Laborarator 10 – Probabilități și Statistică Matematică

REPARTITII CLASICE

Obiectivul acestui laborator este de a prezenta o parte din cele mai cunoscute repartitii continue si de a rezolva câteva probleme cu ajutorul lor. R pune la displozitie majoritatea repartitiilor uzuale. Tabelul de mai jos prezintă numele si parametrii acestora:

Tab. 1: Numele si parametrii repartitiilor uzuale in R

Repartiția	Nume	Parametrii	Valori prestabilite
Uniformă	`unif`	`min`, `max`	`min = 0`, `max = 1`
Normală	`norm`	`mean`, `sd`	`mean = 0`, `sd = 1`
Log-Normală	`lnorm`	`mean`, `sd`	`mean = 0`, `sd = 1`
Exponențială	`exp`	`rate` (=1/mean)	`rate = 1`
Cauchy	`cauchy`	`location`, `scale`	`location = 0`, `scale = 1`
Gamma	`gamma`	`shape`, `rate` (=1/scale)	`rate = 1`
Beta	`beta`	`shape1`, `shape2`	
Student	`t`	`df`	
Chi-Squared	`chisq`	`df`	
Fisher	`f`	`df1`, `df2`	
Weibull	`weibull`	`shape`	

Pentru fiecare repartitie, există patru comenzi în R prefixate cu literele *d, p, q* si *r* si urmate de numele repartitiei (coloana a 2-a). De exemplu *dnorm, pnorm, qnorm* si *rnorm* sunt comenzile corespunzătoare repartit iei normale pe când *dunif, punif, qunif* si *runif* sunt cele corespunzătoare repartit iei uniforme.

• ***dnume***: calculează densitatea atunci când vorbim de o variabilă continuă sau functia de masă atunci când avem o repartitie discretă ($P(X = k)$)

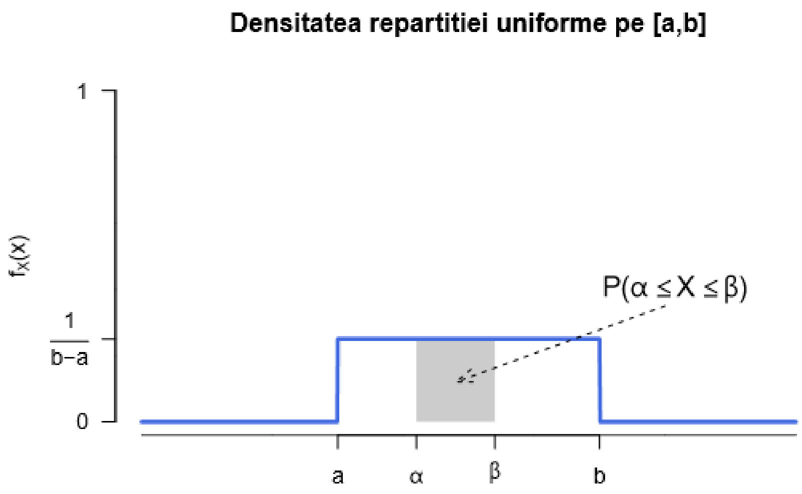
• ***pnume***: calculează functia de repartitie, i.e. $F(x) = P(X \leq x)$

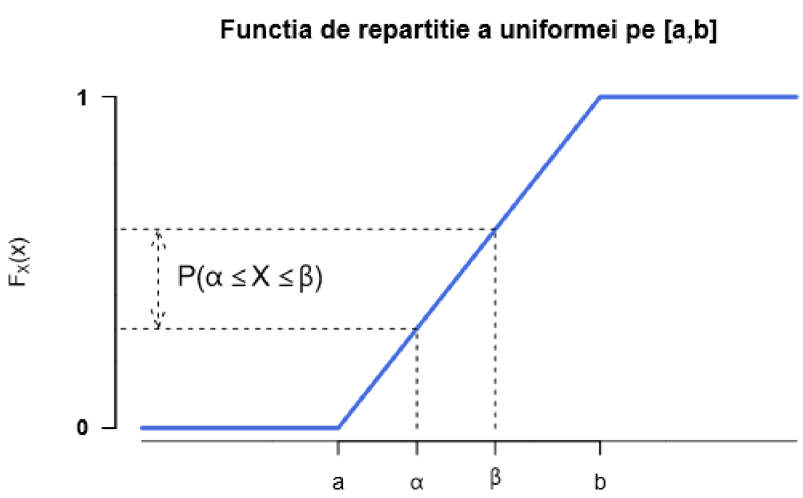
• ***qnume***: reprezintă functia cuantilă, cu alte cuvinte valoarea pentru care functia de repartitie are o anumită probabilitate; în cazul continuu, dacă $pnume(x) = p$ atunci $qnume(p) = x$ iar în cazul discret întoarce cel mai mic întreg u pentru care $P(X \leq u) \geq p$.

• ***rnume***: generează observatii independente din repartitia data.

1. *Repartitia Uniformă*

Variabilele aleatoare repartizate uniform joacă un rol important în teoria simulării variabilelor aleatoare datorită rezultatului lui Paul Levy numit *Teorema de universalitate a repartitiei uniforme*.





În R putem să:

• generăm observatii independente din repartitia $U([a, b])$ $(e.g.\, a = 3\ si\ b = 5)$

runif(10, 3, 5)

[1] 4.534364 4.214851 3.291478 3.192033 3.586674 3.395941 4.534129

[8] 3.262965 3.396016 4.961392

• calculăm densitatea unei variabile aleatoare repartizate uniform pe $[a, b]$ în diferite puncte

dunif(c(3.1, 3.7, 3.95, 4.86), 3, 5)

[1] 0.5 0.5 0.5 0.5

• calculăm functia de repartit ie a unei variabile repartizate uniform pe $[a, b]$ pentru diferite valori

punif(c(3.1, 3.7, 3.95, 4.86), 3, 5)

[1] 0.050 0.350 0.475 0.930

APLICATIE:

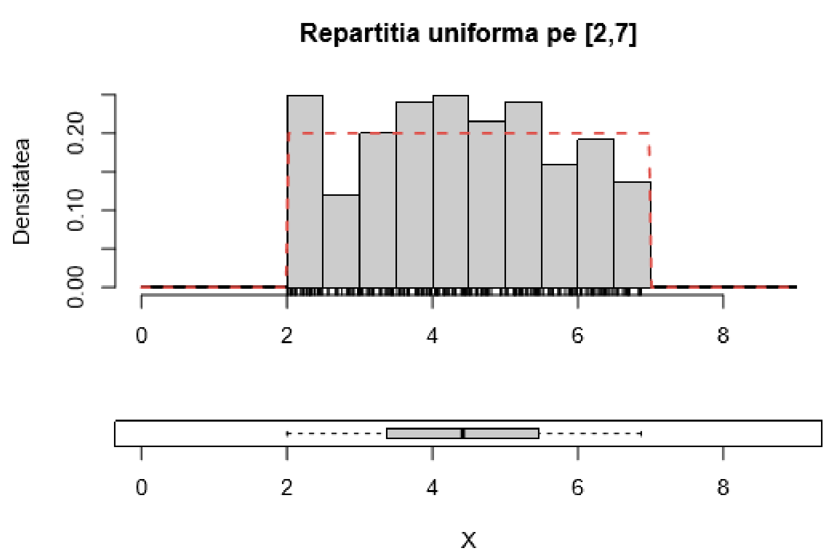
Fie X o variabilă aleatoare repartizată uniform pe [2,7]. Determinati:

a) $P(X \in \{1,2,3,4,5,6,7\})$

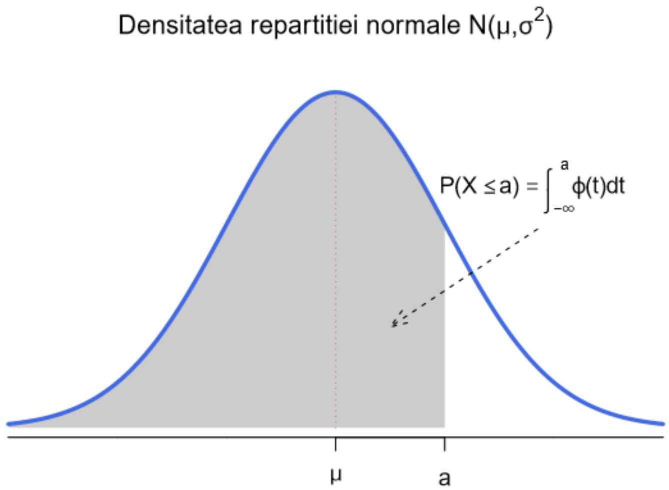
b) $P(X < 3)$ si $P(X \leq 3)$

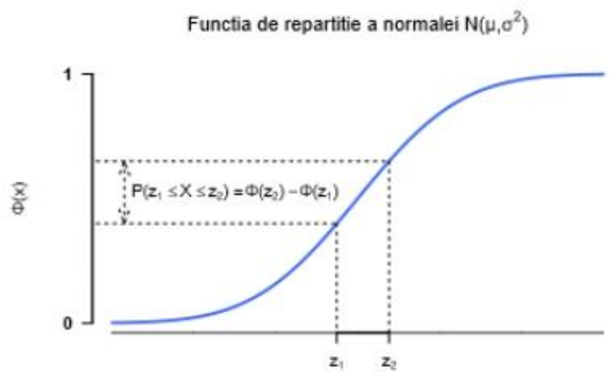
c) $P(X \leq 3 \cup X > 4)$

d) Generati 250 de observatii din repartit ia dată, trasat i histograma acestora si suprapuneti densitatea repartitiei date (vezi figura de mai jos).



2. *Repartitia Normală*





În R putem să:

• generăm observatii independente din repartit ia $N(\mu, \sigma^2)$ $(e.g. \mu\ = 0\ si\ \sigma^2\ = 2$ - în R functiile rnorm, dnorm, pnorm si qnorm primesc ca parametrii media si abaterea standard, σ nu varianta σ^2)

rnorm(10, mean = 0, sd = sqrt(2))

[1] 0.3280245 -3.1136665 0.2715085 -4.1121692 -0.8671217 0.5376793

[7] 1.9818221 0.8563125 1.5201168 1.4278308

• calculăm densitatea unei variabile aleatoare repartizate normal $N(\mu, \sigma^2)$ în diferite puncte

dnorm(seq(-2, 2, length.out = 15), mean = 3, sd = 5)

[1] 0.04839414 0.05115647 0.05390019 0.05660592 0.05925368 0.06182308

[7] 0.06429362 0.06664492 0.06885700 0.07091058 0.07278734 0.07447021

[13] 0.07594361 0.07719368 0.07820854

• calculăm functia de repartitie a unei variabile repartizate normal $N(\mu, \sigma^2)$ pentru diferite valori

pnorm(seq(-1, 1, length.out = 15), mean = 3, sd = 1)

[1] 3.167124e-05 5.736006e-05 1.018892e-04 1.775197e-04 3.033834e-04

[6] 5.086207e-04 8.365374e-04 1.349898e-03 2.137367e-03 3.320943e-03

[11] 5.063995e-03 7.579219e-03 1.113549e-02 1.606229e-02 2.275013e-02

• calculăm cuantilele de ordin $\alpha \in (0,1)$

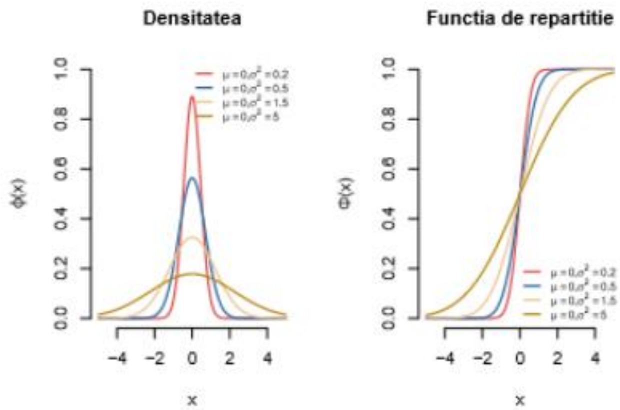
qnorm(c(0.01, 0.025, 0.05, 0.25, 0.5, 0.75, 0.95, 0.975, 0.99), mean = 0, sd = 1)

[1] -2.3263479 -1.9599640 -1.6448536 -0.6744898 0.0000000 0.6744898

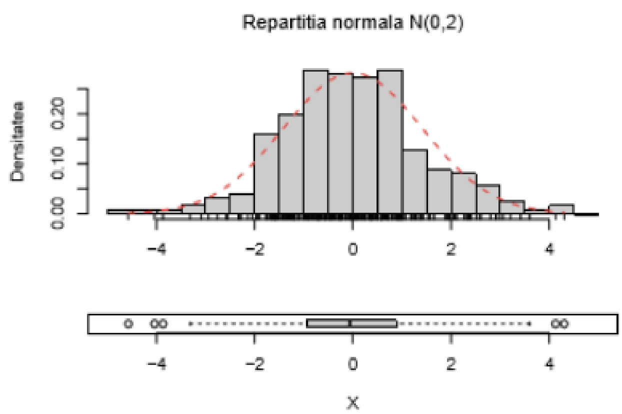
[7] 1.6448536 1.9599640 2.3263479

APLICATII:

1. Fie X o variabilă aleatoare repartizată $N(\mu, \sigma^2)$. Atunci pentru $\mu = 1$ si $\sigma = 3$ calculati:
   a) $P(X\ este\ par)$
   b) $P(X < 3.4)$ si $P(X > 1.3)$
   c) $P(1 < X < 4)$
   d) $P(X \in [2,3] \cup [3.5,5])$
   e) $P(|X - 3| > 6)$
2. Fie X o variabilă aleatoare repartizată $N(\mu, \sigma^2)$. Pentru $\mu = 0$ si $\sigma^2 \in \{0.2,0.5,1.5,5\}$ trasati pe acelasi grafic densitătile repartit iilor normale cu parametrii $N(\mu, \sigma^2)$. Adăugati legendele corespunzătoare. Aceeasi cerintă pentru functiile de repartitie.

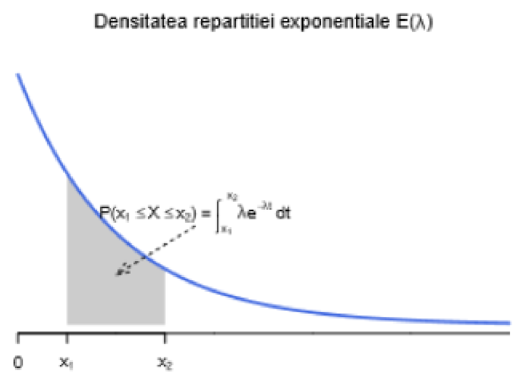


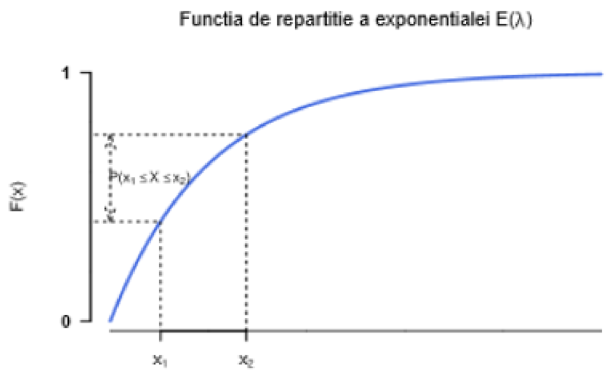
3. Generati 250 de observatii din repartit ia $N(0,2)$, trasati histograma acestora si suprapuneti densitatea repartitiei date (vezi figura de mai jos).



3. *Repartitia Exponentială*

Variabilele aleatoare repartizate exponential sunt utilizate în modelarea fenomenelor care se desfăsoară în timp continuu si care satisfac (aproximativ) proprietatea lipsei de memorie: de exemplu timpul de asteptare la un ghiseu, durata de viată a unui bec sau timpul până la următoarea convorbire telefonică.





În R putem să:

• generăm observatii independente din repartitia $E(\lambda)$ $(e.g. \lambda = 5)$

rexp(15, rate = 5)

[1] 0.13505357 0.15392539 0.25036131 0.15351051 0.00878456 0.07362396

[7] 0.07543271 0.18981181 0.05540771 0.05649451 0.15878039 0.39847262

[13] 0.05191221 0.07776034 0.22483594

• calculăm densitatea unei variabile aleatoare repartizate exponential $E(\lambda)$ în diferite puncte

dexp(seq(0, 5, length.out = 20), rate = 5)

[1] 5.000000e+00 1.341312e+00 3.598237e-01 9.652719e-02 2.589462e-02

[6] 6.946555e-03 1.863500e-03 4.999070e-04 1.341063e-04 3.597568e-05

[11] 9.650925e-06 2.588981e-06 6.945263e-07 1.863153e-07 4.998141e-08

[16] 1.340814e-08 3.596899e-09 9.649130e-10 2.588499e-10 6.943972e-11

• calculăm functia de repartitie a unei variabile repartizate exponential $E(\lambda)$ pentru diferite valori

pexp(seq(0, 5, length.out = 15), rate = 5)

[1] 0.0000000 0.8323228 0.9718843 0.9952856 0.9992095 0.9998675 0.9999778

[8] 0.9999963 0.9999994 0.9999999 1.0000000 1.0000000 1.0000000 1.0000000

[15] 1.0000000

• calculăm cuantilele de ordin α ∈(0,1)

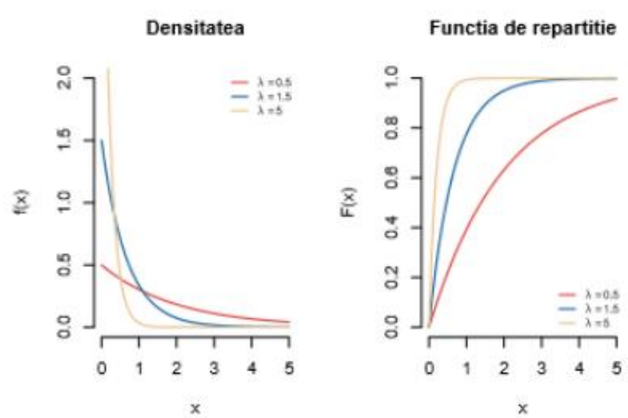
qexp(c(0.01, 0.025, 0.05, 0.25, 0.5, 0.75, 0.95, 0.975, 0.99), rate = 5)

[1] 0.002010067 0.005063562 0.010258659 0.057536414 0.138629436 0.277258872

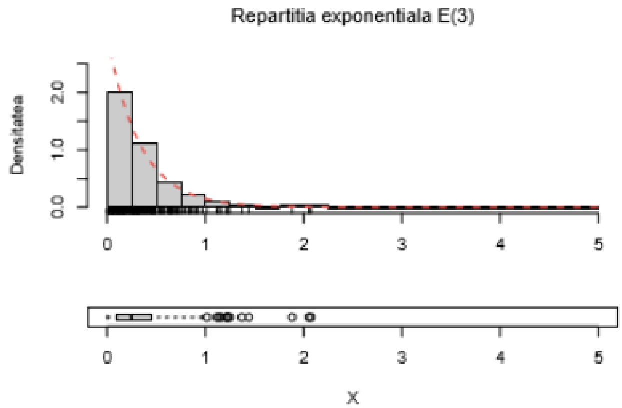
[7] 0.599146455 0.737775891 0.921034037

APLICATII

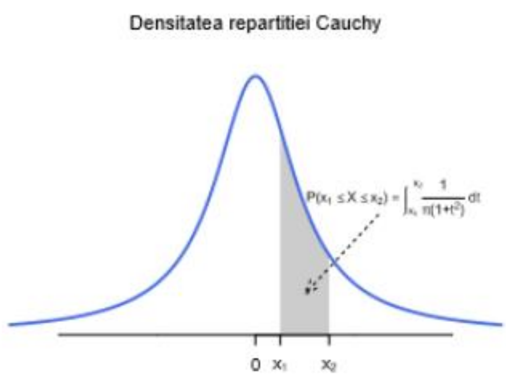
1. Fie X o variabilă aleatoare repartizată $E(\lambda)$. Pentru $\lambda \in \{0.5,1.5,5\}$ trasati pe acelasi grafic densitătile repartitiilor exponentiale de parametru λ. Adăugati legendele corespunzătoare. Aceeasi cerintă pentru functiile de repartitie.

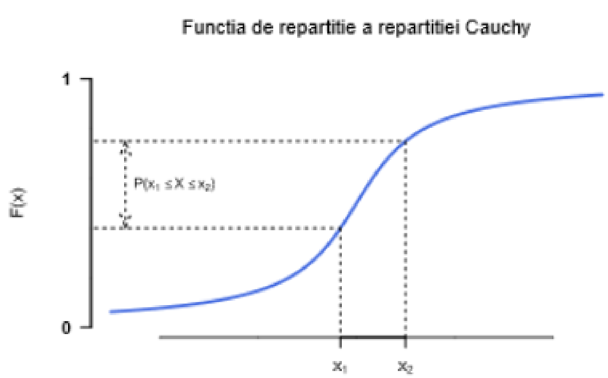


2. Generati 250 de observatii din repartitia $E(3)$, trasati histograma acestora si suprapuneti densitatea repartitiei date (vezi figura de mai jos).



4.Repartitia Cauchy





În R putem să:

• generăm observatii independente din repartitia Cauchy $C(\alpha, \beta)$ $(e.g.\, \alpha = 0, \beta = 2)$

```
rcauchy(15, location = 0, scale = 2)
 [1] -0.5966228 3.7627987 0.6864597 -0.4316018 1.4524446 0.3427032
[7] 8.4285326 3.6056089 2.3506764 -3.5453329 -1.6137218 10.4304800
[13] -0.4449169 2.3005176 -3.6644199
```

 • calculăm densitatea unei variabile aleatoare repartizate Cauchy $C(\alpha, \beta)$ în diferite puncte

```
dcauchy(seq(-5, 5, length.out = 20), location = 1, scale = 3)
[1] 0.02122066 0.02450975 0.02852541 0.03345265 0.03951056 0.04693392
 [7] 0.05591721 0.06648594 0.07825871 0.09012539 0.10006665 0.10558334
 [13] 0.10494052 0.09835367 0.08782920 0.07584810 0.06425529 0.05399054
[19] 0.04532934 0.03819719
```

 • calculăm functia de repartitie a unei variabile repartizate Cauchy $C(\alpha, \beta)$ pentru diferite valori

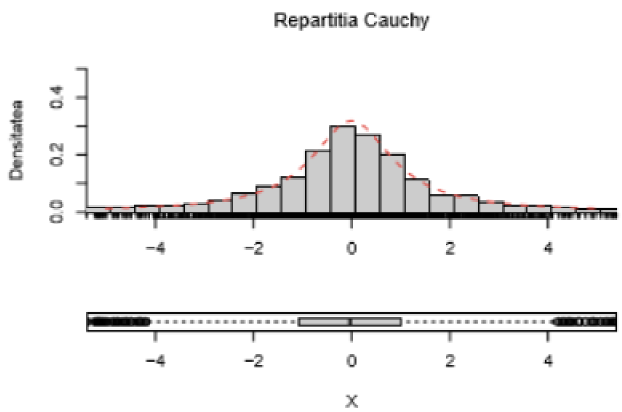
```
pcauchy(seq(-5, 5, length.out = 15), location = 1, scale = 3)
[1] 0.1475836 0.1643213 0.1848605 0.2104166 0.2425988 0.2833834 0.3347507
[8] 0.3975836 0.4697759 0.5451672 0.6158581 0.6764416 0.7255627 0.7644587
[15] 0.7951672
```

• calculăm cuantilele de ordin p $\in(0,1)$

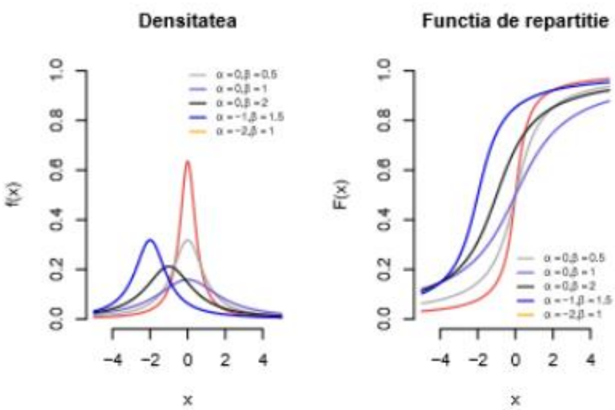
```
 qcauchy(c(0.01, 0.025, 0.05, 0.25, 0.5, 0.75, 0.95, 0.975, 0.99), location = 1, scale = 3)
[1] -94.46155 -37.11861 -17.94125 -2.00000 1.00000 4.00000 19.94125
[8] 39.11861 96.46155
```

APLICATII:

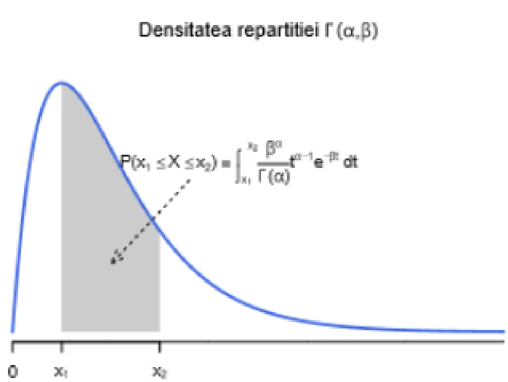
1. Generati 2500 de observatii din repartitia Cauchy, trasati histograma acestora si suprapuneti densitatea repartitiei date pentru intervalul $[-5,5]$ (vezi figura de mai jos).

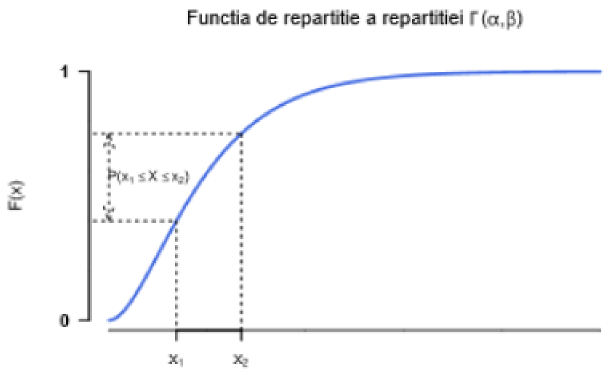


2. Fie X o variabilă aleatoare repartizată Cauchy $C(\alpha,\beta)$. Pentru fiecare pereche de parametrii (α,β) din multimea$\{(0,0.5),(0,1),(0,2),(-1,1.5),(-2,1)\}$trasati pe acelasi grafic densitătile repartitiilor Cauchy cu parametrii (α,β). Adăugati legendele corespunzătoare. Aceeasi cerintă pentru functiile de repartitie.



5.Repartitia Gama





În R putem să:

• generăm observatii independente din repartitia $\Gamma(\alpha,\beta)$ $(e.g.\ \alpha = 2, \beta = 2)$

```
rgamma(15, shape = 2, rate = 2)
 [1] 0.2739606 1.0172288 1.6546379 0.4210210 0.8476985 0.2928765 0.6798413
[8] 1.1393160 1.0763898 1.4411221 0.9500644 0.7387296 0.4159926 0.8942659
[15] 0.8366199
```

• calculăm densitatea unei variabile aleatoare repartizate $\Gamma(\alpha,\beta)$ în diferite puncte

```
dgamma(seq(0, 5, length.out = 20), shape = 1, rate = 3)
[1] 3.000000e+00 1.362251e+00 6.185761e-01 2.808853e-01 1.275455e-01
[6] 5.791632e-02 2.629886e-02 1.194188e-02 5.422615e-03 2.462321e-03
[11] 1.118100e-03 5.077110e-04 2.305433e-04 1.046860e-04 4.753619e-05
[16] 2.158541e-05 9.801583e-06 4.450739e-06 2.021008e-06 9.177070e-07
```

• calculăm functia de repartitie a unei variabile repartizate $\Gamma(\alpha,\beta)$ pentru diferite valori

```
 pgamma(seq(0, 5, length.out = 15), shape = 1, rate = 3)
[1] 0.0000000 0.6574811 0.8826808 0.9598160 0.9862362 0.9952856 0.9983852
 [8] 0.9994469 0.9998106 0.9999351 0.9999778 0.9999924 0.9999974 0.9999991
```

[15] 0.9999997

• calculăm cuantilele de ordin p ∈(0,1)

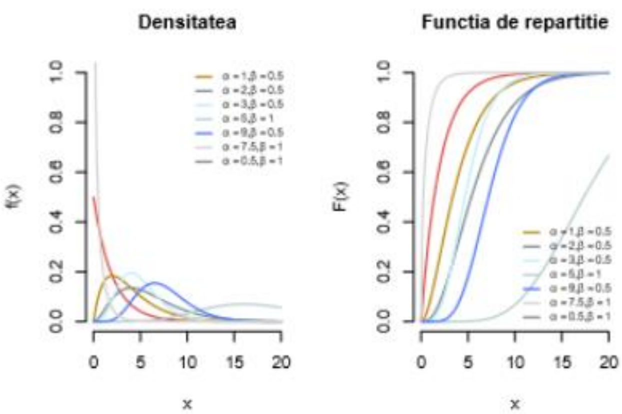
qgamma(c(0.01, 0.025, 0.05, 0.25, 0.5, 0.75, 0.95, 0.975, 0.99), shape = 1, rate = 3)

[1] 0.003350112 0.008439269 0.017097765 0.095894024 0.231049060 0.462098120

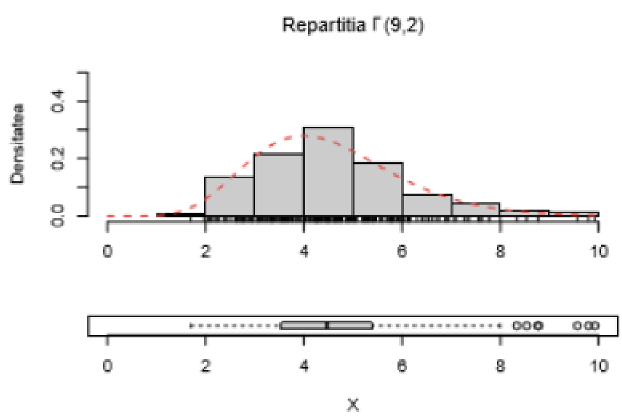
[7] 0.998577425 1.229626485 1.535056729

APLICATII:

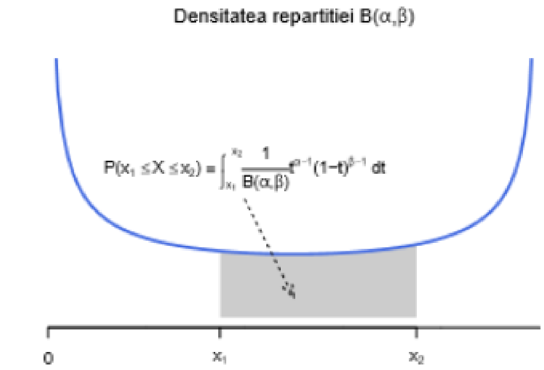
1. Fie X o variabilă aleatoare repartizată $\Gamma(\alpha, \beta)$. Pentru fiecare pereche de parametrii (α, β) din multimea$\{(1,0.5), (2,0.5), (3,0.5), (5,1), (9,0.5), (7.5,1), (0.5,1)\}$ trasati pe acelasi grafic densitătile repartitiilor Gama cu parametrii (α,β). Adăugati legendele corespunzătoare. Aceeasi cerintă pentru functiile de repartitie.

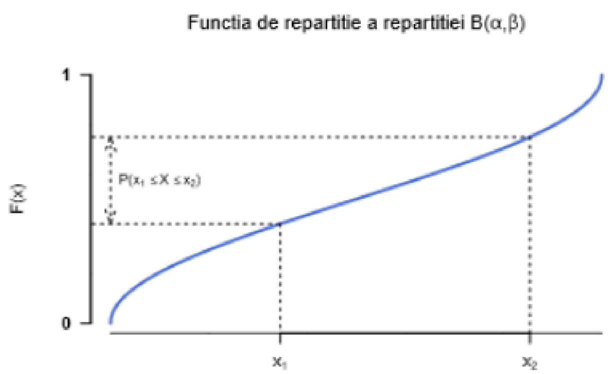


2. Generati 250 de observatii din repartitia $\Gamma(9,2)$, trasati histograma acestora si suprapuneti densitatea repartitiei date (vezi figura de mai jos).



6. *Repartitia Beta*





În R putem să:

• generăm observatii independente din repartitia $B(\alpha, \beta)$ $(e.g. \alpha = 2.5, \beta = 1)$

rbeta(15, shape1 = 2.5, shape2 = 1)

 [1] 0.7945436 0.7609136 0.9265073 0.9309420 0.5621874 0.3664261 0.9694945
[8] 0.5804873 0.9504669 0.9115169 0.8457509 0.6717780 0.7213322 0.9738473
[15] 0.9791769

• calculăm densitatea unei variabile aleatoare repartizate $B(\alpha, \beta)$ în diferite puncte

dbeta(seq(0, 1, length.out = 20), shape1 = 1, shape2 = 3)

 [1] 3.000000000 2.692520776 2.401662050 2.127423823 1.869806094
 [6] 1.628808864 1.404432133 1.196675900 1.005540166 0.831024931
[11] 0.673130194 0.531855956 0.407202216 0.299168975 0.207756233
[16] 0.132963989 0.074792244 0.033240997 0.008310249 0.000000000

• calculăm functia de repartitie a unei variabile repartizate $B(\alpha, \beta)$ pentru diferite valori

pbeta(seq(0, 1, length.out = 15), shape1 = 1, shape2 = 3)

[1] 0.0000000 0.1993440 0.3702624 0.5149417 0.6355685 0.7343294 0.8134111

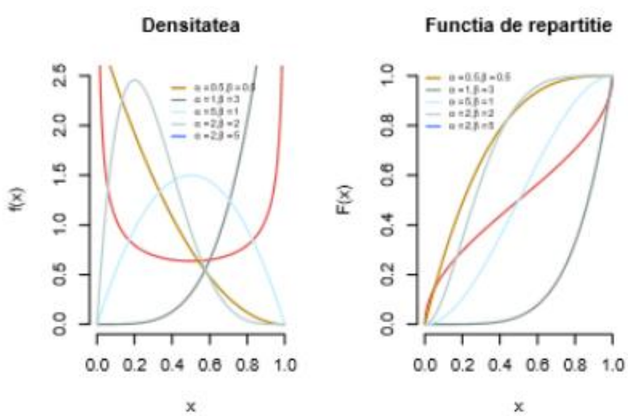
[8] 0.8750000 0.9212828 0.9544461 0.9766764 0.9901603 0.9970845 0.9996356
[15] 1.0000000

• calculăm cuantilele de ordin p ∈(0,1)

```
qbeta(c(0.01, 0.025, 0.05, 0.25, 0.5, 0.75, 0.95, 0.975, 0.99), shape1 = 1, shape2 = 3)
[1] 0.003344507 0.008403759 0.016952428 0.091439704 0.206299474 0.370039475
[7] 0.631596850 0.707598226 0.784556531
```

APLICATII:

1. Fie X o variabilă aleatoare repartizată $B(\alpha,\beta)$. Pentru fiecare pereche de parametrii (α,β) din multimea {(0.5,0.5),(1,3),(5,1),(2,2),(2,5)} trasati pe acelasi grafic densitătile repartitiilor Beta cu parametrii (α,β). Adăugati legendele corespunzătoare. Aceeasi cerintă pentru functiile de repartitie.



2. Generati 250 de observatii din repartit ia $B(3,3)$, trasati histograma acestora si suprapuneti densitatea repartitiei date (vezi figura de mai jos).