Teoria probabilităților - refresher

A, B, C ... - evenimente

$P(A)$ - probabilitatea ca evenimentul A să se întâmple

$0 \leq P(A) \leq 1$

$P(\bar{A}) = 1 - P(A)$ - probabilitatea inversă

A și B → notăm $P(A \cdot B)$

$P(A \cdot B) = P(A|B) \cdot P(B)$

Dacă A și B nu sunt legate cauzal:

$P(A \cdot B) = P(A) \cdot P(B)$

A sau B → $P(A+B)$

$P(A+B) = P(A) + P(B) - P(A \cdot B)$

Dacă A și B se exclud mutual:

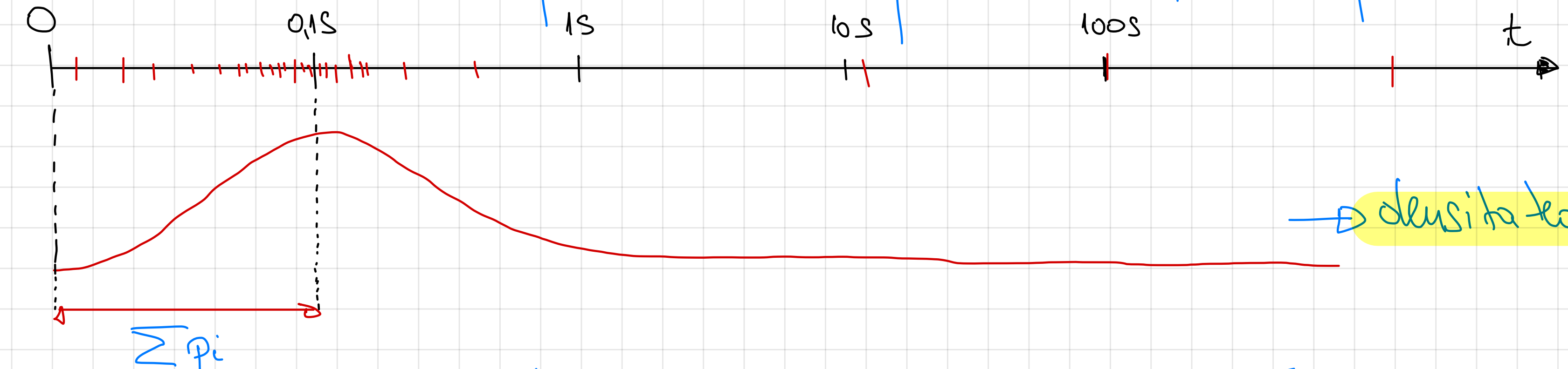
$P(A+B) = P(A) + P(B)$

Distribuții de probabilitate

Variabile aleatoare: X, Y, Z ... în spațiu de reprezentare < discret / continuu

Ex: timpul de răspuns al unui search engine → X

Dacă desenăm pe axa timpului toți timpii de răspuns:



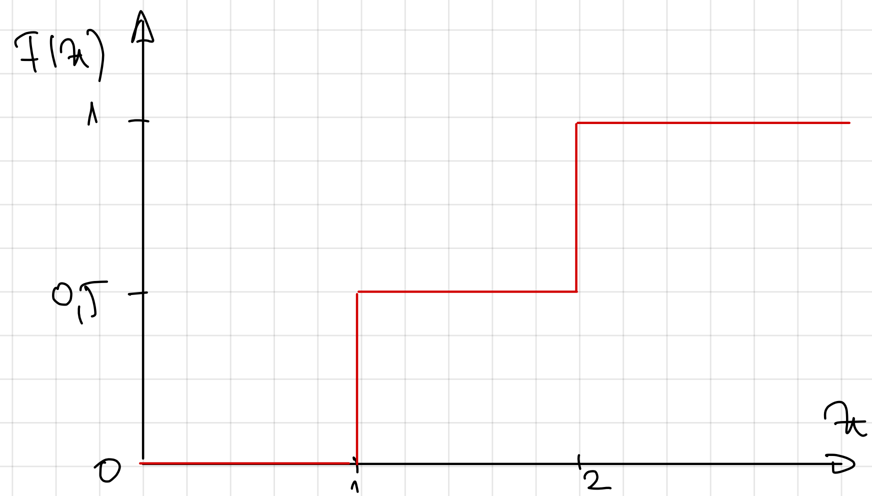
Care este probabilitatea ca timpul de răspuns să fie $\leq 0,1$ s?

Sumă de probabilități → funcție cumulativă de distribuție a prob. (CDF)

CDF: $F_X(x) = P(X \leq x)$

Ex: Aruncarea unei monede

$$F(x) = \begin{cases} 0, & x < 1 \\ 0{,}5, & x \in [1, 2) \\ 1, & x > 2 \end{cases}$$



Este relevant să folosim **timpul** ca variabilă aleatoare

$F(x) = F(t)$, $t \in [0, \infty)$

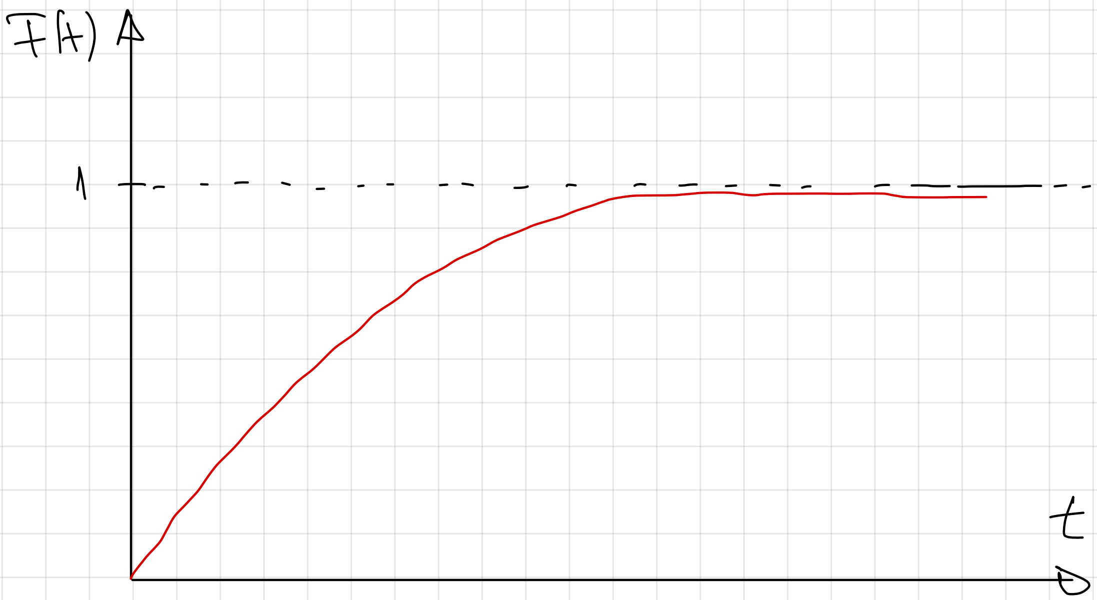
Proprietăți CDF:

$0 \le F(t) \le 1$

$F(0) = 0$

$\lim_{t \to \infty} F(t) = 1$

$F(t)$ este monoton crescătoare



Ex: $F(t) = 1 + 2e^{-3t} - 3e^{-2t}$

Funcția de densitate de probabilitate PDF

Notată $f(x)$ – probabilitatea ca un eveniment să se petreacă în $[t, t+\Delta t]$

Proprietate evidentă : $\int_0^\infty f(x)\,dx = 1$ și $f(x) = \frac{dF(x)}{dx}$

$$F(t) = \int_0^t f(z)\,dz$$

$$P(X \geq t) = \int_t^\infty f(z)\,dz$$

$$P(a \leq X \leq b) = \int_a^b f(z)\,dz = F(b) - F(a)$$

Ex: $F(t) = 1 + 2e^{-3t} - 3e^{-2t} \Rightarrow$

$$f(t) = -6e^{-3t} + 6e^{-2t}$$

Valoarea așteptată

$$E[X] = p_1 x_1 + p_2 x_2 + \dots p_n x_n = \frac{p_1 x_1 + \dots p_n x_n}{1} = \frac{p_1 x_1 + p_2 x_2 + \dots p_n x_n}{p_1 + p_2 + \dots p_n}$$

Ex: Pentru un zar cu 6 fețe:

$$E[X] = \frac{1}{6} \cdot 1 + \frac{1}{6} \cdot 2 + \dots + \frac{1}{6} \cdot 6 = 3,5$$

Dacă avem ca variabilă aleatoare timpul (var. continuă):

$$E[t] = \int_0^\infty t \cdot f(t)\, dt, \ \forall\, t \geq 0$$

Ex: $f(t) = -6e^{-3t} + 6e^{-2t} \Rightarrow E[t] = \int_0^\infty t \cdot \left(-6e^{-3t} + 6e^{-2t}\right) dt =$

$$= -6\int_0^\infty t e^{-3t} dt + 6\int_0^\infty t e^{-2t} dt = \dots = \frac{5}{6}$$

Distribuții de probabilitate

în funcție de tipul variabilei aleatoare
- discrete
- continue

Distribuția binomială

Funcția de probabilitate a distribuției: $p(x)$

$$p(x) = C_n^x \, p^x (1-p)^{n-x}$$

Exemplu: Rată de defect la becuri 5%. Dacă inspectezi 100 de becuri, care e prob. să găsești 2 becuri defecte?

$$P(x=2) = C_{100}^2 \, 0{,}05^2 (1-0{,}05)^{98} \simeq 0{,}081 \; (8{,}1\%)$$

Distributia Poisson

$$P(x) = \frac{\lambda^x e^{-\lambda}}{x!}$$, unde λ este rata de apariție a unui eveniment
$\lambda = p \times n$

Exemplu: pentru ex. cu becurile o să avem $\lambda = 100 \cdot 0{,}05 = 5$

$$P(x=2) = \frac{5^2 e^{-5}}{2!} = \ldots = 0{,}09 \ (9\%)$$

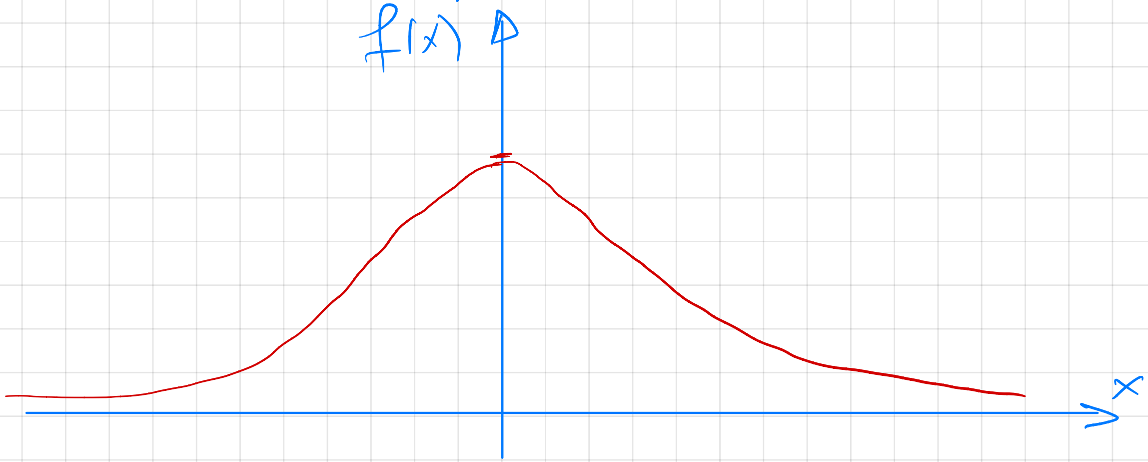
Distribuția normală (gaussiană)

Distribuție continuă → funcție densitate probabilitate

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$$, unde σ – deviația standard
μ – medie

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n-1}}$$

$$\mu = \frac{\sum_{i=1}^{n} x_i}{n}$$

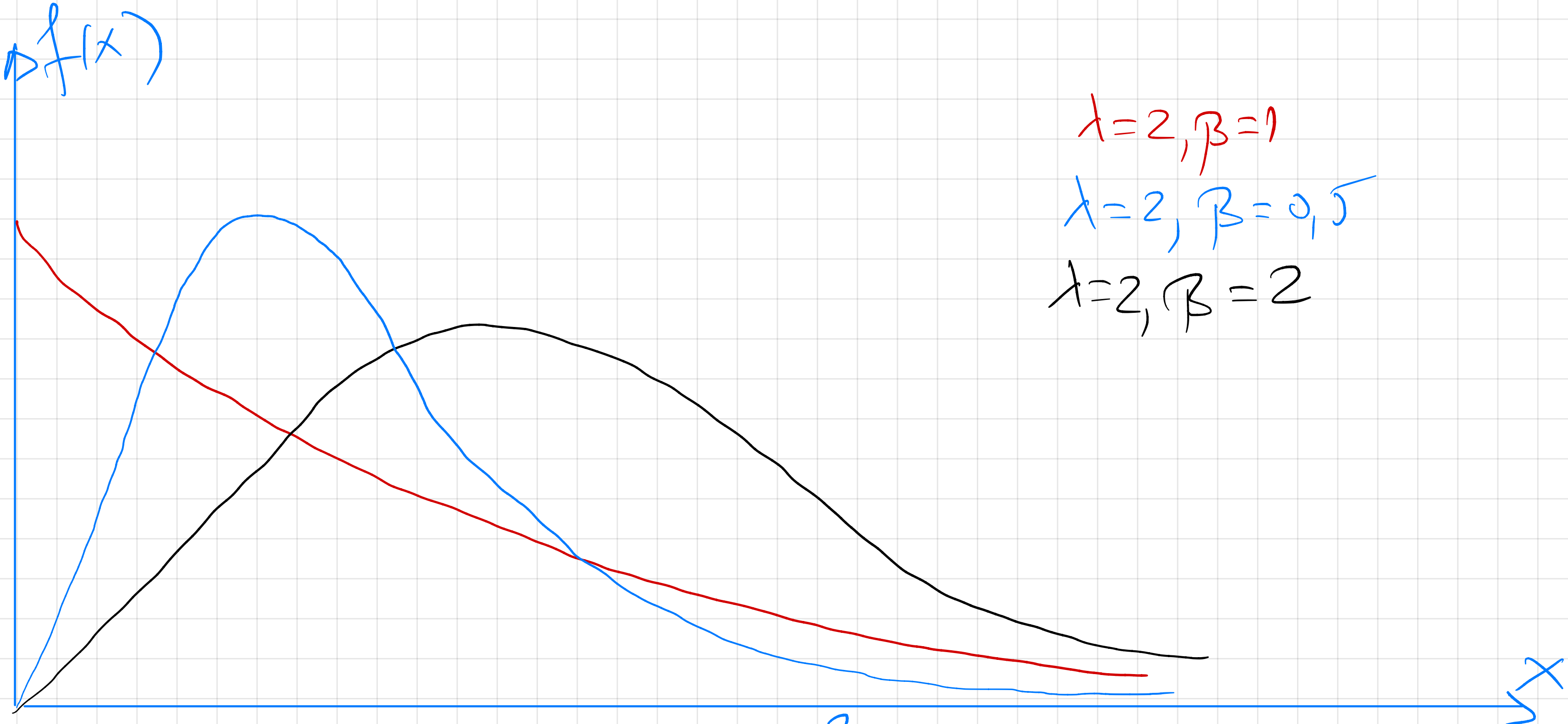


Distribuția Weibull

$$f(x) = \frac{\beta}{\lambda}\left(\frac{x}{\lambda}\right)^{\beta-1} e^{-\left(\frac{x}{\lambda}\right)^{\beta}},$$ unde

$\lambda \to$ parametrul de scală

$\beta \to$ parametrul de formă

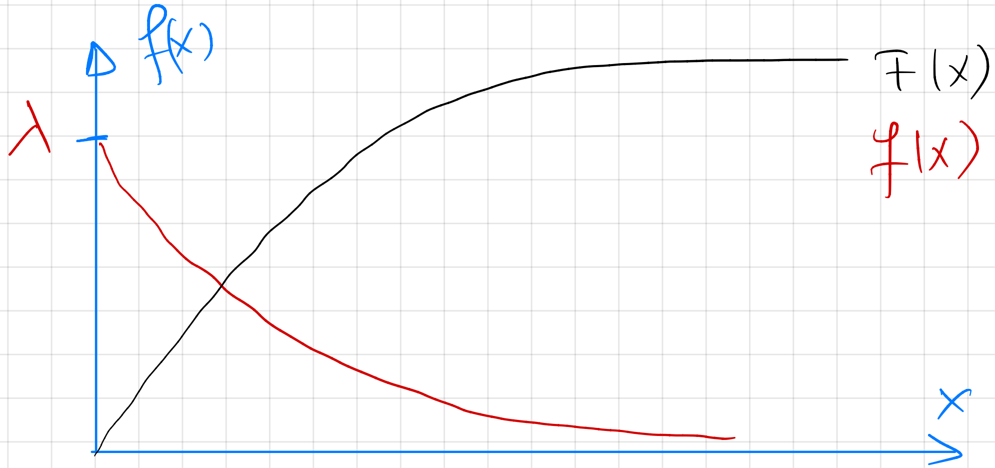


$$F(x) = \int f(x)\,dx = 1 - e^{-\left(\frac{x}{\lambda}\right)^{\beta}}$$

Distribuția exponențială

$f(x) = \lambda e^{-\lambda x}$ → Caz special al Weibull cu $\beta = 1$

$F(x) = 1 - e^{-\lambda x}$



Exemplu: server web, cererile vin aleatoriu. Avem o rată medie de 10 cereri pe minut ($\lambda = 10$).

Probabilitatea să primești o cerere în următorul minut:

$$F(1) = 1 - e^{-\lambda \cdot 1} = 1 - e^{-10} = 0{,}9995 \ (99{,}95\%)$$

$$f(1) = \lambda e^{-\lambda \cdot 1} = 10 e^{-10} \simeq 0{,}0005$$ → prob. cerere la minutul 1 (mică)

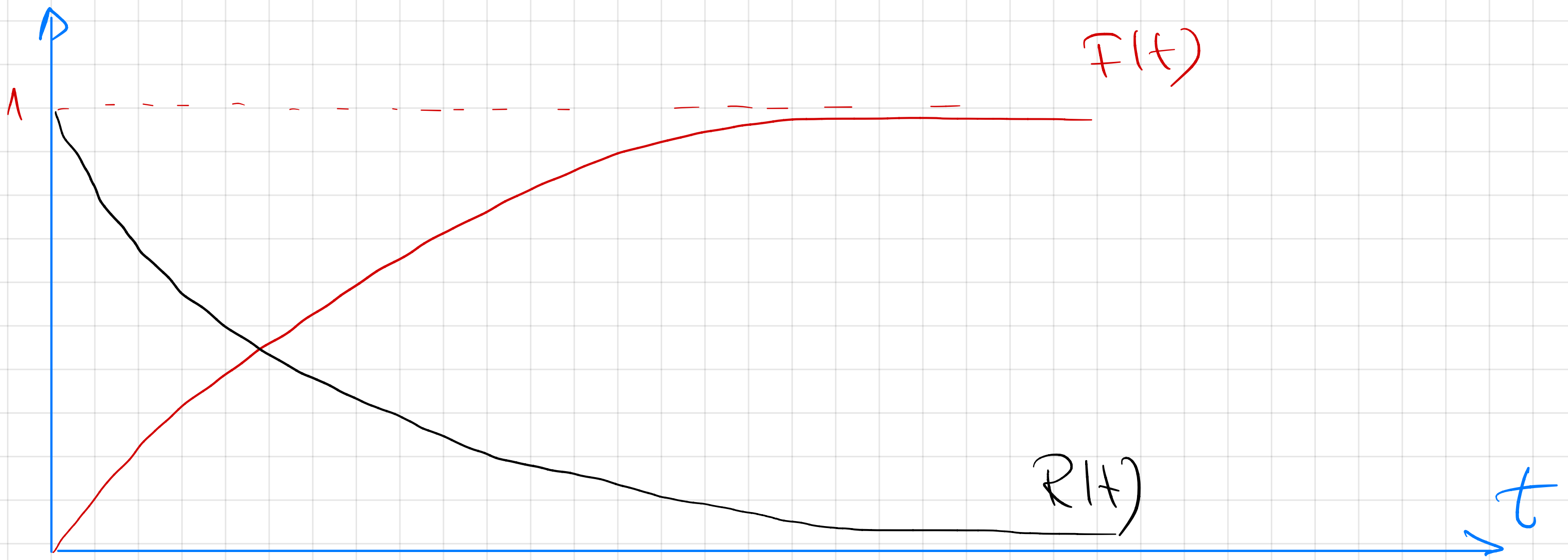
Modelarea fiabilității

Definim o funcție de **fiabilitate $R(t)$**

$R(t) = P($ sistemul este funcțional în $[0,t)$ $)$

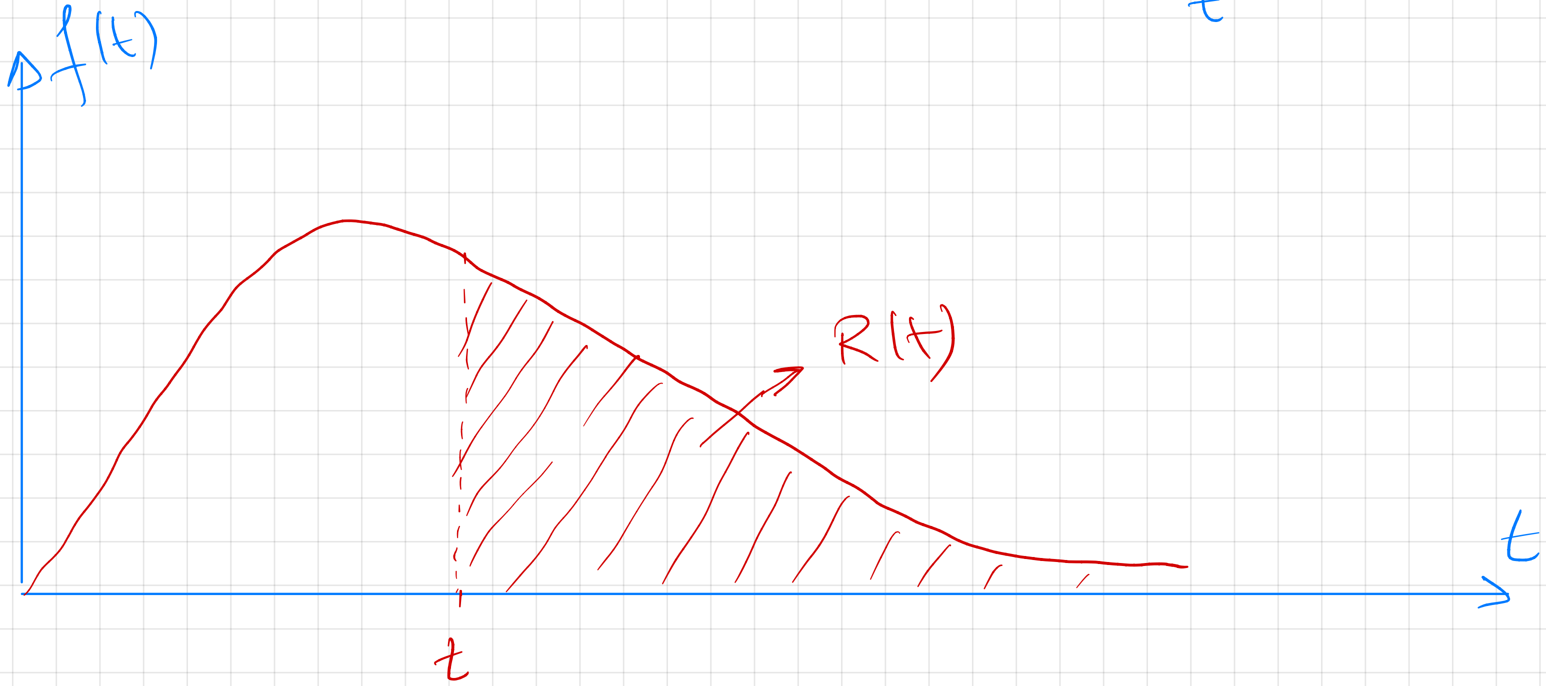
$F(t) \to$ CDF (prob. să apară un defect în $[0,t]$)

$$\left.\begin{aligned} R(t) &= P(X > t) \\ F(t) &= P(X \le t) \end{aligned}\right| \Rightarrow R(t) = 1 - F(t)$$



Reprezentarea grafică a fiabilității

$$F(t) = \int_0^t f(z)\,dz \Rightarrow R(t) = \int_t^{\infty} f(z)\,dz$$



Funcția de fiabilitate poate fi interpretată ca aria de sub graficul funcției de densitate de probabilitate, de la un moment de timp la ∞.

Intensitatea defecțiunilor (failure rate)

Probabilitatea ca sistemul să se defecteze în $[t+\Delta t]$ cu condiția ca sistemul să funcționeze până la t.

$$P(t < X < t+\Delta t \mid X > t) = \frac{P(t<X<t+\Delta t)}{P(X>t)} = \frac{F(t+\Delta t)-F(t)}{1-F(t)}$$

Definim intensitatea defecțiunilor:

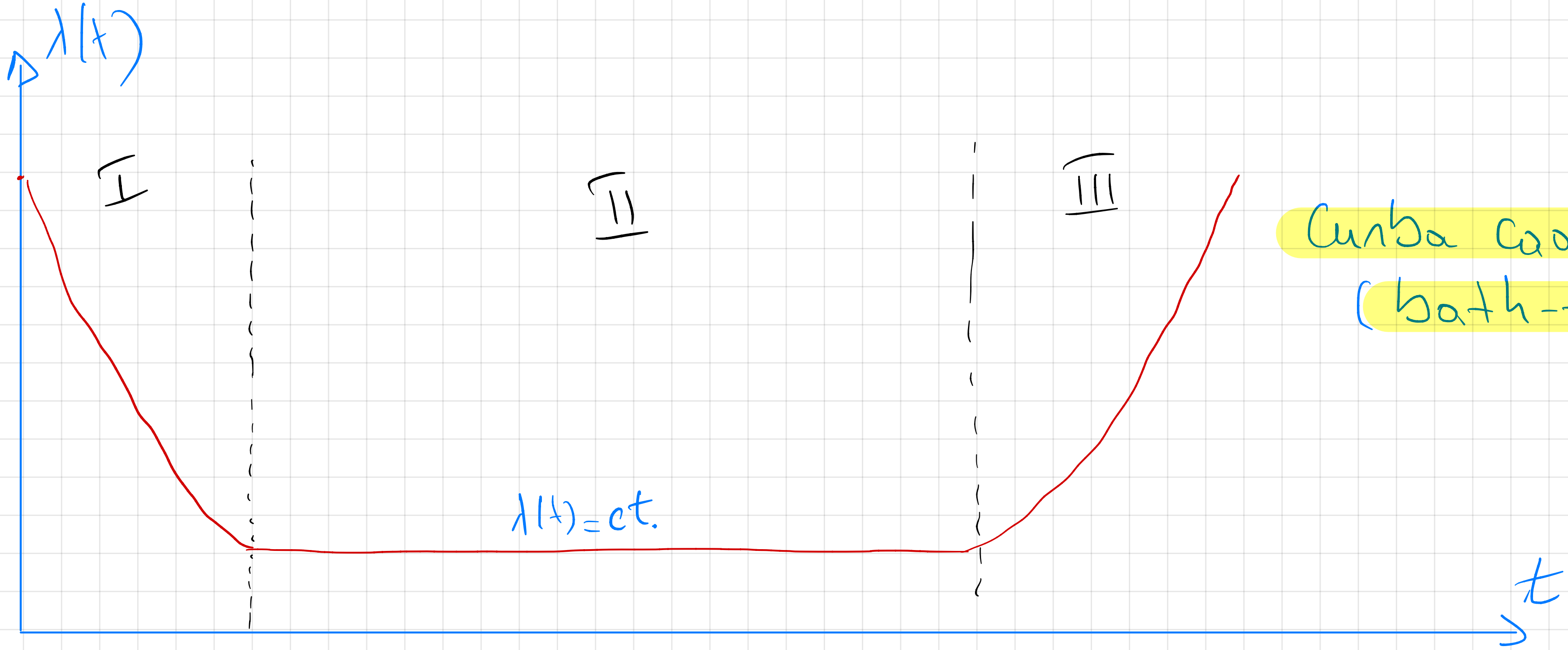
$$\lambda(t) = \lim_{\Delta t \to 0} \frac{F(t+\Delta t)-F(t)}{[1-F(t)]\,\Delta t} = \lim_{\Delta t \to 0} \frac{R(t)-R(t+\Delta t)}{R(t)\,\Delta t}$$

dar $R(t) = 1-F(t)$ $\Rightarrow$

$$\lambda(t) = \frac{1}{R(t)} \lim_{\Delta t \to 0} \frac{R(t)-R(t+\Delta t)}{\Delta t} = \frac{1}{R(t)}\left[-\frac{dR(t)}{dt}\right] = \frac{f(t)}{R(t)}$$

Dar $f(t) = \frac{dF(t)}{dt} = \frac{d(1-R(t))}{dt} = -\frac{dR(t)}{dt}$

Deci $\lambda(t) = \dfrac{f(t)}{R(t)}$, sau $\lambda(t) = -\dfrac{1}{R(t)}\dfrac{dR(t)}{dt}$



Curba cadă de baie
(bath-tub curve)

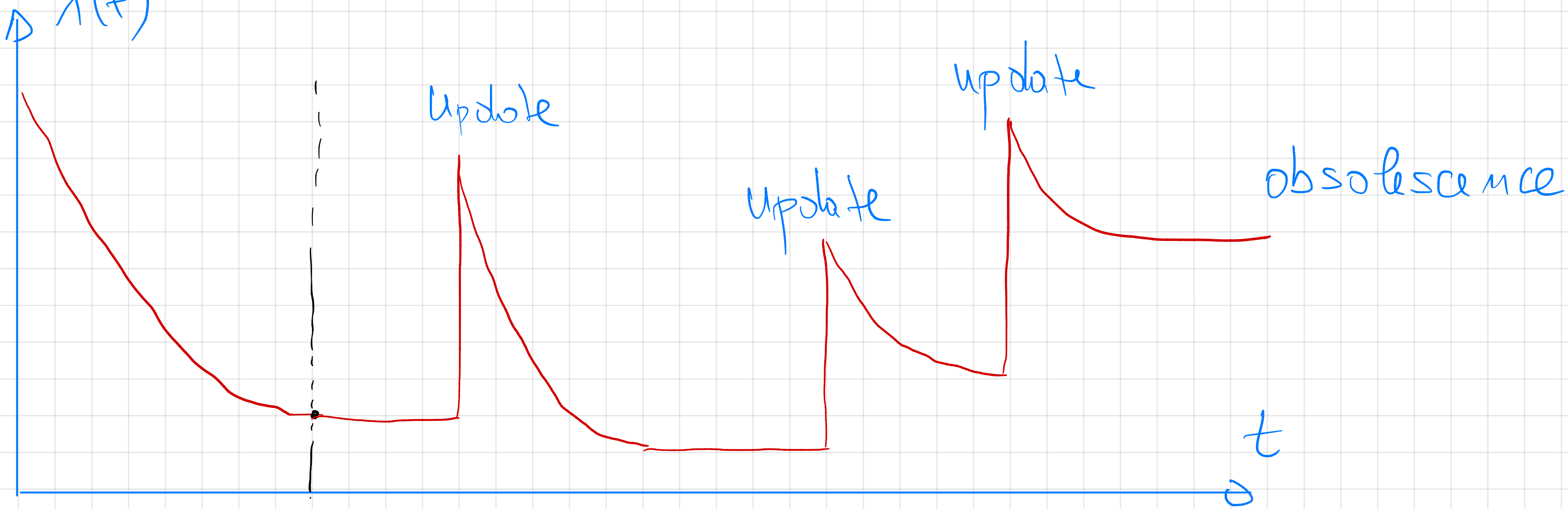
Zone distincte în profilul $\lambda(t)$:

I Mortalitate infantilă

II Zona de utilizare normală a produsului ($\lambda = ct$)

III Îmbătrânire (wear-out)

Intensitatea defectiunilor pentru software



Daca $\lambda(t) = ct. = \lambda \Rightarrow \lambda(t) = \frac{f(t)}{R(t)} = -\frac{1}{R(t)} \frac{dR(t)}{dt} \Rightarrow$

$\Rightarrow \lambda = -\frac{1}{R(t)} \frac{dR(t)}{dt} \Rightarrow -\lambda \, dt = \frac{1}{R(t)} dR(t) \; \Big| \cdot \int \Rightarrow -\lambda \int dt = \int \frac{1}{R(t)} dR(t)$

$\Rightarrow -\lambda(t+c) = \ln(R(t)) \Rightarrow R(t) = e^{-\lambda(t+c)}$

Dar $R(0) = 1 \Rightarrow e^{-\lambda(0+c)} = 1 \Rightarrow e^{-\lambda \cdot c} = 1 \Big|_{\lambda \neq 0} \Rightarrow c = 0 \Rightarrow$ $\boxed{R(t) = e^{-\lambda t}}$

Dacă $\lambda(t)$ nu este constant, modelăm folosind distribuția Weibull

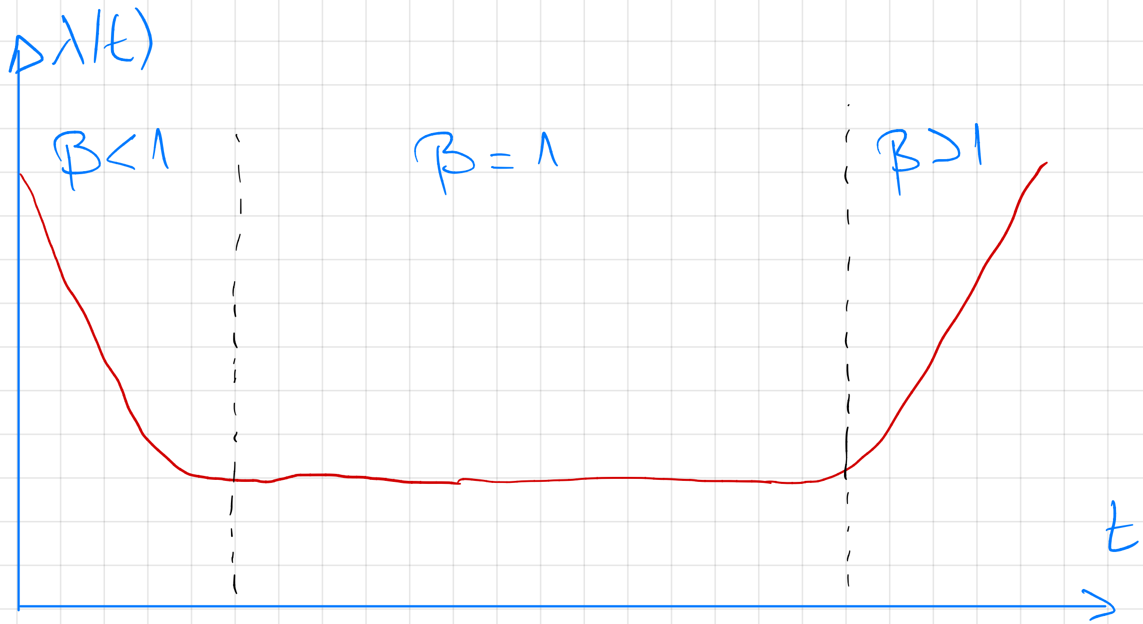
$$\lambda(t) = \lambda \beta t^{\beta-1}$$

$\beta > 1$: $\lambda(t) \nearrow$ – îmbătrânire

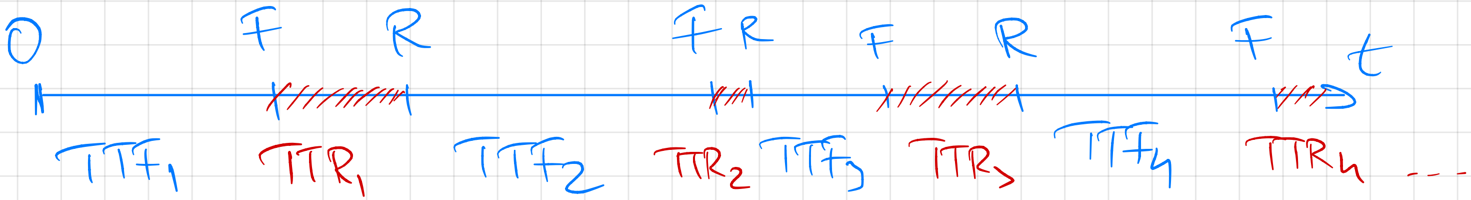
$\beta = 1$: $\lambda(t) = ct = \lambda$

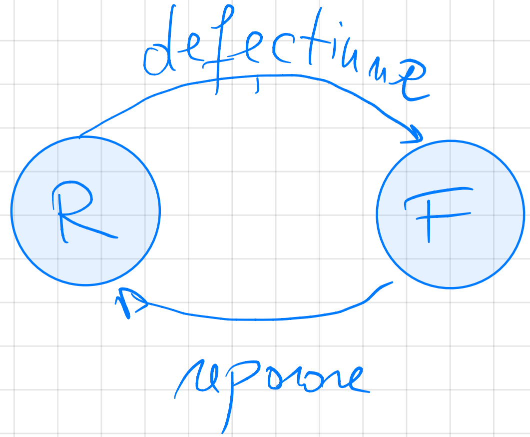
$\beta < 1$: $\lambda(t) \downarrow$ – mortalitate infantilă

$$R(t) = e^{-\lambda t^{\beta}}$$



Media timpului de bună funcționare





$$MTBF = \frac{\sum TTF_i}{n} \qquad MTTR = \frac{\sum TTR_i}{n}$$

A - disponibilitate

$$A = \frac{\sum TTF_i}{\sum TTF_i + \sum TTR_i} = \frac{MTBF}{MTBF + MTTR}$$

Exemplu: MTBF = 2 ore (7200 s), MTTR = 3 s

$$A = \frac{7200}{7200+3} = 99{,}96\%$$

Exemplu:

Server email: 99,99% disponibilitate 4 de 9 → down 52,56 min/an

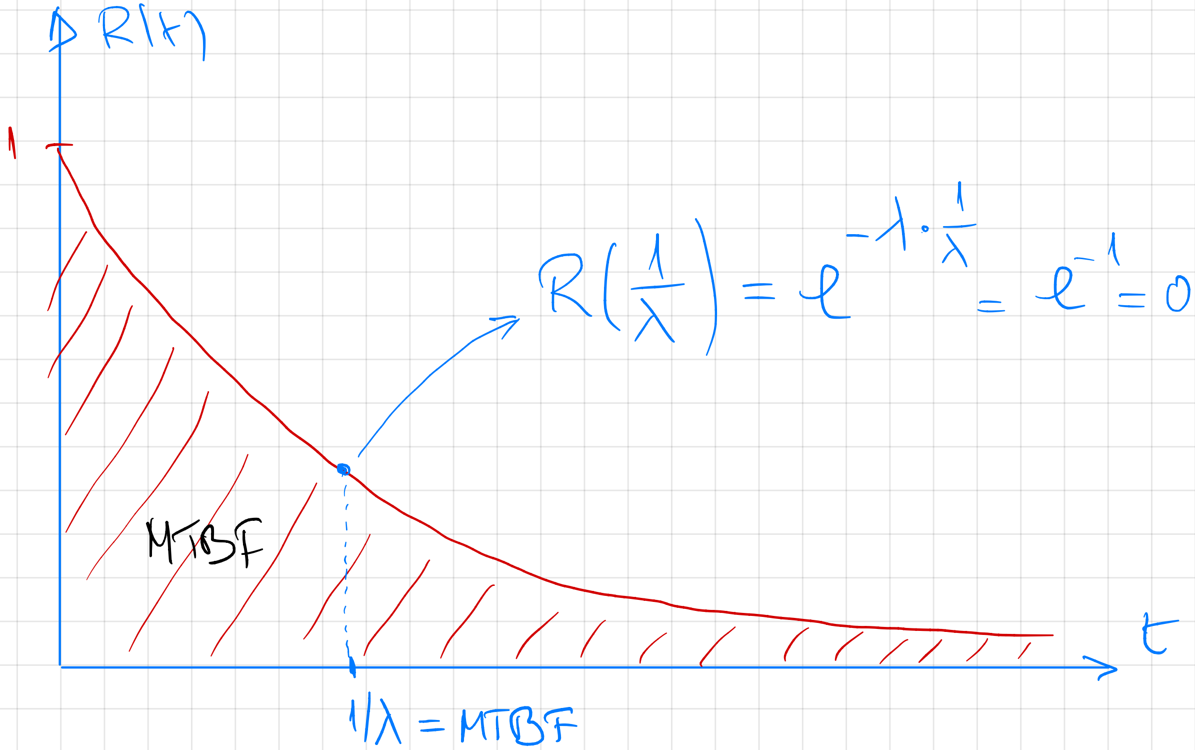
Sisteme critice: 99,9999%, 6 de 9, va fi indisponibil 31,5 s/an

$$MTBF = E[T] = \int_0^\infty t \cdot f(t)\, dt$$

Dan $f(t) = -\dfrac{dR(t)}{dt} \Rightarrow MTBF = \displaystyle\int_0^\infty t \cdot \left(-\frac{dR(t)}{dt}\right) dt =$

$$= -\int_0^\infty t \cdot R'(t)\, dt = -\Big[\underbrace{t \cdot R(t)\Big|_0^\infty}_{=0} - \int_0^\infty R(t)\, dt \Big] = \int_0^\infty R(t)\, dt$$

$$MTBF = \int_0^\infty R(t)\, dt$$



$$R\left(\frac{1}{\lambda}\right) = e^{-\lambda \cdot \frac{1}{\lambda}} = e^{-1} = 0{,}37$$

Dacă $R(t) = e^{-\lambda t} \Rightarrow MTBF = \int_0^\infty e^{-\lambda t} dt = -\frac{1}{\lambda} e^{-\lambda t} \Big|_0^\infty =$

$= -\frac{1}{\lambda}(0 - 1) = \frac{1}{\lambda}$

$$MTBF = \frac{1}{\lambda}$$

Dacă $\lambda(t) \neq ct \Rightarrow R(t) = e^{-\lambda t^\beta}$ (distribuția Weibull)

$$MTBF = \int_0^\infty R(t)\, dt = \int_0^\infty e^{-\lambda t^\beta} dt = \frac{\Gamma(\beta^{-1})}{\beta \lambda^{\beta^{-1}}}$$

$\Gamma(x)$ - funcția gamma - generalizarea factorialei pt $x \in \mathbb{R}$

$\Gamma(0) = \Gamma(1) = 1$

$\Gamma(x+1) = x \cdot \Gamma(x)$, $\forall x > 1$

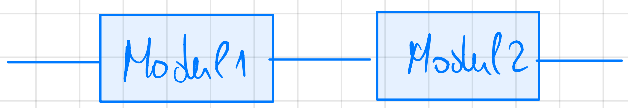
Dacă x este întreg pozitiv $\Gamma(x) = (x-1)!$

$$\Gamma(x) = \int_0^\infty y^{x-1} e^{-y} dy$$

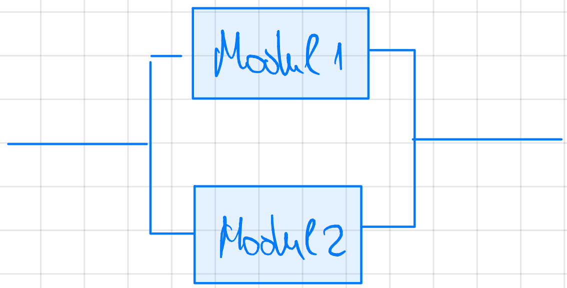
Modelarea fiabilității – diagrame bloc

Folosim diagrame bloc pentru a descrie un sistem din punct de vedere fiabilistic

De exemplu:



structuri serie



Structuri paralele

$R_i(t)$ – fiabilitatea modulului i

$Q_i(t)$ – inversul fiabilității, $Q_i(t) = 1 - R_i(t)$

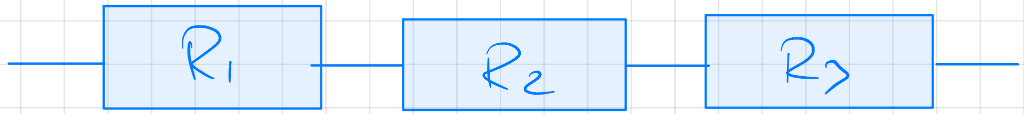
Structura Serie

$$R_{serie} = R_1 \cdot R_2 \cdots R_n = \prod_{i=1}^{n} R_i$$

$$Q_{serie} = 1 - R_{serie} = 1 - \prod_{i=1}^{n} (1 - Q_i) = 1 - \Big(1 - (Q_1 + Q_2 + \dots Q_n) + (Q_1 Q_2 + \dots Q_{n-1} Q_n) - (Q_1 Q_2 Q_3 + \dots) + \dots \prod_{i=1}^{n} Q_i\Big) \simeq 1 - \big(1 - (Q_1 + Q_2 + \dots Q_n)\big) = \sum_{i=1}^{n} Q_i$$

$$Q_{serie} \simeq \sum_{i=1}^{n} Q_i$$

Exemplu: 3 module cu fiabilitate R_1, R_2 și R_3



La un anumit timp t: $R_1 = 0,9$, $R_2 = 0,3$; $R_3 = 0,5$

$$R_{serie} = R_1 \cdot R_2 \cdot R_3 = 0,9 \cdot 0,3 \cdot 0,5 = 0,27 \cdot 0,5 = 0,135$$

Dacă $R_1 = R_2 = R_3 = 0,99 \Rightarrow R_{serie} = 0,99^3 = 0,97$ (97%)

$$Q_{serie} \simeq Q_1 + Q_2 + Q_3 = 0,01 + 0,01 + 0,01 = 0,03$$

Dacă $R_i(t) = e^{-\lambda_i t}$

$$R_{serie} = \prod_{i=1}^{n} R_i(t) = \prod_{i=1}^{n} e^{-\lambda_i t} = e^{-\sum_{i=1}^{n} \lambda_i t} = e^{-\lambda_s t}$$

$$\lambda_{serie} = \sum_{i=1}^{n} \lambda_i$$

$$MTBF_{serie} = \int_0^{\infty} R_{serie}(t)\,dt = \int_0^{\infty} e^{-\sum_{i=1}^{n} \lambda_i t}\,dt = \frac{1}{\sum_{i=1}^{n} \lambda_i} = \frac{1}{\lambda_s}$$

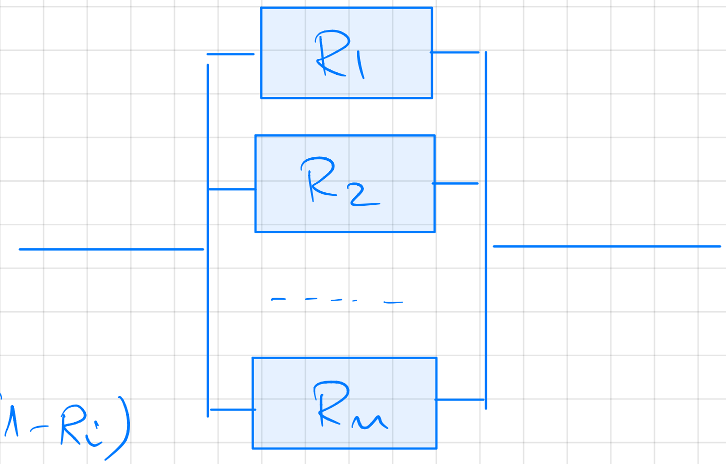
$$MTBF_{serie} = \frac{1}{\lambda_s}$$

Exemplu: 3 module cu $MTBF_1 = 5h$, $MTBF_2 = 7h$ și $MTBF_3 = 2h$

$$MTBF_{serie} = \frac{1}{\lambda_1 + \lambda_2 + \lambda_3} = \frac{1}{\frac{1}{MTBF_1} + \frac{1}{MTBF_2} + \frac{1}{MTBF_3}} = \frac{1}{{}^{14/}\frac{1}{5} + {}^{10/}\frac{1}{7} + {}^{35/}\frac{1}{2}} = \frac{1}{\frac{14+10+35}{70}} =$$

$$= \frac{70}{59} = 1{,}186\,h$$

Structura paralel



$$R_{paralel} = ?$$

$$Q_{paralel} = Q_1 \cdot Q_2 \cdot \ldots \cdot Q_n = \prod_{i=1}^{n} Q_i$$

$$R_{paralel} = 1 - Q_{paralel} = 1 - \prod_{i=1}^{n} Q_i = 1 - \prod_{i=1}^{n} (1 - R_i)$$

Exemplu: $R_1 = 0,9$; $R_2 = 0,7$; $R_3 = 0,5$

$$R_{paralel} = 1 - (1 - R_1)(1 - R_2)(1 - R_3) = 1 - 0,1 \cdot 0,7 \cdot 0,5 = 1 - 0,035 = 0,965$$

96,5 % !

Dacă $R_1 = R_2 = R_3 = 0,99$

$$R_{paralel} = 1 - (1 - 0,99)^3 = 1 - 0,01^3 = 1 - 0,000001 = 99,99\ldots \%$$

Dacă $R_i = e^{-\lambda_i t} \Rightarrow$

$$R_{paralel} = 1 - \prod_{i=1}^{n} \left(1 - e^{-\lambda_i t}\right)$$

$$MTBF_{paralel} = \int_0^\infty R_{paralel}(t)\,dt = \int_0^\infty \left[1 - \prod_{i=1}^{n} \left(1 - e^{-\lambda_i t}\right)\right] dt =$$

$$= \int_0^\infty \sum_{i=1}^{n} e^{-\lambda_i t} dt - \int_0^\infty \sum_{\substack{i,j=1 \\ i \neq j}}^{n} e^{-(\lambda_i+\lambda_j)t} + \int_0^\infty \dots + (-1)^{n+1} \int_0^\infty \prod_{i=1}^{n} e^{-\lambda_i t} =$$

$$= \sum_{i=1}^{n} \frac{1}{\lambda_i} - \sum_{\substack{i,j=1 \\ i \neq j}}^{n} \frac{1}{\lambda_i+\lambda_j} + \dots + (-1)^{n+1} \frac{1}{\sum_{i=1}^{n} \lambda_i}$$

Dacă sistemele sunt identice $\Rightarrow R_i(t) = e^{-\lambda t}$, $MTBF_i = MTBF = \frac{1}{\lambda}$

$$MTBF_{paralel} = \frac{n}{\lambda} - \frac{n}{2\lambda} + \dots (-1)^{n+1} \frac{1}{n\lambda} = \frac{1}{\lambda}\left(n - \frac{n}{2} + \frac{n}{3} - \dots + (-1)^{n+1}\frac{1}{n}\right) =$$

$$= \frac{1}{\lambda} \cdot \underbrace{\sum_{i=1}^{n} (-1)^{i+1} \frac{C_n^i}{i}}_{= \sum_{i=1}^{n} \frac{1}{i}} = \frac{1}{\lambda} \sum_{i=1}^{n} \frac{1}{i} \simeq \frac{\ln 2n}{\lambda}$$

$$MTBF_{paralel} = MTBF \cdot \sum_{i=1}^{n} \frac{1}{i}$$

Exemplu: $MTBF_1 = 5h$, $MTBF_2 = 7h$ și $MTBF_3 = 2h$

$\lambda_1 = 1/MTBF_1 = 1/5$, $\lambda_2 = 1/MTBF_2 = 1/7$ și $\lambda_3 = 1/MTBF_3 = 1/2$

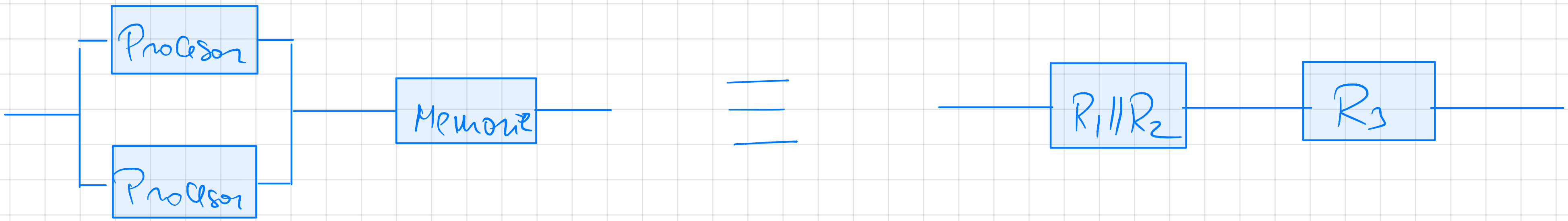
$$MTBF_{paralel} = \left(\frac{1}{\lambda_1} + \frac{1}{\lambda_2} + \frac{1}{\lambda_3}\right) - \left(\frac{1}{\lambda_1+\lambda_2} + \frac{1}{\lambda_1+\lambda_3} + \frac{1}{\lambda_2+\lambda_3}\right) + \frac{1}{\lambda_1+\lambda_2+\lambda_3} =$$

$$= (5+7+2) - \left(\frac{1}{\frac{1}{5}+\frac{1}{7}} + \frac{1}{\frac{1}{5}+\frac{1}{2}} + \frac{1}{\frac{1}{7}+\frac{1}{2}}\right) + \frac{1}{\frac{1}{5}+\frac{1}{7}+\frac{1}{2}} =$$

$$= 14 - \left(\frac{35}{12} + \frac{10}{7} + \frac{14}{9}\right) + \frac{70}{59} = 14 - (2{,}92 + 1{,}43 + 1{,}56) + 1{,}19 = 9{,}28\ h$$

Combinații serie - paralel

Exemplu: procesor dual-core cu o singură memorie RAM



$R_1,\ R_2,\ R_3 \Rightarrow R_{TOTAL} = (R_1 \| R_2) \cdot R_3 = \big(1 - (1-R_1)(1-R_2)\big) \cdot R_3$

$= (R_1 + R_2 - R_1 R_2) R_3 = R_1 R_3 + R_2 R_3 - R_1 R_2 R_3$

$R_1 = 0{,}9$

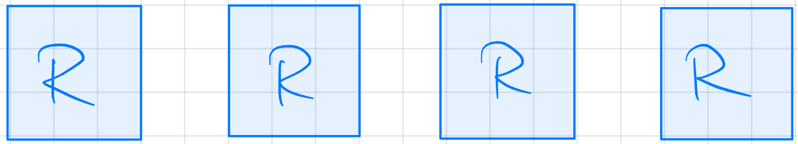
$R_2 = 0{,}3$ $\Rightarrow R_{TOTAL} = \ldots$

$R_3 = 0{,}7$

Structuri k din n

Exemplu: avion pasageri cu 4 motoare, tolerează maxim 2 motoare defecte.

$$R_{2/4} = R^4 + 4R^3(1-R) + 6R^2(1-R)^2$$



Caz general:

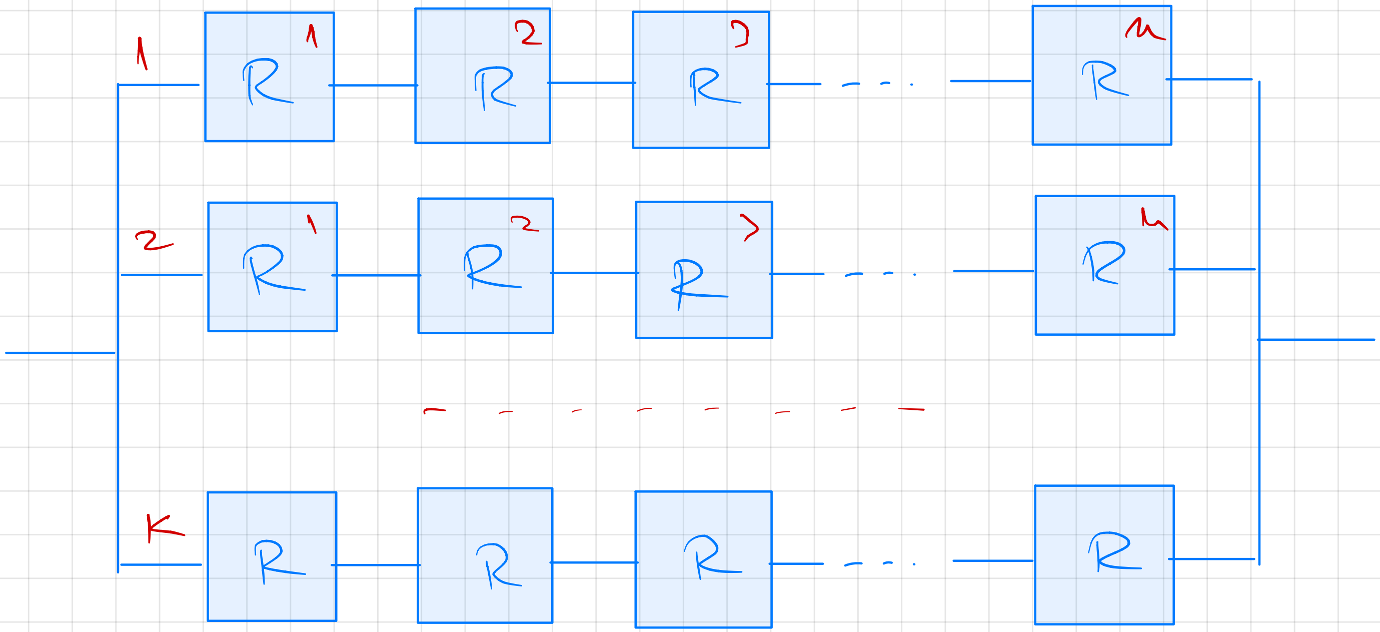
$$R_{k/n} = \sum_{i=k}^{n} C_n^i R^i (1-R)^{n-i}$$

$$R_{1/n} = \sum_{i=1}^{n} C_n^i R^i (1-R)^{n-i}$$ – structura paralel

$$R_{n/n} = R^n$$ – structura serie

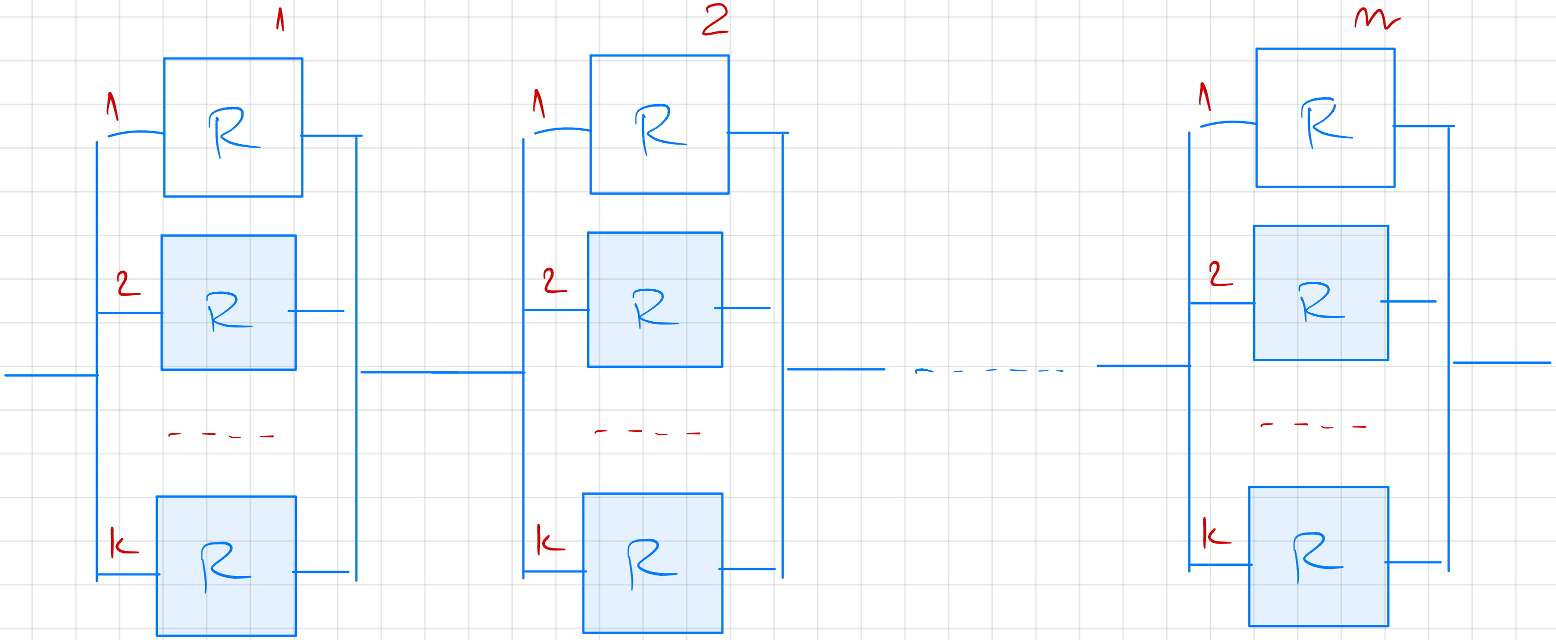
Structuri serie-paralel și paralel-serie

Serie-paralel



$$R_{SP}(t) = \underbrace{R^n \| R^n \| \ldots \| R^n}_{k \text{ ori}} = 1 - \left(1 - R^n(t)\right)^k$$

Paralel-Serie



$$R_{PS}(t) = \left[1-(1-R(t))^k\right]^m$$