Universitatea din Bucuresti

Proiect Probabilitati si Statistica

Distributii si densitati de repartitie

Student:
Alexandru Marin

10 Decembrie, 2013

1 DENSITATI DE REPARTITIE

1.1 Distributia uniforma

1.1.1 Densitatea si functia de repartitie

O variabila aleatoare ξ are o distributie uniforma daca poate lua echiprobabil orice valoare intr-un interval (variabila este uniform distribuita in interval), adica f(x) = ct, cu x$\in$[a,b]. Variabilele aleatoare pot fi considerate ca fiind functii de la un domeniu oarecare catre multimea numerelor reale.

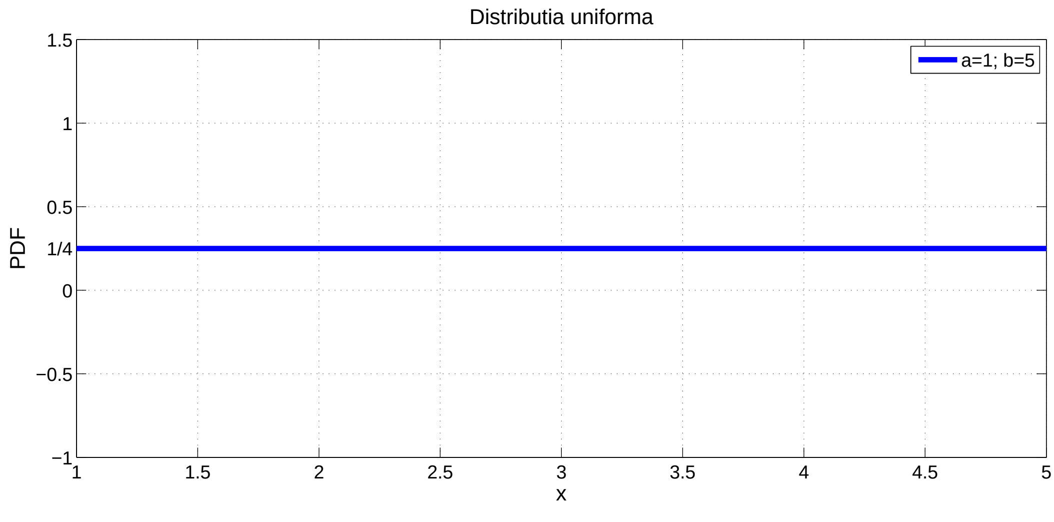
Daca a si b sunt capetele intervalului, atunci **densitatea de probabilitate** va fi:

$$\omega_\xi(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{in rest} \end{cases}$$

Valoarea medie a unei variabile de genul acesta este $\frac{a+b}{2}$.

Pentru a vizualiza graficul densitatii, vom alege intervalul [1,5].
Astfel, pentru:

- $x = 2, \omega_\xi(2) = \frac{1}{5-1} = \frac{1}{4}$;
- $x = 3, \omega_\xi(3) = \frac{1}{5-1} = \frac{1}{4}$;
- $x = 4, \omega_\xi(4) = \frac{1}{5-1} = \frac{1}{4}$;

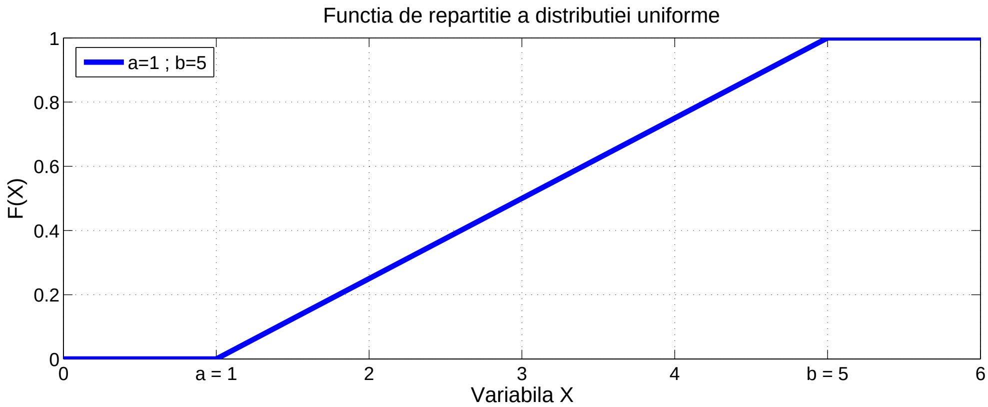


Cod **Matlab**:

```
x = [1:1:5];
y = unifpdf(x,1,5);
plot(x,y)
```

Forma distributiei indica faptul ca o variabila aleatoare uniforma poate lua cu aceeasi probabilitate orice valoare in intervalul [a,b], dar nu poate lua nicio valoare in exteriorul acestuia.

Functia de repartitie a distributie uniforme: $F(x|a,b) = \frac{x-a}{b-a} I_{|}a, b|(x)$.



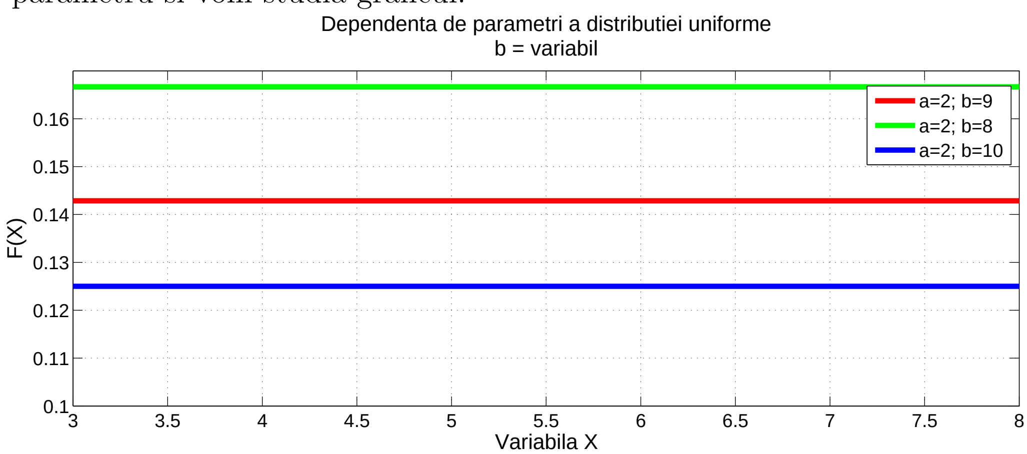
Cod **Matlab**:

```
x = [0:1:6];
y = unifcdf(x,1,5);
plot(x,y)
```

Se poate observa faptul ca probabilitatea ca variabila x sa ia valori mai mari creste liniar in intervalul [1,5], in afara acestuia functia luand valoarea 0, pentru $x < a$, respectiv 1, pentru $x \geq b$.

1.1.2 Dependenta de parametri

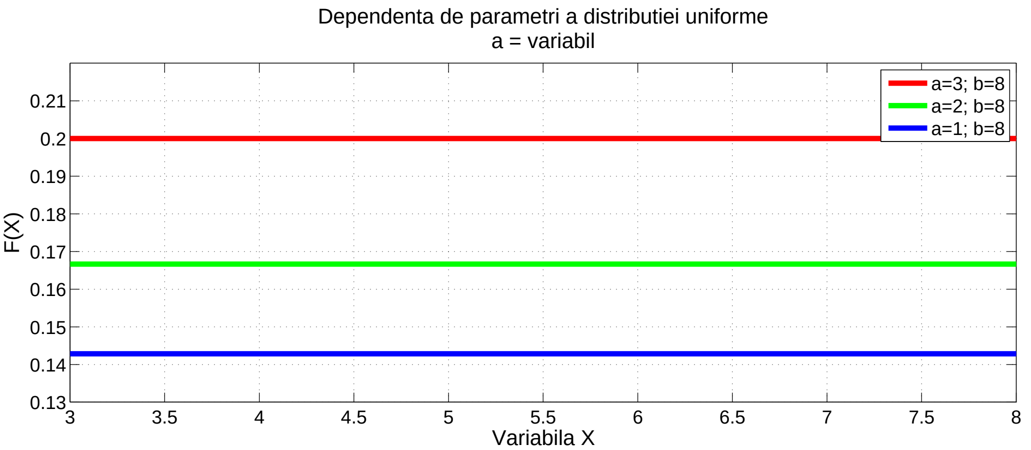
Pentru a determina dependenta distributiei de parametrii a si b, vom schimba, pe rand, fiecare parametru si vom studia graficul.



Cod **Matlab**:

```
x = [3:1:8];
a = unifpdf(x,2,8);
b = unifpdf(x,2,9);
c = unifpdf(x,2,10);
plot(x,a,'green',x,b,'red',x,c,'blue')
```

In cazul $b = variabil$, distributia isi pastreaza proprietatea, astfel variabila x poate lua diferite valori cu aceeasi probabilitate, insa probabilitatea scade odata cu cresterea limitei drepte a intervalului ($punctul\ b$).

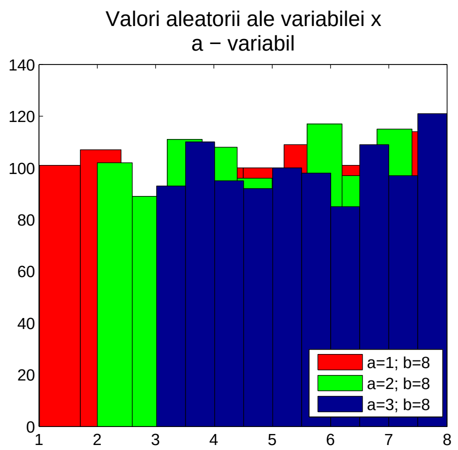


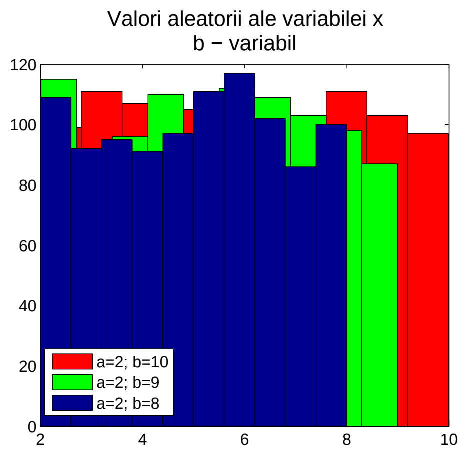
Cod **Matlab**:

```
x = [3:1:8];
a = unifpdf(x,3,8);
b = unifpdf(x,2,8);
c = unifpdf(x,1,8);
plot(x,a,'red',x,b,'green',x,c,'blue')
```

Spre deosebire de cazul precedent, in acest caz probabilitatea creste odata cu scaderea limitei stangi a intervalului ($punctul\ a$).

Pentru valori aleatorii ale variabilei x vom folosi functia *unifrnd* pentru parametrii precedenti si vom reprezenta grafic histogramele valorilor.





Cod **Matlab**:

```
x = unifrnd(2,8,1,1000);
y = unifrnd(2,9,1,1000);
z = unifrnd(2,10,1,1000);
a = unifrnd(1,8,1,1000);
b = unifrnd(2,8,1,1000);
c = unifrnd(3,8,1,1000);
subplot(1,2,1);
hist(x)
hold on
hist(y)
hist(z)
subplot(1,2,2)
hist(c)
hold on
hist(b)
hist(a)
```

Putem considera fluctuatia probabilitatii neglijabila, drept urmare x poate lua cu aproximativ aceeasi probabilitate orice valoare din intervalul specificat, indiferent de parametri.

1.2 Distributia exponentiala

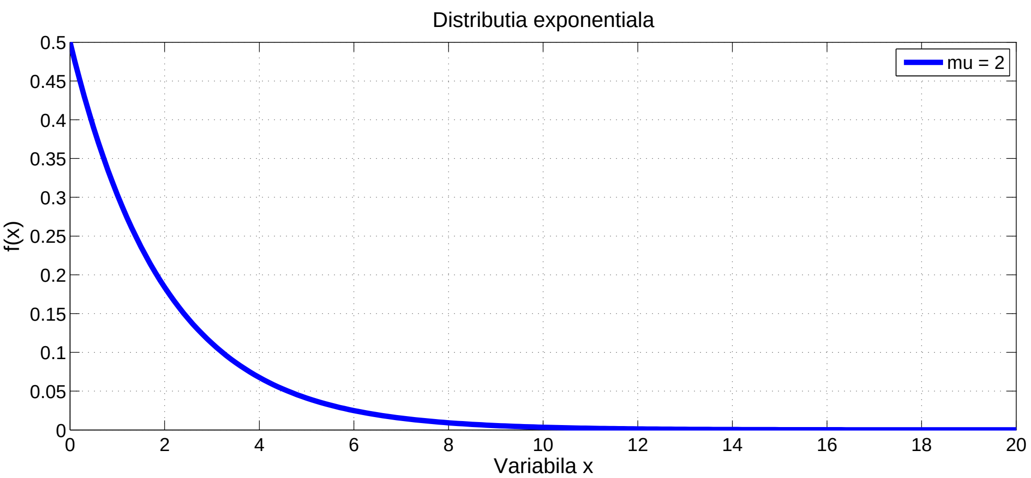
1.2.1 Densitatea si functia de repartitie

O variabila continua X are distributie exponentiala de parametru $\lambda > 0$ daca are densitatea de probabilitate de forma:

$$f(x) = \lambda \cdot e^{-\lambda \cdot x}, x \geq 0$$

, unde $\lambda = \frac{1}{\mu}$.
Valoarea medie a acestei variabile este $\frac{1}{\lambda}$.

Pentru $\mu = 2$, $f(x|\mu = 2) = \frac{1}{2} \cdot e^{-\frac{x}{2}}$.
Pentru $\mu = 3$, $f(x|\mu = 3) = \frac{1}{3} \cdot e^{-\frac{x}{3}}$.

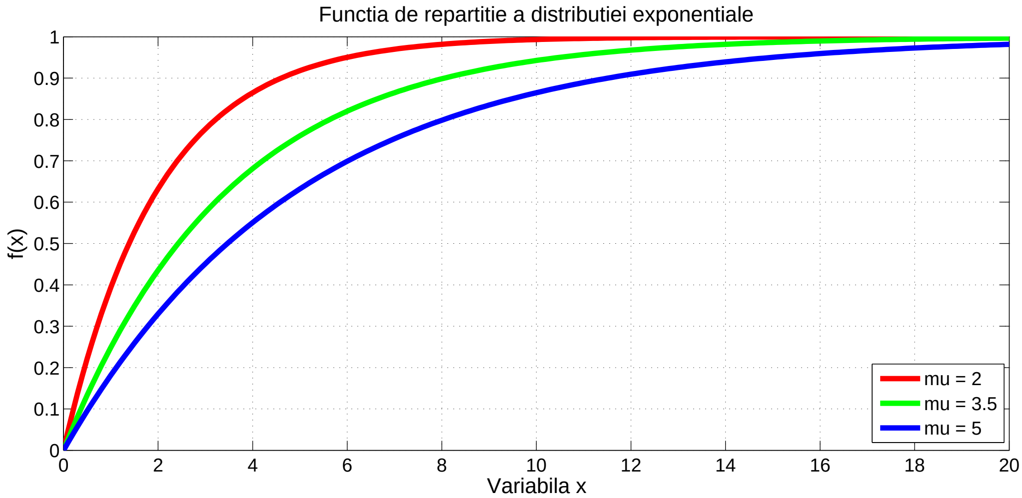


Cod **Matlab**:

```
x = [0:0.1:20];
y = exppdf(x,2);
plot(x,y)
```

Forma distributiei indica faptul ca probabilitatea ca x sa ia valori mici este ridicata, iar probabilitatea ca variabila sa ia valori mici este foarte scazuta.

Functia de repartitie a distributiei exponentiale este: $F(x|\mu) = \int_0^x \frac{1}{\mu} e^{\frac{-t}{\mu}} dt = 1 - e^{\frac{-x}{\mu}}$
.



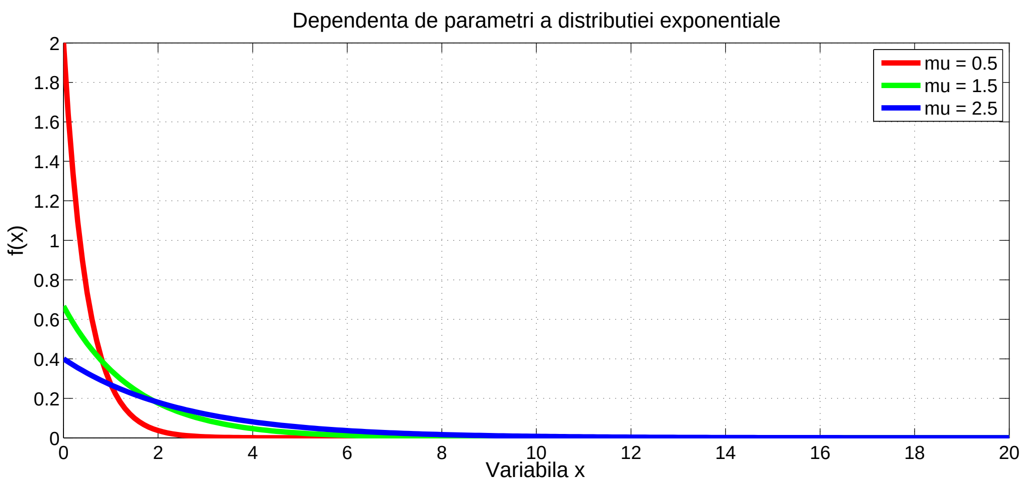
Cod **Matlab**:

```
x = [0:0.1:20];
y = expcdf(x,2);
z = expcdf(x,3.5);
a = expcdf(x,5);
plot(x,y,'r',x,z,'g',x,a,'b')
```

Functia creste exponential si reiese faptul ca x poate lua cu probabilitate ridicata valori mari, in timp ce pentru valori mici exista o probabilitate scazuta.

1.2.2 Dependenta de parametri

Pentru a demonstra dependenta functiei de parametrul μ, vom alege valori diferite si vom reprezenta grafic functia.

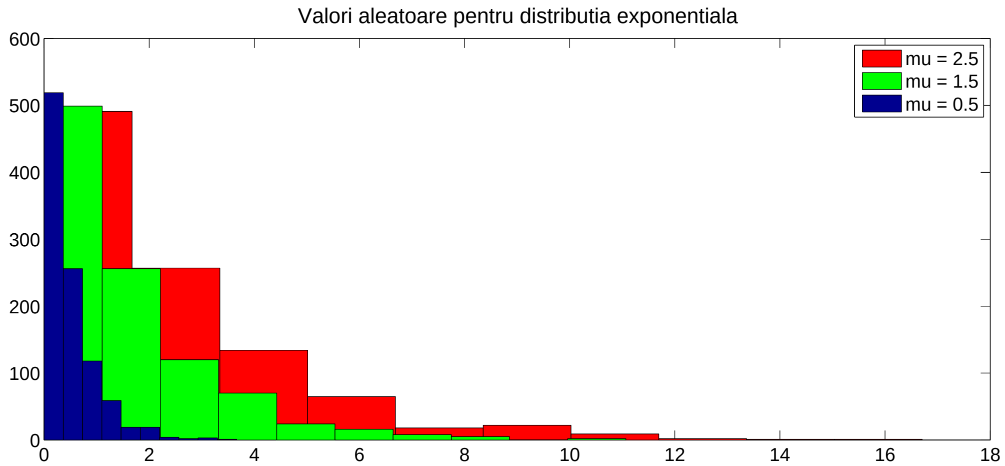


Cod **Matlab**:

```
x = [0:0.1:20];
y = exppdf(x,0.5);
z = exppdf(x,1.5);
a = exppdf(x,2.5);
plot(x,y,'r',x,z,'g',x,a,'b')
```

Din grafic reiese ca, odata cu reducerea valorii lui μ, probabilitatea ca x sa ia valori mici scade, in timp ce probabilitatea ca x sa ia valori mari ramane la fel de scazuta.

Pentru valori aleatorii ale variabilei x vom folosi functia *exprnd* cu parametrii precedenti si vom reprezenta grafic histogramele.



Cod **Matlab**:

```
x = [0:0.1:20];
y = exprnd(0.5,1,1000);
z = exprnd(1.5,1,1000);
a = exprnd(2.5,1,1000);
hist(a)
hold on
hist(z)
hist(y)
```

Dupa cum stiam deja din formula densitatii, valorile mici sunt cele mai probabile, valorile mari avand o probabilitate foarte scazuta.

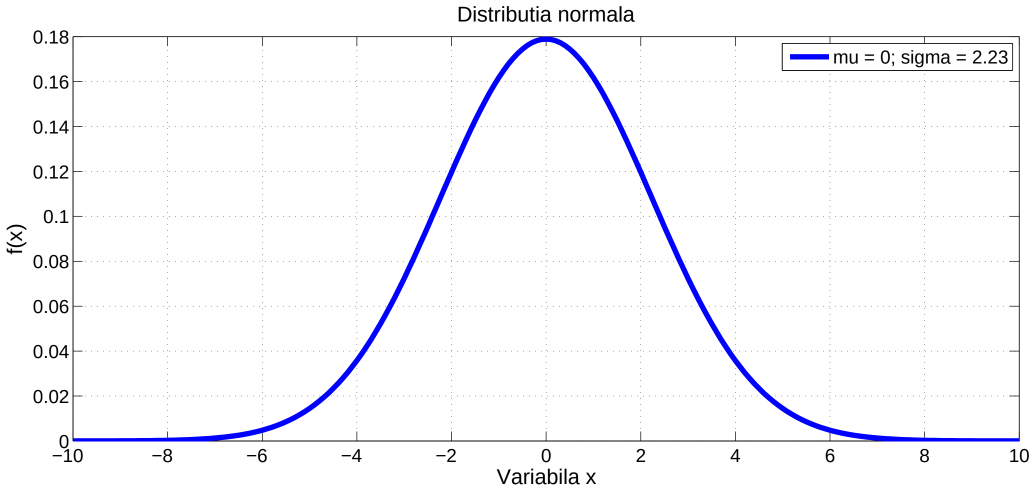
1.3 Distributia normala

Distributia normala este o distributie de probabilitate continua. Este numita de asemenea **distributia Gauss**, deoarece a fost descoperita de catre **Carl Friedrich Gauss**.
Distributia normala standard(cunoscuta si sub numele de **distributie Z**) este distributia normala cu media zero si variatia 1. Aceasta este adesea numita **clopotul lui Gauss**, deoarece graficul densitatii de probabilitate arata ca un clopot.
Se noteaza cu $N(\mu, \sigma)$, unde μ si σ sunt parametrii din functia de distributie care va fi descrisa in continuare.

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

$f(x|\mu = 0, \sigma^2 = 5) = \frac{1}{2.23\sqrt{2\pi}} e^{\frac{-(x)^2}{10}}$

$f(x|\mu = -5, \sigma^2 = 0.2) = \frac{1}{0.44\sqrt{2\pi}} e^{\frac{-(x+5)^2}{0.4}}$



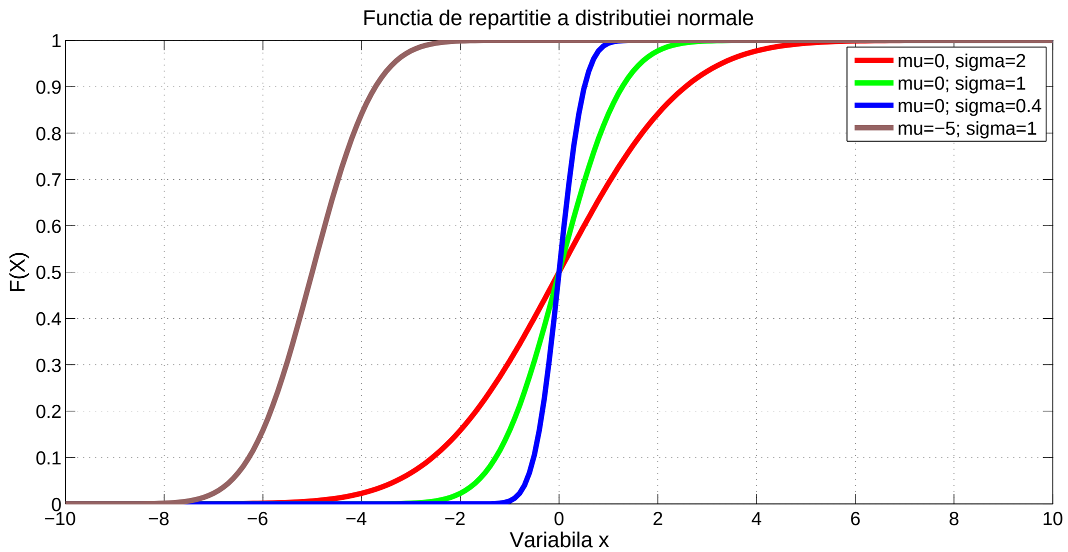
Cod **Matlab**:

```
x = [-10:0.1:10];
y = normpdf(x,0,2.23);
plot(x,y)
```

Densitatea este simetrica in jurul punctului μ (in cazul nostru $\mu = 0$), avand probabilitatea foarte mare de a lua valoarea 0, scazand simetric in afara acestei valori.

Functia de repartitie a distributiei normale este:

$$F(X|\mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{-(t-\mu)^2}{2\sigma^2}} dt$$



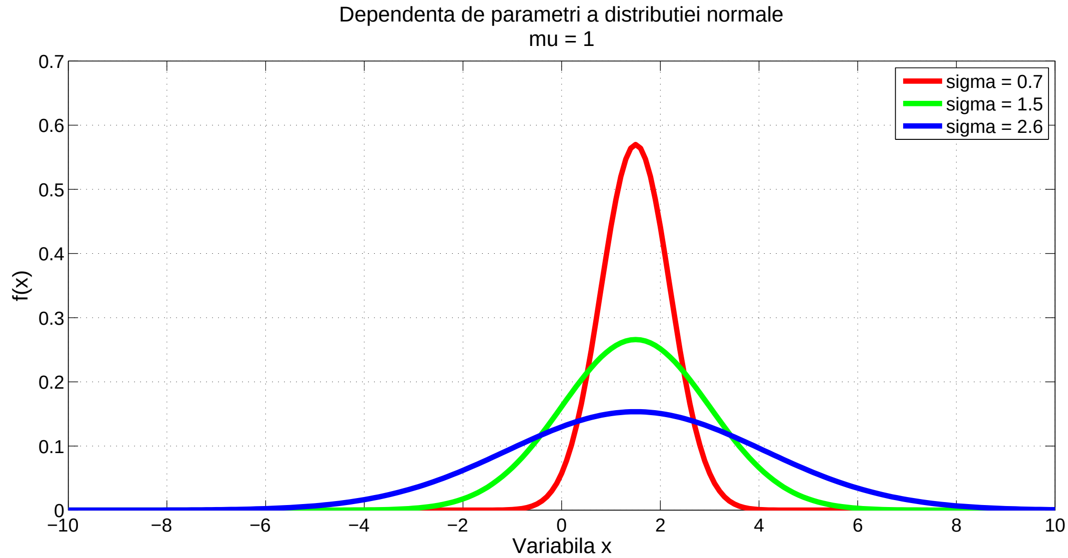
Cod **Matlab**:

```
x = [-10:0.1:10];
y = normcdf(x,0,2);
z = normcdf(x,0,3.5)
a = normcdf(x,0,5)
b = normcdf(x,-5,1)
plot(x,y,'r',x,z,'g',x,a,'b',x,b,'br')
```

Functia variaza in jurul punctului μ; pentru valori $a < \mu$ probabilitatea ca x sa ia valorile a scade, iar pentru valori $b > \mu$ probabilitatea creste.

1.3.1 Dependenta de parametri

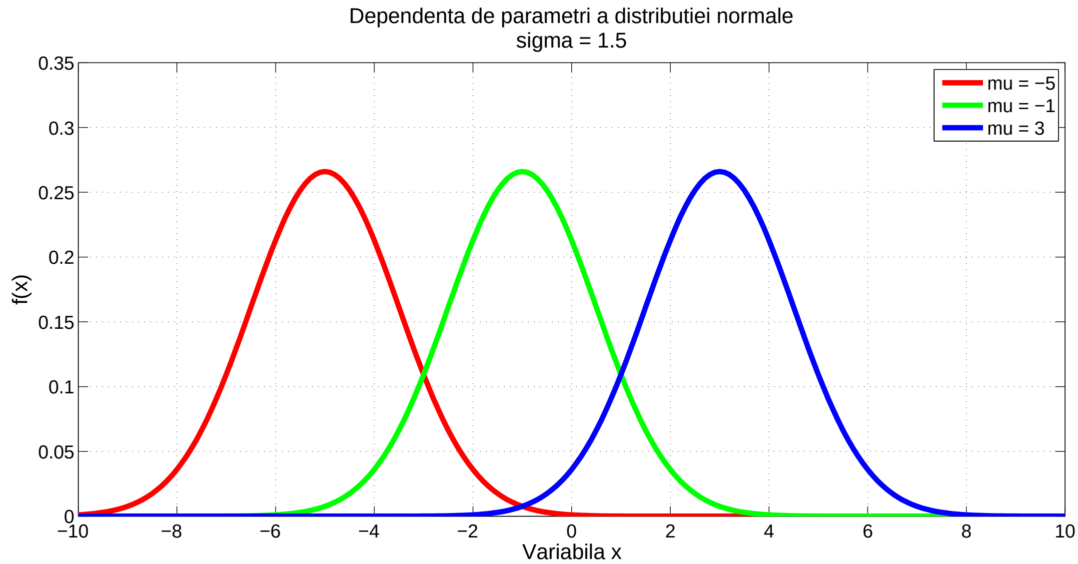
Pentru a demonstra dependenta distributiei de parametrii μ si σ, vom lua valori aleatorii pentru acestea si vom reprezenta grafic distributia.



Cod **Matlab**:

```
x = [-10:0.1:10];
y = normpdf(x,1,0.7);
z = normpdf(x,1,1.5);
a = normpdf(x,1,2.6);
plot(x,y,'r',x,z,'g',x,a,'b')
```

Odata cu cresterea variabilei σ creste si dispersia functiei, x avand probabilitatea mai mare de a lua valori in jurul punctului μ.

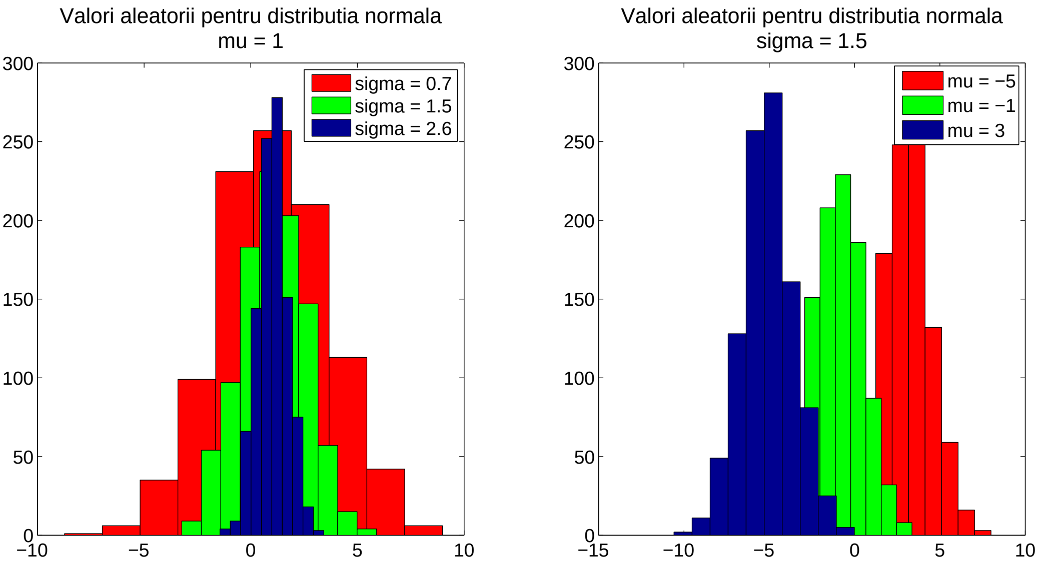


Cod **Matlab**:

```
x = [-10:0.1:10];
y = normpdf(x,-5,1.5);
z = normpdf(x,-1,1.5)
a = normpdf(x,3,1.5)
plot(x,y,'r',x,z,'g',x,a,'b')
```

In acest caz, centrul functiei se deplaseaza odata cu schimbarea variabilei μ, astfel pentru valori mai mici, functia se deplaseaza spre stanga axei, in timp ce dispersia ramane neschimbata.

Vom alege valori aleatorii pentru variabila x folosind functia *normrnd* si vom reprezenta grafic histogramele.



Cod **Matlab**:

```
x = normrnd(1,0.7,1,1000);
y = normrnd(1,1.5,1,1000);
z = normrnd(1,2.6,1,1000);
a = normrnd(-5,1.5,1,1000);
b = normrnd(-1,1.5,1,1000);
c = normrnd(3,1.5,1,1000);
subplot(1,2,1)
hist(z)
hold on
hist(y)
hist(x)
subplot(1,2,2)
hist(c)
hold on
hist(b)
hist(a)
```

Si in cazul valorilor aleatorii se respecta proprietatile enumerate mai sus.

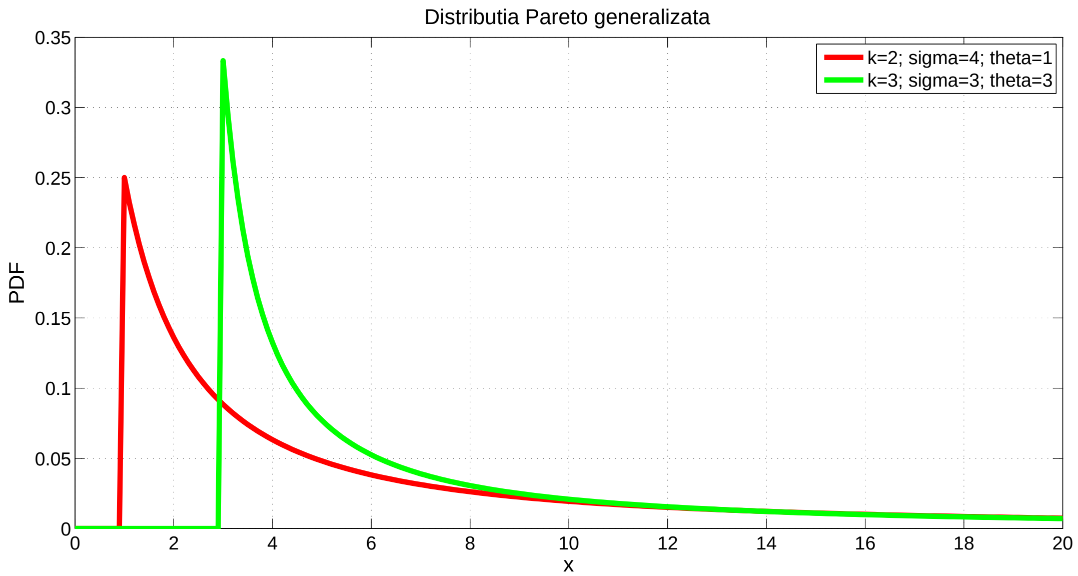
1.4 Distributia Pareto generalizata

1.4.1 Densitatea si functia de repartitie

In statistica, **distributia Pareto generalizata (GPD)** reprezinta o familie de distributii de probabilitate continue. Este specificata de trei parametri: forma k ($k \neq 0$), scala σ si limita θ.

$$f(x|k,\sigma,\theta) = \left(\frac{1}{\sigma}\right)\left(1 + k\frac{(x-\theta)}{\sigma}\right)^{-1-\frac{1}{k}}$$

$f(x|2,4,1) = \left(\frac{1}{4}\right)\left(1 + \frac{(x-1)}{2}\right)^{-\frac{3}{2}}$
$f(x|3,3,3) = \left(\frac{1}{3}\right)(1 + (x-3))^{-\frac{4}{3}}$



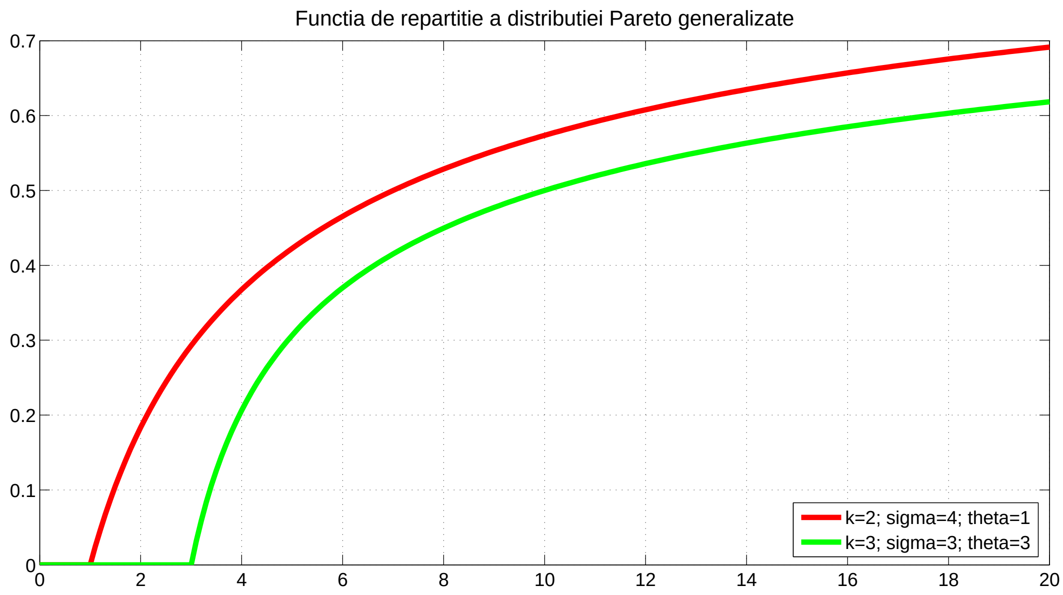
Cod **Matlab**:

```
x = [0:0.1:20]y = gppdf(x,2,4,1);
z = gppdf(x,3,3,3);
plot(x,y,'r',x,z,'g')
```

Avand ca punct de plecare punctul limita θ, densitatea scade exponential pana cand variabila x va avea o probabilitate foarte scazuta de a primi o valoare mare.

Functia de repartitie a **distributiei Pareto generalizate** este:

$$f(x|k,\sigma,\theta) = 1 - \left(1 + k\frac{x-\theta}{\sigma}\right)^{\frac{-1}{k}}$$



Cod **Matlab**:

```
x = [0:0.1:20]y = gpcdf(x,2,4,1);
z = gpcdf(x,3,3,3);
plot(x,y,'r',x,z,'g')
```

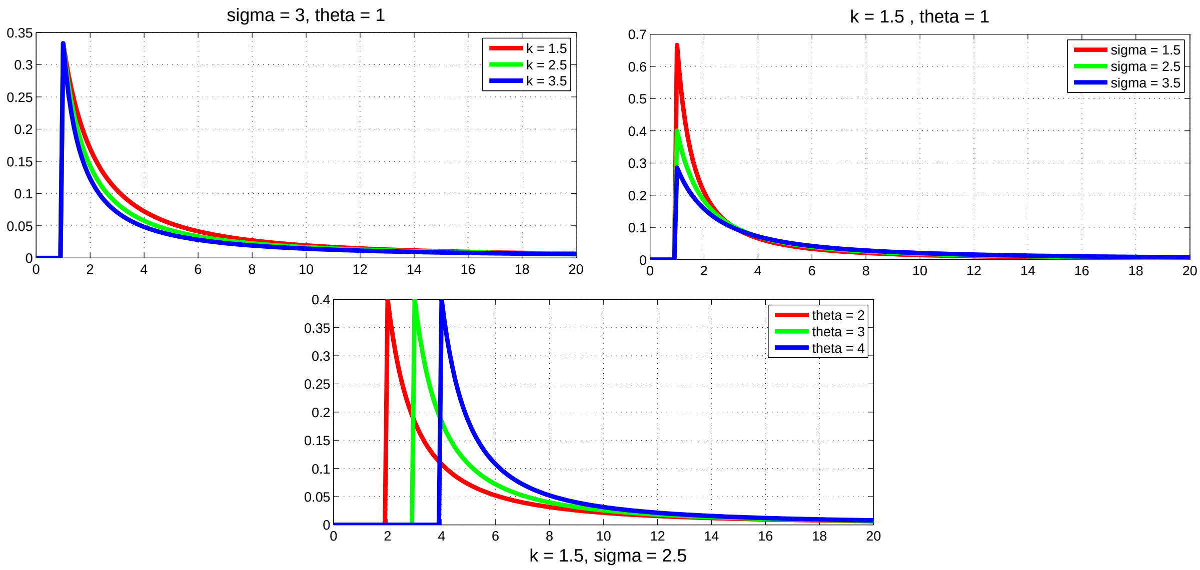
Spre deosebire de densitate, functia de repartitie creste exponential pana in punctul in care variabila x va avea o probabilitate foarte mare de a primi valori mari.

1.4.2 Dependenta de parametri

Pentru a studia dependenta de parametri, vom alege, pe rand, valori diferite pentru k, σ si θ si vom reprezenta grafic rezultatele.

Vom observa ca:

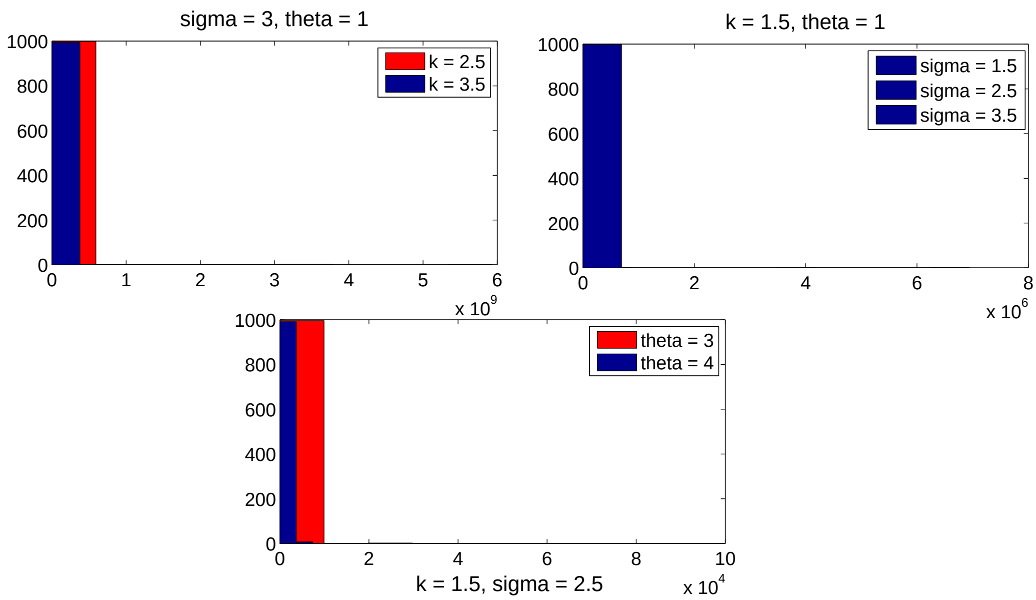
- pentru **k variabil**, cele 3 densitati pornesc cu aceeasi probabilitate, insa odata cu cresterea variabilei k scade dispersia, astfel pentru un k mai mare, probabilitatea ca x sa ia valori mari scade exponential.
- pentru σ **variabil**, odata cu cresterea variabilei scade probabilitatea densitatii in punctul de plecare. Astfel, pentru un σ mai mare, probabilitatea ca x sa ia valori mici scade, iar probabilitatea valorilor mari ramane constanta.
- pentru θ **variabil**, difera punctul de plecare pentru cele 3 densitati. Astfel, pentru un θ mai mare, densitatea isi va muta punctul de plecare spre dreapta, deci va avea probabilitatea mai ridicata de a lua valori mai mari.



Cod **Matlab**:

```
y = gppdf(x,1.5,3,1);
z = gppdf(x,2.5,3,1);
a = gppdf(x,3.5,3,1);
b = gppdf(x,1.5,1.5,1);
c = gppdf(x,1.5,2.5,1);
d = gppdf(x,1.5,3.5,1);
e = gppdf(x,1.5,2.5,2);
f = gppdf(x,1.5,2.5,3);
g = gppdf(x,1.5,2.5,4);
subplot(2,2,1)
plot(x,y,'r',x,z,'g',x,a,'b')
subplot(2,2,2)
plot(x,b,'r',x,c,'g',x,d,'b')
subplot(2,2,3)
plot(x,e,'r',x,f,'g',x,g,'b')
```

Pentru a evidentia densitatea, vom alege valori aleatorii pentru variabila x folosind functia *gprnd* si vom reprezenta grafic dispersia acestora.
Astfel, vom observa ca valorile foarte mici (0-0.5) sunt foarte probabile, in timp ce valorile mai mari au o probabilitate foarte scazuta, indiferent de parametrii folositi.



Cod **Matlab**:

```
x=gprnd(1.5,3,1,1,1000);
y=gprnd(2.5,3,1,1,1000);
z=gprnd(3.5,3,1,1,1000);
a=gprnd(1.5,1.5,1,1,1000);
b=gprnd(1.5,2.5,1,1,1000);
c=gprnd(1.5,3.5,1,1,1000);
d=gprnd(1.5,2.5,2,1,1000);
e=gprnd(1.5,2.5,3,1,1000);
f=gprnd(1.5,2.5,4,1,1000);
subplot(2,2,1)
hist(y)
hold on
hist(z)
subplot(2,2,2)
hist(a)
hold on
hist(b)
hist(c)
subplot(2,2,3)
hist(f)
hold on
hist(e)
```

1.5 Distributia chi-patrat (χ^2)

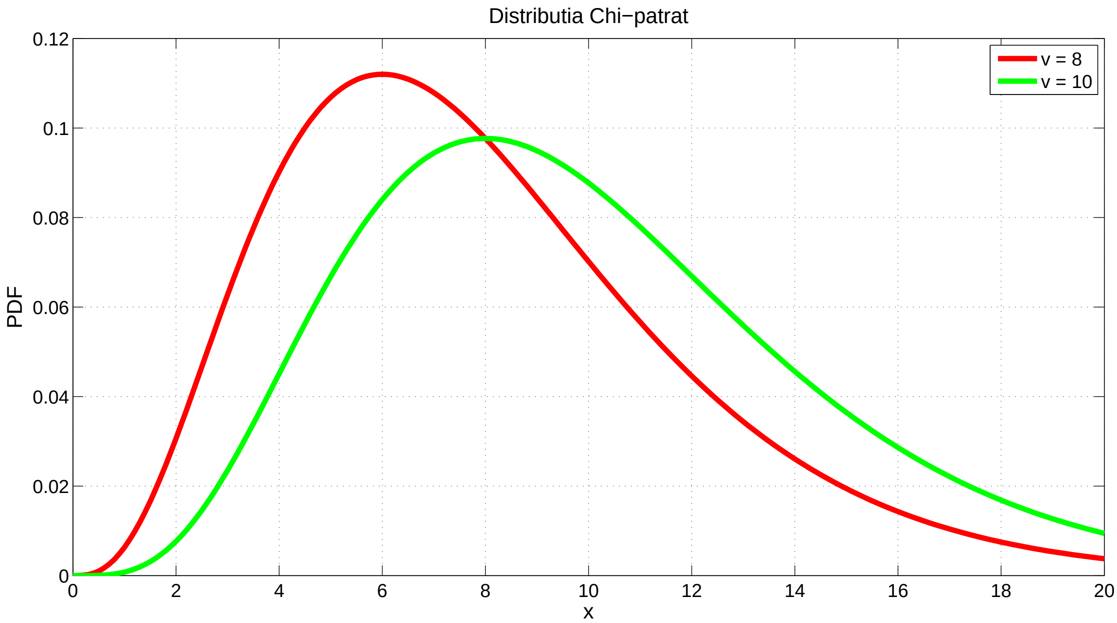
1.5.1 Densitatea si functia de repartitie

In teoria probabilitatilor si statistica, **distributia chi-patrat** cu v grade de libertate reprezinta distributia unei sume de patrate ale v variabile aleatorii, independente si normale. Densitatea distributiei este:

$$f(x|v) = \frac{x^{(v-2)/2}e^{-x/2}}{2^{v/2}\Gamma(v/2)}$$

unde $\Gamma(\,\cdot\,)$ este functia Gamma, iar v,x $\geq$ 0.

Exemplu:
$f(x|8) = \frac{x^3 e^{-4}}{16 \cdot 6} = \frac{x^3 e^{-4}}{96}$, pentru $\Gamma(n) = (n-1)!$
$f(x|10) = \frac{x^4 e^{-5}}{32 \cdot 24} = \frac{x^4 e^{-5}}{768}$



Cod **Matlab**:

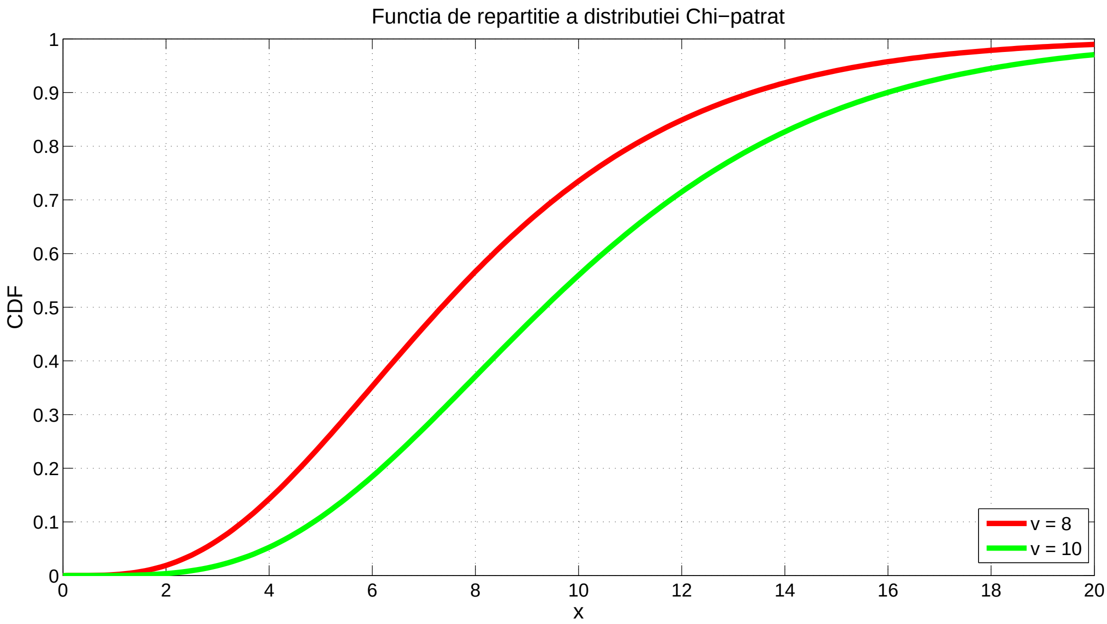
```
x=[0:0.1:20];
y=chi2pdf(x,8);
z=chi2pdf(x,10);
plot(x,y,'r',x,z,'g');
```

Se poate observa faptul ca probabilitatea ca x sa ia valori foarte mici sau foarte mari este foarte scazuta, probabilitatea cea mai mare fiind pentru valorile intermediare.
Functia de repartitie a distributiei chi-patrat:

$$F(x|v) = \int_0^x \frac{t^{(v-2)/2} e^{-t/2}}{2^{v/2}\Gamma(v/2)} dt$$

unde $\Gamma(\,\cdot\,)$ este functia Gamma.

Forma functiei de repartitie ne indica faptul ca x poate lua valori mici cu o probabilitate scazuta, in timp ce pentru valori mari probabilitatea este ridicata.

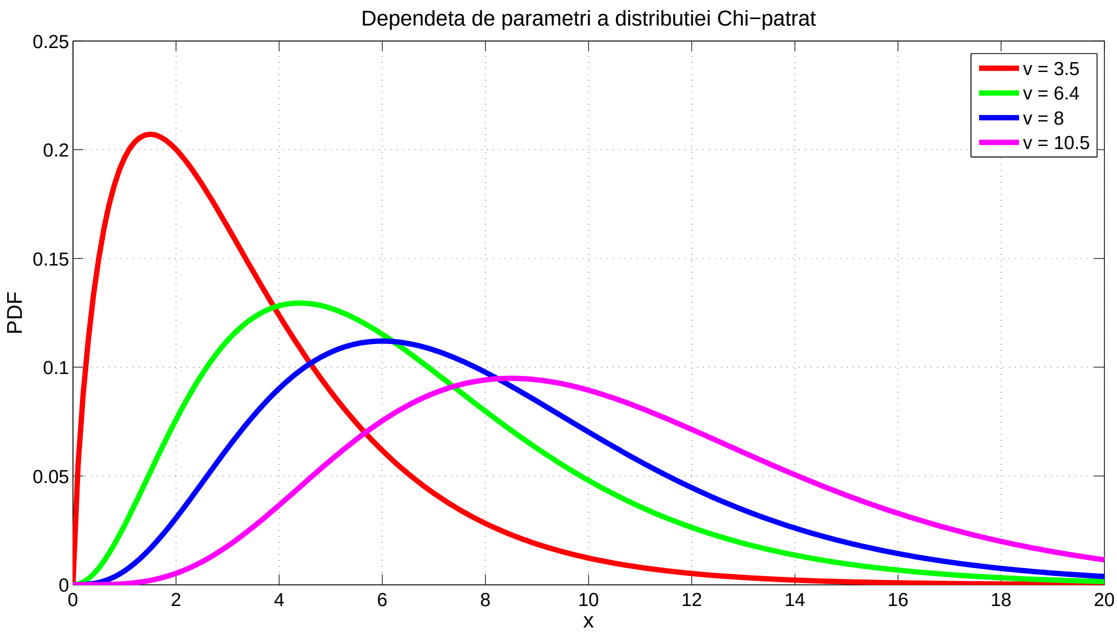


Cod **Matlab**:

```
x=[0:0.1:20];
y=chi2cdf(x,8);
z=chi2cdf(x,10);
plot(x,y,'r',x,z,'g');
```

1.5.2 Dependenta de parametri

Pentru a demonstra dependenta densitatii de parametrul v, vom alege valori diferite pentru acesta si vom reprezenta grafic diferentele.

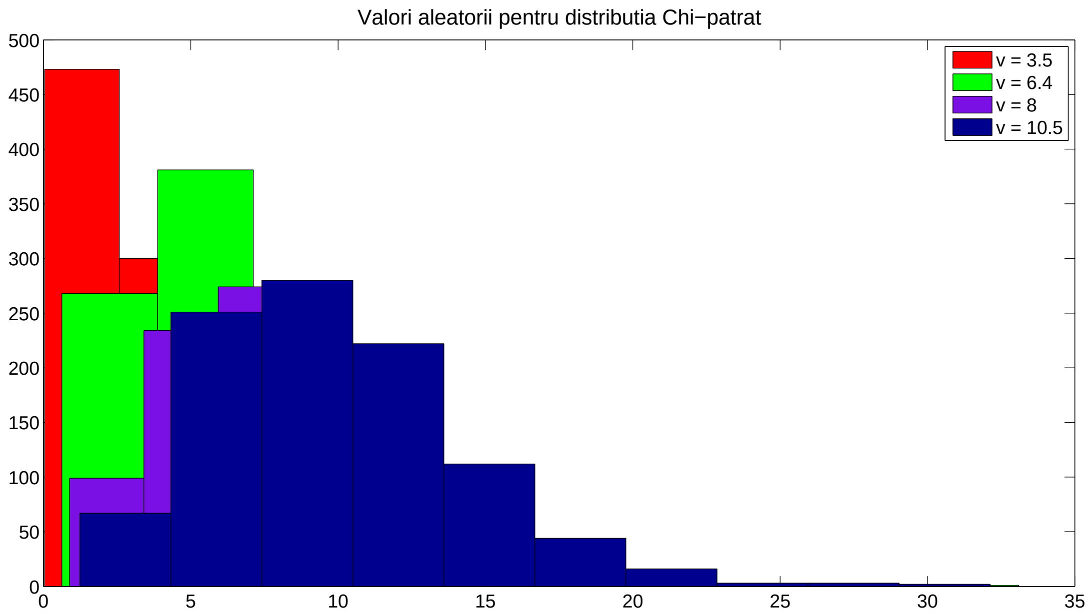


Cod **Matlab**:

```
x=[0:0.1:20];
y=chi2pdf(x,3.5);
z=chi2pdf(x,6.4);
a=chi2pdf(x,8);
b=chi2pdf(x,10.5);
plot(x,y,'r',x,z,'g',x,a'b',x,b,'m');
```

Cu cat variabila v creste, cu atat mai mult creste dispersia functiei, x avand astfel probabilitatea de a lua valori dintr-un interval mai mare.

Pentru a generaliza proprietatile densitatii, vom folosi functia *chi2rnd* pentru a genera valori aleatorii, apoi le vom reprezenta grafic.



Cod **Matlab**:

```
a = chi2rnd(3.5,1,1000);
b = chi2rnd(6.4,1,1000);
c = chi2rnd(8,1,1000);
d = chi2rnd(10,1,1000);
hist(a)
hold on
hist(b)
hist(c)
hist(d)
```

Se verifica astfel faptul ca odata cu v creste si probabilitatea ca x sa ia valori din ce in ce mai mari.

1.6 Distributia Gamma

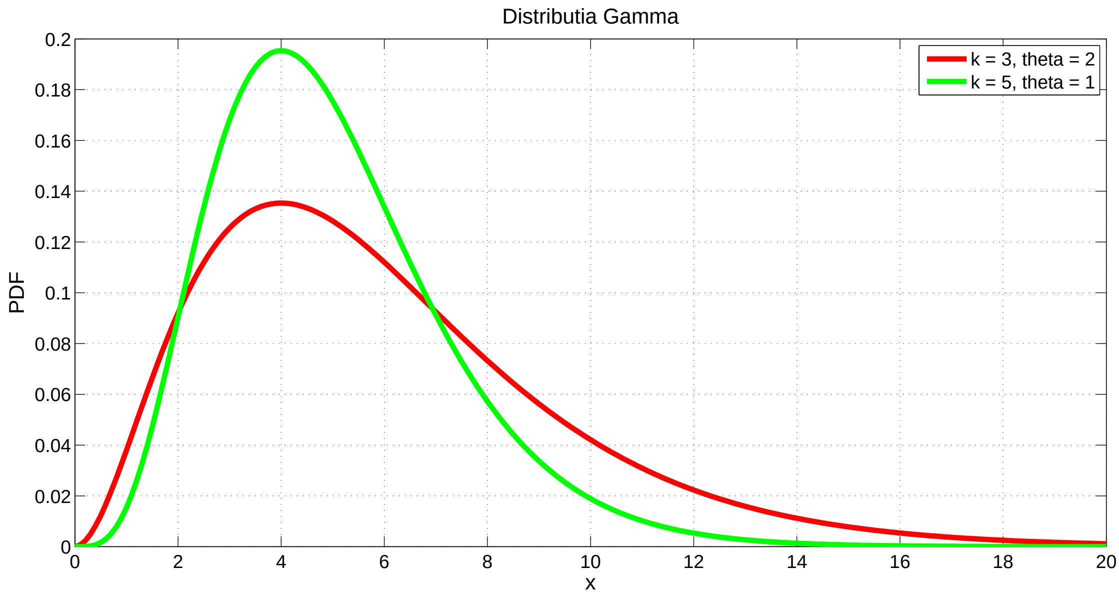
1.6.1 Densitatea si functia de repartitie

In domeniul probabilitatilor si statisticii, **distributia Gamma** este o familie de distributii continue cu doi parametri, unul de forma (k) si unul de scala (θ). Densitatea de repartitie :

$$f(x|k,\theta) = \frac{1}{\theta^k \Gamma(k)} x^{k-1} e^{\frac{-x}{\theta}}$$

pentru x,k,$\theta \geq 0$. Exemplu:

$f(x|3,2) = \frac{1}{16} x^2 e^{\frac{-x}{2}}$
$f(x|5,1) = \frac{1}{6} x^4 e^{-x}$



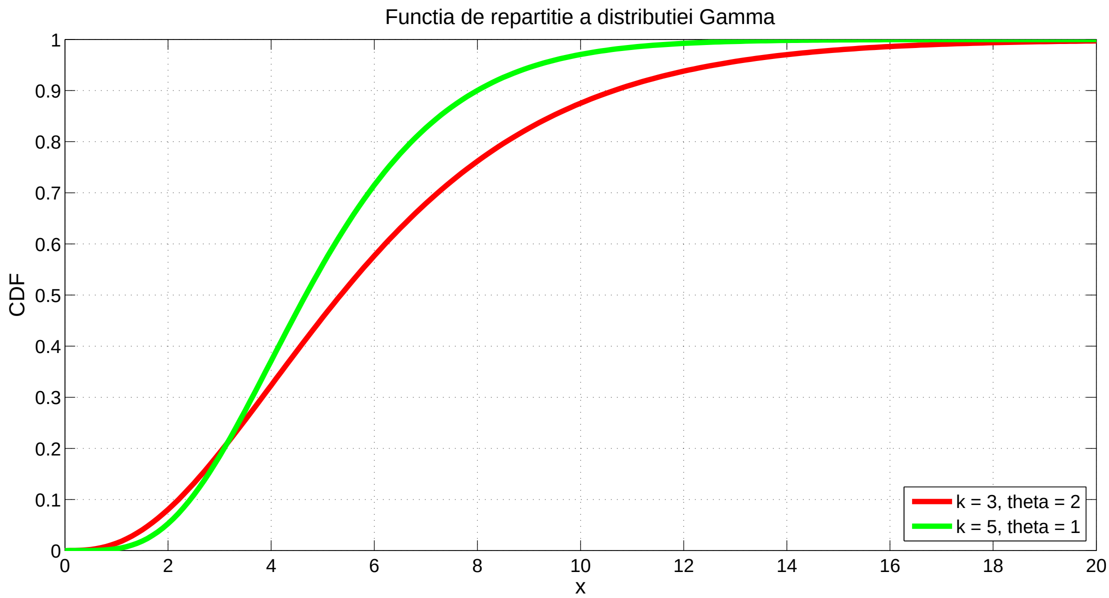
Cod **Matlab**:

```
x = [0:0.1:20];
y = gampdf(x,3,2);
z = gampdf(x,5,1);
plot(x,y,'r',x,z,'g')
```

Forma densitatii ne indica faptul ca valorile mici sunt cele mai probabile, in timp ce probabilitatea valorilor mari scade exponential.

Functia de repartitie a distributiei Gamma :

$$F(x|k,\theta) = \frac{1}{\theta^k \Gamma(k)} \int_0^x t^{k-1} e^{\frac{-t}{\theta}} dt$$



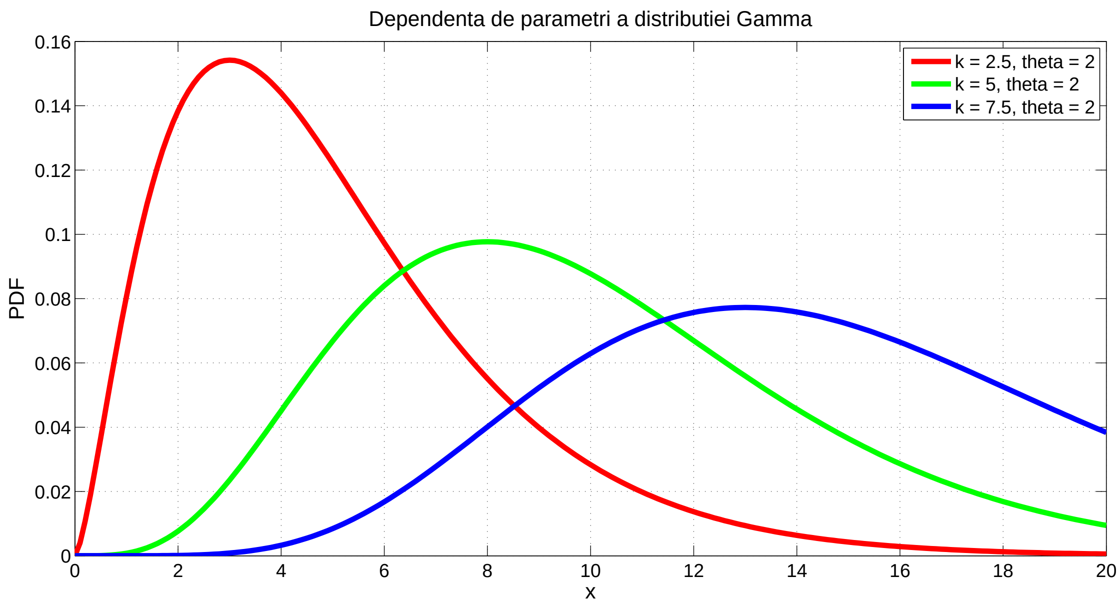
Cod **Matlab**:

```
x = [0:0.1:20];
y = gamcdf(x,3,2);
z = gamcdf(x,5,1);
plot(x,y,'r',x,z,'g')
```

Spre deosebire de densitate, functia de repartitie de indica faptul ca x poate lua cu o probabilitate mare valori mari, in timp ce valorile mici au o probabilitate scazuta.

1.6.2 Dependenta de parametri

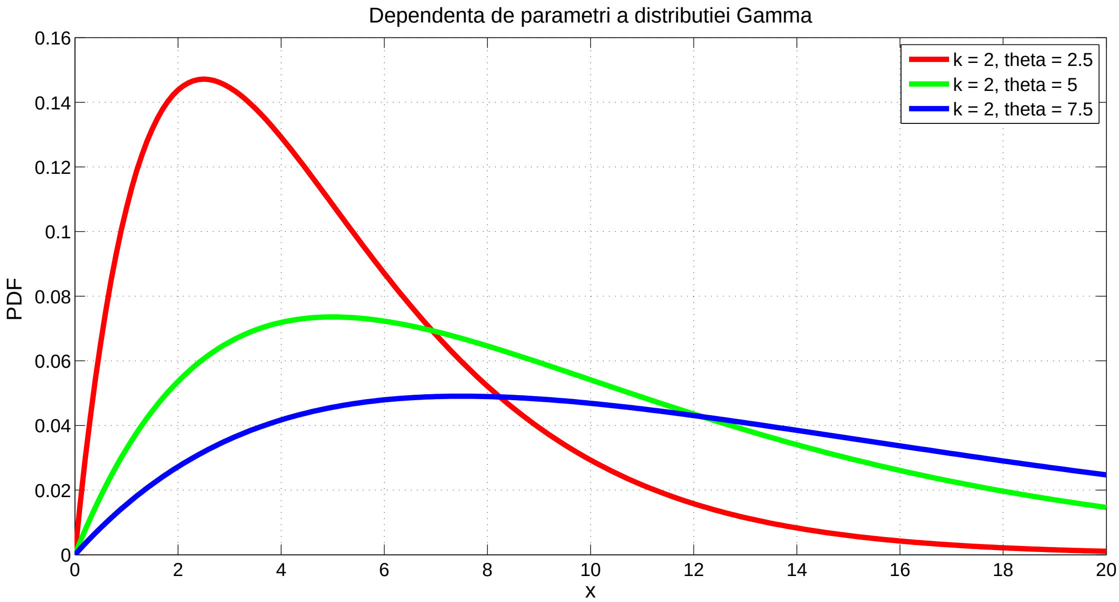
Pentru a demonstra dependenta de parametri, vom alege valori diferite pentru parametrii k si θ si vom reprezenta grafic rezultatele.



Cod **Matlab**:

```
x = [0:0.1:20];a = gampdf(x,2.5,2);
b = gampdf(x,5,2);
c = gampdf(x,7.5,2);
plot(x,a,'r',x,b,'g',x,c,'b')
```

Observam ca, cu cat variabila de forma k creste, cu atat creste si dispersia densitatii, astfel probabilitatea ca x sa ia valori mari este din ce in ce mai ridicata.

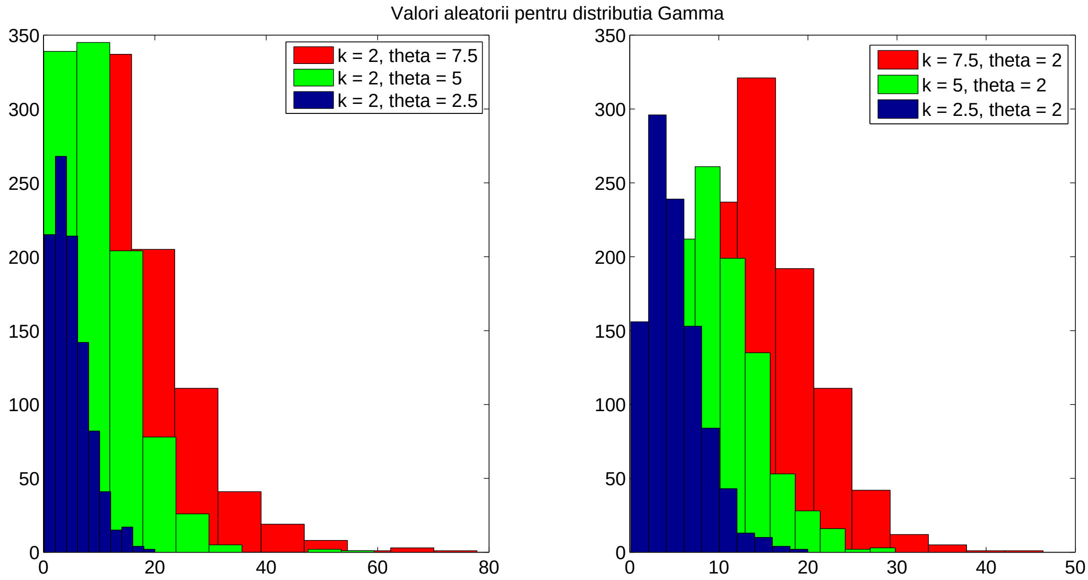


Cod **Matlab**:

```
x = [0:0.1:20];a = gampdf(x,2,2.5);
b = gampdf(x,2,5);
c = gampdf(x,2,7.5);
plot(x,a,'r',x,b,'g',x,c,'b')
```

In acest caz densitatea se comporta exact ca in primul caz, dispersia crescand odata cu valoarea θ.

Pentru a generaliza proprietatile densitatii vom alege valori aleatorii pentru parametrii precedenti folosind functia *gamrnd* si vom reprezenta grafic dispersia acestora.



Cod **Matlab**:

```
a = gamrnd(2,2.5,1,1000);
b = gamrnd(2,5,1,1000);
c = gamrnd(2,7.5,1,1000);
d = gamrnd(2.5,2,1,1000);
e = gamrnd(5,2,1,1000);
f = gamrnd(7.5,2,1,1000);
subplot(1,2,1)
hist(c)
hold on
hist(b)
hist(a)
subplot(1,2,2)
hist(f)
hold on
hist(e)
hist(d)
```

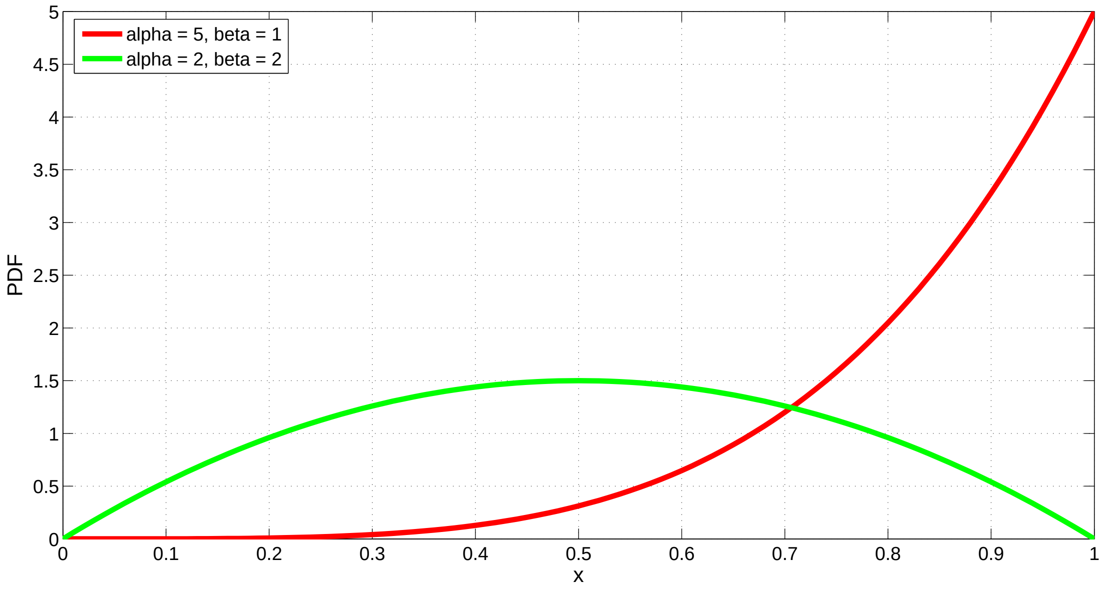
1.7 Distributia Beta

1.7.1 Densitatea si functia de repartitie

Distributia Beta este o familie de distributii continue de probabilitate definite pe intervalul [0,1], parametrizate de doi parametri pozitivi de forma, notati α si β si care apar ca exponenti ai variabilei aleatoare, controland forma distributiei.
Densitatea de repartitiei se defineste prin:

$$f(x|\alpha,\beta) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}$$

pentru x,$\alpha,\beta \geq 0$. unde B ($\cdot$) este functia Beta, $B(x,y) = \frac{(x-1)!(y-1)!}{x+y-1)}!$.
Exemplu:
$f(x|5,1) = \frac{x^4}{0.2}$
$f(x|2,2) = \frac{x(1-x)}{0.16}$

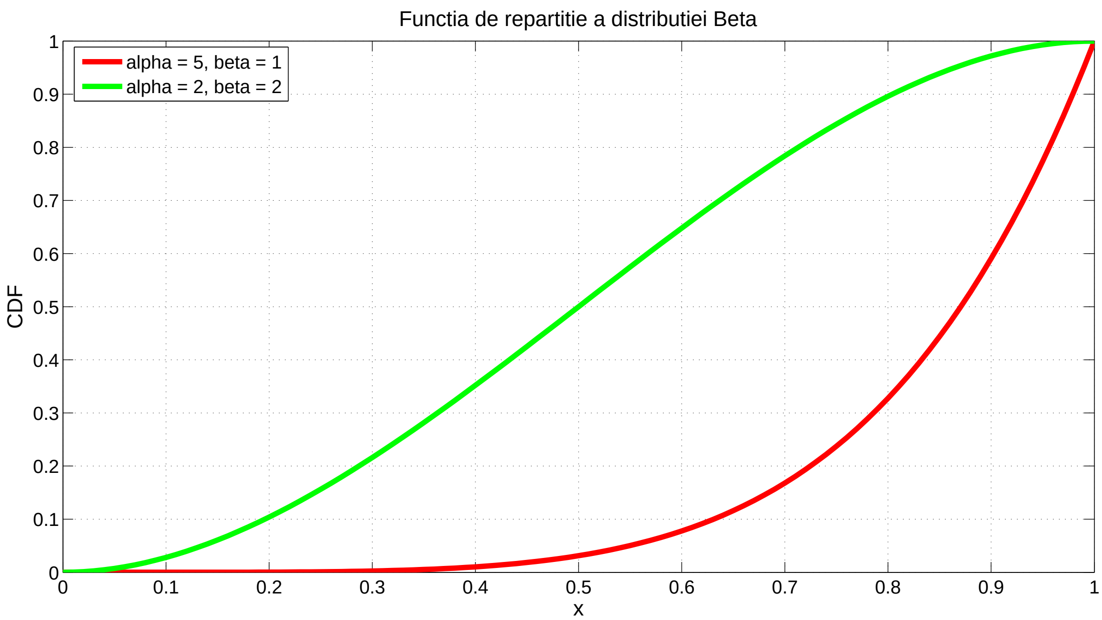


Cod **Matlab**:

```
x = [0:0.01:1];a = betapdf(x,5,1);
b = betapdf(x,2,2);
plot(x,y,'r',x,z,'g')
```

Observam ca pentru α mai mare, x are probabilitatea mai mare de a lua valori mari, in timp ce pentru β mai mare, probabilitatea ca x sa ia valori mari este scazuta. **Functia de repartitie** a distributiei Beta:

$$F(x|\alpha,\beta) = \frac{1}{B(\alpha,\beta)} \int_0^x t^{\alpha-1}(1-t)^{\beta-1}dt$$



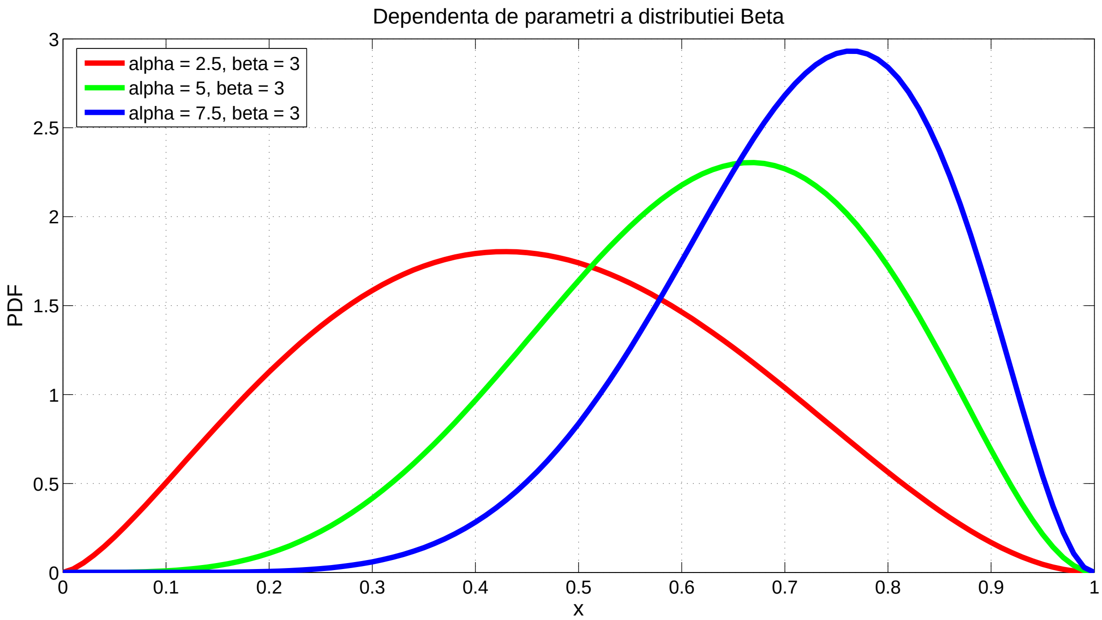
Cod **Matlab**:

```
x = [0:0.01:1];a = betacdf(x,5,1);
b = betacdf(x,2,2);
plot(x,y,'r',x,z,'g')
```

Putem observa ca variabila α este responsabila forma curbei pe care functia o ia, iar variabila β este responsabila de inaltimea functiei, astfel crescand si probabilitatea ca x sa ia anumite valori.

1.7.2 Dependenta de parametri

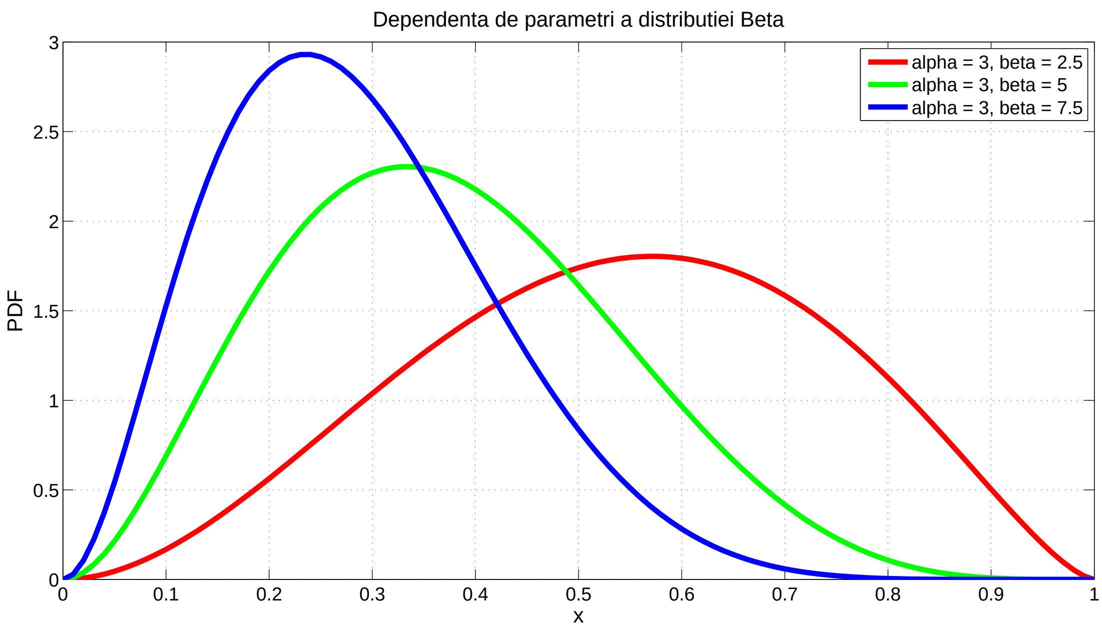
Pentru a demonstra dependenta de parametri vom alege valori diferite pentru parametrii α si β si le vom reprezenta grafic.



Cod **Matlab**:

```
x = [0:0.01:1];a = betapdf(x,2.5,3);
b = betapdf(x,5,3);
c = betapdf(x,7.5,3);
plot(x,a,'r',x,b,'g',x,c,'b')
```

Observam faptul ca variabila α controleaza dispersia densitatii, astfel pentru un α mai mare, x are probabilitatea mare de a lua valori dintr-un interval mai restrans.

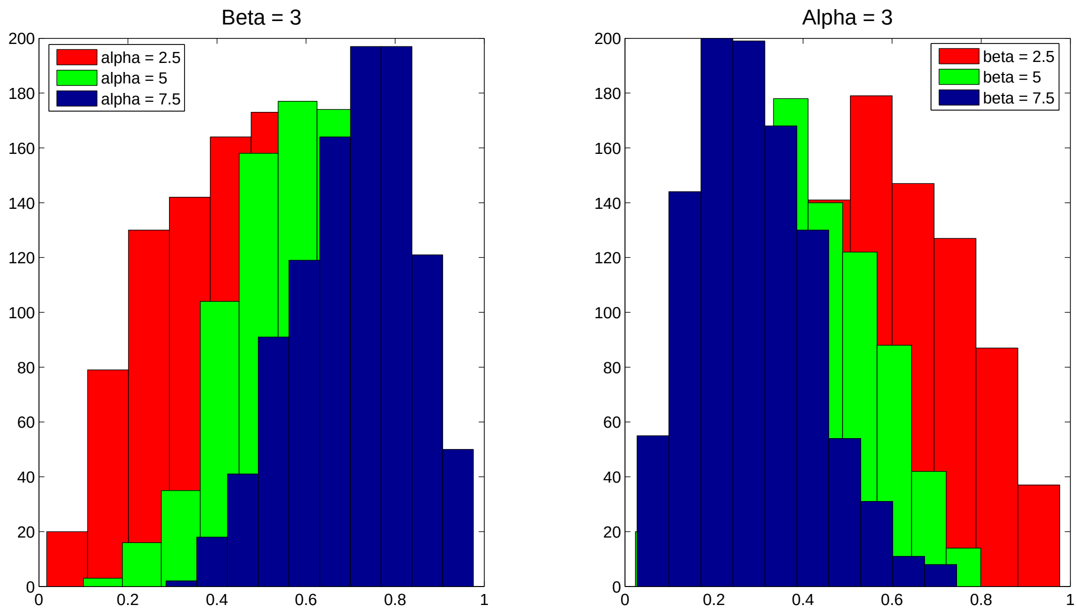


Cod **Matlab**:

```
x = [0:0.01:1];a = betapdf(x,3,2.5);
b = betapdf(x,3,5);
c = betapdf(x,3,7.5);
plot(x,a,'r',x,b,'g',x,c,'b')
```

In acest caz, cu cat β este mai mic, cu atat dispersia densitatii creste, x avand probabilitatea mai mare de a lua valori dintr-un interval mai larg.

Pentru a generaliza proprietatile distributiilor vom alege valori aleatorii pentru parametrii precedenti folosind functia *betarnd* si vom reprezenta grafic dispersia acestora.



Cod **Matlab**:

```
a = betarnd(2.5,3,1,1000);
b = betarnd(5,3,1,1000);
c = betarnd(7.5,3,1,1000);
d = betarnd(3,2.5,1,1000);
e = betarnd(3,5,1,1000);
f = betarnd(3,7.5,1,1000);
subplot(1,2,1)
hist(a)
hold on
hist(b)
hist(c)
subplot(1,2,2)
hist(d)
hold on
hist(e)
hist(f)
```

Pentru α mai mare, x ia cu o probabilitate mai mare valori mari, in timp ce pentru β mai mic, x are o probabilitate mai mare sa ia valori mici.

2 Distributii discrete si distributii continue

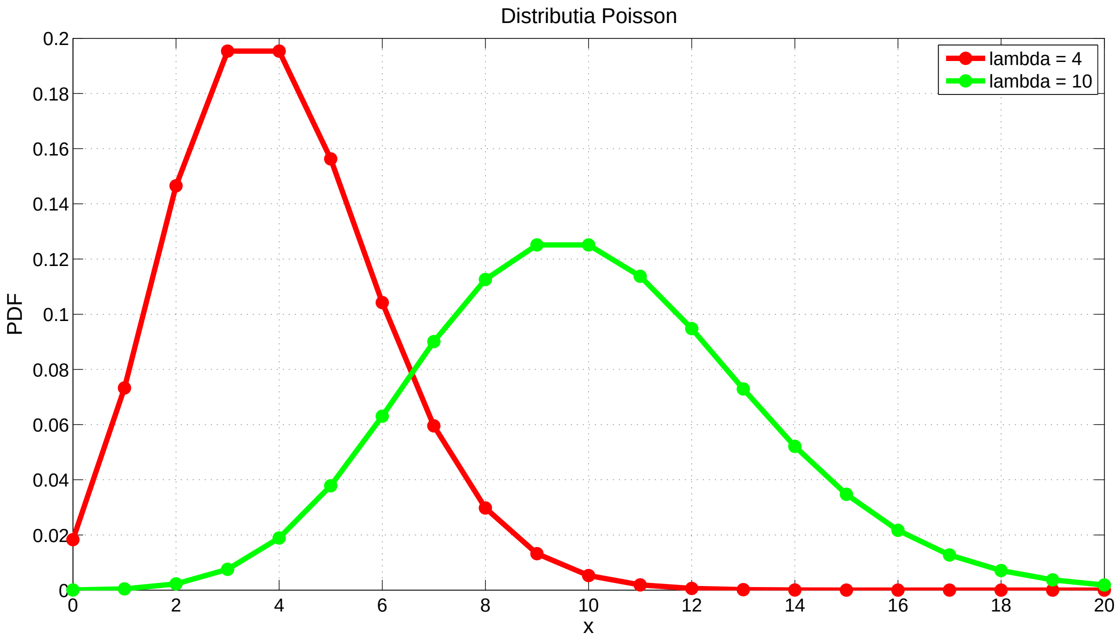
2.1 Distributia discreta Poisson

Distributia Poisson, denumita dupa matematicianul francez **Simon Denis Poisson**, este o distributie discreta de probabilitate care exprima probabilitatea unui numar dat de evenimente ce au loc intr-un interval fix de timp si/sau spatiu, cunoscand media acestora si timpul ultimului eveniment.
Densitatea de repartitie:

$$f(x|\lambda) = \frac{\lambda^x}{x!} e^{-\lambda} I_{0,1,\ldots}(x)$$

pentru $\lambda, x \geq 0$, x intreg.
Exemplu:

$f(x|4) = \frac{4^x}{24} e^{-4}$
$f(x|10) = \frac{10^x}{10!} e^{-10}$



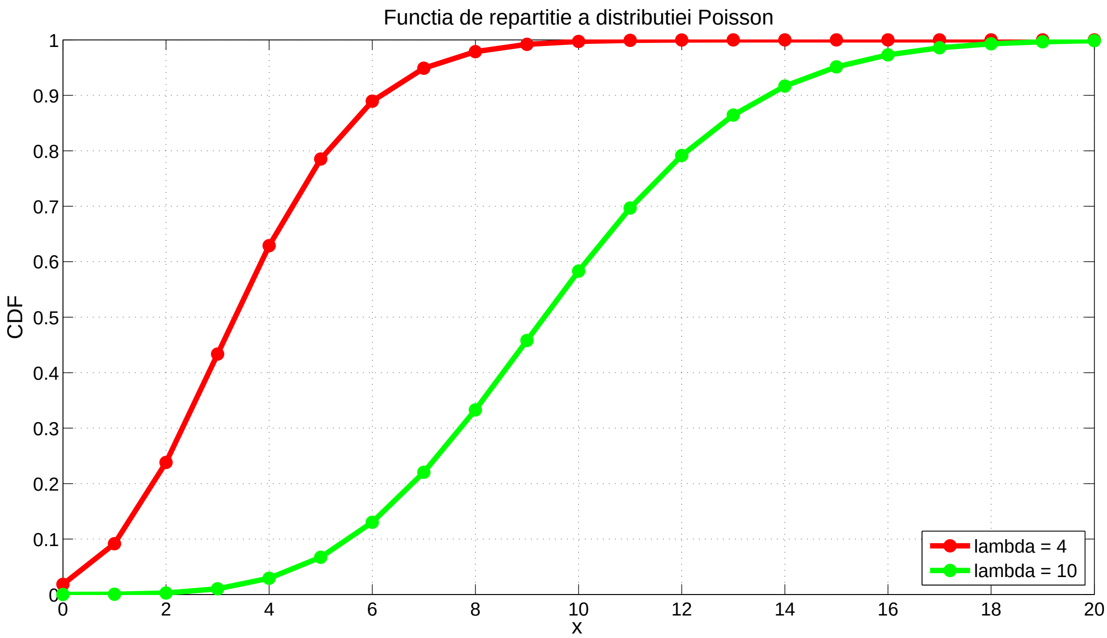
Cod **Matlab**:

```
x=[0:1:20];
y=poisspdf(x,4);
z=poisspdf(x,10);
plot(x,y,'r',x,z,'g')
```

Pe axa y se afla indicele de ocurenta, punctele fiind unite cu o linie continua doar pentru a oferi un grafic usor interpretabil. Astfel, pentru λ mai mic, probabilitatea ca x sa ia valori mici este ridicata, in caz contrar avand probabilitatea mare de a lua valori mai mari.

Functia de repartitie:

$$F(x|\lambda) = e^{-\lambda} \sum_{i=0}^{[x]} \frac{\lambda^i}{i!}$$



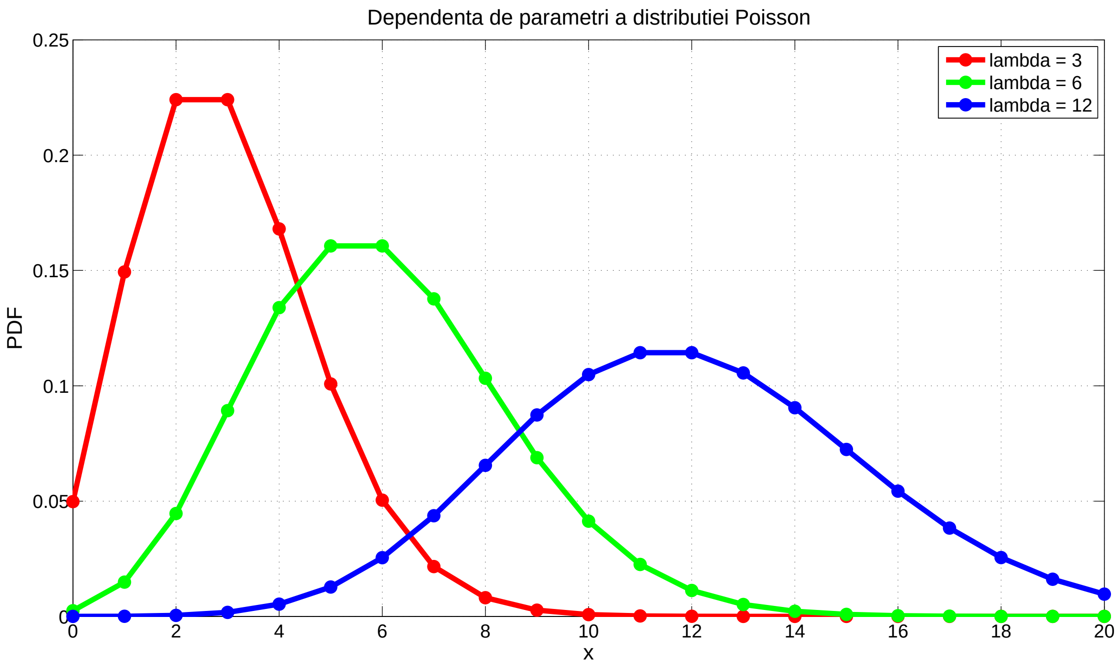
Cod **Matlab**:

```
x=[0:1:20];
y=poisscdf(x,4);
z=poisscdf(x,10);
plot(x,y,'r',x,z,'g')
```

Pentru λ mic, x poate lua cu probabilitate mare valori mici, in caz contrar probabilitatea fiind redusa pentru astfel de valori.

2.1.1 Dependenta de parametri

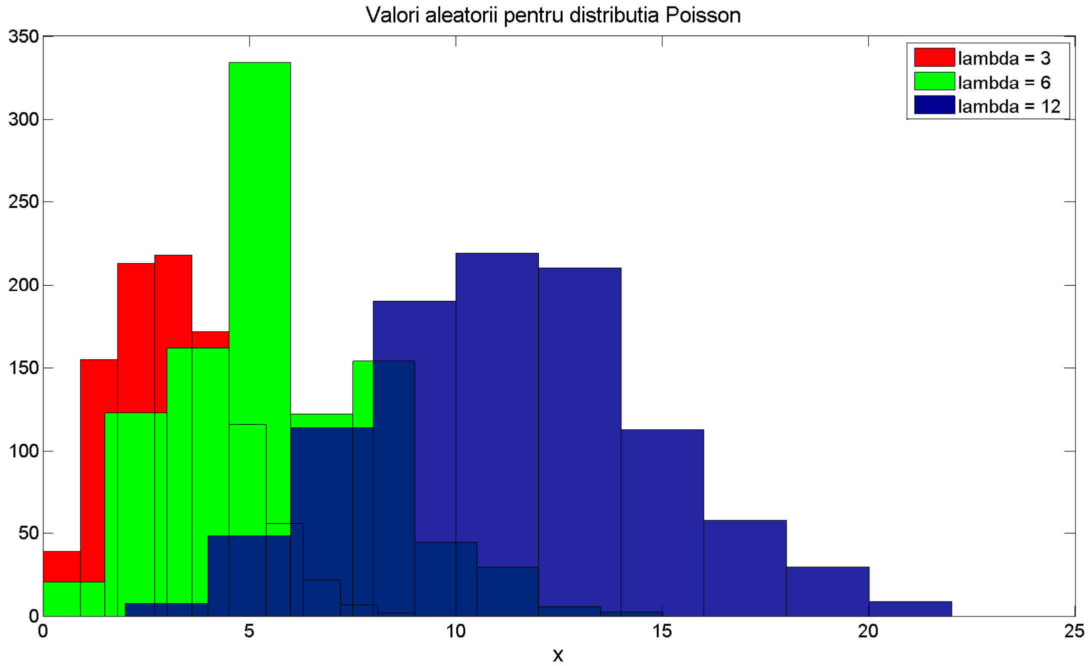
Pentru a reprezenta dependenta distributiei de parametrul λ, vom alege valori diferite pentru acesta si vom reprezenta grafic densitatile.



Cod **Matlab**:

```
x=[0:1:20];
y=poisspdf(x,3);
z=poisspdf(x,6);
a=poisspdf(x,12);
plot(x,y,'r',x,z,'g',x,a,'b')
```

Forma graficelor ne indica faptul ca, cu cat variabila λ este mai mare, cu atat dispersia e mai mare, iar variabila x poate lua cu o probabilitate mai mare valori mari.
Vom alege, in continuare, valori aleatorii pentru parametrul λ folosind functia *poissrnd* si vom reprezenta grafic dispersia acestora.



Cod **Matlab**:

```
x=poissrnd(3,1,1000);
y=poissrnd(6,1,1000);
z=poissrnd(12,1,1000);
hist(x)
hold on
hist(y)
hist(z)
```

Se respecta proprietatile generale: pentru λ mai mare, probabilitatea ca x sa ia valori mai mari este ridicata.

2.2 Distributia continua Rayleigh

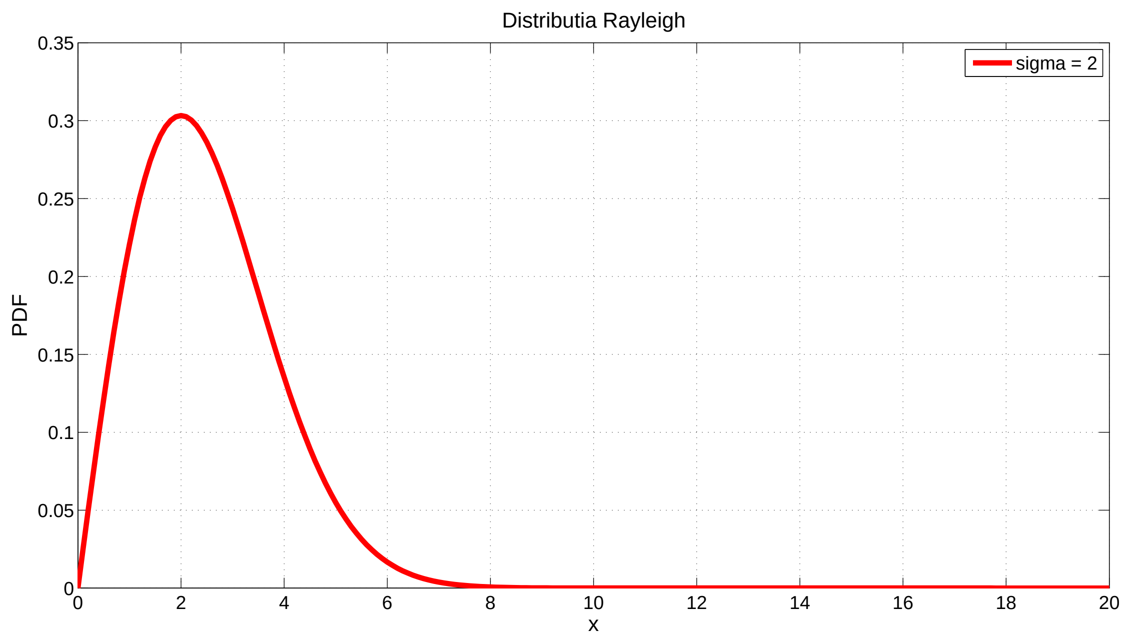
2.2.1 Densitatea si functia de repartitie

In statistica si teoria probabilitatilor, **distributia Rayleigh** este o distributie de probabilitate continua. Ea poate aparea cand un vector bidimensional are elemente ce sunt intr-o distributie normala, necorelate si cu varianta egala. Modulul vectorului va avea in acest caz o distributie Rayleigh. **Densitatea de repartitie**:

$$f(x|\sigma) = \frac{x}{\sigma^2} e^{\frac{-x^2}{2\sigma^2}}$$

pentru $x \in [0, \infty)$ si $\sigma > 0$.
Exemplu:

$f(x|2) = \frac{x}{4} e^{\frac{-x^2}{8}}$
$f(x|4) = \frac{x}{16} e^{\frac{-x^2}{32}}$



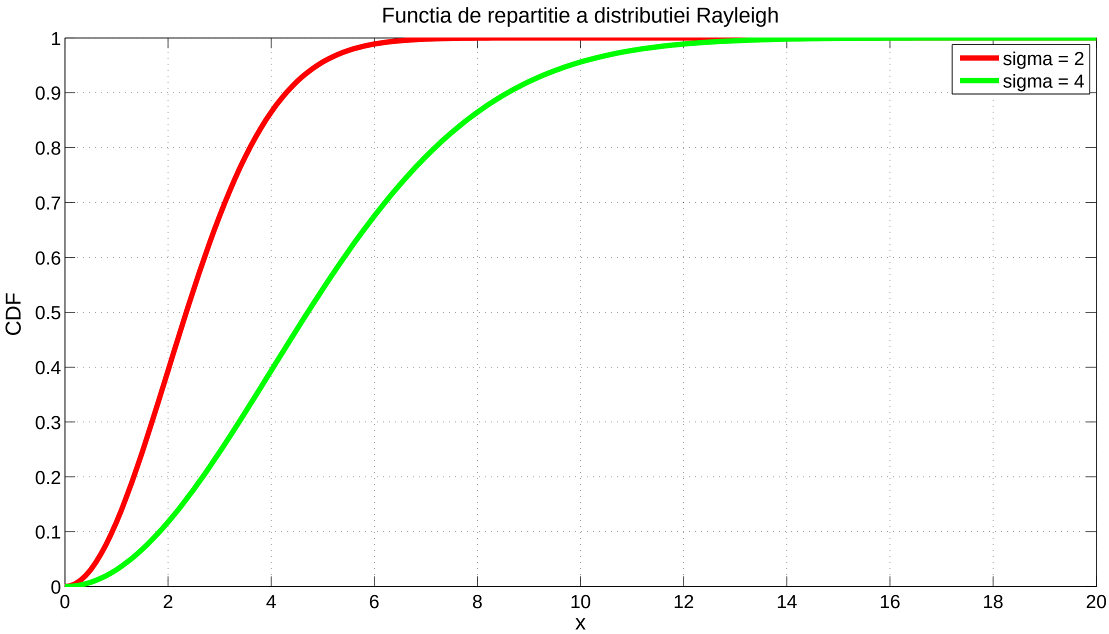
Cod **Matlab**:

```
x=[0:0.1:20];
y=raylpdf(x,2);
plot(x,y,'r')
```

Forma densitatii ne indica faptul ca x poate lua cu o probabilitate mare valori mici, in timp ce pentru valorile mari probabilitatea este redusa.
Functia de repartitie pentru distributia Rayleigh:

$$F(x|b) = \int_0^x \frac{t}{\sigma^2} e^{\frac{-t^2}{2\sigma^2}} dt = 1 - e^{\frac{-x^2}{2\sigma^2}}$$

pentru $x \in [0, \infty)$.



Cod **Matlab**:

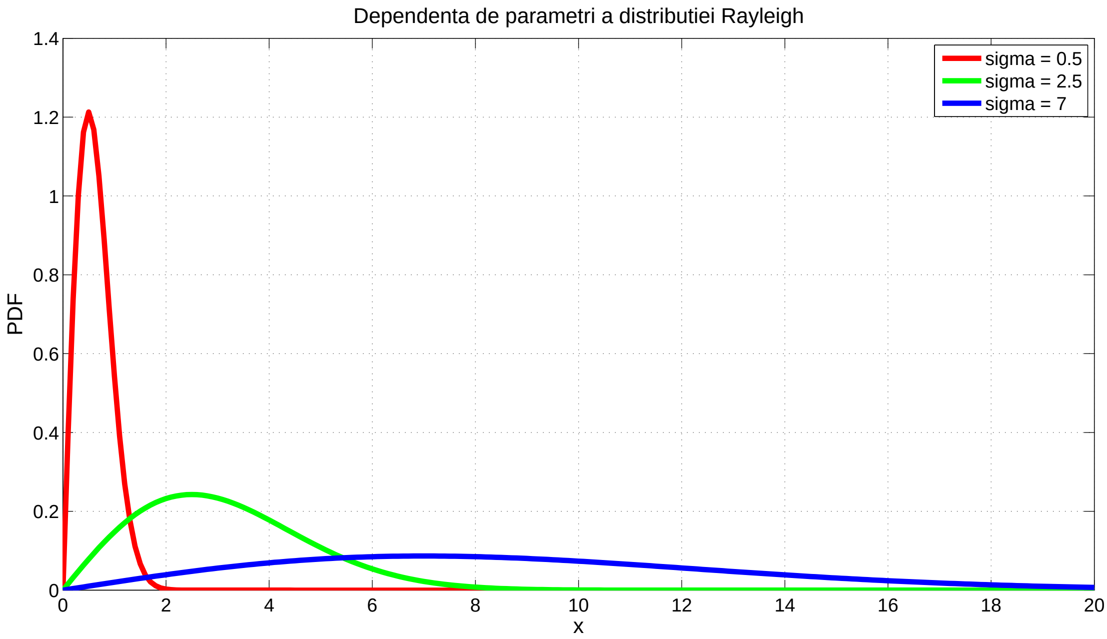
```
x=[0:0.1:20];
y=raylcdf(x,2);
z=raylcdf(x,4);
plot(x,y,'r',x,z,'g')
```

Observam ca functia de repartitie are o probabilitate ridicata pentru valori mari.

2.2.2 Dependenta de parametri

Pentru a demonstra dependeta de parametri, vom alege valori diferite pentru parametrul σ si vom reprezenta grafic diferentele.

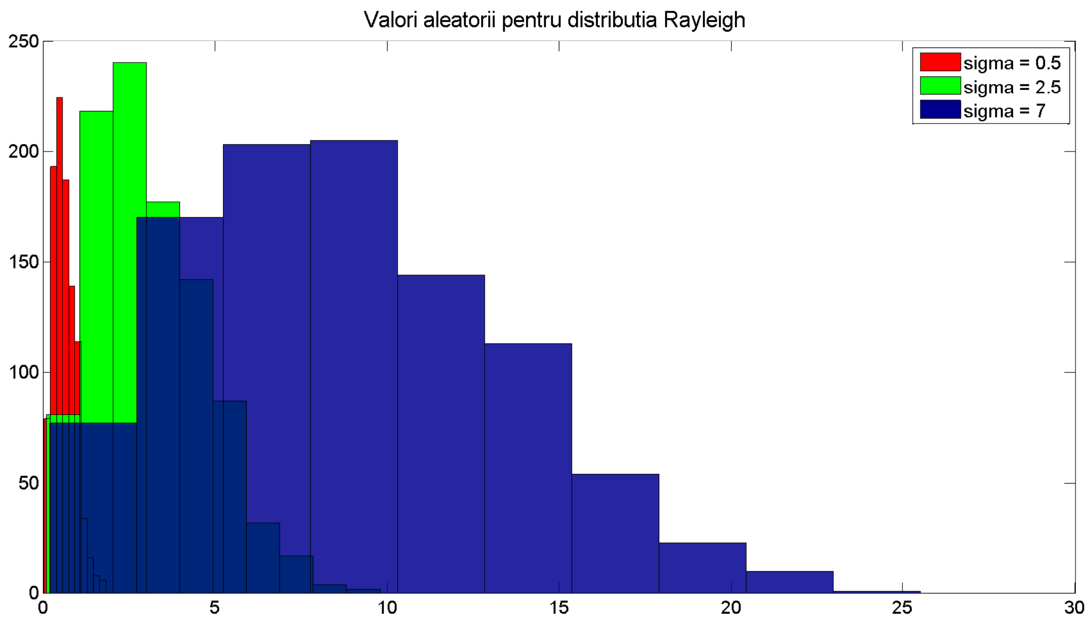
Vom observa faptul ca, atunci cand σ ia valori mici, probabilitatea ca x sa ia valori mici este ridicata. Atunci cand σ ia valori mari, x are probabilitatea redusa de a lua atat valori mici, cat si valori mari.



Cod **Matlab**:

```
x=[0:0.1:20];
y=raylpdf(x,0.5);
z=raylpdf(x,2.5);
a=raylpdf(x,7);
plot(x,y,'r',x,z,'g',x,a,'b')
```

Pentru a generaliza proprietatile distributiei vom alege valori aleatorii pentru parametrul σ si vom reprezenta grafic dispersia acestora.



Cod **Matlab**:

```
a=raylrnd(0.5,1,1000);
b=raylrnd(2.5,1,1000);
c=raylrnd(7,1,1000);
;hist(a)
hold on
hist(b)
hist(c)
```

2.3 Exercitii

Pentru fiecare subpunct, verificati proprietatile densitatii de repartitie, calculati functia de repartitie, calculati valoarea medie M(X) si varianta Var(X).

2.3.1 X $\sim$ Unif(a=1,b=5)

Densitatea de repartitie:

$$\omega_\xi(x) = \begin{cases} \frac{1}{b-a}, & a \le x \le b \\ 0, & \text{in rest} \end{cases}$$

$\int_{-\infty}^{\infty} f_x(x)dx = \int_a^b \frac{1}{b-a}dx = \frac{1}{b-a}\int_a^b dx = \frac{b-a}{b-a} = \frac{5-1}{5-1} = \frac{4}{4} = 1$

Functia de repartitie:

- daca $x < a$, atunci : $F_X(x) = P(X \le x) = 0$.
- daca $a \le x \le b$, atunci: $F_X(x) = P(X \le x) = \int_{-\infty}^{x} f_X(t)dt = \int_a^x \frac{1}{b-a}dt = \frac{1}{b-a} \cdot t\Big|_a^x = (x-a)(b-a) = (x-1)(5-1) = 4(x-1)$.
- daca $x > b$, atunci : $F_X(x) = P(X \le x) = 1$.

M[X] $= \int_{-\infty}^{\infty} x f_X(x)dx = \int_a^b x\frac{1}{b-a}dx = \frac{1}{b-a}\int_a^b xdx = \frac{1}{b-a}(\frac{1}{2}x^2)\Big|_a^b = \frac{1}{b-a}\frac{1}{2}(b^2-a^2) = \frac{(b-a)(b+a)}{2(b-a)} = \frac{b+a}{2}$

Var[X] $= M[X^2] - M[X]^2$
$M[X^2] = \frac{b^2+ab+a^2}{3}$
$M[X]^2 = \frac{b^2+2ab+a^2}{4}$
$Var[X] = \frac{b^2+ab+a^2}{3} - \frac{b^2+2ab+a^2}{4} = \frac{4b^2+4ab+fa^2-3b^2-6ab-3a^2}{12} = \frac{(4-3)b^2+(4-6)ab+(4-3)a^2}{12} = \frac{b^2-2ab+a^2}{12} = \frac{(b-a)^2}{12}$

2.3.2 X $\sim$ Exp($\lambda = 0.5$)

Densitatea de repartitie:

$$f_x(x) = \begin{cases} \lambda e^{-\lambda x}, & x \ge 0 \\ 0, & x < 0 \end{cases}$$

$\int_{-\infty}^{\infty} f_x(x)dx = \int_0^{\infty} \lambda e^{-\lambda x}dx = -e^{-\lambda x}\Big|_0^{\infty} = 0 - (-1) = 1.$

$f_x(x|0.5) = \frac{1}{2}e^{\frac{-x}{2}}.$

Functia de repartitie:

- $x < 0$, atunci: $F_x(X) = P(X \leq x) = 0$
- $x \geq 0$, atunci: $F_x(X) = P(X \leq x) = \int_0^{-\infty} x f_x(t)dt = \int_0^x \lambda e^{-\lambda t} dt = -e^{-\lambda t}\Big|_0^x = 1 - e^{-\lambda x}$

$F_x(X|0.5) = 1 - e^{\frac{-x}{2}}$

$\mathbf{M[X]} = \int_0^\infty x f(x)dx = \int_0^\infty x\lambda e^{-\lambda x} dx = (-xe^{-\lambda x})\Big|_o^\infty + \int_0^\infty e^{-\lambda x} dx = (0 - 0) + (-\frac{1}{\lambda} e^{-\lambda x})\Big|_0^\infty = 0 + (0 + \frac{1}{\lambda}) = \frac{1}{\lambda}$

$M[X|0.5] = \frac{1}{\frac{1}{2}} = 2.$

Var[X] $= \int_0^\infty x^2 \lambda e^{-\lambda x} dx - \frac{1}{\lambda^2} = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$
Var[X|0.5]$=\frac{1}{\frac{1}{4}} = 4.$

2.3.3 X $\sim$ Norm(m=0, $\sigma = 1$)

Densitatea de repartitie:

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x-m^2}{2\sigma^2}}$$

$f_X(x|0,1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x-1}{2}}.$

$\int_{-\infty}^\infty f_X(x)dx = (2\pi)^{-1/2)} \int_{-\infty}^\infty e^{-\frac{1}{2}x^2} dx = (2\pi)^{-1/2} \int_0^\infty e^{-\frac{1}{2}x^2} dx =$
$(2\pi)^{-1/2} 2 \left(\int_0^\infty e^{-\frac{1}{2}x^2} dx \int_0^\infty e^{-\frac{1}{2}x^2} dx \right)^{1/2} = (2\pi)^{-1/2} 2 \left(\int_0^\infty e^{-\frac{1}{2}x^2} dx \int_0^\infty e^{-\frac{1}{2}y^2} dy \right)^{1/2} =$
$(2\pi)^{-1/2} \left(\int_0^\infty \int_0^\infty e^{-\frac{1}{2}(x^2+y^2)} dydx \right)^{1/2} = (2\pi)^{-1/2} 2 \left(\int_0^\infty \int_0^\infty e^{-\frac{1}{2}(x^2+s^2x^2)} xdsdx \right)^{1/2} =$
$(2\pi)^{-1/2} 2 \left(\int_0^\infty \int_0^\infty e^{-\frac{1}{2}x^2(1+s^2)} xdxds \right)^{1/2} = (2\pi)^{-1/2} 2 \left(\int_0^\infty \left[-\frac{1}{1+s^2} e^{-\frac{1}{2}x^2(1+s^2)} \right]\Big|_0^\infty ds \right)^{1/2} =$
$(2\pi)^{-1/2} 2 \left(\int_0^\infty \left(0 + \frac{1}{1+s^2} \right) ds \right)^{1/2} = (2\pi)^{-1/2} 2 \left(\int_0^\infty \frac{1}{1+s^2} ds \right)^{1/2} = (2\pi)^{-1/2} 2 (arctan(s)\Big|_0^\infty)^{1/2} =$
$(2\pi)^(-1/2) 2 (arctan(\infty) - arctan(0))^{1/2} = (2\pi)^{-1/2} 2 \left(\frac{\pi}{2} - 0 \right)^{1/2} = 2^{-1/2} \pi^{-1/2} 2\pi - 1/22 - 1/2 =$
$1.$

Functia de repartitie:

$$F_X(x) = \int_{-\infty}^x f_X(t)dt$$

Integrala nu poate fi exprimata in functii elementare, astfel fiind nevoie de un tabel de valori pentru a calcula functia.

M[X] $= \int_{-\infty}^{\infty} x f_X(x)dx = (2\pi)^{-1/2}\int_{-\infty}^{\infty} xe^{-\frac{1}{2}x^2}dx = (2\pi)^{-1/2}\int_{-\infty}^{0} xe^{-\frac{1}{2}x^2}x + (2\pi)^{-1/2}\int_{0}^{\infty} xe^{-\frac{1}{2}x^2}dx = (2\pi)^{-1/2}(-e^{-\frac{1}{2}x^2}\big|_{-\infty}^{0} + (2\pi)^{-1/2}(-e^{-\frac{1}{2}x^2}\big|_{0}^{\infty} = (2\pi)^{-1/2} - (2\pi)^{-1/2} = 0$

Var[X] $= M[X^2] - M[X]^2 = \int_{-\infty}^{\infty} f_X(x)dx - 0^2 = 1 - 0 = 1$

2.3.4 X ∼ Binomial(n=1, p=5)

Densitatea de repartitie:

$p_X(x) = \begin{cases} \binom{n}{x}p^x(1-p)^{n-x}, & x \in \{0, 1, ..., n\} \\ 0, & \text{in rest} \end{cases}$ unde $\binom{n}{x} = \frac{n!}{x!(n-x)!}$ este coeficientul binomial.

$p_X(x) = \binom{1}{x}5^x(-4)^{1-x} = \frac{1}{x!(1-x)!}5^x(-4)^{1-x}$.

$\sum_{x \in R_x} p_X(x) = \sum_{x=0}^{n}\binom{n}{x}p^x(1-p)^{n-x} = [p + (1-p)]^n = 1^n = 1$
unde am folosit formula expansiunii binomiale: $(a+b)^n = \sum_{x=0}^{n}\binom{n}{x}a^x b^{n-x}$.

M[X]$=M\left[\sum_{i=1}^{n} Y_i\right] = \sum_{i=1}^{n} M[Y_i] = \sum_{i=1}^{n} p = np.$

M[X|1,5] $= 1 \cdot 5 = 5$

Var[X]$=Var\left[\sum_{i=1}^{n} Y_i\right] = \sum_{i=1}^{n} Var[Y_i] = \sum_{i=1}^{n} p(1-p) = np(1-p)$

Var[X | 1,5]$=1 \cdot 5(1-5) = -20$

2.3.5 X ∼ Poisson($\lambda = 4$)

Densitatea de repartitie:

$$p_X(x) = \begin{cases} e^{-\lambda}\frac{1}{x!}\lambda^x, & x \in \mathbb{Z} \\ 0, & \text{in rest} \end{cases}$$

$p_X(x) = e^{-4}\frac{1}{x!}4^x$

$P(X \geq x) = P(\tau_1 + ... + \tau_x \leq 1)$
$P(X \geq x) = 1 - P(X < x) = 1 - P(x \leq x-1) = 1 - \sum_{j=0}^{x-1} P(X = j) = 1 - \sum_{j=0} x - 1p_X(j) = 1 - \sum_{j=0}^{x-1}\frac{\lambda^j}{j!}e^{-\lambda} = P(\tau_1 + ... + \tau_x \leq 1).$

Functia de repartitie:

$$F_X(x) = \begin{cases} e^{-\lambda} \sum_{s=0}^{[x]} \frac{1}{s!} \lambda^s, & x \geq 0 \\ 0, & \text{in rest} \end{cases}$$

$F_X(x|4) = e^{-4} \sum_{s=0}^{[x]} \frac{1}{s!} 4^s$.

$F_X(x) = P(X \leq x) = \sum_{s=0}^{[x]} P(X = s) = \sum_{s=0}^{[x]} p_X(s) = \sum_{s=0}^{[x]} e^{-\lambda} \frac{1}{s!} \lambda^s = e^{-\lambda} \sum_{s=0}^{[x]} \frac{1}{s!} \lambda^s$

M[X] $= \sum_{x \in R_x} x p_X(x) = \sum_{x=0}^{\infty} x e^{-\lambda} \frac{1}{x!} \lambda^x = 0 + \sum_{x=1}^{\infty} x e^{-\lambda} \frac{1}{x!} \lambda^x = \sum_{y=0}^{\infty} (y+1) e^{-\lambda} \frac{1}{(y+1)!} \lambda^{y+1} = \sum_{y=0}^{\infty} (y+1) e^{-\lambda} \frac{1}{(y+1)y!} \lambda \lambda^y = \lambda \sum_{y=0}^{\infty} e^{-\lambda} \frac{1}{y!} \lambda^y = \lambda \sum_{y=0}^{\infty} p_y(y) = \lambda$.

M(X| 4)= 4.

Var[X]$=M[X^2] - M[X]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$.

Var[X|4] = 4.