Pravděpodobnost a matematická statistika

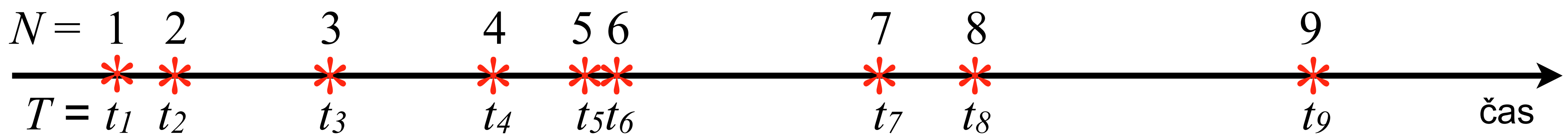
prof. RNDr. Gejza Dohnal, CSc.
dohnal@nipax.cz

II. Náhodná veličina

https://sms.nipax.cz/pas

Náhodné události v čase

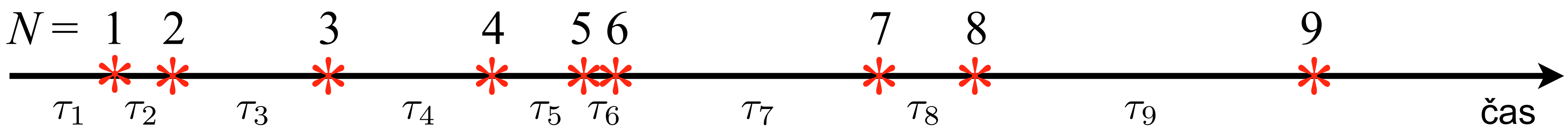
počet událostí v čase t: $N(t)$ diskrétní náhodná veličina



čas výskytu i-té události: T_i spojitá náhodná veličina

Náhodné události v čase

počet událostí v čase t: $N(t)$ ← diskrétní náhodná veličina



doby mezi událostmi: $\tau_i = t_i - t_{i-1}$

spojitá náhodná veličina

$$\lambda(t) = \frac{N(t)}{t} \xrightarrow[t \to \infty]{} \lambda = \lim_{t\to\infty} \frac{N(t)}{t}$$

střední počet událostí za jednotku času

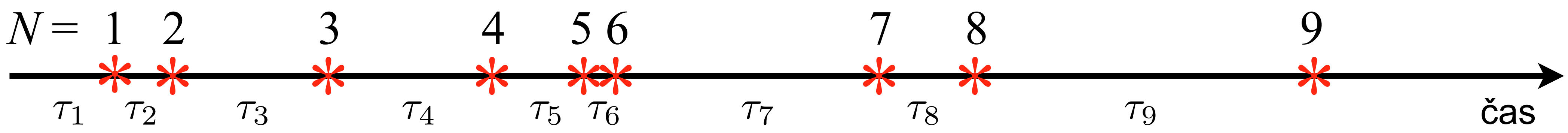
$$\bar{T}_i = \frac{t_i}{i} \xrightarrow[t \to \infty]{} \bar{T} = \lim_{t\to\infty} \frac{t}{N(t)} = \frac{1}{\lambda}$$

střední doba mezi událostmi

λ je intenzita událostí, $1/\lambda$ je střední doba mezi událostmi

Náhodné události v čase

počet událostí v čase t: $N(t)$



doby mezi událostmi: $\tau_i = t_i - t_{i-1}$

Jaká je pravděpodobnost, že doba mezi událostmi nepřekročí hodnotu t ?

$$P(\tau \leq t) = ?$$

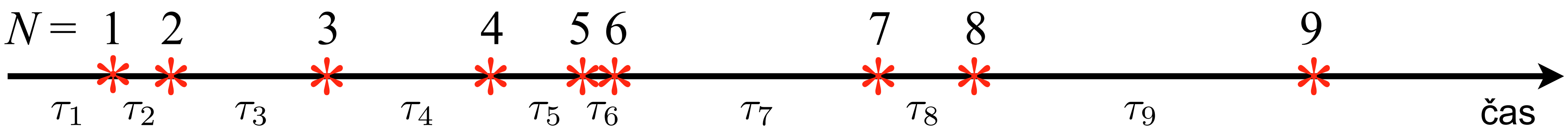
Jaká je pravděpodobnost, že za dobu t nastane k událostí?

$$P(N(t) = k) = ?$$

λ je intenzita událostí, $1/\lambda$ je střední doba mezi událostmi

Náhodné události v čase

počet událostí v čase t: $N(t)$



doby mezi událostmi: $\tau_i = t_i - t_{i-1}$

Jaká je pravděpodobnost, že doba mezi událostmi nepřekročí hodnotu t ? $P(\tau \leq t) =?$

Exponenciální rozdělení s parametrem λ :

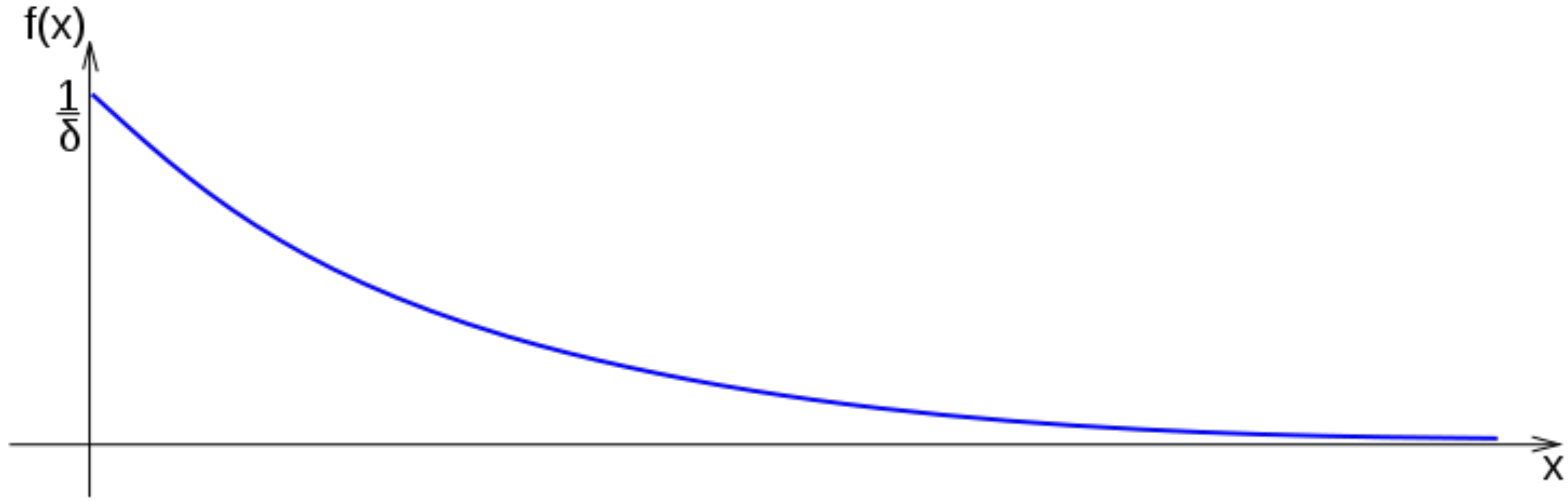
$$P(\tau \leq t) = 1 - \mathrm{e}^{\lambda t}, \quad t \geq 0$$

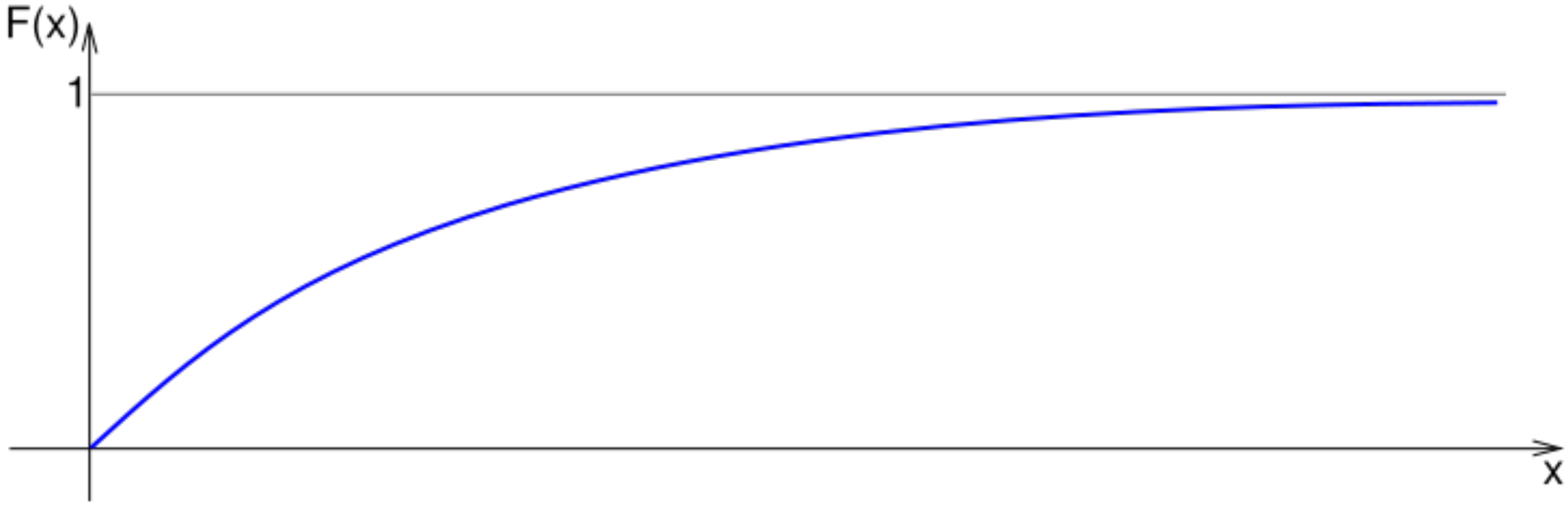
$$F(t) = 1 - \mathrm{e}^{\lambda t} \;\Rightarrow\; f(t) = \lambda \mathrm{e}^{\lambda t},\; t \geq 0, \qquad E(\tau) = \frac{1}{\lambda}, \quad Var(\tau) = \lambda^2$$

λ je intenzita událostí, $1/\lambda$ je střední doba mezi událostmi

Exponenciální rozdělení

$$f(t) = \begin{cases} 0, & t < 0 \\ \lambda e^{-\lambda t}, & t \geq 0 \end{cases} \qquad F(t) = \begin{cases} 0, & t < 0 \\ 1 - e^{-\lambda t}, & t \geq 0 \end{cases}$$





Exponenciální rozdělení

$$f(t) = \begin{cases} 0, & t < 0 \\ \lambda e^{-\lambda t}, & t \geq 0 \end{cases} \qquad F(t) = \begin{cases} 0, & t < 0 \\ 1 - e^{-\lambda t}, & t \geq 0 \end{cases}$$

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

$$f(t) = \begin{cases} 0, & t < 0 \\ \frac{1}{\delta} e^{-\frac{t}{\delta}}, & t \geq 0 \end{cases} \qquad F(t) = \begin{cases} 0, & t < 0 \\ 1 - e^{-\frac{t}{\delta}}, & t \geq 0 \end{cases}$$

Exponenciální rozdělení je rozdělení pravděpodobnosti náhodné veličiny, která popisuje dobu mezi nezávislými, náhodně se vyskytujícími událostmi v čase.

Přitom střední doba mezi těmito událostmi je rovna δ a střední počet těchto událostí za jednotku času je λ.

Exponenciální rozdělení

$$\boxed{P(\tau \leq t + s | \tau \geq s)} = \frac{P\big((\tau \leq t + s) \ \& \ (\tau \geq s)\big)}{P(\tau \geq s)}$$

$$= \frac{P\big(\tau \in \langle s, t + s\rangle\big)}{P(\tau \geq s)} = \frac{F(t + s) - F(s)}{1 - F(s)}$$

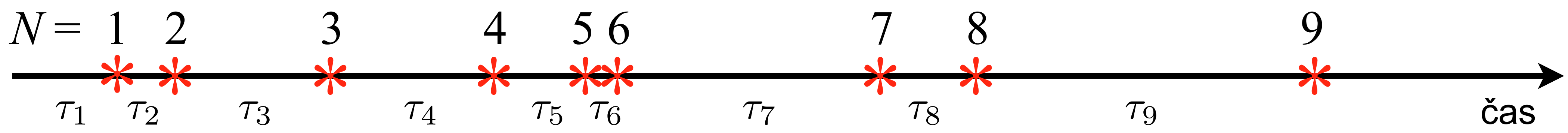
$$= \frac{1 - e^{-\frac{t+s}{\delta}} - 1 + e^{-\frac{s}{\delta}}}{e^{-\frac{s}{\delta}}} = \frac{e^{-\frac{s}{\delta}} - e^{-\frac{t}{\delta}} e^{-\frac{s}{\delta}}}{e^{-\frac{s}{\delta}}}$$

$$= 1 - e^{-\frac{t}{\delta}} = F(t) \boxed{= P(\tau \leq t)}$$

Exponenciální rozdělení nemá paměť!

Náhodné události v čase

počet událostí v čase t: $N(t)$



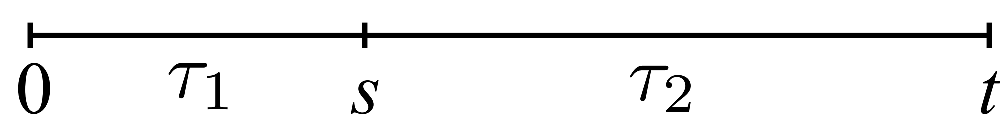
doby mezi událostmi: $\tau_i = t_i - t_{i-1}$

Za jak dlouho nastane k událostí ?

Jaké je rozdělení časů t_k ? $P(t_k \leq t) =?$

k=2: $P(t_2 \leq t) =?$

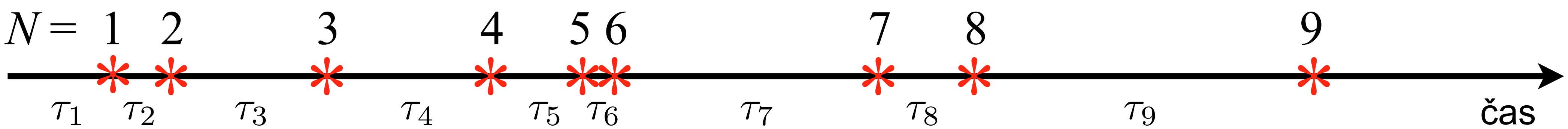
$t_2 = \tau_1 + \tau_2$



$$P(t_2 \leq t) = \sum_{i=1}^{n} P(\tau_1 \in \langle s_{i-1}, s_i \rangle) P(\tau_2 \leq t - s_i \,|\, \tau_1 \in \langle s_{i-1}, s_i \rangle))$$

Náhodné události v čase

počet událostí v čase t: $N(t)$



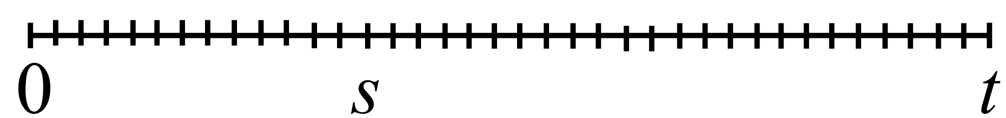
doby mezi událostmi: $\tau_i = t_i - t_{i-1}$

Za jak dlouho nastane k událostí ?

Jaké je rozdělení časů t_k ?

k=2: $P(t_2 \le t) = ?$

$t_2 = \tau_1 + \tau_2$



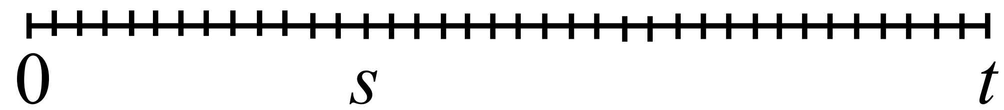
$$P(t_2 \le t) = \sum_{i=1}^{n} P(\tau_1 \in \langle s_{i-1}, s_i \rangle) P(\tau_2 \le t - s_i \,|\, \tau_1 \in \langle s_{i-1}, s_i \rangle))$$

$$P(t_2 \le t) = \sum_{i=1}^{n} P(\tau_1 \in \langle s_{i-1}, s_i \rangle) P(\tau_2 \le t - s_i\ )$$

Součet náhodných veličin s Exponenciálním rozdělením

k=*2*: $P(t_2 \le t) = ?$

$t_2 = \tau_1 + \tau_2$



$$P(t_2 \le t) = \sum_{i=1}^{n} P(\tau_1 \in \langle s_{i-1}, s_i \rangle) P(\tau_2 \le t - s_i \,|\, \tau_1 \in \langle s_{i-1}, s_i \rangle))$$

$$P(t_2 \le t) = \sum_{i=1}^{n} P(\tau_1 \in \langle s_{i-1}, s_i \rangle) P(\tau_2 \le t - s_i\ )$$

$$= \sum_{i=0}^{n-1} P(\tau_1 \in \langle s_i, s_i + ds\ \rangle) P(\tau_2 \le t - s_i - ds\ )$$

$$\doteq \sum_{i=0}^{n-1} f(s_i).ds\, F(t - s_i - ds\ ) \qquad \to \int_0^t F(t-s) f(s)\, \mathrm{d}s$$

$$\int_0^t (1 - \mathrm{e}^{-\lambda(t-s)}) \lambda \mathrm{e}^{-\lambda s} \mathrm{d}s = 1 - \mathrm{e}^{-\lambda t} - \lambda - \mathrm{e}^{-\lambda t} = 1 - \mathrm{e}^{-\lambda t}(1 + \lambda)$$

Součet náhodných veličin s Exponenciálním rozdělením

k=2: $F_2(t) = 1 - \mathrm{e}^{-\lambda t}(1 + \lambda)$ $\qquad f_2(t) = \lambda \mathrm{e}^{-\lambda t}(1 + \lambda)$

k=3: $P(t_3 \le t) = \sum_{i=1}^{n} P(\tau_3 \in \langle s_{i-1}, s_i \rangle) P(\tau_1 + \tau_2 \le t - s_i)$

$$F_3 = \int_0^t F_2(t-s) f(s) \, \mathrm{d}s$$

k: $P(t_k \le t) = \sum_{i=1}^{n} P(\tau_k \in \langle s_{i-1}, s_i \rangle) P(\tau_1 + \cdots + \tau_{k-1} \le t - s_i)$

$$F_k = \int_0^t F_{k-1}(t-s) f(s) \, \mathrm{d}s$$

k-tá konvoluční mocnina F(t)

$$F_k(t) = 1 - \mathrm{e}^{-\lambda t} \sum_{i=0}^{k-1} \frac{(\lambda t)^i}{i!}, \quad \text{pro } t \ge 0$$

$$F_k(t) = 0, \quad \text{pro } t < 0$$

Erlangovo rozdělení

Rozdělení součtu k náhodných veličin s exponenciál- ním rozdělením s parametrem se nazývá Erlangovo:

$$F_k(t) = 0, \quad \text{pro } t < 0$$

$$F_k(t) = 1 - \mathrm{e}^{-\lambda t} \sum_{i=0}^{k-1} \frac{(\lambda t)^i}{i!}, \quad \text{pro } t \geq 0$$

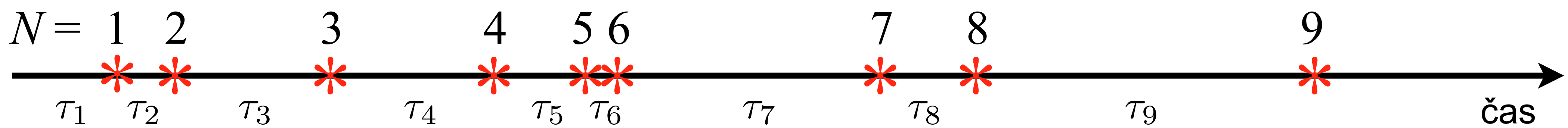
$$f_k(t) = 0, \quad \text{pro } t < 0$$

$$f_k(t) = \lambda \mathrm{e}^{-\lambda t} \frac{(\lambda t)^{k-1}}{(k-1)!}, \quad \text{pro } t \geq 0$$

$$E(X) = \frac{k}{\lambda}, \quad Var(X) = \frac{k}{\lambda^2}.$$

Náhodné události v čase

počet událostí v čase t: $N(t)$



doby mezi událostmi: $\tau_i = t_i - t_{i-1}$

Za jak dlouho nastane k událostí ?

Jaké je rozdělení časů t_k ? $P(t_k \leq t) =?$

Jaká je pravděpodobnost, že za dobu t nastane k událostí? $P(N(t) = k) =?$

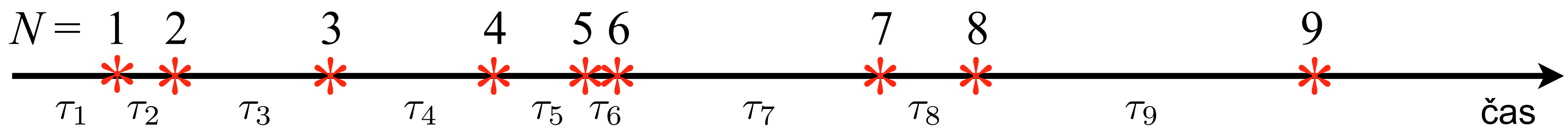
$$P(N(t) < k) = P(t_k \geq t) = 1 - F_k(t) = \mathrm{e}^{-\lambda t} \sum_{i=0}^{k-1} \frac{(\lambda t)^i}{i!}, \quad \text{pro } t \geq 0$$

odtud:

$$P(N(t) = k) = P(N(t) < k+1) - P(N(t) < k) = \mathrm{e}^{-\lambda t} \frac{(\lambda t)^k}{k!}$$

Náhodné události v čase

počet událostí v čase t: $N(t)$



doby mezi událostmi: $\tau_i = t_i - t_{i-1}$

Za jak dlouho nastane k událostí ?

Jaké je rozdělení časů t_k ? $P(t_k \leq t) = ?$

Jaká je pravděpodobnost, že za dobu t nastane k událostí? $P(N(t) = k) = ?$

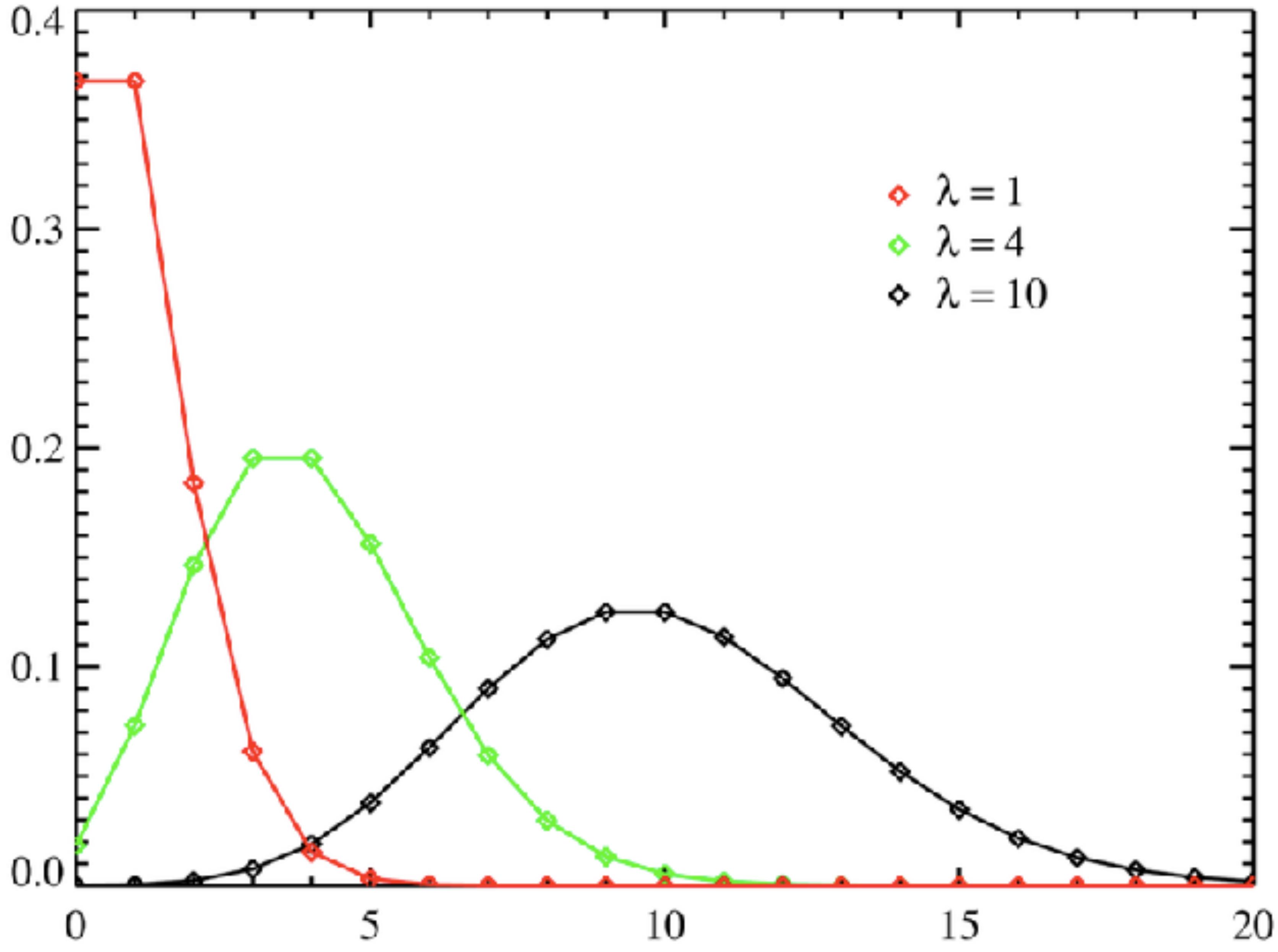
$$P(N(t) = k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \quad k = 0, 1, \ldots$$

$$E(N(t)) = Var(N(t)) = \lambda t$$

Poissonovo rozdělení

$$P(N(t) = k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \quad k = 0, 1, \ldots$$

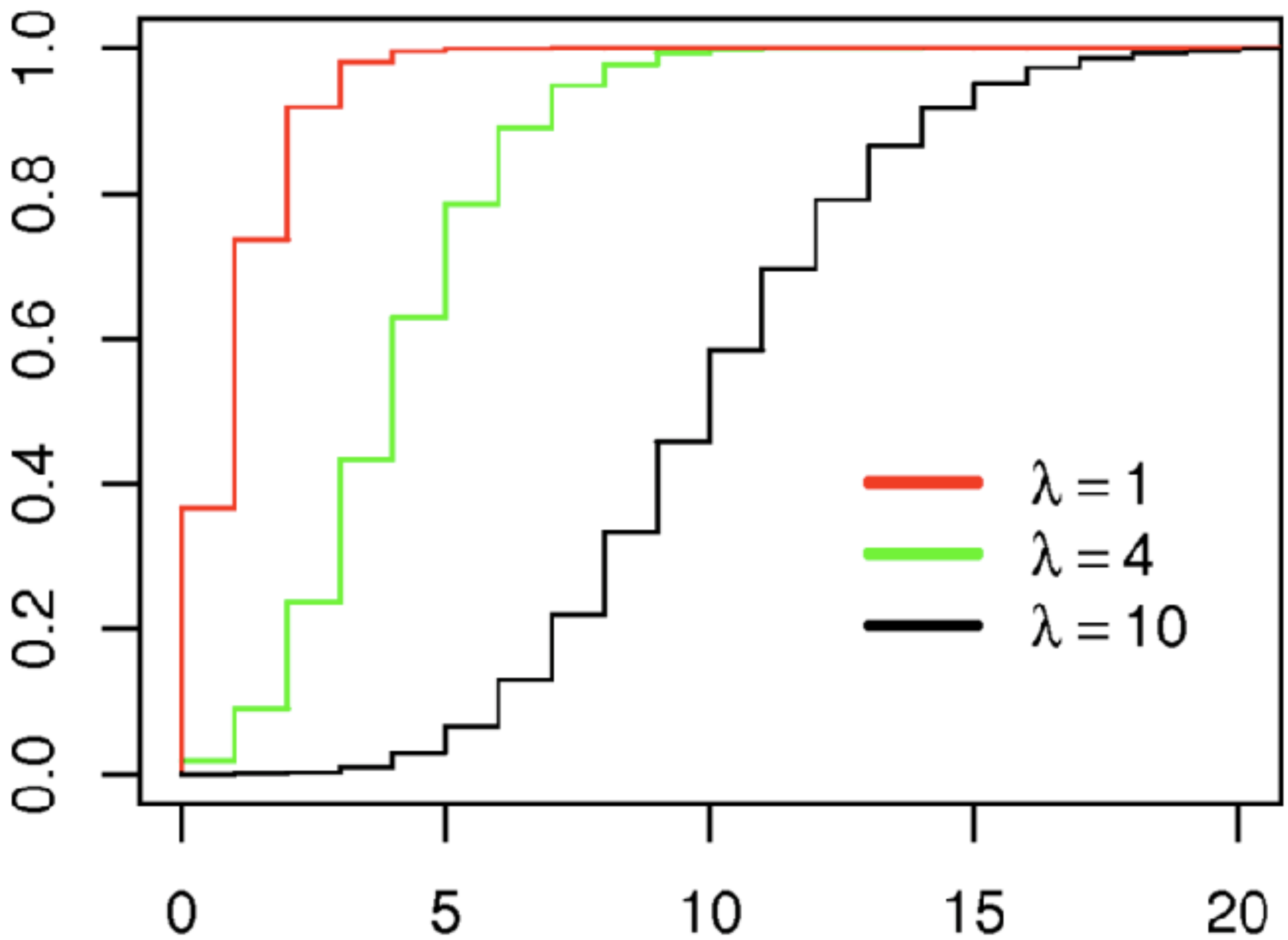
$$E(N(t)) = Var(N(t)) = \lambda t$$



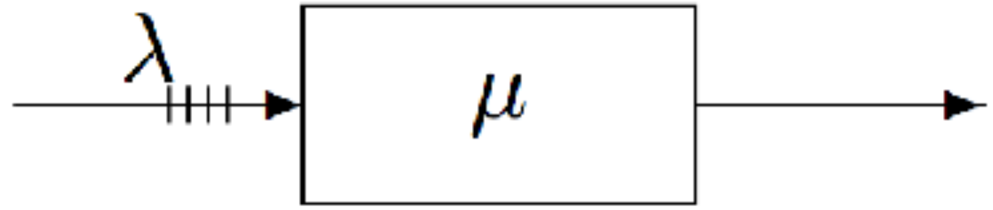
Poissonovo rozdělení

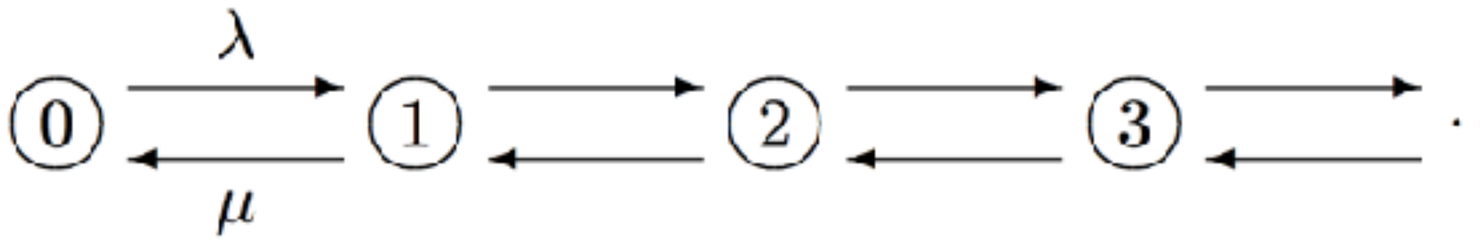
$$P(N(t) = k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \quad k = 0, 1, \ldots$$

$$E(N(t)) = Var(N(t)) = \lambda t$$



Systém hromadné obsluhy





P(v čase $(t,t+h\rangle$ přijde 1 zákazník) $= 1 - \mathrm{e}^{-\lambda h} = \lambda h + o(h), \quad h \to 0$

P(v čase $(t,t+h\rangle$ nepřijde 1 zákazník) $= \mathrm{e}^{-\lambda h} = 1 - \lambda h + o(h), \quad h \to 0$

P(v čase $(t,t+h\rangle$ odejde 1 zákazník) $= 1 - \mathrm{e}^{-\mu h} = \mu h + o(h), \quad h \to 0$

P(v čase $(t,t+h\rangle$ přijdou 2 zákazníci) $= (1 - \mathrm{e}^{-\lambda h})(1 - \mathrm{e}^{-\mu h}) = o(h), \quad h \to 0$

P(v čase $(t,t+h\rangle$ se počet zákazníků nezmění) $= P(X(t+h) = j \mid X(t) = j)$

= P(v čase $(t,t+h\rangle$ žádný zákazník nepřijde a žádný neodejde) $= \mathrm{e}^{-\lambda h}\,\mathrm{e}^{-\mu h}$

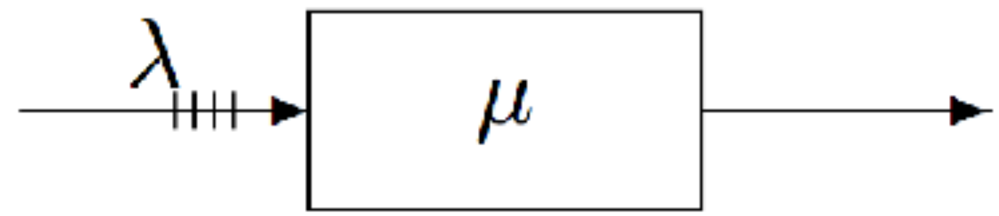
$$= 1 - \lambda h - \mu h + o(h), \quad h \to 0$$

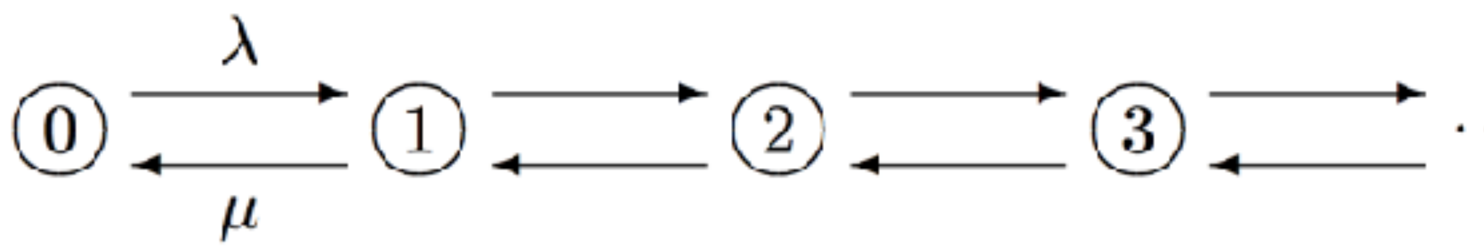
P(v čase $(t,t+h\rangle$ se počet zákazníků zvýší o 1) $= P(X(t+h) = j+1 \mid X(t) = j)$

= P(v čase $(t,t+h\rangle$ jeden zákazník přijde a žádný neodejde) $= (1 - \mathrm{e}^{-\lambda h})\mathrm{e}^{-\mu h}$

$$= \lambda h + o(h), \quad h \to 0$$

Systém hromadné obsluhy





$p_k(t)$ = P(v čase t je systém ve stavu k) = P(v čase t je v systému k zákazníků)

$$p_0(t+h) = p_0(t)(1-\lambda h) + p_1\mu h + o(h)$$

$$p_k(t+h) = p_k(t)(1-\lambda h - \mu h) + p_{k-1}(t)\lambda h + p_{k+1}\mu h + o(h), \quad k>0$$

$$p_0'(t) = -\lambda p_0(t) + \mu p_1(t)$$

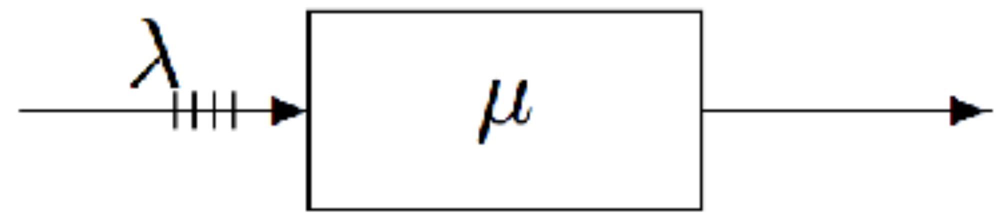
$$p_k'(t) = -(\lambda+\mu)p_k(t) + \lambda p_{k-1}(t) + \mu p_{k+1}(t)$$

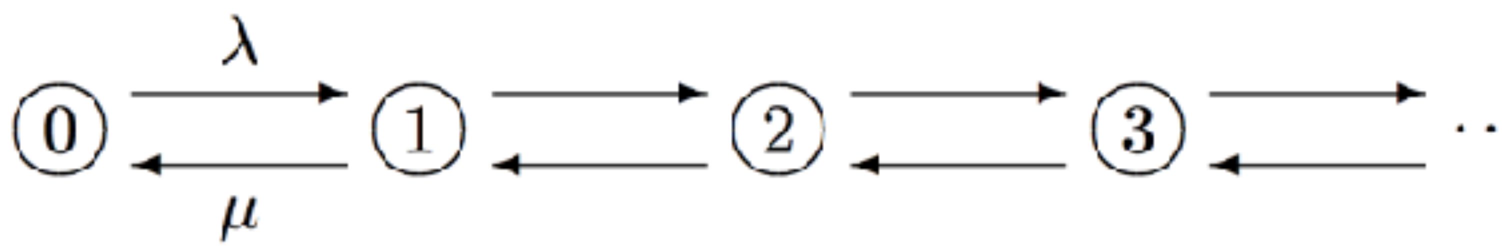
$$0 = -\lambda p_0 + \mu p_1$$

$$0 = -(\lambda+\mu)p_k + \lambda p_{k-1} + \mu p_{k+1}$$

$$\sum_{k=0}^{\infty} p_k = 1$$

Systém hromadné obsluhy





$$0 = -\lambda p_0 + \mu p_1$$
$$0 = -(\lambda + \mu) p_k + \lambda p_{k-1} + \mu p_{k+1}$$

$$\sum_{k=0}^{\infty} p_k = 1$$

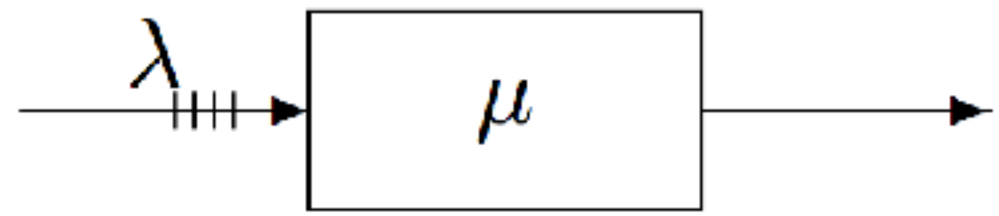
$$z_k = -\lambda p_k + \mu p_{k+1}$$

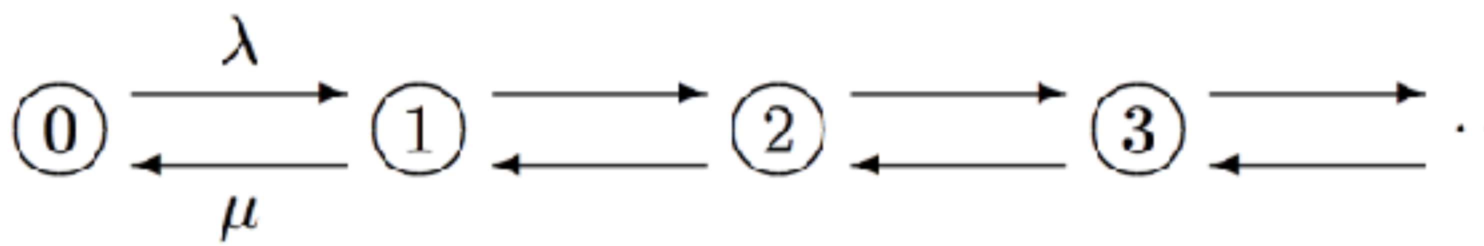
$$0 = z_0$$
$$0 = z_{k-1} + z_k, \quad k = 1, 2, \ldots$$

$$\lambda p_k = \mu p_{k+1} \;\Rightarrow\; \frac{p_{k+1}}{p_k} = \frac{\lambda}{\mu} \;\Rightarrow\; p_k = \left(\frac{\lambda}{\mu}\right)^k p_0$$

$$\rho = \frac{\lambda}{\mu} < 1 \;\Rightarrow\; \sum_{k=0}^{\infty} p_0 \left(\frac{\lambda}{\mu}\right)^k = \frac{p_0}{1-\rho} = 1 \;\Rightarrow\; p_0 = 1 - \rho$$

Systém hromadné obsluhy





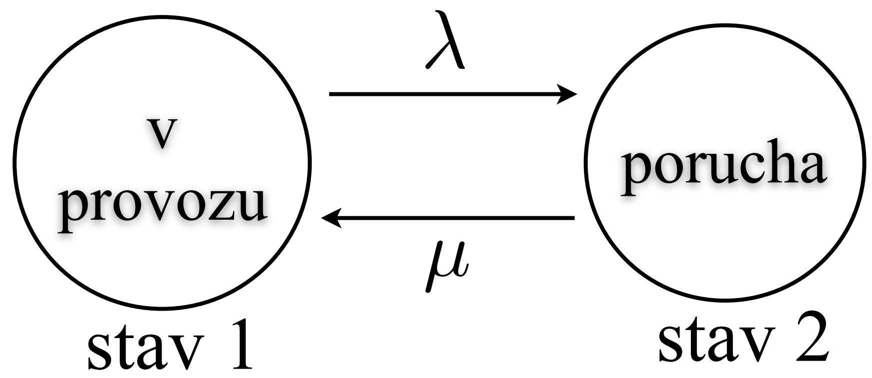
$$p_k = \left(1 - \frac{\lambda}{\mu}\right)\left(\frac{\lambda}{\mu}\right)^k, \quad k = 0, 1, 2, \ldots$$

$$\rho = \left(\frac{\lambda}{\mu}\right)$$ - intenzita obsluhy

Otázky: 1. Jaký je průměrný počet zákazníků v systému?

2. Jaký je průměrný počet zákazníků ve frontě?

Nespolehlivý systém



$$p_1(t+\triangle) = p_1(t)(1-\lambda\triangle) + p_2(t)\mu\triangle$$
$$p_2(t+\triangle) = p_1(t)\lambda\triangle + p_2(t)(1-\mu\triangle)$$

$$p_1(0) = 1, \quad p_2(0) = 0$$

$$\frac{dp_1(t)}{dt} = -\lambda p_1(t) + \mu p_2(t)$$
$$\frac{dp_2(t)}{dt} = \quad \lambda p_1(t) - \mu p_2(t)$$
$$p_1(0) = 1, \quad p_2(0) = 0$$

$$\boxed{\begin{aligned} p_1(t) &= \frac{\mu}{\lambda+\mu} + \frac{\lambda}{\lambda+\mu}e^{-(\lambda+\mu)t} \\ p_2(t) &= \frac{\lambda}{\lambda+\mu}\left(1-e^{-(\lambda+\mu)t}\right) \end{aligned}}$$

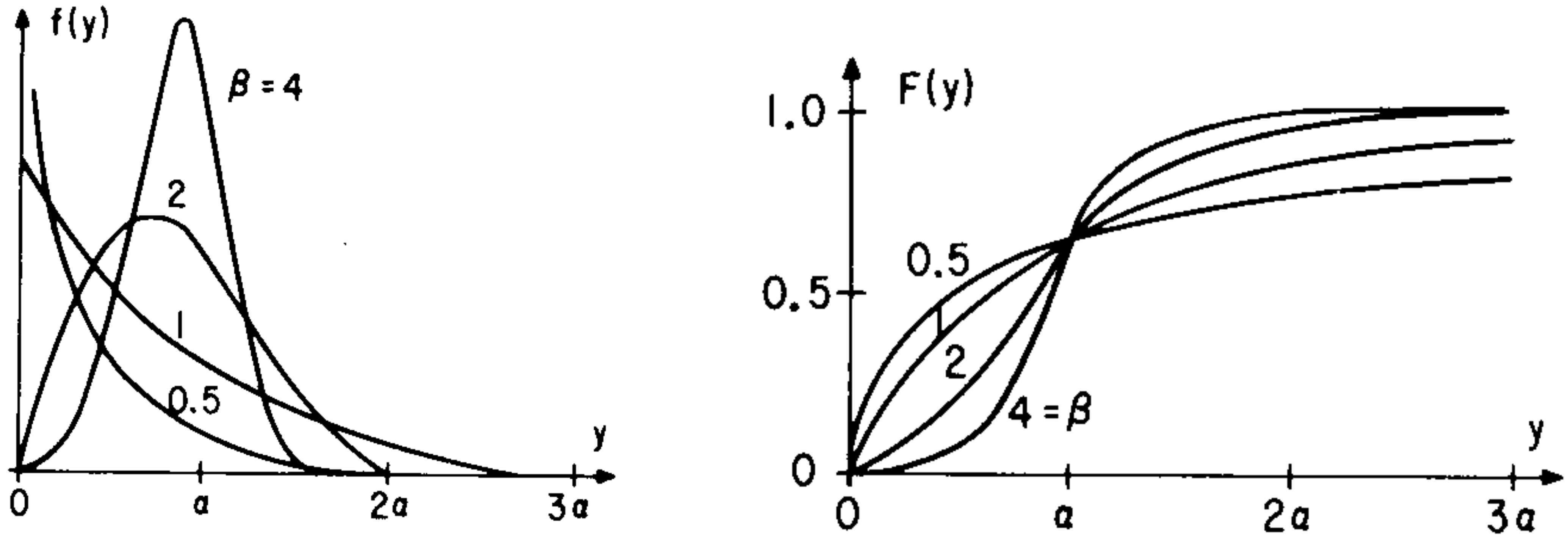
$$\lim_{t\to\infty} p_1(t) = \frac{\mu}{\lambda+\mu}, \qquad \lim_{t\to\infty} p_2(t) = \frac{\lambda}{\lambda+\mu}$$

limitní pohotovost   *limitní nepohotovost*

Weibullovo rozdělení

$$F(t) = 1 - e^{-\left(\frac{t}{\delta}\right)^{\beta}}, \quad t \geq 0$$

$$f(t) = \left(\frac{t}{\delta}\right)^{\beta-1} \frac{\beta}{\delta} e^{-\left(\frac{t}{\delta}\right)^{\beta}}, \quad t \geq 0$$

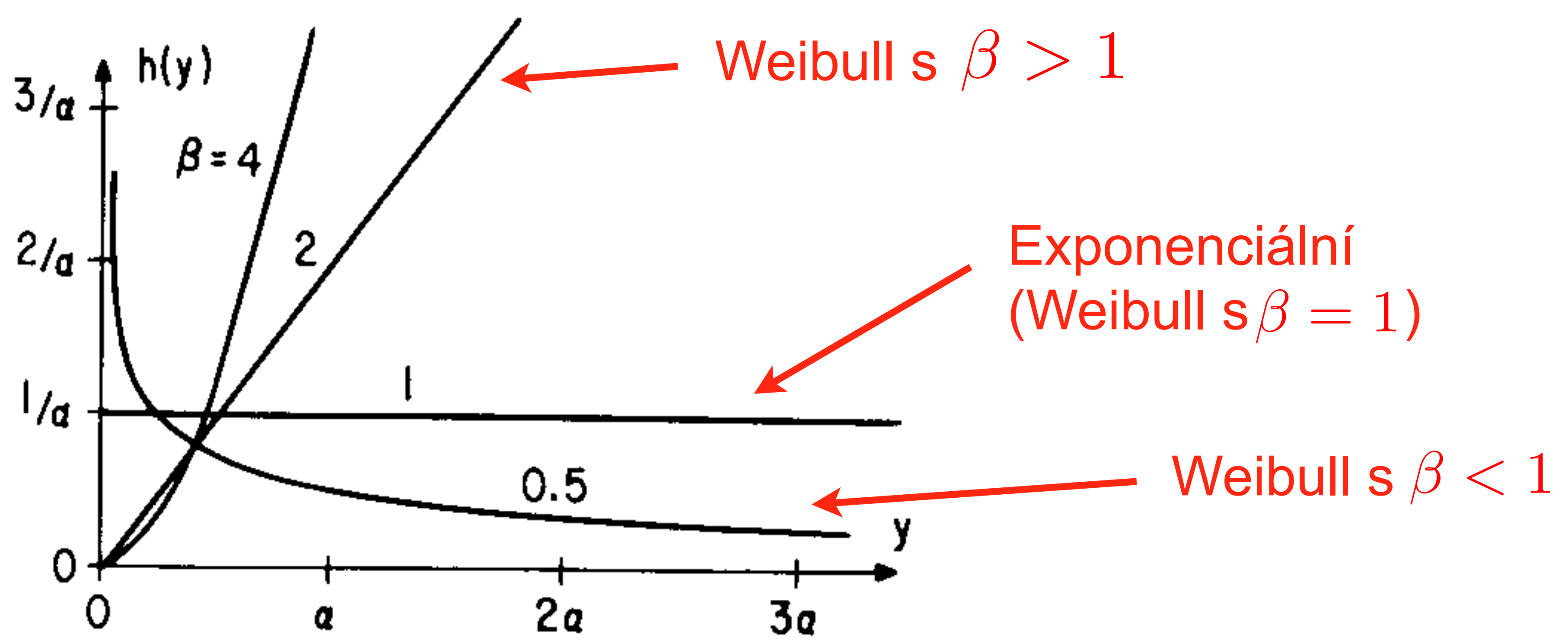


$$E(T) = \delta\Gamma(\frac{1}{\beta} + 1), \quad Var(T) = \delta^2\left\{\Gamma(\frac{2}{\beta} + 1) - \left[\Gamma(\frac{1}{\beta} + 1)\right]^2\right\}$$

Intenzita poruchy

$$\lambda(t) = \frac{f(t)}{1 - F(t)}, \quad t \in R$$

Pravděpodobnost, že se zařízení porouchá v čase $t + \triangle$, když víme, že se do doby t neporouchalo (pro hodně malá $\triangle$) je rovna přibližně $\triangle\, \lambda(t)$.

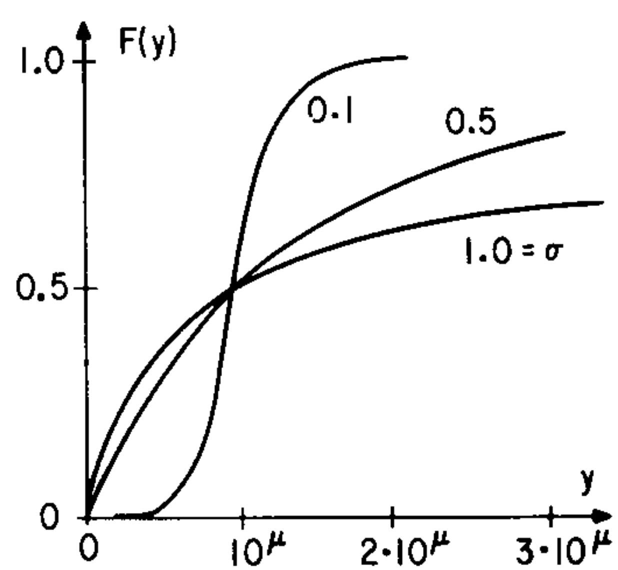


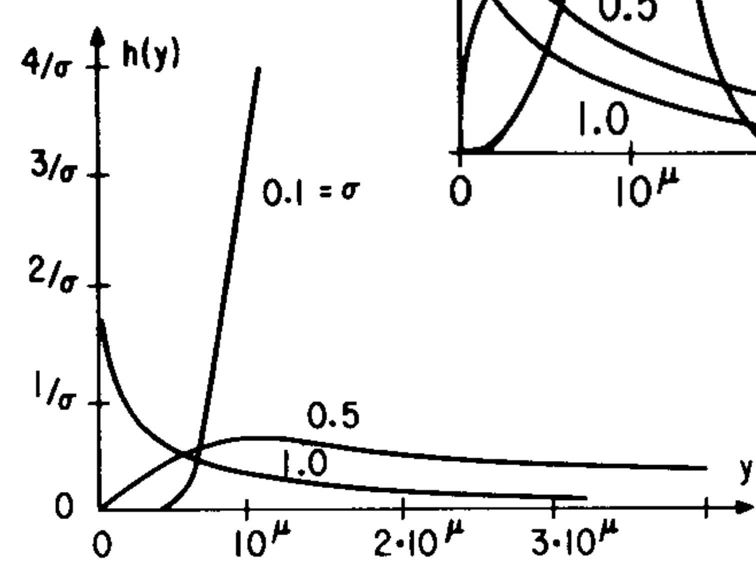
Logaritmicko-normální rozdělení

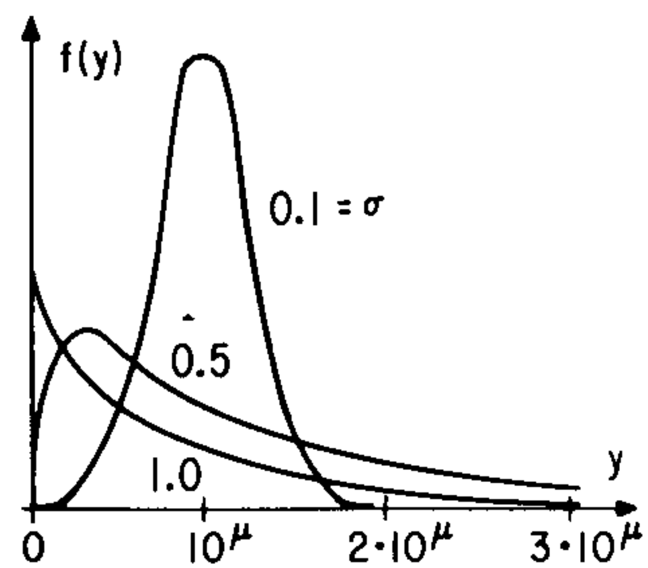
$$f(x) = \begin{cases} 0, & x < 0 \\ \dfrac{1}{\sigma x \sqrt{2\pi}} \exp\left(-\dfrac{(\ln x - \mu)^2}{2\sigma^2} \right), & x \geq 0 \end{cases}$$

$$Y \sim N/\mu, \sigma^2)$$
$$X = \exp(Y)$$
$$X \sim LN(\mu, \sigma^2)$$

$$E(X) = \exp\left(\mu + \frac{\sigma^2}{2} \right), \quad Var(X) = \exp\left(2\mu + \sigma^2\right).\left(\exp \sigma^2 - 1\right)$$







Charakteristiky doby života

Příklad: Životnost ventilátoru dieselového motoru můžeme popsat pomocí exponenciálního rozdělení se střední dobou θ = 28.700 hod.

1. S jakou pravděpodobností se ventilátor porouchá během prvních 100 hodin?

$$P(X \leq 100) = \int_0^{100} \frac{1}{28700} \exp(-\frac{x}{28700}) dx = 1 - e^{-\frac{100}{28700}} = 0{,}003478$$

2. Jaká je pravděpodobnost, že ventilátor vydrží bez poruchy záruční dobu 8.000 hodin?

$$P(X > 8000) = \int_{8000}^{\infty} \frac{1}{28700} \exp(-\frac{x}{28700}) dx = e^{-\frac{8000}{28700}} = 0{,}76$$

3. Jaká je intenzita poruchy?

Intenzita poruchy je potom λ=1/ θ=1/28.700=34,8 ppm.

$$P(X \leq 100) \cong \lambda .100 = 100.0{,}00003484 = 0{,}003484$$

Charakteristiky doby života

Příklad: Životnost ventilátoru dieselového motoru můžeme popsat pomocí exponenciálního rozdělení se střední dobou θ = 28.700 hod.

4. Jaká je „typická“ délka života ventilátoru?

$$\text{median} = \tilde{x}_{0,5} = -28.700 \ln(1-0.5) = 19.900$$

5. Jaká je střední doba života ventilátoru?

$$E(X) = \theta = 28.700$$

6. Do jaké doby se porouchá v průměru 90% všech ventilátorů?

$$\tilde{x}_{0,9} = -28.700 \ln(1-0.9) = 66.084$$