეული სიდიდეების ხდომილობის ალბათობას. MATLAB ფუნქცია **rand** გვაძლევს შემთხვევით სიდიდეთა ერთობლიობას ისე, რომ თითოეულის მნიშვნელობა თანაბრად ალბათურია ინტერვალში [0,1]. მისი სიმკვრივის ფუნქცია მოცემული მე-5 ნახაზზე. სიმკვრივის ფუნქცია უჩვენებს შემთხვევით სიდიდეთა ზედა და ქვედა ზღვარს (0 და 1) x ღერძზე. ფუნქციის (y ღერძი) მართკუთხა ფორმა მიუთითებს, რომ 0 სა და 1 შორის ყველა სიდიდის ხდომილობის ალბათობა ერთნაირია. სიმკვრივის ფუნქციით შემოსაზღვრული ფართი ერთის ტოლია.



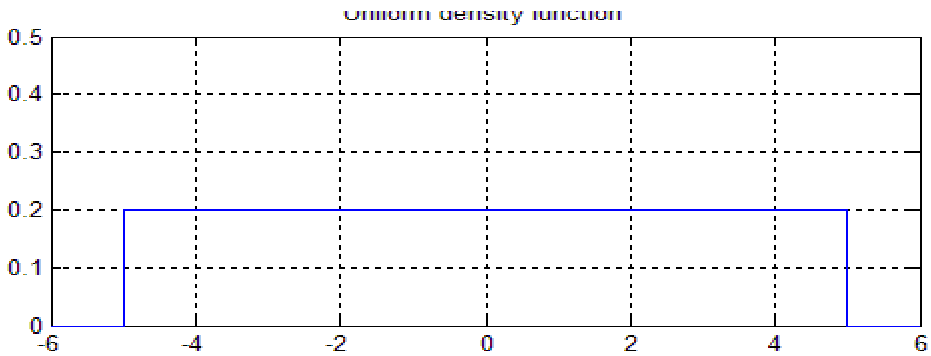
ნახ. 5.5 თანაბარი განაწილების სიმკვრივის ფუნქცია

შემდეგი ბრძანებები ქმნის თანაბრად განაწილებული 10 სიდიდის ერთობლიობას, რომელთა მნიშვნელობები 0 და 1 შორისაა:

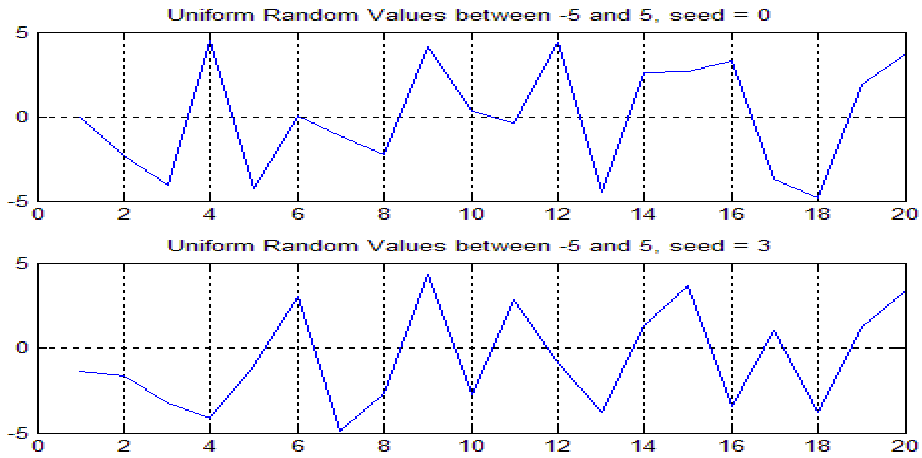
rand(10,1) აწარმოებს თანაბრად განაწილებულ 10 შემთხვევით სიდიდეს [0,1] ინტერვალში.

ხშირად საჭიროა შემთხვევითი სიდიდეები არა მხოლოდ [0 1] ინტერვალში. მაგალითად მე-6 ნახაზზე მოცემულია თანაბარი განაწილების ფუნქციაა –5-სა და 5-ს შორის.

თუ გვაქვს r შემთხვევითი სიდიდე, რომელიც ვაწარმოეთ თანაბარი განაწილების განერატორით [0 1] შორის, მაშინ მისგან ადვილად შეგვიძლია ვაწარმოოთ თანაბრად განაწილებული შემთხვევითი სიდიდე, რომლის მნიშვნელიბაც იქნება სხვა საზღვრებში. პირველ რიგში r-ს გავამრავლებთ სიმკვრივის ფუნქციის სიგანეზე. სიგანე გამოითვლება ზედა და ქვედა საზღვრებს შორის სხვაობით. მიღებულ შედეგს დავუმატებთ ქვედა საზღვრის მნიშვნელობას, რომ მივიღოთ საჭირო ინტერვალი. მაგალითად, დავუშვათ გვსურს მივიღოთ შემთხვევითი სიდიდეთა ერთობლიობა ინტერვალში –5, 5. პირველ რიგში ვაწარმოებთ შემთხვევით სიდიდეებს 0-სა და 1 შორის, შემდეგ მათ გავამრავლებთ 10-ზე და დავუმატებთ –5, შედეგად მივიღებთ შემთხვევით სიდიდეთა ერთობლიობას, რომელიც განაწილებულია თანაბრად ინტერვალში [–5 5]. ამგვარად, თუ გვსურს [0 1] ინტერვალში თანაბრად განაწილებული სიდიდეები გადავიყვანოთ ისეთ სიდიდეებში, რომელნიც თანაბრად იქნებიან განაწილებულნი ინტერვალში [a b], უნდა ვისარგებლოთ ფორმულით: x = (b - a) * r + a



ნახ. 6 სიმკვრივის ფუნქცია თანაბარი განაწილებისათვის ინტერვალში –5 და 5



nax. 7 შემთხვევით სიდიდეთა მიმდევრობა ინტერვალში –5 და 5 **seed** სხვადასხვა მნიშვნელობით

სავარჯიშო

მიეცით MATLAB–ს ბრძანება, რომ შექმნათ შემთხვევით რიცხვთა 10 წევრიანი მიმდევრობა შემდეგი მახასიათებლებით:

1. მიმდევრობა თანაბარი განაწილებით ინტერვალში 0 – 10.0.
2. მიმდევრობა თანაბარი განაწილებით ინტერვალში -1 – 1.
3. მიმდევრობა თანაბარი განაწილებით ინტერვალში -20 – 10.
4. მიმდევრობა თანაბარი განაწილებით ინტერვალში 4.5 – 5.
5. მიმდევრობა თანაბარი განაწილებით ინტერვალში π - -π.

ნორმალური ანუ გაუსის განაწილება

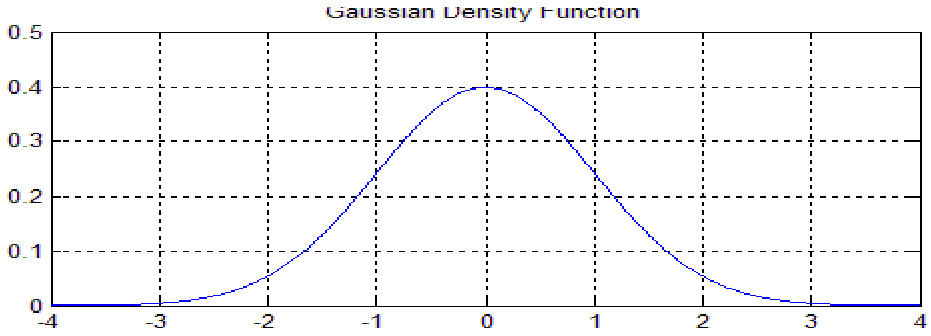
თანაბარი განაწილების დროს ყველა სიდიდის ხდომილების ალბათობა ერთნაირია. რეალურ ცხოვრებაში აბსოლუტურად შემთხვევით სიდიდეებს ახასიათებთ ისეთი განაწილება, სადაც სიდიდეთა რომელიღაც მნიშვნელობები უფრო ხშირად გვხვდება სხვა მნიშვნელობებთან შედარებით. მაგალითად რაიმე სიდიდის გაზომვათა ცდომილებები ჯგუფდებიან საშუალოს მნიშვნელობის გარშემო.

შემთხვევით სიდიდეთა ასეთი მიმდევრობები აღიწერება გაუსის ნორმალური განაწილების კანონის მეშვეობით. მისი სიმკვრივის ფუნქციაა:

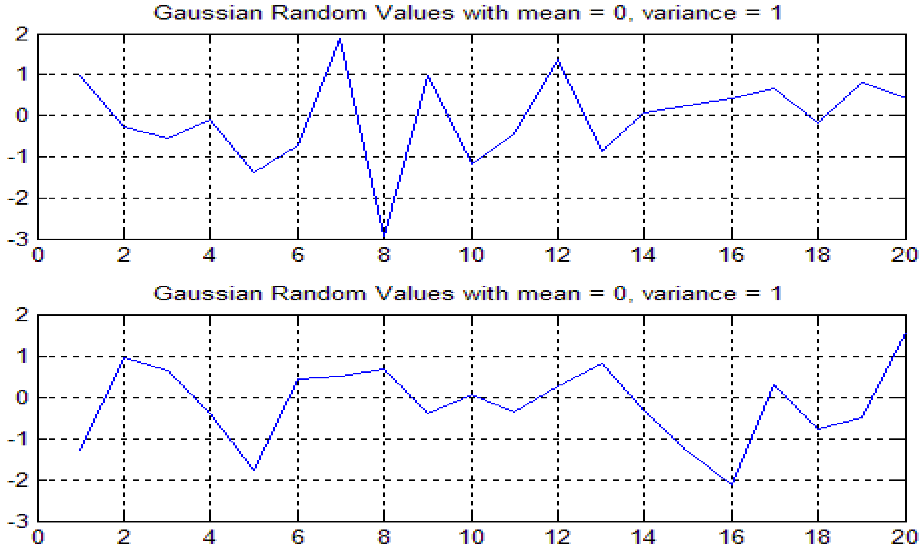
$$f(x)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-(x-\mu)^2/(2\sigma^2)}$$

სადაც μ საშუალოა, ხოლო σ^2 დისპერსია. მაგალითი μ=0 და σ^2=1 მახასიათებლების მქონე გაუსის განაწილების სიმკვრივის ფუნქციისა, ნაჩვენებია ნახ. 8-ზე.

ასეთი განაწილების საშუალო მნიშვნელობა შეესაბამება სიმკვრივის ფუნქციის მაქსიმუმის წერტილის x კოორდინატს. სიმკვრივის ფუნქციის გრაფიკის მიხედვით შეამჩნევთ, რომ საშუალო მნიშვნელობების მახლობელი სიდიდეების გენერირების ალბათობა უფრო დიდია. შევნიშნავთ, რომ თანაბარ განაწილებას ახასიათებს მნიშვნელობათა ზედა და ქვედა საზღვარი, ხოლო გაუსის განაწილებას ასეთი საზღვრები არ გააჩნია. გაუსის შემთხვევოთი სიდიდეების უმრავლესობა საშუალოდან მცირედ განსხვავებულ მნიშვნელობებს იღებს, თუმცა ზოგიერთი მათგანი შესაძლოა საკმარისად დაშორდეს საშუალო მნიშვნელობას. მე-9 ნახაზზე წარმოადგენილია ნულოვანი საშუალოსა და ერთეულოვანი დისპერსიის მქონე ნორმალურად განაწილებული შემთვევითი სიდიდეების ორი განსხვავებული რეალიზაცია. (შესაბამისი განაწილების სიმკვრივის ფუნქციის გარაფიკი მოცემულია ნახ. 8-ზე.)

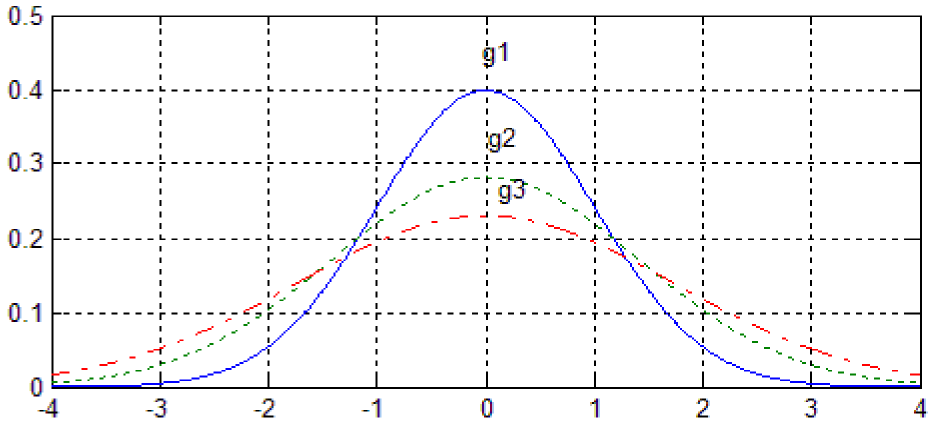


ნახ. 8 ნორმალური განაწილების სიმკვრივის ფუნქცია $\mu = 0$, $\sigma^2 = 1$.



ნახ. 9 ნორმალურად განაწილებულ შემთხვევით სიდიდეთა ორი მიმდევრობა. $\mu=0$, $\sigma^2=1$.

ნახ. 10-ზე წარმოდგენილია გაუსის განაწილების სიმკვრივის სამი სხვადასხვა ფუნქცია. სამივე მათგანის საშუალო მნიშვნელობაა 0, მაგრამ მათი სტანდარტული გადახრა და დისპერსია სხვადასხვაა. g1 განაწილებას აქვს ყველაზე მცირე დისპერსია, g3 – ყველაზე დიდი.



ნახ. 10 გაუსის სამი სხვადასხვა განაწილების სიმკვრივის ფუნქცია

სტატისტიკიდან ცნობილია, რომ გაუსის განაწილებისათვის სიდიდეთა 68% ღებულობს მნიშვნელობებს საშუალოდან $\pm\sigma$ ინტერვალში, 95% $\pm2\sigma$ ინტერვალში, ხოლო 99% - $\pm3\sigma$ ინტერვალში.

MATLAB ფუნქცია **randn** ქმნის გაუსის ნორმალური კანონის მიხედვით განაწილებულ შემთხვევით სიდიდეთა მიმდევრობას საშუალოთი 0 და დისპერსიით 1. იმისათვის რომ შევქმნათ განაწილება განსხვავებული პარამეტრებით, ეს სიდიდე უნდა

გავამრავლოთ შესაბამის სტანდარტულ გადახრაზე და დავუმატოთ საშუალო მნიშვნელობა. ამგვარად, თუ r შემთხვევითი სიდიდეა საშუალოთი 0 და სტანდარტული გადახრით 1, შემდეგი განტოლება აწარმოებს შემთხვევით სიდიდეს x-ს საშუალო მნიშვნელობით b და სტანდარტული გადახრით a: $x = a \cdot r + b$. მაგალითად შემდეგი ბრძანებები გვაძლევს შემთხვევით სიდიდეთა მიმდევრობას გაუსის ანუ ნორმალური განაწილებით, რომლის საშუალოა 5 და დისპერსია 2:

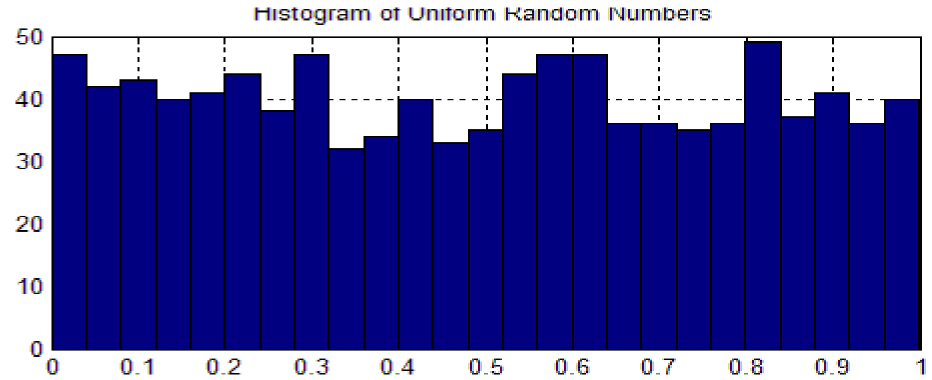
```
s = sqrt(2)*randn(10,1) + 5.
```

სიმკვრივის ფუნქცია

ბრძანება **hist** შეიძლება გამოვიყენოთ სიგნალის სიმკვრივის ფუნქციის შესაფასებლად. თუ გვაქვს 1000 შემთხვევითი სიდიდე და ავაგებთ მათ ჰისტოგრამას, ფაქტიურად ჩვენ ვაგებთ მათი განაწილების სიმკვრივის ფუნქციის გრაფიკს. მაგალითად MATLAB საშუალებით შევქმნით თანაბრად განაწილებული 1000 რიცხვის მიმდევრობა 0 და 1 შორის, ჩავწერეთ ისინი სვეტ ვექტორში u_values. შეგვიძლია ვისარგებლოთ **hist** ბრძანებით იმისათვის, რომ ავაგოთ სათანადო განაწილების ფუნქცია 25 სვეტით:

```
u_values = rand(1000,1);
hist(u_values,25)
```

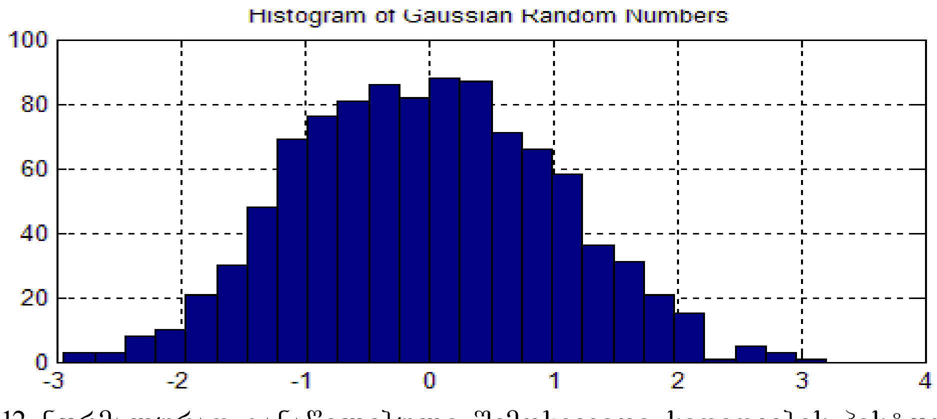
შედეგად მივიღებთ მე-11 ნახაზზე წარმოდგენილ გრაფიკს. როგორც მოსალოდნელი იყო, სიდიდეები განაწილებული არიან 0 და 1 შორის და განაწილება შედარებით თანაბარია. ახლა გავიმეოროთ იგივე ნორმალური განაწილებისათვის საშუალოთი 0 და გადახრით 1.



ნახ. 11 თანაბრად განაწილებული შემთხვევითი სიდიდეების ჰისტოგრამა

```
g_values = randn(1000,1);
hist(u_values,25)
```

მივიღებთ მე-12 ნახაზზე წარმოდგენილ გრაფიკს. როგორც მოსალოდნელი იყო, განაწილების გრაფიკის პიკია 0, საშუალო მნიშვნელობა და მონაცემთა უმრავლესობის მნიშვნელობები მოთავსებულია ინტერვალში $\pm 2\sigma$.



ნახ. 12 ნორმალურად განაწილებული შემთხვევითი სიდიდეების ჰისტოგრამა

სავარჯიშო

MATLAB-ის საშუალებით აწარმოეთ 1000 შემთხვევითი სიდიდე განსაზღვრული მახასიათებლებით. გამოთვალეთ მისი საშუალო და სტანდარტული გადახრა. ააგეთ მათთვის ჰისტოგრამა 25 სვეტით:

1. შემთხვევით სიდიდეთა გაუსის განაწილება საშუალოთი 1 და სტანდარტული გადახრით 0.5
2. შემთხვევით სიდიდეთა გაუსის განაწილება საშუალოთი –5.5 და სტანდარტული გადახრით 0.25
3. შემთხვევით სიდიდეთა გაუსის განაწილება საშუალოთი –5.5 და სტანდარტული გადახრით 1.25
4. შემთხვევით სიდიდეთა გაუსის განაწილება საშუალოთი π და სტანდარტული გადახრით π/8.

თანაფარდობა სიგნალი/ხმაური

ხშირად საჭიროა ისეთი სიგნალის მოდელირება, რომელსაც ახლავს შემთხვევითი ხმაური. ხმაური შემთხვევით სიდიდეთა ერთობლიობაა და მისი წილი მიღებულ საგნალში განისაზღვრება თანაფარდობით სიგნალი/ხმაური – SNR (Signal to Noise Ratio). იგი სიგნალის სიმძლავრის ერთეულებში გამოისახება.

სიგნალის სიმძლავრე

სიმძლავრე – ეს არის სიგნალის საშუალოს გარშემო ცვლილების ზომა. რაც მეტია სიგნალის ამოლიტუდა, მით მეტია მისი სიმძლავრე. რამდენადაც ამპლიტუდა შეიძლება იყოს დადებითიც და უარყოფითიც, სიმძლავრე განისაზღვრება ამპლიტუდის კვადრატით, ისე რომ სიმძლავრის მნიშვნელობა ყოველთვის დადებითი სიდიდეა. x ვექტორით წარმოდგენილი სიგნალის სიმძლავრე შეიძლება შევაფასოთ სიგნალის მნიშვნელობათა კვადრატების საშუალო მნიშვნელობით:

$$power \approx \frac{\sum_{1}^{N} x_k^2}{N}$$

MATLAB–ში ეს სიდიდე გამოითვლება **sum** ფუნქციის საშუალებით:

power = sum(x.^2)/N; (1)

შესაძლებელია აგრეთვე ვაჩვენოთ, რომ სიგნალის სიმძლავრე ტოლია დისპერსიისა და საშუალო მნიშვნელობის კვადრატთა ჯამისა:

$$power = \sigma^2 + \mu^2$$

MATLAB საშუალებით ეს შეიძლება შემდეგი ბრძანებებით გამოითვალოს:

power = std(x)^2 + mean(x)^2; (2)

თუ სიგნალი სინუსოიდაა, ადვილად შეიძლება ვაჩვენოთ, რომ მისი სიმძლავრე ტოლია მისივე ამპლიტუდის კვადრატის ნახევრის. მაგალითად სინუსოიდის 4sin2πt სიმძლავრე ტოლია 16/2, ანუ 8-ის.

სავარჯიშო

მოცემული პარამეტრების საშუალებით აწარმოეთ 100 შემთხვევითი სიდიდე. გამოთვალეთ სიმძლავრე ორი სხვადასხვა გზით (მნიშვნელობათა კვადრატების საშუალო და საშუალოსა და დისპერსიის მნიშვნელობათა მიხედვით) და შეადარეთ ერთმანეთს.

1. თანაბარად განაწილებული შემთხვევითი სიდიდეები ინტერვალში 0 და 10.
2. თანაბარად განაწილებული შემთხვევითი სიდიდეები ინტერვალში -2 და 4.
3. თანაბარად განაწილებული შემთხვევითი სიდიდეები საშუალოთი 0 და დისპერსიით 1.0
4. თანაბარად განაწილებული შემთხვევითი სიდიდეები საშუალოთი –0.5 და დისპერსიით 4.0

5. ნორმალურად განაწილებული შემთხვევითი სიდიდეები საშუალოთი 0 და დისპერსიით 2.0
6. შემთხვევოთი სიდიდეები ნორმალური განაწილებით საშუალოთი 0 და დისპერსიით 0.5
7. შემთხვევოთი სიდიდეები ნორმალური განაწილებით საშუალოთი –2.5 და დისპერსიით 0.5

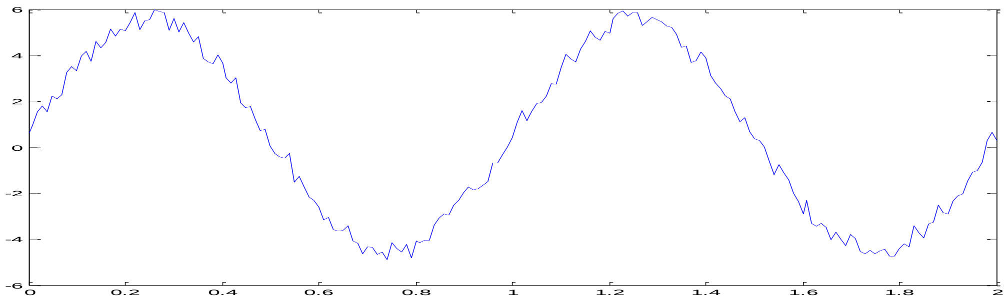
სიგნალის ხმაურთან შეფარდების გამოთვლა

თანაფარდობა სიგნალი/ხმაური არის სიგნალის სიმძლავრის თანაფარდობა ხმაურის სიმძლავრესთან. საილუსტრაციოდ შევქმნათ 201 ელემენტიანი ვექტორი სიგნალი, რომელიც წარმოადგენსა სინუსოიდისა და თანაბრად განაწილებული ხმაურის ჯამს (ნახ. 13)

```
t=(0:0.01:2); N=length(t);
s1=5*sin(2*pi*t);
s2= rand(1,201);
s=s1+s2;
plot(t,s);
```

ახლა შეგვიძლია გამოვთვალოთ სიგნალის ხმაურთან თანაფარდობა, ამისათვის ცალკე ვიანგარიშებთ სინუსოიდალური სიგნალის სიმძლავრეს და ცალკე ხმაურისას და გამოვიანგარიშებთ მათ შეფარდებას:

snr= sum(s1.^2)/N/(sum(s2.^2)/N)



ნახ. 13 სინუსოიდური სიგნალი ხმაურით.

განხილული ბრძანებები და ფუნქციები

cumprod	განსაზღვრავს კუმულაციურ ნამრავლს
cumsum	განსაზღვრავს კუმულაციურ ჯამს
hist	აგებს ჰისტოგრამას
max	განსაზღვრავს უდიდეს მნიშვნელობას
mean	განსაზღვრავს საუალო მნიშვნელობას
median	განსაზღვრავს მედიანას მნიშვნელობას
min	განსაზღვრავს უმცირეს მნიშვნელობას
prod	განსაზღვრავს მონაცემთა ნამრავლს
rand	განსაზღვრავს შემთხვევით რიცხვებს
sort	ზრდის მიხედვით დაალაგებს მონაცემებს
std	გამოითვლის სტანდარტულ გადახრას
sum	განსაზღვრავს მონაცემთა ჯამს