

**ZÁPADOČESKÁ UNIVERZITA V PLZNI
FAKULTA ELEKTROTECHNICKÁ**

KATEDRA ELEKTROENERGETIKY A EKOLOGIE

DIPLOMOVÁ PRÁCE

Statistické vyhodnocení dat v technice vysokého napětí

František Kováč

2015

ZÁPADOČESKÁ UNIVERZITA V PLZNI
Fakulta elektrotechnická
Akademický rok: 2014/2015

ZADÁNÍ DIPLOMOVÉ PRÁCE
(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení: **Bc. František KOVÁČ**
Osobní číslo: **E13N0075P**
Studijní program: **N2612 Elektrotechnika a informatika**
Studijní obor: **Elektroenergetika**
Název tématu: **Statistické vyhodnocení dat v technice vysokého napětí**
Zadávací katedra: **Katedra elektroenergetiky a ekologie**

Z á s a d y p r o v y p r a c o v á n í :

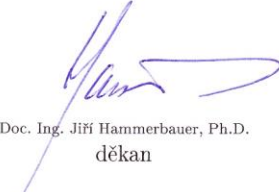
1. Vypracujte odborný text věnující se těm částem z oblasti statistické matematiky, které je potřeba znát a uplatnit při vyhodnocování experimentálně získaných dat z oblasti techniky vysokých napětí.
2. V odborném textu práce se zabývejte jednotlivými vysokonapěťovými testovacími metodami (testovací procedura, vlastnosti získaných dat), které se používají při experimentálních a zkušebních testech.
3. Vytvořte aplikaci určenou pro statistické vyhodnocení naměřených dat. Vlastnosti a možnosti této aplikace navrhňte tak, aby vyhovovaly potřebám výzkumu v oblasti oboru techniky vysokého napětí.
4. Ve své práci prezentujte, jako výstup vytvořené aplikace, výsledky statistického vyhodnocení vybraných experimentálně získaných dat.

Rozsah grafických prací: podle doporučení vedoucího
Rozsah pracovní zprávy: 30 - 40 stran
Forma zpracování diplomové práce: tištěná/elektronická
Seznam odborné literatury:

1. W. Mosch, W. Hauschild: Statistical Techniques for High-Voltage Engineering, The Institution of Engineering and Technology, Lightning Source UK Ltd., Milton Keynes UK, Reprinted 2007 (First published in German by VEB Verlag Technik, Berlin 1984)
2. ČSN EN 60060-1, Technika zkoušek vysokým napětím - Část 1: Obecné definice a požadavky na zkoušky

Vedoucí diplomové práce: Ing. Petr Martínek, Ph.D.
Katedra elektroenergetiky a ekologie

Datum zadání diplomové práce: 15. října 2014
Termín odevzdání diplomové práce: 11. května 2015


Doc. Ing. Jiří Hammerbauer, Ph.D.
děkan




Doc. Ing. Karel Noháč, Ph.D.
vedoucí katedry

V Plzni dne 15. října 2014

Abstrakt

Předkládaná diplomová práce je zaměřena na metody statistického vyhodnocování a jejich aplikace na měřicí metody využívané v technice vysokého napětí. Pro potřeby laboratoří vysokého napětí byl vytvořen program pro statistické vyhodnocování zkoušek více hladinami a postupným namáháním.

Klíčová slova

Normální rozdělení, Laplaceovo rozdělení, statistika, technika vysokého napětí, zkouška více hladinami, zkouška postupným namáháním, aplikace, Visual Studio, C#

Abstract

This diploma thesis is focused on the methods of statistical evaluation and their application on the methods of measurement used in the high voltage engineering. For the needs of high voltage laboratories, was created software, which can statistically evaluate constant voltage test and rising voltage test.

Key words

Normal distribution, Laplace distribution, Statistics, High Voltage Engineering, Constant Voltage Test, Rising Voltage Test, Application, Visual Studio, C#

Prohlášení

Prohlašuji, že jsem tuto diplomovou práci vypracoval samostatně, s použitím odborné literatury a pramenů uvedených v seznamu, který je součástí této diplomové práce.

Dále prohlašuji, že veškerý software, použitý při řešení této diplomové práce, je legální.

V Plzni dne 5.5.2015

František Kováč

.....

podpis

Poděkování

Tímto bych rád poděkoval vedoucímu diplomové práce Ing. Petru Martínkovi, Ph.D za cenné profesionální rady během konzultací, odborné připomínky a metodické vedení práce.

Obsah

OBSAH	1
SEZNAM OBRÁZKŮ, TABULEK A GRAFŮ	2
SEZNAM SYMBOLŮ A ZKRATEK	3
ÚVOD	5
1 STATISTIKA	6
1.1 ZÁKLADNÍ PRINCIPY STATISTIKY	6
1.1.1 <i>Náhodný pokus a náhodný jev</i>	6
1.1.2 <i>Relativní frekvence a pravděpodobnost</i>	6
1.2 DISTRIBUČNÍ FUNKCE	7
1.2.1 <i>Odhad parametrů</i>	11
1.2.2 <i>Binomické rozdělení</i>	13
1.2.3 <i>Normální rozdělení</i>	16
1.2.4 <i>Laplaceovo rozdělení</i>	18
1.3 LINEÁRNÍ REGRESE	20
1.3.1 <i>Metoda nejmenších čtverců</i>	20
1.3.2 <i>Metoda pravděpodobnostního papíru</i>	21
1.4 KOLMOGOROVŮV TEST	22
2 METODY MĚŘENÍ	23
2.1 ZKOUŠKA VÍCE HLADINAMI	23
2.1.1 <i>Ověření nezávislosti</i>	24
2.1.2 <i>Statistické vyhodnocení měření</i>	25
2.2 ZKOUŠKY POSTUPNÝM NAMÁHÁNÍM	26
2.2.1 <i>Ověření nezávislosti</i>	28
2.2.2 <i>Statistické vyhodnocení měření</i>	28
3 NÁSTROJ PRO STATISTICKÉ VYHODNOCOVÁNÍ TESTŮ	30
3.1 SOFTWARE	30
3.2 PROGRAM	30
3.2.1 <i>Interface programu</i>	30
3.2.2 <i>Intervaly spolehlivosti</i>	33
3.2.3 <i>Multi Fit</i>	34
3.3 VZOROVÁ EXPERIMENTÁLNÍ MĚŘENÍ	34
ZÁVĚR	37
SEZNAM LITERATURY A INFORMAČNÍCH ZDROJŮ	38
PŘÍLOHY	39

Seznam obrázků, tabulek a grafů

OBR. 1.1 ZÁVISLOST RELATIVNÍ FREKVENCE NA POČTU PROVEDENÝCH TESTŮ [1]	7
OBR. 1.2 DISKRÉTNÍ HUSTOTA PRAVDĚPODOBNOSTI [1].....	9
OBR. 1.3 DISKRÉTNÍ DISTRIBUČNÍ FUNKCE [1]	9
OBR. 1.4 DISTRIBUČNÍ FUNKCE [1].....	10
OBR. 1.5 HUSTOTA PRAVDĚPODOBNOSTI [1]	11
OBR. 1.6 ILUSTRACE PARAMETRŮ [1].....	13
OBR. 1.7 BINOMICKÉ ROZDĚLENÍ S PARAMETRY $p=0,2$ A $n=10$	14
OBR. 1.8 INTERVALY SPOLEHLIVOSTI DLE NEUMANA S KOEFICIENTEM SPOLEHLIVOSTI $e=0,95$ [1]	15
OBR. 1.9 DISTRIBUČNÍ FUNKCE (PLNÁ ČÁRA) A FUNKCE HUSTOTY PRAVDĚPODOBNOSTI (ČERCHOVANÁ ČÁRA) STANDARDNÍHO NORMÁLNÍHO ROZDĚLENÍ ($\mu=0; \sigma^2=1$) [1]	17
OBR. 1.10 DISTRIBUČNÍ FUNKCE A HUSTOTA PRAVDĚPODOBNOSTI LAPLACEOVA ROZDĚLENÍ S PARAMETRY $\gamma = 1, \eta = 0$ [1].....	19
OBR. 1.11 KORELAČNÍ KOEFICIENT PRO RŮZNÉ SOUBORY DAT [1]	21
OBR. 1.12 VYTVOŘENÍ PRAVDĚPODOBNOSTNÍHO PAPIRU PRO NORMÁLNÍ ROZDĚLENÍ [1]	21
OBR. 2.1 GRAF ZÁVISLOSTI INTERVALŮ SPOLEHLIVOSTI NA POČTU TESTOVACÍCH IMPULZŮ n [1]	24
OBR. 2.2 A) VHODNÉ PRO PROLOŽENÍ DISTRIBUČNÍ FUNKCÍ; B) NEVHODNÉ PRO PROLOŽENÍ DISTRIBUČNÍ FUNKCÍ [1].....	26
OBR. 2.3 KROKOVÉ ZVYŠOVÁNÍ NAPĚTÍ [1].....	26
OBR. 2.4 PLYNULÉ ZVYŠOVÁNÍ NAPĚTÍ [1].....	27
OBR. 2.5 GRAF PRO VYHODNOCENÍ ZÁVISLOSTI DAT [1]	28
OBR. 3.1 HLAVNÍ NABÍDKA	30
OBR. 3.2 NASTAVENÍ.....	30
OBR. 3.3 VÝBĚR MÍSTA ROZDĚLENÍ ZADANÝCH DAT	31
OBR. 3.4 INTERFACE PRO ZKOUŠKU VÍCE HLADINAMI S JEDNODUCHÝM PROLOŽENÍM.....	31
OBR. 3.5 PODROBNÉ VÝSLEDKY	32
OBR. 3.6 INTERFACE PRO ZKOUŠKU POSTUPNÝM NAMÁHÁNÍM S DVOJITÝM PROLOŽENÍM.....	32
OBR. 3.7 PRAVDĚPODOBNOSTNÍ GRAF, MĚŘENÍ 1	35
TAB. 1.1 HODNOTY Y PRO NORMÁLNÍ ROZDĚLENÍ	21
TAB. 1.2 VZORCE PRO VÝPOČET KRITICKÉ HODNOTY PRO JEDNOTLIVÉ HLADINY	22
TAB. 2.1 ILUSTRACE ZÁZNAMU ZKOUŠKY VÍCE HLADINAMI [1]	23
TAB. 2.2 MĚŘENÍ PŘI ATMOSFÉRICKÝCH PODMÍNKÁCH (NEZÁVISLÉ)	24
TAB. 2.3 MĚŘENÍ V UZAVŘENÉ NÁDOBĚ (ZÁVISLÉ).....	25
TAB. 2.4 ZAPSANÍ VÝSLEDKŮ MĚŘENÍ VÍCE HLADINAMI.....	25
TAB. 2.5 ZAPSANÍ VÝSLEDKŮ ZKOUŠKY POSTUPNÝM NAMÁHÁNÍM.....	29
TAB. 3.1 NAMĚŘENÉ HODNOTY, ZKOUŠKA VÍCE HLADINAMI , MĚŘENÍ 1	34
TAB. 3.2 SUBSTITUOVANÉ HODNOTY	35
TAB. 3.3 NAMĚŘENÉ HODNOTY, MĚŘENÍ 2	35
TAB. 3.4 POROVNÁNÍ VÝSLEDKŮ, MĚŘENÍ 2	36
TAB. 3.5 NAMĚŘENÉ HODNOTY, MĚŘENÍ 3	36
TAB. 3.6 POROVNÁNÍ VÝSLEDKŮ, MĚŘENÍ 3	36

Seznam symbolů a zkratk

C	Eulerova konstanta
$D(X)$	Rozptyl náhodné proměnné X
$E(X)$	Střední hodnota náhodné proměnné X
$F(x)$	Distribuční funkce náhodné proměnné X
$F_{m_1; m_2; q}$	Kvantil F rozdělení (řádu q ; stupňů volnosti m_1 a m_2)
$f(x)$	Funkce hustoty pravděpodobnosti náhodné proměnné X
g_u	Spodní hodnota intervalu spolehlivosti
g_o	Horní hodnota intervalu spolehlivosti
h	Relativní frekvence
$h_{\Sigma k}$	Kumulativní relativní frekvence
k	Počet pozorovaných událostí
m	Počet pozorovaných hodnot; počet napět'ových hladin
m_k	Empirický k -tý moment
n	Velikost vzorku, počet nezávislých pokusů
P	Pravděpodobnost jevu
R	Variační rozpětí
r	Korelační koeficient
s	Výběrová směrodatná odchylka
$t_{m; q}$	Kvantil Studentova t rozdělení (řádu q ; stupně volnosti m)
u_0	Počáteční napětí
v	Variační koeficient
v_u	Rychlost zvyšování napětí
X	Náhodná proměnná
$X_{m; q}^2$	Kvantil Chí-kvadrát rozdělení (řádu q ; stupně volnosti m)
α	Hladina významnosti
γ	Rozptyl Laplaceovo rozdělení
$\Delta t, \Delta t_p$	Doba mezi jednotlivými testy
Δt_s	Doba trvání kroku
Δt_z	Doba zvyšování napětí mezi úrovněmi
Δu	Napět'ový krok

Δu_r	Rozsah testovaného napětí
ε	Koeficient spolehlivosti
η	63% kvantil; parametr Laplaceova rozdělení
λ_q	Kvantil Normálního rozdělení
μ	střední hodnota, parametr normálního rozdělení
σ	Rozptyl, parametr normálního rozdělení
$\Phi(x)$	Distribuční funkce
$\varphi(x)$	Funkce hustoty pravděpodobnosti
CIL	Common Intermediate Language
CDF	Kumulativní distribuční funkce
invCDF	Inverzní kumulativní distribuční funkce

Úvod

Tato práce se zabývá problematikou vysokonapěťových měření. Většina vysokonapěťových dějů nelze popsat standardními vzorci, jelikož se jedná o náhodné děje. Z toho důvodu se využívají k popisu těchto jevů různá pravděpodobnostní rozdělení. V práci jsou popsány postupy vedoucí k získání daných rozdělení a jejich parametrů z naměřených hodnot.

Práce je rozdělena do tří částí. První část se zabývá distribučními funkcemi, odhady parametrů a matematickým aparátem potřebným k aplikaci na získaná data z měření. Druhá část se zabývá měřicími metodami využívanými v technice vysokého napětí včetně volby parametrů měření, metod ověření nezávislosti naměřených dat a postupů pro vyhodnocení měření. V třetí části je popsán vytvořený program, který aplikuje popsané metody a řeší se zde problematika zkoušky více hladinami.

1 Statistika

1.1 Základní principy statistiky

1.1.1 Náhodný pokus a náhodný jev

Náhodný pokus je experiment, jehož výsledek je při stejných počátečních podmínkách nejistý. Výsledek nejsme schopni přesně předpovědět. Množina všech možných výsledků náhodného pokusu se označuje Ω [1][7][9].

Během pokusu nastávají náhodně procesy, díky kterým je výsledkem pokusu náhodný jev. Náhodný jev je podmnožina množiny Ω a označuje se velkým písmenem např. A, B, C atd. [1][7][9].

1.1.2 Relativní frekvence a pravděpodobnost

Pokud se pokus opakuje n -krát a náhodný jev A nastane m -krát, můžeme získat relativní frekvenci náhodného jevu A jako

$$h_n(A) = \frac{m}{n}. \quad (1.1)$$

V praxi, kde je množství n pokusů omezeno časem se používá vztah

$$h_n(A) = \frac{m}{n+1}, \quad (1.2)$$

který se využívá v případě nízkého počtu n . Pro n větší než 50 se využívá rovnice (1.1) [1][7][9].

Jelikož vždy platí $0 \leq m \leq n$. Relativní frekvence může nabývat pouze hodnot

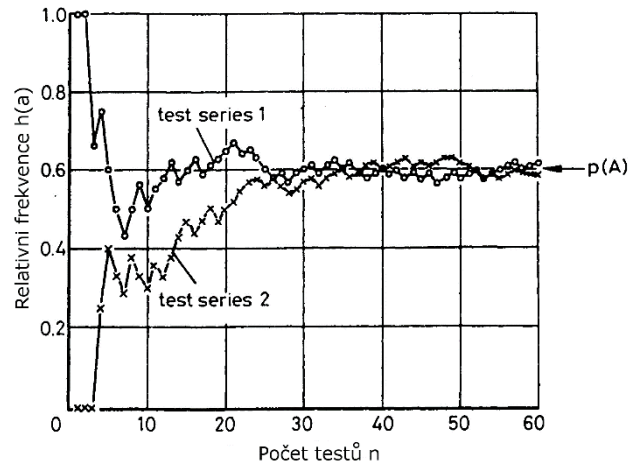
$$0 \leq h_n \leq 1. \quad (1.3)$$

Z toho vyplývá, že $h_n(I) = 1$ je jev I jistý a $h_n(P) = 0$ je jev P nemožný. V technice vysokého napětí se především využívá vztahu dvou komplementárních jevů A a \bar{A}

$$h_n(\bar{A}) = 1 - h_n(A). \quad (1.4)$$

Pokud testovací série obsahující n náhodných pokusů je opakována s různým počtem n náhodných pokusů. Výsledná relativní frekvence se bude měnit v závislosti na počtu n

náhodných pokusů. Při porovnání dvou testovacích sérií si můžeme všimnout, že se rozdíl mezi jednotlivými relativními frekvencemi postupně zmenšuje. Jak je znázorněno na Obr. 1.1 a postupně se bude blížit k dané hodnotě pravděpodobnosti [1][7][9].



Obr. 1.1 Závislost relativní frekvence na počtu provedených testů [1]

$$\lim_{x \rightarrow \infty} h_n(A) = p(A) \quad (1.5)$$

Pokud se n bude blížit k nekonečnu, získáme pravděpodobnost náhodného jevu. Jelikož v praxi nelze dosáhnout tak vysokých hodnot n , je tato pravděpodobnost odhadována pomocí relativní frekvence [1][7][9].

1.2 Distribuční funkce

Distribuční funkce je funkce, která v bodě x udává pravděpodobnost, s níž proměnná X bude mít hodnotu menší než x . Je definována jako [1][7][9]

$$F(x) = P(X < x). \quad (1.6)$$

Z její definice vyplývají určité vlastnosti:

- I. Distribuční funkce je nezáporné číslo menší nebo rovno jedné

$$0 \leq F(x) \leq 1. \quad (1.7)$$

II. Distribuční funkce je funkce neklesající

$$F(x_1) \leq F(x_2). \quad (1.8)$$

III. Distribuční funkce je spojitá zprava

$$\lim_{x \rightarrow -\infty} F(x) = F(-\infty) = 0, \quad (1.9)$$

$$\lim_{x \rightarrow \infty} F(x) = F(+\infty) = 1. \quad (1.10)$$

IV. Každá distribuční funkce definovaná rovnicí (1.6) je spojitá zleva.

Proto se každá funkce, která splňuje tyto pravidla, dá označit jako distribuční funkce. Z praktického hlediska se využívají jen určité empirické distribuční funkce. Rozlišujeme mezi distribučními funkcemi diskrétními a spojitými [1][7][9].

Distribuční funkci diskrétní proměnné (Obr. 1.3) můžeme zapsat

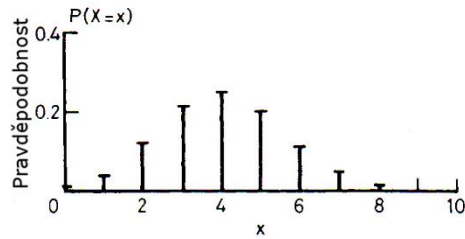
$$F(x) = F(X < x) = \sum_{x_i < x} P(X = x_i) = \sum_{x_i < x} p_i, \quad (1.11)$$

kde x_i jsou diskrétní hodnoty, kterých může proměnná dosáhnout a p_i jsou jejich dané pravděpodobnosti. Místo diskrétní distribuční funkce můžeme využít diskrétní funkci hustoty pravděpodobnosti danou (Obr. 1.2)

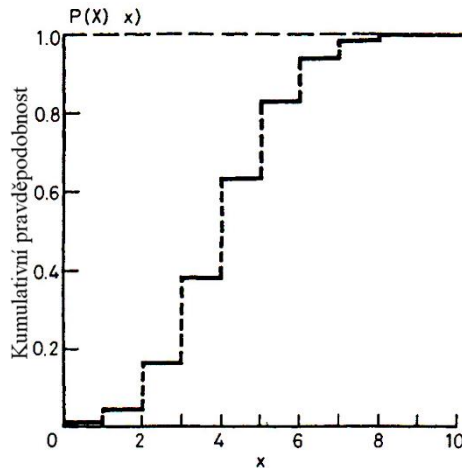
$$f(x) = P(X = x_i) = p_i, \quad (1.12)$$

nebo pravděpodobnostní tabulku (1.13). Při výpočtech je často výhodné pracovat s pravděpodobnostmi z tabulky a ne s jednotlivými funkcemi [1][7][9].

$$X: \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 0,006 & 0,04 & 0,121 & 0,215 & 0,251 & 0,201 & 0,112 & 0,043 & 0,011 & 0,002 & 0 \end{pmatrix} \quad (1.13)$$



Obr. 1.2 Diskrétní hustota pravděpodobnosti [1]



Obr. 1.3 Diskrétní distribuční funkce [1]

U distribučních funkcí se sledují především dva parametry, a to jsou střední hodnota a rozptyl. Střední hodnota je dána jako

$$\mu = E(X) = \sum_i x_i p_i \quad (1.14)$$

a rozptyl

$$\sigma = D(X) = \sum_i (x_i - \mu)^2 p_i. \quad (1.15)$$

Distribuční funkce spojitého rozdělení je zapsána

$$F(x) = P(X < x) = \int_{-\infty}^x f(t) dt, \quad (1.16)$$

kde výraz

$$f(x) = \frac{d}{dx}F(x) \quad (1.17)$$

je nazýván jako hustota pravděpodobnosti

$$F(x) \geq 0 \quad (1.18)$$

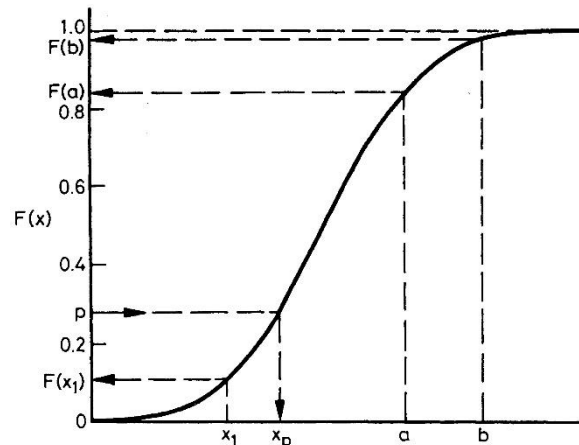
pro všechny x , a

$$\int_{-\infty}^{+\infty} f(x)dx = 1. \quad (1.19)$$

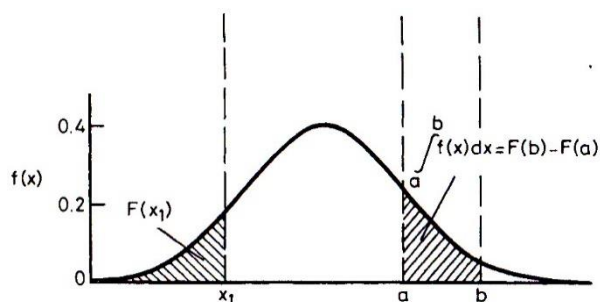
Každá funkce, která má tyto vlastnosti, může být hustota pravděpodobnosti spojité proměnné. V technice vysokého napětí se spíše upřednostňuje distribuční funkce namísto funkce hustoty pravděpodobnosti [1][7][9].

Pravděpodobnost spojité proměnné X v intervalu $[a,b]$ je dána

$$P(a \leq X \leq b) = \int_a^b f(x)dx = F(b) - F(a). \quad (1.20)$$



Obr. 1.4 Distribuční funkce [1]



Obr. 1.5 Hustota pravděpodobnosti [1]

1.2.1 Odhad parametrů

Parametry se často využívají k popisu distribučních funkcí. Musí se však rozlišovat mezi parametry obecnými (popisující soubor dat), jako střední hodnota, rozptyl atd. a parametry vyskytujícími se ve vzorci pro dané teoretické rozdělení. Mezi těmito druhy parametrů je blízký vztah a v některých případech jsou tyto parametry identické [1][7][9].

Základní obecné parametry jsou popsány v kapitole 1.2. Další parametry jsou odvozené pomocí momentů, které vycházejí z rovnice [1][7][9]

$$m_k = \frac{1}{n} \sum_{i=1}^n (x_i - a)^k, \quad (1.21)$$

kde $k = 1, 2, 3, \dots$

Následující parametry mohou být použity jako střední hodnota:

- i. Aritmetický průměr

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i. \quad (1.22)$$

- ii. Medián \tilde{x} , je hodnota dělící seřazený soubor dat na dvě stejné skupiny. Pokud celkový počet hodnot je sudý, určí se medián jako aritmetický průměr hodnot na pozici $n/2$ a $n/2 + 1$.
- iii. Modus \hat{x} , je hodnota, která se vyskytuje nejčastěji v souboru dat. Lze určit z funkce hustoty pravděpodobnosti, kde tvoří její vrchol. U rozdělení, která mají pouze jeden vrchol např. normálové rozdělení se modus rovná mediánu.

iv. Harmonický průměr

$$\bar{x} = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}}. \quad (1.23)$$

Následující parametry mohou být použity jako rozptyl:

i. Variační rozpětí

$$R = x_{max} - x_{min}, \quad (1.24)$$

u malého počtu dat (<5) je spojováno s rozptylem. Při větším počtu dat ($10 \leq n \leq 100$) se začne vyjadřovat směrodatnou odchylkou s , vztahem

$$R \approx 4s. \quad (1.25)$$

ii. Rozdíl kvantilů

$$I = x_p - x_q, \quad (1.26)$$

standardně se využívá rozdíl mezi 90% a 10% kvantilem. Pomocí rozdílu kvantilů lze také odhadnout směrodatnou odchylku jako

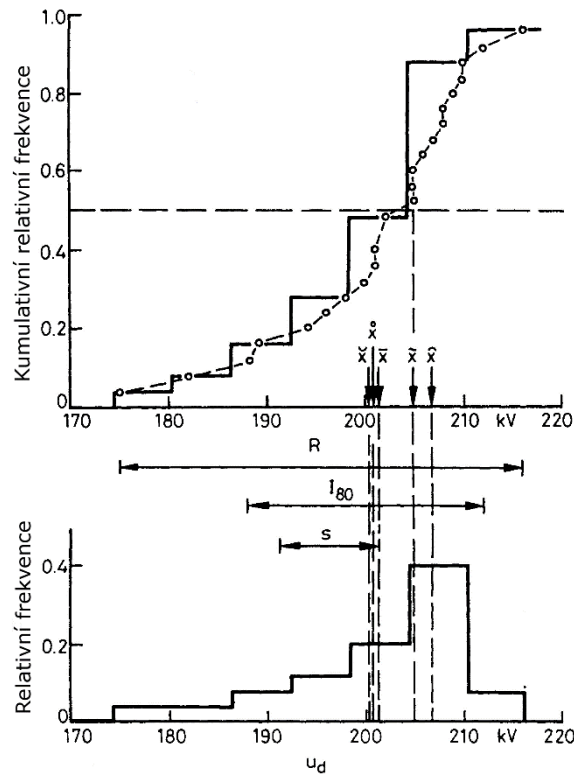
$$s = I_{50-16} = x_{50} - x_{16}. \quad (1.27)$$

iii. Výběrová směrodatná odchylka

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n x_i^2 - n\bar{x}^2} \quad (1.28)$$

Pro rychlé porovnání souborů dat se zavedl variační koeficient v , který se počítá jako podíl střední hodnoty a rozptylu a vyjadřuje fluktuaci naměřených dat [1][7][9]

$$v = \frac{s}{\bar{x}}. \quad (1.29)$$



Obr. 1.6 Ilustrace parametrů [1]

1.2.2 Binomické rozdělení

Binomické rozdělení popisuje četnost výskytu náhodného jevu v n nezávislých pokusech, kde má jev stále stejnou pravděpodobnost výskytu. Jedná se o nespojitě rozdělení [1][7][9].

Pravděpodobnost, že jev nastane k -krát z n pokusů při dané pravděpodobnosti je dána rozdělením (hustota pravděpodobnosti)

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}, \quad (1.30)$$

kumulativní pravděpodobnost (distribuční funkce) je definována jako [1][7][9]

$$P(X \leq k) = \sum_{m=0}^k \binom{n}{m} p^m (1 - p)^{n-m}. \quad (1.31)$$

Parametry: n - počet nezávislých pokusů
 p - pravděpodobnost výskytu jevu A

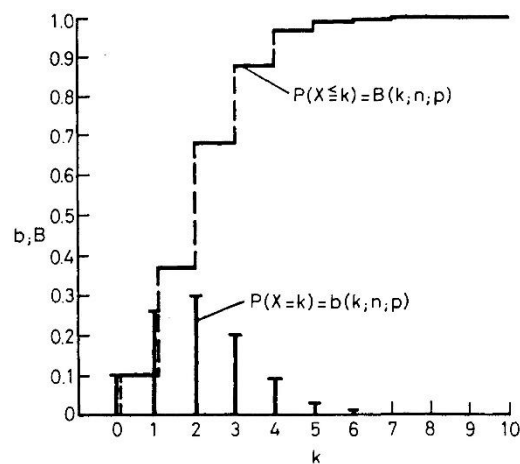
Střední hodnota je

$$E(X) = np. \quad (1.32)$$

Rozptyl je

$$D(X) = np(1 - p). \quad (1.33)$$

Pravděpodobnost jednotlivých hodnot odvozujeme pomocí relativní frekvence, viz kapitola 1.1.2 [1][7][9].



Obr. 1.7 Binomické rozdělení s parametry $p=0,2$ a $n=10$

Získaná relativní frekvence je pouze bodový odhad, který je pouze jedním číslem. Namísto toho interval spolehlivosti je interval, ve kterém by měla daná charakteristika ležet s určitou pravděpodobností určenou koeficientem spolehlivosti ε . Z toho důvodu pro bodové odhady počítáme intervaly spolehlivosti. K intervalům spolehlivosti pro pravděpodobnost binomického rozdělení můžeme přistupovat více způsoby [1][5][6]:

- i. Pomocí F rozdělení (Fisherovo-Snedecorovo rozdělení) získáme jednostranný interval spolehlivosti. Pro horní limitní hodnotu $[-\infty; p_o]$ platí

$$p_o = \frac{(k + 1)F_{m_1; m_2; \varepsilon}}{n - k + (k + 1)F_{m_1; m_2; \varepsilon}}, \quad (1.34)$$

kde $F_{m_1; m_2; \varepsilon}$ je kvantil F rozdělení řádu $q = \varepsilon$ (koeficient spolehlivosti) a stupni volnosti $m_1 = 2(k + 1)$ a $m_2 = 2(n - k)$ ($i = [0; p_o]$). Nižší

limitní hodnotu $[p_u; +\infty]$ získáme

$$p_u = \frac{k}{k + (n - k + 1)F_{m_3; m_4; \varepsilon}}, \quad (1.35)$$

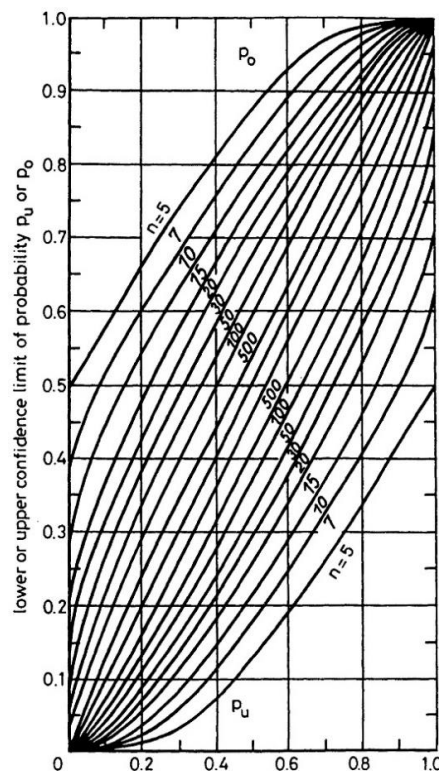
kde $m_3 = 2(n - k + 1)$ a $m_4 = 2k$ ($i = [p_u; 1]$). F rozdělení je blíže popsáno v [1] [15].

- ii. Oboustranný interval spolehlivosti $[p_u; p_o]$ lze získat z tabulkových hodnot pro hodnoty $n = 1$ až $n = 30$. Pro větší hodnoty lze interval spolehlivosti získat z

$$\frac{p_u}{p_o} = \frac{1}{n + \lambda_q^2} \left(k + \frac{\lambda_q^2}{2} \mp \lambda_q \sqrt{\frac{k(n - k)}{n} + \frac{\lambda_q^2}{4}} \right), \quad (1.36)$$

kde λ_q je kvantil normalizovaného normálního rozdělení řádu $q = \frac{1+\varepsilon}{2}$.

Hodnoty intervalů spolehlivosti v závislosti na n jsou znázorněny v Obr. 1.8[1][7][9].



Obr. 1.8 Intervaly spolehlivosti dle Neumana s koeficientem spolehlivosti $\varepsilon=0,95$ [1]

1.2.3 Normální rozdělení

Náhodný proces se řídí dle normálního rozdělení v případě, kdy vzniká vlivem velkého počtu nezávislých jevů, a to pouze pokud je vliv jednotlivých efektů zanedbatelný [1][7][9].

Funkce hustoty pravděpodobnosti

$$\varphi(x; \mu; \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x-\mu}{2\sigma^2}}. \quad (1.37)$$

Distribuční funkce (kumulativní pravděpodobnost)

$$\Phi(x; \mu; \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x e^{-\frac{(t-\mu)^2}{2\sigma^2}}. \quad (1.38)$$

Parametry: μ - střední hodnota, medián, modus
 σ^2 - rozptyl, směrodatná odchylka

Střední hodnota je

$$E(X) = \mu. \quad (1.39)$$

Rozptyl je

$$D(X) = \sigma^2. \quad (1.40)$$

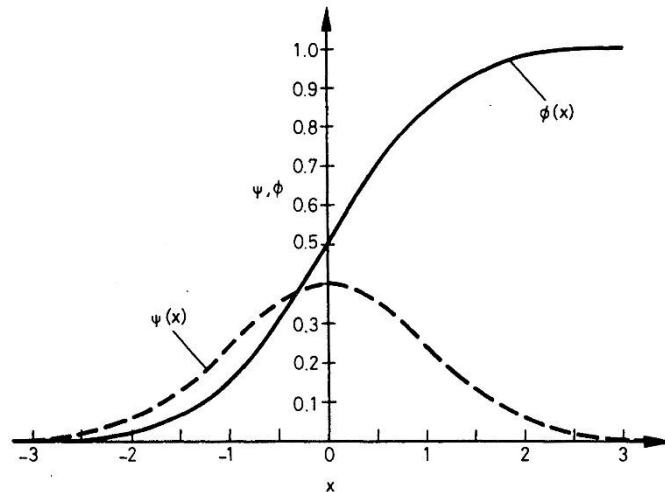
Odhad střední hodnoty μ

$$\mu^* = \frac{1}{n} \sum_{i=1}^n x_i = x_{50}^*. \quad (1.41)$$

Odhad rozptylu σ^2

$$\sigma^* = s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = (x_{50}^* - x_{16}^*)^2. \quad (1.42)$$

Hodnoty x_{50}^* a x_{16}^* jsou kvantily distribuční funkce, které můžeme využít k odhadu parametrů v případě využití regrese [1][7][9].



Obr. 1.9 Distribuční funkce (plná čára) a funkce hustoty pravděpodobnosti (čerchovaná čára) standardního normálního rozdělení ($\mu=0; \sigma^2=1$) [1]

Intervaly spolehlivosti jsou u normálního rozdělení počítány pro střední hodnotu a rozptyl. Intervaly spolehlivosti pro střední hodnotu můžeme počítat jako dva jednostranné intervaly nebo jeden oboustranný interval [1][6]:

- i. Horní hodnota jednostranného intervalu $[-\infty; g_o]$ je dána

$$g_o = \bar{x} + t_{n-1; \varepsilon} \frac{s}{\sqrt{n}} \quad (1.43)$$

spodní hodnota $[g_u; +\infty]$ je

$$g_u = \bar{x} - t_{n-1; \varepsilon} \frac{s}{\sqrt{n}} \quad (1.44)$$

- ii. Oboustranný interval spolehlivosti $[g_u; g_o]$ je dán vztahy

$$g_u = \bar{x} - t_{n-1; (1+\varepsilon)/2} \frac{s}{\sqrt{n}} \quad (1.45)$$

a

$$g_o = \bar{x} + t_{n-1; (1+\varepsilon)/2} \frac{s}{\sqrt{n}} \quad (1.46)$$

Oba způsoby výpočtu intervalů spolehlivosti pro střední hodnotu využívají Studentova t rozdělení. Studentovo t rozdělení je blíže popsáno v [1] [15].

Na rozdíl od výpočtu střední hodnoty se rozptyl počítá pomocí chí-kvadrát rozdělení [1][6].

- i. Horní hodnota jednostranného intervalu $[-\infty; g_o]$ je dána

$$g_o = \frac{(n-1)s^2}{X_{n-1;1-\varepsilon}^2}, \quad (1.47)$$

spodní hodnota je

$$g_u = \frac{(n-1)s^2}{X_{n-1;\varepsilon}^2}. \quad (1.48)$$

- ii. Oboustranný interval spolehlivosti $[g_u; g_o]$ je dán vztahy

$$g_u = \frac{(n-1)s^2}{X_{n-1;(1+\varepsilon)/2}^2} \quad (1.49)$$

a

$$g_o = \frac{(n-1)s^2}{X_{n-1;(1-\varepsilon)/2}^2}. \quad (1.50)$$

1.2.4 Laplaceovo rozdělení

Laplaceovo nebo také dvojitě exponenciální rozdělení patří do skupiny extrémních rozdělení. Základní model je stejný jako u Weibullova rozdělení [1][7][9].

Funkce hustoty pravděpodobnosti

$$f_d(x) = \frac{1}{\gamma} \exp\left(\frac{x-\eta}{\gamma} - \exp\left(\frac{x-\eta}{\gamma}\right)\right). \quad (1.51)$$

Distribuční funkce (kumulativní pravděpodobnost)

$$F_d(x) = 1 - \exp\left(-\exp\left(\frac{x-\eta}{\gamma}\right)\right). \quad (1.52)$$

Parametry: η - 63% kvantil, modus
 γ - rozptyl

Střední hodnota je

$$E(X) = \eta - \gamma C, \quad (1.53)$$

kde C je Eulerova konstanta $C = 0,5772$.

Rozptyl je

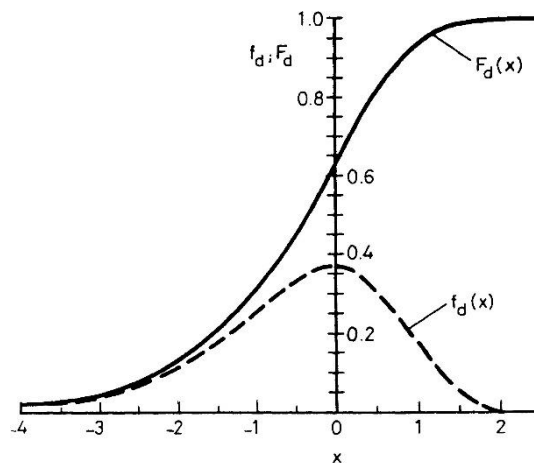
$$D(X) = \frac{1}{6} \pi^2 \gamma^2. \quad (1.54)$$

Odhad η

$$\eta^* = x_{63}. \quad (1.55)$$

Odhad γ

$$\gamma^* = \frac{1}{3} (x_{63} - x_{05}). \quad (1.56)$$



Obr. 1.10 Distribuční funkce a hustota pravděpodobnosti Laplaceova rozdělení
s parametry $\gamma = 1$, $\eta = 0$ [1]

1.3 Lineární regrese

Lineární regrese je metoda využívaná pro proložení bodů přímkou. Předpokládáme, že závislost Y na X lze vyjádřit přímkou a hodnoty X jsou přesné na rozdíl od hodnot Y , které mohou být ovlivněny určitou chybou (např. chyba měření) [1][7].

1.3.1 Metoda nejmenších čtverců

Metoda nejmenších čtverců je jedna z numerických metod sloužících k nalezení funkce k proložení dat. Pracuje na principu hledání nejmenší kvadratické odchylky. Lze využít i k hledání složitých funkcí, pro naše potřeby nám vyhovuje proložení přímkou, viz kapitola 1.3.2. Hledaná přímka má tvar [1][3][7]

$$y = kx + q. \quad (1.57)$$

Parametry k a q můžeme najít pomocí operace s maticemi nazývané pseudoinverze

$$x = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T y, \quad (1.58)$$

kde x je $x = \begin{pmatrix} k \\ q \end{pmatrix}$, matice \mathbf{A} obsahuje sloupec hodnot x a sloupec jedniček a sloupcový vektor y obsahuje hodnoty y [1][3][7].

Nebo můžeme využít ekvivalentní vzorce, které nevyužívají maticové operace

$$k = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}, \quad (1.59)$$

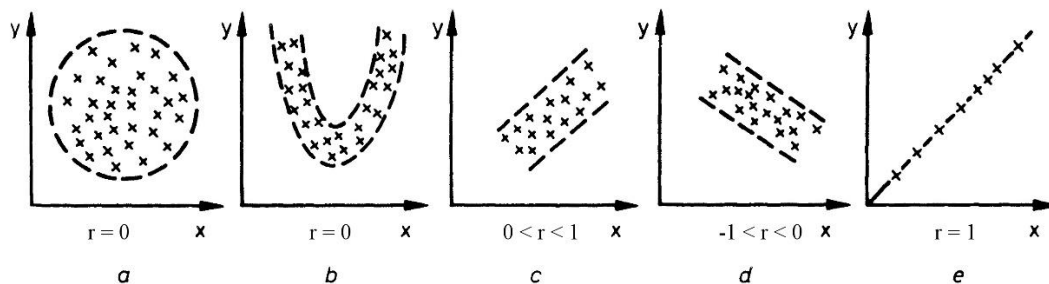
$$q = \frac{\sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i - \sum_{i=1}^n x_i \sum_{i=1}^n x_i y_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}, \quad (1.60)$$

kde n je počet bodů [1][3][7].

Vhodnost vypočítané přímky ověříme pomocí korelačního koeficientu, který je dán

$$r = \frac{\sum_{i=1}^n x_i y_i - n \bar{x} \bar{y}}{\sqrt{(\sum_{i=1}^n x_i^2 - n \bar{x}^2)(\sum_{i=1}^n y_i^2 - n \bar{y}^2)}}. \quad (1.61)$$

Korelační koeficient nabývá hodnot $-1 \leq r \leq 1$, kdy hodnoty blízké jedné, resp. mínus jedné, naznačují lineární závislost veličiny [1][3][7].



Obr. 1.11 Korelační koeficient pro různé soubory dat [1]

1.3.2 Metoda pravděpodobnostního papíru

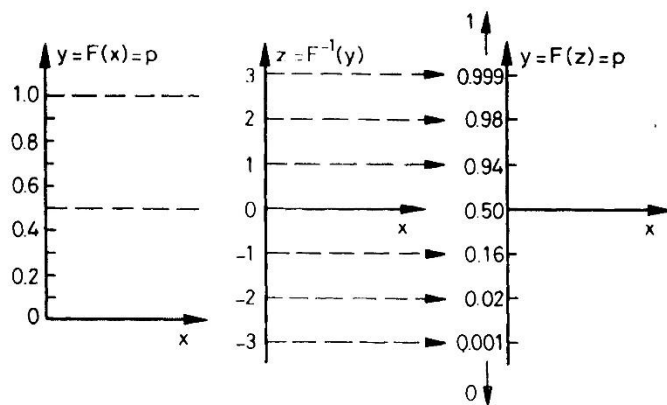
Jedná se o grafickou metodu, která využívá transformaci osy Y, dle zkoumaného typu distribuční funkce. Hodnoty osy Y odpovídají inverzní funkci daného rozdělení

$$F^{-1}(y) = z. \quad (1.62)$$

Tímto dosáhneme změny zkoumané distribuční funkce na přímku. Tento postup lze aplikovat na všechna rozdělení viz Tab. 1.1 a Obr. 1.12, kde je zobrazen postup pro normální rozdělení.

$z = 0$	$y = \Phi(0) = 0,5$
$z = 1$	$y = \Phi(1) = 0,841$
$z = -1$	$y = \Phi(-1) = -0,159$

Tab. 1.1 Hodnoty y pro normální rozdělení



Obr. 1.12 Vytvoření pravděpodobnostního papíru pro normální rozdělení [1]

Díky zjednodušení problému můžeme rozeznat, zda naměřená data podléhají navrhovanému rozdělení či nikoli. Při vyhodnocování je pro ověření výsledů vhodné využít více typů pravděpodobnostních papírů [1].

1.4 Kolmogorovův test

Kolmogorovův test je metoda matematické statistiky, která dovoluje testovat, zda dvě jednorozměrné náhodné proměnné pocházejí ze stejného rozdělení pravděpodobnosti [1].

Jednovýběrový test ověřuje, zda se rozdělení náhodné veličiny liší od určitého teoretického rozdělení.

Nulová hypotéza předpokládá, že testovaný výběr odpovídá vybranému teoretickému rozdělení. Pokud platí

$$D_t < D_{krit}, \quad (1.63)$$

nulovou hypotézu přijmeme, v opačném případě ji zamítneme. Hodnotu D_{krit} pro nízké hodnoty n získáme z tabulek, které nalezneme v [1] a [16]. Pokud $n \geq 35$ můžeme využít vzorce uvedené v tabulce Tab. 1.2. Hladina významnosti α nám udává pravděpodobnost chybného zamítnutí hypotézy.

Hodnota D_t se vypočítá vzorcem

$$D_t = \max(F_n(x) - F_0(x)) \quad (1.64)$$

kde $F_n(x)$ je zkoumané rozdělení a $F_0(x)$ je teoretické rozdělení.

Hladina významnosti α	D_{krit}
20%	$\frac{1,07}{\sqrt{n}}$
10%	$\frac{1,22}{\sqrt{n}}$
5%	$\frac{1,36}{\sqrt{n}}$
2%	$\frac{1,52}{\sqrt{n}}$
1%	$\frac{1,63}{\sqrt{n}}$

Tab. 1.2 Vzorce pro výpočet kritické hodnoty pro jednotlivé hladiny

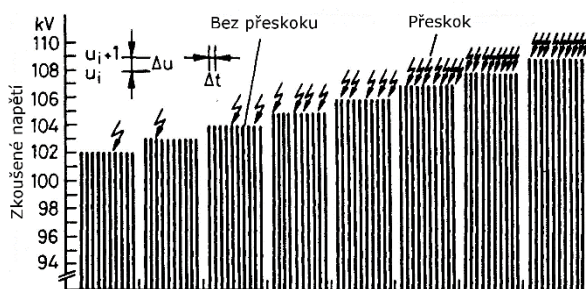
2 Metody měření

2.1 Zkouška více hladinami

Při zkoušce více hladinami přikládáme n napětí, která způsobí k přeskoků na m napět'ových hladinách. Ve většině případů se využívá impulzní napětí, lze použít i střídavé a stejnosměrné napětí s předepsanou dobou namáhání. Dle normy ČSN EN 60060-1 se doporučují následující parametry [1] [12]:

$$m \geq 5; n \geq 10; \Delta U = (0,01 \text{ až } 0,06)U_{50}$$

V normě není stanoven minimální čas mezi jednotlivými impulzy Δt . Tento parametr je ovlivněn především vlastnostmi zkoušeného prvku a musí se určit individuálně u každého měření [1] [12].



Tab. 2.1 Ilustrace záznamu zkoušky více hladinami [1]

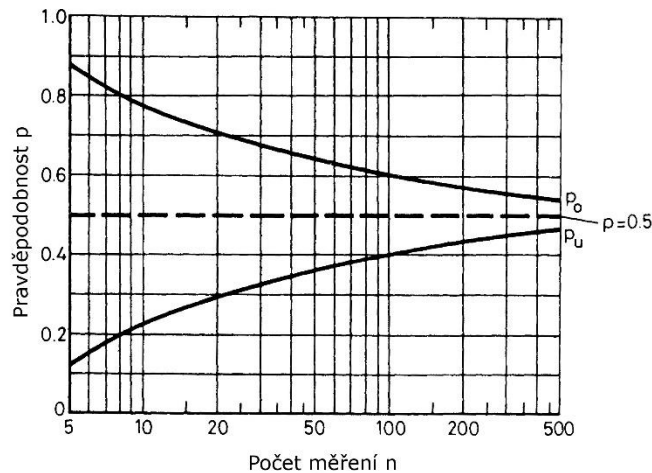
Při volbě počtu napět'ových hladin m a napět'ového kroku Δu musíme zohlednit rozsah testovaného napětí Δu_r . Tento rozsah by měl obsahovat pouze napětí, u kterých je pravděpodobnost přeskoků větší než nula a menší než 1 ($0 \leq p(u_d) \leq 1$). Nesmíme zapomenout zohlednit testovací zařízení a jeho nejmenší krok při nastavování napětí Δu_p . Počet napět'ových hladin by měl být menší než m_t , kde [1] [12]

$$m_t = \frac{\Delta u_r}{\Delta u_p}. \quad (2.1)$$

Následně můžeme vypočítat vhodný napět'ový krok Δu jako

$$\Delta u = \frac{\Delta u_r}{m}. \quad (2.2)$$

Volba počtu přiložení napětí n má největší vliv na přesnost výsledků. Počet napěťových impulzů n volíme vzhledem k intervalům spolehlivosti. Můžeme využít Obr. 1.8 nebo rovnice (1.34) až (1.36) k vytvoření grafu závislosti, ze kterého můžeme následně určit vhodný počet napěťových impulzů n [1] [12].



Obr. 2.1 Graf závislosti intervalů spolehlivosti na počtu testovacích impulzů n [1]

2.1.1 Ověření nezávislosti

Test probíhá na vybrané napěťové hladině, pro kterou zvolíme určitý počet přiložení napětí n_t . Hodnota n_t by neměla být menší než 10, doporučuje se alespoň 20. Při tomto testu musíme zaznamenávat jednotlivé události chronologicky. Získaný soubor dat rozdělíme na menší, stejně velké intervaly, pro které vypočítáme relativní frekvence přeskoků. Pokud se relativní frekvence jednotlivých intervalů pohybuje v blízkosti celkové relativní frekvence, jsou data nezávislá a mohou být statisticky vyhodnocena (Tab. 2.2). Pokud není nezávislost zřejmá z relativních frekvencí, musíme využít některou z matematických metod, které jsou popsány v [1].

Měření ve volném prostoru																h_{20}	h_{100}				
-	x	x	x	-	-	-	x	-	x	x	x	x	x	x	-	x	-	x	0,65	0,68	
x	-	x	-	-	x	x	x	x	x	-	x	x	-	x	x	x	x	x	0,75		
-	x	x	x	x	-	x	x	x	x	x	-	x	-	x	x	x	-	-	0,7		
x	x	x	-	x	x	-	x	-	x	-	-	-	x	-	x	x	x	-	x		0,6
x	x	-	x	x	x	x	-	-	x	-	x	-	x	x	x	x	x	-	x		0,7

Tab. 2.2 Měření při atmosférických podmínkách (Nezávislé)

Měření v uzavřené nádobě																		h ₂₀	h ₁₀₀		
x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	-	x	0,95	0,48
x	-	x	-	x	x	-	x	x	-	-	x	x	x	-	x	-	-	x	-	0,55	
-	-	-	x	x	x	-	x	-	-	-	x	-	x	x	-	x	-	x	x	0,5	
-	x	x	-	-	-	-	-	-	x	-	-	-	x	-	x	-	-	x	x	0,35	
-	-	-	-	-	-	x	-	-	-	-	-	-	-	-	-	-	-	-	-	0,05	

Tab. 2.3 Měření v uzavřené nádobě (Závislé)

2.1.2 Statistické vyhodnocení měření

Po ověření nezávislosti dat můžeme přistoupit ke statistickému vyhodnocení měření. Prvním krokem je výpočet relativních frekvencí na jednotlivých napěťových hladinách dle rovnice (1.1) a pomocí rovnic (1.34) až (1.36) vypočítat jejich intervaly spolehlivosti. Výsledné hodnoty je vhodné zapsat do tabulky (Tab. 2.4) [1] [12].

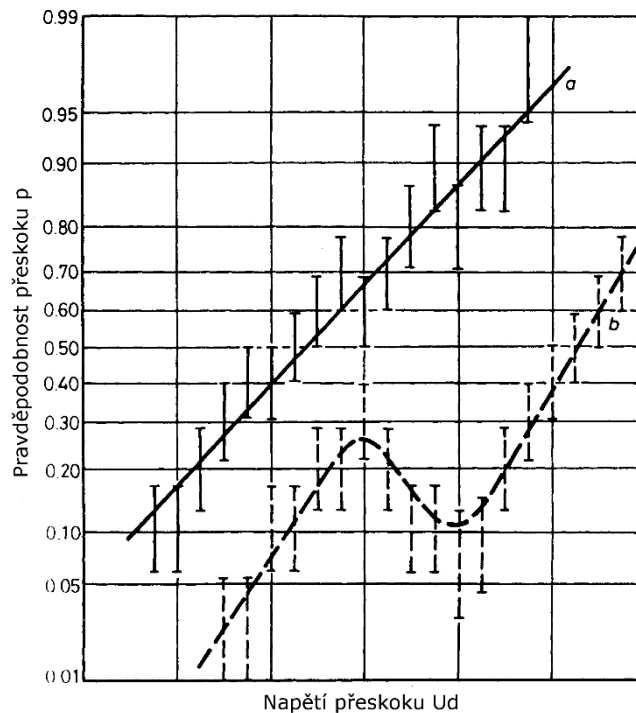
Napětí [kV]	Počet přeskoků	Relativní frekvence	Interval spolehlivosti $\varepsilon=0,95$	
170	1	0,04	0,002	0,176
175	5	0,2	0,082	0,375
180	10	0,4	0,236	0,583
185	13	0,52	0,341	0,695
190	19	0,76	0,58	0,89

Tab. 2.4 Zapsání výsledků měření více hladinami

Distribuční funkci lze získat více způsoby. Jedním z nich je vynesení relativních frekvencí na pravděpodobnostní papír a následně jejich proložení přímkou pomocí pravítka. V případě velkého rozptylu bodů musíme vynést na pravděpodobnostní papír i intervaly spolehlivosti (Nejčastěji používané hodnoty koeficientu spolehlivosti ε 90 % a 95 %) a ověřit zdali je zvolená přímka vhodná. Další způsob je pomocí metody nejmenších čtverců (kapitola 1.3.1), kdy relativní frekvence nejdříve převedeme (kapitola 1.3.2) pomocí inverzní funkce zvolené distribuční funkce a následně aplikujeme metodu nejmenších čtverců. Ověřit vhodnost vypočítané přímky můžeme opět pomocí intervalů spolehlivosti nebo pomocí Kolmogorova testu (kapitola 1.4) [1] [12].

Parametry distribuční funkce můžeme získat pomocí kvantilů z nalezené přímky nebo pomocí empirických vzorců pro každé rozdělení (kapitoly 1.2.3 a 1.2.4). Pro získané parametry musíme opět vypočítat jejich intervaly spolehlivosti (kapitola 1.2.3). Získaná distribuční funkce se dá využít jako performance function (Performance function popisuje pravděpodobnosti přeskoků pro jednotlivá napětí v celém rozsahu) [1] [12].

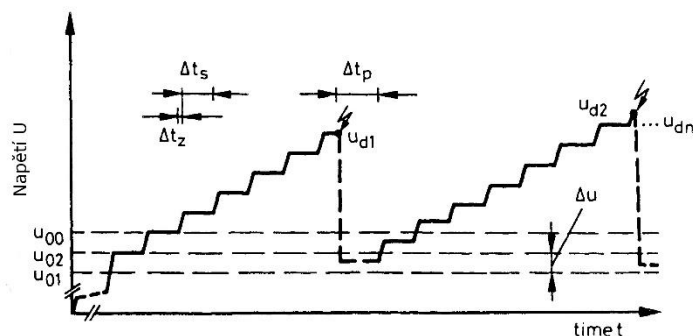
Proložení distribuční funkcí (přímkou) v případě výrazné změny relativní frekvence (funkce b na Obr. 2.2) není vhodné, jelikož distribuční funkce je neklesající [1] [12].



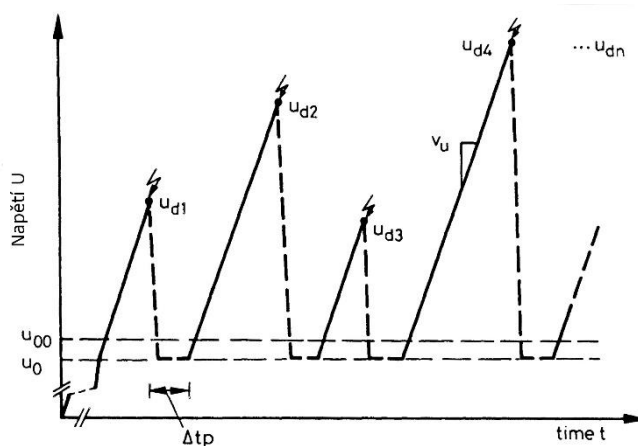
Obr. 2.2 a) vhodné pro proložení distribuční funkcí; b) nevhodné pro proložení distribuční funkcí [1]

2.2 Zkoušky postupným namáháním

U zkoušky postupným namáháním se postup, vedoucí vždy k průraznému výboji na zkoušeném objektu, provede n -krát. Zkouška lze provádět stejnosměrným, střídavým nebo postupným zvyšováním impulzního napětí. Napětí lze zvyšovat plynule (Obr. 2.4) nebo po krocích (Obr. 2.3). Výsledkem měření je n hodnot přesokového napětí [1] [12].



Obr. 2.3 Krokové zvyšování napětí [1]



Obr. 2.4 Plynulé zvyšování napětí [1]

Sledované parametry u postupného zvyšování napětí jsou rychlost zvyšování napětí v_u , celkový počet přeskoků (počet měření) n , doba mezi jednotlivými testy Δt_p a počáteční napětí u_0 . Při zvedání napětí po krocích sledujeme navíc velikost kroku Δu , čas trvání jednoho kroku Δt_s a dobu zvyšování napětí Δt_z . Při volbě počátečního napětí u_0 musíme zohlednit nejmenší napětí přeskoků u_{dmin} , měla by být splněna podmínka [1] [12]

$$u_{dmin} > 1,1u_0. \quad (2.3)$$

Velikost kroku Δu , je omezena nejmenším krokem při nastavování napětí Δu_p a neměla by být větší než směrodatná odchylka s . Ideálně se snažíme přiblížit k

$$\Delta u \approx 0,5s. \quad (2.4)$$

V případě postupného zvyšování napětí je rychlost zvyšování napětí dána

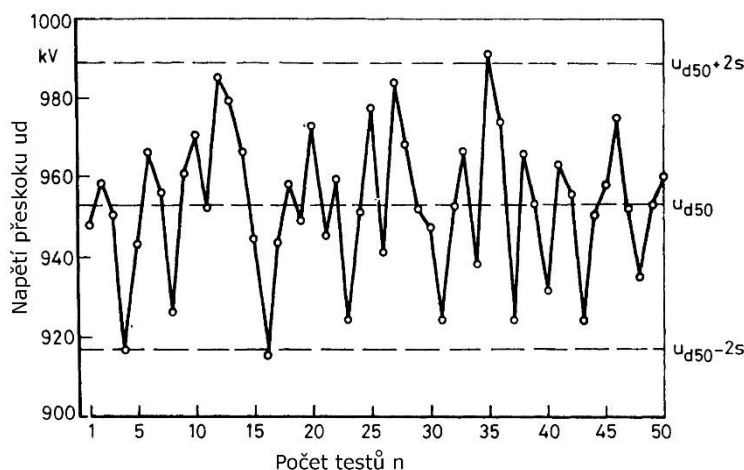
$$v_u = \frac{\Delta u}{\Delta t_s}. \quad (2.5)$$

Velikost v_u musí být udržena téměř konstantní, aby nedocházelo k ovlivnění výsledků.

Počet individuálních testů n je opět dán požadavky na přesnost měření. Při volbě můžeme postupovat jako v kapitole 2.1 [1] [12].

2.2.1 Ověření nezávislosti

U této metody nemusíme provádět vlastní měření pro ověření nezávislosti, ve většině případů stačí pouze grafická metoda ověření (Obr. 2.5). Ta spočívá v chronologickém vykreslení napětí do grafu, kdy hodnoty by neměly vykazovat žádný trend (náhodně rozložené) a měly by se pohybovat v okolí střední hodnoty v maximální odchylce $2s$, kde s je směrodatná odchylka [1] [12].



Obr. 2.5 Graf pro vyhodnocení závislosti dat [1]

2.2.2 Statistické vyhodnocení měření

Získaná přeskoková napětí nejdříve vzestupně seřadíme, pokud došlo k více přeskokům na stejném napětí, zapíšeme dané napětí pouze jednou a počet opakování zaznamenáme (Tab. 2.5). Relativní frekvenci vypočítáme pomocí rovnice (1.2), kde n je počet měření a m je počet opakování daného napětí. V případě zkoušky postupným namáháním nemůžeme získat distribuční funkci pouze z relativní frekvence. Proto použijeme vzorec

$$h_{\Sigma k} = \sum_{j=1}^k h_j, \quad (2.6)$$

kterým získáme kumulativní relativní frekvenci [1] [12].

Distribuční funkci můžeme získat více způsoby. Jeden ze způsobů je vynesení kumulativní relativní frekvence na pravděpodobnostní papír a následně proložit přímkou pomocí pravítka. Další způsob je za pomoci metody nejmenších čtverců (kapitola 1.3.1). Kumulativní relativní frekvence nejdříve převedeme (kapitola 1.3.2) pomocí inverzní

funkce zvolené distribuční funkce a následně aplikujeme metodu nejmenších čtverců. Ověřit vhodnost vypočítané přímky můžeme pomocí Kolmogorova testu (kapitola 1.4) .

Parametry distribuční funkce můžeme získat pomocí kvantilů z nalezené přímky nebo pomocí empirických vzorců pro každé rozdělení (kapitoly 1.2.3 a 1.2.4). Pro získané parametry musíme opět vypočítat jejich intervaly spolehlivosti (kapitola 1.2.3) [1] [12].

Na rozdíl od zkoušky více hladinami, distribuční funkce zde neodpovídá skutečnosti a dále se přepočítává dle napěťového kroku Δu a použitého teoretického rozdělení na performance function. Pokud nevyžadujeme vysokou přesnost výsledků, můžeme využít, za cenu předimenzování, i nepřečítanou distribuční funkci [1] [12].

Přeskokové napětí [kV]	Počet opakování	Relativní frekvence	Kumulativní relativní frekvence
106,4	2	0,12	0,12
116,2	1	0,06	0,19
123	4	0,25	0,44
125,9	1	0,06	0,5
130,4	1	0,06	0,56
134,2	1	0,06	0,62
137,6	1	0,06	0,69
144,2	1	0,06	0,75
148	1	0,06	0,81
156,2	1	0,06	0,88
163,7	1	0,06	0,94

Tab. 2.5 Zapsání výsledků zkoušky postupným namáháním

3 Nástroj pro statistické vyhodnocování testů

3.1 Software

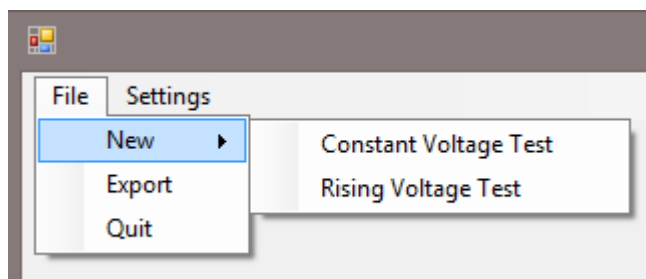
Nástroj pro statistické vyhodnocování testů, jsem vytvořil v jazyce C#. Jedná se o CIL jazyk (Common Intermediate Language). Jako vývojové prostředí jsem využil Visual Studio 2013 od společnosti Microsoft. Program obsahuje matematickou volně šiřitelnou knihovnu Math.NET [13]. Z důvodu jednoduššího šíření programu jsem využil aplikaci ILMerge, která dovoluje přidat externí knihovny do exe souboru.

Ke správné funkci programu je zapotřebí mít nainstalovaný Microsoft .NET Framework alespoň ve verzi 4.0 a Microsoft Excel pro exportování do xls souborů.

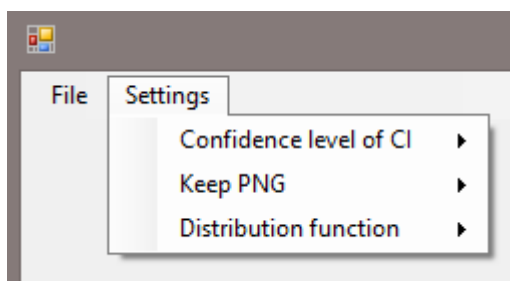
3.2 Program

3.2.1 Interface programu

V horní nabídce najdeme položky pro vytvoření nového měření, export (Obr. 3.1) a nastavení (Obr. 3.2). V nastavení se nachází položky pro volbu hledané distribuční funkce, nastavení koeficientu spolehlivosti a ponechání výsledného grafu v samostatném souboru ve formátu png.



Obr. 3.1 Hlavní nabídka



Obr. 3.2 Nastavení

Po zvolení daného typu měření se zobrazí nové rozložení interface (Obr. 3.4 a Obr. 3.6). Výběrem dané záložky (Single Fit a Multiple Fit) se zvolí žádaný způsob hledání

distribuční funkce. Vkládání dat probíhá pomocí dataGridView, který je umístěn vlevo. V případě využití Multiple Fit musí být vybrán způsob rozdělení zadaných dat (Obr. 3.3). Program nabízí ruční nebo automatické rozdělení.

Dividing dataset

Dividing point

Obr. 3.3 Výběr místa rozdělení zadaných dat

	Voltage [kV]	Number of breakdowns
	1	1
	2	2
	3	3
	4	4
	5	5
	6	6
	7	7
	8	8
	9	9
	10	10
▶▶		

Number of discharges

Kolmogov test

Confidence Level

Results

Median

Standart deviation

Obr. 3.4 Interface pro zkoušku více hladinami s jednoduchým proložením

Podrobné výsledky jsou zobrazeny v záložce Results (Obr. 3.5), kde jsou vypsány i intervaly spolehlivosti pro získané parametry. Vypočítaná data a výsledný graf lze vyexportovat do Excel souboru. Screenshot celého programu naleznete v Příloze A.

Results		Confidence limits	
Median	8,1	6,3	9,9
Standart deviation	4,6	3,6	6,3
Correlation coefficient	0,986		
Median	14,6	13,8	15,3
Standart deviation	1,9	1,5	2,7
Correlation coefficient	0,995		

Comment: Vzorová data

Obr. 3.5 Podrobné výsledky

Single Fit		Multiple Fit		Results	
	Voltage of Breakdown [kV]	Kolmogov test			
	1	Confidence Level	0,2	Passed	
	2	Confidence Level	0,2	Passed	
	1	Dividing dataset: Auto			
	2	Results			
	3	Median	3		
	4	Standart deviation	1,9		
	5	Median	4,2		
	5	Standart deviation	3,6		
	6				
	7				
	8				
	9				
	»»				

Obr. 3.6 Interface pro zkoušku postupným namáháním s dvojitým proložením

3.2.2 Intervaly spolehlivosti

Interval spolehlivosti pro střední hodnotu μ (Normální rozdělení) se v programu počítá jako oboustranný interval spolehlivosti podle rovnic (1.45) a (1.46). Pro výpočet invCDF (Inverzní kumulativní distribuční funkce) Studentova t rozdělení využívám funkce z knihovny Math.Net. Při výpočtu oboustranného intervalu spolehlivosti pro rozptyl σ (Normální rozdělení), viz rovnice (1.49) a (1.50) se využívá Chí-kvadrát rozdělení, v tomto případě jsem nemohl využít přímo funkci invCDF, která v knihovně Math.Net chybí. Tento problém jsem vyřešil použitím funkce využívající CDF (kumulativní distribuční funkce) Chí-kvadrát rozdělení.

```
public static double InvChiSquare(int degfreed,double prob)
{
    double Inv = 0;
    double test = 0;
    int i = 0;

    while(i<130000)
    {
        test=Math.Round(ChiSquared.CDF(degfreed,Inv),3);
        i++;

        if (prob == test)
        {
            return Inv;
        }

        else if (prob > test)
        {
            Inv=Inv+0.001;
        }

        else if (prob < test)
        {
            Inv=Inv-0.001;
        }
    }

    return Inv;
}
```

Výpočet touto funkcí je přesný pro měření, kde n je menší než 100. Tato hodnota byla zvolena s dostatečnou rezervou, jelikož provádění měření s vysokými hodnotami n je časově náročné.

V případě výpočtu intervalů spolehlivosti pro pravděpodobnosti program využívá rovnice (1.34) a (1.35). Jedná se o jednostranné intervaly, které se počítají pomocí invCDF F rozdělení.

Intervaly spolehlivosti pro parametry Laplaceova rozdělení není možné počítat s dostatečnou přesností, proto se v programu nachází pouze bodové odhady parametrů.

3.2.3 Multi Fit

Pomocí této funkce můžeme lépe popsat děje, které přesně neodpovídají jednoduchým pravděpodobnostním rozdělením. Funguje na principu rozdělení vstupního souboru dat na dva samostatné soubory dat, které se vyhodnocují individuálně. Místo rozdělení lze zadat ručně (Obr. 3.3) nebo využít automatický režim, který pomocí porovnávání odchylek mezi nalezenou distribuční funkcí a vstupními daty pro jednotlivá místa rozdělení hledá ideální místo rozdělení dat.

3.3 Vzorová experimentální měření

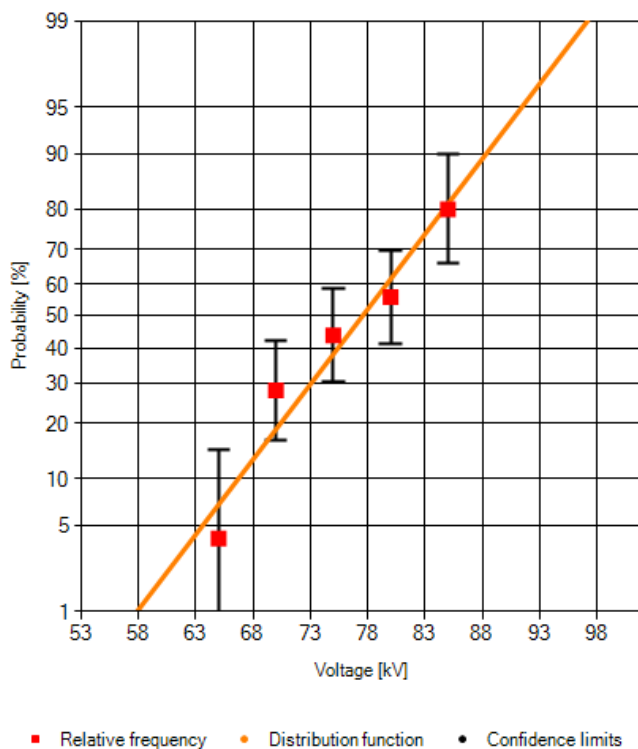
K nastínění různých přístupů při vyhodnocování budu v této části využívat vlastní program na modelových měřeních. Počet impulzů u všech měření je 25.

První případ jsou ideální data. Naměřené hodnoty jsou uvedeny v Tab. 3.1. Výsledný pravděpodobnostní graf Obr. 3.7. a výsledné parametry normálního rozdělení jsou $\mu = 77,52 \text{ kV} [74,65; 80,40]$ a $\sigma = 8,4 \text{ kV} [6,82; 11,07]$. Z výsledného pravděpodobnostního grafu je patrné, že nalezená distribuční funkce je vhodná jako performance function.

Napětí [kV]	Počet přeskoků	Relativní frekvence	Interval spolehlivosti $\varepsilon=0,90$	
65	1	0,04	0,004	0,147
70	7	0,28	0,163	0,426
75	11	0,44	0,301	0,587
80	14	0,56	0,413	0,699
85	20	0,8	0,66	0,899

Tab. 3.1 Naměřené hodnoty, zkouška více hladinami , měření 1

Při zkoušce více hladinami můžeme naměřit hodnoty, které se nedají přímo vyhodnotit. Jedná se o pravděpodobnosti 0 % a 100 %. Tyto hodnoty není možné přímo zobrazit na pravděpodobnostní papír, jelikož jejich daná pravděpodobnost výskytu se nachází v $-\infty$ respektive $+\infty$. Můj program je uzpůsoben na automatické substituování těchto limitních hodnot v případě zadání (Hodnoty popsány v Tab. 3.2), proto pokud se uživatel rozhodne zanedbat tyto hodnoty, nesmí je zadávat do programu.



Obr. 3.7 Pravděpodobnostní graf, měření 1

	Normální rozdělení	Laplaceovo rozdělení
0%	0,13%	0,70%
100%	99,87%	99,90%

Tab. 3.2 Substituované hodnoty

Při vyhodnocování takových měření musíme dávat pozor na umístění těchto dat ve vzorku. Pokud se napěťová hladina s pravděpodobností 0 % nebo 100 % nachází na začátku, respektive na konci, měřicího intervalu, můžeme tyto hodnoty zanedbat, aniž bychom se dopustili velké chyby (Tab. 3.3, Tab. 3.4).

Napětí [kV]	Počet přeskoků	Relativní frekvence	Interval spolehlivosti $\varepsilon=0,90$	
165	0	0	0,001	0,088
170	1	0,04	0,004	0,147
175	5	0,2	0,101	0,34
180	10	0,4	0,265	0,548
185	13	0,52	0,375	0,662
190	19	0,76	0,617	0,869

Tab. 3.3 Naměřené hodnoty, měření 2

	Nezanedbané	Zanedbané	ΔU
U_{10} [kV]	174,45 kV	172,63 kV	1,82 kV
U_{50} [kV]	183,55 kV	183,6 kV	0,05 kV

Tab. 3.4 Porovnání výsledků, měření 2

V případě, kdy se napěťová hladina s pravděpodobností 0 % nebo 100 % nenachází na začátku respektive na konci měřicího rozsahu, je ve většině případů vhodné použít vyhodnocení pomocí více distribučních funkcí, jelikož distribuční funkce je neklesající. V Tab. 3.5 je příklad měření, kdy je vhodné popsat měření více distribučními funkcemi. Při použití pouze jedné distribuční funkce je z pravděpodobnostního grafu (Příloha B) v obou případech patrná nevhodnost tohoto proložení (intervaly spolehlivosti bodů se nedotýkají nalezené distribuční funkce).

Napětí [kV]	Počet přeskoků	Relativní frekvence	Interval spolehlivosti $\varepsilon=0,90$	
90	1	0,04	0,004	0,147
95	3	0,12	0,045	0,248
100	7	0,28	0,163	0,426
105	9	0,36	0,23	0,508
110	10	0,4	0,265	0,548
115	0	0	0,001	0,088
120	4	0,16	0,072	0,295
125	16	0,64	0,492	0,77
130	20	0,8	0,66	0,899

Tab. 3.5 Naměřené hodnoty, měření 3

Proložení více funkcemi bylo dosaženo rozdělením dat mezi hladinami 110 kV a 115 kV. K určení bodu rozdělení byl použit automatický režim, který hledá řešení s nejmenší možnou odchylkou. Pokud porovnáme výsledné grafy proložení více funkcemi (Příloha C) s předchozím výsledkem (Příloha B), můžeme jednoznačně určit, že řešení s více funkcemi je vhodnější pro tento typ dat. V Tab. 3.6 jsou shrnuty hodnoty, které nás zajímají v praxi a vycházejí z daných postupů.

	Jednoduché proložení		Dvojitě proložení	
	Zanedbaná	Nezanedbaná	Zanedbaná	Nezanedbaná
U_{10} [kV]	92,88	96,8	94,08	94,08
U_{50} [kV]	119,65	130,03	110,81	110,81

Tab. 3.6 Porovnání výsledků, měření 3

Závěr

Při vyhodnocování a plánování měření si nejdříve musíme uvědomit, zda potřebujeme získat celou performance function nebo nám stačí pouze důležité hodnoty jako U_{10} a U_{50} . Z Tab. 3.6 je patrné, že i když funkce přesně nepopisují celý rozsah performance function, u hodnot U_{10} se liší minimálně, a proto je zbytečné v takových případech provádět rozsáhlá a časově náročná měření. Ty jsou zapotřebí v případech zkoumání celé performance function.

V kapitole 3.3 je nastíněna problematika s extrémními hodnotami pravděpodobnosti při provádění zkoušky více hladinami. Těmto problémům se nedá předejít lepším plánováním měření, zvláště v případech, kdy se dané hodnoty objeví uvnitř měřeného rozsahu (Tab. 3.5; hodnoty jsou způsobené charakterem sledovaného jevu). Při pochybnostech je vhodné provést znovu test nezávislosti nebo provést měření s větší dobou mezi impulzy Δt . Zda se rozhodneme tyto hodnoty využít a jak, záleží čistě na dalších okolnostech. Pravděpodobně nejefektivnějším způsobem vypořádání se s daným problémem je změna testovací metody na zkoušku postupným namáháním, kde díky jiné metodě výpočtu relativní frekvence nemohou nastat problémy s danými pravděpodobnostmi. Bohužel tato metoda je více náročná na testovací zdroj, kde rychlost zvyšování napětí v_u musí být téměř konstantní. Nekonstantní v_u může způsobit závislost měření a způsobit nepřesnosti při převodu kumulativní pravděpodobnosti na performance function.

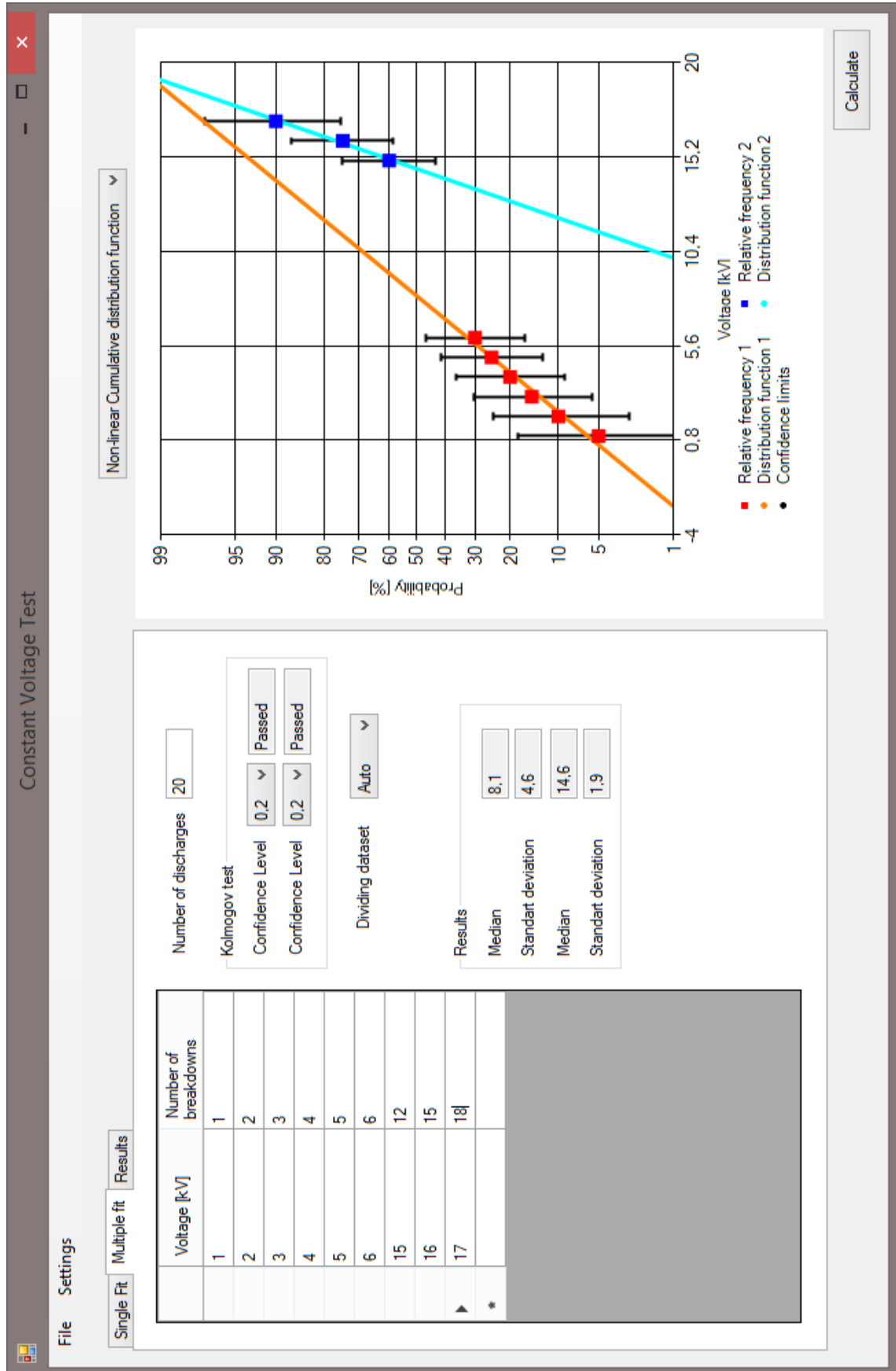
Statistické vyhodnocování není čistě exaktní věda. Proto se při hledání nejhodnějšího výsledku často využívá metoda pokus omyl. Z toho důvodu je vhodné mít nástroj dostatečně variabilní, aby mohl poskytnout co nejvíce variant, které by mohly být využity při vyhodnocování měření. Proto můj nástroj obsahuje nejen Normální rozdělení, ale také zástupce extrémních rozdělení, kterým je Laplaceovo rozdělení (Dvojitě exponenciální rozdělení). V některých případech, jako měření číslo 3 (Tab. 3.5) si nevystačíme ani s různými distribučními funkcemi a musíme měření popsat více funkcemi najednou (Multi fit). Můj program neobsahuje pouze automatický režim rozdělování dat, ale dovoluje i možnost vybrání místa pro rozdělení dat. Při využití programu odpadají rutinní výpočty a při výpočtech se omezí lidský faktor, aby nedošlo k nechtěným chybám. Program je vybaven statistickými nástroji pro porovnání výsledků jako je Kolmogorův test a korelační koeficient. Výsledný program, který byl hlavním cílem práce, naleznete na přiloženém CD.

Seznam literatury a informačních zdrojů

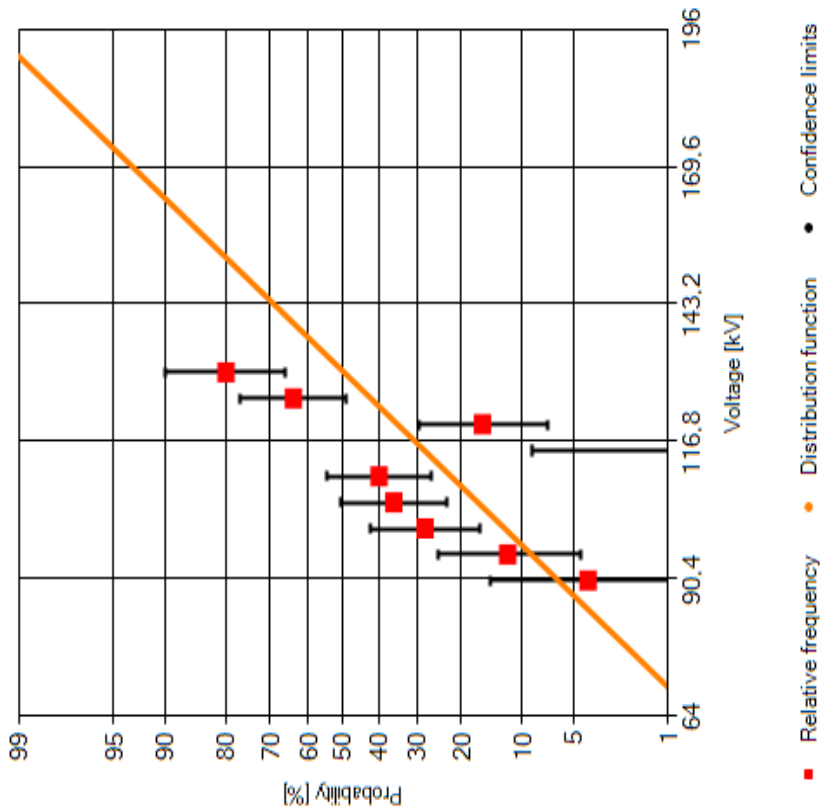
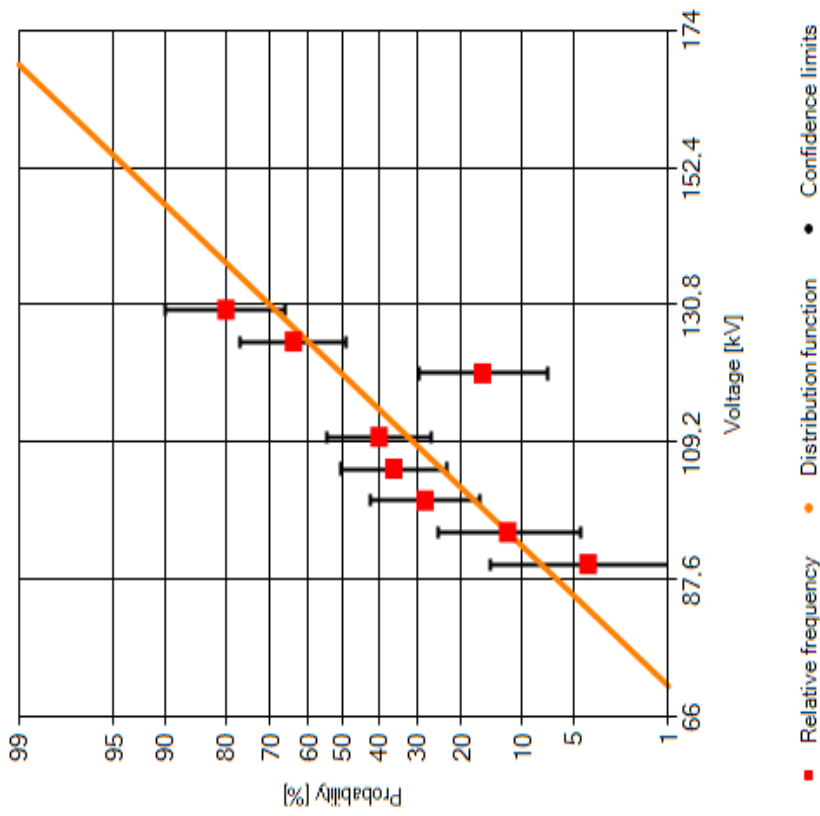
- [1] HAUSCHILD, Wolfgang a Wolfgang MOSCH. *Statistical techniques for high-voltage engineering*. English ed. London: P. Peregrinus on behalf of the Institution of Electrical Engineers, c1992, xii, 310 p. ISBN 086341205x.
- [2] MIČKA, Pavel. Metoda nejmenších čtverců. In: *Algoritmy.net* [online]. 209 [cit. 2015-04-29]. Dostupné z: <http://www.algoritmy.net/article/1592/Nejmensi-ctverce>
- [3] Metoda nejmenších čtverců. In: *Wikipedia: the free encyclopedia* [online]. San Francisco (CA): Wikimedia Foundation, 2001- [cit. 2015-04-29]. Dostupné z: http://cs.wikipedia.org/wiki/Metoda_nejmen%C5%A1%C3%ADch_%C4%8Dtverc%C5%AF
- [4] RŮŽIČKOVÁ, Irena a Rudolf HLAVIČKA. *Numerické metody*. Brno, Vysoké Učení Technické v Brně. Dostupné z: <http://physics.ujep.cz/~jskvor/NME/DalsiSkripta/Numerika.pdf>
- [5] MAYFIELD, Philip. SIGMAZONE. *Understanding Binomial Confidence Intervals* [online]. 2011 [cit. 2015-04-29]. Dostupné z: http://www.sigmazone.com/binomial_confidence_interval.htm
- [6] Confidence interval. In: *Wikipedia: the free encyclopedia* [online]. San Francisco (CA): Wikimedia Foundation, 2001- [cit. 2015-04-29]. Dostupné z: http://en.wikipedia.org/wiki/Confidence_interval
- [7] ŠEDIVÁ, Blanka, Tomáš ŤOUPAL a Eva WÁGNEROVÁ. *Pravděpodobnost a statistika pro FEL*. Plzeň, 2011. Dostupné z: http://home.zcu.cz/~sediva/pse/hlavni_text_PSE.pdf
- [8] SOUKUP, Petr. UNIVERZITA KARLOVA V PRAZE. *Nesprávná užívání statistické významnosti a jejich možná řešení*. Praha, 2010.
- [9] ŘEZANKOVÁ, Hana, Luboš MAREK a Michal VRABEC. *INTERAKTIVNÍ UČEBNICE STATISTIKY* [online]. 2001 [cit. 2015-04-29]. Dostupné z: <http://iastat.vse.cz/>
- [10] MICROSOFT. *Documentation Visual Studio* [online]. 2015 [cit. 2015-04-29]. Dostupné z: <https://msdn.microsoft.com/library/vstudio>
- [11] MATHWORKS, Inc. *Documentation Matlab* [online]. 1994 [cit. 2015-04-29]. Dostupné z: <http://www.mathworks.com/help/matlab/index.html>
- [12] ČSN EN 60060-1 Technika zkoušek vysokým napětím Část 1: Obecné definice a požadavky na zkoušky. Praha: Úřad pro technickou normalizaci, metrologii a státní zkušebnictví, 2011
- [13] *Math.NET Numerics* [online]. 2009 [cit. 2015-04-29]. Dostupné z: <http://www.mathdotnet.com/>
- [14] BARNETT, Mike. ILMerge. *Microsoft.com* [online]. 2009 [cit. 2015-04-29]. Dostupné z: <http://research.microsoft.com/en-us/people/mbarnett/ilmerge.aspx>
- [15] MÜLLER, Paul Heinz, Peter NEUMANN a Regina STORM. *Tafeln der mathematischen Statistik*. 2., verb. Aufl. Wien: Hanser, 1977, 275 p. ISBN 3446121412.
- [16] Table 7: Kolmogorov-Smirnov Test. ECONOMIC RESEARCH INSTITUTE. *Distance Learning Center* [online]. 2015 [cit. 2015-04-29]. Dostupné z: <http://dlc.eriери.com/onlinetextbook/index.cfm?fuseaction=textbook.appendix&FileName=Table7>

Přílohy

Příloha A

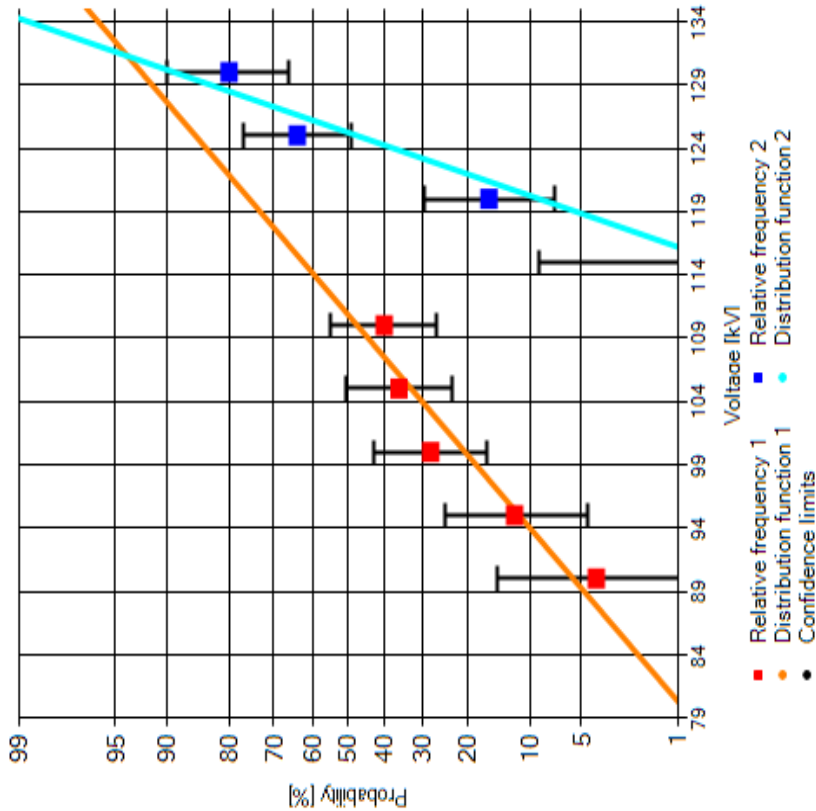
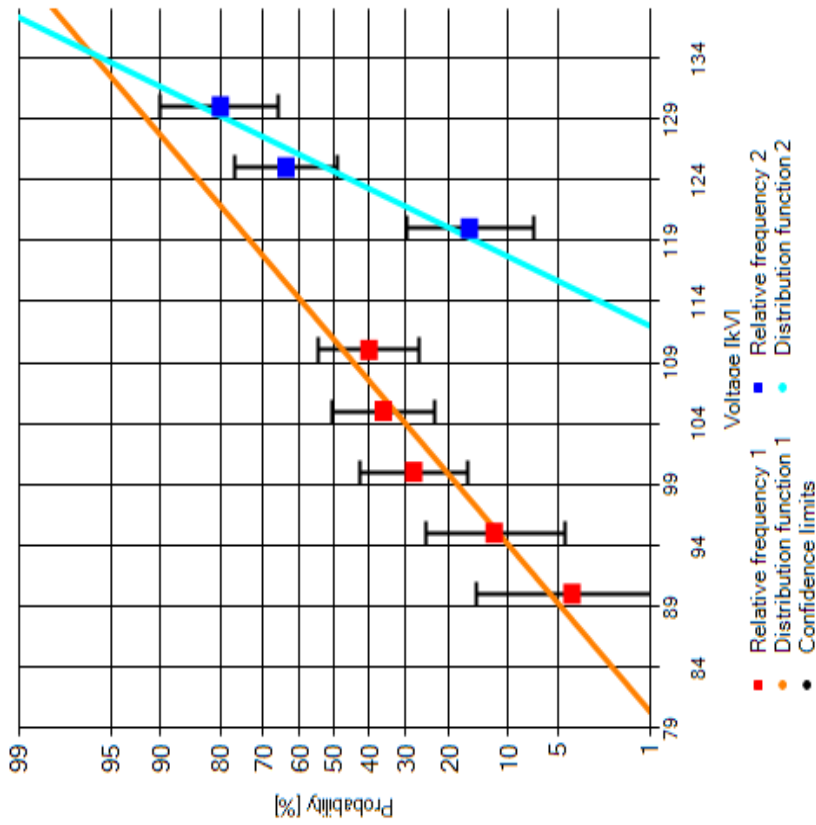


Příloha B



- a) Nezanedbaná hladina 115 kV $\mu = 130,03 \text{ kV}$, $\sigma = 25,93 \text{ kV}$
 b) Zanedbaná hladina 115 kV $\mu = 119,65 \text{ kV}$, $\sigma = 20,89 \text{ kV}$

Příloha C



- a) Nezanedbaná hladina 115 kV $\mu_1 = 110,81 \text{ kV}$, $\sigma_1 = 13,05 \text{ kV}$, $\mu_2 = 125,21 \text{ kV}$, $\sigma_2 = 3,86 \text{ kV}$
 b) Zanedbaná hladina 115 kV $\mu_1 = 110,81 \text{ kV}$, $\sigma_1 = 13,05 \text{ kV}$, $\mu_2 = 124,63 \text{ kV}$, $\sigma_2 = 5,42 \text{ kV}$