Curs 3

Cristian Niculescu

1 Dispersia (varianţa) variabilelor aleatoare discrete

1.1 Scopurile învăţării

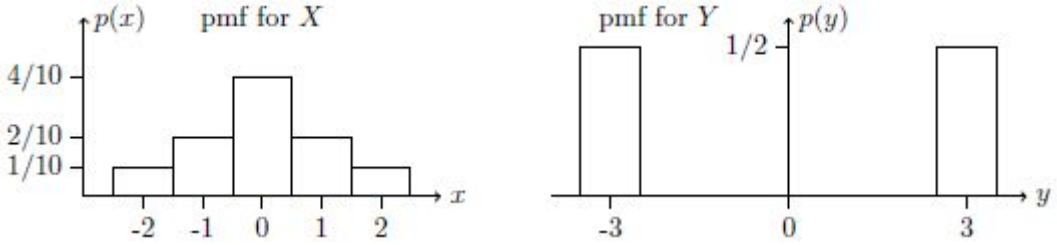
1. Să poată să calculeze dispersia şi deviaţia standard a unei variabile aleatoare discrete.
2. Să înţeleagă că deviaţia standard este o măsură a scalării sau împrăştierii.
3. Să poată să calculeze dispersia folosind proprietăţile de scalare şi liniaritate.

1.2 Împrăştierea

Media unei variabile aleatoare este o măsură a tendinţei centrale. Dacă ar trebui să rezumăm cu un singur număr o variabilă aleatoare, media ar fi o alegere bună. Totuşi, media nu cuprinde toată informaţia. De exemplu, variabilele aleatoare X şi Y de mai jos au ambele media 0, dar masa lor de probabilitate este împrăştiată în jurul mediei complet diferit.

$$X \sim \begin{pmatrix} -2 & -1 & 0 & 1 & 2 \\ 1/10 & 2/10 & 4/10 & 2/10 & 1/10 \end{pmatrix} \quad Y \sim \begin{pmatrix} -3 & 3 \\ 1/2 & 1/2 \end{pmatrix}.$$

Pentru a vedea împrăştierile diferite, reprezentăm pmf-urile. Folosim bare în loc de puncte pentru a da un sens mai bun masei.



Pmf-urile pentru 2 repartiţii diferite, ambele cu media 0

1.3 Dispersia şi deviaţia standard

Luând media centrul repartiţiei unei variabile aleatoare, dispersia este o măsură a cât de mult masa de probabilitate este împrăştiată în jurul acestui centru.

Definiţie. Dacă X este o variabilă aleatoare cu media $E(X) = \mu$, atunci dispersia lui X este

$$Var(X) = E((X - \mu)^2).$$

Deviaţia standard σ a lui X este

$$\sigma = \sqrt{Var(X)}.$$

Dacă variabila aleatoare relevantă este clară din context, atunci dispersia şi deviaţia standard sunt adesea notate cu σ^2 şi σ ("sigma"), la fel cum media este μ ("miu").

Rescriem definiţia explicit ca o sumă. Dacă X ia valorile $x_1, x_2, ..., x_n$ cu pmf p, atunci

$$Var(X) = E((X - \mu)^2) = \sum_{i=1}^{n} p(x_i)(x_i - \mu)^2.$$

Formula pentru $Var(X)$ este o medie ponderată a pătratelor distanţelor la medie. Prin ridicare la pătrat, ne asigurăm că mediem numai valori nenegative, astfel încât împrăştierea la dreapta mediei să nu se anuleze cu cea de la stânga. Prin folosirea mediei, ponderăm valorile cu probabilitate mai mare mai mult decât valorile cu probabilitate mai mică.

Observaţii despre unităţi de măsură:

1. σ are aceleaşi unităţi de măsură ca X.
2. $Var(X)$ are aceleaşi unităţi de măsură ca pătratul lui X. Astfel, dacă X este în metri, atunci $Var(X)$ este în metri pătraţi.

Deoarece σ şi X au aceleaşi unităţi de măsură, deviaţia standard este o măsură firească a împrăştierii.

Exemplul 1. Calculaţi media, dispersia şi deviaţia standard ale variabilei aleatoare

$$X \sim \begin{pmatrix} 1 & 3 & 5 \\ 1/4 & 1/4 & 1/2 \end{pmatrix}.$$

Răspuns: Întâi calculăm $E(X) = 1 \cdot (1/4) + 3 \cdot (1/4) + 5 \cdot (1/2) = 7/2$. Apoi completăm tabelul cu $(X - 7/2)^2$.

valoarea x	1	3	5
$p(x)$	1/4	1/4	1/2
$(x - 7/2)^2$	25/4	1/4	9/4

Acum calculul dispersiei este similar celui al mediei:

$$Var(X) = \frac{25}{4} \cdot \frac{1}{4} + \frac{1}{4} \cdot \frac{1}{4} + \frac{9}{4} \cdot \frac{1}{2} = \frac{11}{4}.$$

Extrăgând radicalul, avem deviaţia standard $\sigma = \sqrt{11}/2$.

Exemplul 2. Pentru fiecare variabilă aleatoare X, Y, Z şi W reprezentaţi pmf şi calculaţi media şi dispersia.

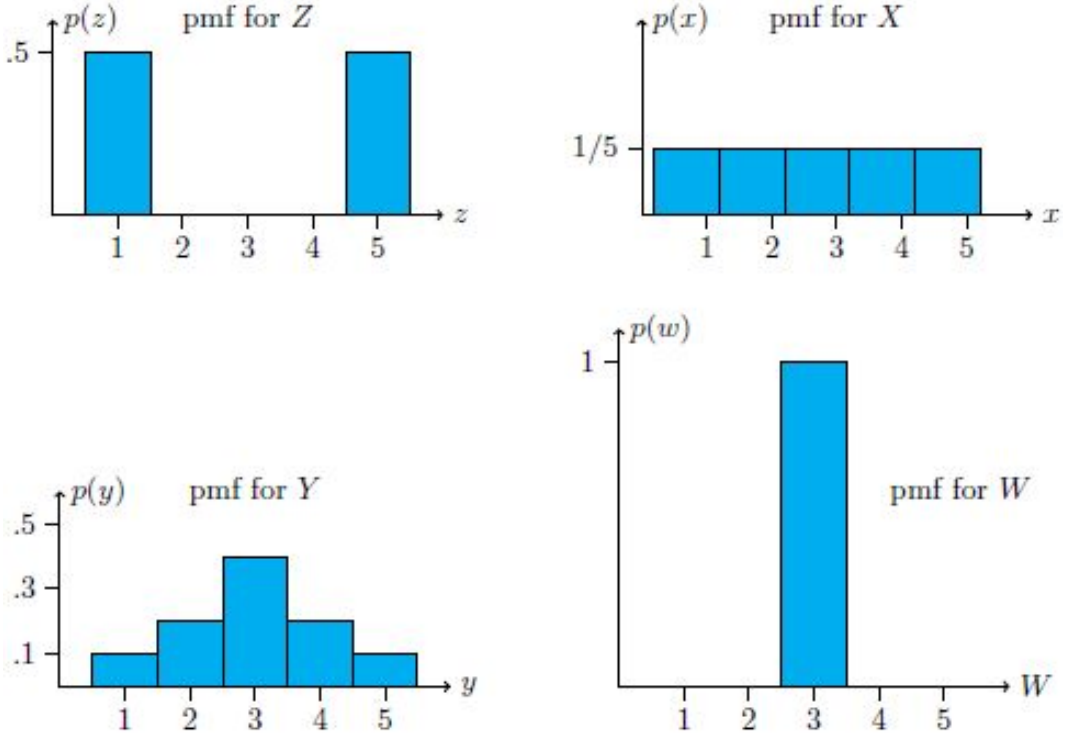
(i) $X \sim \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1/5 & 1/5 & 1/5 & 1/5 & 1/5 \end{pmatrix}$;

(ii) $Y \sim \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1/10 & 1/5 & 2/5 & 1/5 & 1/10 \end{pmatrix}$;

(iii) $Z \sim \begin{pmatrix} 1 & 5 \\ 1/2 & 1/2 \end{pmatrix}$;

(iv) $W \sim \begin{pmatrix} 3 \\ 1 \end{pmatrix}$.

Răspuns: Fiecare variabilă aleatoare are media 3, dar probabilitatea este împrăştiată diferit. În reprezentările de mai jos, le ordonăm de la cea mai mare la cea mai mică dispersie: Z, X, Y, W.



Calculăm dispersia fiecărei variabile. Toate au media $\mu = 3$. Deoarece dispersia este definită ca o medie, o putem calcula folosind tabele.

(i)

valoarea x	1	2	3	4	5
pmf $p(x)$	1/5	1/5	1/5	1/5	1/5
$(X-\mu)^2$	4	1	0	1	4

$Var(X) = E((X-\mu)^2) = \frac{4}{5} + \frac{1}{5} + 0 + \frac{1}{5} + \frac{4}{5} = 2.$

(ii)

valoarea y	1	2	3	4	5
pmf $p(y)$	1/10	1/5	2/5	1/5	1/10
$(Y-\mu)^2$	4	1	0	1	4

$Var(Y) = E((Y-\mu)^2) = \frac{2}{5} + \frac{1}{5} + 0 + \frac{1}{5} + \frac{2}{5} = 1.2.$

(iii)

valoarea z	1	5
pmf $p(y)$	1/2	1/2
$(Z-\mu)^2$	4	4

$Var(Z) = E((Z-\mu)^2) = 2 + 2 = 4.$

(iv)

valoarea w	3
pmf $p(y)$	1
$(W-\mu)^2$	0

$Var(W) = E((W-\mu)^2) = 0$. Observăm că W nu variază, de aceea are dispersia 0.

1.3.1 Dispersia unei variabile aleatoare Bernoulli(p)

Dacă $X \sim$ Bernoulli(p), atunci

$$Var(X) = p(1-p).$$

Demonstraţie: Ştim că $E(X) = p$. Calculăm $Var(X)$ folosind un tabel.

valorile lui X	0	1
pmf $p(x)$	$1-p$	p
$(X-\mu)^2$	$(0-p)^2$	$(1-p)^2$

$Var(X) = (1-p)p^2 + p(1-p)^2 = p(1-p)(p+1-p) = p(1-p)$.
Gândiţi: Pentru ce valoare a lui p Bernoulli(p) are cea mai mare dispersie? Încercaţi să răspundeţi aplicând inegalitatea mediilor sau aflând maximul funcţiei de gradul 2 $f(p) = p - p^2$ pe $(0,1)$.

1.3.2 Variabile aleatoare discrete independente

Până acum am folosit noţiunea de variabile aleatoare independente fără să o definim riguros. De exemplu, o repartiţie binomială este sumă de probe Bernoulli independente. Sigur, avem un sens intuitiv despre ce înseamnă independenţa pentru probe experimentale. Avem de asemenea sensul probabilistic că variabilele aleatoare X şi Y sunt independente când cunoaşterea

valorii lui X nu ne dă informaţii despre valoarea lui Y.
Definiţie: Variabilele aleatoare discrete X şi Y sunt independente $\iff$

$$P(X = a, Y = b) = P(X = a)P(Y = b), \forall a, b.$$

1.3.3 Proprietăţile dispersiei

Cele mai utile 3 proprietăţi pentru calculul dispersiei sunt:
1. Dacă X şi Y sunt independente, atunci $Var(X+Y) = Var(X)+Var(Y)$.
2. Pentru constantele a şi $b, Var(aX + b) = a^2Var(X)$.
3. $Var(X) = E(X^2) - E(X)^2$.
Pentru proprietatea 1, observăm cu atenţie cerinţa că X şi Y sunt independente.
Proprietatea 3 dă o formulă pentru $Var(X)$ care este adesea mai uşor de utilizat în calcule.
Exemplul 3. Presupunem că X şi Y sunt independente şi $Var(X) = 3$ şi $Var(Y) = 5$. Aflaţi:
(i) $Var(X + Y)$,
(ii) $Var(3X + 4)$,
(iii) $Var(X + X)$,
(iv) $Var(X + 3Y)$.
Răspuns: Pentru a calcula aceste dispersii, folosim proprietăţile 1 şi 2.
(i) Deoarece X şi Y sunt independente, $Var(X+Y) = Var(X)+Var(Y) = 8$.
(ii) Folosind proprietatea 2, $Var(3X + 4) = 9Var(X) = 27$.
(iii) Nu vă lasaţi păcăliţi! Proprietatea 1 nu poate fi aplicată deoarece sigur X nu este independentă de ea însăşi. Putem folosi proprietatea 2: $Var(X + X) = Var(2X) = 4Var(X) = 12$. (Observaţie: dacă din greşală am fi folosit proprietatea 1, am fi răspuns greşit 6.)
(iv) Folosim ambele proprietăţi 1 şi 2.

$$Var(X + 3Y) = Var(X) + Var(3Y) = 3 + 9 \cdot 5 = 48.$$

Exemplul 4. Folosiţi proprietatea 3 pentru a calcula dispersia lui $X \sim$ Bernoulli(p).
Răspuns: Din tabelul

X	0	1
$p(x)$	$1-p$	p
X^2	0	1

avem $E(X^2) = p$. Deci proprietatea 3 dă

$$Var(X) = E(X^2) - E(X)^2 = p - p^2 = p(1 - p).$$

Rezultatul coincide cu cel din calculul anterior.
Exemplul 5. Refaceţi exemplul 1 folosind proprietatea 3.
Răspuns: Din tabelul

X	1	3	5
$p(x)$	1/4	1/4	1/2
X^2	1	9	25

avem $E(X)$=7/2 şi

$$E(X^2) = 1 \cdot \frac{1}{4} + 9 \cdot \frac{1}{4} + 25 \cdot \frac{1}{2} = \frac{60}{4} = 15.$$

Deci $Var(X) = 15 - (7/2)^2 = 11/4$, ca în exemplul 1.

1.3.4 Dispersia pentru binomial(n, p)

Presupunem $X \sim$ binomial(n, p). Deoarece X este sumă de n variabile Bernoulli(p) *independente* şi fiecare variabilă Bernoulli are dispersia $p(1-p)$, avem

$$X \sim \text{binomial}(n, p) \Rightarrow Var(X) = np(1-p).$$

1.3.5 Demonstraţiile proprietăţilor 2 şi 3

Demonstraţia proprietăţii 2: Rezultă din proprietăţile lui $E(X)$.
Fie $\mu = E(X)$. Atunci $E(aX + b) = a\mu + b$ şi

$$\begin{aligned} Var(aX+b) = E((aX+b-(a\mu+b))^2) = E((aX-a\mu)^2) = E(a^2(X-\mu)^2) \\ = a^2 E((X-\mu)^2) = a^2 Var(X). \end{aligned}$$

Demonstraţia proprietăţii 3: Folosim proprietăţile lui $E(X)$. Reamintim că μ este o constantă şi că $E(X) = \mu$.

$$\begin{aligned} E((X-\mu)^2) &= E(X^2 - 2\mu X + \mu^2) \\ &= E(X^2) - 2\mu E(X) + \mu^2 \\ &= E(X^2) - 2\mu^2 + \mu^2 \\ &= E(X^2) - \mu^2 \\ &= E(X^2) - E(X)^2, \text{q.e.d.} \end{aligned}$$

1.4 Tabele de repartiţii şi proprietăţi

Repartiţia	valorile lui X	pmf $p(x)$	media $E(X)$	dispersia $Var(X)$
Bernoulli(p)	0,1	$p(0) = 1-p, p(1) = p$	p	$p(1-p)$
Binomial(n, p)	$0, 1, ..., n$	$p(k) = C_n^k p^k (1-p)^{n-k}$	np	$np(1-p)$
Uniform(n)	$1, 2, ..., n$	$p(k) = \frac{1}{n}$	$\frac{n+1}{2}$	$\frac{n^2-1}{12}$
Geometric(p)	0,1,2,...	$p(k) = p(1-p)^k$	$\frac{1-p}{p}$	$\frac{1-p}{p^2}$

Fie X o variabilă aleatoare discretă cu valorile $x_1, x_2, ...$ şi pmf $p(x_j)$.

Media:	**Dispersia:**
Sinonime:	varianţă
Notaţii: $E(X), \mu, m$	$Var(X), \sigma^2$
Definiţie: $E(X) = \sum_j x_j p(x_j)$	$E((X-\mu)^2) = \sum_j p(x_j)(x_j - \mu)^2$
Scalare, translatare: $E(aX+b) = aE(X) + b$	$Var(aX+b) = a^2 Var(X)$
Aditivitate: $E(X+Y) = E(X) + E(Y)$	X, Y ind.$\Rightarrow Var(X+Y) = Var(X) + Var(Y)$
Funcţii de X: $E(h(X)) = \sum_j h(x_j) p(x_j)$	
Formulă alternativă:	$Var(X) = E(X^2) - E(X)^2 = E(X^2) - \mu^2$

2 Variabile aleatoare continue

2.1 Scopurile învăţării

1. Să ştie definiţia unei variabile aleatoare continue.
2. Să ştie definiţia funcţiei densitate de probabilitate (pdf) şi funcţiei de distribuţie cumulativă (cdf).
3. Să poată să explice de ce folosim densitatea de probabilitate pentru variabilele aleatoare continue.

2.2 Introducere

Variabilele aleatoare atribuie un număr fiecărui rezultat posibil dintr-un spaţiu al probelor. În timp ce variabilele aleatoare discrete iau o mulţime discretă de valori posibile, variabilele aleatoare continue au o mulţime continuă de valori.
Computaţional, pentru a merge de la discret la continuu, pur şi simplu înlocuim sumele cu integrale.
Exemplul 1. Deoarece timpul este continuu, cantitatea de timp cu care Ion vine mai devreme (sau întârzie) la curs este o variabilă aleatoare continuă.
Presupunem că măsurăm cât de devreme soseşte Ion la curs în fiecare zi (în minute). Adică, rezultatul unei probe din experimentul nostru este un timp

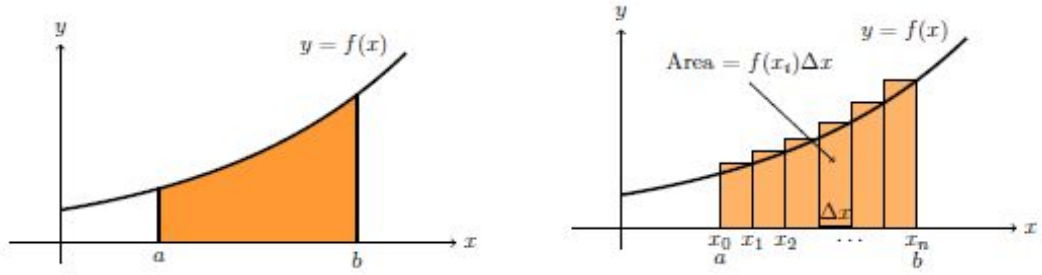
în minute. Presupunem că sunt fluctuaţii aleatoare în timpul exact în care el apare. Deoarece în principiu Ion ar putea sosi, să zicem, cu 3.43 minute mai devreme, sau cu 2.7 minute mai târziu (corespunzând rezultatului -2.7), sau la orice alt timp, spaţiul probelor constă din toate numerele reale. Deci variabila aleatoare care dă rezultatul are ea însăşi un domeniu continuu de valori posibile.

2.3 Încălzire de analiză matematică

Dacă f este o funcţie integrabilă şi $f(x) \geq 0, \forall x \in [a,b]$, atunci
1. $\int_a^b f(x)dx =$ aria de sub curba $y = f(x)$.
2. $\int_a^b f(x)dx =$ "suma din $f(x)dx$".
Conexiunea dintre cele 2 este:

$$\text{aria} \approx \text{suma ariilor dreptunghiurilor} = f(x_1)\Delta x + f(x_2)\Delta x + ... + f(x_n)\Delta x = \sum_{i=1}^{n} f(x_i)\Delta x.$$

Pe măsură ce lungimea Δx a intervalelor devine mai mică, aproximarea devine mai bună.



Aria este aproximativ suma ariilor dreptunghiurilor

Observaţie: Interesul nostru în integrale vine în primul rând din interpretarea lor ca "sumă" şi într-o mai mică măsură din interpretarea lor ca arie.

2.4 Variabile aleatoare continue şi funcţii densitate de probabilitate

O variabilă aleatoare continuă ia un domeniu de valori a cărui lungime poate fi finită sau infinită. Iată câteva exemple de domenii de valori: $[0,1], [0,\infty), \mathbb{R}, [a,b]$.
Definiţie: O variabilă aleatoare X este continuă $\iff \exists$ o funcţie $f(x)$ astfel încât $\forall c \leq d$ avem

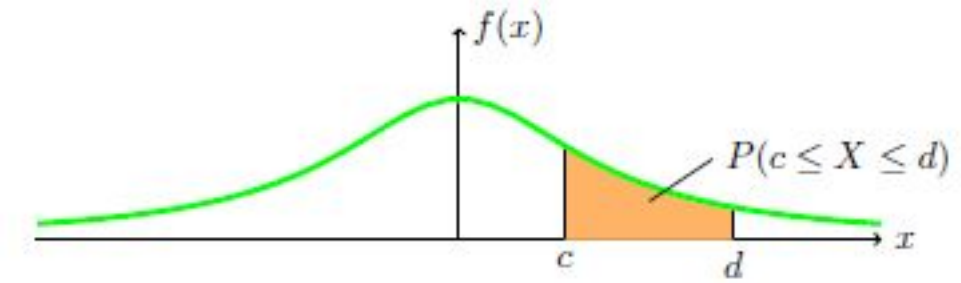
$$P(c \leq X \leq d) = \int_c^d f(x)dx. \tag{1}$$

Funcţia f este numită funcţia densitate de probabilitate (**pdf**).
Pdf satisface totdeauna următoarele proprietăţi:
1. $f(x) \geq 0$ (f este nenegativă).
2. $\int_{-\infty}^{\infty} f(x)dx = 1$ (Aceasta este echivalentă cu: $P(-\infty < X < \infty) = 1$).
Funcţia densitate de probabilitate f a unei variabile aleatoare continue este analoaga funcţiei masă de probabilitate p a unei variabile aleatoare discrete. Iată 2 diferenţe importante:
1. Spre deosebire de p, pdf f nu este o probabilitate. Trebuie s-o integrăm pentru a obţine probabilitate.
2. Deoarece f nu este o probabilitate, nu există restricţia $f(x) \leq 1$.
Observaţie: În proprietatea 2, am integrat de la $-\infty$ la ∞ deoarece n-am ştiut domeniul de valori luate de X. Formal, aceasta are sens deoarece definim $f(x) = 0$ în afara domuniului de valori al lui X. În practică, vom integra între marginile date de domeniul de valori al lui X.

2.4.1 Vederea grafică a probabilităţii

Dacă facem graficul pdf a unei variabile aleatoare continue X, atunci

$$P(c \leq X \leq d) = \text{ aria de sub graficul pdf dintre } c \text{ şi } d.$$



Gândiţi: Cât este aria totală de sub pdf f?

2.4.2 Termenii "masă de probabilitate" şi "densitate de probabilitate"

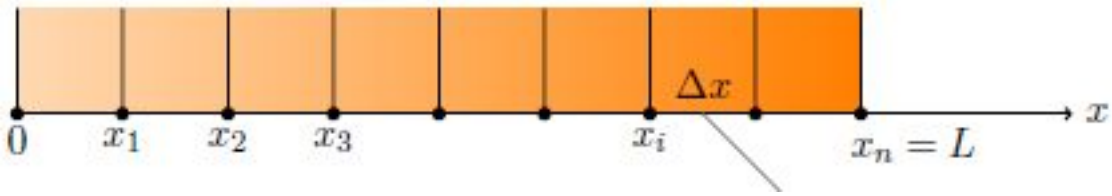
De ce folosim termenii masă şi densitate pentru a descrie pmf şi pdf? Care este diferenţa dintre cele 2? Răspunsul simplu este că aceşti termeni sunt complet analogi masei şi densităţii din fizică.
Masa ca o sumă:
Dacă masele m_1, m_2, m_3 şi m_4 sunt puse pe o linie în poziţiile x_1, x_2, x_3 şi x_4, atunci masa totală este $m_1 + m_2 + m_3 + m_4$.

Putem defini o "funcţie de masă" p cu $p(x_j) = m_j$ pentru $j = 1, 2, 3, 4$ şi $p(x) = 0$ altfel. Cu această notaţie, masa totală este $p(x_1) + p(x_2) + p(x_3) + p(x_4)$.

Funcţia masă de probabilitate se comportă în exact acelaşi mod, cu excepţia faptului că are dimensiunea probabilităţii în loc de cea a masei.

Masa ca o integrală a densităţii:

Presupunem că avem o tijă de lungime L metri cu densitate variabilă $f(x)$ kg/m. (Observăm că unităţile de măsură sunt masă/lungime.)



masa celei de-a i-a bucăţi $\approx f(x_i)\Delta x$

Dacă densitatea variază continuu, trebuie să aflăm masa totală a tijei prin integrare:

$$\textbf{masa totală} \ = \int_0^L f(x)dx.$$

Această formulă provine din împărţirea tijei în bucăţi mici şi "adunarea" masei fiecărei bucăţi. Adică:

$$\textbf{masa totală} \ \approx \sum_{i=1}^{n} f(x_i)\Delta x.$$

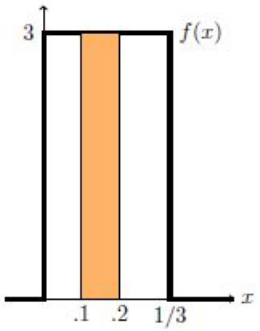
La limită când Δx tinde la 0, suma tinde la integrală.

Funcţia densitate de probabilitate se comportă exact la fel, exceptând faptul că are ca unităţi de măsură probabilitea/(unitatea lui x) în loc de kg/m. Într-adevăr, relaţia (1) este exact analoaga integralei de mai sus pentru masa totală.

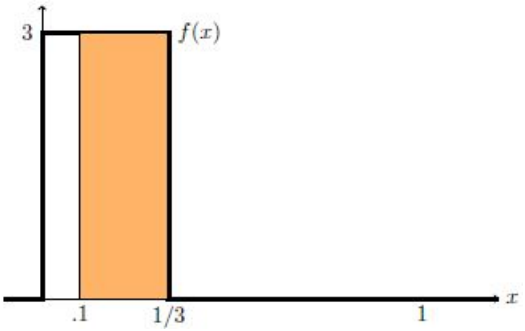
Pentru ambele variabile aleatoare, discretă şi continuă, media este centrul masei sau punctul de echilibru.

Exemplul 2. Presupunem că X are pdf $f(x) = 3$ pe $[0, 1/3]$ (aceasta înseamnă că $f(x) = 0$ în afara lui $[0, 1/3]$). Reprezentaţi pdf şi calculaţi $P(0.1 \leq X \leq 0.2)$ şi $P(0.1 \leq X \leq 1)$.

Răspuns:



$P(0.1 \leq X \leq 0.2)$



$P(0.1 \leq X \leq 1)$

$$P(0.1 \leq X \leq 0.2) = \int_{0.1}^{0.2} f(x)dx = \int_{0.1}^{0.2} 3dx = 3x|_{0.1}^{0.2} = 0.6 - 0.3 = 0.3.$$

Sau putem afla probabilitatea geometric:

$$P(0.1 \leq X \leq 0.2) = \text{ aria dreptunghiului } = 3 \cdot 0.1 = 0.3.$$

Deoarece $P(0.1 \leq X \leq 1)$ este aria de sub $f(x)$ de la 0.1 doar până la 1/3, avem $P(0.1 \leq X \leq 1) = 3(1/3 - 0.1) = 3 \cdot (7/30) = 0.7$.

Gândiţi: În exemplul anterior $f(x)$ ia valori mai mari ca 1. De ce aceasta nu încalcă regula că probabilităţile sunt totdeauna între 0 şi 1?

Observaţie asupra notaţiei. Putem defini o variabilă aleatoare dându-i domeniul de valori şi funcţia densitate de probabilitate. De exemplu, putem spune: fie X o variabilă aleatoare cu domeniul de valori $[0, 1]$ şi pdf $f(x) = 2x$. Implicit, aceasta înseamnă că X nu are densitate de probabilitate în afara domeniului de valori dat. Dacă am fi vrut să fim absolut riguroşi, am fi scris explicit şi că $f(x) = 0$ în afara lui [0,1], dar în practică aceasta nu este necesar.

Exemplul 3. Fie X o variabilă aleatoare cu domeniul de valori [0,1] şi pdf $f(x) = Cx^2$. Cât este C?

Răspuns: Deoarece probabilitatea totală trebuie să fie 1, avem

$$\int_0^1 f(x)dx = 1 \Leftrightarrow \int_0^1 Cx^2dx = 1.$$

Calculând integrala, ultima relaţie devine

$$((Cx^3)/3)|_0^1 = 1 \Rightarrow C/3 = 1 \Rightarrow C = 3.$$

Observaţie: Avem nevoie de constanta C de mai sus pentru a norma densitatea astfel încât probabilitatea totală să fie 1.

Exemplul 4. Fie X variabila aleatoare din exemplul 3. Aflaţi $P(X \leq 1/2)$.
Răspuns: $P(X \leq 1/2) = \int_0^{1/2} 3x^2 dx = x^3|_0^{1/2} = \frac{1}{8}$.
Gândiţi: Pentru acest X (sau orice variabilă aleatoare continuă):

- Cât este $P(a \leq X \leq a)$?
- Cât este $P(X = 0)$?
- $P(X = a) = 0$ înseamnă că X nu poate fi niciodată egal cu a?

În cuvinte, întrebările de mai sus duc la faptul că probabilitatea ca înălţimea unei persoane aleatoare să fie exact 1.7526 m (la precizie infinită, i.e. fără rotunjiri!) este 0. Totuşi, este posibil ca înălţimea cuiva să fie exact 1.7526 m. Deci răspunsurile la întrebările de gândire sunt 0, 0 şi Nu.

2.4.3 Funcţia de distribuţie cumulativă (funcţia de repartiţie)

Funcţia de distribuţie cumulativă (**cdf**) a unei variabile aleatoare continue X este definită în exact acelaşi mod ca cdf a unei variabile aleatoare discrete.

$$F(b) = P(X \leq b).$$

Observăm că definiţia este despre probabilitate. Când folosim cdf ar trebui să ne gândim mai întâi la ea ca la o probabilitate. Apoi, când o calculăm putem folosi

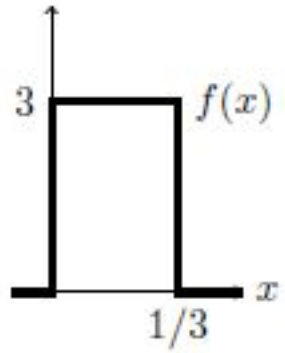
$$F(b) = P(X \leq b) = \int_{-\infty}^{b} f(x)dx, \text{ unde } f(x) \text{ este pdf a lui } X.$$

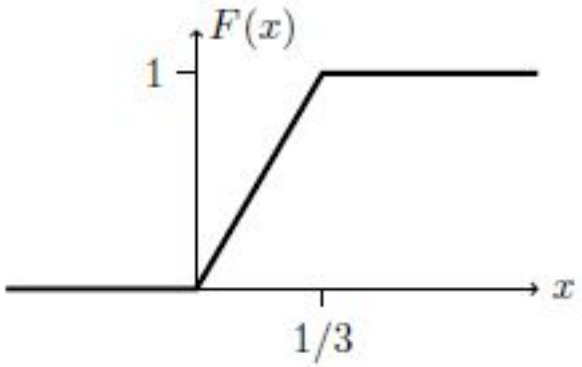
Observaţii:
1. Pentru variabile aleatoare discrete, am definit funcţia de distribuţie cumulativă, dar nu am avut prea mult ocazia să o folosim. Cdf joacă un rol mult mai proeminent pentru variabile aleatoare continue.
2. Ca mai înainte, am integrat de la $-\infty$ deoarece n-am ştiut precis domeniul de valori al lui X. Formal, aceasta are sens deoarece $f(x) = 0$ în afara domeniului de valori al lui X. În practică, ştim domeniul de valori şi integrăm de la începutul lui.
3. În practică spunem adesea "X are repartiţia $F(x)$" mai degrabă decât "X are funcţia de distribuţie cumulativă $F(x)$".
Exemplul 5. Aflaţi cdf pentru densitatea din exemplul 2.
Răspuns: Pentru $a \in [0, 1/3]$ avem $F(a) = \int_0^a f(x)dx = \int_0^a 3dx = 3x|_0^a = 3a$.

Deoarece $f(x) = 0$ în afara lui [0,1/3] avem $F(a) = P(X \leq a) = 0$ pentru $a < 0$ şi $F(a) = 1$ pentru $a > 1/3$. Punând toate acestea împreună, avem

$$F(a) = \begin{cases} 0, & \text{dacă } a < 0, \\ 3a, & \text{dacă } 0 \leq a \leq 1/3, \\ 1, & \text{dacă } a > 1/3. \end{cases}$$

Iată graficele lui f şi F.





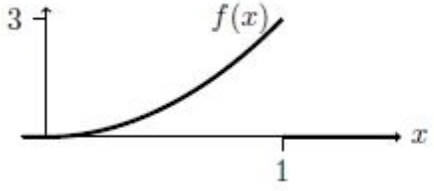
Observaţi scalele diferite pe axele verticale. Reamintim că axa verticală pentru pdf reprezintă densitatea de probabilitate, iar cea pentru cdf reprezintă probabilitatea.

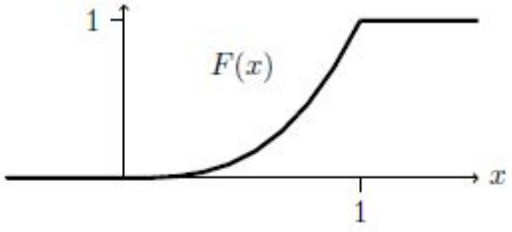
Exemplul 6. Aflaţi cdf pentru pdf din exemplul 3, $f(x) = 3x^2$ pe [0,1]. Presupunem că X este o variabilă aleatoare cu această repartiţie. Aflaţi $P(X < 1/2)$.

Răspuns: $f(x) = 3x^2$ pe $[0, 1] \Rightarrow F(a) = \int_0^a 3x^2 dx = x^3|_0^a = a^3$ pe [0,1]. De aceea,

$$F(a) = \begin{cases} 0, & \text{dacă } a < 0, \\ a^3, & \text{dacă } 0 \leq a \leq 1, \\ 1, & \text{dacă } a > 1. \end{cases}$$

Astfel, $P(X < 1/2) = F(1/2) = (1/2)^3 = 1/8$. Iată graficele lui f şi F:





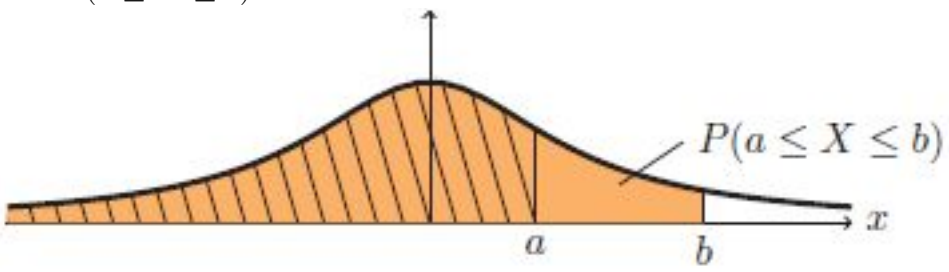
2.4.4 Proprietăţi ale funcţiei de distribuţie cumulative

Iată un rezumat al celor mai importante proprietăţi ale funcţiei de distribuţie cumulative (cdf) pentru o variabilă aleatoare continuă:
1. (Definiţia) $F(x) = P(X \leq x)$.
2. $0 \leq F(x) \leq 1, \forall x \in \mathbb{R}$.
3. F este crescătoare, i.e. dacă $a \leq b$, atunci $F(a) \leq F(b)$.
4. $\lim_{x \to -\infty} F(x) = 0$ şi $\lim_{x \to \infty} F(x) = 1$.
5. $P(a \leq X \leq b) = F(b) - F(a)$.
6. $F'(x) = f(x), \forall x$ în care F este derivabilă.

Proprietăţile 2, 3 şi 4 sunt identice cu cele pentru repartiţii discrete. Graficele din exemplele anterioare le ilustrează.
Proprietatea 5 poate fi demonstrată:

$$\begin{aligned} &\int_{-\infty}^{b} f(x)dx = \int_{-\infty}^{a} f(x)dx + \int_{a}^{b} f(x)dx \\ \Leftrightarrow &\int_{a}^{b} f(x)dx = \int_{-\infty}^{b} f(x)dx - \int_{-\infty}^{a} f(x)dx \\ \Leftrightarrow &P(a \leq X \leq b) = F(b) - F(a). \end{aligned}$$

Proprietatea 5 poate fi de asemenea văzută geometric. Regiunea colorată de mai jos reprezintă $F(b)$ şi regiunea haşurată reprezintă $F(a)$. Diferenţa lor este $P(a \leq X \leq b)$.



Proprietatea 6 este o teoremă fundamentală a analizei matematice.

2.4.5 Densitatea de probabilitate ca o tablă de darts

Ne gândim la prelevarea de valori ale unei variabile aleatoare continue ca la aruncarea de darts-uri într-o tablă de darts. Considerăm regiunea de sub graficul pdf ca o tablă de darts. Împărţim regiunea în pătrate mici de aceeaşi mărime şi presupunem că, atunci când aruncăm un dart, este egal probabil ca el să aterizeze în orice pătrat. Probabilitatea ca dartul să aterizeze într-o regiune dată este fracţia din totalul ariei de sub curbă preluată de regiune.

Deoarece totalul ariei este 1, această fracţie este chiar aria regiunii. Dacă X reprezintă coordonata x a dartului, atunci probabilitatea ca dartul să aterizeze cu coordonata x între a şi b este chiar

$$P(a \leq X \leq b) = \text{ aria de sub } f(x) \text{ dintre } a \text{ şi } b = \int_a^b f(x)dx.$$

3 Galerie a variabilelor aleatoare continue

3.1 Scopurile învăţării

1. Să poată da exemple de ceea ce modelează repartiţiile uniformă, exponenţială şi normală.
2. Să poată da domeniile de valori şi pdf-urile repartiţiilor uniformă, exponenţială şi normală.

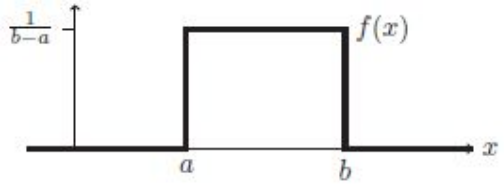
3.2 Introducere

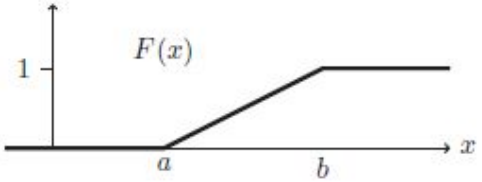
Introducem câteva repartiţii continue fundamentale. Pentru fiecare repartiţie dăm domeniul de valori, pdf, cdf şi o scurtă descriere a situaţiilor pe care le modelează. Toate aceste repartiţii depind de parametri, pe care îi specificăm.
Cu toate că o abordăm spre final, repartiţia normală este cea mai importantă din cele definite aici. Când dăm funcţia de repartiţie (cdf, prescurtat "repartiţia") se subînţelege, dacă este cazul, că la stânga intervalului pe care este dată este 0, iar la dreapta este 1.

3.3 Repartiţia uniformă

1. Parametri: $a < b$.
2. Domeniu de valori: $[a, b]$.
3. Notaţie: uniform(a, b) sau $U(a, b)$.
4. Densitate: $f(x) = \frac{1}{b-a}$ pentru $a \leq x \leq b$.
5. Repartiţia: $F(x) = (x - a)/(b - a)$ pentru $a \leq x \leq b$.
6. Modele: Toate rezultatele din domeniul de valori au probabilitate egală (mai precis, toate rezultatele au aceeaşi densitate de probabilitate).

Grafice pentru pdf şi cdf:





Exemple. 1. Presupunem că avem o ruletă marcată în milimetri. Dacă măsurăm (până la cel mai apropiat marcaj) lungimea unor articole care au aproximativ 1 metru, eroarea de rotunjire va fi repartizată uniform între -0.5 şi 0.5 milimetri.
2. Multe jocuri de tablă folosesc săgeţi care se învârt pentru a introduce hazardul. Când este învârtită, săgeata se opreşte la un unghi care este repartizat uniform între 0 şi 2π radiani.
3. În cele mai multe generatoare de numere pseudo-aleatoare, generatorul de bază simulează o repartiţie uniformă şi toate celelalte repartiţii sunt construite transformând generatorul de bază.
Teorema de universalitate a repartiţiei uniforme (Paul Levy)

X variabilă aleatoare

$Y \sim U(0,1)$

$F_X^{-1} : (0,1) \to \mathbb{R}$, $F_X^{-1}(y) = \inf\{x \in \mathbb{R} | F_X(x) \geq y\}, \forall y \in (0,1)$ - funcţia cuantilă (inversa generalizată) asociată lui F_X

$\Longrightarrow$ X şi $F_X^{-1}(Y)$ sunt repartizate la fel.

Observaţie. Acest rezultat joacă un rol important în simulare.

El ne permite să generăm observaţii independente din orice repartiţie, dacă ştim funcţia cuantilă.

$x_1, ..., x_n$, observaţii independente repartizate $U(0,1)$ generate prin `runif(n,0,1)`.

$F_X^{-1}(x_1), ..., F_X^{-1}(x_n)$ sunt observaţii independente repartizate la fel ca X.

3.4 Repartiţia exponenţială

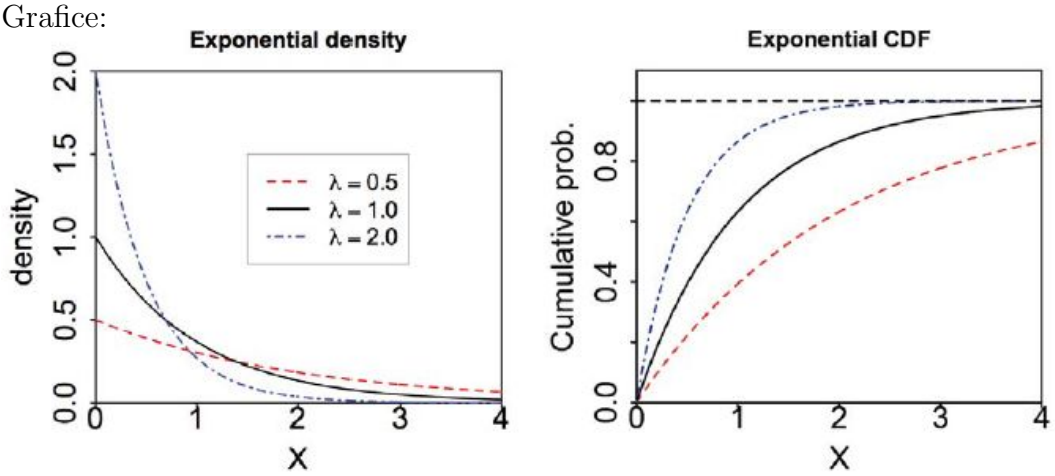
1. Parametru: $\lambda > 0$.
2. Domeniu de valori: $[0, \infty)$.
3. Notaţie: exponential(λ) sau $exp(\lambda)$.
4. Densitate: $f(x) = \lambda e^{-\lambda x}$ pentru $x \geq 0$.
5. Repartiţie: $F(x) = 1 - e^{-\lambda x}$ pentru $x \geq 0$.
6. *Repartiţia cozii drepte*: $P(X > x) = 1 - F(x) = e^{-\lambda x}$.
7. Modele: Timpul de aşteptare pentru un proces continuu de schimbare a stării.

Exemple. 1. Dacă ies afară după curs şi aştept un taxi, timpul meu de aşteptare în minute este repartizat exponenţial. În acest caz λ este dat de 1 supra numărul mediu de taxiuri care trec pe minut (în această perioadă de timp din zilele lucrătoare).
2. Repartiţia exponenţială modelează timpul de aşteptare până când un izotop instabil suferă o descompunere nucleară. În acest caz λ este legat de timpul de înjumătăţire al izotopului.
Proprietatea de lipsă a memoriei: Sunt şi alte repartiţii care modelează

timpii de aşteptare, dar repartiţia exponenţială are proprietatea adiţională că este lipsită de memorie. Iată ce înseamnă aceasta în contextul exemplului 1. Presupunem că probabilitatea ca un taxi să sosească în primele 5 minute este p. Dacă aştept 5 minute şi de fapt nu soseşte niciun taxi, atunci probabilitatea ca un taxi să sosească în următoarele 5 minute este tot p.
Din contra, să presupunem că merg la o staţie de metrou şi aştept următorul tren. Deoarece trenurile sunt coordonate să urmeze un program (de exemplu, cel mult 12 minute între trenuri), dacă aştept 5 minute fără să văd un tren, atunci este o probabilitate mult mai mare ca trenul să sosească în următoarele 5 minute. În particular, timpul de aşteptare pentru metrou nu este fără memorie şi un model mai bun pentru el ar fi repartiţia uniformă pe intervalul [0,12].
Lipsa de memorie a repartiţiei exponenţiale este analoagă lipsei de memorie a repartiţiei geometrice (discrete), unde faptul că am aruncat 5 reversuri la rând nu ne dă nicio informaţie despre următoarele 5 aruncări.
Într-adevăr, repartiţia exponenţială este precis perechea continuă a repartiţiei geometrice, care modelează timpul de aşteptare pentru un proces discret de schimbare de stări. Mai formal, lipsa de memorie înseamnă că probabilitatea de a aştepta mai mult de t minute este neafectată de faptul că am aşteptat deja s minute fără ca evenimentul să se producă. În simboluri, $P(X > s+t|X > s) = P(X > t)$.
Demonstraţia lipsei de memorie: Deoarece $(X > s+t) \cap (X > s) = (X > s+t)$, avem

$$P(X > s+t|X > s) = \frac{P(X > s+t)}{P(X > s)} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} = e^{-\lambda t} = P(X > t), \text{q.e.d.}$$

Grafice:



3.5 Repartiţia normală

În 1809, Carl Friedrich Gauss a publicat o monografie introducând câteva noţiuni care au devenit fundamentale în statistică: repartiţia normală, estimarea de verosimilitate maximă şi metoda celor mai mici pătrate. Din acest motiv, repartiţia normală mai este numită şi repartiţia *Gaussiană*. Ea este cea mai importantă repartiţie continuă.

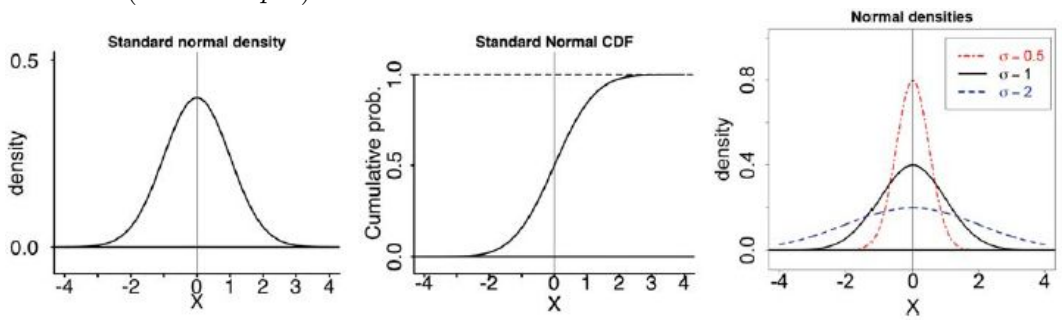
1. Parametri: $\mu \in \mathbb{R}, \sigma > 0$.
2. Domeniul de valori: $\mathbb{R}$.
3. Notaţie: normal(μ, σ^2) sau $N(\mu, \sigma^2)$.
4. Densitate: $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$.
5. Repartiţie: $F(x)$ nu are formulă, aşa că folosim tabele sau comenzi ca `pnorm` în R pentru a calcula $F(x)$.
6. Modele: măsurarea erorii, inteligenţei/abilităţii, înălţimii, mediilor loturilor de date.

Repartiţia normală standard $N(0, 1)$ are media 0 şi dispersia 1. Rezervăm Z pentru o variabilă aleatoare normală standard, $\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$ pentru densitatea normală standard şi $\Phi(z)$ pentru repartiţia normală standard.

Observaţie: Media şi dispersia variabilelor aleatoare continue au aceeaşi interpretare ca în cazul discret. Repartiţia normală $N(\mu, \sigma^2)$ are media μ, dispersia σ^2 şi deviaţia standard σ.

Iată câteva grafice ale repartiţiei normale. Observăm că densităţile au graficele curbe în formă de *clopot*. Mai observăm că pe măsură ce σ creşte, ele devin din ce în ce mai împrăştiate.

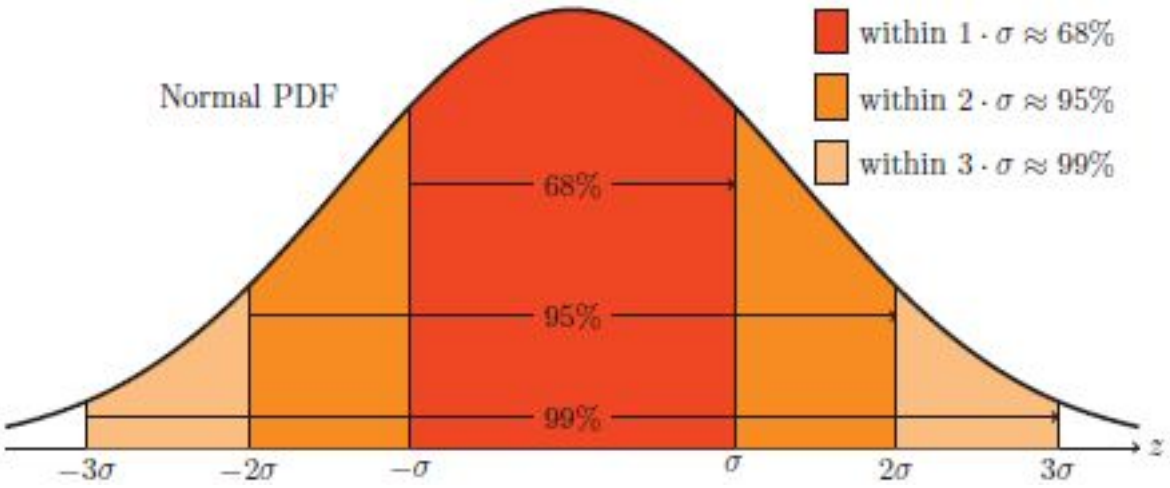
Grafice: (*curba clopot*):



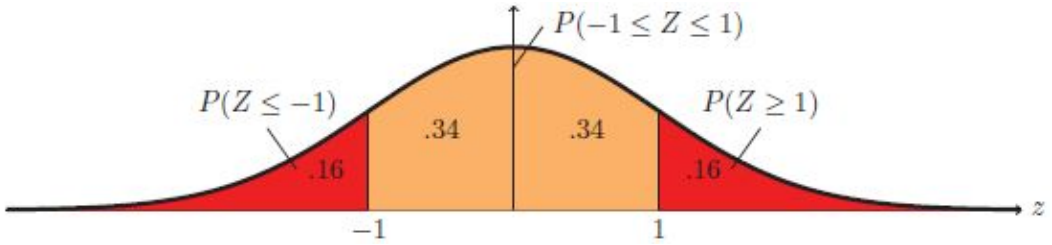
3.5.1 Probabilităţi normale

Pentru a face aproximări este util să ne reamintim următoarea regulă a degetului mare pentru 3 probabilităţi aproximative

$P(-1 \leq Z \leq 1) \approx 0.68,\ P(-2 \leq Z \leq 2) \approx 0.95,\ P(-3 \leq Z \leq 3) \approx 0.99.$



Calcule cu simetrie
Putem folosi simetria graficului densităţii normale standard faţă de dreapta $x = 0$ pentru a face unele calcule.
Exemplul 1. Regula degetului mare spune $P(-1 \leq Z \leq 1) \approx 0.68$. Folosiţi aceasta pentru a estima $\Phi(1)$.
Răspuns: $\Phi(1) = P(Z \leq 1)$.



În figură, cele 2 cozi (în roşu) au împreună aria $1 - 0.68 = 0.32$. Din simetrie, coada stângă are aria 0.16 (jumătate din 0.32), deci $P(Z \leq 1) \approx 0.16 + 0.68 = 0.84$.

3.5.2 Folosirea lui R pentru a calcula $\Phi(z)$

```
# Folosim funcţia R pnorm(x, μ, σ) pentru a calcula F(x) pentru N(μ, σ²).
pnorm(1,0,1)
[1] 0.8413447
pnorm(0,0,1)
[1] 0.5
pnorm(1,0,2)
[1] 0.6914625
pnorm(1,0,1)-pnorm(-1,0,1)
[1] 0.6826895
```

```
pnorm(5,0,5)-pnorm(-5,0,5)
[1] 0.6826895
# Desigur, z poate fi un vector de valori:
pnorm(c(-3,-2,-1,0,1,2,3),0,1)
[1] 0.001349898 0.022750132 0.158655254 0.500000000 0.841344746 0.977249868
[7] 0.998650102
```

Observaţie: Funcţia R pnorm(x, μ, σ) utilizează σ în timp ce notaţia noastră pentru repartiţia normală $N(\mu, \sigma^2)$ foloseşte σ^2.
Iată un tabel de valori cu o acurateţe cu mai puţine zecimale:

z:	-2	-1	0	.3	.5	1	2	3
$\Phi(z)$:	0.0228	0.1587	0.5000	0.6179	0.6915	0.8413	0.9772	0.9987

Exemplul 2. Utilizaţi R pentru a calcula $P(-1.5 \leq Z \leq 2)$.
Răspuns: $P(-1.5 \leq Z \leq 2) = \Phi(2) - \Phi(-1.5) =$ `pnorm(2,0,1)-pnorm(-1.5,0,1)=0.9104427`.

3.6 Repartiţia Pareto

1. Parametri: $m > 0$ şi $\alpha > 0$.
2. Domeniul de valori: $[m, \infty)$.
3. Notaţie: Pareto(m, α).
4. Densitate: $f(x) = \frac{\alpha m^\alpha}{x^{\alpha+1}}$.
5. Repartiţia: $F(x) = 1 - \frac{m^\alpha}{x^\alpha}$, pentru $x \geq m$.
6. Repartiţia coadă: $P(X > x) = m^\alpha / x^\alpha$, pentru $x \geq m$.
7. Modele: Repartiţia Pareto modelează o **lege de putere**, unde probabilitatea ca un eveniment să aibă loc variază ca o putere a unui atribut al evenimentului. Multe fenomene urmează o lege de putere, ca mărimea meteoriţilor, nivelurile veniturilor într-o populaţie şi nivelurile populaţiei în oraşe. Vezi `http://en.wikipedia.org/wiki/Pareto_distribution#Applications`.