

Návrh a vyhodnocování experimentů

Petr Holub a Radka Svobodová-Vařeková

<hopet@ics.muni.cz>, <4056@mail.muni.cz>

DUVOD

2010-12-07

Přehled přednášky

Motivace

Měření

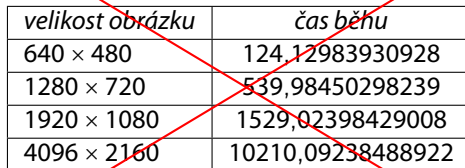
Zpracování měření

Návrh experimentu

Literatura

Délka zpracování obrázku

<i>velikost obrázku</i>	<i>čas běhu</i>
640 × 480	124,12983930928
1280 × 720	539,98450298239
1920 × 1080	1529,02398429008
4096 × 2160	10210,09238488922



<i>velikost obrázku</i>	<i>čas běhu</i>
640 × 480	124,12983930928
1280 × 720	539,98450298239
1920 × 1080	1529,02398429008
4096 × 2160	10210,09238488922

Měříme délku výpočtu v Javě

```
~$ R
```

```
...
```

```
> library(psych)
```

```
> runlength <- read.csv(file="java-example.table", head=FALSE, sep=",")
```

```
> summary(runlength$V1)
```

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
	92.08	104.70	108.80	166.80	187.20	594.70

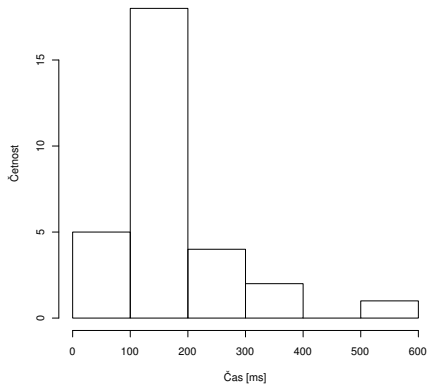
```
> describe(runlength$V1)
```

	var	n	mean	sd	median	trimmed	mad	min	max	range	skew
1	1	30	166.82	113.67	108.78	142.1	20.88	92.08	594.71	502.63	2.14
	kurtosis		se								
1			4.55	20.75							

$$\begin{aligned}
 N &= 30 \\
 \bar{x} &= 166,82 \\
 s_x &= 113,67 \\
 s_{\bar{x}} &= \frac{s_x}{\sqrt{N}} = 20,75 \\
 t_{0,05;29} &= 2,045
 \end{aligned}$$

$$\bar{x} \pm t_{0,05;N-1} s_{\bar{x}} = 167 \pm 42 \text{ ms}$$

Javové měření



Javové měření

$$N = 30$$

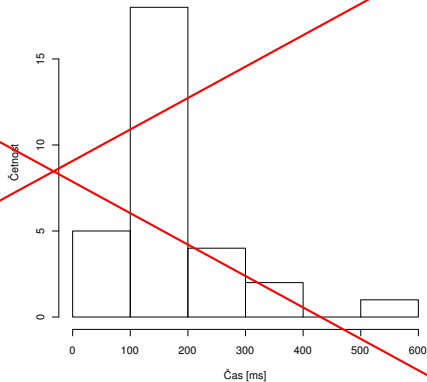
$$\bar{x} = 166,82$$

$$s_x = 113,67$$

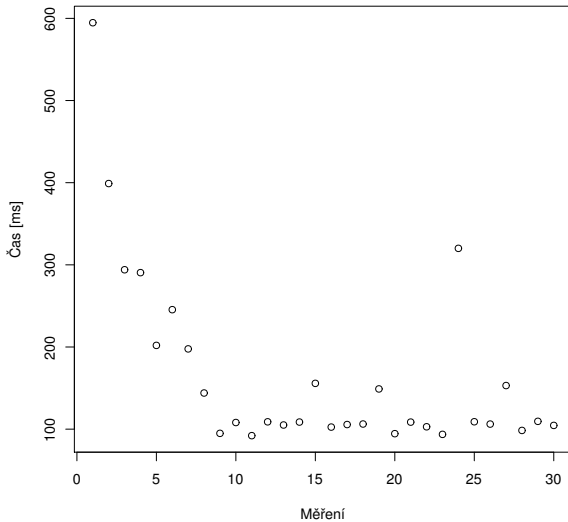
$$s_{\bar{x}} = \frac{s_x}{\sqrt{N}} = 20,75$$

$$t_{0,05;29} = 2,045$$

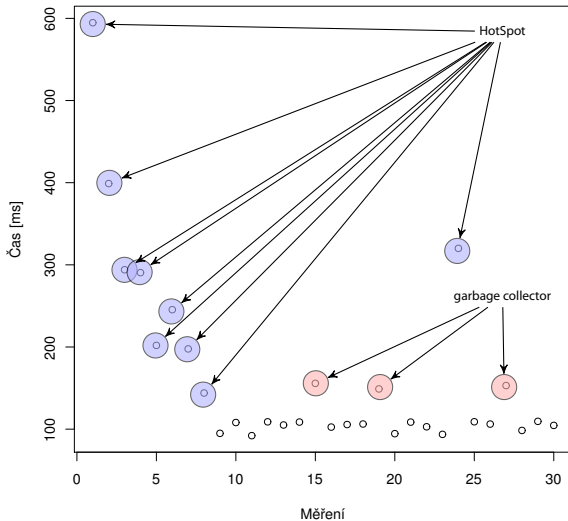
$$\bar{x} \pm t_{0,05;N-1} s_{\bar{x}} = 167 \pm 42 \text{ ms}$$



Javové měření



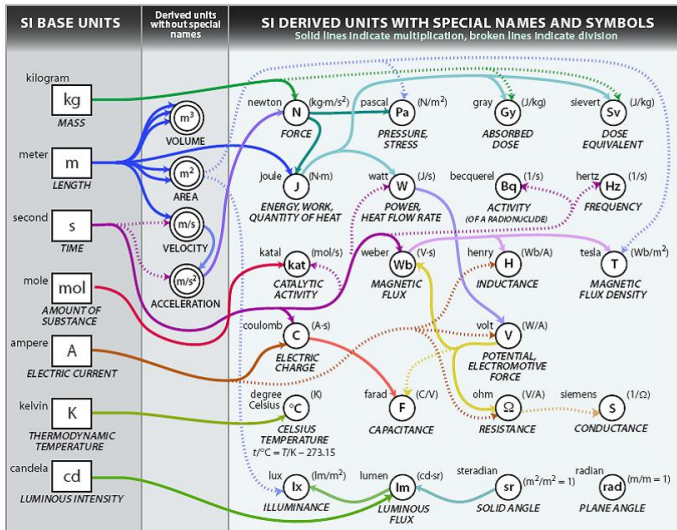
Javové měření



Proč experimenty?

- Informatika má silné nástroje pro zjišťování faktů
 - důkazy
 - výpočty
 - simulace
- Praktické studium vlastností systémů
 - některé vlastnosti neumíme nebo z důvodu obtížnosti nemůžeme simulovat
- Podpoření nebo vyvrácení hypotézy
 - pozor... nedokazujeme!

Soustava jednotek pro informatiky



Zdroj: [http://www.icrf.nl/Portals/106/SI_units_diagram\(1\).jpg](http://www.icrf.nl/Portals/106/SI_units_diagram(1).jpg)

Soustava jednotek pro informatiky

- Předpony nejen speciálně inforatické

yocto-	10^{-24}	y	–	–	–
zepto-	10^{-21}	z	–	–	–
atto-	10^{-18}	a	–	–	–
femto-	10^{-15}	f	–	–	–
pico-	10^{-12}	p	–	–	–
nano-	10^{-9}	n	–	–	–
micro-	10^{-6}	μ	–	–	–
milli-	10^{-3}	m	–	–	–
kilo-	10^3	k	kibi	2^{10}	Ki
mega-	10^6	M	mebi	2^{20}	Mi
giga-	10^9	G	gibi	2^{30}	Gi
tera-	10^{12}	T	tebi	2^{40}	Ti
peta-	10^{15}	P	pebi	2^{50}	Pi
exa-	10^{18}	E	exbi	2^{60}	Ei
zetta-	10^{21}	Z	zebi	2^{70}	Zi
yotta-	10^{24}	Y	yobi	2^{80}	Yi

Amendment 2 to "IEC 60027-2: Letter symbols to be used in electrical technology – Part 2: Telecommunications and electronics" (1999)

Typy měřících metod

- **Subjektivní × objektivní metody**
 - subjektivní: působí bezprostředně na lidské smysly
 - objektivní: působí na měřící zařízení
- **Přímé × nepřímé metody**
 - přímé: přímé srovnání se známou hodnotou veličiny
 - nepřímé: na základě jiných veličin, pomocí nichž lze měřenou veličinu spočítat
- **Absolutní × relativní metody**
 - absolutní: měření přímo v příslušné jednotce
 - relativní: měření srovnáním
- **Statické × dynamické metody**
 - statické: z klidového stavu přístroje
 - dynamické: na základě dynamiky měřícího přístroje

Výsledky měření

- Rozlišení měření
- Chyby měření
 - skládání většího počtu mikroskopických jevů
 - subjektivní vliv u měřících metod
- Jedno číslo zdaleka nepostihuje tyto informace

Výsledky měření

$$x = (\hat{\mu}_x \pm z_x) \text{ [jednotka]}$$

- $\hat{\mu}_x$... nejpravděpodobnější hodnota měřené veličiny
- z_x ... interval spolehlivosti / přesnost
- jak tyto věci spočítat / odhadnout?

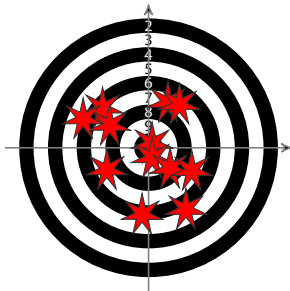
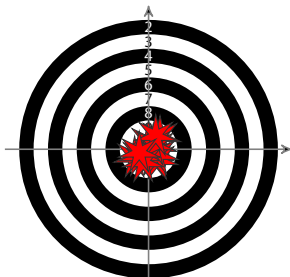
Výsledky měření

- Protokolování podmínek, na nichž měření probíhalo
 - zachycení všech podmínek, které mohou mít na měření vliv
 - ◆ konfigurace hardware
 - ◆ popis síťové topologie
 - ◆ instalovaný operační systém
 - ◆ instalovaný software
 - ◆ popis konfigurace a souběžně běžících procesů
 - ◆ uschování vlastního měřeného software/hardware
 - ◆ přesný popis použitých měřících metod
 - ◆ přesná identifikace měřících nástrojů/přístrojů
 - důležité pro reprodukovatelnost měření

Chyby měření

- Klasifikace chyb podle místa vzniku
 - instrumentální (přístrojové) chyby
 - metodické chyby
 - teoretické chyby (principy, model)
 - chyby zpracování
- Klasifikace chyb podle původu
 - hrubé (omyly)
 - systematické
 - náhodné

Chyby měření



Chyby měření dle norem

- Metrologické normy

ČSN 01 0250 Statistické metody v průmyslové praxi. Všeobecné základy

ČSN 01 0251 Vzájemná shoda výsledků zkušebních metod. Stanovení opakovatelnosti a reprodukovatelnosti normalizované zkušební metody pomocí mezilaboratorních zkoušek

ČSN 25 0008 Metrologia. Chyby primárních etalónov. Spôsoby vyjadrovania

ČSN 25 1202 Posuvná měřidla. Technické požadavky

ČSN 25 1401 Mikrometrická měřidla na vnější měření. Technické požadavky

ČSN 25 8304 Provozní termoelektrické snímače teploty

ČSN 25 8305 Prevádzkové termoelektrické snímače teploty. Metody skúšania pri úradnom overovaní

ČSN 25 8306 Provozní odporové snímače teploty

ČSN 25 8307 Prevádzkové odporové snímače teploty. Metody overovania

ČSN 35 6505 Elektronické měřicí přístroje. Všeobecné technické podmínky

... a mnoho dalších

Chyby měření dle norem

- Termíny z ČSN 35 6505

Chyba rozdíl mezi údajem přístroje a skutečnou hodnotou

Absolutní chyba

Relativní chyba v %

Vztažná hodnota k níž se vztahuje relativní chyba

Základní chyba stanovená v referenčních podmínkách

Přídavná chyba jedna z hodnot nabývá libovolné hodnoty, ostatní jsou mají referenční hodnoty (a pak se neuvažuje základní chyba)

Chyba stálosti (stabilita) průběh chyby vytvářené samotným přístrojem v čase

Meze chyb maximální hodnoty chyb pro jakýkoli parametr ve stanovených podmínkách (referenčních, jmenovitých, pracovních, ...)

Chyby měření dle norem

- Termíny z ČSN 35 6505

Naměřená hodnota

Referenční podmínky souhrn podmínek a rozsahů pro parametry a ovlivňující veličiny, při nichž přístroj splňuje ustanovení o dovolených chybách, při kterých se u přístroje ověřuje základní chyba a/nebo se přístroje nastavují.

Jmenovitý rozsah použití rozsah hodnot, u nichž přístroj splňuje požadavky na chyby

Jmenovité pracovní podmínky souhrn pracovních hodnot, rozsahů, parametrů a ovlivňujících veličin, pro něž jsou udány technické vlastnosti přístroje

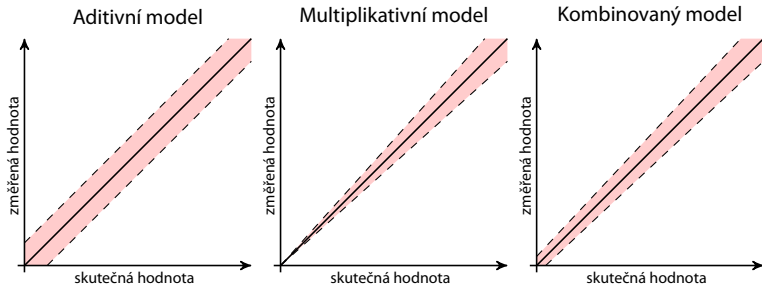
Doba náběhu přístroje

Přesnost měřících nástrojů

Přesnost přístroje ... náhodná chyba

Správnost přístroje ... systematická chyba

- Aditivní vs. multiplikativní chyby
- Mezní hodnota chyb
- Třída přesnosti přístroje



Přesnost měření času v počítači

- `gettimeofday()` – unixové API
 - přesnost závislá na použitém HW
 - v patologických případech může přeskočit i zpět
 - potenciální režie systémového volání
- `clock_gettime()` – POSIXové API
 - přesnost lze zjistit pomocí `clock_getres()`
 - `CLOCK_REALTIME` ve standardu
 - `CLOCK_MONOTONIC` jsou běžně dostupné
 - různé systémy poskytují různá rozšíření typu hodin (např. `CLOCK_REALTIME_FAST`, `CLOCK_REALTIME_PRECISE`, `CLOCK_PROCESS_CPUTIME_ID`, `CLOCK_THREAD_CPUTIME_ID`)
 - potenciální režie systémového volání

Přesnost měření času v počítači

- TSC – přístup k čítači x86 procesorů (od Pentia)
 - frekvence tiků rovna frekvenci procesoru
 - nízká režie – přímý přístup k čítači z ASM
 - problém absence synchronizace mezi procesory
 - ◆ nastavit afinitu
 - problém s dynamickou změnou frekvence procesoru
 - ◆ příznak `constant_tsc` v `/proc/cpuinfo` na Linuxu
 - problém s out-of-order vykonáváním instrukcí
 - ◆ předřadit serializující CPUID instrukci
 - problém resetu při uspání
 - ne všechny procesory jej mají (např. Cyrix 6x86)

Přesnost měření času v počítači

- `QueryPerformanceCounter` – Windows
 - frekvenci lze zjistit pomocí `QueryPerformanceFrequency`
 - opět třeba zamknout na procesor
- `System.currentTimeMillis()` – Java
 - ekvivalent `gettimeofday()/clock_gettime(CLOCK_REALTIME)`
 - nominální rozlišení 1 ms, fakticky i 10 ms v závislosti na OS
- `System.nanoTime()` – Java
 - přidání od JDK 1.5
 - aproximace TSC

Náhodné chyby

aneb proč se běžně pracuje s normálním rozdělením chyb?

- Hypotéza elementárních chyb [1]
 - každá náhodná chyba v měření je složena z řady malých chyb
 - při velkém počtu měření se vyskytne zhruba stejný počet chyb kladných i záporných a malé chyby jsou početnější než velké
- 1. m elementárních náhodných vlivů
- 2. každý elementární vliv generuje chybu α (dále označováno jako případ a) nebo $-\alpha$ (dále případ b)
- 3. chyby a a b jsou stejně časté
- dostáváme binomické rozdělení kumulace vlivů elementárních chyb

$$\binom{m}{0}a^m, \binom{m}{1}a^{m-1}b, \dots, \binom{m}{l}a^{m-l}b^l, \dots, \binom{m}{m}b^m$$

$$P(0) = \frac{1}{2^m} \binom{m}{m/2} \quad P(\varepsilon_l) = \frac{1}{2^m} \binom{m}{l}, \quad \varepsilon_l = (l - (m-l))\alpha = (2l - m)\alpha = 2s\alpha$$

Náhodné chyby

aneb proč se běžně pracuje s normálním rozdělením chyb?

- Co se stane, pokud $m \rightarrow \infty$?
 - pro sudá $m = 2k \implies k \rightarrow \infty$ (sudá, abychom měli $P(0)$)

$$P(\varepsilon) = P(2s\alpha) = \frac{1}{2^{2k}} \binom{2k}{k+s}$$

$$\frac{P(2s\alpha)}{P(0)} = \frac{\binom{2k}{k+s}}{\binom{2k}{k}} = \frac{k(k-1)\dots(k-s+1)}{(k+1)(k+2)\dots(k+s)} = \frac{\left(1 - \frac{1}{k}\right)\left(1 - \frac{2}{k}\right)\dots\left(1 - \frac{s-1}{k}\right)}{\left(1 + \frac{1}{k}\right)\left(1 + \frac{2}{k}\right)\dots\left(1 + \frac{s}{k}\right)}$$

- pro $s \ll k$

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \dots \approx x$$

$$\ln \frac{P(2s\alpha)}{P(0)} = -\frac{1}{k} - \frac{2}{k} - \dots - \frac{s-1}{k} - \frac{1}{k} - \frac{2}{k} - \dots - \frac{s}{k} = -\frac{2}{k} \frac{s(s-1)}{2} - \frac{s}{k} = -\frac{s^2}{k}$$

$$P(2s\alpha) = P(0)e^{-\frac{s^2}{k}} = P(0)e^{-\frac{\varepsilon^2}{4k\alpha^2}}$$

Náhodné chyby

aneb proč se běžně pracuje s normálním rozdělením chyb?

- Co se stane, pokud $m \rightarrow \infty$?
 - převod na spojitě rozdělení

$$h^2 = \frac{1}{\lim_{k \rightarrow \infty} 4k\alpha^2}, \quad \eta(\varepsilon) = \frac{h}{\sqrt{\pi}} e^{-h^2 \varepsilon^2}$$

- šikmost binomického rozdělení

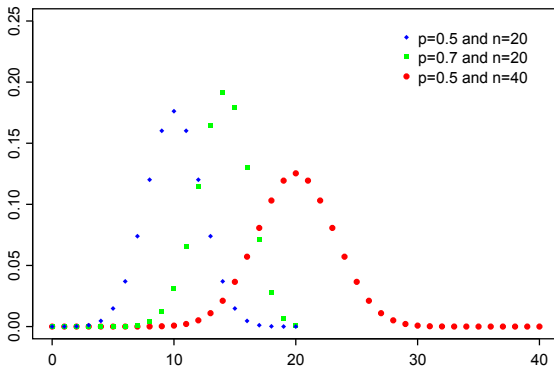
$$\frac{1 - 2p}{\sqrt{np(1-p)}}, \quad \lim_{n \rightarrow \infty} \frac{1 - 2p}{\sqrt{np(1-p)}} = 0$$

a také 0 pro $p = 0,5$

- další studium: Central Limit Theorem, [2]

Binomické vs. normální rozdělení

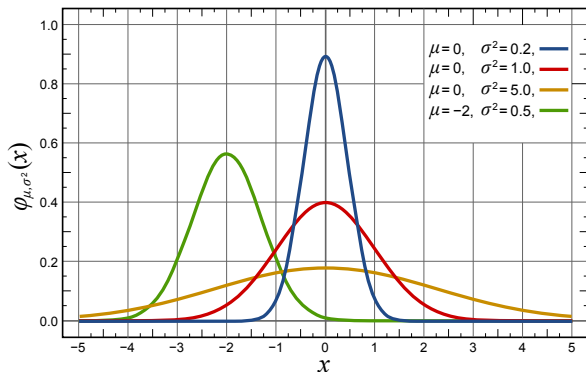
Binomické rozdělení $\binom{n}{k} p^k (1-p)^{n-k}$



Zdroj: http://en.wikipedia.org/wiki/File:Binomial_distribution_pmf.svg

Binomické vs. normální rozdělení

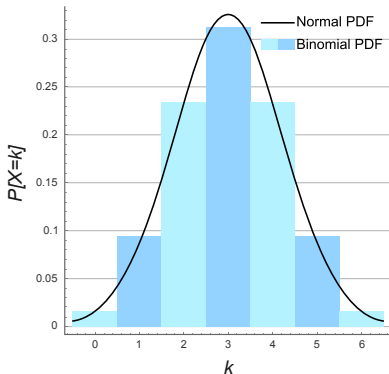
Normální rozdělení $\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$



Zdroj: http://en.wikipedia.org/wiki/File:Normal_Distribution_PDF.svg

Binomické vs. normální rozdělení

Srovnání binomického a normálního rozdělení pro $p = 0,5$ a $n = 6$



Zdroj: http://en.wikipedia.org/wiki/File:Binomial_Distribution.svg

Studentovo rozdělení t

- Používá se pro normální rozdělení při malém vzorku

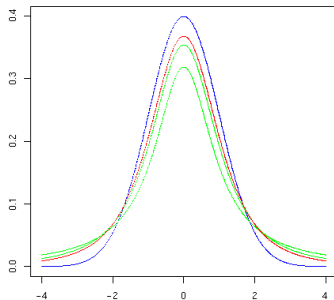
$$f(t) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{t^2}{\nu}\right)^{-(\nu+1)/2}$$

kde ν je počet stupňů volnosti.

- odhad průměrů a chyby
- t-test – odlišení průměrů

Studentovo rozdělení t

Srovnání s normálním rozdělením (modré)
počet stupňů volnosti $\nu = 3$



Zdroj: http://en.wikipedia.org/wiki/File:T_distribution_3df.png

Poissonovo rozdělení

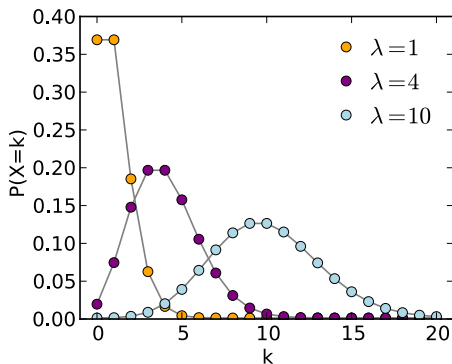
- Počty událostí v daném časovém okně, odehrávají se nezávisle se známou průměrnou rychlostí

$$f(k; \lambda) = \frac{\lambda^k e^{-\lambda}}{k!}$$

k ... počet výskytů událostí, λ ... očekávaný počet událostí ve studovaném intervalu

- Příklady – Poissonovské procesy
 - počet telefonních hovorů na ústředně za minutu
 - počet přístupů k webovému serveru (nemění-li se λ v čase – předpoklad homogenity)
 - radioaktivní rozpad atomů
- Pro $\lambda \rightarrow \infty$ je opět dobrou aproximací normální rozdělení

Poissonovo rozdělení



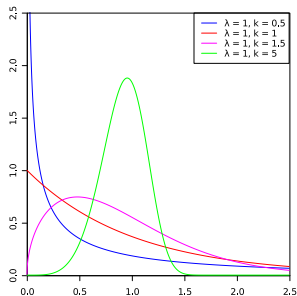
Zdroj: http://en.wikipedia.org/wiki/File:Poisson_pmf.svg

Jiná rozdělení

- Weibullovo

$$f(x; \lambda, k) = \begin{cases} \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k} & x \geq 0, \\ 0 & x < 0. \end{cases}$$

- používá se k popisu času do selhání, úmrtí
- $k < 1$ – rychlost selhání klesá v čase, jak z vzorku mizí kusy (např. úmrtnost novorozenců)
- $k = 1$ – rychlost selhání je konstantní v čase, typicky způsobena vnějšími vlivy (např. úmrtnost vojáků ve válce)
- $k > 1$ – selhání vzrůstá v čase, typický proces stárnutí komponent



Zdroj:

http://en.wikipedia.org/wiki/File:Weibull_PDF.svg

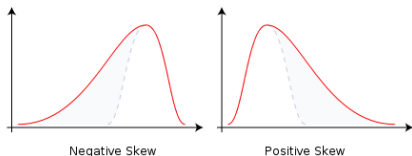
Jiná rozdělení

- modely rozdělení specifické pro aplikace
 - Rayleighovo rozdělení – rychlost větru ve 2D složkách
 - záření černého tělesa
- nejedná se o „chyby“, ale o charakteristiku

Normalizace rozdělení chyb

- Ověření normality rozdělení
 - vizuální
 - šikmost vzorku (sample skewness)

$$g_1 = \frac{\frac{1}{n} \sum_{i=1} N(x_i - \bar{x})^3}{\left(\frac{1}{n} \sum_{i=1} N(x_i - \bar{x})^2\right)^{3/2}}$$



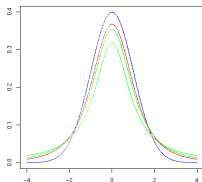
Zdroj: http://en.wikipedia.org/wiki/File:Skewness_Statistics.svg

Normalizace rozdělení chyb

- Ověření normality rozdělení
 - špičatost vzorku (sample kurtosis)

$$g_2 = \frac{\frac{1}{n} \sum_{i=1}^n N(x_i - \bar{x})^4}{\left(\frac{1}{n} \sum_{i=1}^n N(x_i - \bar{x})^2\right)^2} - 3$$

lehké konce (leptokurtic), $g_2 > 0$ × těžké konce (platycurtic), $g_2 < 0$



Zdroj: http://en.wikipedia.org/wiki/File:T_distribution_3df.png

Normalizace rozdělení chyb

- Techniky normalizace

- šikmá rozdělení $g_1 > 0$: transformace hodnot
 - ◆ $\sqrt[n]{x}$
 - ◆ $\log(x)$
 - ◆ $\frac{1}{x}$
- šikmá rozdělení $g_1 < 0$: převrácení hodnot (reflection)
 - ◆ $-x + c$ s vhodně zvolenou konstantou c
- špičatá rozdělení: problém
- další čtení: [3]

Odhad spolehlivosti

$$x = (\hat{\mu}_x \pm z_x) \text{ [jednotka]}$$

Statistická definice [4]: Je-li výsledek měření $\hat{\mu}_x$ a z_x je chyba tohoto měření odpovídající míře jistoty p , pak skutečná hodnota měřené veličiny leží v intervalu $(\hat{\mu}_x \pm z_x)$ s pravděpodobností p .

- Intervaly
 - 0,68 – střední kvadratická chyba
 - 0,95
 - 0,99 – krajní chyba
- Zaokrouhlování
 - z_x nejvýše na 2 platná místa
 - $\hat{\mu}_x$ podle z_x

Odhad spolehlivosti

$$x = (\hat{\mu}_x \pm z_x) [\text{jednotka}]$$

Pro normální rozdělení chyby

- $\hat{\mu}_x = \bar{x} = \frac{\sum_{i=1}^N x_i}{n}$
- s směrodatná odchylka jednoho měření, D rozptyl

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

- $s_{\bar{x}} = \sqrt{\sum_{i=1}^N \left(\frac{1}{n}\right)^2 s_{x_i}}$ a protože měření byly prováděny za stejných podmínek

$$s_{\bar{x}} = \frac{s_x}{\sqrt{n}} = \sqrt{\frac{\sum_{i=1}^N (\bar{x} - x_i)^2}{n(n - 1)}}$$

Odhad spolehlivosti

$$x = (\hat{\mu}_x \pm z_x) [\text{jednotka}]$$

Pro normální rozdělení chyby

- $z_x = t_{(p;n-1)} s_{\bar{x}}$

n \ P	P			n \ P	P		
	0,683	0,954	0,99		0,683	0,954	0,99
1	1,8395	13,8155	63,6567	16	1,0329	2,1633	2,9208
2	1,3224	4,5001	9,9248	18	1,0292	2,1433	2,8784
3	1,1978	3,2923	5,8409	20	1,0263	2,1276	2,8453
4	1,1425	2,8585	4,6041	30	1,0176	2,0817	2,75
5	1,1113	2,6396	4,0321	40	1,0133	2,0595	2,7045
6	1,0913	2,5084	3,7074	50	1,0108	2,0463	2,6778
7	1,0775	2,4214	3,4995	60	1,0091	2,0377	2,6603
8	1,0673	2,3594	3,3554	70	1,0078	2,0315	2,6479
9	1,0594	2,3131	3,2498	80	1,0069	2,0269	2,6387
10	1,0533	2,2773	3,1693	90	1,0062	2,0234	2,6316
12	1,0441	2,2253	3,0545	100	1,0057	2,0206	2,6259
14	1,0377	2,1895	2,9768				

Odhad spolehlivosti

$$x = (\hat{\mu}_x \pm z_x) [\text{jednotka}]$$

Příklad – měření výšky válečku [4]:

výška v [mm]	4,6	4,5	4,7	4,4	4,5	4,6	4,4	4,4	4,3	4,5
----------------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

- $n = 10$
- $\bar{v} = 4,49$ [mm]
- $s_{\bar{v}} = 0,038$ [mm]
- $t_{(0,68;9)} = 1,059$
- $t_{(0,99;9)} = 3,250$

$$v = (4,49 \pm 0,04) \text{ mm} \quad \text{pro } p = 0,68$$

$$v = (4,49 \pm 0,12) \text{ mm} \quad \text{pro } p = 0,99$$

Chyba nepřímo měřené veličiny

- K odhadu střední hodnoty a rozptylu lze použít
 - Taylorův rozvoj funkce

$$f(x + \varepsilon) = f(x) + \sum_{n=1}^{\infty} \frac{f^{(n)}(x)}{n!} \varepsilon^n$$

kde $f^{(n)}(x)$ je n -tá derivace f ,

- dvoubodovou aproximaci $y = f(x_1, \dots, x_m)$

$$\bar{y} = \sum_{i=1}^m \frac{f(\bar{x}_i + s_{\bar{x}_i}) + f(\bar{x}_i - s_{\bar{x}_i})}{2m} \quad s_y^2 = \sum_{i=1}^m \frac{[f(\bar{x}_i + s_{\bar{x}_i}) - f(\bar{x}_i - s_{\bar{x}_i})]^2}{4m}$$

- Monte Carlo simulace

Zákon přenosu chyb

- Na základě Taylorova rozvoje do druhého členu

$$s_z^2 = \sum_{i=1}^N \left(\frac{\partial z}{\partial x_i} \right)^2 s_{x_i}^2 + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N \frac{\partial z}{\partial x_i} \frac{\partial z}{\partial x_j} s_{x_i} s_{x_j} \rho_{ij},$$

kde $s_{x_i}^2$ je rozptyl (variance) x_i a ρ_{ij} je kovariance x_i a x_j .

Pro jednoduché případy, kdy x a y jsou nezávislé ($\rho_{ij} = 0$):

- aditivní funkce $z = ax \pm by$

$$s_z = \sqrt{a^2 s_x^2 + b^2 s_y^2}, \quad (1)$$

- multiplikativní funkce $z = ax^b y^c$

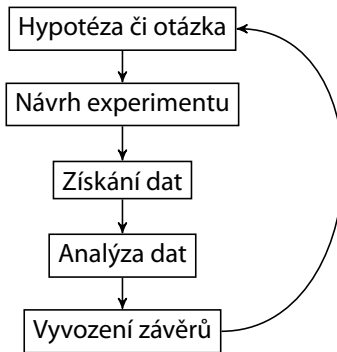
$$s_z = \bar{z} \sqrt{\left(\frac{b s_x}{x} \right)^2 + \left(\frac{c s_y}{y} \right)^2}. \quad (2)$$

kde $\bar{z} = a \bar{x}^b \bar{y}^c$, protože

$$\sum_{i=1}^N \left(\frac{\partial z}{\partial x_i} \right)^2 s_i^2 = \left(\frac{abx^b y^c s_x}{x} \right)^2 + \left(\frac{ax^b cy^c s_y}{y} \right)^2 = z^2 \left(\left(\frac{b s_x}{x} \right)^2 + \left(\frac{c s_y}{y} \right)^2 \right)$$

- Příklad použití: <http://www.phy.ohiou.edu/~murphy/courses/sample.pdf>

Fáze experimentu



Hypotéza

- Přesná formulace hypotézy či otázky
 - do experimentu je třeba vstupovat s hypotézou, kterou se snažíme podpořit nebo vyvrátit
 - případně s otázkou, kterou se snažíme zodpovědět
 - ... avšak bez očekávání výsledku!

Návrh experimentu

- Volba nezávislých a závislých proměnných: (x_i, y_i)
 - nezávislé proměnné x_i
 - ◆ měníme v průběhu experimentu
 - ◆ nabývají minimálně 2 hodnot
 - závislé proměnné y_i
 - ◆ mění se v závislosti na změně nezávislých proměnných
 - ◆ měříme je

Příklad:

- *Otázka:* Jaký vliv má alkohol na reakční dobu?
- *Nezávislá proměnné:* Množství alkoholu pro jednotlivce/skupiny (např. 0 dcl – referenční skupina, 2 dcl, 4 dcl, ...).
- *Závislé proměnné:* Reakční doba.

Návrh experimentu

- Proměnné subjektu
 - vlastnosti subjektu, které nemohou být změněny experimentátorem.
 - např. věk, pohlaví a IQ osob, technické parametry počítačů, ...
- Externí (kontrolované) proměnné
 - proměnné, které nejsou v rámci experimentu studovány, ale musí být pod kontrolou, protože jejich změna může ovlivnit průběh experimentu.
 - např: procesy běžící na počítači, přenosy běžící na síti

Příklad:

- *Otázka:* Jaký vliv má alkohol na reakční dobu?
- *Proměnné subjektu:* Hmotnost člověka, stav jater, ...
- *Externí proměnné:* Doba od podání alkoholu do okamžiku měření reakční doby, současně brané farmaka, způsob podání alkoholu, ...

Návrh experimentu

- Více nezávislých proměnných
 - současná změna × nezávislá změna
 - jaké jsou vztahy mezi proměnnými?

Příklad (do not try this! ;)):

- *Otázka:* Jaký vliv má alkohol na reakční dobu při současném požívání tuků a diazepamu?
- *Nezávislé proměnné:* Množství alkoholu, množství tuku, množství diazepamu.
- *Kvalitativní úroveň:* nic, alkohol, tuk, diazepam, alkohol+tuk, alkohol+diazepam, tuk+diazepam, alkohol+tuk+diazepam

Návrh experimentu

- Volba měřící metody / měřících metod
 - ovlivnění měřeného systému samotnou měřící metodou
 - ◆ např. zpomalení voláním `clock_gettime()`
 - ovlivnění měřící metody měřeným systémem mimo předpokládané vlivy
 - ◆ např. interakce se plánovače procesů, obsluhy přerušení a časovačů
 - vzájemné interakce měřících metod
 - ◆ např. profilování kódu (počítání průchodů větvemi) ovlivňuje měření času (a můžeme potřebovat obé současně)

Návrh experimentu

- Volba měřící metody / měřících metod – příklad pro Javu
 - Problém garbage collection
 - ◆ `-verbose:gc`
 - ◆ krátká měření: vybrat pouze běhy, v nichž nedošlo ke GC
 - ◆ dlouhé běhy: dostatečně dlouhé, aby se přítomnost GC projevila reprezentativně
 - Problém HotSpot kompilace
 - ◆ `-XX:+PrintCompilation`
 - ◆ dostatečný warm-up (minuty!)
 - ◆ mohou se vyskytovat rekompilace (optimalizace, nahrání nové třídy která zruší dosavadní předpoklady)
 - ◆ housekeeping tasks: oddělení nesouvisejících měření pauzou nebo restartem JVM

Návrh experimentu

- Opakování měření

- teoreticky můžeme zvyšováním počtu stejně přesných měření dosáhnout libovolné přesnosti

$$s_{\bar{x}} = \frac{s}{\sqrt{n}}$$

- alespoň 3 měření pro odhalení hrubé chyby
- běžný počet opakování je 5–20
- velký počet opakování: pokles chyby zpomaluje a je obtížné zajistit stálost vnějších podmínek v čase
- zjištění vlivu jednotlivých nezávislých proměnných na chybu

Model

- Mapování matematického modelu na naměřené hodnoty
 - hledáme parametry modelu
 - minimalizujeme odchylky (rezidua) modelu od naměřených dat

$$r_i(\mathbf{x}) = \|y_i - M(\mathbf{x})\|$$

příp. vyjádřeno jako minimalizace normy vektoru

$$\mathbf{r}(\mathbf{x}) = (r_1(\mathbf{x}), \dots, r_m(\mathbf{x}))^T$$

- nejčastěji pracujeme s euklidovskou L_2 normou (metoda nejmenších čtverců)

$$f(\mathbf{x}) = \mathbf{r}(\mathbf{x})^T \mathbf{r}(\mathbf{x}) = \sum_{i=1}^m r_i(\mathbf{x})^2$$

- lze použít např. i L_1 (součet absolutních hodnot – méně citlivé na data s větší kumulací chyb, příp. zatížená hrubou chybou) či L_∞ (maximum z absolutních hodnot)

Model

- Metoda nejmenších čtverců

- mějme data (x_i, y_i) , kde x_i je nezávislá proměnná a y_i je závislá (měřená proměnná)
- minimalizujeme $S = \sum_{i=1}^n r_i^2 = \sum_{i=1}^n (y_i - f(x_i, \mathbf{c}))^2$, kde \mathbf{c} je vektor parametrů
- hledáme minimum vzhledem k \mathbf{c} , tedy

$$\frac{\partial S}{\partial c_j} = 2 \sum_i r_i \frac{\partial r_i}{\partial c_j} = -2 \sum_i \frac{\partial f(x_i, \mathbf{c})}{\partial c_j} r_i = 0 \quad j = 1, \dots, m$$

Model

- Lineární kombinace elementárních funkcí

$$f(x_i, \mathbf{c}) = \sum_{j=1}^m c_j \phi_j(x_i)$$

- ϕ_j mohou být polynomy, podíly polynomů, trigonometrické funkce, exponenciální funkce, ...

$$X_{ij} = \frac{\partial f(x_i, \mathbf{c})}{\partial c_j} = \phi_j(x_i)$$

$$\hat{\mathbf{c}} = (X^T X)^{-1} X^T \mathbf{y}$$

Model

- Příklad lineární funkce
 - minimalizujeme

$$f(x_i, (a, b)) = a + bx_i$$

$$Q = \sum_{i=1}^n (y_i - a - bx_i)^2$$

$$\frac{\partial Q}{\partial a} = 2na + \sum_{i=1}^n (-2y_i + 2bx_i) = 0 \quad (3)$$

$$\frac{\partial Q}{\partial b} = \sum_{i=1}^n (-2y_i x_i + 2ax_i + 2bx_i^2) = 0 \quad (4)$$

- dvě rovnice (3) a (4) o dvou neznámých a a b

$$a = - \frac{- \sum_{i=1}^n y_i \sum_{i=1}^n x_i^2 + \sum_{i=1}^n x_i \sum_{i=1}^n y_i x_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}$$

$$b = \frac{n \sum_{i=1}^n y_i x_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}$$

Hodnocení modelu

- Pearsonův korelační koeficient

$$r_{x,y} = \frac{\sum_{i=1}^n ((x_i - \bar{x})(y_i - \bar{y}))}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot \sum_{i=1}^n (y_i - \bar{y})^2}}$$

- lineární závislost dvou veličin x a y a nabývá hodnot $[-1;1]$
 - 1 ... přesná souhlasná závislost,
-1 ... přesná inverzní závislost,
0 nezávislé
 - využívá se často jako $r_{x,y}^2$
- Root mean square deviation – RMSD

$$\text{RMSD}_{x,y} = \sqrt{\frac{\sum_{i=1}^n (x_i - y_i)^2}{n}}$$

- srovnání mezi získaným modelem a originálními hodnotami

Prezentace výsledků



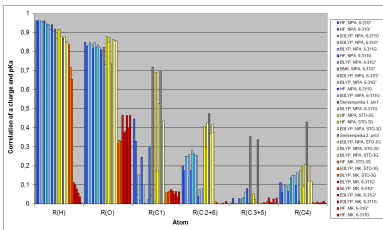
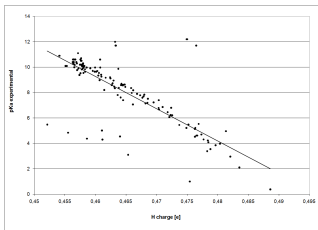
SINGLE RESPONSIBILITY PRINCIPLE

Just Because You Can, Doesn't Mean You Should

Zdroj: <http://www.lostechies.com/blogs/derickbailey/archive/2009/02/11/solid-development-principles-in-motivational-pictures.aspx>

Prezentace výsledků

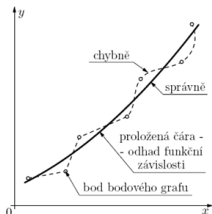
- Co dělat
 - popisovat osy grafů i s jednotkami
 - popisovat sloupce/řádky v tabulkách, nezapomínat na jednotky



Prezentace výsledků

- Co nedělat

- neprokládat n bodů polynomem $n - 1$ řádu
- nepoužívat v grafech spojnice (zejm. spline) mezi jednotlivými body vždy a za všech okolností



- být opatrný s rozlišováním v grafech pomocí barev – řada článků se pořád tiskne černobíle

Literatura I



Zdeněk Horák.

Praktická fyzika.

Státní nakladatelství technické literatury, Praha, 1958.



Patrick L. Brockett.

On the misuse of the central limit theorem in some risk calculations.

The Journal of Risk and Insurance, 50(4):727–731, December 1983.

<http://www.jstor.org/stable/pdfplus/252712.pdf>.



Jason W. Osborne.

Normalizing data transformations. ERIC digest.

Technical report, ERIC Clearinghouse on Assessment and Evaluation

College Park MD, August 2002.

<http://www.ericdigests.org/2003-3/data.htm>.

Literatura II



František Šťastný.

Zpracování experimentálních dat.

Katedra obecné fyziky PřF MU, Brno, 1997.

http://amper.ped.muni.cz/jenik/nejistoty/frst_zed.pdf.



Milan Meloun and Jiří Militký.

Data analysis in the chemical laboratory part 1. analysis of indirect measurements.

Analytica Chimica Acta, 293(1-2):183–189, 1994.

<http://www.sciencedirect.com/science/article/B6TF4-44HT11Y-6D/2/eb0dc71f565eaf9211806cb31425a66a>.

Tento projekt je spolufinancován Evropským sociálním fondem a státním rozpočtem České republiky.



INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

Inovace doktorského studia na Fakultě informatiky MU (IDSnaFI)
(CZ.1.07/2.2.00/15.0196)