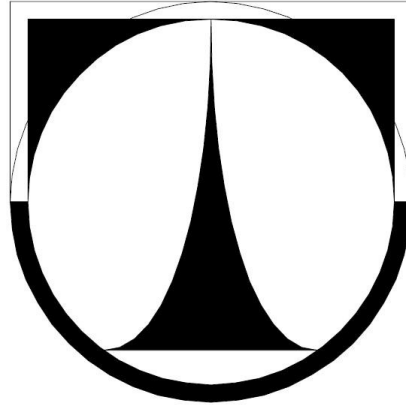


**Technická univerzita v Liberci**

Fakulta mechatroniky a mezioborových inženýrských studií



**Zjištění použitelnosti modifikovaného Weibullova  
rozdělení**

Usability of modified Weibull distribution.

Bakalářská práce

## PROHLÁŠENÍ

Byl jsem seznámen s tím, že na moji bakalářskou práci se plně vztahuje zákon č. 121/2000 o právu autorském, zejména § 60 (školní dílo).

Beru na vědomí, že TUL má právo na uzavření licenční smlouvy o užití mé bakalářské práce, a prohlašuji, že **s o u h l a s í m** s případným užitím mého bakalářského projektu (prodej, zapůjčení apod.).

Jsem si vědom toho, že užití svoji bakalářskou práci či poskytnout licenci k jeho využití mohu jen se souhlasem TUL, která má právo ode mne požadovat přiměřený příspěvek na úhradu nákladů vynaložených univerzitou na vytvoření díla (až do jejich skutečné výše).

Bakalářskou práci jsem vypracoval samostatně s použitím uvedené literatury a na základě konzultací s vedoucím bakalářské práce a konzultantem.

Datum:

Podpis:

## ABSTRAKT

V dnešní době jsou nejpoužívanější statistické modely, popisující dobu do poruchy, exponenciální, s distribuční funkcí  $F(t) = 1 - e^{-\lambda.t}$ , a Weibullův, s distribuční funkcí  $F(t) = 1 - e^{-\left(\frac{t}{\alpha}\right)^\beta}$ , popisují průběh pravděpodobnosti, že po určité době dojde k poruše. U exponenciálního se předpokládá, že počet poruch ve zvoleném intervalu o délce  $\langle t_1, t_2 \rangle$  je nezávislý na volbě času  $t_1$ . Weibullovo rozdělení vyhovuje pro výrobky, které selhávají vlivem opotřebení.

V kapitole 4 je popsáno modifikované Weibullovo rozdělení, které má distribuční funkci ve tvaru  $F(t) = 1 - e^{-at^\beta e^{\lambda t}}$ . Podle článku [1] [3] lze toto rozdělení využít (při dostatku dat o poruchách systému) při popisu všech tří částí vanové křivky (popsané v kapitole 2.3). Cílem této práce je ověřit na základě vstupních dat, jakým způsobem lze aplikovat toto rozdělení a jaké jsou jeho výsledky.

Pro ověření použitelnosti tohoto rozdělení je nutné mít velké množství dat o poruchách. Jako vhodná data byly použity informace o době do úmrtí člověka. I když tato data mohou být vnímána jako neetická, z pohledu doby do poruchy, mají několik výhod. Stáří člověka při úmrtí je přesně známo. Je relativně snadné získání velkého množství vstupních dat. Intenzita úmrtí člověka v nízkém věku se výrazně snižuje, později je přibližně konstantní a ke konci života se výrazně zvyšuje. Což reprezentuje tvar vanové křivky.

Cílem této práce je ukázat použitelnost daného rozdělení a dát návod, jak získat jeho jednotlivé parametry.

**Klíčová slova:** Weibullovo rozdělení, exponenciální rozdělení, modifikované Weibullovo rozdělení

## ABSTRACT

Nowadays the most used statistic models are exponential distribution, with distribution function  $F(t) = 1 - e^{-\lambda.t}$ , which describes time to failure and Weibull distribution, with distribution function  $F(t) = 1 - e^{-\left(\frac{t}{\alpha}\right)^\beta}$ , which describes process of probability that after a specific time leads to failure. By the exponential distribution, it is supposed that the count of failures in chosen interval of length  $< t_1, t_2 >$  is independent on the option of time  $t_1$ . Weibull distribution corresponds for products, which are malfunctioned by the effect of attrition.

Modified Weibull distribution is described in chapter 4, which has distribution function  $F(t) = 1 - e^{-\alpha t^\beta e^{\lambda t}}$ . According to the article [1] [3], it is possible to use this distribution (with enough dates of system failures) in description of all three parts of bathtub curve (described in chapter 2.3). The aim of this work is to verify, on the basis of input dates, what way it is possible to apply this distribution and what its results are.

To verify the applicability of this distribution, it is necessary to have a large amount of data failure. Information about the time to death of a person was used as a suitable data. Even if these dates might be perceived as unethical, from the perspective of time to failure, they have a few advantages. The age of death of a person is exactly known. It is relatively easy to obtain a large number of input dates. The intensity of human deaths reduces significantly in young age, later it is approximately constant and at the end of life it increases markedly. It is represented by the bathtub curve.

The objective of this bachelor thesis is to demonstrate the applicability of the selected division and to give guidance for getting its various parameters.

**Key words:** Weibull distribution, exponential distribution, modified Weibull distribution

Chtěl bych poděkovat vedoucímu bakalářské práce Ing. Josefu Chudobovi, za podporu při vedení této práce, za jeho ochotu, konzultace a případné připomínky při vypracovávání této bakalářské práce. Dále bych chtěl poděkovat rodičům za podporu a umožnění studií.

## OBSAH:

SEZNAM OBRÁZKŮ A TABULEK: .....	9
1. ÚVOD .....	10
2. TEORETICKÁ ČÁST .....	11
2.1 Náhodná proměnná .....	11
2.2 Možnosti popisu rozdělení pravděpodobnosti náhodné veličiny .....	11
2.2.1 Distribuční funkce .....	12
2.2.2 Hustota pravděpodobnosti .....	13
2.2.3 Intenzita náhodného jevu .....	14
2.3 Vanová křivka .....	15
2.4 Přehled relací mezi $F(t)$ , $f(t)$ , $\lambda(t)$ a $R(t)$ .....	16
2.5 Test dobré shody $\chi^2$ .....	16
3. EXPONENCIÁLNÍ ROZDĚLENÍ .....	18
4. WEIBULLOVO ROZDĚLENÍ .....	21
5. MODIFIKOVANÉ WEIBULLOVO ROZDĚLENÍ .....	24
6. PŘÍKLAD - VSTUPNÍ DATA .....	27
6.1 Získání vektoru vstupních dat .....	28
6.2 Výhody a nevýhody vstupních dat .....	29
7. POSTUP ZÍSKÁNÍ BODOVÝCH ODHADŮ PARAMETRŮ MODIFIKOVANÉHO WEIBULLOVA ROZDĚLENÍ .....	31
7.1 Zdrojový kód .....	31
7.1.1 Deklarace proměnných .....	31
7.1.2 Popis omezení oblasti .....	32
7.1.3 Popis stanovení parametru $\lambda$ a $\beta$ .....	34
7.1.4 Popis řešení rovnic (5.6) a (5.7) .....	34
7.1.5 Výpis výsledků na monitor .....	35
7.2 Získané výsledky .....	36

7.3	Výpočet parametru $\alpha$ .....	36
7.4	Grafické znázornění výsledků .....	37
8.	TEST DOBRÉ SHODY.....	39
8.1	Vstupní data pro test dobré shody .....	39
8.2	Výpočet .....	39
8.3	Závěry k výsledkům testu dobré shody.....	41
9.	ZÁVĚR .....	42
10.	LITERATURA .....	43

## SEZNAM OBRÁZKŮ A TABULEK:

Obr. 1: Distribuční funkce spojité náhodné proměnné .....	13
Obr. 2: Hustota pravděpodobnosti spojité náhodné proměnné .....	14
Obr. 3: Vanová křivka $\lambda(t)$ s třemi typickými obdobími .....	15
Obr. 4: Graf hustoty pravděpodobnosti exponenciálního rozdělení pro dané parametry..	19
Obr. 5: Graf distribuční funkce exponenciálního rozdělení pro dané parametry.....	19
Obr. 6: Graf intenzity poruch exponenciálního rozdělení pro dané parametry.....	20
Obr. 7: Graf distribuční funkce Weibullova rozdělení pro dané parametry .....	22
Obr. 8: Graf hustoty pravděpodobnosti Weibullova rozdělení pro dané parametry ..	23
Obr. 9: Graf intenzity poruch Weibullova rozdělení pro dané parametry .....	23
Obr. 10: Intenzita poruch modifikovaného Weibullova rozdělení.....	26
Obr. 11: Ukázka zjišťování optimálních parametrů $\beta$ a $\lambda$ .....	33
Obr. 12: Intenzita poruch modifikované Weibullova rozdělení v období 0. - 5. roku.....	37
Obr. 13: Intenzita poruch modifikované Weibullova rozdělení v období 0. - 20. roku.....	38
Obr. 14: Intenzita poruch modifikované Weibullova rozdělení v období 0. - 50. roku .....	38
Obr. 15: Intenzita poruch modifikované Weibullova rozdělení v období 0. - 104. roku ...	38
Tab. 1: Matematické převodní vztahy mezi základními funkčními ukazateli bezporuchovosti .....	16
Tab. 2: Ukázka podrobné úmrtnostní tabulky.....	27
Tab. 3: Ukázka získaných dat pro období mezi 1. a 2. věkem.....	29
Tab. 4: Výsledky testu dobré shody.....	40



## 1. ÚVOD

Cílem této práce je na základě reálných dat ověření použitelnosti Weibullova modifikovaného rozdělení s distribuční funkcí  $F(t) = 1 - e^{-\alpha t^\beta e^{\lambda t}}$  a zjistit jeho parametry. K popisu doby do poruchy komponent, ve všech fázích technického života, se používá Weibullovo rozdělení s distribuční funkcí  $F(t) = 1 - e^{-\left(\frac{t}{\alpha}\right)^\beta}$  a exponenciální rozdělení s  $F(t) = 1 - e^{-\lambda t}$ . Weibullovo modifikované rozdělení by podle článku [1] a [3] mělo umožnit popis celého průběhu komponenty. Rozdělení je tříparametrické, proto je potřebné pro zjištění dat velké množství vstupních dat o poruchách komponent.

Použitá data o úmrtí člověka mají vhodné vlastnosti k použití pro popis modifikovaného Weibullova rozdělení. Z počátku života člověka je úmrtnost vysoká, která postupně klesá a ustaluje se. Ke konci se úmrtnost opět zvyšuje (je dáno vysokým věkem). Vstupní data získaná ze statistických ročenek [8] popisují úmrtnost v jednotlivých letech podle dosaženého věku. Výhoda těchto statistických dat je, že je jich dostatek a jsou přesná. Zvolená tabulka úmrtnosti nám velice dobře reprezentuje vstupní data doby do poruchy, protože je přesně znám časový údaj o úmrtí - z pohledu spolehlivosti doba do poruchy. Pro potřebu bakalářské práce bylo získáno 100 000 vstupních dat. Autor předpokládá, že toto množství vstupních dat je dostatečné k prokázání použitelnosti či nepoužitelnosti modifikovaného Weibullova rozdělení. Popis získání dat je uveden v kapitole 6.1.

V kapitole 2 se seznámíme se základními pojmy spojenými s danou tematikou, která je nezbytná pro pochopení a orientaci v této práci. V práci se uvádí již odvozené vzorce, jejichž odvození se zde nepopisuje, jsou přebrány z [1], [2], [3], [4], [5] a [7].

Výpočty parametrů modifikovaného Weibullova rozdělení jsou prováděny pomocí programu Matlab. Podrobný popis a postup výpočtu je popsán v kapitole 7. Pro výsledné parametry se bude ověřovat testem dobré shody (vysvětlené v kapitole 2.5), zda je možné toto rozdělení použít pro popis doby do úmrtí člověka.

## 2. TEORETICKÁ ČÁST

### 2.1 Náhodná proměnná

[6], [11]

Náhodnou proměnnou nazveme takovou proměnnou:

- jejíž každá hodnota je jednoznačně určena výsledkem náhodného pokusu;
- která může nabývat libovolné hodnoty z definovaného oboru hodnot, vždy však pouze s určitou pravděpodobností, kterou lze vyjádřit jistým zákonem rozdělení pravděpodobnosti, např. *distribuční funkcí, hustotou pravděpodobnosti apod.*

Náhodné proměnné budeme dále označovat velkými písmeny z konce abecedy ( $X, Y, Z, \dots$ ) a jejich možné hodnoty odpovídajícími malými písmeny ( $x, y, z, \dots$ ).

Při praktických aplikacích se zpravidla setkáváme s náhodnými proměnnými dvojího typu:

- *Diskrétní náhodná proměnná  $T$*  je taková náhodná proměnná, která může nabývat jen hodnot z nějaké konečné, nebo spočetné množiny  $\{t_1, t_2, t_3, \dots\}$ .
- *Spojité náhodná proměnná  $T$*  je taková náhodná proměnná, která může nabývat všech hodnot z určitého intervalu. Například doba bezporuchového provozu systému  $T$  může nabývat hodnot  $t \in (0, \infty)$ .

V bakalářské práci budu používat především spojitou náhodnou proměnnou. Proměnná  $t$  bude představovat čas.

### 2.2 Možnosti popisu rozdělení pravděpodobnosti náhodné veličiny

[6], [11]

Pro praktickou práci s náhodnými proměnnými potřebujeme znát pravidla, podle kterých je možné hodnotám náhodné proměnné přiřazovat odpovídající pravděpodobnosti. Například s jakou pravděpodobností se porouchá sledovaná součástka do nějakého času  $t$ .

Je-li takový vztah mezi jevy a pravděpodobnostmi jejich nastoupení znám, říkáme, že je dán zákon rozdělení pravděpodobnosti příslušné náhodné proměnné.

### 2.2.1 Distribuční funkce

[6], [11]

Každá náhodná proměnná je charakterizovaná především svojí distribuční funkcí. Distribuční funkcí náhodné proměnné  $T$  v intervalu  $< 0; \infty$ ) rozumíme funkci  $F(t)$  definovanou vztahem:

$$F(t) = P(T \leq t) \tag{2.1}$$

Hodnota funkce  $F(t)$  v bodě  $t$  tedy vyjadřuje pravděpodobnost toho, že náhodná proměnná  $T$  nabude hodnoty menší nebo rovné  $t$ . Distribuční funkce má následující vlastnosti:

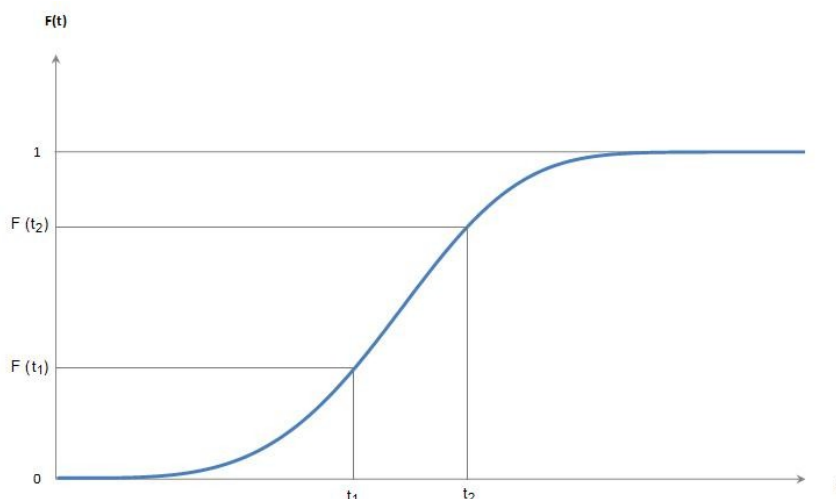
- $0 \leq F(t) \leq 1$ ,
- $F(t)$  je neklesající a zleva spojitá funkce,
- $\lim_{t \rightarrow 0} F(t) = 0$ ,
- $\lim_{t \rightarrow \infty} F(t) = 1$ .

Na základě uvedených vlastností distribuční funkce lze také odvodit důležitý vztah pro pravděpodobnost toho, že náhodná proměnná nabude hodnoty z jistého intervalu  $t \in (t_1, t_2)$  :

$$P(t_1 < T \leq t_2) = F(t_2) - F(t_1) \tag{2.2}$$

Tento vztah říká, že pravděpodobnost toho, že náhodná proměnná nabude hodnoty z daného intervalu je rovna rozdílu hodnot distribuční funkce  $F(t)$  v krajních bodech tohoto intervalu.

Na Obr. 1 je zobrazen příklad distribuční funkce. Distribuční funkce interpretuje pravděpodobnost, že se komponenta porouchá do času  $t$ , za předpokladu, že v čase  $t = 0$  byla komponenta v provozuschopném stavu.



**Obr. 1:** Distribuční funkce spojitě náhodné proměnné

Ve spolehlivosti se často také používá komplement k distribuční funkci, který nazýváme funkce spolehlivosti (bezporuchovosti), protože vyjadřuje pravděpodobnost toho, že jev (např. porucha) do okamžiku  $t$  nenastane:

$$R(T) = 1 - F(T) = P(T > t) \quad (2.3)$$

### 2.2.2 Hustota pravděpodobnosti

[4], [6]

Dalším ukazatelem rozdělení pravděpodobnosti spojitě náhodné proměnné může být také vyjádřen pomocí hustoty pravděpodobnosti.

Náhodná proměnná  $T$  má rozdělení spojitěho typu, existuje-li nezáporná reálná funkce  $f(t)$  taková, že pro všechna kladná  $t$  se dá distribuční funkce  $F(t)$  vyjádřit ve tvaru:

$$F(t) = \int_0^t f(t) \cdot dt \quad (2.4)$$

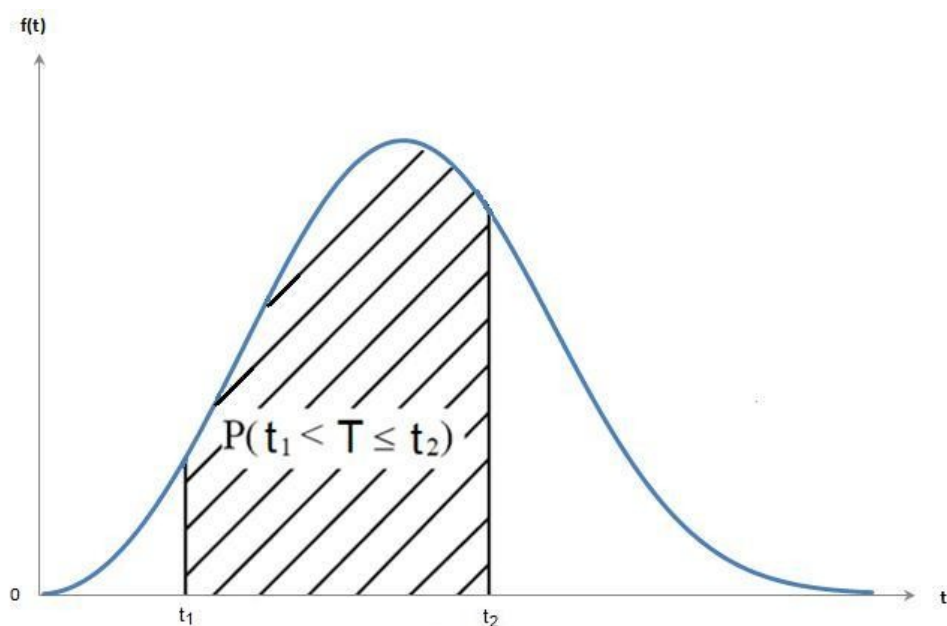
kde funkce  $f(t)$  se nazývá hustota pravděpodobnosti náhodné proměnné  $t$ . Tuto funkci je možno vyjádřit vztahem:

$$f(t) = \frac{dF(t)}{dt}$$

(2.5)

Hustota pravděpodobnosti má následující vlastnosti:

- $\int_0^{\infty} f(t) \cdot dt = 1$ ,
- $P(t_1 < T \leq t_2) = F(t_2) - F(t_1) = \int_{t_1}^{t_2} f(t) \cdot dt$ , pro všechna reálná čísla  $t_1 < t_2$ ,
- $P(T = t) = 0$ .



**Obr. 2:** Hustota pravděpodobnosti spojité náhodné proměnné

### 2.2.3 Intenzita náhodného jevu

[6], [11]

Intenzita náhodného jevu je definována jako podmíněná pravděpodobnost toho, že jev nastane během nekonečně malého intervalu  $dt$  za podmínky, že do okamžiku  $t$  jev nenastal. Tuto podmíněnou pravděpodobnost můžeme vyjádřit následujícím vztahem:

$$\lambda(t) = \frac{f(t)}{1 - F(t)} \tag{2.6}$$

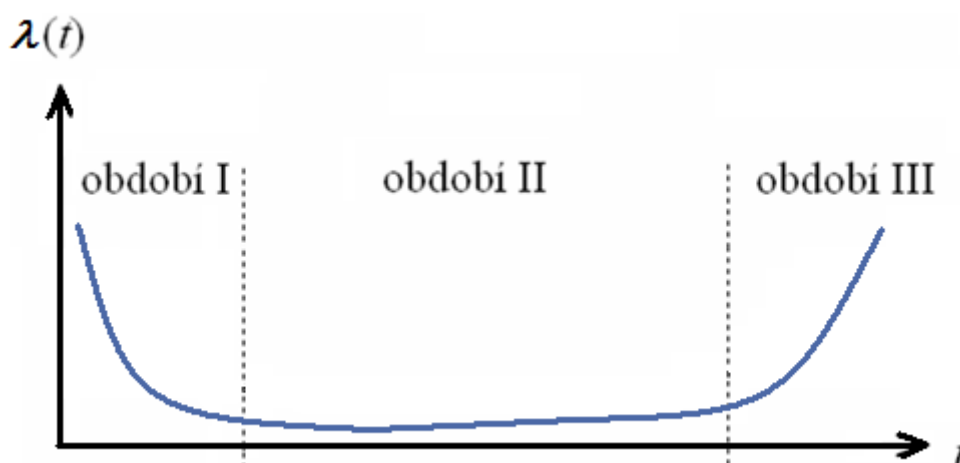
### 2.3 Vanová křivka

[4], [12]

Vanová křivka popisuje okamžitou intenzitu poruch komponenty  $\lambda(t)$ . U mnoha objektů funkce intenzity poruch závislá na čase klesá na určitou hodnotu, která je pak po nějakou dobu přibližně konstantní, a poté zase narůstá. Z těchto důvodů se jí říká křivka vanová, a její tři oblasti jsou charakterizovány v následujících odstavcích.

První období I se nazývá obdobím časných poruch a je typická jejich zvýšeným množstvím (období dětských nemocí během zavádění komponenty do provozu apod.). První období technického života se popisuje Weibullovým rozdělením s parametrem  $\beta < 1$ . Tyto poruchy jsou zaviněny různými výrobními vadami při montáži i chybami při samotném návrhu. Po jejich odstranění však intenzita poruch klesá, protože došlo k záběhu výrobku a následně k poruchám dochází z vnějších příčin. Intenzita poruch je téměř konstantní a toto období II se nazývá obdobím normálního využívání a popisuje ho exponenciální rozdělení.

Na Obr. 3 je uvedena typická vanová křivka komponenty popisující změny intenzity poruch v čase. Období uvedené na obrázku jsou uvedené výše.



**Obr. 3:** Vanová křivka  $\lambda(t)$  s třemi typickými obdobími

Třetí období vanové křivky je charakterizováno procesy stárnutí, projevuje se degradace materiálu a intenzita poruch  $\lambda(t)$  opět narůstá. Toto období se nazývá období poruch dožitím. Popisuje Weibullovo rozdělení s parametrem  $\beta > 1$ .

Vanovou křivku obvykle nejsme schopni modelovat nějakou jednoduchou analytickou funkcí. Zpravidla ji modelujeme různými funkcemi v jednotlivých obdobích, a to tak, aby pokud možno byla dostatečně hladká.

#### 2.4 Přehled relací mezi $F(t)$ , $f(t)$ , $\lambda(t)$ a $R(t)$

V Tab. 1 jsou uvedeny relace mezi distribuční funkcí  $F(t)$ , pravděpodobností bezporuchového provozu  $R(t)$ , hustotou pravděpodobnosti  $f(t)$  a intenzitou poruch  $\lambda(t)$ . Převody mezi jednotlivými ukazateli jsou pro uvedená rozdělení jednoznačné.

**Tab. 1:** Matematické převodní vztahy mezi základními funkčními ukazateli bezporuchovosti [5]

	$F(x)$	$R(x)$	$f(x)$	$h(x)$
$F(x)$		$1 - F(x)$	$\frac{dF(x)}{dx}$	$\frac{f(x)}{1 - F(x)}$
$R(x)$	$1 - R(x)$		$-\frac{dR(x)}{dx}$	$-\frac{1}{R(x)} \cdot \frac{dR(x)}{dx}$
$f(x)$	$\int_0^x f(u) \cdot du$	$\int_x^\infty f(u) \cdot du$		$\frac{f(x)}{\int_x^\infty f(u) \cdot du}$
$h(x)$	$1 - \exp\left[-\int_0^x h(u) \cdot du\right]$	$\exp\left[-\int_0^x h(u) \cdot du\right]$	$h(x) \cdot \exp\left[-\int_0^x h(u) \cdot du\right]$	

#### 2.5 Test dobré shody $\chi^2$

[9]

Test dobré shody se využívá, jestliže se testuje, zda získaná statistická data pocházejí či nepocházejí z předpokládaného rozdělení. Soubor dat se roztrídí do několika skupin. Pomocí testu dobré shody se zjišťuje pomocí vzorce (2.7) hodnota testové statistiky, která se limitně blíží  $\chi^2$  rozdělení s  $k - h - 1$  stupni volnosti,

kde  $k$  představuje počet jednotlivých skupin a  $h$  počet neznámých odhadovaných parametrů.

Při výpočtu se pro ověřování sledovaného rozdělení testují následující hypotézy:

- $H_0$ : Výběr pochází z populace, v níž jsou relativní četnosti jednotlivých variant rovny číslům  $\pi_{0,1}$ ;  $\pi_{0,2}$ ; ...;  $\pi_{0,k}$  (populace musí být roztržitelná

podle nějakého znaku do  $k$  skupin). Například číslo  $\pi_{0,2}$  představuje teoretickou pravděpodobnost v druhé skupině.

- $H_0$ : Výběr pochází z rozdělení určitého typu (např. normální), jehož parametry jsou dány (úplně specifikovaný model).
- $H_0$ : Výběrový soubor pochází z rozdělení určitého typu (např. normální), (neúplně specifikovaný model - neověřujeme informace o parametrech rozdělení, parametry modelu odhadujeme).

Oproti hypotéze  $H_1$ , která je negací hypotézy  $H_0$ .

### Volba testové statistiky

Jako testovou statistiku volíme statistiku  $G$ , která má pro dostatečný rozsah výběru asymptoticky  $\chi_{k-h-1}^2$  rozdělení:

$$G = \sum_{i=1}^k \frac{(n_i - n \cdot \pi_{0,i})^2}{n \cdot \pi_{0,i}} \rightarrow \chi_{k-h-1}^2$$

(2.7)

kde

$n$  - je rozsah výběru vstupních dat,

$n_i$  - je skutečná četnost v  $i$ -té skupině,

$k$  - je počet skupin,

$h$  - je počet odhadovaných parametrů modelu,

$\pi_{0,i}$  - jsou očekávané relativní četnosti v  $i$ -té skupině (tj. relativní četnosti, jichž by měly nabýt jednotlivé varianty v případě, že je splněna nulová hypotéza).

$n \cdot \pi_{0,i}$  jsou tedy očekávané četnosti jednotlivých variant (tj. četnosti, jichž by měly nabýt jednotlivé varianty v případě, že je splněna nulová hypotéza) a  $(n_i - n \cdot \pi_{0,i})$  pak jsou odchylky očekávaných četností od četností skutečných.

Při testu dobré shody je nutné splnit podmínku, že všechny očekávané četnosti jsou vyšší než 5 - ( $n \cdot \pi_{0,i} > 5$  ( $i = 1, 2, \dots, k$ )).

Je-li vypočítaná hodnota  $G$ , podle vzorce (2.7), větší než tabulková, potom zamítneme  $H_0$  - tj. zamítneme přijetí hypotézy, že data pocházejí z daného rozdělení. Je-li zjištěná hodnota  $G$  menší než tabulková nezamítne se hypotéza  $H_0$ , že data pocházejí z daného rozdělení.



### 3. EXPONENCIÁLNÍ ROZDĚLENÍ

Jedná se o jednoparametrické rozdělení s parametrem  $\lambda$ , náhodná veličina  $T$  má exponenciální rozdělení s parametrem  $\lambda$ , jestliže hustota pravděpodobnosti je dána vztahem:

$$f(t) = \lambda e^{-\lambda \cdot t} \quad \text{kde } \lambda > 0, t \geq 0. \quad (3.1)$$

Pro distribuční funkci platí vztah:

$$F(t) = 1 - e^{-\lambda \cdot t} \quad (3.2)$$

Pro intenzitu náhodného jevu platí<sup>1</sup>:

$$\lambda(t) = \lambda \quad (3.3)$$

Funkce spolehlivosti je dána vztahem:

$$R(t) = e^{-\lambda \cdot t} \quad (3.4)$$

Střední hodnota:

$$E(T) = \frac{1}{\lambda} \quad (3.5)$$

Rozptyl:

$$D(T) = \frac{1}{\lambda^2} \quad (3.6)$$

Exponenciální rozdělení se často používá ve výzkumu spolehlivosti, v teorii hromadné obsluhy, v teorii obnovy apod. Využívá se zejména pro popis bezporuchovosti těch objektů, kde se neprojevuje vliv postupné degradace součástí, jako např. stárnutí, koroze, opotřebení apod. Konstantní intenzita poruch se například běžně předpokládá u vysoce spolehlivých, složitých technických systémů. [6], [10]

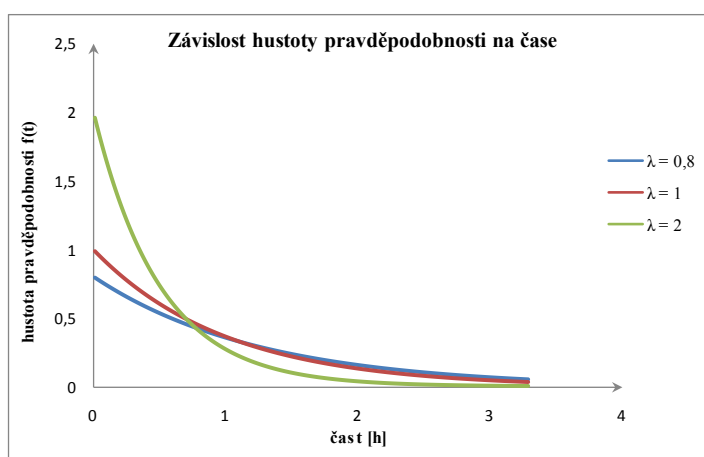
---

<sup>1</sup> Pokud jsou sledovanými náhodnými jevy poruchy, hovoříme o intenzitě poruch.

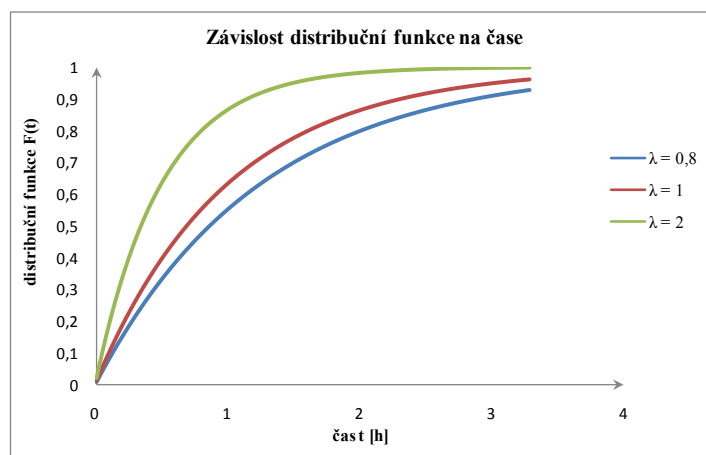
Na Obr. 4 až 6 jsou zobrazeny grafy:

- hustoty pravděpodobnosti  $f(t)$  – Obr. 4,
- distribuční funkce  $F(t)$  – Obr. 5,
- intenzity poruch  $\lambda(t)$  – Obr. 6

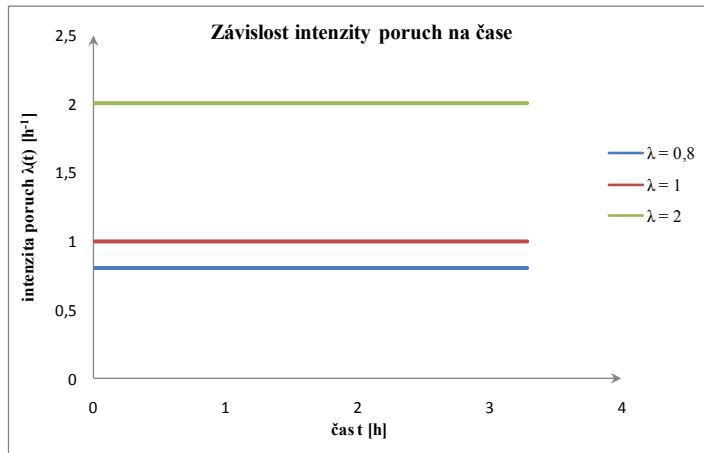
na čase  $t$ . Jsou uvedeny tři modelové případy, kdy intenzita poruch je buď  $\lambda = 0,8$ ;  $\lambda = 1$ ;  $\lambda = 2$ . tj. střední doba do poruchy je buď  $E(T) = 1,25$ ;  $E(T) = 1$ ;  $E(T) = 0,5$ .



**Obr. 4:** Graf hustoty pravděpodobnosti exponenciálního rozdělení pro dané parametry



**Obr. 5:** Graf distribuční funkce exponenciálního rozdělení pro dané parametry



**Obr. 6:** Graf intenzity poruch exponenciálního rozdělení pro dané parametry

#### 4. WEIBULLOVO ROZDĚLENÍ

Jedná se o dvouparametrické rozdělení s parametry  $\alpha$  a  $\beta$ , náhodná veličina  $T$  má Weibulovo rozdělení s parametry  $\alpha$  a  $\beta$ , jestliže hustota pravděpodobnosti je dána vztahem:

$$f(t) = \frac{\beta}{\alpha^\beta} \cdot t^{\beta-1} \cdot e^{-\left(\frac{t}{\alpha}\right)^\beta} \quad (4.1)$$

Kde  $\alpha$  - parametr polohy rozdělení,  $\beta$  - parametr tvaru rozdělení.

Definiční obor veličin:  $t \geq 0$ ;  $\alpha, \beta > 0$ .

Pro distribuční funkci platí vztah:

$$F(t) = 1 - e^{-\left(\frac{t}{\alpha}\right)^\beta} \quad (4.2)$$

Funkci spolehlivosti je dána vztahem:

$$R(t) = e^{-\left(\frac{t}{\alpha}\right)^\beta} \quad (4.3)$$

Pro intenzitu poruch platí vztah:

$$\lambda(t) = \frac{\beta \cdot t^{\beta-1}}{\alpha^\beta} \quad (4.4)$$

Střední hodnota:

$$E(T) = \alpha \cdot \Gamma\left(1 + \frac{1}{\beta}\right) \quad (4.5)$$

Rozptyl:

$$D(T) = \alpha^2 \cdot \left[ \Gamma\left(1 + \frac{2}{\beta}\right) - \Gamma^2\left(1 + \frac{1}{\beta}\right) \right] \quad (4.6)$$

Kde symbol  $\Gamma(x)$  zobrazuje gama funkci.

Weibullovo rozdělení se používá tehdy, když nelze přijmout předpoklad o konstantní intenzitě jevu. Ve spolehlivosti je toto rozdělení široce využíváno pro popis dob spojených s poruchami tak i dob nápravné údržby. Rozdělení s parametrem  $\beta > 1$  umožňuje dobrý popis bezporuchovosti a životnosti objektů

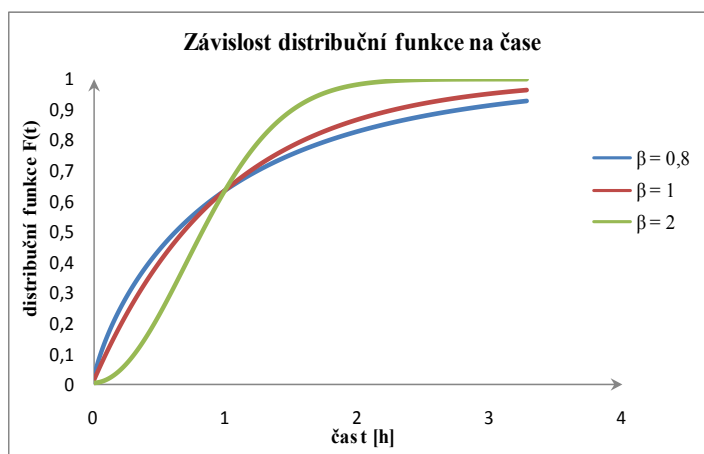
u kterých se výrazně projevuje vliv opotřebení, únavy, koroze a dalších degradačních procesů. Rozdělení s parametrem  $\beta < 1$  umožňuje popis bezporuchovosti v počátečních fázích provozu, kdy se projevují výrobní vady. Je-li parametr  $\beta = 1$  Weibullovo rozdělení přechází v exponenciální. [6], [10]

Na Obr. 7 až 9 jsou zobrazeny grafy:

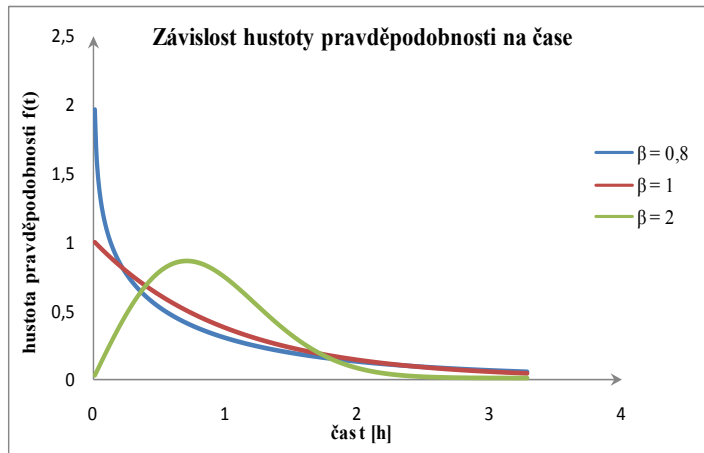
- distribuční funkce  $F(t)$  – Obr. 7,
- hustoty pravděpodobnosti  $f(t)$  – Obr. 8,
- intenzity poruch  $\lambda(t)$  – Obr. 9

na čase  $t$ . Jsou uvedeny tři modelové případy, kdy parametr degradace  $\beta$  je buď 0,8; 1 nebo 2.

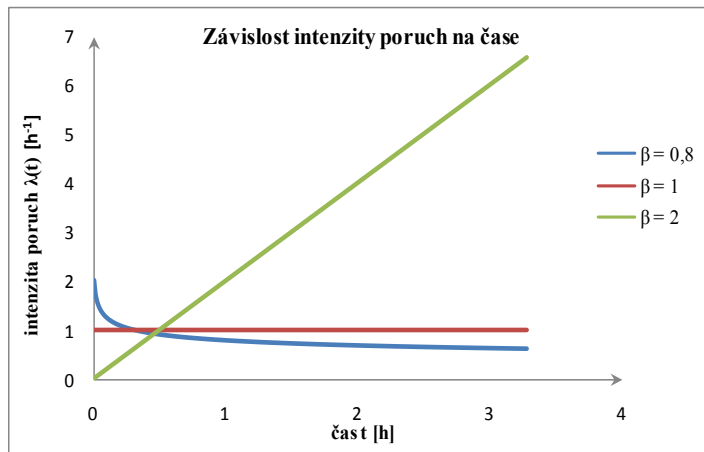
V každém z těchto tří případů se předpokládá, že parametr  $\alpha = 1$ . První případ, kdy  $\beta = 0,8$  lze použít pro popis prvního období vanové křivky. Jestliže je parametr  $\beta = 1$  potom se systém popisuje pomocí exponenciálního rozdělení a zobrazuje chování při běžném opotřebení. Třetím případem kdy parametr  $\beta = 2$  lze popsat období degradace komponenty.



**Obr. 7:** Graf distribuční funkce Weibullova rozdělení pro dané parametry



**Obr. 8:** Graf hustoty pravděpodobnosti Weibullova rozdělení pro dané parametry



**Obr. 9:** Graf intenzity poruch Weibullova rozdělení pro dané parametry

## 5. MODIFIKOVANÉ WEIBULLOVO ROZDĚLENÍ

Jedná se o tříparametrické rozdělení s parametry  $\alpha$ ,  $\beta$  a  $\lambda$ , náhodná veličina  $T$  má modifikované Weibullovo rozdělení s parametry  $\alpha$  a  $\beta$  a  $\lambda$ , jestliže hustota pravděpodobnosti je dána vztahem [1], [3] a [7]:

$$f(t) = \alpha(\beta + \lambda t) \cdot t^{\beta-1} \cdot e^{\lambda t} \cdot e^{-\alpha t^{\beta} e^{\lambda t}} \quad (5.1)$$

kde  $\alpha > 0$  a  $\beta, \lambda \geq 0$

Pro distribuční funkci platí vztah:

$$F(t) = 1 - e^{-\alpha t^{\beta} e^{\lambda t}} \quad (5.2)$$

Pro intenzitu poruch platí vztahy:

$$\lambda(t) = \alpha(\beta + \lambda t) \cdot t^{\beta-1} \cdot e^{\lambda t} \quad (5.3)$$

Funkce spolehlivosti:

$$R(t) = e^{-\alpha t^{\beta} e^{\lambda t}} \quad (5.4)$$

Výhoda modifikovaného Weibullova rozdělení je, podle textu [1], [3], že popisuje všechny tři období života komponenty. Je to dáno tím, že v exponenciální funkci jsou uvedeny dvě funkce -  $e^{\lambda t}$ ,  $\alpha t^{\beta}$ , kde jedna obvykle popisuje první období vanové křivky, zatímco druhá vystihuje poruchovost komponenty ve třetím období vanové křivky.

Nevýhoda modifikovaného Weibullova rozdělení je, že je potřeba velké množství vstupních dat pro určení bodových odhadů neznámých parametrů  $\alpha, \beta, \lambda$ . Další nevýhoda je v obtížnějším použitém matematickém aparátu. Třetí nevýhodou je, že jednotlivé bodové odhady parametrů nemají fyzikální základ a nelze vzájemným porovnáváním určit spolehlivostní vlastnosti komponenty.

Metodou maximální věrohodnosti<sup>2</sup>, jejíž popis zde není uváděn pro svou rozsáhlost a složitost, popsané v [2] a [5] lze získat bodové odhady parametrů  $\alpha, \beta, \lambda$  modifikovaného Weibullova rozdělení.

Věrohodnostní funkce modifikovaného Weibullova rozdělení je: [1], [3]

$$l(\alpha, \beta, \lambda; t) = n \log \alpha + \sum_{i=1}^n \log(\beta + \lambda t_i) + \\ + (\beta - 1) \cdot \sum_{i=1}^n \log t_i + \lambda \sum_{i=1}^n t_i - \alpha \sum_{i=1}^n t_i^\beta e^{\lambda t_i} \quad (5.5)$$

Z věrohodnostní funkce se podle jednotlivých parciálních derivací podle proměnných obdrží bodové odhady parametrů  $\beta, \lambda$  a  $\alpha$ . Bodový odhad parametru  $\alpha$  lze získat z bodových odhadů parametru  $\beta$  a  $\lambda$ . Bodové odhady parametru  $\beta$  a  $\lambda$  nejsou závislé na parametru  $\alpha$ .

#### **Bodové odhady parametrů – $\beta, \lambda$ :**

Bodové odhady parametrů  $\beta$  a  $\lambda$  modifikovaného Weibullova rozdělení se získají řešením rovnic (5.1) a (5.2) vyplývajících z metody největší věrohodnosti.

$$\sum_i \frac{1}{\beta + \lambda t_i} + \sum_i \ln t_i - \frac{n \sum_i t_i^\beta e^{\lambda t_i} \ln t_i}{\sum_i t_i^\beta e^{\lambda t_i}} = 0 \quad (5.6)$$

$$\sum_i \frac{t_i}{\beta + \lambda t_i} + \sum_i t_i - \frac{n \sum_i t_i^{\beta+1} e^{\lambda t_i} \ln t_i}{\sum_i t_i^\beta e^{\lambda t_i}} = 0 \quad (5.7)$$

---

<sup>2</sup> Stručné naznačení postupu: Jednotlivé hustoty pravděpodobnosti s dosazením sledovaného času poruchy se vzájemně vynásobí. Hledá se lokální extrém této funkce, což může být obtížné, proto se daná funkce zlogaritmuje viz. rovnice (5.5). Logaritmováním se nezmění hodnota lokálního extrému. Aby existoval lokální extrém, musí parciální derivace podle jednotlivých parametrů být rovny nule. Tyto parciální derivace jsou uvedeny v rovnicích (5.6) a (5.7).



Následně se stanoví z bodových odhadů parametrů  $\beta$  a  $\lambda$  parametr  $\alpha$ .

$$\alpha = \frac{n}{\sum_i t_i^\beta e^{\lambda t_i}} \quad (5.8)$$

Kde

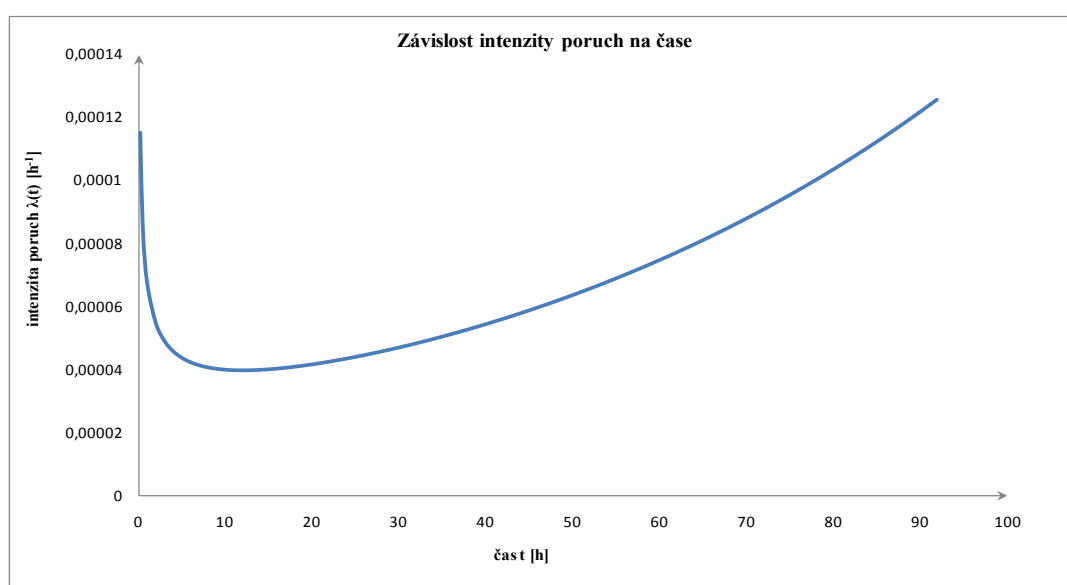
- $\alpha, \beta, \lambda$  - bodové odhady modifikovaného Weibullova rozdělení,
- $t_i$  - stáří při úmrtí  $i$ -tého člověka,
- $n$  - počet vstupních dat.

Cílem bakalářské práce bylo popsat a zjistit bodové odhady parametrů modifikované Weibullova rozdělení. Zdrojový text pro zjištění parametrů je uveden v kapitole 7.1.

Na Obr. 10 je ukázka grafu intenzity poruch modifikovaného Weibullova rozdělení pro parametry

- $\alpha = 0,0001$ ,
- $\beta = 0,65$ ,
- $\lambda = 0,013$ .

Z grafu je evidentní, že skutečně popisuje všechny tři období vanové křivky.



**Obr. 10:** Intenzita poruch modifikovaného Weibullova rozdělení

## 6. PŘÍKLAD - VSTUPNÍ DATA

Jako vstupních dat, pro vyhodnocení bodových odhadů parametrů modifikovaného Weibullova rozdělení  $\alpha, \beta, \lambda$ , byly použity údaje o počtu úmrtí mužů v roce 2001 a 2002. Tato data jsou získána ze stránek Českého statistického úřadu (<http://www.czso.cz/>). Vycházelo se z tabulky, která je pojmenována „Tab. 7 2001-2002 Podrobné úmrtnostní tabulky - město – muži“ [8]. Celá tabulka s názvem „tab\_data.xls“ je uvedena v příloze 1. Její dílčí část je zobrazena v Tab. 2.

**Tab. 2:** Ukázka podrobné úmrtnostní tabulky

<b>Tab. 7.</b>	<b>2001-2002</b>			
<b>Podrobné úmrtnostní tabulky - město muži</b>				
věk	Qx	px	lx	dx
0	0,004477	0,995523	100000	448
1	0,000688	0,999312	99552	68
2	0,000264	0,999736	99484	26
3	0,000137	0,999863	99458	14
4	0,000163	0,999837	99444	16
5	0,000213	0,999787	99428	21
6	0,000237	0,999763	99407	24
7	0,000191	0,999809	99383	19
8	0,000176	0,999824	99364	17
9	0,000134	0,999866	99347	13
10	0,000144	0,999856	99333	14
11	0,000141	0,999859	99319	14
12	0,000155	0,999845	99305	15
13	0,000160	0,999840	99290	16
14	0,000214	0,999786	99274	21
15	0,000296	0,999704	99253	29
16	0,000496	0,999504	99223	49
17	0,000650	0,999350	99174	64
18	0,000830	0,999170	99110	82
19	0,000940	0,999060	99027	93

Tab. 2 obsahuje podrobná data o úmrtnosti mužů mezi rokem 2001 a 2002. Počet sledovaných dat o úmrtí je 100 000. Úmrtnost je evidována v městech. Statistický úřad nabízí velké množství obdobných statistik vázaných např. k jednotlivým oblastem, úmrtnosti žen a mužů. Pro náš případ není podstatné, jestli se jedná o tu nebo onu lokalitu, protože nám tyto data reprezentují dobu do poruchy. A naším úmyslem je pouze ověřit na příkladě použitelnost modifikovaného Weibullova rozdělení pro popis doby do poruchy silně degradující komponenty.

Jednotlivé sloupce v Tab. 2 označují:

- $věk$  - obsahuje hodnoty o dosaženém věku,
- $qx$  - vyjadřuje pravděpodobnost nepřežití o jeden rok, za předpokladu, že se člověk dožil začátku,
- $px$  - pravděpodobnost přežití člověka o jeden rok,
- $lx$  - počet přežilých do daného roku,
- $dx$  - počet zemřelých v jednom roce.

Významný je pro nás sloupec  $dx$  (počet zemřelých v daném období). Pro každý interval bylo vygenerováno náhodnými čísly z rovnoměrného rozdělení přesný věk úmrtí, čímž jsme získali vektor dat od 0 do 104 obsahující 100 000 dat reprezentující přesný věk úmrtí jednotlivce. Předpokládá se, že 100 000 vstupních dat je pro tříparametrické rozdělení dostačující. Způsob získání vstupních dat je uvedeno v kapitole 6.1.

### 6.1 Získání vektoru vstupních dat

Vstupní data, uvedena v Tab. 2, jsou zpracována pomocí programu Microsoft Excel pomocí funkce:

$$=NÁHČÍSLO()+x', \quad (6.1)$$

kde  $x$  udává číslo z intervalu věku. Z vygenerovaného náhodného čísla se například pomocí funkce  $=náhčíslo()+1$  získá dosažený věk při úmrtí v intervalu 1 až 2 roky. Velikost daného intervalu nám udá příslušná hodnota  $dx$  pro věk 1.

Tedy interval mezi 1. a 2. rokem má 68 záznamů určujících přesný časový údaj úmrtí jedince v daném období. Vstupní data se upravila do formy setříděného vektoru a vyexportovala do textového souboru s názvem „data.txt“. V Tab. 3 je uvedena část výstupu z textové souboru uvedeného v příloze 1.

**Tab. 3:** Ukázka získaných dat pro období mezi 1. a 2. věkem

1,01830	1,20771	1,34209	1,50407	1,73698
1,02240	1,22635	1,34909	1,52716	1,74268
1,03548	1,22972	1,37652	1,56875	1,83425
1,03714	1,23066	1,40260	1,56929	1,83536
1,04491	1,23926	1,41590	1,57436	1,85848
1,05108	1,24616	1,44077	1,58356	1,86497
1,05751	1,24803	1,44759	1,62258	1,91743
1,06002	1,24855	1,45189	1,63769	1,93958
1,09909	1,25717	1,45699	1,64428	1,94728
1,11015	1,28562	1,46047	1,64652	1,96480
1,13815	1,28724	1,46551	1,64833	1,97571
1,16609	1,29703	1,47838	1,65173	1,99206
1,19155	1,30755	1,49031	1,66985	
1,19388	1,32505	1,49534	1,70939	

## 6.2 Výhody a nevýhody vstupních dat

Výhoda modifikovaného Weibullova rozdělení je, podle textu [1], [3], že popisuje všechny tři období života komponenty. Je to dáno tím, že v exponenciální funkci jsou uvedeny dvě funkce -  $e^{\lambda t}$ ,  $at^{\beta}$ , kde jedna obvykle popisuje první období vanové křivky, zatímco druhá vystihuje poruchovost komponenty ve třetím období vanové křivky.

Nevýhoda modifikovaného Weibullova rozdělení je, že je potřeba velké množství vstupních dat pro určení bodových odhadů neznámých parametrů

$\alpha, \beta, \lambda$ . Další nevýhoda je v obtížnějším použitém matematickém aparátu. Třetí nevýhodou je, že jednotlivé bodové odhady parametrů nemají fyzikální základ, a nelze vzájemným porovnáváním určit spolehlivostní vlastnosti komponenty.

V našem případě se velké množství vstupních dat získává pomocí doby do úmrtí člověka. To může vyvolávat etické problémy, protože pro některé lidi může být neetické mluvit o úmrtí jako o době do poruchy. Výhodou, ale i zároveň nevýhodou vstupních dat je, že lidé se snaží, aby se co nejvíce oddálit okamžik smrti každého jedince.

## 7. POSTUP ZÍSKÁNÍ BODOVÝCH ODHADŮ PARAMETRŮ MODIFIKOVANÉHO WEIBULLOVA ROZDĚLENÍ

Výpočet bodových odhadů parametrů modifikovaného Weibullova rozdělení lze provést například pomocí programu Matlab, který umožňuje efektivní počítání složitých matematických operací. Výpočet bodových odhadů parametrů spočívá v nalezení parametrů  $\lambda$  a  $\beta$ , tak aby výsledky rovnic (5.6), (5.7) byly rovné nule.

Vstupních dat o poruchách je 100 000. Hledání bodových odhadů parametrů bylo časově náročné i přesto, že bylo použito numerických metod.

### 7.1 Zdrojový kód

Zdrojový kód se napsal pomocí skriptu v programu Matlab. Skript má název „*vypocet. m*“ a je uveden v příloze 2.

#### 7.1.1 Deklarace proměnných

*clear all;* Vymazání hodnot všech proměnných.  
*clc;* Vyčištění příkazového okna.

```
suma1=0;  
suma2=0;  
suma3=0;  
suma4=0;  
suma11=0;  
suma22=0;  
suma33=0;  
suma44=0;
```

Proměnné *suma1*, *suma2*, *suma3*, *suma4*, *suma11*, *suma22*, *suma33*, *suma44*, jsou vynulovány. Tyto proměnné se využívají pro zapsání rovnic (5.6), (5.7), kde představují jednotlivé sumy v rovnicích. Blíže viz. vysvětlení při jejich použití ve *for* cyklech.

*n=100000;* Do proměnné *n* se při spuštění načte počet dat o poruše.  
V našem případě je to 100 000.

`load C:\Work\ti.txt;`

Příkazem `load` se načte obsah textového souboru v daném místě uložení. Soubor obsahuje na každém řádku jeden záznam o době do poruchy. Do proměnné  $n$  je vloženo 100 000, což představuje počet vstupních dat. Tyto načtená data se uloží do pole o velikosti  $1 \times \text{počet prvků}$  (v našem případě  $1 \times 100\,000$ ). V našem případě tedy uložena do pole s názvem „*ti.txt*“, který je uveden v příloze 2.

### 7.1.2 Popis omezení oblasti

V tomto odstavci bude uveden způsob zjištění parametrů  $\lambda$  a  $\beta$ . Podle teorie v kapitole 5 musí být parametry  $\beta \geq 0$  a  $\lambda \geq 0$ . Řeší se soustava rovnic (5.6), (5.7). Při optimálním řešení je pravá strana těchto rovnic rovna nule. Bodové odhady parametrů se určují pomocí *for* cyklu, kdy se zvolí náhodně hodnota parametrů  $\lambda$  a  $\beta$ . Sledují se hodnoty pravých stran rovnic (5.6) a (5.7). Protože náhodné určování parametrů  $\lambda$  a  $\beta$  nevede úspěšně k cíli, volí se parametry v určitém intervalu s konstantní diferencí. Na monitor se zobrazují parametry  $\lambda$  a  $\beta$  jestliže pravá strana rovnic (5.6) a (5.7), dané proměnnou *vysledek1* a *vysledek2*, je menší než námi zvolená hodnota. Například při zadání rozsahu od 0 do 2 v tisících krocích se zjišťuje hodnota pravé strany rovnic pro parametry 0,002; 0,004; 0,006 ... 1,998; 2.

Na Obr. 11 jsou vyznačeny body, ve kterých je prováděn výpočet pravých stran. Křížkem je zobrazen bod, kdy obě rovnice jsou rovny nule. Elipsou je vyznačena oblast, kdy pravé strany dosahují nižších hodnot než námi zvolené. Průnik bodů, ve kterých jsou počítány pravé strany a vnitřní body elipsy představují řešení naší úlohy.

Při tomto zadání by výpočet trval roky. Například v našem případě při 1 000 krocích pro parametr  $\lambda$  a 1 000 krocích pro parametr  $\beta$  by se provádělo  $10^6$  výpočtů. V každém výpočtu se řeší výpočet sum, uvedených v rovnicích (5.6) a (5.7), představujících  $10^5$  násobení. Řádově by bylo nutné provést  $10^{12}$  násobení nebo sčítání.

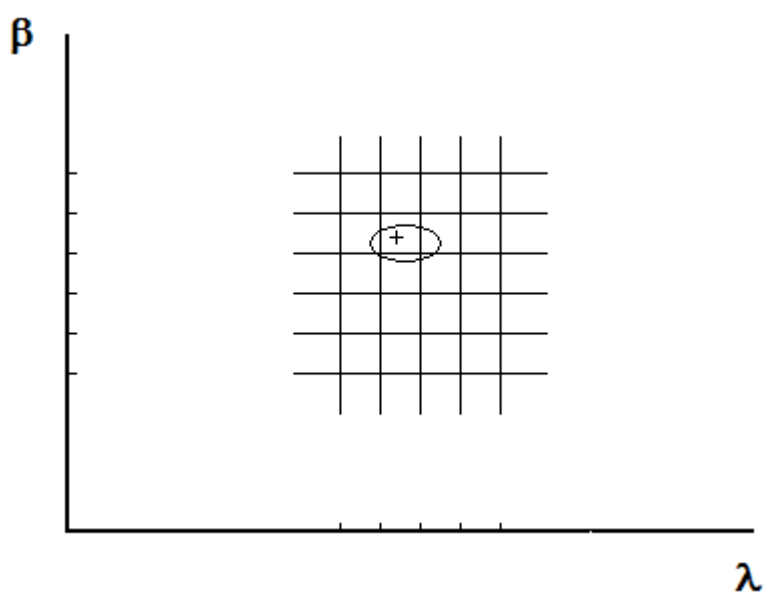
Odstranění tohoto problému bylo vyřešeno několikasupňovým omezováním plochy, na které se může vyskytovat dané řešení rovnic (5.6) a (5.7). Při omezování plochy se vycházelo z oblasti hodnot zobrazených na monitoru.

Rovnice (5.6) a (5.7) mají v oblasti  $\beta \geq 0$  a  $\lambda \geq 0$  maximálně jedno řešení. Předpokládá se, že rovnice budou rovny nule v oblasti zobrazených hodnot na monitoru.

Příklad na omezení oblasti. Parametr  $\beta$  je omezen hodnotami 0 až 2 s konstantním krokem 0,002. Parametr  $\lambda$  je také omezen hodnotami 0 až 2 s konstantním krokem 0,002. Proměnná *vysledek1* a *vysledek2*, představující pravou stranu rovnic (5.6) a (5.7), je nastavena na hodnotu  $10^4$  a  $10^6$ . V případě, že obě dvě proměnné jsou v absolutní hodnotě menší než uvedené konstanty, výsledek se znázorní na monitoru. Zobrazeny jsou například následující hodnoty:

$\beta = 0,186$	0,188	0,188	0,188	0,190	0,190
$\lambda = 0,532$	0,530	0,532	0,534	0,530	0,532

Protože pravděpodobně řešení rovnic je v oblasti pro  $\beta \in \langle 0,186; 0,190 \rangle$  a  $\lambda \in \langle 0,530; 0,534 \rangle$  lze omezit oblast například následovně. Pro parametr  $\beta \in \langle 0,182; 0,194 \rangle$  a pro parametr  $\lambda \in \langle 0,526; 0,538 \rangle$ . Šířka oblasti pro parametr  $\beta$  je 0,012 proto při 1000 krocích bude diference rovna  $\frac{0,012}{1000}$ , což je  $12 \cdot 10^{-6}$ . Jednotlivé body se zjišťují pomocí *for* cyklu, kdy parametr  $\beta$  nabývá hodnot  $0,182 + i \cdot 0,000012$ . Obdobně pro parametr  $\lambda$ . Následně se zmenší hodnota podmínek vypisovaných parametrů *vysledek1* a *vysledek2* na monitor.



**Obr. 11:** Ukázka zjišťování optimálních parametrů  $\beta$  a  $\lambda$



### 7.1.3 Popis stanovení parametru $\lambda$ a $\beta$

*for j=1:20*

*for k=1:50*

$b=0.195+0.0001*j;$

$lambda=0.0869169+0.00000002*k;$

Pomocí cyklu *for j=1:20* se vytvoří rozsah pro parametru  $\beta$ . Pomocí cyklu *for k=1:50* se vytvoří rozsah pro parametru  $\lambda$ . Postupným omezováním oblasti se v poslední fázi se parametr  $\beta$  stanovoval pomocí rovnice  $\beta = 0,195 + j \cdot 0,0001$ . A parametr  $\lambda$  pomocí rovnice  $\lambda = 0,0869169 + k \cdot 2 \cdot 10^{-8}$ . Parametr  $\lambda$  byl přesněji vypočítán, protože obě rovnice byly na něm mnohem více závislé. Celý výpočet byl proveden v 6 samostatných fázích.

### 7.1.4 Popis řešení rovnic (5.6) a (5.7)

*for i=1:n*

Cyklus *for i=1:n* nám nabývá hodnot 1 až počet dat o poruchách, což je v našem případě 100 000.

$suma1=suma1+(1/(b+lambda*ti(i)));$

$suma2=suma2+log(ti(i));$

$suma3=suma3+((ti(i)^b)*exp(lambda*ti(i))*log(ti(i)));$

$suma4=suma4+((ti(i)^b)*exp(lambda*ti(i)));$

$suma11=suma11+((ti(i))/(b+lambda*ti(i)));$

$suma22=suma22+(ti(i));$

$suma33=suma33+((ti(i)^(b+1))*exp(lambda*ti(i)));$

$suma44=suma44+((ti(i)^b)*exp(lambda*ti(i)));$

Jednotlivé proměnné představují:

- $suma1$  – v rovnici (5.6)  $\sum_i \frac{1}{\beta + \lambda t_i}$ ,
- $suma2$  – v rovnici (5.6)  $\sum_i \ln t_i$ ,
- $suma3$  – v rovnici (5.6)  $\sum_i t_i^\beta e^{\lambda t_i} \ln t_i$ ,
- $suma4$  – v rovnici (5.6)  $\sum_i t_i^\beta e^{\lambda t_i}$

- *suma11* – v rovnici (5.7)  $\sum_i \frac{t_i}{\beta + \lambda t_i}$ ,
- *suma22* – v rovnici (5.7)  $\sum_i t_i$ ,
- *suma33* – v rovnici (5.7)  $\sum_i t_i^{\beta+1} e^{\lambda t_i} \ln t_i$ ,
- *suma44* – v rovnici (5.7)  $\sum_i t_i^\beta e^{\lambda t_i}$

Výpočet probíhá tak, že v prvním kroku se do *sumy1* přičte předchozí výsledek *sumy1* a výsledek  $\frac{1}{b + \lambda \cdot ti(i)}$ , kde *ti(i)* je *i*-tá hodnota z pole *ti*. Obdobně se vypočítávají i *suma2*, *suma3*, *suma4*, *suma11*, *suma22*, *suma33*, *suma44* s tím rozdílem, že dané proměnné se vždy připočítává jiná část rovnice.

```
vysledek1 = suma1+suma2-((n*suma3)/suma4);
vysledek2 = suma11+suma22-((n*suma33)/suma44);
end;
```

Výsledná podoba rovnic je zapsána v proměnných *vysledek1* (reprezentující rovnici 5.6) a *vysledek2* (rovnice 5.7).

### 7.1.5 Výpis výsledků na monitor

```
if(vysledek1<3 & vysledek1 >-3)
    if(vysledek2<3 & vysledek2 >-3)
        j
        k
        vysledek1
        vysledek2
    end;
end;
```

Podmíněné příkazy *if* omezují rozsah vypisovaných výsledků. Parametry  $\lambda$  a  $\beta$  se na monitor zobrazují tehdy, jestliže jsou splněny podmínky pro omezení pravých stran pomocí proměnných *vysledek1* a *vysledek2*. Odhady parametrů  $\lambda$  a  $\beta$  se nevypisovali přímo, ale zjišťovali se pomocí parametrů *j* a *k*. Z nich se stanovily parametry  $\lambda$  a  $\beta$ .

## 7.2 Získané výsledky

Z dat získaných v kapitole 6 jsme obdrželi bodové odhady parametrů s využitím metody popsané v kapitole 7.1.

Výsledky bodových odhadů jsou následovné:

Parametr  $\lambda = 0,086917$

Parametr  $\beta = 0,196$

Výsledky pro tyto parametry nejsou přesně nulové, ale při takovém množství dat jsou brány jako úspěšné. Výsledek pravé strany rovnice (5.6) je roven 0,7955, rovnice (5.7) 0,0916.

## 7.3 Výpočet parametru $\alpha$

Parametr  $\alpha$  modifikovaného Weibullova rozdělení se zjistí pomocí rovnice (5.8) s využitím bodových odhadů parametrů  $\lambda$  a  $\beta$ . Protože stanovení bodového odhadu je obtížné, byl vytvořen skript v softwarovém prostředí Matlab s názvem „alfa.m“, který je uveden v příloze 2.

Zdrojový kód:

```
clear all;
clc;
suma5=0;
b=0.196;           Nastavení vypočteného odhadu parametru  $\beta$ .
lambda=0.086917;  Nastavení vypočteného odhadu parametru  $\lambda$ .
n=100000;         Počet vstupních dat.
load C:\work\ti.txt;   Načtení vstupního souboru dat.
for i=1:n
    suma5=suma5+((ti(i)^b)*exp(lambda*ti(i)));
    vysledek3 = n/suma5;
end;
vysledek3
```

Pomocí proměnné *suma5* se reprezentovalo  $\sum_i t_i^\beta e^{\lambda t_i}$  z rovnice (5.8). A výsledný tvar rovnice byl vypočten pomocí proměnné *vysledek3*. *Vysledek3* vrátí hodnotu parametru  $\alpha$ .

Hodnota parametru  $\alpha = 0,00045186$ .

#### 7.4 Grafické znázornění výsledků

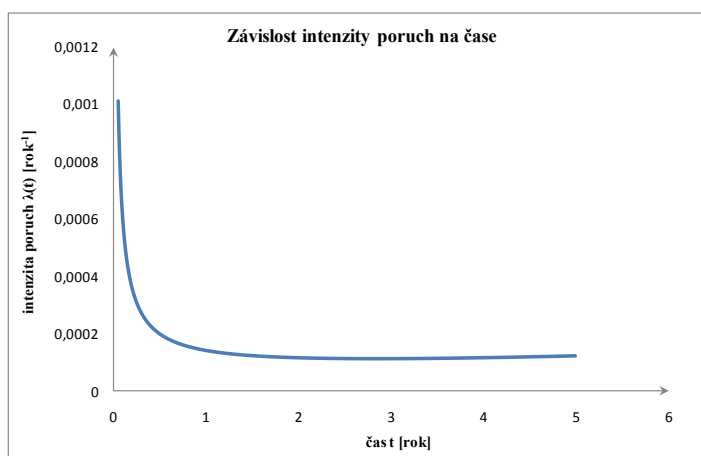
Na Obr. 11 až 14 jsou uvedeny grafické výstupy předpokládané intenzity poruch v čase. Na Obr. 11 je zobrazena intenzita úmrtí 0 až 5 let. Je zřejmé, že intenzita úmrtí v daném období klesá. V raném období je úmrtnost vyšší, protože je způsobena úmrtími při porodu a dětskými nemocemi. Je zde vidět předpokládané první období vanové křivky, kdy intenzita náhodného jevu klesá.

Na Obr. 12 je uveden obdobný graf, kdy je zobrazena intenzita úmrtí až do 20 let. Je zde vidět, že intenzita úmrtí mezi pátým rokem a dvacátým rokem je rostoucí funkce, ale zároveň lze předpokládat, že v daném intervalu je funkce konstantní.

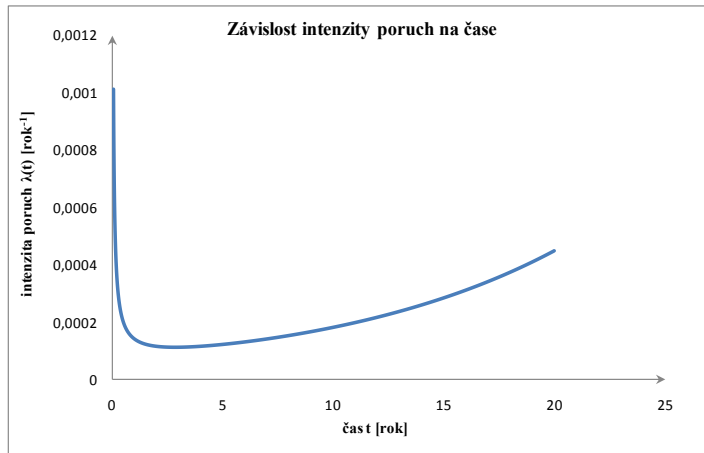
Na Obr. 13 a 14 je zobrazena intenzita úmrtí buď do 50. roku nebo 104. roku života. Intenzita úmrtí v těchto obdobích má výrazní rostoucí charakter. Funkce popisuje třetí období vanové křivky, tedy období degradace.

Z grafů je zřejmé, že dané rozdělení popisuje všechny tři období vanové křivky. Tvar funkce je závislá na vypočtených bodových odhadech parametru  $\alpha, \beta, \lambda$ .

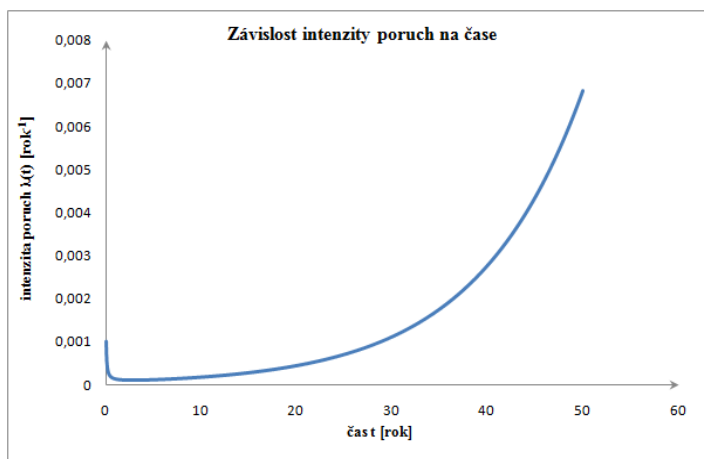
V našem případě byly odhadnuty parametry  $\alpha = 0,00045186$ ,  $\lambda = 0,086917$  a  $\beta = 0,196$ .



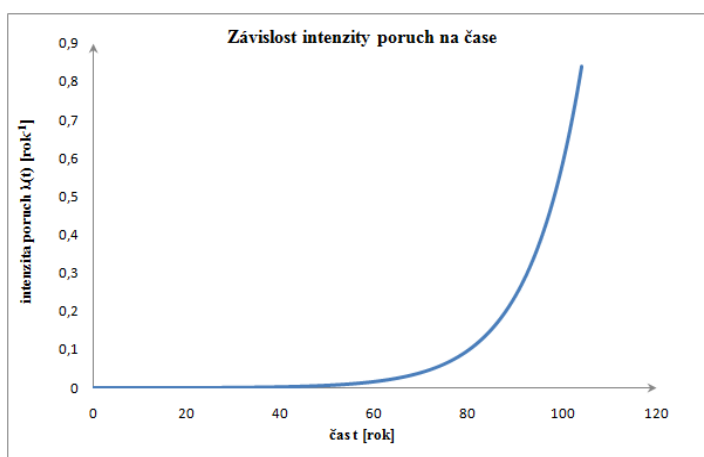
**Obr. 12:** Intenzita poruch modifikované Weibullova rozdělení v období 0. - 5. roku



**Obr. 13:** Intenzita poruch modifikované Weibullova rozdělí v období 0. - 20. roku



**Obr. 14:** Intenzita poruch modifikované Weibullova rozdělí v období 0. - 50. roku



**Obr. 15:** Intenzita poruch modifikované Weibullova rozdělí v období 0. - 104. roku

## 8. TEST DOBRÉ SHODY

### 8.1 Vstupní data pro test dobré shody

Test dobré shody se využívá, jestliže lze předpokládat, že by mohla naměřená data pocházet z určitého předem známého statistického rozdělení. Test dobré shody je vysvětlen v kapitole 2.5.

Zjištěná data o dobách do úmrtí lidí jsou uvedena ve statistických ročenkách [8] a v příloze 1. Část vstupních dat je uvedena v Tab. 2. Ze vstupních dat lze zjistit podíl lidí, kteří zemřou podle dosaženého věku.

Pomocí testu dobré shody se rozdělí data do několika skupin, porovná se skutečný počet poruch s teoretickou hodnotou, podle vypočtených bodových odhadů parametrů teoretického rozdělení a následně se vypočte testová statistika pomocí rovnice (2.7). Jestliže je hodnota testové statistiky menší než  $\chi_{k-h-1}^2$  potom hypotézu, že data pocházejí z daného rozdělení, nezamítáme.

Je popsáno několik pravidel, podle kterých lze rozčlenit vstupní soubor dat do daného počtu dílčích souborů. Jsou to například následující vzorce:  $\sqrt{n}$  nebo  $3,3 \ln n$ . V prvním případě by bylo potřeba mít 316 intervalů, v druhém případě by bylo potřeba mít 38 intervalů. Tab. 2 je rozčleněna na 104 intervalů. Tento počet intervalů lze použít pro analýzu, protože leží mezi počty intervalů vypočítaných podle teoretických vzorců.

Nutná podmínka pro analýzu je, aby v každém intervalu bylo očekáváno alespoň 5 dat podle daného teoretického rozdělení.

V kapitole 7 byly zjištěny bodové odhady parametrů modifikovaného Weibullova rozdělení  $\alpha, \beta, \lambda$ , které jsou následující:

- $\alpha = 0,00045186$ ,
- $\lambda = 0,086917$ ,
- $\beta = 0,196$ .

Pomocí testu dobré shody budou odhadovány tři parametry, tedy  $h$  ve vzorci (2.7) bude rovno 3.

### 8.2 Výpočet

Pro test dobré shody bylo stanoveno, že délka každého intervalu je 1 rok. Celkem byly doby do úmrtí seříděny do 104 intervalů.

**Tab. 4:** Výsledky testu dobré shody

i	$n \cdot \pi_{0,i}$	$n_i$	$\frac{(n_i - n \cdot \pi_{0,i})^2}{n \cdot \pi_{0,i}}$
1	49,28	448	354,53
2	12,29	68	46,09
3	11,14	26	8,66
4	11,20	14	0,42
5	11,71	16	1,24
6	12,47	21	3,57
7	13,42	24	4,35
8	14,54	19	1,05
9	15,81	17	0,15
10	17,23	13	1,15
11	18,82	14	1,44
12	20,57	14	3,11
13	22,50	15	3,32
14	24,63	16	4,76
15	26,97	21	1,54
16	29,54	29	0,00
17	32,36	49	5,75
18	35,45	64	13,08
19	38,83	82	22,91
20	42,54	93	27,42
21	46,61	95	24,30
22	51,05	99	23,06
23	55,92	104	21,91
24	61,25	101	15,93
25	67,07	98	9,67
26	73,44	92	3,67
27	80,41	90	0,93
28	88,02	91	0,13
29	96,34	93	0,09
30	105,43	98	0,51
31	115,36	103	1,57
32	126,21	113	1,44
33	138,04	123	1,86
34	150,96	124	5,85
35	165,05	130	9,17
36	180,42	136	14,37
37	197,17	147	17,11
38	215,43	158	20,54
39	235,31	181	16,41
40	256,96	195	19,43
41	280,52	223	14,55
42	306,14	253	11,31
43	333,98	284	8,61
44	364,22	310	9,42
45	397,05	359	4,10
46	432,65	389	4,94
47	471,22	422	5,69
48	512,97	463	5,39
49	558,13	522	2,54
50	606,90	584	0,88
51	659,52	647	0,23
52	716,21	701	0,34

i	$n \cdot \pi_{0,i}$	$n_i$	$\frac{(n_i - n \cdot \pi_{0,i})^2}{n \cdot \pi_{0,i}}$
53	777,19	761	0,33
54	842,68	813	1,06
55	912,89	903	0,11
56	988,00	994	0,04
57	1068,17	1072	0,01
58	1153,54	1177	0,48
59	1244,19	1277	0,83
60	1340,15	1320	0,30
61	1441,40	1421	0,29
62	1547,81	1507	1,08
63	1659,17	1613	1,30
64	1775,15	1764	0,07
65	1895,29	1884	0,07
66	2018,97	1982	0,70
67	2145,39	2114	0,47
68	2273,56	2207	2,01
69	2402,27	2339	1,74
70	2530,10	2428	4,26
71	2655,35	2538	5,40
72	2776,09	2624	8,86
73	2890,15	2765	5,62
74	2995,11	2878	4,75
75	3088,35	3003	2,40
76	3167,10	3101	1,40
77	3228,48	3186	0,56
78	3269,61	3246	0,18
79	3287,71	3303	0,07
80	3280,21	3336	0,93
81	3244,97	3339	2,64
82	3180,37	3304	4,63
83	3085,52	3234	6,85
84	2960,44	3128	8,99
85	2806,17	2985	10,69
86	2624,92	2805	11,60
87	2420,08	2593	11,49
88	2196,22	2352	10,27
89	1958,93	2089	8,09
90	1714,63	1813	5,34
91	1470,17	1533	2,61
92	1232,51	1260	0,61
93	1008,18	1003	0,02
94	802,83	771	1,32
95	620,84	570	4,61
96	464,97	403	9,58
97	336,26	271	15,47
98	234,08	173	21,27
99	156,29	104	25,92
100	99,71	59	28,58
101	60,53	31	28,87
102	34,80	15	26,89
103	18,86	7	23,22
104	9,58	4	8,07

V Tab. 4 jsou uvedeny výsledky testu dobré shody. V prvním sloupci tabulky je zobrazena  $i$ -tá položka intervalu. V druhém sloupci, označeném  $n \cdot \pi_{0,i}$ , je vypočtena teoretická četnost zjištěná z bodových odhadů parametrů modifikovaného Weibullova rozdělení. V třetím sloupci  $n_i$  je skutečný počet úmrtí v daném intervalu. Poslední čtvrtý sloupec  $\frac{(n_i - n \cdot \pi_{0,i})^2}{n \cdot \pi_{0,i}}$  označuje hodnotu testové statistiky testu dobré shody pro daný interval.

Tab. 4 není nutné upravit, protože v každém intervalu je teoretická četnost větší než 5.

### 8.3 Závěry k výsledkům testu dobré shody

Při zahrnutí všech intervalů je výsledek testu dobré shody roven 1103,5, zatímco tabulková hodnota pro test je  $\chi_{100}^2(0,95) = 124,34$ . Lze tedy zamítnout, že vstupní data pocházejí z daného rozdělení.

Po odstranění krajních intervalů  $t \in (0; 3 > a \ t > 94$  let je výsledek testu dobré shody roven 501,7. Tabulková hodnota pro test  $\chi_{87}^2(0,95) = 107,8$ . Je evidentní, že skutečná hodnota testu se výrazně snížila, tedy dané rozdělení není schopno popsat výrazně snižující se intenzitu úmrtí v intervalu 0 až 3 roky a zároveň výrazně zvyšující se intenzitu v období vyšším než 94 let.

Z výsledků v Tab. 4 je dále vidět, že určité delší intervaly dávají dobré výsledky při popisu pomocí modifikovaného Weibullova rozdělení. Jedná se například o následující intervaly:  $\langle 3; 16 \rangle$ ,  $\langle 27; 33 \rangle$ ,  $\langle 47; 69 \rangle$ .



## 9. ZÁVĚR

Cílem této práce bylo zhodnotit použitelnost modifikovaného Weibullova rozdělení při popisu doby do poruchy komponenty. Jako vstupních dat bylo užito statistických ročenek [8] popisujících dobu do úmrtí člověka. Výhoda těchto dat je, že je jich mnoho a jsou přesná.

Z grafů uvedených na Obr. 12 až 15 je vidět graf intenzity poruch modifikovaného Weibullova rozdělení. Z grafů je zřejmé, že popisuje všechny tři období vanové křivky. Pravděpodobné důvody, proč toto rozdělení nepopisuje vstupní data použita v této práci, jsou následující:

- Vstupních dat je 100 000 a rozdělení má pouze tři parametry.
- Z etických důvodů je snaha co nejvíce oddálit okamžik úmrtí, tím se výrazně zvyšuje intenzita úmrtí ve vyšším věku. Většina technických výrobků se při výrazně zvyšující se intenzitě poruch nahradí za nový.
- Mnoho komplexních technických výrobků je testováno výrobcem na funkčnost, z toho důvodu se neprojevuje tak výrazně první období vanové křivky jako v porovnání s úmrtností v raném věku člověka.

Na základě zjištěných výsledků je autor této práce přesvědčen, že toto rozdělení je schopno popsat popis doby do poruchy velkého množství technických výrobků tam, kde nelze použít jednodušší modely jako například exponenciální či Weibullův.

## 10. LITERATURA

- [1] LAI C. D., XIE M., MURTHY D. N. P., *A Modified Weibull Distribution*, IEEE Transactions on Reliability, Vol. 52 No. 1, 2003 [cit. 22. prosince 2008].
- [2] KOPŘIVA P., LAUSCHMANN H., *Vybrané problémy teorie spolehlivosti*. Praha. 1990 [cit. 11. října 2008].
- [3] NG H. K. T., *Parameter Estimation for a Modified Weibull Distribution for Progressively Type-II Censored Samples*, IEEE Transactions on Reliability, Vol. 54 No. 3, 2005 [cit. 22. prosince 2008].
- [4] BRIŠ R., *Inovační metody pro ocenění spolehlivosti prvků a systémů*. Monografie 1. vydání. VŠB-Technická univerzita Ostrava. Ostrava. 2007. ISBN 978-80-248-1596-1 [cit. 10. listopadu 2008].
- [5] BRIŠ R., LITSCHMANNOVÁ M., *Statistika II.* Ostrava. 2007, první vydání [cit. 2. prosince 2008].
- [6] FUCHS P., VALIŠ D., *Metody analýzy a řízení rizika*. Liberec. 2004 [cit. 7. ledna 2009].
- [7] AARSET M. V., *How to Identify Bathtub Hazard Rate*, IEEE Transactions on Reliability. 1987 [cit. 21. listopadu 2008].
- [8] *Tab. 7 2001-2002 Podrobné úmrtnostní tabulky - město – muži*. [online]. 2008 [cit. 4. prosince 2008]. Dostupné na WWW: <http://www.czso.cz/csu/2003edicniplan.nsf/p/4022-03>.
- [9] BRIŠ R., LITSCHMANNOVÁ M., *STATISTIKA I.* [online]. [cit. 4. 12. 2008]. Dostupné na WWW: <http://www.am.vsb.cz/litschmannova/STA1/Skripta/PDF/>.
- [10] FUCHS P. *Využití spolehlivosti v provozní praxi*. Technická univerzita v Liberci [cit. 13. října 2008].
- [11] HOLUB R., VINTR Z., *Spolehlivost letadlové techniky (elektronická učebnice)*. [online]. Brno. 2007 [cit. 7. ledna 2009]. Dostupné na WWW: <http://lu.fme.vutbr.cz/files/SpolehlivostLetadloveTechniky.pdf>.
- [12] CHUDOBA J., *Zpracování dat o poruchách z provozu*. Technická univerzita v Liberci. Liberec. 2006 [cit. 1. prosince 2008]. Dostupné na WWW: <http://flow.kmo.tul.cz/~www/czech/seminare/2006-03-09-chudoba.pdf>.