



**ČESKÁ SPOLEČNOST PRO JAKOST**  
Novotného lávka 5, 116 68 Praha 1

**NEPRÁVEM OPOMÍJENÉ  
METODY ANALÝZY  
SPOLEHLIVOSTI**



**Materiály z 31. setkání  
odborné skupiny pro spolehlivost**

Praha, květen 2008

# OBSAH

<b>NEPARAMETRICKÉ STATISTICKÉ METODY HODNOCENÍ SPOLEHLIVOSTI</b> <i>doc. RNDr. Miroslav Koucký, CSc.</i>	<b>3</b>
<b>VÝBĚR Z METOD RŮSTU SPOLEHLIVOSTI</b> <i>prof. Ing. Rudolf HOLUB, CSc.</i>	<b>14</b>
<b>SOFT METODY ANALÝZY SPOLEHLIVOSTI – FUZZY LOGIKA A JEJÍ APLIKACE V ANALÝZÁCH SPOLEHLIVOSTI</b> <i>Ing. David VALIŠ, Ph.D.</i>	<b>37</b>

# NEPARAMETRICKÉ STATISTICKÉ METODY HODNOCENÍ SPOLEHLIVOSTI

Doc. RNDr. Miroslav Koucký, CSc., Katedra aplikované matematiky, Technická univerzita  
v Liberci

[miroslav.koucky@tul.cz](mailto:miroslav.koucky@tul.cz)

## 1. Úvod

Hlavní typy dat, které je třeba v oblasti statistické analýzy spolehlivosti vyhodnocovat jsou - jednoduchá, cenzorovaná a oříznutá, resp. jejich kombinace.

- **Jednoduchá data**  
Jde o základní, nejjednodušší typ statistických dat, kde zjištěné údaje  $t_1, t_2, \dots, t_n$  tvoří náhodný výběr z rozdělení pravděpodobnosti doby do poruchy  $T$ . Získaná data tak reprezentují realizace nezávislých a stejně rozdělených náhodných veličin.
- **Cenzorovaná data**  
Data jsou obvykle ve tvaru  $(t_1, d_1), \dots, (t_n, d_n)$ , kde  $t_i = \min(T, C)$ ,  $T$  je náhodná veličina vyjadřující dobu do poruchy,  $C$  je doba cenzorování a  $d_i$  je indikátor cenzorování definovaný vztahem

$$d_i = \begin{cases} 1 & \text{, jestliže } t_i \text{ je doba do poruchy,} \\ 0 & \text{, v ostatních případech (cenzorování).} \end{cases}$$

Základní typy cenzorování - s pevným časem ( $C$  je pevně daný čas) a s náhodným časem ( $C$  je náhodná veličina s daným rozdělením pravděpodobnosti s hustotou  $f_C(c)$ ). Tento typ spolehlivostních dat je v praxi velmi rozšířený a vyskytuje se zpravidla v situacích, kdy doba sledování je ukončena po uplynutí určité doby.

- **Oříznutá data** – jde o variantu, kdy poruchy jsou registrovány až po uplynutí jisté doby. V praxi se lze s tímto typem dat setkat nejčastěji v situaci, kdy údaje o poruchách nejsou v počátečním období registrovány.

Klasifikace statistických metod používaných v oblasti statistické analýzy spolehlivosti:

- **Parametrické metody**  
Tyto metody vychází z předpokladu, že napozorovaná data tvoří náhodný výběr z určitého konkrétního typu rozdělení (např. exponenciální, Weibullovo, gama apod.). Hlavním úkolem pak bývá určit model pro pravděpodobnost bezporuchového provozu  $R_T(t; \theta)$ , resp. hustotu pravděpodobnosti  $f_T(t; \theta)$  náhodné veličiny  $T$ , kde  $\theta = (\theta_1, \theta_2, \dots, \theta_r)^T$  je vektor parametrů. Základní úlohy pak spadají do oblasti teorie odhadu (určení „správné“ hodnoty parametru  $\theta$ , resp. jeho funkce  $\tau = \tau(\theta)$ ; postup - klasický Neyman-Pearsonův, Fischerův maximálně věrohodnostní a Bayesův) a do oblastí testování hypotéz (Neyman-Pearson, Fischer).
- **Neparametrické metody**

Tyto metody nepředpokládají žádný konkrétní typ rozdělení dat a jsou proto „univerzálně“ použitelnou alternativou parametrických metod<sup>†</sup>. Lze je aplikovat i na náhodné výběry malých rozsahů pocházející z výrazně nenormálních rozdělení – jedna z jejich hlavních předností, spolu s jednoduchým použitím. Jejich nevýhodou bývá pochopitelně menší síla při porovnání s parametrickými metodami. Mezi základní typy úloh patří - odhady hodnot spolehlivostních charakteristik (pravděpodobnost poruchy, intenzita poruchy apod.) nebo testy různých typů hypotéz týkajících se např. hodnot spolehlivostních charakteristik, nezávislosti, shody dvou rozdělení apod.

- **Semiparametrické metody**

Tyto metody lze považovat za jistý kompromis mezi parametrickými a neparametrickými metodami, neboť vyžadují pouze „částečnou“ specifikaci rozdělení. Pro „důležité“ kovariáty je zaveden parametrický a pro „nepodstatné“ neparametrický model (viz např. [3]). Doposud však neexistuje jednoznačná odpověď na otázku, kdy a jak tyto metody efektivně použít.

### Přehled značení

$T$	... náhodná veličina vyjadřující dobu do poruchy,
$t_1, t_2, \dots, t_n$	... náhodný výběr z rozdělení veličiny $T$ , tj. naměřené doby do poruchy, resp. údaje o případném cenzorování,
$t_{(1)}, t_{(2)}, \dots, t_{(n)}$	... uspořádaný náhodný výběr, tj. hodnoty $t_1, t_2, \dots, t_n$ seřazené neklesajícím způsobem (včetně cenzorování),
$t_{[1]}, t_{[2]}, \dots, t_{[m]}$	... uspořádaný náhodný výběr dob do poruchy (bez údajů o cenzorování),
$\Lambda_T(t), \Lambda^*_T(t)$	... kumulativní intenzita poruchy, resp. její bodový odhad,
$R_T(t), R^*_T(t)$	... pravděpodobnost bezporuchového provozu, resp. její bodový odhad,
$E[\bullet]$	... střední hodnota veličiny $\bullet$ , resp. jeho bodový odhad,
$\text{var}[\bullet]$	... rozptyl veličiny $\bullet$ , resp. jeho bodový odhad.

## 2. Neparametrické metody - odhady základních ukazatelů bezporuchovosti

V následující části je uveden stručný výběr základních neparametrických metod z oblasti teorie odhadu a testování hypotéz, které lze využít při statistické analýze spolehlivostních dat. Jde především o Kaplan-Meierův odhad pravděpodobnosti bezporuchového provozu  $R_T(t)$ , Nelson-Altschulterův odhad okamžité kumulativní intenzity poruchy  $\Lambda_T(t)$  a Mantelův test shody dvou funkcí přežití doby  $t$ .

V následující části předpokládáme, že spolehlivostní data jsou ve formě náhodného výběru  $t_1, t_2, \dots, t_n$  vyjadřujícího doby do poruchy (resp. mezi poruchami), včetně případného cenzorování.

### a) Kaplan-Meierův odhad (K\_M)

---

<sup>†</sup> Neparametrické modely mohou být považovány za parametrické s nekonečně mnoha parametry.

K\_M bodový odhad  $R^*_T(t)$  pravděpodobnosti bezporuchového provozu  $R_T(t)$  je počítán ze vztahu

$$R^*_T(t) = \prod_{\substack{i \\ t_{[i]} \leq t}} \left(1 - \frac{m_{[i]}}{r_{[i]}}\right), \quad (1)$$

kde  $t_{[i]}$  je  $i$ -tá složka uspořádaného náhodného výběru dob do poruchy (nezahrnujeme časy cenzorování!),

$m_{[i]}$  je četnost hodnoty  $t_{[i]}$ ,

$r_{[i]}$  je velikost tzv. „risk“ množiny, tj. počet objektů, které byly v okamžiku  $t_{[i]}$  ještě sledovány.

Pro stanovení intervalu spolehlivosti K\_M odhadu lze využít Greenwoodovu formuli

$$\text{var}(\log(R^*_T(t))) = \sum_{\substack{i \\ t_{[i]} \leq t}} \frac{m_i}{r_i(r_i - m_i)},$$

umožňující odhadnout rozptyl veličiny  $\log(R^*_T(t))$ . Jako odhad rozptylu odhadu pravděpodobnosti bezporuchového provozu  $R^*_T(t)$  tak dostáváme

$$\text{var}[R^*_T(t)] = (R^*_T(t))^2 \text{var}(\log(R^*_T(t)))$$

a pro  $(1-\alpha)\%$  interval spolehlivosti veličiny  $R^*_T(t)$

$$\left(R^*_T(t) - u_{1-\alpha/2} \sqrt{\text{var}[R^*_T(t)]}, R^*_T(t) + u_{1-\alpha/2} \sqrt{\text{var}[R^*_T(t)]}\right), \quad (2)$$

kde  $u_\alpha$  označuje  $\alpha\%$  kvantil normovaného normálního rozdělení.

b) Nelson-Altschulerův odhad (N\_A)

N\_A bodový odhad kumulativní intenzity poruchy je počítán ze vztahu

$$\Lambda^*_T(t) = \sum_{\substack{i \\ t_{[i]} \leq t}} \frac{m_{[i]}}{r_{[i]}}, \quad (3)$$

kde význam jednotlivých symbolů odpovídá významu symbolů z Kaplan-Meierova odhadu (1).

K určení rozptylu a tedy i stanovení intervalu spolehlivosti pro  $\Lambda^*_T(t)$  lze využít asymptotický vztah

$$\text{var}[\Lambda^*_T(t)] = \sum_{\substack{i \\ t_{[i]} \leq t}} \frac{m_{[i]}}{r_{[i]}^2}.$$

Pro  $(1-\alpha)\%$  interval spolehlivosti veličiny  $\Lambda^*_T(t)$  tak dostáváme

$$\left(\Lambda^*_T(t) - u_{1-\alpha/2} \sqrt{\text{var}[\Lambda^*_T(t)]}, \Lambda^*_T(t) + u_{1-\alpha/2} \sqrt{\text{var}[\Lambda^*_T(t)]}\right). \quad (4)$$

Jako bodový odhad pravděpodobnosti bezporuchového provozu pak zřejmě dostáváme

$$R_T^*(t) = e^{-\Lambda_T^*(t)} \quad (5)$$

(analogicky lze určit i intervalový odhad).

c) Mantelův test

Tento neparametrický test umožňuje pro dva nezávislé náhodné výběry ověřit shodu jejich funkce bezporuchovosti. Testované hypotézy mají tvar:

$$H_0: R_1(t) = R_2(t) \quad \text{proti alternativě} \quad H_1: R_1(t) = R_2(t)^{\exp(\beta)}, \text{ pro } \beta \neq 0$$

a příslušná testová statistika

$$M = U^2 / I, \quad (6)$$

$$\text{kde } U = \sum_{i=1}^{n_0} \left( m_{[2i]} - \frac{m_{[i]} r_{[2i]}}{r_{[i]}} \right),$$

$$I = \sum_{i=1}^{n_0} \frac{m_{[i]} (r_{[i]} - m_{[i]}) r_{[1i]} r_{[2i]}}{n_{[i]}^2 (r_{[i]} - 1)},$$

$$m_{[i]} = m_{[1i]} + m_{[2i]}, \quad r_{[i]} = r_{[1i]} + r_{[2i]}.$$

Vzhledem k tomu, že při platnosti nulové hypotézy má testová statistika  $M$  rozdělení  $\chi^2(1)$ , zamítneme nulovou hypotézu na hladině  $\alpha$ , jestliže  $M \leq \chi^2_{\alpha/2}$  nebo  $M \geq \chi^2_{1-\alpha/2}$ .

Na závěr poznamenejme, že existuje celá řada dalších neparametrických metod vhodných pro analýzu spolehlivostních dat. Jde např. o neparametrické odhady hustoty obnovy, funkce obnovy apod. Podrobnosti lze nalézt v [1-6].

### 3. Aplikace neparametrických odhadů při analýze spolehlivostních dat

Následující část demonstruje použití výše uvedených neparametrických metod pro vyhodnocení spolehlivostních dat.

a) Kaplan-Meierův odhad

Použitá primární data jsou uvedena v příloze v odstavci Data\_1 a obsahují kumulativní doby do poruchy registrované u pěti „stejných“ obnovovaných systémů (systémy byly uváděny do provozu postupně v 500 hodinových intervalech). Pro potřeby analýzy je třeba určit doby mezi poruchami, resp. doby do cenzorování – viz následující tabulka č. 1 (červené údaje nejsou doby do poruchy, ale doba provozu od poslední poruchy do ukončení sledování).

System	$t_{(i)}$ – doba do poruchy, resp. <b>cenzorování</b>							
A	452	300	215	289	176	567	384	<b>617</b>
B	233	69	208	401	806	390	<b>393</b>	
C	783	1022	<b>195</b>					
D	783	<b>717</b>						
E	<b>1000</b>							

Tabulka č. 1 – spolehlivostní data

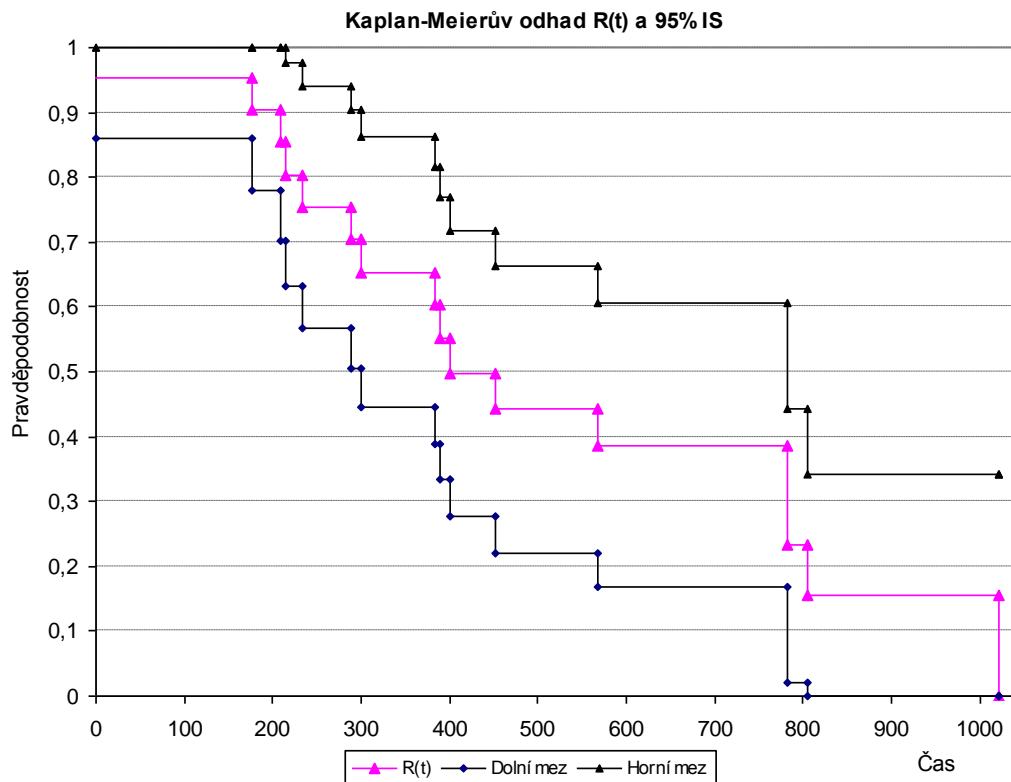
Následující tabulka č. 2 ilustruje postup výpočtu Kaplan-Meierova odhadu pravděpodobnosti bezporuchového provozu  $R^*_T(t_{[i]})$  a jeho 95% intervalu spolehlivosti. Význam jednotlivých údajů v tabulce č. 2 je následující:

- Událost ( $i$ ) – pořadové číslo události (poruchy nebo cenzorování),
- Čas  $t_{(i)}$  – doba do poruchy (černě) nebo cenzorování (červeně) seříděné vzestupně,
- Porucha  $[i]$  – pořadové číslo poruchy,
- Četnost  $m_{[i]}$  – počet událostí odpovídající času  $t_{[i]}$ ,
- Risk set  $r_{[i]}$  – velikost „risk“ množiny v čase  $t_{[i]}$ ,
- $R^*_T(t_{[i]})$  - bodový odhad pravděpodobnosti bezporuchového provozu pro časový interval  $(t_{[i-1]}, t_{[i]})$  vypočtený dle vztahu (1);
- $D_{95\%}(R_T)$ ,  $H_{95\%}(R_T)$  – dolní, resp. horní mez 95% intervalu spolehlivosti pro  $(t_{[i-1]}, t_{[i]})$  vypočtené dle (2).

Událost ( $i$ )	Čas $t_{(i)}$	Porucha $[i]$	Četnost $m_{[i]}$	Risk set $r_{[i]}$	Kaplan-Meier		
					$R_T(t_{[i]})$	$D_{95\%}(R_T)$	$H_{95\%}(R_T)$
1	69	1	1	21	0,95238	0,86130	1,04346
2	176	2	1	20	0,90476	0,77921	1,03031
3	<b>195</b>	-----	-----	-----	-----	-----	-----
4	208	3	1	18	0,85450	0,70210	1,00690
5	215	4	1	17	0,80423	0,63187	0,97659
6	233	5	1	16	0,75397	0,56633	0,94161
7	289	6	1	15	0,70370	0,50438	0,90303
8	300	7	1	14	0,65344	0,44543	0,86145
9	384	8	1	13	0,60317	0,38910	0,81725
10	390	9	1	12	0,55291	0,33518	0,77064
11	<b>393</b>	-----	-----	-----	-----	-----	-----
12	401	10	1	10	0,49762	0,27633	0,71890
13	452	11	1	9	0,44233	0,22068	0,66398
14	567	12	1	8	0,38704	0,16820	0,60587
15	<b>617</b>	-----	-----	-----	-----	-----	-----
16	<b>718</b>	-----	-----	-----	-----	-----	-----
17-18	783	13	2	5	0,23222	0,02042	0,44403
19	806	14	1	3	0,15481	0,00000	0,34265
20	<b>1000</b>	-----	-----	-----	-----	-----	-----
21	1022	15	1	1	0,00000	0,00000	0,34265

Tabulka č. 2 – Výpočet Kaplan-Meierova odhadu

V následujícím grafu je zobrazen průběh Kaplan-Meierova bodového a intervalového odhadu  $R^*_{T(t_{[i]})}$  s typickým schodovitým tvarem.



Graf č. 1 – Kaplan-Meierův odhad pravděpodobnosti bezporuchového provozu

b) Nelson-Altschulerův odhad

Následující tabulka č. 3 ilustruje výpočet Nelson-Altschulerova odhadu  $\Lambda^*_{T(t)}$  na stejných datech jako v odstavci a). Popis sloupců Událost – Risk set odpovídá tabulce č. 2. Hodnoty ve sloupci  $\Lambda^*_{T(t_{[i]})}$  jsou počítány dle vztahu (3) a jsou bodovým odhadem v intervalu  $(t_{[i-1]}, t_{[i]})$ .

Událost (i)	Čas $t_{(i)}$	Porucha $[i]$	Četnost $m_{[i]}$	Risk set $r_{[i]}$	Nelson-Altschuler	
					$\Lambda^*_{T(t_{[i]})}$	$R^*_{T(t_{[i]})}$
1	69	1	1	21	0,04762	0,95350
2	176	2	1	20	0,09762	0,90699
3	195	-----	-----	-----	-----	-----
4	208	3	1	18	0,15317	0,85798
5	215	4	1	17	0,21200	0,80897
6	233	5	1	16	0,27450	0,75995
7	289	6	1	15	0,34116	0,71094
8	300	7	1	14	0,41259	0,66193
9	384	8	1	13	0,48952	0,61292
10	390	9	1	12	0,57285	0,56392
11	393	-----	-----	-----	-----	-----
12	401	10	1	10	0,67285	0,51025

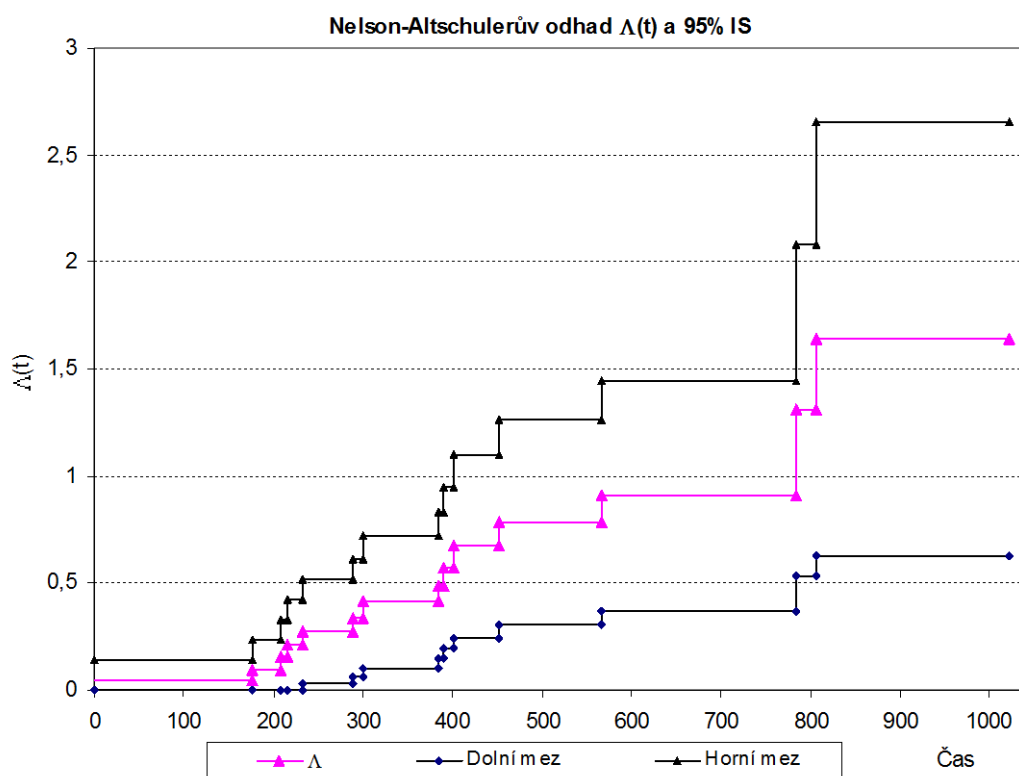


13	452	11	1	9	0,78396	0,45659
14	567	12	1	8	0,90896	0,40294
15	617	-----	-----	-----	-----	-----
16	718	-----	-----	-----	-----	-----
17-18	783	13	2	5	1,30896	0,27010
19	806	14	1	3	1,64229	0,19354
20	1000	-----	-----	-----	-----	-----
21	1022	15	1	1	2,64229	0,07120

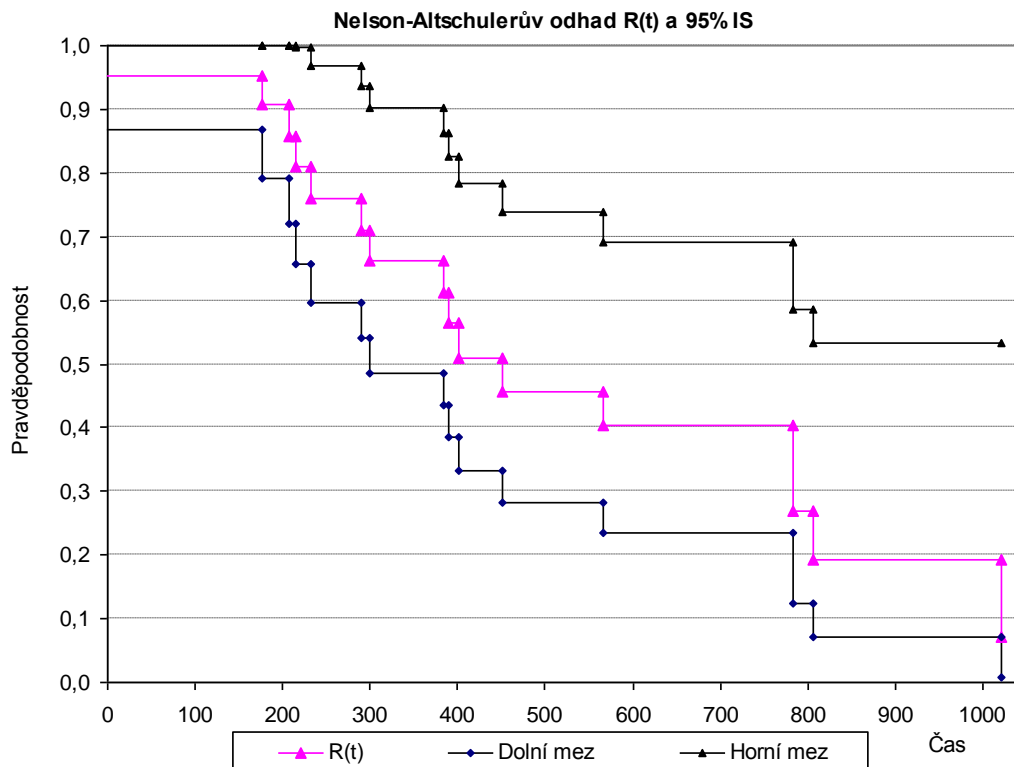
Tabulka č. 3 – Výpočet Nelson-Altschulerova odhadu

Nelson-Altschulerův odhad lze využít také k výpočtu odhadu pravděpodobnosti bezporuchového provozu. Tyto N-A odhady jsou obsaženy ve sloupci nadepsaném  $R^*_\tau(t_{[i]})$  a výpočet probíhá dle vztahu (5).

Následující grafy č. 2 a 3 znázorňují průběhy odhadů  $\Lambda^*_\tau(t)$ ,  $R^*_\tau(t)$  a jejich 95% intervalů spolehlivosti.



Graf č. 2 Nelson-Altschulerův odhad kumulativní intenzity poruchy



Graf č. 3 Nelson-Altschulerův odhad pravděpodobnosti bezporuchového provozu

Výše uvedené průběhy dokládají obecně známou skutečnost, že Nelson-Altschulerův odhad pravděpodobnosti bezporuchového provozu je systematicky optimističtější než odhad Kaplan-Meierův.

c) Mantelův test

Použitá primární data jsou uvedena v příloze v odstavci Data\_2 a obsahují údaje o počtech nabíjecích cyklů baterií téhož typu dvou různých výrobců. Aplikace Mantelova testu umožňuje rozhodnout, zda se rozdělení životnosti (vyjádřeno počtem nabíjecích cyklů) baterií obou výrobců statisticky významně liší.

Následující tabulka č. 4 demonstruje výpočet testové statistiky  $M$  dle vztahu (6).

Čas $t_{(i)}$	Risk_1 $r_{[1i]}$	Četnost_1 $m_{[1i]}$	Risk_2 $r_{[2i]}$	Četnost_2 $m_{[2i]}$	Risk $r_{[i]}$	Četnost $m_{[i]}$	$U_i$	$I_i$
76	15	0	20	1	35	1	0,4286	0,2449
82	15	0	19	1	34	1	0,4412	0,2465
164	15	2	18	0	33	2	-1,0909	0,4804
210	13	0	18	1	31	1	0,4194	0,2435
218	13	1	17	0	30	1	-0,5667	0,2456
230	12	1	17	0	29	1	-0,5862	0,2426
263	11	1	17	0	28	1	-0,6071	0,2385

315	10	0	17	1	27	1	0,3704	0,2332
385	10	0	16	1	26	1	0,3846	0,2367
412	10	0	15	1	25	1	0,4000	0,2400
467	10	1	14	0	24	1	-0,5833	0,2431
491	9	0	14	1	23	1	0,3913	0,2382
504	9	0	13	1	22	1	0,4091	0,2417
522	9	0	12	1	21	1	0,4286	0,2449
538	9	1	11	0	20	1	-0,5500	0,2475
639	8	1	11	0	19	1	-0,5789	0,2438
646	-----	-----	-----	-----	---	-----	-----	-----
669	7	1	10		17	1	-0,5882	0,2422
678	6	0	10	1	16	1	0,3750	0,2344
775	6	0	9	1	15	1	0,4000	0,2400
884	6	0	8	1	14	1	0,4286	0,2449
917	6	1	7	0	13	1	-0,5385	0,2485
1131	5	0	7	1	12	1	0,4167	0,2431
1148	5	1	6	0	11	1	-0,5454	0,2479
1446	4	0	6	1	10	1	0,4000	0,2400
1678	-----	-----	-----	-----	---	-----	-----	-----
1824	0	0	5	1	5	1	0	0
1827	0	0	4	1	4	1	0	0
2248	0	0	3	1	3	1	0	0
2385	0	0	2	1	2	1	0	0
3077	0	0	1	1	1	1	0	0

Součet: -0,5421 6,0520

Tabulka č. 4 – Výpočet testové statistiky Mantelova testu

Vzhledem k hodnotě testové statistiky

$$M = U^2 / I = \frac{(-0,5421)^2}{6,0520} = 0,0486$$

je zřejmé, že na hladině  $\alpha = 0,05$  nulovou hypotézu nezamítneme.

#### 4. Závěr

Kromě výše uvedeného Mantelova testu existuje celá řada dalších neparametrických testů. Následující tabulka č. 5 obsahuje základní přehled obecně používaných parametrických testů a jejich univerzálněji<sup>†</sup> použitelných neparametrických „ekvivalentů“.

<sup>†</sup> Univerzálnost neparametrických metod však neznamená, že nejsou založeny na žádných předpokladech týkajících se rozdělení náhodného výběru. Např. jednovýběrový Wilcoxonův test je relativně citlivý na porušení symetrie, Kruskal-Wallisův test na odlehlá pozorování apod.

Parametrický test	Neparametrický test
Jednovýběrový t-test	Znaménkový test
	Jednovýběrový Wilcoxonův test
Dvouvýběrový t-test	Mann-Whitneyův test
Párový t-test	Wilcoxonův párový test
ANOVA, jednoduché třídění	Kruskal-Wallisův test
	Moodův mediánový test
ANOVA, dvojné třídění	Friedmanův test

Tabulka č. 5 – Přehled ne/parametrických testů

### Základní kritéria pro výběr:

- |   |  |
|---|--|
| a) parametrické metody  | b) neparametrické metody   |
| - splněny předpoklady týkající se rozdělení náhodného výběru (např. normalita),                       | - náhodný výběr malého rozsahu (neumožňuje aplikovat centrální limitní větu na výběrový průměr), |
| - dostatečně velký rozsah výběru, který dovoluje aplikovat centrální limitní větu na výběrový průměr. | - výběr z výrazně nenormálního rozdělení,  |
|   | - nominální nebo ordinální data.   |

Príspevek vzniknul v rámci výzkumného centra „Centrum pro jakost a spolehlivost výroby“ financovaného MŠMT ČR, č. 1M06047.

### Literatura

- [1] Ansell J.I, Phillips M.J.: Practical Methods for Reliability Data Analysis. Clarendon Press, Oxford, 1994.
- [2] Mantel N.: Evaluation of survival data and two new rank order statistics arising from its consideration. Cancer Chemother. Rep., 50, 163-170.
- [3] J.I. Ansell, L.A. Walls: Dependency analysis in reliability studies. IMA J. Math. Appl. Bus. Ind., 1992 (3), 333-348.
- [4] D.R. Cox: Regression models and life tables (with Discussion). J. R. Stat. Soc. B, 1972(34), 187-220.
- [5] M.J. Crowder, A.C. Kimber: Statistical Analysis of Reliability Data. Chapman and Hall, London, 1991.
- [6] W. Nelson: Hazard plotting of left truncated life data. J. Qual. Tech., 1990(22), 230-238.
- [7] D.W. Newton: Some pitfalls in reliability data analysis. Reliab. Eng., 1991(34), 7-21.

## Příloha

### Data\_1

Následující tabulka obsahuje časové údaje o vzniku poruch u pěti obnovovaných identických systémů.

<b>Systém</b>	<b>Čas poruchy</b>							<b>Celkem</b>
<b>A</b>	452	752	967	1 256	1 432	1 999	2 383	<b>3 000</b>
<b>B</b>	233	302	510	911	1 717	2 107		<b>2 500</b>
<b>C</b>	783	1 805						<b>2 000</b>
<b>D</b>	783							<b>1 500</b>
<b>E</b>								<b>1 000</b>

Čas poruchy ... doba vzniku poruchy (v hodinách),

Celkem ... celková doba sledování systému (pravostranné cenzorování časem).

### Data\_2

Následující tabulka obsahuje data o době do poruchy (počet nabíjecích cyklů) baterií téhož typu dvou různých výrobců.

<b>Série 1</b>	164	164	218	230	263	467	538	639	669	917
	1 148	1 678+	1 678+	1 678+	1 678+					
<b>Série 2</b>	76	82	210	315	385	412	491	504	522	646+
	678	775	884	1 131	1 446	1 824	1 827	2 248	2 385	3 077

Doby označené + jsou cenzorovány zprava časem.

# VÝBĚR Z METOD RŮSTU SPOLEHLIVOSTI - MOŽNOSTI OVĚŘOVÁNÍ SPOLEHLIVOSTI PROTOTYPŮ VÝROBKŮ V ETAPĚ JEJICH VÝVOJE.

*Prof. Ing. Rudolf HOLUB, CSc.<sup>1</sup>*

## 1 Úvod

Obecným cílem všech zkoušek spolehlivosti výrobků je experimentální určení nebo ověření ukazatelů spolehlivosti, definovaných v normách nebo v technických specifikacích. Z tohoto důvodu se rozlišují dva typy zkoušek, určující a ověřovací. Z hlediska praktické proveditelnosti zkoušky se předpokládá, že soubor ukazatelů spolehlivosti je pro výrobek zvolen podle zásady nezávislosti, praktičnosti a prokazatelnosti.

Zvláštní problémy jsou spojeny se zkouškami prototypů nových nebo modernizovaných výrobků. Je to dáno řadou skutečností, specifických právě pro prototypy. Jsou k dispozici obvykle ve velmi omezeném počtu kusů, v průběhu vývojových prací a tedy i zkoušek se na nich provádí řada technických nebo technologických změn, které mění i jejich spolehlivostní vlastnosti. Tyto a další skutečnosti vedou často mnoho odborníků k názoru, že u prototypů nelze věrohodně ověřovat vývoj jejich růstu spolehlivosti (Reliability Growth). Proto se často doporučuje provádět pouze věcnou analýzu vzniklých poruch, provést nápravná opatření a na základě jejich realizace usoudit, zda prototyp splnil požadavky na spolehlivost. Provedení vlastní zkoušky růstu spolehlivosti se potom předpokládá až v provozních podmínkách u výrobků z ověřovací série případně sériových výrobků.

Předložený příspěvek se snaží poukázat na objektivně existující možnosti realizace zkoušky růstu spolehlivosti již v etapách vývoje prototypů a naznačuje praktické postupy a metody, které je možné k danému účelu použít. Autor se opírá ve svých názorech o vlastní zkušenosti, o existující národní i mezinárodní standardy, např. [3], [5], [6], [7], [8], [9] v nichž jsou zásady těchto zkoušek popsány pro systémy technických oborů.

## 2 Vybrané pojmy

### ***Bezporuchovost (reliability)***

*Schopnost objektu plnit požadovanou funkci v daných podmínkách a v daném časovém intervalu.*

### ***Růst bezporuchovosti (reliability growth)***

Okolnost, charakterizovaná postupným zlepšováním ukazatelů bezporuchovosti objektu s časem (dobou provozu).

### ***Ukazatel bezporuchovosti (measure)***

*Funkce nebo číselná hodnota (v pravděpodobnostním pojetí) používaná pro popis náhodné proměnné nebo náhodného procesu.*

---

<sup>1</sup> Prof. Ing. Rudolf HOLUB, CSc.; Fakulta vojenských technologií; Univerzita obrany v Brně, - emeritní profesor; Kounicova 65; 612 00 Brno; tel.: +420543245765; E-mail: [rudolf.holub@seznam.cz](mailto:rudolf.holub@seznam.cz).

## **Porucha (failure)**

*Jev spočívající v ukončení schopnosti objektu plnit požadovanou funkci. Kritéria poruchy musí být vymezena v technických podmínkách.*

### **Kategorizace poruch podle závažnosti důsledků**

1. **Závažná porucha (relevant failure):**

Je porucha, která musí být zahrnuta do hodnocení spolehlivosti podle kritérií k tomu účelu zvláště vypracovaných;

2. **Nezávažná porucha (non-relevant failure):**

Je porucha, která podle dohodnutých kritérií může být vyloučena z hodnocení spolehlivosti.

### **Kategorizace poruch podle vlivu na růst úrovně spolehlivosti**

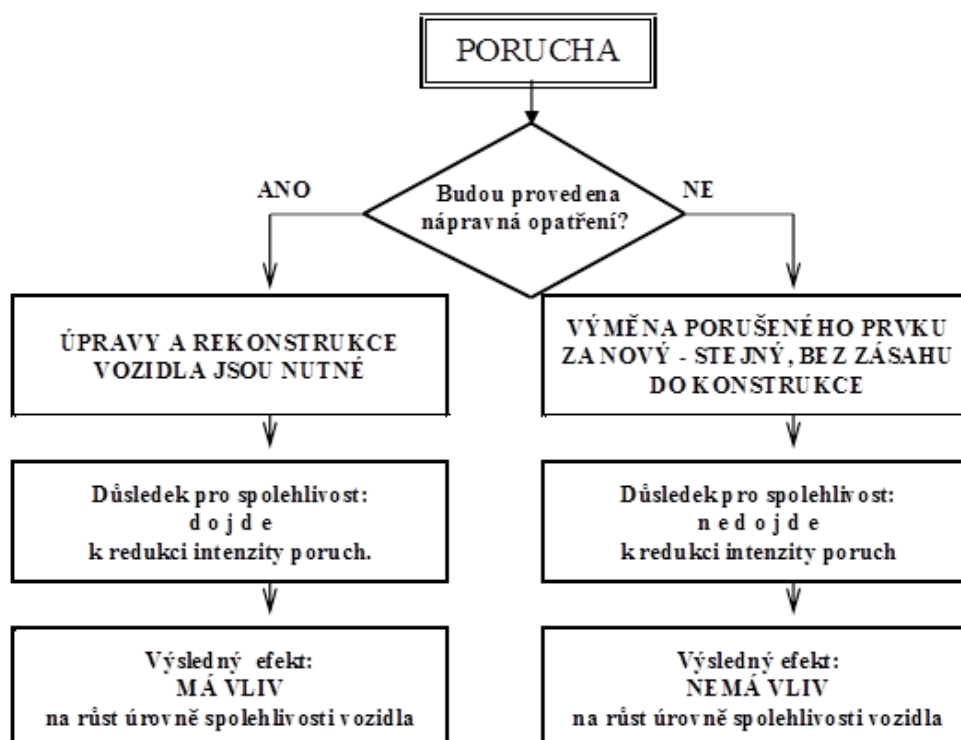
1. **Systematická porucha - defekt (systematic defect):**

Defekt známé (pozané) povahy, který může být vhodným zásahem do konstrukce, technologie nebo výrobního procesu odstraněn nebo jeho nepříznivý účinek zmírněn. Tento typ poruchy má vliv na růst úrovně spolehlivosti.

2. **Reziduální porucha- defekt (residual defect):**

Defekt, který není systematické povahy. Tento defekt je z konstrukce neodstranitelný nebo neodstraňovaný. Tento typ poruchy nemá vliv na růst úrovně spolehlivosti.

3. **Závislá (druhotná) porucha.** Tento typ poruchy se do hodnocení nezapočítává.



**Obr. 2.1 Schéma třídění poruch při sledování růstu bezporuchovosti**

### 3 Výběr metod růstu spolehlivosti - obecná charakteristika

Ze všech druhů zkoušek je možné pro prototypy doporučit tři možné postupy hodnocení spolehlivosti. Jsou to postupy, vhodné pro složité obnovované objekty na nichž se provádí průběžně konstrukční nebo technologické změny nebo je monitorován vývoj spolehlivosti v době jejich záběhu. Postupy jsou demonstrovány na případech výrobků střední složitosti.

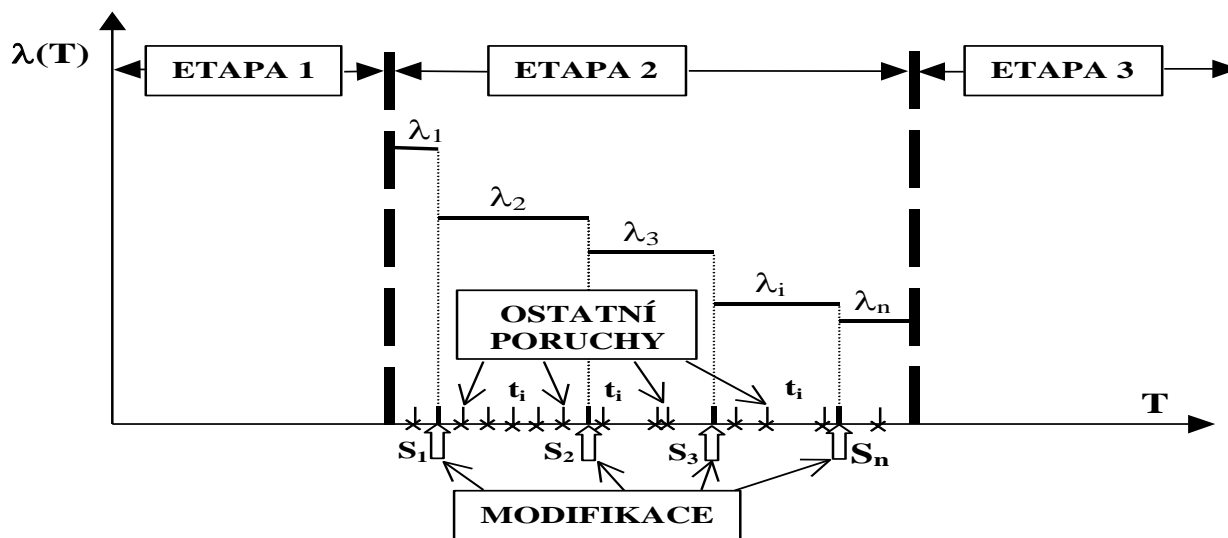
První dva postupy jsou známy jako Modely DUAN a AMSAA. Jsou založeny na empirické zkušenosti o možnosti modelovat závislost mezi dobou zkoušky  $T$  a vybranými parametry bezporuchovosti (MTBF;  $\lambda$ ).

V případě DUAN se předpokládá, že proces zlepšování úrovně spolehlivosti se řídí Homogenním Poissonovým Procesem s konstantní změnou intenzity poruch Exponenciálního typu. Podle toho se tato metoda někdy zkráceně označuje symbolicky jako metoda H.P.P.

V případě AMSAA se předpokládá, že proces zlepšování úrovně spolehlivosti se řídí Nehomogenním Poissonovým Procesem s proměnlivou změnou intenzity poruch Weibullovského typu. Podle toho se tato metoda někdy zkráceně označuje symbolicky jako metoda N.H.P.P.

Oba modely vývoje ukazatelů jsou založeny na empirické zkušenosti, že existuje možnost „modelovat“ (matematicky popsat) závislosti mezi vybranými parametry bezporuchovosti a dobou trvání zkoušky  $T$ . Tato skutečnost se využívá při praktickém postupu odhadu parametrů. V metodě AMSAA se předpokládá, že proces zlepšování úrovně parametrů spolehlivosti se řídí nehomogenním Poissonovým procesem s intenzitou poruch Weibullovského typu s parametry  $\alpha$ ,  $\eta$ . Podle charakteru procesu se tato metoda symbolicky označuje jako metoda N.H.P.P. (Nehomogenní Poissonův Proces).

Zkušební programy se v etapě vývoje sestávají ze zkušebních fází (etap), časově stejné nebo i nestejně délky, v nichž ale zkoušený objekt existuje v jedné konstrukčně-technologické variantě.



Obr.3.1 Obvyklé vývojové etapy projektu



Modely monitorují „změny (růst) parametrů bezporuchovosti“ dané varianty objektu uvnitř jedné zkušební fáze (etapy). Změny jsou vyvolány tím, že v konkrétních časových okamžicích  $S_1$ ,  $S_2$ , atd. dochází ke změnám v konstrukci takové povahy, že to má následně vliv na změny hodnot parametrů bezporuchovosti. Očekává se, že změny vedou ke „zvyšování“ úrovně ukazatelů bezporuchovosti. Celkovou situaci demonstruje Obr .3.1 v němž jsou symbolicky naznačeny možné změny v hodnotách intenzity poruch.

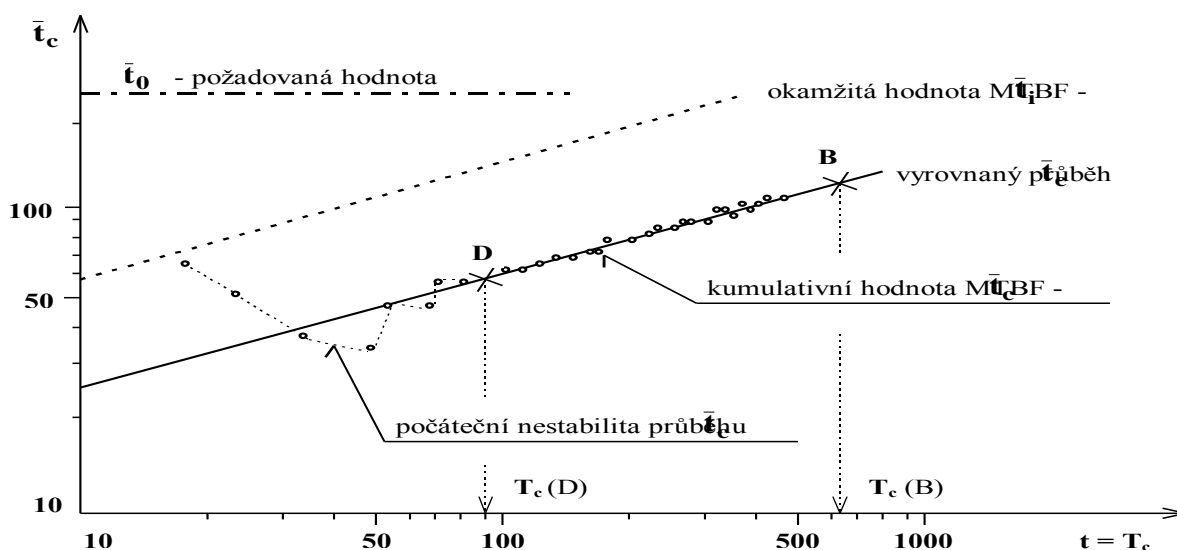
Cílovou informací získanou z obou modelů je „kumulativní“ a "okamžitá" hodnota ukazatelů spolehlivosti (např.  $\lambda_i$ , resp. MTBF) v daném okamžiku zkoušky. Tyto hodnoty potom lze porovnávat s požadavky v TP a podle okolností přijímat odpovídající závěry. Metody umožňují navíc prognózovat vývoj ukazatelů spolehlivosti do budoucího období.

Třetí postup je vhodný pro techniku s relativně již ukončeným vývojem. Je založen na možnosti stanovení rozsahu zkoušky a odhadu ukazatelů spolehlivosti na libovolně zvolené konfidenční úrovni  $C$ , tj. prakticky odhadovat zaručované nebo mezní hodnoty ukazatelů. K demonstraci je použita grafická interpretace účelově nazvané jako REGULAČNÍ DIAGRAM zkoušky. Postup je velmi názorný a umožňuje provádět hodnocení spolehlivosti především obnovovaných výrobků na libovolně zvolené konfidenční úrovni. Postup též umožňuje dopředu odhadnout očekávaný rozsah zkoušky a to v pojmech: minimální doba trvání zkoušky a maximální počet poruch ve zkoušce. Dále umožňuje prognózovat budoucí vývoj zkoušky a odhadnout rizika splnění/nesplnění požadavků na spolehlivost. Dá se použít pro celý systém i pro jeho jednotlivé soustavy. Podle konkrétního průběhu zkoušky lze přijímat organizační opatření pro případné úpravy zkoušky.

## **4 Duanův model vývoje ukazatelů spolehlivosti**

### **4.1 Popis metody**

Duanova metoda je grafickou technikou. Používá se při analýze vývojových tendencí ukazatelů spolehlivosti v etapě vývoje nového výrobku, kdy se na výrobku často provádí technické a technologické změny. Tyto změny mají významný vliv i na ukazatele spolehlivosti. K objektivizaci těchto změn slouží právě Duanova metoda. Graf na obr.4.1 znázorňuje empirickou zkušenost, že po počáteční nestabilitě v průběhu sledovaného ukazatele  $\bar{t}_c$  (kumulativní hodnota MTBF) se jeho další vývoj v průběhu pokračující zkoušky stabilizuje a vykazuje v log-log souřadnicích lineární závislost  $\bar{t}_c$  na době provozu  $T_c$ . Z této lineární závislosti se potom dá objektivně usoudit, zda dochází ve vývoji ukazatelů spolehlivosti k tendencím, zaručujícím splnění požadavků na spolehlivost či nikoliv. Grafické znázornění vývoje spolehlivosti umožňuje též prognózovat vývoj do budoucna, případně odhadnout potřebný rozsah zkoušky. Stejně tak umožňuje odhalit včas problémy se zabezpečením požadované úrovně spolehlivosti.



Obr. 4.1 Typický průběh linearizované závislosti  $\bar{t}_c$  a  $T_c$  v Duanově modelu

Symbols, použité v rovnicích Duanova modelu jsou:

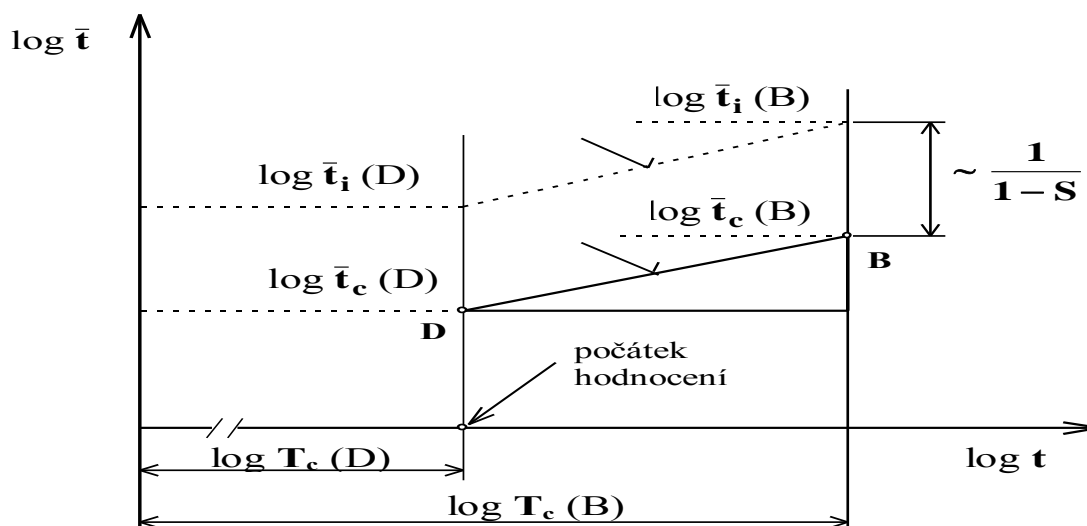
$t$  obecné označení pro dobu provozu (časovou osu),

$T_c$  celková (kumulativní) doba trvání zkoušky,

$\bar{t}_c$  kumulativní hodnota MTBF,

$\bar{t}_i$  okamžitá hodnota MTBF,

$U, S$  parametry Duanova modelu.



Obr.4.2 Linearizovaný model závislosti  $\bar{t}_c$  a  $T_c$  (Duanův model)

## 4.2 Definice, pojmy a výpočtové vztahy Duanova modelu

### Základní definice a pojmy v Duanově modelu

Duanův model řeší důležitý vztah mezi kumulativní a okamžitou hodnotou ukazatele spolehlivosti (nejčastěji MTBF -  $\bar{t}_c$ ,  $\bar{t}_i$ ).

#### Kumulativní hodnota ukazatele spolehlivosti

Je číselná hodnota ukazatele spolehlivosti (např.  $\bar{t}_c$ ), stanovená pro daný okamžik doby provozu  $t$  pomocí celkové (kumulativní) doby trvání zkoušky  $T_c$  uvažované od počátku zkoušky a celkového počtu významných poruch  $r$  k nimž v průběhu zkoušky došlo. Obsahuje v sobě celou předchozí "historii" vývoje spolehlivosti.

#### Okamžitá hodnota ukazatele spolehlivosti

Je číselná hodnota ukazatele spolehlivosti (např.  $\bar{t}_i$ ), stanovená pro daný okamžik doby provozu  $t$  (zkoušky) z nekonečně malého intervalu doby provozu  $\langle t; t + dt \rangle$ . Neobsahuje v sobě již žádnou informaci z předchozí "historie" vývoje spolehlivosti takže objektivně postihuje "dopad" všech změn, k nimž v průběhu vývoje u výrobku došlo. Je to nejdůležitější informace o právě dosažené úrovni ukazatelů spolehlivosti.

#### Výpočtové vztahy

Pro kumulativní hodnotu MTBF Duan definoval model závislosti ve tvaru:

$$\bar{t}_c = U \cdot T_c^S \quad (4.1)$$

resp. pro kumulativní intenzitu poruch ve tvaru:

$$\lambda_c = \frac{1}{\bar{t}_c} = \frac{1}{U} \cdot T_c^{-S} \quad (4.2)$$

V souladu s obr.4.1 a obr.4.2 lze vybudovat linearizovaný model a odvodit všechny další potřebné výpočtové vztahy.

Vztah mezi kumulativní a okamžitou hodnotou ukazatelů spolehlivosti se stanovuje pomocí následujících rovnic:

Pro okamžitou hodnotu MTBF, tj.  $\bar{t}_i$  je možné odvodit vztah:

$$\bar{t}_i = \frac{\bar{t}_c}{1-S} = \bar{t}_c \cdot \frac{1}{1-S} \quad (4.3)$$

Tím je dáno, že okamžitá hodnota ukazatele spolehlivosti  $\bar{t}_i$  (kterou nelze experimentálně určit) může být zobrazena v diagramu druhou přímkovou závislostí (čárkovane v obr. 4.1, 4.2 a 4.3), vzdálenou od ní o hodnotu  $1/(1-S)$ .

Z obr.4.2 pro parametry  $S$  a  $U$  pak lze odvodit vztahy:

$$S = \frac{\log \bar{t}_c(B) - \log \bar{t}_c(D)}{\log T_c(B) - \log T_c(D)} \quad (4.4)$$

$$U = \frac{\bar{t}_c(D)}{[T_c(D)]^S} \text{ nebo } U = \{\bar{t}_c\}_{T_c=1} \quad (4.5)$$

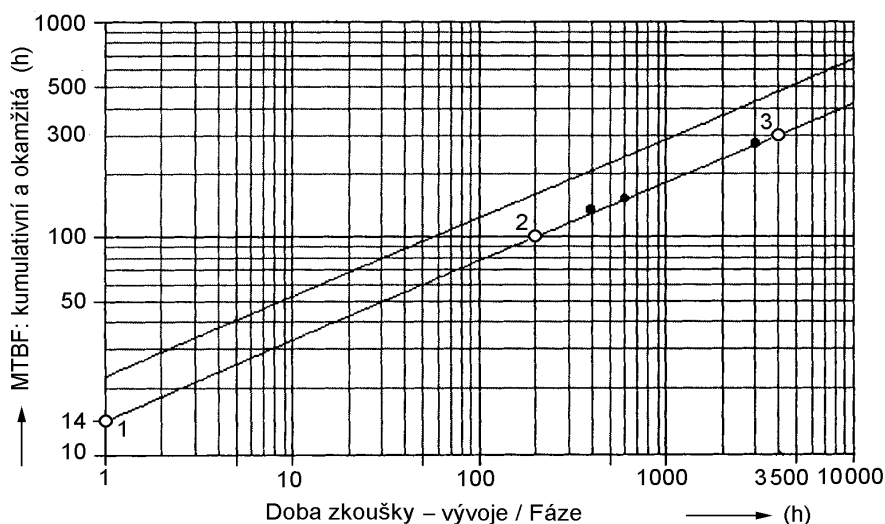
### 4.3 Odhad parametrů Duanova modelu

Odhady parametrů Duanova modelu S a U lze provádět metodou:

1. Grafickou;
2. Nejmenších čtverců;
3. Maximální věrohodnosti.

#### *Příklad jednoduché metody grafické*

Grafická metoda linearizace závislosti mezi ukazatelem bezporuchovosti a dobou provozu je vhodná při malém počtu poruch. Odhad parametru S se provede ze dvou bodů přímky graficky vyrovnané závislosti (viz např. body 2 a 3 v Obr. 4.3) podle vztahu (4.4). Odhad parametru U se může provést buď podle vztahu (4.5), nebo se jeho hodnota odečte z Obr.4.3 (bod 1) na ose MTBF pro T = 1.



**Obr.4.3 Duanův model - Kumulativní a okamžitá hodnota MTBF - příklad**

Zbylé dvě metody odhadu jsou metody analytické a jsou proto vhodné pro větší počet poruch (pro počet stupňů volnosti  $\nu > 30$ ). Podrobněji je o těchto postupech pojednáno v další části příspěvku, včetně možnosti intervalového odhadu ukazatelů spolehlivosti (vývoje ukazatelů).

## 5 Metoda AMSAA

### 5.1 Metoda AMSAA - obecná charakteristika metody

Metoda AMSAA představuje druhou progresivní a velmi používanou metodu monitorování vývoje ukazatelů spolehlivosti objektů v etapě jejich vývojových zkoušek nebo záběhu. Její název je odvozen od původce svého vzniku, kterým je: US Army Materiel Systems Analysis Activity.

Metoda dává objektivní informace o změnách v úrovni ukazatelů bezporuchovosti objektu v průběhu jeho vývojových zkoušek nebo záběhu. Změny ukazatelů v této etapě jsou přirozeným výsledkem konstrukčních a technologických změn, případně vlivu časných poruch a provedených oprav (úprav) u dané verze objektu, zkoušeného v přesně vymezené době / etapě zkoušek. Není určena k dodatečnému stanovení úrovně ukazatelů spolehlivosti u objektu až po dodatečně provedených konstrukčních, technologických či koncepčních změnách, prováděných jednorázově na konci vývojové etapy.

### 5.2 Teoretické nástroje metody AMSAA

Teoretické nástroje (výpočtový model a výpočtové vztahy) metody AMSAA jsou postaveny na těchto předpokladech:

Základní představa vychází z praktické zkušenosti o průběhu programu vývoje každého složitěho obnovovaného výrobku, probíhajícího v několika etapách jako v případě DUAN podle Obr. 3.1. Předpokládá se, že každá etapa nové varianty objektu začíná svůj technický život v čase  $T = 0$  a okamžiky  $0 < S_1 < S_2 < \dots < S_n$  jsou časové okamžiky provádění modifikačních zásahů (úprav, rekonstrukcí,....) na objektu uvnitř jedné zkušební fáze.

Podle typu poruch a charakteru jejich oprav (způsobu eliminace jejich příčin) se předpokládá, že každá jednotlivá vývojová etapa probíhá v důsledku prováděných konstrukčních a technologických změn v okamžicích  $S_1, S_2, \text{atd.}$ , Obr. 3.1, jako nehomogenní náhodný Poissonův proces s nekonztantní výslednou intenzitou poruch Weibullovskeho typu.

Naproti tomu mezi jednotlivými časovými okamžiky (periodami) změn v konstrukci  $[S_{i-1}; S_i; S_{i+1}; \text{atd.}]$  se předpokládá, že intenzity poruch  $\lambda_i$  vyhovují podmínce homogenního náhodného Poissonova procesu s konstantní výslednou intenzitou poruch Exponenciálního typu.

#### Základní výpočtové vztahy

V této části jsou (bez důkazu) uvedeny základní výpočtové vztahy pro nejčastěji požadované ukazatele bezporuchovosti objektu a jeho částí. Konkrétně pro předepsané ukazatele MTBF, zde označené jako  $m_C(T); m_i(T)$ , resp.  $\lambda_C(T); \lambda_i(T)$  a  $\Theta(t)$ , odhadované metodou AMSAA.

#### Výpočtové vztahy pro ukazatele bezporuchovosti

Metodika umožňuje vyhodnotit průběh těchto ukazatele bezporuchovosti:

1. Kumulativní a okamžitou hodnotu střední doby mezi poruchami;

$$m_C(T); m_i(T) \quad (\text{obecné označení - MTBF})$$

2. Kumulativní a okamžitou hodnotu intenzity poruch

$$\lambda_c(T); \lambda_i(T);$$

3. Očekávaný (střední) počet poruch za dobu T

$$\Theta(T)$$

### Formulace výpočtových vztahů

Rozhodující funkcí je funkce intenzity poruch je tvaru

$$h(T) = \frac{\beta}{\eta^\beta} \cdot T^{(\beta-1)}; \quad (5.1)$$

kde  $\beta, \eta$  jsou parametry Weibullova rozdělení.

Jestliže pro zjednodušení zápisu vztahů označíme dále symbolicky

$$\tilde{\lambda} = \frac{1}{\eta^\beta},$$

potom je funkce intenzity poruch  $h(T)$ , resp. okamžitá intenzita poruch  $\lambda_i(T)$  definována vztahem

$$\lambda_i(T) = \tilde{\lambda} \cdot \beta \cdot T^{(\beta-1)} \quad \text{kde: } T > 0; \quad \lambda > 0; \quad \beta > 0 \quad (5.2)$$

A očekávaný počet poruch za dobu T (resp. v intervalu  $[T(S_1), T(S_2)]$ , atd.) vyjadřuje vztah

$$\Theta(T) = \int_0^T \lambda_i(T) dT = \int_0^T \tilde{\lambda} \beta T^{(\beta-1)} dT = \tilde{\lambda} T^\beta \quad (5.3)$$

Kumulativní intenzita poruch  $\lambda_c$  za dobu T bude

$$\lambda_c = \tilde{\lambda} T^\beta \cdot T^{-1} = \tilde{\lambda} \cdot T^{(\beta-1)} \quad (5.4)$$

Snadno je možné dále odvodit vztahy pro MTBF:

Kumulativní hodnota  $MTBF_c$  bude

$$m_c(T) = MTBF_c = \frac{1}{\tilde{\lambda}} \cdot T^{(1-\beta)} \quad (5.5)$$

Okamžitá hodnota  $MTBF_i$  potom bude

$$m_i(T) = MTBF_i = \frac{1}{\tilde{\lambda} \cdot \beta} \cdot T^{(1-\beta)} \quad (5.6)$$

Pro Weibullův zákon rozdělení, parametr obecně označovaný jako  $\tilde{\lambda}$ , je parametrem měřítka protože závisí na počtu poruch v průběhu  $T$ . Parametr  $\beta$  je parametrem tvaru a charakterizuje tvar intenzity poruch v závislosti na době zkoušky  $T$ .

K odhadu uvedených parametrů mohou být v konkrétních případech použity následující postupy:

**Pro bodové odhady:**

1. Grafické;
2. Analytické (metodou maximální věrohodnosti);
3. Lineární regrese (metodou minimálních čtverců);

**Pro intervalové odhady potom:**

4. Metoda odhadu konfidenčních mezí, využívající vlastností chí-kvadrát rozdělení;
5. Metoda maximálně věrohodných odhadů konfidenčních mezí, využívající vlastností Fišerovy kovariační matice (metoda viz. [2], ČSN 010611).

Výběr konkrétní metody pro odhad parametrů musí být v každém jednotlivém případě proveden v souladu s charakterem souboru zpracovávaných údajů a jeho rozsahem.

### 5.3 Bodové odhady parametrů v metodě AMSAA

#### A. Grafická metoda

Grafická interpretace výsledků zkoušky je jedním z možných způsobů zpracování naměřených údajů. Metoda může být použita jako nástroj pro odhad parametrů především v procesu předběžného monitorování „růstu“ nebo „poklesu“ ukazatelů bezporuchovosti. Základní informací je histogram četnosti poruch v jednotlivých zvolených časových intervalech zkoušky (okamžité intenzity poruch -  $\lambda_i$ ).

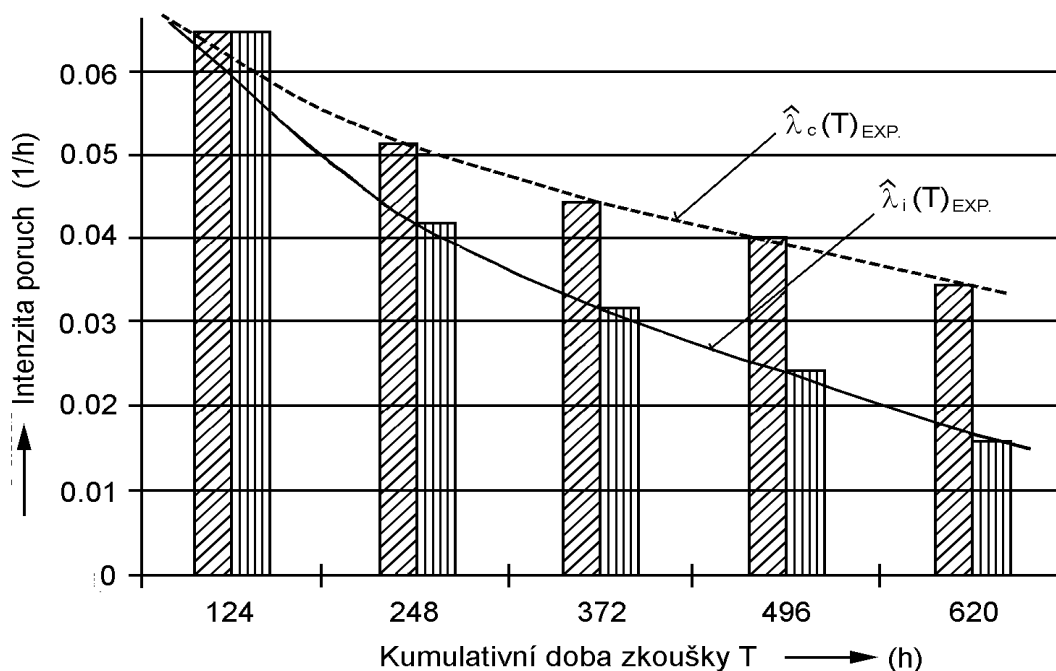
Celková doba trvání zkoušky je rozdělena do několika časových intervalů. Evidovaný počet poruch v každém intervalu je podělen „délkou“ (dobou) intervalu  $\Delta T_i$ , tím se dostane odhad okamžité intenzity poruch v každém intervalu, který se „blíží“ (tím více, čím je šířka intervalu menší) odhadu skutečné okamžité intenzity poruch  $\hat{\lambda}_i(T)$ .

Takže platí vztah 
$$\hat{\lambda}_i = \frac{\Delta n_i}{\Delta T_i} \quad (5.7)$$

Kumulativní intenzity poruch -  $\lambda_c$  se vypočte podle vztahu

$$\hat{\lambda}_c(T) = \frac{\sum n_i(T)}{\sum T_i}, \quad (5.8)$$

Grafická interpretace je patrná z příkladu na Obr. 5.1.



**Obr.5.1 Histogram četnosti poruch a experimentální odhad parametrů**

#### **Graficko analytický odhad parametrů**

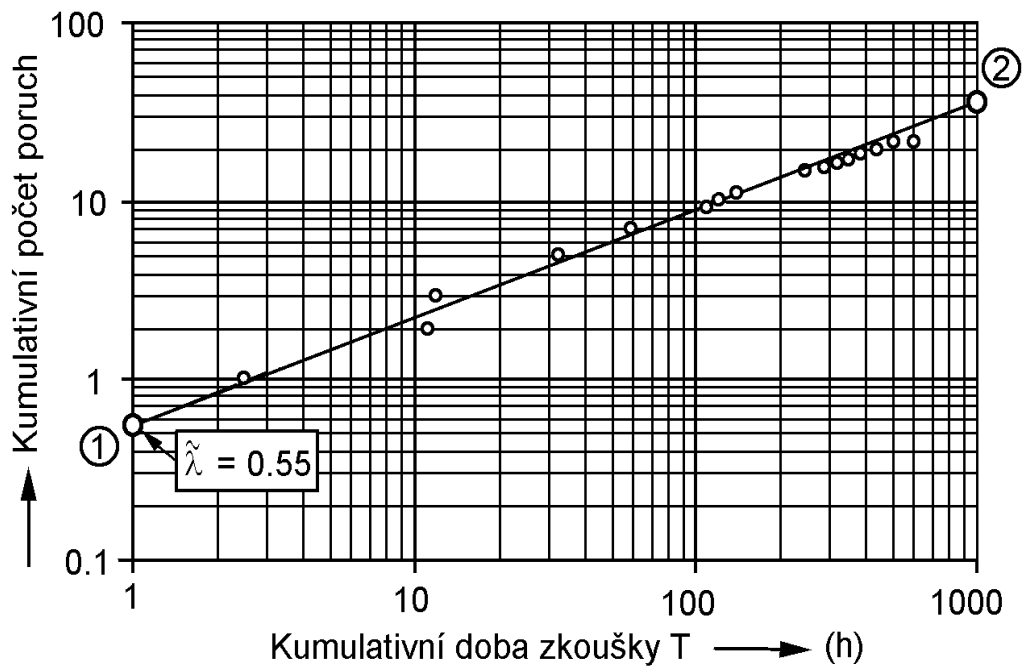
Tento postup je kombinací analytické metody, (analyticky známého modelu závislosti zákona rozdělení intenzity poruch na době trvání zkoušky) a grafické interpretace této závislosti, pomocí které je možné snadno odhadnout hledané parametry tohoto zákona.

Možný způsob odhadu je vybudován na platnosti rovnice (5.3) resp. (5.9), udávající závislost očekávaného (středního) počtu poruch za dobu trvání zkoušky  $T$ . Jejím logaritmováním dostaneme následující rovnici

$$\ln \Theta(T) = \ln \tilde{\lambda} + \beta \cdot \ln T \quad (5.9)$$

Je to rovnice přímky v  $\log(T)$  a  $\log(\Theta(T))$  souřadnicích. Pokud do této souřadné soustavy vyneseme naměřená data ze zkoušky, můžeme je lineárně aproximovat přímkou, která potom umožňuje provést odhad parametrů  $\tilde{\lambda}$  a  $\beta$  (viz Obr.5.2).





Obr. 5.2 AMSAA model – Kumulativní počet poruch vs. doba zkoušky

Směrnici přímkové závislosti snadno vypočteme ze dvou libovolně zvolených bodů ( $T_2 > T_1$ ) v grafu. Směrnice udává hodnotu druhého parametru  $\beta$  a vypočte se ze vztahu

$$\beta = \frac{\ln[\Theta(T_2) - \Theta(T_1)]}{\ln[T_2 - T_1]} \quad (5.10)$$

Dá se dokázat, že pořadnice bodu na přímkové závislosti, odečtená pro  $T = 1$  udává hodnotu druhého parametru  $\tilde{\lambda}$ .

Uvedené způsoby odhadu jsou poměrně jednoduché a postačují pro rychlé, předběžné odhady. Pro korektnost informace je nutné poznamenat, že graficko-analytický způsob odhadu parametrů je značně citlivý na přesnost odhadu parametrů  $\tilde{\lambda}$  a  $\hat{\beta}$ .

### B. Maximálně věrohodný odhad parametrů

Je další možný postup odhadu parametrů modelu monitorování „růstu ukazatelů“.

#### Popis a analytický model metody

Jak bylo dříve zdůvodněno, bude se při monitorování vývoje ukazatelů bezporuchovosti obecně předpokládat platnost nehomogenního Poissonova procesu vzniku poruch. U tohoto modelu platí pro popis hustoty pravděpodobnosti poruch Weibullova typu rovnice

$$f(T) = \frac{\beta}{\eta} \cdot \left( \frac{T^*}{\eta} \right)^{\beta-1} \cdot e^{-\left(\frac{T}{\eta}\right)^\beta} \quad (5.11)$$

Jestliže označíme jako dříve symbolicky

$$\tilde{\lambda} = \frac{1}{\eta^\beta}$$

potom je hustota pravděpodobnosti Weibullova typu dána vztahem

$$f(T) = \tilde{\lambda} \cdot \beta \cdot T^{*(\beta-1)} \cdot e^{-\tilde{\lambda} \cdot T^\beta} \quad (5.12)$$

Nyní je možné zformulovat funkci věrohodnosti, potřebnou pro odhad:

$$L = \tilde{\lambda}^n \cdot \beta^n \cdot e^{-\tilde{\lambda} T^\beta} \cdot \prod_{i=1}^n T_i^{*(\beta-1)}$$

Nebo v logaritmickém tvaru:

$$\Lambda = n \ln \tilde{\lambda} + n \ln \beta - \tilde{\lambda} T^{*\beta} + (\beta - 1) \cdot \sum_{i=1}^n \ln T_i \quad (5.13)$$

Derivace rovnice (5.13) podle  $\tilde{\lambda}$  bude:

$$\frac{\partial \Lambda}{\partial \tilde{\lambda}} = \frac{n}{\tilde{\lambda}} - T^{*\beta}$$

A z podmínky, že tato rovnice je rovna nule dostaneme vztah pro odhad  $\tilde{\lambda}$ , tedy:

$$\tilde{\lambda} = \frac{n}{T^{*\beta}} \quad (5.14)$$

Kde:  $T^* = T_n$ , když zkouška je ukončena poruchou (necenzurovaná zkouška);

$T^* = T < T_n$  když zkouška není ukončena poruchou (cenzurovaná zkouška);

Nyní derivujeme-li rovnici (5.13) podle  $\beta$ , potom bude:

$$\frac{\partial \Lambda}{\partial \beta} = \frac{n}{\beta} - \tilde{\lambda} T^{*\beta} \cdot \ln T^* + \sum_{i=1}^n \ln T_i$$

A z podmínky, že tato rovnice je rovna nule dostaneme vztah pro odhad  $\hat{\beta}$ , tedy:

$$\hat{\beta} = \frac{n}{n \ln T^* - \sum_{i=1}^n \ln T_i} \quad (5.15)$$

Maximálně věrohodné odhady parametrů  $\tilde{\lambda}$  a  $\hat{\beta}$  získáme řešením rovnic (5.14) a (5.15)

### **C. Odhad parametrů metodou nejmenších čtverců**

Je třetí možný metodologický postup odhadu parametrů ukazatelů bezporuchovosti v metodě AMSAA.

#### ***Popis metody***

Při odhadu parametrů touto metodou se vychází z rovnice, obsahující hledané parametry, přičemž lineární regrese se provádí analyticky metodou nejmenších čtverců.

Platí výchozí výpočtový vztah (5.9) ve tvaru:

$$\ln \Theta(T) = \ln \tilde{\lambda} + \beta \cdot \ln T \quad (5.9)$$

Pro zjednodušení výpočtu zavedeme substituční vztahy:

$$\ln \Theta_i(T) = Y_i \quad ; \quad \ln \tilde{\lambda} = a; \quad \beta = b; \quad \ln T_i = X_i$$

pomocí kterých rovnice (5.9) přejde na tvar:

$$Y_i = \hat{a} + \hat{b}X_i \quad (5.16)$$

což je rovnice přímky v souřadnicích (X, Y).

Soubor informací ze zkoušky má vždy charakter párových údajů typu: (X<sub>1</sub>, Y<sub>1</sub>); (X<sub>2</sub>, Y<sub>2</sub>); .....;(X<sub>N</sub>, Y<sub>N</sub>). K odhadu parametrů se použije metoda minima čtverců, která minimalizuje svislé odchylky experimentálních bodů od lineárně regresního modelu, nahrazujícího soubor těchto dat lineárním modelem (přímkovou závislostí proměnných (X<sub>i</sub>, Y<sub>i</sub>).

Po jednoduchých úpravách budou odhady hledaných parametrů:

$$\tilde{\lambda} = \exp \left[ \frac{1}{N} \left( \sum_{i=1}^N \ln \Theta_i - \beta \cdot \sum_{i=1}^N \ln T_i \right) \right] \quad (5.17)$$

a pro druhý parametr

$$\hat{\beta} = \frac{\sum_{i=1}^N \ln T_i \cdot \ln \Theta_i - \frac{1}{N} \cdot \left( \sum_{i=1}^N \ln T_i \cdot \sum_{i=1}^N \ln \Theta_i \right)}{\sum_{i=1}^N (\ln T_i)^2 - \frac{1}{N} \cdot \left( \sum_{i=1}^N \ln T_i \right)^2} \quad (5.18)$$

Odhad parametrů  $\tilde{\lambda}$  a  $\hat{\beta}$  získáme řešením rovnic (5.17) a (5.18).

#### 5.4 Intervalové odhady parametrů v metodě AMSAA

V této části jsou popsány prakticky použitelné postupy odhadu konfidenčních intervalů parametrů bezporuchovosti. Výpočtové postupy jsou účelovou aplikací výpočtových postupů, uvedených v ČSN IEC 605-4 a ČSN 010611. K odhadům můžeme alternativně použít dva metodologické přístupy:

*První postup:* bude použit pro soubory s četností údajů menších než 15 (počet stupňů volnosti menším než 30). Je založen na známém způsobu odhadu konfidenčního intervalu parametru bezporuchovosti s využitím chi-kvadrát rozdělení.

*Druhý postup:* bude použit pro soubory s četností údajů větších než 15 (počet stupňů volnosti větší než 30). Použitý postup je založen na poznatku, že u metody maximálně věrohodných odhadů parametrů větších souborů se jejich rozptyl (resp. p100%-ní kvantily) řídí standardním normovaným normálním rozdělením.

#### Odhad konfidenčních mezí pomocí chí-kvadrát rozdělení

Tento postup odhadu konfidenčních mezí se použije, jak již bylo uvedeno, v případech, kdy počet údajů o poruchách v hodnoceném souboru je menší než 15, protože konfidenční interval v těchto případech je nesouměrný (vychýlený), čemuž lépe vyhovuje model s použitím vlastností chí-kvadrát rozdělení.

Výpočtové vztahy pro konfidenční meze pro intenzitu poruch  $\lambda(T)$ , resp. MTBF, pro zvolenou konfidenční úroveň  $C = 1 - \gamma$  jsou následující, (hodnoty chí-kvadrát rozdělení jsou tabelovány):

U podnikových zkoušek se použijí následující výpočtové vztahy pro bodový odhad parametrů:

$$\lambda(T) = \frac{\chi_{C;(2r+2)}^2}{2T^*} \quad \text{pro odhad } \lambda \quad (5.19)$$

$$m(T) = \frac{2T^*}{\chi_{C;(2r+2)}^2} \quad \text{pro odhad MTBF} \quad (5.20)$$

Kde:  $T^*$  je celková doba trvání zkoušky (kumulativní doba zkoušky všech objektů);

$\chi^2$  je hodnota chí-kvadrát rozdělení pro příslušné  $C$  a  $v$ ;

$C = 1 - \gamma$  je konfidenční úroveň, na které se požaduje vyhodnocení zkoušky;

$\gamma$  - je nejvyšší přípustná chyba ve výsledku zkoušky;

$v = (2r+2)$  - počet stupňů volnosti;

$r$  - počet poruch ve zkoušeném souboru;

Je možné odhadovat jednostranné nebo i dvoustranné konfidenční intervaly parametrů podle vtaů:

Jednostranný konfidenční interval s horní mezí pro  $\lambda$ :

$$\lambda \langle \hat{\lambda} \frac{\chi_{c;(2r+2)}^2}{2r} ; \quad \text{nebo} \quad \lambda \langle \frac{\chi_{c;(2r+2)}^2}{2T^*} \quad (5.21)$$

Nebo dvoustranný konfidenční interval pro  $\lambda$  se vypočte ze vztahu

$$\frac{\chi_{(1-\gamma/2);(2r)}^2}{2T^*} \langle \lambda \langle \frac{\chi_{\gamma/2;(2r+2)}^2}{2T^*} \quad (5.22)$$

Jednostranný konfidenční interval s dolní mezí pro MTBF :

$$m \rangle \hat{m} \frac{2r}{\chi_{c;(2r+2)}^2} ; \quad \text{nebo} \quad m \rangle \frac{2T^*}{\chi_{c;(2r+2)}^2} \quad (5.23)$$

Dvoustranný konfidenční interval pro MTBF se vypočte ze vztahu:

$$\frac{2T^*}{\chi_{\gamma/2;(2r+2)}^2} \langle m \langle \frac{2T^*}{\chi_{(1-\gamma/2);(2r)}^2} \quad (5.24)$$

Hodnoty chí-kvadrát rozdělení jsou uvedeny v tabulkách.

Pro počet stupňů volnosti větším než 30 se doporučuje použít přibližného vztahu:

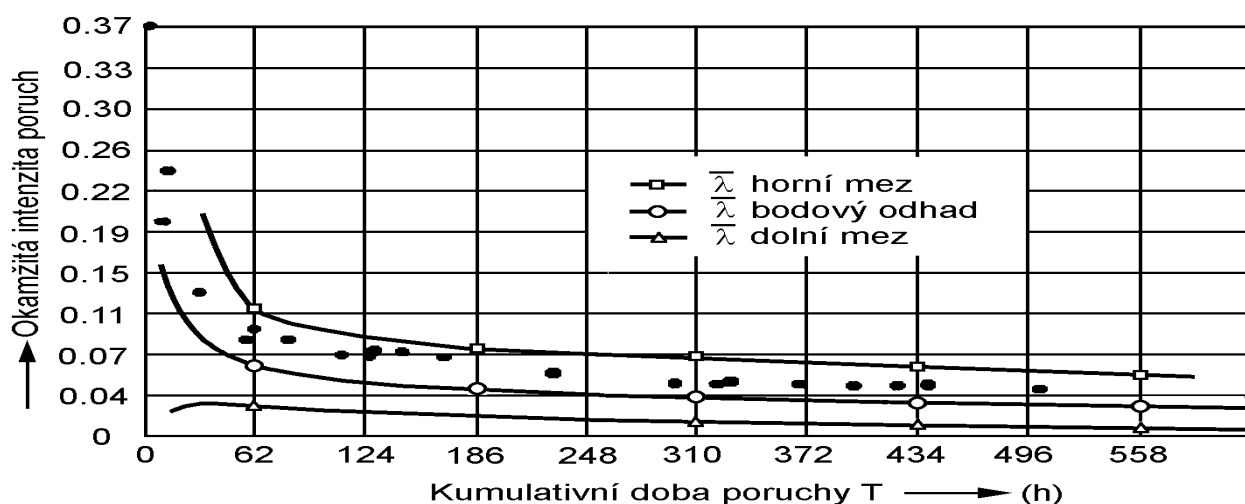
$$\chi_{c;v}^2 = \frac{1}{2} \left[ (z_p + \sqrt{2v-1})^2 \right];$$

Kde  $v = (2r+2)$  - počet stupňů volnosti;

$r$  - počet hodnocených poruch v souboru;

$z_p$  - je  $p$ -%ní kvantil normovaného normálního rozdělení

Demonstrativní příklad vývoje intenzity poruch je na Obr. 5.3



Obr.5.3 90%-ní konfidenční interval pro okamžitou intenzitu poruch, stanovenou pomocí chí-kvadrát rozdělení.

#### Odhad konfidenčních mezí metodou maximální věrohodnosti

Maximálně věrohodné odhady je možné použít pouze v případech, kdy počet údajů v souboru je větší než 15 (počet stupňů volnosti větší než 30). Pokud nastane taková situace a k odhadu parametrů bude použita metoda maximální věrohodnosti, postup výpočtu konfidenčních mezí intenzity poruch, nebo MTBF bude proveden za předpokladu normálního rozdělení hodnot pro daný časový okamžik  $T$ . Tento předpoklad je korektní, protože konfidenční intervaly pro parametry jsou pro každé  $T$  již prakticky souměrné vzhledem ke střední hodnotě bodových odhadů a jsou již prakticky shodné s odhady pomocí chí-kvadrát rozdělení.

Takže pro dvoustranný odhad konfidenčních mezí (CM) parametrů budou použity následující výpočtové vztahy:

$$CM(\lambda) = \hat{\lambda}(T) \pm z_p \sqrt{\text{Var}(\hat{\lambda}(T))} \quad - \text{ pro intenzitu poruch} \quad (5.25)$$

$$CM(m) = \hat{m}(T) \pm z_p \sqrt{\text{Var}(m(T))} \quad - \text{ pro MTBF} \quad (5.26)$$

kde  $z_p$  je  $p$ - kvantil normálního normovaného rozdělení, daného v Příloze č. 2.

#### Odhad rozptylu náhodné veličiny:

Např. pro intenzitu poruch, označený symbolicky  $\text{Var}(\hat{\lambda}(T))$  bude dán vztahem:

$$\begin{aligned} \text{Var}(\hat{\lambda}(T)) = & \left( \frac{\partial \lambda(T)}{\partial \beta} \right)^2 \cdot \text{Var}(\hat{\beta}) + \left( \frac{\partial \lambda(T)}{\partial \lambda} \right)^2 \cdot \text{Var}(\hat{\lambda}) + \\ & + 2 \left( \frac{\partial \lambda(T)}{\partial \beta} \right) \cdot \left( \frac{\partial \lambda(T)}{\partial \lambda} \right) \cdot \text{Cov}(\hat{\beta}, \hat{\lambda}) \end{aligned} \quad (5.27)$$

Odhadované hodnoty jsou:  $\beta = \hat{\beta}$ ;  $\lambda = \hat{\lambda}$ .

Hodnoty rozptylů,  $\text{Var}(\hat{\lambda})$ ;  $\text{Var}(\hat{\beta})$ ; a  $\text{Cov}(\hat{\beta}, \hat{\lambda})$ , potřebných pro odhad konfidenčních mezí parametrů, můžeme vypočítat použitím Fišerovy matice a její inverzní hodnoty [2].

Rovnice pro funkci věrohodnosti má tvar.

$$\Lambda = n \ln \lambda + n \ln \beta - \lambda T^{*\beta} + (\beta - 1) \cdot \sum_{i=1}^n \ln T_i \quad (5.28)$$

Pro  $\Lambda$  funkci věrohodnosti, potom platí:

$$\left[ \begin{array}{cc} -\frac{\partial^2 \Lambda}{\partial \lambda^2} & -\frac{\partial^2 \Lambda}{\partial \lambda \partial \beta} \\ -\frac{\partial^2 \Lambda}{\partial \lambda \partial \beta} & -\frac{\partial^2 \Lambda}{\partial \beta^2} \end{array} \right]_{\beta=\hat{\beta}; \lambda=\hat{\lambda}}^{-1} = \left[ \begin{array}{cc} \text{Var}(\hat{\lambda}) & \text{Cov}(\hat{\beta}, \hat{\lambda}) \\ \text{Cov}(\hat{\beta}, \hat{\lambda}) & \text{Var}(\hat{\beta}) \end{array} \right] \quad (5.29)$$

Vztahy pro jednotlivé parciální derivace budou ve tvaru:

$$\frac{\partial \Lambda}{\partial \lambda} = \frac{n}{\lambda} - T^{*\beta}; \quad \frac{\partial^2 \Lambda}{\partial \lambda^2} = -\frac{n}{\lambda^2}; \quad \frac{\partial^2 \Lambda}{\partial \lambda \partial \beta} = -T^\beta \cdot \ln T;$$

$$\frac{\partial \Lambda}{\partial \beta} = \frac{n}{\beta} - \lambda T^{*\beta} \cdot \ln T^* + \sum_{i=1}^n \ln T_i; \quad \frac{\partial^2 \Lambda}{\partial \beta^2} = -\frac{n}{\beta^2} - \lambda T^\beta \cdot (\ln T)^2;$$

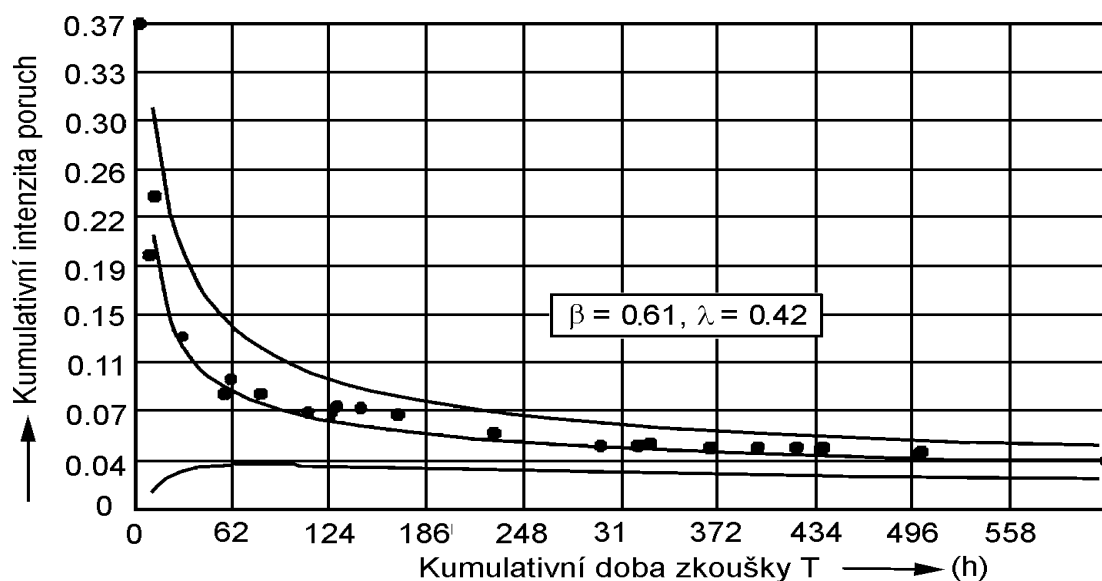
$$\frac{\partial \lambda_i(T)}{\partial \beta} = \lambda T^{(\beta-1)} + \lambda \beta T^{(\beta-1)} \cdot \ln T; \quad \frac{\partial \lambda_i(T)}{\partial \lambda} = \beta T^{(\beta-1)};$$

$$\frac{\partial \lambda_c(T_i)}{\partial \beta} = \lambda T^{(\beta-1)} \cdot \ln T; \quad \frac{\partial \lambda_c(T)}{\partial \lambda} = T^{(\beta-1)};$$

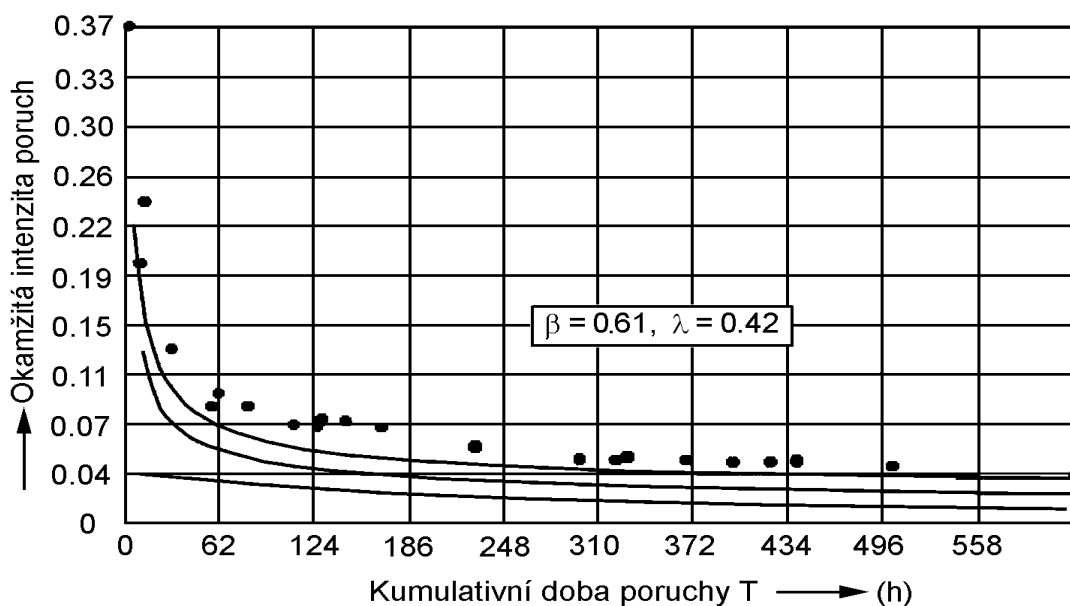
S použitím těchto rovnic můžeme potom hodnoty rozptylů vypočítat postupem uvedeným v [2], takže např. pro intenzitu poruch bude:

$$\text{CM}(\lambda) = \hat{\lambda}(T) \pm z_p \sqrt{\text{Var}(\lambda(T))}$$

Demonstrativní příklady konfidenčních intervalů jsou na Obr. 5.4 a Obr. 5.5.



*Obr.5.4 90%-ní konfidenční intervaly pro kumulativní intenzitu poruch*



*Obr.5.4.5 90%-ní konfidenční intervaly pro okamžitou intenzitu poruch*

Podobně jak byl prováděn odhad konfidenčních mezí pro kumulativní a okamžitou intenzitu poruch budou provedeny odhady i pro MTBF.

## 6 Regulační diagram průběhu zkoušky

Je grafickou technikou umožňující vyhodnotit ukazatele spolehlivosti (např. MTBF a j.) na libovolně zvolené konfidenční úrovni  $C$  (např. jako „zaručované“ nebo „mezí“ hodnoty



ukazatelů). Je to postup vhodný pro výrobky u nichž se již v průběhu zkoušek neodstraňují příčiny systematických poruch a které mají proto úroveň spolehlivosti již stabilizovanou. Technika umožňuje též dopředu stanovit očekávaný rozsah zkoušky.

Vychází se z grafické interpretace známého vztahu (6.1) pro odhad dolní konfidenční meze odhadovaného ukazatele spolehlivosti, např.  $\bar{t}_D = MTBF_D$ :

Pro odhad rozsahu zkoušky a pro vyhodnocení jejího průběhu použijeme vztah pro dolní (index D) konfidenční mez parametru  $\bar{t}$  ve tvaru:

$$\bar{t}_D \geq \frac{2 \cdot T_c}{\chi_{2\nu, C}^2} \quad (6.1)$$

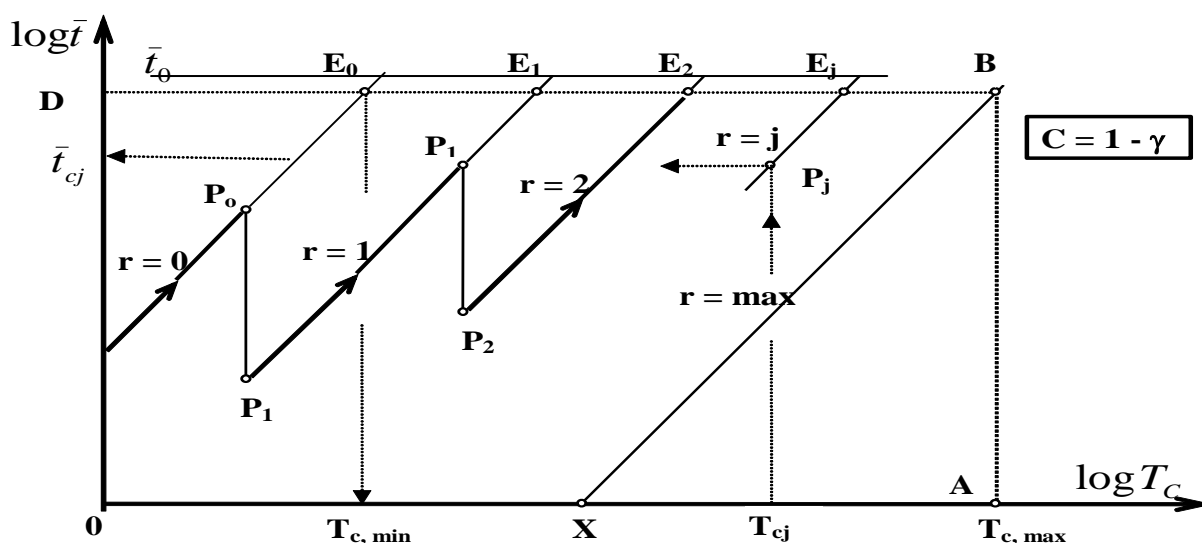
kde:  $2\nu = 2(r + 1)$ ,  $r$  - počet poruch,  $C$  - je zadaná konfidenční úroveň zkoušky

### 6.1 Diagram znázorňující průběh zkoušky spolehlivosti.

Vztah (10) umožňuje sestavit tzv. "Regulační diagram zkoušky spolehlivosti", který je v log-log souřadnicích znázorněn na obr.6.1 Diagram je sestaven pro jednu zvolenou konfidenční úroveň  $C$ . Průběh zkoušky je v daném případě vymezen těmito okrajovými veličinami:

- maximální disponibilní dobou pro zkoušku  $T_{c, \max}$  (bod A na vodorovné ose),
- velikostí požadovaného ukazatele bezporuchovosti  $\bar{t}_0$  (bod D na svislé ose).

Těmito veličinami je v obrázku vymezen obdélník OABD a úsečka BX vycházející z bodu B. Rovnoběžné přímky s parametrem  $r$  vyhovují v  $\log T_c - \log \bar{t}$  souřadnicích rovnici (6.1). Skutečná zkouška pak probíhá za kompromisních podmínek, které vyplývají z obr.6.1.



Obr.6.1 Diagram zkoušky spolehlivosti

## 6.2 Průběh zkoušky.

- Zkouška je zahájena v čase  $T_c = 0$  a v okamžiku kdy byly ověřeny všechny jeho funkce a proveden nezbytný "záběh". Ten by neměl trvat déle, než je nezbytně nutné.
- Od okamžiku zahájení zkoušky je třeba vést veškerou předepsanou dokumentaci.
- V případě výskytu poruch je třeba provést jejich klasifikaci podle Obr. 2.1.
- Započitatelné poruchy postupně vynášíme do připraveného diagramu, Obr. 6.1.
- Okamžitá kumulativní hodnota doby trvání zkoušky  $T_c$  na vodorovné ose v diagramu a její průsečík s přímkou, odpovídající počtu dosud vzniklých poruch  $r = 0,1,2,\dots$  udává bod  $P_j$ , který definuje hodnotu **právě prokázané úrovně** ukazatele bezporuchovosti  $\bar{t}_{cj}$ .
- V diagramu může být vyznačena i požadovaná hodnota ukazatele bezporuchovosti (např. z Technických podmínek  $\bar{t}_0$ ).
- Příмка znázorňující zkoušku, která probíhá bez poruchy ( $r = 0$ ), musí protínat úsečku **BD** v bodě  $E_0$  ležícím nalevo od bodu **B**. Jinak by zkouška neměla smysl, protože by nedošlo v rámci zadaného rozsahu zkoušky  $T_{c\max}$  k ověření požadovaného ukazatele  $\bar{t}$  ani při bezporuchové zkoušce.
- Pokud zkouška probíhá bez poruchy, může být ukončena již v bodě  $E_0$ , tj. po uplynutí  $T_{c\min} < T_{c\max}$ , protože v bodě  $E_0$  došlo k ověření požadované hodnoty  $\bar{t}$ . Časové úspory jsou dány rozdílem  $T_{c\max} - T_{c\min}$ .
- Pokud v průběhu zkoušky nastanou poruchy, přechází se postupně z přímkou  $r = 0$  na přímkou  $r = 1$  při první poruše (body  $P_0 - P_1$ ), z  $r = 1$  na  $r = 2$  při druhé poruše atd., až nejvýše na přímkou vycházející z bodu **B** a znamenající  $r = \max$ , tj. maximální přípustný počet poruch v rámci vymezených podmínek zkoušky. Tato příмка vymezení na ose T bod **X**. Do té doby je buď:
  - protnuta úsečka **BD** v bodě  $E_j$  ( $j = 1,2, \dots, r = \max$ ) a zadaný parametr je ověřen a zkouška může být ukončena v plánovaném čase, nebo
  - je protnuta úsečka **BX** přechodem z bodu  $P_j$  do bodu  $P_{j+1}$ , tj. je překročen  $r = \max$  a zkouška nemůže být ukončena v plánovaném termínu nebo požadovaný parametr  $\bar{t}$  nebude stanoven s požadovanou konfidencí. Pouze za předpokladu, že bude prodloužena doba trvání zkoušky  $T_{c\max}$  nebo snížena konfidenční úroveň, může zkouška pokračovat dále.

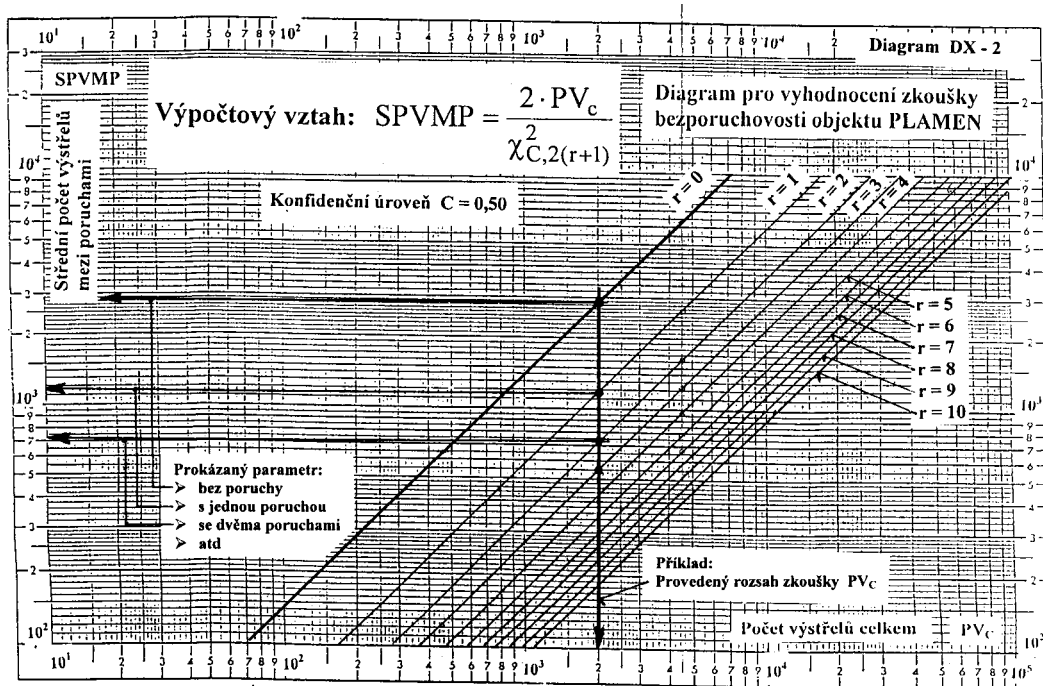
**Zkouška končí dosažením některého z bodů  $E_j$ , nebo protnutím spojnice  $BX = r_{\max}$ .**

- Protože z povahy použitého modelu a vlastností náhodných veličin vyplývá, že v průběhu zkoušky dochází ke kumulaci doby provozu u všech zkoušených výrobků, platí, že  $T_c$  ve výrazu (6.2) nepředstavuje dobu trvání zkoušky měřenou v reálném čase (na "hodinách"), ale je to celková kumulativní doba zkoušky  $T_c$  která je závislá na počtu zkoušených výrobků, počtu poruch a typu výběrového plánu.
- Při počtu výrobků ve zkoušce  $n$  můžeme dobu zkoušení  $T_c$  stanovit podle vztahu :

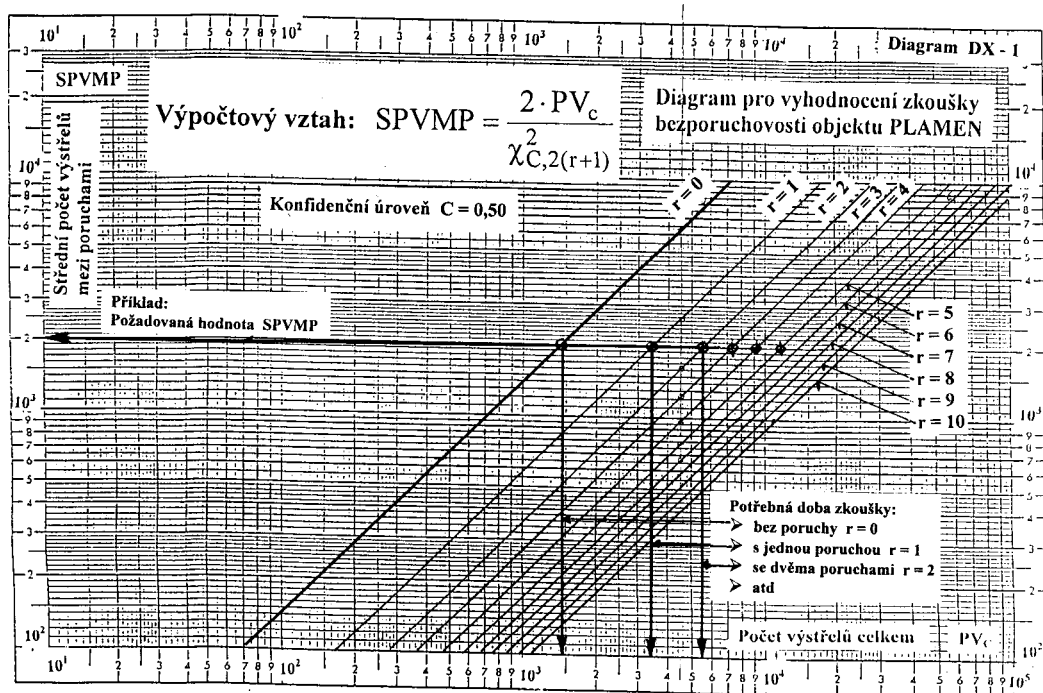
$$T_c = \sum_{i=1}^n T_{ci} \quad (6.2)$$

Postup zkoušení je použitelný i u neobnovovaných výrobků, tj. ve zkoušce, při které zkoušíme  $n$  výrobků, každý do první poruchy, po které se prvek ze zkoušky vyřazuje. Předpokládá se přitom, že doba do poruchy se u všech výrobků řídí společným zákonem rozdělení. I v tomto případě lze kumulovat dobu zkoušení jednotlivých výrobků do společného souboru pozorovaných veličin.

Příklad praktického použití regulačního diagramu zkoušky.



Obr. 6.2 Příklad stanovení prokázaného ukazatele MTBF z dané doby zkoušky  $PV_c$  v závislosti na počtu poruch ve zkoušce



Obr. 6.3 Příklad stanovení potřebné doby zkoušky  $PV_c$  pro průkaz zadané hodnoty MTBF v závislosti na počtu poruch ve zkoušce.

## 7 Použité podklady

- [1] ČSN IEC 50(191) Medzinárodný elektrotechnický slovník. Kapitola 191: Spôľahlivosť a akosť služieb
- [2] ČSN 010611 Spolehlivost v technice. Pravidla pro stanovení bodových a intervalových ukazatelů spolehlivosti. Parametrické metody.
- [3] MIL-HDBK-189 Reliability Growth Management
- [4] ČSN IEC 300-3-2. Sběr dat o spolehlivosti v provozu.
- [5] ČSN IEC 1014 Programy růstu bezporuchovosti.
- [6] ČSN IEC 1164 Růst bezporuchovosti – Metody statistických testů a odhadů.
- [7] MIL-STD-1635 Reliability growth testing
- [8] MIL-HDBK-189.: Reliability Growth Management; Reliability Growth Models. 1981.
- [9] IEC 61014, Ed.2.: Programmes for Reliability Growth; 2001.
- [10] MIL-HDBK-781A Reliability Test Methods, Plans and Environments for Engineering Development, Qualification and Production
- [11] HOLUB R.; VINTR Z.; Metodika hodnocení bezporuchovosti kompletu PLAMEN v průběhu podnikových a kontrolních zkoušek; Výzkumná zpráva; VA Brno, 2002.
- [12] HOLUB R.; VINTR Z.; Návod y a postupy pro zkoušky spolehlivosti. Metodická pomůcka pro Škoda Transportation s.r.o.; Univerzita obrany Brno; Výzkumná zpráva; 2006.

# SOFT METODY ANALÝZY SPOLEHLIVOSTI – FUZZY LOGIKA A JEJÍ APLIKACE V ANALÝZÁCH SPOLEHLIVOSTI

*Ing. David VALIŠ, Ph.D., Fakulta vojenských technologií, Univerzita obrany v Brně*  
[david.valis@unob.cz](mailto:david.valis@unob.cz)

## 1 ÚVOD

Neurčitost v lidském životě s sebou nese mnohá úskalí a její stanovení není úplně jednoduché. Fuzzy logika je matematickou disciplínou, která si získala značnou popularitu na začátku devadesátých let minulého století. Pravděpodobně pro své fascinující aplikace uskutečněné v Japonsku a později také v jiných zemích. Nejúspěšnější příklady aplikací jsou v řízení a regulaci, existují ovšem i mnohé další jako například v rozpoznávání obrazu, klasifikaci, rozhodování, a další. Hlavním zdrojem úspěchu je fakt, že fuzzy logika dokáže zahrnout nepřesnost a poměrně jednoduchým způsobem pracovat s významy slov přirozeného jazyka (kvalitativním popisem), který patří k přirozeným a nejdůležitějším částem lidského života. To je asi důvod, proč fuzzy logika zasahuje do tolika oblastí lidské činnosti a stále nachází další možnosti svého využití.

Tento materiál je určen pro využití v rámci analýz spolehlivosti, kde známe již některé další „soft metody“ analýzy. Kromě zde zmiňované fuzzy logiky známe a využíváme ještě např. metody:

- teorie možnosti;
- teorie nepřesnosti v pravděpodobnosti;
- Dempster-Shafer teorie;
- teorie sítí;
- teorie grafů, aj.

V dalším pojednání se budeme soustředit na teorii fuzzy množin a její možné využití právě v analýzách spolehlivosti a bezpečnosti/rizika.

Často se ale objevují pochybnosti typu, jestli to není pouze předstírání, anebo proč tomu tak je? Respektive jestli nejde jenom o matení pojmů? Předně je nutné zdůraznit, že již existence mnohých úspěšných aplikací vyvrací pochybnosti o nějakém předstírání. Fuzzy logika nabídla překvapivě jednoduchá řešení, která opravdu fungují. Častá námitka, že to, co je řešeno pomocí fuzzy logiky lze řešit i bez ní, neobstojí. Příkladů může být velké množství. Jedním z nich je rameno, které bylo použito k opravě Hubbleova vesmírného teleskopu. Bylo vyvinuto za 14 dní a vedlo k plně uspokojivému chování. Za stejnou dobu by byl klasický PID regulátor nepoužitelný. Není pravda, že by to pomocí PID nešlo vyřešit, ale šlo o úspory času a financí, a navíc byl výsledek rovnocenný.

Proč tedy fuzzy logika funguje? Odpověď je poměrně překvapivá. Protože využívá vágně charakterizované expertní znalosti. Tedy pravý opak toho, co se vždy požadovalo – větší přesnost. Narážíme zde na reálný rozpor, jehož řešení neexistuje. Jde o klasický vztah mezi relevancí a přesností informace. Zde jsme narazili na dvě zásadní věci. Jednak na potřebu vyjádřit relevantní informaci, což je nezbytné učinit pomocí přirozeného jazyka. Protože jazyk je dosud jediný a velmi účinný prostředek, který nám umožňuje efektivně pracovat s vágními pojmy. Druhá věc se týká samotné podstaty přesnosti. Ukazuje se, že přesnost je pouze iluze, neboť je principiálně nedosažitelná. Pokud by byla, pak by to znamenalo, že při libovolně

zvolené hranici přesnosti dostaneme stále stejný výsledek. Problém se dá ilustrovat na příkladu srovnání objektů, například objemu dvou nádob. Pokud jsou obě absolutně přesně stejné, pak to znamená, že i jejich objem je stejný (ať měřen v  $m^3$ , nebo  $mm^3$  a dále v jejich libovolných zlomcích bilióntiny, kvadrilióntiny, apod.). Potom se ovšem dostaneme do konceptuálních problémů, protože bychom asi nebyli schopni rozlišit, který atom či dokonce elementární částice je součástí nádoby, a který není. Snaha o absolutní přesnost nás vždy dovede do sporu. Není však třeba litovat, neboť vágnost, kterou přirozený jazyk umí tak dokonale využít, je jeho hlavní silou, nikoliv nedostatkem.

Všechna uvedená fakta stojí v pozadí úvah zakladatelů fuzzy logiky (hlavním představitelem je profesor kalifornské university v Berkeley, Lofti A. Zadeh). Fuzzy logika, která vychází z teorie fuzzy množin, se zaměřuje na vágnost, kterou se snaží matematicky zachytit.

Základním pojemem v teorii pravděpodobnosti je *rozdělení pravděpodobnosti*. To charakterizuje způsob nastoupení jevů vybíraných z nějaké množiny různých jevů, o nichž víme jenom to, že jeden z nich určitě nastane. Pravděpodobnost nám potom dává informaci o tom, zda nastoupení daného jevu můžeme očekávat s větší jistotou, než nastoupení jiného jevu.

Dalším důležitým pojemem je *nezávislost jevů*. Pokud jsou jevy nezávislé, potom pravděpodobnost toho, že nastanou současně, je rovna *součinu* jejich pravděpodobností.

Naproti tomu uvažujme např. červené objekty. Pak si musíme nejdříve odpovědět na otázku co to je „červený“, aniž bychom hovořili o tom, zda nějaký jev nastane či ne. Jde totiž o *vymezení* jevu a nikoliv charakterizace toho, zda nastane či ne. Základním pojemem je zde fuzzy množina objektů a stupeň příslušnosti objektu do ní. Stupně příslušnosti, stejně jako pravděpodobnosti, mohou být čísla z intervalu  $[0,1]$ . To je však jen vnějšková shoda. Interpretace obou pojmů je zcela rozdílná, stejně jako manipulace s nimi. U stupňů příslušnosti nemá smysl hovořit o závislosti a základní operací je hledání minima, nikoliv násobení.

V obou případech se setkáváme s obecnějším jevem, který můžeme nazvat *neurčitost*. Neurčitost má (nejméně) dvě vzájemně komplementární stránky: *vágnost a nejistota*. Vágnost lze modelovat pomocí teorie fuzzy množin, zatímco nejistotu pomocí teorie pravděpodobnosti a popř. dalších teorií, jako jsou např. teorie možnosti, věrohodnosti, apod. S trochou nadsázky lze tedy říci, že pravděpodobnost nám odpovídá na otázku, zda „něco nastane“, zatímco teorie fuzzy množin nám odpovídá na otázku „co vlastně nastalo“?

## 2 Obecné základy teorie fuzzy množin a jazyková proměnná

V této části bych chtěl stručně pojednat o teorii fuzzy množin a modelování sémantiky přirozeného jazyka pomocí fuzzy množin. V teorii fuzzy množin hrají významnou úlohu uspořádané množiny, popř. svazy. Proto budeme často používat běžné symboly pro svazové operace. Ústředním pojemem ve fuzzy logice je fuzzy množina, což je zobecněným pojemem množiny. Představíme si například, že někdo chce, aby se specifikovali množinu výšek všech *velkých* lidí. Nejprve můžeme říci, že každý vysoký člověk má výšku např. mezi 160 cm a 240 cm. Odrazovým můstkem bude množina  $U = (160; 240)$  cm. Dále ovšem narazíme na nepřekonatelné potíže. Zjistíme totiž, že nejsme schopni specifikovat velké lidi přesně. Např. hranice vysokých lidí 175 cm může apriori vylučovat osoby s výškou 174,6 cm, které nejsme schopni podle oka odlišit od osob vysokých právě 175 cm, a tak dále...

V teorii fuzzy množin musíme vyjít z množiny všech myslitelných výšek  $U = (40; 240)$  [cm], kterou nazveme *universum*. Každé v praxi myslitelné výšce uvažovaného universa přiřadíme číslo z intervalu  $[0,1]$ , které bude vyjadřovat *stupeň pravdivosti* tvrzení, že daná výška označuje

„velkého člověka“. Stupeň pravdivosti „0“ vyjadřuje naprostou nepravdu (naprostý nesouhlas) zatímco stupeň „1“ znamená naprostou pravdu (bezvýhradný souhlas). Čísla mezi těmito dvěma hodnotami vyjadřují částečný souhlas, který je tím větší, čím je větší stupeň pravdivosti. Definovaný stupeň pravdivosti je tedy zároveň stupněm příslušnosti dané výšky do fuzzy množiny všech výšek velkých lidí. Podle našeho příkladu můžeme tedy říci, že např. „165 cm vysoký člověk“ je pravda ve stupni „0,3“, zatímco „190 cm vysoký člověk“ je pravda ve stupni „1“ (někdy také absolutní pravda, nebo pravda ve 100%). Ukazuje se totiž, že je často velmi užitečné vyjadřovat stupně pravdivosti (příslušnosti) v jednotkách „%“.

Fuzzy množinu „výšek lidí“ můžeme charakterizovat pomocí následující funkce, která libovolné výšce (v cm) přiřadí stupeň pravdivosti podle tohoto předpisu:

$$A_{\text{malý}}(x) = \begin{cases} 1, & \text{jestliže } x \leq 165, \\ 0, & \text{jestliže } x > 185, \\ 1 - \frac{1}{2} \left( \frac{x-165}{10} \right)^2 \}, & \text{jestliže } 165 < x < 175, \\ \frac{1}{2} \left( \frac{185-x}{10} \right)^2 \}, & \text{jestliže } 175 \leq x \leq 185, \end{cases}$$

Podobnou úvahu lze provést i s jinými slovy přirozeného jazyka, např. výška, věk, nebo libovolné slovo, jehož obsahem jsou nějaké změřené hodnoty. Toto je zvláště zajímavé pro fuzzy regulaci (Zde je nutné dodat, že termín fuzzy množin je pochopitelně obecnější, avšak pro potřeby fuzzy regulace se zpravidla omezujeme jen na množiny definované na universu tvořeném čísly).

Při definici fuzzy množin tedy postupujeme takto. Nejprve definujeme množinu  $U$  nazývanou *universum diskurzu* nebo stručně *universum*. To může být množina prvků libovolného druhu. Např.. množina rostlin, lidí, nebo velmi často jistá množina čísel. Tento případ je významný zejména pro fuzzy regulaci, kde se pracuje s pojmy „odchylka“, „změna odchylky“ nebo „akční zásah“, a to jsou jistá čísla představující výsledky měření. *Fuzzy množina* je z matematického pohledu funkce

$$A: U \rightarrow [0,1] \tag{1}$$

Řečeno slovy, fuzzy množina je tvořena prvky  $x$  vybíranými z množiny  $U$ ,  $x \in U$ , z nichž každý má přiřazeno číslo  $a \in [0,1]$  nazývané *stupeň příslušnosti* prvku  $x$  do fuzzy množiny  $A$ . Zároveň je  $A(x)$  stupeň pravdivosti toho, že  $x$  patří do  $A$ . Stupně pravdivosti a stupně příslušnosti jsou tedy ztotožněny. Funkce uvedená ve výrazu (1) se někdy nazývá *funkce příslušnosti*. To znamená, že fuzzy množina je *ztotožněna* se svou funkcí příslušnosti. Stupeň příslušnosti prvku  $x \in U$  do fuzzy množiny  $A$  se zapisuje jako funkční hodnota  $A(x)$ . V odborné literatuře se někdy pro funkci příslušnosti používá speciální symbol  $\mu$ . Potom se stupeň příslušnosti prvku  $x$  do fuzzy množiny  $A$  zapisuje jako  $\mu_A(x)$ . To je však jednak nepřesné a matoucí, a navíc velmi nepraktické, neboť např. u složitěji definovaných fuzzy množin se nepřehledný výraz zapisuje do indexu (např.

$$\mu_{(A \cup B) \cap (B \cup \bar{D})}).$$

Fuzzy množina je tedy zobecněním klasické množiny také v následujícím smyslu. V teorii množin se definuje tzv. *charakteristická funkce*  $\chi_A: U \rightarrow \{0,1\}$  množiny  $A$  vzhledem k  $U$  takto:

$$\chi_A(x) = \begin{cases} 1, & \text{jestliže } x \in A, \\ 0, & \text{jestliže } x \notin A. \end{cases}$$

To znamená, že  $\chi_A(x) = 1$ , jestliže prvek  $x$  patří do množiny  $A$  a  $\chi_A(x) = 0$ , pokud do ní nepatří. Je ihned vidět, že funkce příslušnosti fuzzy množiny je zobecněním charakteristické funkce. Ztotožnění fuzzy množiny se svou funkcí příslušnosti je přirozené a není v rozporu s chápáním klasických množin, které jsou také často ztotožňovány se svými charakteristickými funkcemi. Všimněte si, že při definici fuzzy množiny vycházíme z universa, což je klasická množina. Tím je tedy navíc prokázáno, že fuzzy množiny rozšiřují nikoliv popírají pojem množiny.

Fakt, že  $A$  je fuzzy množina v universu  $U$  definovaná v (1) často zapisujeme symbolem  $A \subseteq U$ .

Explicitně se fuzzy množiny zapisují takto:

$$A = \{a_1 / x_1, \dots, a_n / x_n\}, \quad (2)$$

kde  $x_1, \dots, x_n \in U$  jsou prvky, kterým jsou přiřazeny stupně příslušnosti  $a_1, \dots, a_n \in (0, 1]$ , tzn. že prvky se stupně příslušnosti „0“ nejsou zahrnuty.

Příklad může být následující:

Uvažujme universum  $U = \{0, 1, \dots, 10\}$ . Pak

$$A = \{0.4/1, 0.7/2, 0.5/4, 1/6, 1/7, 0.1/9\} \quad (3)$$

je fuzzy množina  $U$  do níž číslo „1“ patří se stupněm příslušnosti 0.4, číslo 2 se stupněm příslušnosti 0.7, atd. Čísla z  $U$ , která nejsou ve výrazu (3) uvedena, mají stupeň příslušnosti roven „0“, tj. do  $A$  nepatří.

Není-li universum konečná množina a prvky fuzzy množiny nelze zapsat výčtem dle výrazu (2), zapisujeme fuzzy množiny takto:

$$A = \{a_i / x_i \mid i \in I\}, \quad (4)$$

kde  $i$  je určitá indexová množina, popř. lze podrobněji specifikovat vlastnosti  $x_i$  a  $a_i$ . Jsou-li například prvky  $x$  reálná čísla a stupně příslušnosti jsou dány nějakou funkcí, lze napsat fuzzy množinu takto:

$$A = \{f(x) / x \mid x \in R\}$$

kde  $R$  je množina všech reálných čísel. V odborné literatuře se můžeme setkat také s následujícím zápisem:

$$A = \int_{x \in R} f(x) / x.$$

Symbol integrálu je zde použit ve významu sjednocení a nikoliv ve svém původním významu. Smysl tohoto zápisu je v tom, že se na fuzzy množinu lze také dívat jako na sjednocení tzv. fuzzy jednoprvkových množin. V určitých publikacích je symbol integrálu nahrazen symbolem tzv. velkého sjednocení, potom je výraz psán ve tvaru:

$$A = \bigcup_{x \in R} f(x) / x.$$



Nejpřesnější je však přece jenom zápis (2) resp. (4). Důležitou roli v teorii fuzzy množin mají následující tři klasické množiny:

a) Nosič

$$Supp(A) = \{x | A(x) > 0\}$$

tj. nosič fuzzy množiny  $A$  je množina všech prvků universa, jejichž stupeň příslušnosti do  $A$  je nenulový. Tato množina je velmi důležitá, protože obsahuje všechny prvky, které jsou pro nás zajímavé (prvky se stupněm příslušnosti „0“ nejsou zajímavé, neboť mohou být docela libovolné).

Příkladem může být třeba nosič fuzzy množiny  $A$  ve výrazu (3), který je definován jako

$$Supp(A) = \{1, 2, 4, 6, 7, 9\}.$$

b)  $a$ -řez

$$A_a = \{x | A(x) \geq a\} \tag{5}$$

tj.  $a$ -řez je množina prvků majících stupeň příslušnosti větší nebo roven zadanému stupni  $a$ . Tuto množinu získáme z fuzzy množiny  $A$  „oddělením“ všech prvků se stupněm příslušnosti menším než  $a$ . Jako příklad můžeme využít výše uvedenou fuzzy množinu. Potom můžeme uvést:

$$A_{0.5} = \{2, 4, 6, 7\}$$

$$A_{0.7} = \{2, 6, 7\}$$

Pro  $a$ -řezy fuzzy množiny platí tento jednoduchý, avšak důležitý vztah:

$$\text{Jestliže } a \leq b, \text{ pak } A_b \subseteq A_a$$

Vztahy mezi fuzzy množinou a jejími řezy je velmi úzký. Lze dokázat následující rovnost:

$$A(x) = \bigvee_{x \in A_a} a \tag{6}$$

Rovnost ve výrazu (6) se nazývá *věta o reprezentaci* fuzzy množiny a znamená, že stupeň příslušnosti prvku  $x$  do fuzzy množiny  $A$  je roven supremu všech indexů  $a$  řezů, do nichž patří. Podle této věty tedy lze fuzzy množinu chápat jako posloupnost jejích  $a$ -řezů.

c) Jádro

$$Ker(A) = \{x | A(x) = 1\},$$

tj. jádro je množina těch prvků, které určitě patří do fuzzy množiny  $A$ . Představují typické prvky (prototypy) pro danou fuzzy množinu, např. typicky „velký“, „malý“, „dobrý“, apod. Ve výše uvedeném příkladu by typicky „velcí lidé“ byli lidé řekněme vyšší než 185 cm.

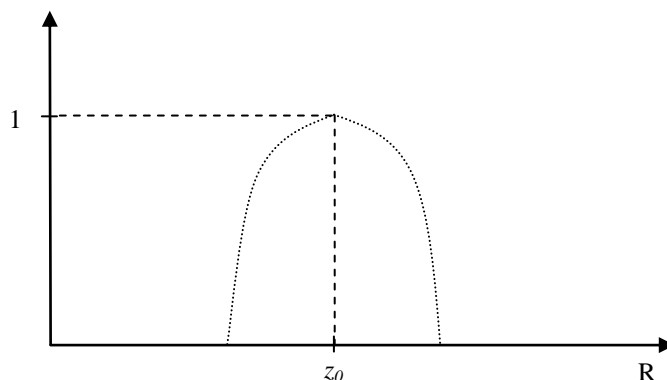
Příkladem jádra fuzzy množiny uvedené ve výrazu (2) je

$$Ker(A) = \{6, 7\}$$

Z uvedeného je zřejmé, že jádro fuzzy množiny je její 1-řez. Dále hovoříme o tom, že fuzzy množina je *normální* pokud  $Ker(A) \neq \emptyset$ . Fuzzy množina, která není normální, se nazývá *subnormální*. Je zřejmé, že stupeň příslušnosti všech jejích prvků jsou menší než 1.

Základní operace s fuzzy množinami jsou následující. S fuzzy množinami lze, podobně jako s klasickými množinami, definovat základní operace sjednocení, průniku a doplňku. Kromě nich však lze definovat ještě řadu dalších operací, které v klasické teorii množin buď nemají smysl nebo dávají výsledek, který je ekvivalentní s některou ze základních operací. To značně rozšiřuje možnosti teorie fuzzy množin.

Dalším důležitým pojmem jsou fuzzy čísla, která jsou v podstatě speciálními fuzzy množinami v množině reálných čísel  $\mathbb{R} = (-\infty; +\infty)$ . Zpravidla se předpokládá, že mají speciální tvar (viz Obr. 1Obr. 1:).



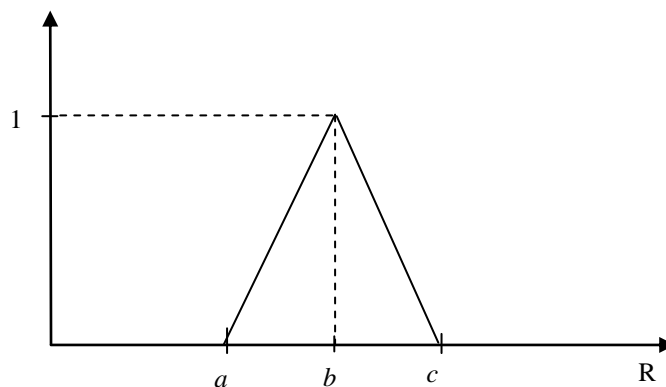
**Obr. 1:** Funkce příslušnosti fuzzy čísla „asi  $z_0$ “

Fuzzy číslo intuitivně reprezentuje hodnotu, která je nepřesná, tj. hodnotu, kterou lze slovně charakterizovat jako „asi  $z_0$ “, „zhruba  $z_0$ “, apod. Typické příklady jsou „asi 5“, „zhruba 1205“, „přibližně 1m“, apod. Taková čísla nejsou nijak výjimečná. Právě naopak, v praxi pracujeme téměř výhradně s čísly, která jsou fuzzy. Pouze zřídka máme na mysli přesná čísla. Z matematického hlediska jsou fuzzy čísla konvexní fuzzy množiny se spojitou funkcí příslušnosti a jednoprvkovým jádrem. Speciálním a nejpoužívanějším typem fuzzy čísel jsou tzv. *trojúhelníková fuzzy čísla*. Není to nic jiného, než fuzzy čísla, jejichž funkce příslušnosti má tvar trojúhelníka, tj. obecný popis následující.

$$Z(x, a, b, c) = \begin{cases} 0, & x < a, \text{ nebo } x > c \\ \frac{x-a}{b-a}, & a \leq x \leq b, \\ \frac{c-x}{c-b}, & b \leq x \leq c, \\ 1, & x = b, \end{cases}$$

kde  $a, b, c$  jsou parametry znázorněné na Obr. 2.

Zpravidla jsou parametry  $a, c$  umístěny symetricky kolem hodnoty  $b$ . To znamená, že funkce příslušnosti tvoří rovnoramenný trojúhelník. Výjimečná nejsou ani fuzzy čísla uspořádaná do tvaru lichoběžníka (viz Obr. 3). V teorii fuzzy množin je rozvinuta základní aritmetika s fuzzy čísly (sčítání, odečítání, násobení a dělení) a rovněž některé složitější operace.

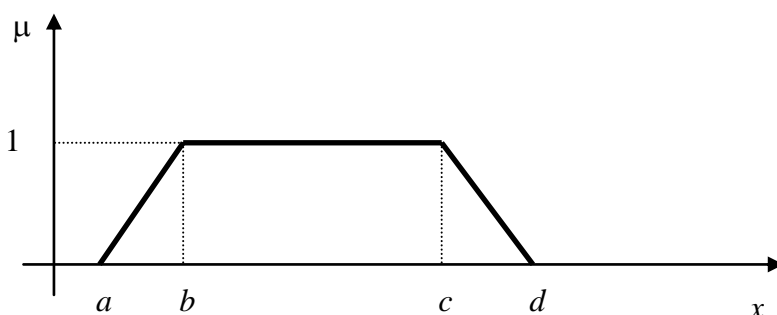


**Obr. 2:** Funkce příslušnosti trojúhelníkového fuzzy čísla

### 3 PŘÍKLADY VYUŽITÍ A APLIKACE NĚKTERÝCH FUZZY ALGORITMŮ

Jak již bylo řečeno výše, využití fuzzy logiky a množin je možné v mnoha oblastech (rozpoznání obrazu, klasifikace, rozhodování, aj.) a v mnohých jiných. Kromě jiných našla tato teorie své místo i v oblasti analýz spolehlivosti, které se podobnými aspekty klasifikace, hodnocení, rozhodování a mnohé jiné řídí poměrně často a v mnoha aplikacích. Dále jsou uvedeny příklady využití fuzzy logiky při klasifikaci závažností důsledků poruch. Jedná se o jednu z klíčových etap různých analýz, která je ale tímto zatížena vysokou neurčitostí (danou především pozicí hodnotitele – lidský faktor). Fuzzy logika napomáhá verifikovat provedené analýzy a výsledky jsou tedy poměrně validní v porovnání s jinými postupy. Jako příklad je zde uvedena klasifikace závažností důsledků poruch u vozidlových skupin.

Závažnosti důsledků poruch jednotlivých skupin budou touto metodou přiřazeny do fuzzy množin. Zde se omezuje na to, že jednotlivé fuzzy podmnožiny jsou sestaveny z koeficientů závažnosti důsledků poruchy. Kritéria výpočtu těchto koeficientů jsou popsána níže. Pro sestavení fuzzy podmnožin používáme metodu „Fuzzifikace veličin“. Konkrétní pozorované hodnoty fyzikálních veličin jsou ohraničené a vyjadřujeme je pomocí reálných čísel. Proto jako universum fuzzy čísel reprezentujících vágní pojmy souvisejících s klasifikací důsledků poruchy postačí vhodný uzavřený interval pro každou z nich. Pro praktické aplikace a také názorné grafické vyjádření plně vyhovují fuzzy lichoběžníková čísla, viz Obr. 30, kde  $\mu$  vyjadřuje funkci příslušnosti a  $x$  zvažované fuzzy číslo.



**Obr. 3:** Příklad lichoběžníkového fuzzy čísla

Abychom stanovili konkrétní funkci příslušnosti pro fuzzifikovanou hodnotu zvolené fyzikální veličiny, stačí zjistit, v jakém intervalu se tato hodnota obvykle vyskytuje. Tento interval je pak jádrem hledaného fuzzy čísla a značíme ho  $\langle b, c \rangle$ . Pro náš případ je toto jádro vždy vyjádřeno okrajovými hodnotami pro jednotlivé závažnosti důsledků poruch na skupině (viz níže). Dále zjistíme, jakých hodnot veličina zaručeně nenabývá. Množinu těchto hodnot předpokládáme ve

tvaru  $(-\infty; a) \cup (d; \infty)$ , přičemž  $a < b \leq c < d$ . Potom interval  $\langle a; d \rangle$  je support-množina „A“ hledaného fuzzy čísla.

Funkci příslušnosti hledaného fuzzy čísla do množiny „A“ pak vyjádříme ve tvaru:

$$\mu_A(x, a, b, c, d) = \max\left(\min\left(\frac{x-a}{b-a}, \frac{x-d}{c-d}, 1\right), 0\right) \approx \begin{cases} 1, & b \leq x \leq c \\ 0, & x < a \text{ nebo } x > d \\ \frac{x-a}{b-a}, & a \leq x \leq b \\ \frac{x-c}{c-d}, & c \leq x \leq d \end{cases}$$

### 3.1 Určení koeficientu důležitosti a kategorizace závažnosti důsledku poruchy

Velikost výsledného koeficientu důležitosti posuzované skupiny či soustavy je dána skalárním součinem dílčích koeficientů důležitosti uvedených v následujících kategoriích pro funkci, bezpečnost a náklady spojené s obnovou. Tyto kategorie jsou záměrně voleny tak, aby pojímaly maximální možný rozsah důsledků poruchy na významné oblasti související s vozidlem (v tomto případě s bojovým vozidlem). Výsledný koeficient důležitosti „D“ je dán vztahem:

$$D = D_1 \cdot D_2 \cdot D_3 \quad (7)$$

kde:  $D_1$  – koeficient důležitosti v závislosti na závažnosti důsledku poruchy u kategorie – funkce, (viz níže)  
 $D_2$  – koeficient důležitosti v závislosti na závažnosti důsledku poruchy u kategorie – bezpečnost, (viz níže)  
 $D_3$  – koeficient důležitosti v závislosti na závažnosti důsledku poruchy u kategorie – náklady spojené s obnovou, (viz níže)

Výsledný koeficient důležitosti může nabývat mezních hodnot  $D_{\min} = 1$ ;  $D_{\max} = 64$ .

1. FUNKCE – kategorizace závažnosti důsledku poruchy objektu z hlediska schopnosti plnit požadovanou funkci.

Kategorie	Definice	Koeficient důležitosti „D <sub>1</sub> “
1	Objekt je i po vzniku poruchy schopen funkce bez změněných vlastností. Důsledek takové poruchy nijak neovlivňuje funkční předpoklad pro úspěšné splnění „mission“.	1
2	Důsledek takové poruchy částečně omezí splnění bojového úkolu, přičemž tuto funkci nemůže zabezpečit žádná jiná entita (skupina). Důsledek je však zvládnutelný osádkou.	2
3	Důsledek takovéto poruchy významně omezí splnění bojového úkolu, přičemž tento důsledek není možné zvládnout silami osádky. Při pokračování důsledku takové poruchy v čase může tento vést k nesplnění bojového úkolu.	3
4	Důsledkem takové poruchy je okamžité zamezení schopnosti splnit bojový úkol. Vozidlo s tímto důsledkem už není schopno bojový úkol splnit.	4

2. BEZPEČNOST – kategorizace závažnosti důsledku poruchy objektu ve vztahu k bezpečnosti. Tato bezpečnost souvisí se „Schopností splnit bojový úkol“.

Kategorie	Definice	Koeficient důležitosti „ $D_2$ “
1	Důsledek takové poruchy nemá žádný vliv na bezpečnost vozidla, osádky i okolí.	1
2	Důsledkem takové poruchy je snížení bezpečnosti skupiny (entity) nebo vozidla jak pro osádku, tak pro okolí.	2
3	Důsledkem takové poruchy je stav, kdy se vozidlo, skupiny (entita) stávají nebezpečné jak pro osádku, tak pro okolí.	3
4	Důsledkem takové poruchy na skupině (entitě) je ztráta na lidských životech nebo velmi závažná hmotná škoda.	4

3. NÁKLADY SPOJENÉ S OBNOVOU – kategorizace závažnosti důsledku poruchy objektu ve vztahu k nákladům spojených s obnovou.

Kategorie	Definice	Koeficient důležitosti „ $D_3$ “
1	Důsledek takovéto poruchy na entitě nebude vyžadovat vyšší náklady spojené s obnovou než jmenovité pořizovací náklady porouchané entity.	1
2	Důsledek takovéto poruchy vzniklé na entitě (skupině) bude vyžadovat náklady složené z jmenovitých pořizovacích nákladů entity (skupiny) a nákladů vynaložených na provedení obnovy.	2
3	Důsledek takové poruchy entity (skupiny) bude vyžadovat náklady na obnovu i u jiných entit, které v součtu převýší jmenovité pořizovací náklady na porouchanou entitu (skupinu) plus náklady vynaložené na provedení obnovy.	3
4	Důsledek takové poruchy entity (skupiny) bude vyžadovat náklady na obnovu i u jiných entit, přičemž součet těchto nákladů vícenásobně převýší jmenovité pořizovací náklady na porouchanou entitu (skupinu) plus náklady vynaložené na provedení obnovy.	4

Podle těchto zmíněných kritérií je posuzována závažnost důsledků vzniklých poruch na jednotlivých skupinách nebo prvcích ve vozidle. V následujícím hodnocení se soustředíme pouze na ty skupiny a prvky, které přímo ovlivňují požadavky na zabezpečení taktické a operační mobility z hlediska nároků „Schopnosti splnit bojový úkol“. Pro toto hodnocení používáme podobný nástroj, který byl uveden výše. Toto hodnocení se liší v matematickém modelu fuzzy množin, který je v tomto případě rozčleněn do čtyř skupin fuzzy podmnožin z hlediska důležitosti (postradatelnosti) jednotlivých skupin, resp. prvků, ve vztahu ke „Schopnosti splnit bojový úkol“. Těmto pozmeněným fuzzy podmnožinám přísluší také pozmeněné funkce příslušnosti k jednotlivým fuzzy množinám, které odvodíme níže. Jak již bylo zmíněno, uvádíme čtyři fuzzy podmnožiny. Jejich jádra tvoří okrajové hodnoty závažností důsledků přiřazené k jednotlivým kategoriím tak, jako je tomu v hodnotícím aparátu.

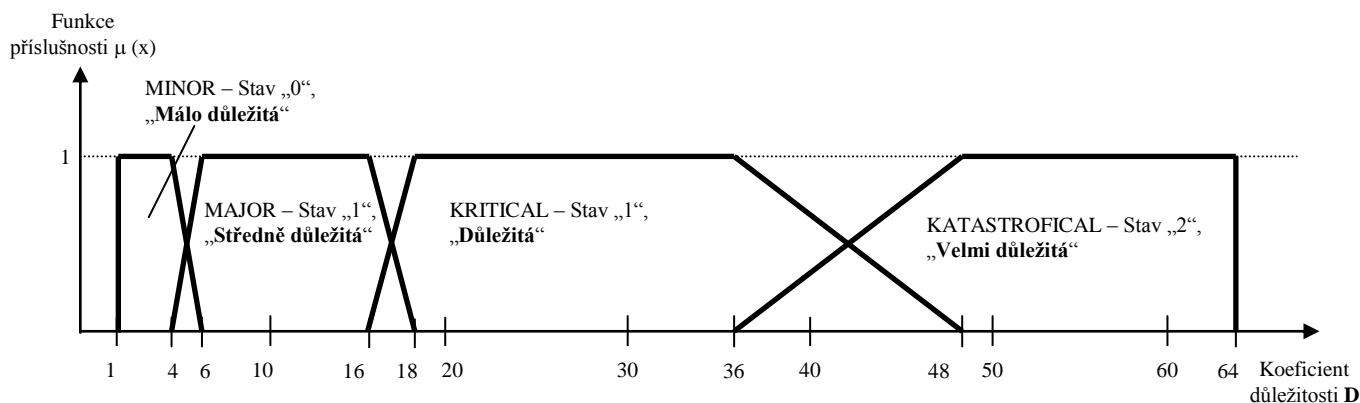
Nejdřív provedeme hodnocení těchto skupin a vyčleníme ty, které jsou zařazeny do kategorií nebo fuzzy množin, abychom zdůraznili jejich důležitost. S ohledem na dříve zmíněné definice a použití teorie Fuzzy množin, byly ke kategorizačním poruch přiřazeny i další jazykové náhodné proměnné. Podoba jednotlivých kategorií a fuzzy množin má následující formu:

**MINOR:** – důsledek nesníží ani neovlivní funkční schopnosti, efektivnost a výkon pod limitní hodnoty; nemá žádný vliv na úspěšné splnění bojového úkolu. Tato kategorie odpovídá stavu „0“ a fuzzy množině „**málo důležitá (MD)**“. Jestliže výsledný koeficient důležitosti nabude hodnot z intervalu  $\langle 1;4 \rangle$ , posuzovaná skupina, resp. objekt s velikostí funkce příslušnosti **1** bude patřit do této fuzzy množiny, viz 0.

**MAJOR:** – důsledek by mohl snížit funkční schopnosti, efektivnost a výkon pod přijatelnou limitní hodnotu, ale stále je zvládnutelný obsluhou. Je tedy možné plnit bojový úkol, ale s mírnými obtížemi a ne v plném rozsahu. Tato kategorie odpovídá stavu „1“ a fuzzy množině „**středně důležitá (SD)**“. Jestliže výsledný koeficient důležitosti nabude hodnot z intervalu  $\langle 6;16 \rangle$ , posuzovaná skupina, resp. objekt s velikostí funkce příslušnosti **1** bude patřit do této fuzzy množiny, viz 0.

**CRITICAL:** – tento důsledek může snížit funkční schopnosti a může vést ke katastrofické poruše, nejsou-li včas přijata nápravná opatření. V této situaci lze ještě plnit bojový úkol, avšak na pokraji všech možných funkcí a ne v plném rozsahu požadavků. Tato kategorie ještě odpovídá stavu „1“, ale fuzzy množině „**důležitá (D)**“. Jestliže výsledný koeficient důležitosti nabude hodnot z intervalu  $\langle 18;36 \rangle$ , posuzovaná skupina, resp. objekt s velikostí funkce příslušnosti **1**, bude patřit do této fuzzy množiny, viz 0.

**KATASTROFIC:** – tento důsledek může způsobit vážné poškození objektu do té míry, že může dojít k ukončení funkce nebezpečným způsobem a může zapříčinit zranění lidí nebo způsobit velké materiální škody. V této situaci vůbec nelze plnit bojový úkol. Tato kategorie odpovídá stavu „2“ a fuzzy množině „**velmi důležitá (VD)**“. Jestliže výsledný koeficient důležitosti nabude hodnot z intervalu  $\langle 48;64 \rangle$ , posuzovaná skupina, resp. objekt s velikostí funkce příslušnosti **1** bude patřit do této fuzzy množiny, viz Obr. 4.



**Obr. 4:** Fuzzy množiny závažnosti důsledků poruch a důležitosti jednotlivých funkcí ve vztahu k mobilitě

Tímto způsobem je možné klasifikovat závažnosti důsledků poruch u vozidla. Stanovení hodnotících škál není dogma, ačkoliv v tomto případě si autor vzal jako inspiraci čtyřstupňovou škálu z normy IEC 60 812. Existují ale i jiné normy (např. SAE J1739) kde je možné nalézt i jiné hodnotící škály se svými jazykovými popisy.

Dalším příkladem aplikace fuzzy logiky a množin je nejenom v oblasti spolehlivosti, resp., bezpečnosti v případě sledování např. dlouhodobě stálých a požadovaných veličin z hlediska spolehlivosti. Například verifikace a validace signálu o sledované náhodné veličině (měřicí obvody elektráren, zpracovatelské provozy, energetické soustavy, prvky kritické infrastruktury, apod.).

Výchozí stav:

Při zpracování elektrických signálů můžeme nalézt velké množství různých přístupů. Každý přístup má své výhody i nevýhody a většinou je upřednostňován zadavatelem, resp. uživatelem měřicího obvodu. V rámci aplikací v energetických zařízeních můžeme také nalézt širokou škálu měřících modifikací, které umožňují pracovat se získaným signálem v dalších fázích procesu různě.

V případě verifikace a validace signálů se jedná především o násobné snímání sledovaných fyzikálních veličin, tzv. elektrické měření neelektrických veličin. Princip verifikace a především validace spočívá především ve výběru vhodných (pravdivých a relevantních) signálů, které jsou pro další fáze procesu významné. V první fázi potřebujeme mít pravdivý/platný obraz o skutečném stavu (potřebujeme mít validní údaj) a ve druhé fázi jej potřebujeme libovolně mnohokrát (v různém čase, podmínkách, sledech) ověřit (verifikovat). Pro výběr toho pravého signálu potřebujeme využít vhodný algoritmus, který zajistí s určitou mírou jistoty, určitosti a konfidence správnou volbu. Některé algoritmy pracují v základních principech logiky za využití určitého náhodného počtu výběru  $m$  prvků z  $n$ . My pro další postup navrhuje jiný přístup.

Na základě provedené analýzy stávajících způsobů zpracování signálů a informací o měřené veličině z hlediska spolehlivosti této informace, jsme dospěli k názoru, že poměr míry nejistoty a neurčitosti o měřené veličině jednoznačně evokuje fuzzy přístup k danému problému.

Navrhovaný postup řešení:

Na základě provedené analýzy, viz, výše jsme se rozhodli pro využití fuzzy logiky při rozhodnutí o validním signálu měřené fyzikální veličiny. Jelikož se jedná o disciplínu, která umožňuje zahrnout nepřesnost a navíc dovoluje pracovat i z nepřesnostmi přirozeného jazyka, který je v různých aplikacích hojně užíván při kvalifikaci procesů. Podstatou fuzzy logiky je princip využití vágně definovaných expertních znalostí a potřeb. Je tedy pravým opakem standardních potřeb o „větší přesnosti“ a „spolehlivosti“ výstupních údajů. Pokud bychom se zabývali podrobněji podstatou fuzzy logiky (což není úmyslem v obsahu této zprávy) došli bychom k mnohým zajímavým souvislostem, které nám jednoznačně potvrzují naše rozhodnutí o vhodnosti jejich použití.

Jednou z disciplín, kde je možné fuzzy logiku aplikovat je výběr vhodných reprezentantů v určitém počtu prvků nějaké množiny. Ve správné terminologii fuzzy logiky se tento případ nazývá fuzzy clustering, neboli fuzzy shluková analýza.

Tato analýza nám umožňuje nalézt shluky v nějakém smyslu podobných dat. Tento problém se vyskytuje velmi často (viz. případ naší aplikace), a to zejména v metodách rozpoznávání obrazů, datových množin, shluků měřených signálů, apod. V klasickém pojetí totiž pracujeme se shluky, jako s disjunktními podmnožinami nějaké množiny bodů. Základním pojmem ve shlukové analýze je *vzdálenost* mezi objekty. V našem případě jejich příbuznost. Může se jednat o

absolutní hodnotu rozdílu, euklidovskou vzdálenost, nebo jinou libovolnou funkci  $d$ , přičemž platí že  $V \times V \rightarrow R$ . Základními vlastnostmi funkce jsou tyto axiomy:

$$d(x,x) = 0; \quad (8)$$

$$d(x,y) = d(y,x); \quad (9)$$

$$d(x,z) \leq d(x,y) + d(y,z). \quad (10)$$

pro všechna  $x, y, z \in V$ . Podmínka (10) se nazývá tzv. *trojúhelníková nerovnost*. Právě způsobem definice vzdálenosti se od sebe liší různé metody shlukové analýzy.

Matematickými vyjádřeními jednotlivých postupů při hledání správných shluků bychom se zabývali později. V současném stavu můžeme volit, na základě získaných informací, ze dvou základních případů a principů provedení fuzzy shlukové analýzy. První z nich je tzv. „Algoritmus fuzzy c-means“ a druhým z nich je „Gustavson-Kesselův algoritmus“. Zatímco na základě získaných údajů můžeme pomocí algoritmu „c-means“ zpracovat pouze data, které jsou z algebraického hlediska sloučena do kruhových množin, v případě „Gustavson-Kesselova algoritmu“ lze zpracovat i data jiných tvarů množin. Oba způsoby umožňují pracovat s jinými typy funkcí příslušnosti a také, a to je podstatné, různé příbuznosti jednotlivých dat.

Výsledkem při použití kteréhokoliv algoritmu je potom sestavení fuzzy pravidel pro rozhodnutí o akceptování nebo zamítnutí významnosti získaného shluku, resp. středu, který reprezentuje měřenou hodnotu signálu. Všechny tyto principy nám mohou dále umožnit aplikovat tzv. fuzzy regulátory, tedy elementy, jež umožňují rozhodnout a použít určitou bázi znalostí ve prospěch rozhodovacího procesu.

Další postup:

V dalším postupu se budeme zabývat, na základě dokončené analýzy o zpracování signálu, zcela konkrétní možnou cestou vhodnou pro aplikaci fuzzy algoritmu do podmínek zpracování signálů o měřené veličině.

Základním pojmem ve shlukové analýze je *vzdálenost* mezi objekty. Může to být absolutní hodnota rozdílu, euklidovská vzdálenost, nebo libovolná jiná funkce  $d: V \times V \rightarrow R$ , která má vlastnosti uvedené ve vztazích (8), (9) a (10):

Z čistě matematického hlediska je fuzzy shluková analýza hledáním vhodného rozkladu množiny  $V$ , tj. hledáním takového systému podmnožin (*shluků*)

$$P_c = \{R_j | j = 1, \dots, c, R_j \subset V\}$$

$$\bigcup_{j=1}^c R_j = V$$

$$R_i \cap R_j = \emptyset, \quad i \neq j, \quad 1 \leq i, j \leq c$$

a který zároveň splňuje požadavek, aby prvky jednoho shluku  $R_j$ ,  $j = 1, \dots, c$  byly vzájemně „blíže“, než prvky z různých jiných shluků. Obvykle se tohoto dosahuje tím, že se definují jistý funkcionář  $J(P_c)$ , který se postupně minimalizuje.

Ve fuzzy shlukové analýze jsou  $R_j \subseteq V$  fuzzy množiny. Navíc musí splňovat následující podmínky (hovoříme zde o fuzzy pseudorozkladu): nechť  $V = \{v_1, \dots, v_r\}$ . Pak

$$\sum_{j=1}^c R_j(v_i) = 1, \quad i = 1, \dots, r \quad (11)$$



$$0 < \sum_{i=1}^r R_j(v_i) < r, \quad j = 1, \dots, c. \quad (12)$$

První podmínka nám říká, že každý prvek  $v_i \in V$  musí patřit alespoň do jednoho shluku, přičemž součet všech stupňů jeho příslušnosti do všech shluků musí být roven 1.

Druhá podmínka nám říká, že do žádného shluku nesmí v maximálním stupni příslušnosti patřit všechny prvky množiny  $V$  a každý shluk musí být neprázdný.

### Příklad 1:

Nechť  $V = \{v_1, \dots, v_5\}$ . Pak příkladem fuzzy pseudorozkladu je  $P_3 = \{R_1, R_2, R_3\}$ , kde

$$\begin{aligned} R_1 &= \{0,3/v_1; 0,5/v_2; 0,7/v_3; 1/v_4\} \\ R_2 &= \{0,2/v_1; 0,4/v_2; 0,1/v_3; 0,9/v_5\} \\ R_3 &= \{0,5/v_1; 0,1/v_2; 0,2/v_3; 0,1/v_5\} \end{aligned}$$

Data pro fuzzy shlukovou analýzu jsou reálná čísla v tomto tvaru

$$\begin{matrix} & \varphi_1 & \varphi_2 & \dots & \dots & \dots & \varphi_n \\ \begin{matrix} o_1 \\ o_2 \\ \vdots \\ o_r \end{matrix} & \begin{bmatrix} u_{11} & u_{12} & \dots & u_{1n} \\ u_{21} & u_{22} & \dots & u_{2n} \\ \vdots & \vdots & & \vdots \\ u_{r1} & u_{r2} & \dots & u_{rn} \end{bmatrix} \end{matrix} \quad (13)$$

V matici (13) jsou symboly v pozicích  $o_1, \dots, o_r$  objekty na nichž jsou měřeny charakteristiky  $\varphi_1, \dots, \varphi_n$ . Objekty mohou být libovolné předměty, obrazové pixely, časové okamžiky a podobně.

Měřené charakteristiky mohou být naproti tomu jednak fyzikální veličiny (hmotnost, velikost, délka, intenzita, apod.) a jednak také veličiny abstraktní (jako např. zisk, nějaký stupeň na škále, apod.).

Data lze charakterizovat jako  $r$  bodů v  $n$ -rozměrném prostoru, kde prvky  $u_{i1}, \dots, u_{in}, i = 1, \dots, r$  jsou souřadnice těchto bodů. Každý bod je tedy v datech (13) reprezentován jedním řádkem. Pro další výklad označíme tyto řádky symboly  $\mathbf{u}_i, i = 1, \dots, r$  a budeme je chápat jako řádkové vektory

$$\mathbf{u}_i = (u_{i1}, \dots, u_{in})$$

Je-li  $H$  matice typu  $(m;n)$ , pak vektor  $H^T$  označuje její transpozici, tj. matici typu  $(n;m)$ . Pak  $\mathbf{u}_i^T$  je sloupcový vektor.

### Použitelné algoritmy fuzzy shlukování

Základní a patrně nejpoužívanější fuzzy shlukovací metodou je fuzzy *c-means* (fuzzy *c-průměrů*) kde  $c$  označuje počet shluků. Vznikla jako zobecnění známé metody ISODATA, přičemž ostatní metody jsou od této metody nějakým způsobem odvozeny.

Metody fuzzy shlukování vycházejí z takového základního předpokladu, že je zadán počet shluků, které mají být nalezeny. Předpokládejme, že už jsou shluky  $P_c = \{R_1, \dots, R_c\}$  známy. Pro každý shluk se vypočte jeho *centrum*. U fuzzy *c-means* se k tomu použije vztah:

$$s_j = \frac{\sum_{i=1}^r (R_j(u_i))^m u_i}{\sum_{i=1}^r (R_j(u_i))^m}, j = 1, \dots, c. \quad (14)$$

kde  $m > 1$  je předem zvolený parametr. Výrazu (14) je nutné rozumět tak, že souřadnice centra  $s_j = \langle s_{j1}, \dots, s_{jn} \rangle$  jsou pomocí (14) počítány po jednotlivých složkách.

Optimální řešení fuzzy shlukové analýzy je dosaženo, je-li minimalizován funkcionář  $J_m(P_c)$ , který u fuzzy c-means má tvar

$$J_m(P_c) = \sum_{i=1}^r \sum_{j=1}^c (R_j(u_i))^m d^2(u_i, v_j), \quad (15)$$

kde  $d$  je nějaká předem definovaná vzdálenost.

### Algoritmus fuzzy c-means

Vlastní algoritmus počítá funkce příslušnosti fuzzy shluků iterativně, tedy dokud není dosažena požadovaná přesnost.

1. Zvolíme parametr  $m$ , přesnost  $\varepsilon$  a zadáme počáteční fuzzy pseudorozklad  $P_c^{(0)}$ .
2. Vypočteme centra  $s_1^{(1)}, \dots, s_c^{(1)}$  pomocí (14).
3. Pro každé  $i = 1, \dots, r$  upravíme fuzzy pseudorozklad takto:
  - a) Necht' pro všechna  $j = 1, \dots, c$  je  $d^2(u_i, s_j) > 0$ . Pak

$$R_j^{(l+1)}(u_i) = \left[ \sum_{k=1}^c \left( \frac{d(u_i, s_j^{(l)})}{d(u_i, s_k^{(l)})} \right)^{\frac{2}{m-1}} \right]^{-1}, \quad (16)$$

- b) Necht' pro některá  $j, j = j_1, \dots, j_p, p \leq c$ , je  $d(u_i, s_j) = 0$ . Pak pro tato  $j$  položíme  $R_j^{(l+1)}(u_i)$  rovna libovolným nezáporným číslům tak, aby

$$\sum_{j=j_1}^{j_p} R_j^{(l+1)}(u_i) = 1, \quad (17)$$

a pro ostatní  $j$  položíme  $R_j^{(l+1)}(u_i) = 0$ .

4. Srovnáme fuzzy pseudorozklady  $P_c^{(l+1)}$  a  $P_c^{(l)}$  prostřednictvím vzdálenosti

$$\|P_c^{(l+1)}, P_c^{(l)}\| = \max_{\substack{j=1, \dots, c \\ i=1, \dots, r}} |R_j^{(l+1)}(u_i) - R_j^{(l)}(u_i)|. \quad (18)$$

Následující příklad ilustruje použití fuzzy c-means shlukování.

**Příklad 2:**

Předpokládejme, že máme 10 objektů, u kterých měříme 2 charakteristiky, tj.  $r = 10$  a  $n = 2$ . Tato data jsou výsledkem experimentálního měření nějaké funkční závislosti  $y = f(x)$  v deseti bodech. Potom data dle vztahu (12) mají následující tvar:

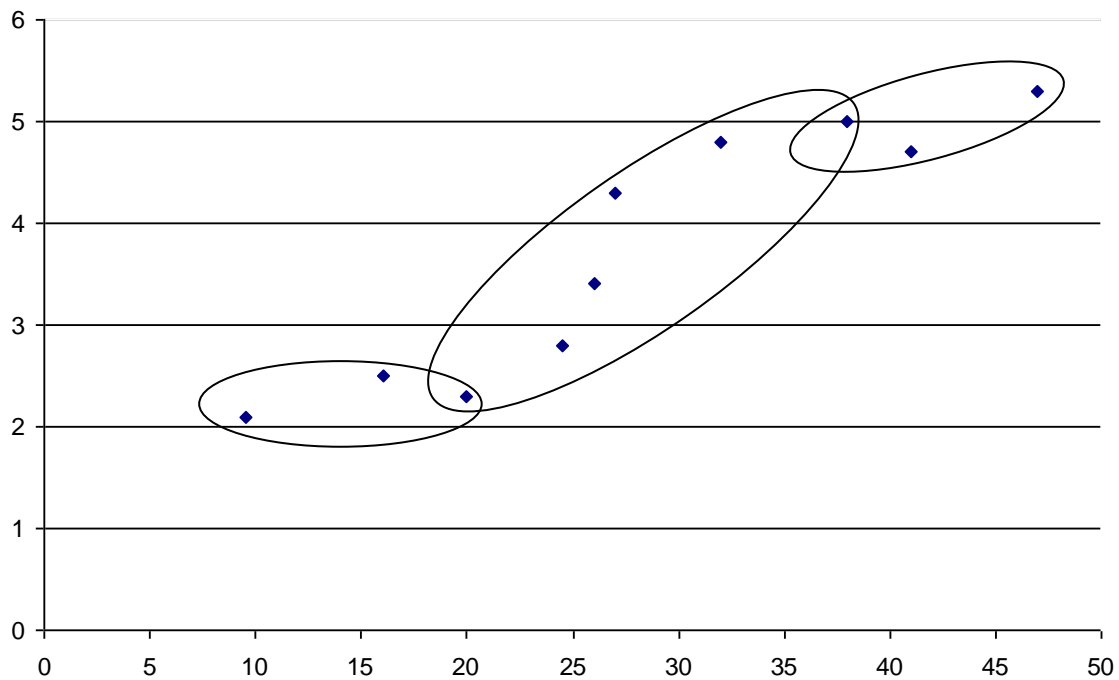
$$R = \begin{matrix} & \begin{matrix} u_i & v_i \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \\ 10 \end{matrix} & \begin{bmatrix} 9,5 & 2,1 \\ 16,0 & 2,5 \\ 20,0 & 2,3 \\ 24,5 & 2,8 \\ 26,0 & 3,4 \\ 27,0 & 4,3 \\ 32,0 & 4,8 \\ 38,0 & 5,0 \\ 41,0 & 4,7 \\ 47,0 & 5,3 \end{bmatrix} \end{matrix} \quad (19)$$

Data lze chápat jako 10 vektorů tvaru  $(u_i, v_i)$ ,  $i = 1, \dots, 10$  a jsou znázorněna graficky na Obr. 5. Z obrázku vidíme, že data lze rozdělit do tří oblastí – levá rovnější část, střední rostoucí a pravá opět rovnější. Z tohoto důvodu budeme hledat tři shluky ( $c = 3$ ). Vzdálenost mezi objekty a centry bude euklidovská, tj.

$$d(u_i, s_j^{(l)}) = \sqrt{(u_i - s_{j1})^2 + (v_i - s_{j2})^2},$$

kde  $s_{j1}$ ,  $s_{j2}$  jsou souřadnice  $j$ -tého shluku. Dále zvolíme parametr  $m = 2$  (obvyklá hodnota), přesnost  $\varepsilon = 0,05$  a počáteční fuzzy pseudorozklad na tři shluky:

	1	2	3	4	5	6	7	8	9	10
$R_1$	0,6	0,6	0,6	0,3	0,3	0,3	0,3	0,3	0,3	0,3
$R_2$	0,3	0,3	0,3	0,6	0,6	0,5	0,5	0,3	0,3	0,3
$R_3$	0,1	0,1	0,1	0,1	0,1	0,2	0,2	0,4	0,4	0,4



**Obr. 5:** Data pro fuzzy shlukovou analýzu

Vypočtená centra shluků v nulté iteraci jsou podle (14) tato:

$$s_1 = (22; 3), \quad s_2 = (27,5; 3,6), \quad s_3 = (43,4; 5,1).$$

Po šesti iteracích již klesla chyba a tedy dostaneme následující centra shluků

$$s_1 = (14,2; 2,3), \quad s_2 = (26,5; 3,5), \quad s_3 = (42; 5).$$

Výsledkem jsou následující fuzzy shluky:

	1	2	3	4	5	6	7	8	9	10
R <sub>1</sub>	0,91	0,97	0,54	0,04	0	0	0,07	0,02	0	0,02
R <sub>2</sub>	0,07	0,03	0,42	0,94	1	1	0,66	0,11	0,01	0,05
R <sub>3</sub>	0,02	0	0,04	0,02	0	0	0,27	0,87	0,99	0,93

Vidíme, že se oblast vskutku rozdělila na tři poměrně zřetelné shluky. Se stupněm příslušnosti vyšším než 0,5 patří do prvního shluku body 1, 2, 3, do druhého body 3, 4, 5, 6, 7, a do třetího body 8, 9 a 10.

### Gustavson – Kesselův algoritmus

Algoritmus fuzzy c-means umožňuje detekovat shluky kruhových tvarů. Níže popsáný algoritmus je modifikací algoritmu fuzzy c-means. Jeho podstatnou výhodou je, že umožňuje najít shluky, které mají různý tvar a různou orientaci. Vlastní výpočet je však ale náročnější.

Shluky ve fuzzy c-means jsou charakterizovány pouze jejich centrem. V Gustavson-Kesselově algoritmu je každý shluk navíc charakterizován speciální (pozitivně definitní) maticí  $H_j$ ,  $j = 1, \dots, c$ . Význam prvků této matice je v deformaci souřadnic, což umožňuje specifikovat jiný tvar shluků, než kružnice. V principu použití algoritmu se k výpočtu matice  $H_j$  používá tzv. *kovarianční matice*.

$$S_j = \frac{\sum_{i=1}^r (R_j^{(l)}(u_i))^m (u_i - s_j^{(l)})^T (u_i - s_j^{(l)})}{\sum_{i=1}^r (R_j^{(l)}(u_i))^m}, \quad j = 1, \dots, c \quad (20)$$

Poté definujeme

$$H_j = (\det(S_j))^{-\frac{1}{n}} S_j^{-1}. \quad (21)$$

Pomocí matice  $H_j$  je definována norma  $\|u\|_{H_j} = \sqrt{u H_j u^T}$  a dále se definuje vzdálenost

$$d_{H_j}(u, v) = \|u - v\|_{H_j}$$

Kromě toho se také definuje omezení, aby determinant matice  $H$  byl  $\det(H_j) = \rho$ , kde  $\rho$  je nějaké číslo určené pro každou matici. Při interpretaci to znamená, že je specifikován konstantní objem (velikost) shluků, avšak nikoliv jejich tvar. Nevýhodou je nutnost mít nějakou apriorní informaci o shlucích.

Vlastní algoritmus je opět iterativní a končí při dosažení přesnosti  $\varepsilon$ . Postup provedený je uveden následně:

1. Zvolíme parametr  $m$ , přesnost  $\varepsilon$  a zadáme počáteční fuzzy pseudorozklad  $P_c^{(0)}$ .
2. Vypočteme centra  $s_1^{(l)}, \dots, s_c^{(l)}$  pomocí vztahu (7).
3. Vypočteme všechny kovarianční matice  $S_j, j = 1, \dots, c$  dle postupu (13).
4. Pro každé  $i = 1, \dots, r$  vypočteme vzdálenosti

$$d_H(u_i, s_j^{(l)}) = \sqrt{(u_i - s_j^{(l)}) \det(S_j)^{-\frac{1}{n}} S_j^{-1} (u_i - s_j^{(l)})^T}, \quad j = 1, \dots, c. \quad (22)$$

5. Pro každé  $i = 1, \dots, r$  upravíme fuzzy pseudorozklad takto:
  - a) Nechť pro všechna  $j = 1, \dots, c$  je  $d_H(u_i; s_j) > 0$ . Pak

$$R_j^{(l+1)}(u_i) = \left[ \sum_{k=1}^c \left( \frac{d_H(u_i, s_j^{(l)})}{d_H(u_i, s_k^{(l)})} \right)^{\frac{2}{m-1}} \right]^{-1}. \quad (23)$$

- b) Nechť pro některá  $j, j = j_1, \dots, j_p, p \leq c$ , je  $d_H(u_i; s_j) = 0$ . Pak pro tato  $j$  položíme  $R_j^{(l+1)}(u_i)$  rovna libovolným nezáporným číslům tak, aby

$$\sum_{j=j_1}^{j_p} R_j^{(l+1)}(u_i) = 1 \quad (24)$$

a pro ostatní  $j$  položíme  $R_j^{(l+1)}(u_i) = 0$ .

6. Srovnáme fuzzy pseudorozklady  $P_c^{(l+1)}$  a  $P_c^{(l)}$  prostřednictvím vzdálenosti

$$\|P_c^{(l+1)}, P_c^{(l)}\| = \max_{\substack{j=1, \dots, c \\ i=1, \dots, r}} |R_j^{(l+1)}(u_i) - R_j^{(l)}(u_i)|. \quad (25)$$

Pokud je  $\|P_c^{(l+1)}, P_c^{(l)}\| \leq \varepsilon$ , proces ukončíme. V opačném případě zvýšíme  $l$  a vrátíme se k bodu 2.

### Určení počtu shluků

Nevýhodou výše uvedených metod je, že vyžadují znát předem počet  $c$  hledaných shluků. Pokud  $c$  nejsme s to předem stanovit nabízejí se nám dvě metody pro jeho odhad.

1. Sestrojíme nějakou globální charakteristiku získaného fuzzy pseudorozkladu  $G(P_c)$ . Vlastní algoritmus pak probíhá takto. Nejprve zvolíme nějakou maximální hranici  $c^{max}$ . Dále realizujeme shlukovací algoritmus postupně pro  $c = 1, \dots, c^{max}$  a získáme fuzzy pseudorozklad  $P_c$ . V každém kroku vypočteme charakteristiku  $G(P_c)$ . Nedošli-li už k jejímu zlepšení považujeme za nejlepší fuzzy pseudorozklad  $G(P_{c-1})$ . Existuje řada charakteristik  $G(P_c)$ . Některé se minimalizují, jiné se maximalizují. Problém je v tom, že neexistuje žádná optimální charakteristika, a proto je vhodné je dovedně kombinovat. Podrobný rozbor lze najít v některých literárních zdrojích. Příkladem charakteristiky  $G(P_c)$  je *koeficient rozkladu*

$$PC(P_c) = \frac{\sum_{i=1}^r \sum_{j=1}^c R_j(u_i)^2}{r}. \quad (26)$$

Tato charakteristika nabývá hodnot  $\frac{1}{c} \leq PC(P_c) \leq 1$ , přičemž vyšší hodnoty znamenají zlepšování.

2. Druhou možností je posuzovat jednotlivé shluky individuálně podle jejich „kvality“ opět posuzované pomocí nějaké lokální charakteristiky. V případě že jeden ze dvou sousedních shluků je výrazně horší než druhý, oba spojíme do jednoho. Protože tato analýza je složitější, do naší knihy ji už nezahrneme a odkazujeme na speciální literaturu.

### Sestavení fuzzy pravidel na základě nalezených shluků

Jak jsme se již zmínili v části výše lze fuzzy shlukovou analýzu použít pro nalezení fuzzy pravidel typu JESTLIŽE-PAK na základě zadaných dat za předpokladu, že jejich interpretace je funkcionální nebo jde o pravidla typu Takali-Sugeno.

Vraťme se ale nyní k předpokladu, kdy pracujeme v reálném světě:

$$v = \langle (V_1, h_1), \dots, (V_n, h_n) \rangle, \quad (27)$$

kde množiny  $V_k$  reprezentují jazykový kontext pro jednotlivé proměnné. Jsou to například intervaly

$$V_k = [{}^l v_k, {}^r v_k] \quad \text{kde } k = 1, \dots, n. \quad (28)$$

Přitom jazykový popis, jehož interpretaci se na základě shlukovací analýzy budeme snažit odvodit, je dán těmito vztahy:

$$R_1 := X_1 \text{ je } A_{11}A \cdots AX_{n-1} \text{ je } A_{1n-1}AY \text{ je } B_1,$$

.

.

.

$$R_c := X_1 \text{ je } A_{c1}A \cdots AX_{n-1} \text{ je } A_{cn-1}AY \text{ je } B_c.$$

Všimněte si, že počet pravidel je zde označen písmenem  $c$ . Z toho je hned zřejmé, že počet odvozených pravidel bude roven počtu nalezených fuzzy shluků.

Dále předpokládáme, že jsou dána data (viz. vztah (13)). Sloupce ve výrazu (13) odpovídají proměnným  $X_1, \dots, X_{n-1}, Y$  viz. (22). Zároveň tato data odpovídají hodnotám v možném světě  $v$  viz. vztah (20), tj.  $u_{1k}, \dots, u_{rk} \in V_k, k = 1, \dots, n$ . Protože nemáme k dispozici žádnou další informaci, můžeme rovnou předpokládat, že

$$l_{v_k} = \min\{u_{1k}, \dots, u_{rk}\},$$

$$r_{v_k} = \max\{u_{1k}, \dots, u_{rk}\}.$$

Nechť nyní jsou pomocí fuzzy shlukové analýzy vypočteny shluky  $R_1, \dots, R_c$ . Musíme si uvědomit, že každý fuzzy shluk, je vlastně fuzzy relací.

$$R_j(u_{i1}, \dots, u_{in}) = R_j(u_i), \quad i = 1, \dots, r$$

Pak fuzzy množiny  $A_{jk}$  a  $B_j, j = 1, \dots, c$  chápané jako extenze jazykových výrazů  $A_{jk} = \text{Ext}(A_{jk}), k = 1, \dots, n - 1, B_j = \text{Ext}(B_j)$  vyskytujících se v (28) se vypočtou jako projekce na  $k$ -tou proměnnou, tj.

$$A_{jk}(u_{ik}) = \underset{\substack{u_p, p=1, \dots, r \\ u_{pk}=u_{ik}}}{V} R_j(u_p) \quad (29)$$

$$B_j(u_{ik}) = \underset{u_p}{V} R_j(u_p) \quad (30)$$

Kde symbol  $\underset{\substack{u_p, p=1, \dots, r \\ u_{pk}=u_{ik}}}{V}$  znamená (maximum) uvažované přes všechny řády  $\mathbf{u}_p, p=1, \dots, r$  v datech

(13), které mají  $k$ -tou hodnotu  $u_{pk}$  rovnou hodnotě  $u_{ik}$ . Existují metody, jak získané fuzzy množiny  $A_{jk}, B_j$  vhodně aproximovat např. trojúhelníkovými, trapezoidními či jinými funkcemi, aby byly ve shodě s obecně přijatými tvary.

Pokud sestavujeme pravidla Takali-Sugeno, pak hledáme pouze fuzzy množiny  $A_{j1}, \dots, A_{jn-1}$  a namísto fuzzy množin  $B_j, j=1, \dots, c$  hledáme regresní konstanty  $a_j, b_{j1}, \dots, b_{jn-1}$ .

**Příklad 3:** Vyjdeme ze stejného zadání jako v příkladu 2. Budeme hledat tři pravidla typu Takali-Sugeno, které charakterizují zadanou funkční závislost. Protože data (19) byla naměřena při specifikaci funkce  $j$ , předpokládejme, že význam hledaných pravidel budeme definovat v modelu

$$\nu = \langle ([0;50], h_1), ([0,6], h_2) \rangle$$

Hranice intervalů  $[^l\nu_1; ^r\nu_1]$  a  $[^l\nu_2; ^r\nu_2]$  v  $\nu$  byly určeny jako dolní resp. horní meze možných hodnot argumentů  $x$  funkčních hodnot  $f(x)$  funkce  $f$ .

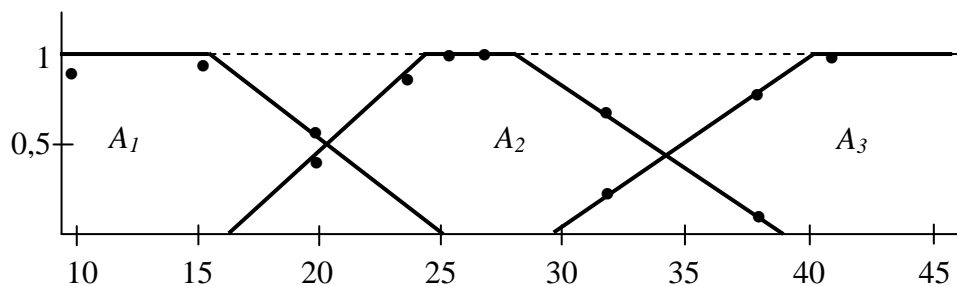
Na základě vypočtených funkcí příslušnosti shluků  $R_1, R_2, R_3$  odvodíme pomocí (29) funkce příslušnosti fuzzy množin  $A_1, A_2, A_3 \stackrel{\sim}{=} [^l\nu_1; ^r\nu_1]$ . Protože výpočty funkcí příslušnosti shluků ve fuzzy shlukové analýze jsou poměrně složité, nemá dobrý smysl dále pracovat s malými hodnotami, které jsou výsledkem. Proto budeme všechny stupně příslušnosti menší než 0,1 považovat za nulové. Pak dostaneme

$$A_1 = \{0,91/9,5; 0,97/16; 0,54/20\}, \quad (31)$$

$$A_2 = \{0,42/20; 0,94/24,5; 1/26; 1/27; 0,66/32; 0,11/38\}, \quad (32)$$

$$A_3 = \{0,27/32; 0,87/38; 0,99/41; 0,93/47\}. \quad (33)$$

Tyto fuzzy množiny budeme vhodně aproximovat trapezoidními funkcemi, tj. např. tak, jak je to znázorněno na Obr. 6.



**Obr. 6:** Aproximace fuzzy množin  $A_1, A_2, A_3$  vypočtených pomocí fuzzy shlukové analýzy (jsou zakresleny tučnými body) trapezoidními funkcemi příslušnosti.

Nakonec vypočteme regresní koeficienty podle známých vztahů

$$a_1 = 1,884; \quad b_1 = 0,029; \quad (34)$$

$$a_2 = -1,965; \quad b_2 = 0,207; \quad (35)$$

$$a_3 = 3,516; \quad b_3 = 0,035. \quad (36)$$

Dále již nebudeme hledat žádné jazykové výrazy pro nalezené fuzzy množiny a přiřadíme jim pouze symboly  $A_1, A_2$  a  $A_3$ . Pak dostaneme výsledná fuzzy pravidla v tomto tvaru:

$$R_1 := \text{JESTLIŽE } X \text{ je } A_1 \text{ PAK } Y = 1,884 + 0,029X$$

$$R_2 := \text{JESTLIŽE } X \text{ je } A_2 \text{ PAK } Y = -1,965 + 0,207X$$



$R_3 := \text{JESTLIŽE } X \text{ je } A_3 \text{ PAK } Y = 3,516 + 0,035X$

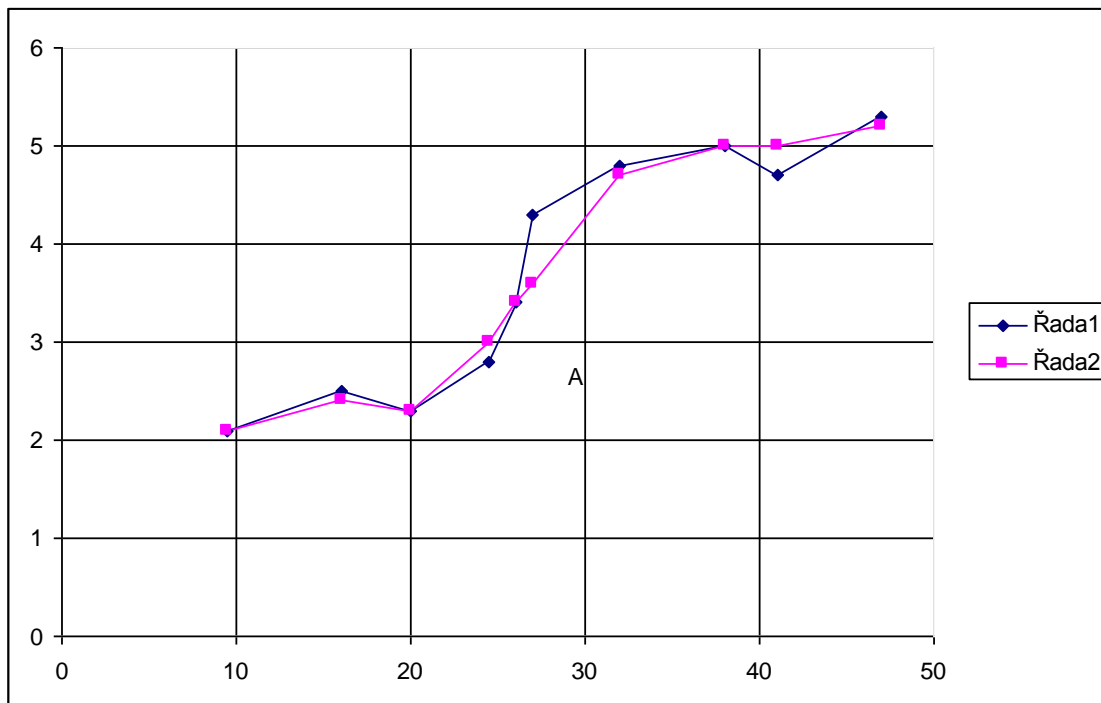
Nakonec podle vztahu pro výpočet přibližné dedukce realizované na základě nalezených pravidel

$$f^A(u) = \frac{\sum_{j=1}^m A_j(u) \cdot (a_j + b_j u)}{\sum_{j=1}^m A_j(u)} \quad (37)$$

vypočteme teoretické hodnoty  $v_i^{teor}$ :

$u_i$	$v_i^{teor}$
9,5	2,16
16,0	2,35
20,0	2,31
24,5	3,11
26,0	3,42
27,0	3,63
32,0	4,74
38,0	4,98
41,0	4,97
47,0	5,18

Celá situace je názorněji zobrazena na Obr. 7. Zde jsou uvedena původní data (pro přehlednost spojena úsečkami) a výsledek přibližné dedukce pomocí fuzzy pravidel typu Takagi-Sugeno vygenerovaných výše. Výsledná aproximující funkce je fuzzy model odvozený na základě dat.



**Obr. 7:** Fuzzy aproximace dat pomocí vygenerovaných Takagi-Sugeno pravidel (čtverečky jsou označením pro původní data a kosočtverečky výsledky aproximace).

Vidíme, že pomocí Takagi-Sugeno pravidel lze dostat dobré výsledky při hledání nejvýstižnějšího průběhu neznámé funkce. Protože obdobnou úlohu řeší také klasická regresní analýza, chápe se někdy právě popsany postup jako jedna z metod *fuzzy regresní analýzy*.

#### 4 ZÁVĚR

Tento příspěvek byl zpracován za podpory výzkumného záměru FVT 0000 401 a dále za podpory výzkumného centra „Progresivní technologie a systémy pro energetiku“, projekt MŠMTV č. 1M06059. Tato pasáž má napomoci při orientaci v problematice klasifikace poruch, verifikace deterministických metod analýzy spolehlivosti, bezpečnosti/rizika a zpracování a vyhodnocování informace o měřené veličině z hlediska spolehlivosti.

Zde je prezentováno několik přístupů a aplikačních postupů, které umožňují s využitím fuzzy logiky nalézt optimální hodnoty analyzovaných veličin. Vzhledem k míře vágnosti a neurčitosti, kterými jsou deterministické metody analýz a procesy měření zatíženy, byly zvoleny právě tyto přístupy.

#### 5 Použitá literatura

1. *Mezinárodní elektrotechnický slovník (International Electro-technical Vocabulary)*, Praha: Český normalizační institut 1993, ČSN IEC 50 (191).
2. NOVÁK, V.: *Fuzzy množiny a jejich aplikace (Fuzzy Sets and Their Application)*, Praha: Matematický seminář SNTL 1986.
3. NOVÁK, V.: *Základy fuzzy modelování (Fuzzy Modelling Foundations)*. Praha: BEN 2000.
4. VALIŠ, D.: *Analýzy důsledků aplikace vetroniky na dependability bojových vozidel (Vetronics Effects Analysis on Combat Vehicle Dependability)*, Brno: VA v Brně Application 2003, Disertační práce.
5. VALIŠ, D., FUCHS, P.: *Metody analýzy a řízení risk (Risk Management Analysis Methods)*, Liberec: TUL 2004.
6. VINTR, Z.: *Specifikace požadavků na bezporuchovost technických objektů (Technical Object Reliability Requirement Specification)*. Brno: VA v Brně 1998, inaugural dissertation.
7. VINTR, Z.: *Reliability and safety providing for railway applications*, In: “Trans & Motauto 2005”, Bulgaria, Veliko Turnovo.