

Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

DIPLOMOVÁ PRÁCE



Jméno a příjmení autora

Název práce

Katedra

Vedoucí diplomové práce: RNDr. Jméno Vedoucí, Ph.D.

Studijní program: Matematika

Studijní obor: Obor

Praha 2013

Poděkování (nepovinné).

Prohlašuji, že jsem tuto diplomovou práci vypracoval(a) samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona v platném znění, zejména skutečnost, že Univerzita Karlova v Praze má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle §60 odst. 1 autorského zákona.

V dne

Podpis autora

Název práce: Název práce v češtině dle zadání

Autor: Jméno a příjmení autora

Katedra: Název katedry či ústavu, kde byla práce oficiálně zadána

Vedoucí diplomové práce: RNDr. Jméno Vedoucí, Ph.D., pracoviště

Abstrakt: Český abstrakt v rozsahu 80–200 slov. Nejedná se o opis zadání diplomové práce

Klíčová slova: 3 až 5 klíčových slov

Title: Název práce v angličtině dle SISu

Author: Jméno a příjmení autora

Department: Department

Supervisor: RNDr. Jméno Vedoucí, Ph.D., Department

Abstract: Anglický abstrakt v rozsahu 80–200 slov. Nejedná se o překlad zadání diplomové práce.

Keywords: 3 až 5 klíčových slov v angličtině

Názov práce: Názov práce preložený do slovenčiny

Autor: Meno a priezvisko autora

Katedra: Názov katedry či ústavu, kde byla práce oficiálně zadána

Vedúci diplomovej práce: RNDr. Jméno Vedoucí, Ph.D., pracoviště

Abstrakt: Slovenský abstrakt v rozsahu 80–200 slov. Nejedná sa o preklad zadania diplomovej práce. Táto stránka sa vkladá iba do slovenských prác.

Klíčové slová: 3 až 5 klíčových slov vo slovenčině

Obsah

1	Sazba matematického textu	2
1.1	Několik jednoduchých ukázek	2
1.2	Matematické vzorce a výrazy	2
1.3	Definice, věty, důkazy,	4
2	Odkazy na literaturu	5
2.1	Několik ukázek	5
3	Tabulky, obrázky, softwarový kód	6
3.1	Tabulky	6
3.2	Obrázky	7
3.3	Softwarový kód	7
	Literatura	12
	Seznam obrázků	13
	Seznam tabulek	14

Kapitola 1

Sazba matematického textu

1.1 Několik jednoduchých ukázek

Číslo v matematickém režimu s desetinnou čárkou: $\pi \doteq 3,141\,592\,653\,589$.

Test na hladině 5 % (mezera mezi 5 a %), ale 95% (není mezera mezi 95 a %)
interval spolehlivosti.

Platí: $\text{var}(X) = E X^2 - (E X)^2$.

1.2 Matematické vzorce a výrazy

Nechť

$$\mathbb{X} = \begin{pmatrix} \mathbf{x}_1^\top \\ \vdots \\ \mathbf{x}_n^\top \end{pmatrix}.$$

Povšimněme si tečky za maticí. Byť je matematický text vysázen ve specifickém prostředí, stále je gramaticky součástí věty a tudíž je zapotřebí neopomenout patričná interpunkční znaménka. Výrazy, na které chceme později odkazovat, je vhodné očíslovat:

$$\mathbb{X} = \begin{pmatrix} \mathbf{x}_1^\top \\ \vdots \\ \mathbf{x}_n^\top \end{pmatrix}. \quad (1.1)$$

Výraz (1.1) definuje matici \mathbb{X} . Pro lepší čitelnost a přehlednost textu je vhodné číslovat pouze ty výrazy, na které se autor někde v další části textu odkazuje. To jest, nečíslujte automaticky všechny výrazy vysázené některým z matematických prostředí.

Zarovnání vzorců do několika sloupečků:

$$\begin{aligned} S(t) &= P(T > t), & t > 0 & \quad (\text{zprava spojitá}), \\ F(t) &= P(T \leq t), & t > 0 & \quad (\text{zprava spojitá}). \end{aligned}$$

Dva vzorce se spojovníkem:

$$\left. \begin{aligned} S(t) &= P(T > t) \\ F(t) &= P(T \leq t) \end{aligned} \right\} t > 0 \quad (\text{zprava spojité}). \quad (1.2)$$

Dva centrovane nečíslované vzorce:

$$\mathbf{Y} = \mathbb{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

$$\mathbb{X} = \begin{pmatrix} 1 & \mathbf{x}_1^\top \\ \vdots & \vdots \\ 1 & \mathbf{x}_n^\top \end{pmatrix}.$$

Dva centrovane číslované vzorce:

$$\mathbf{Y} = \mathbb{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \quad (1.3)$$

$$\mathbb{X} = \begin{pmatrix} 1 & \mathbf{x}_1^\top \\ \vdots & \vdots \\ 1 & \mathbf{x}_n^\top \end{pmatrix}. \quad (1.4)$$

Definice rozdelená na dva případy:

$$P_{r-j} = \begin{cases} 0, & \text{je-li } r-j \text{ liché,} \\ r!(-1)^{(r-j)/2}, & \text{je-li } r-j \text{ sudé.} \end{cases}$$

Všimněte si použití interpunkce v této konstrukci. Čárky a tečky se dávají na místa, kam podle jazykových pravidel patří.

$$\begin{aligned} x &= y_1 - y_2 + y_3 - y_5 + y_8 - \dots = && \text{z (1.3)} \\ &= y' \circ y^* = && \text{podle (1.4)} \\ &= y(0)y' && \text{z Axiomu 1.} \end{aligned} \quad (1.5)$$

Dva zarovnané vzorce nečíslované:

$$L(\boldsymbol{\theta}) = \prod_{i=1}^n f_i(y_i; \boldsymbol{\theta}),$$

$$\ell(\boldsymbol{\theta}) = \log\{L(\boldsymbol{\theta})\} = \sum_{i=1}^n \log\{f_i(y_i; \boldsymbol{\theta})\}.$$

Dva zarovnané vzorce, první číslovaný:

$$L(\boldsymbol{\theta}) = \prod_{i=1}^n f_i(y_i; \boldsymbol{\theta}), \quad (1.6)$$

$$\ell(\boldsymbol{\theta}) = \log\{L(\boldsymbol{\theta})\} = \sum_{i=1}^n \log\{f_i(y_i; \boldsymbol{\theta})\}.$$

Vzorec na dva řádky, první řádek zarovnaný vlevo, druhý vpravo, nečíslovaný:

$$\begin{aligned} \ell(\mu, \sigma^2) &= \log\{L(\mu, \sigma^2)\} = \sum_{i=1}^n \log\{f_i(y_i; \mu, \sigma^2)\} = \\ &= -\frac{n}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \mu)^2. \end{aligned}$$

Vzorec na dva řádky, zarovnaný na =, číslovaný uprostřed:

$$\begin{aligned}\ell(\mu, \sigma^2) &= \log\{L(\mu, \sigma^2)\} = \sum_{i=1}^n \log\{f(y_i; \mu, \sigma^2)\} = \\ &= -\frac{n}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \mu)^2.\end{aligned}\tag{1.7}$$

1.3 Definice, věty, důkazy, ...

Konstrukce typu definice, věta, důkaz, příklad, ... je vhodné odlišit od okolního textu a případně též číslovat s možností použití křížových odkazů. Pro každý typ těchto konstrukcí je vhodné mít v hlavním souboru (`BcPrace.tex`) nadefinované jedno prostředí, které zajistí jak vizuální odlišení od okolního textu, tak automatickou tvorbu čísel s možností křížově odkazovat.

Definice 1. *Nechť náhodné veličiny X_1, \dots, X_n jsou definovány na témž pravděpodobnostním prostoru (Ω, \mathcal{A}, P) . Pak vektor $\mathbf{X} = (X_1, \dots, X_n)^\top$ nazveme náhodným vektorem.*

Definice 2 (náhodný vektor). *Nechť náhodné veličiny X_1, \dots, X_n jsou definovány na témž pravděpodobnostním prostoru (Ω, \mathcal{A}, P) . Pak vektor $\mathbf{X} = (X_1, \dots, X_n)^\top$ nazveme náhodným vektorem.*

Definice 1 ukazuje použití prostředí pro sazbu definice bez titulku, definice 2 ukazuje použití prostředí pro sazbu definice s titulkem.

Věta 1. *Náhodný vektor \mathbf{X} je měřitelné zobrazení prostoru (Ω, \mathcal{A}, P) do $(\mathbb{R}_n, \mathcal{B}_n)$.*

Lemma 2 (Anděl, 2007, str. 29). *Náhodný vektor \mathbf{X} je měřitelné zobrazení prostoru (Ω, \mathcal{A}, P) do $(\mathbb{R}_n, \mathcal{B}_n)$.*

Důkaz. Jednotlivé kroky důkazu jsou podrobně popsány v práci Anděl (2007, str. 29). □

Věta 1 ukazuje použití prostředí pro sazbu matematické věty bez titulku, lemma 2 ukazuje použití prostředí pro sazbu matematické věty s titulkem. Lemmata byla zavedena v hlavním souboru tak, že sdílejí číslování s větami.

Kapitola 2

Odkazy na literaturu

Odkazy na literaturu vytváříme nejlépe pomocí příkazů `\citet`, `\citep` atp. (viz L^AT_EXový balíček `natbib`) a následného použití `BibTEXu`. V matematickém textu obvykle odkazujeme stylem „Jméno autora/autorů (rok vydání)“, resp. „Jméno autora/autorů [číslo odkazu]“. V českém/slovenském textu je potřeba se navíc vypořádat s nutností skloňovat jméno autora, respektive přechylovat jméno autorky. Je potřeba mít na paměti, že standardní příkazy `\citet`, `\citep` produkují referenci se jménem autora/autorů v prvním pádě a jména autorek jsou nepřechýlena.

2.1 Několik ukázek

Mezi nejvíce citované statistické články patří práce Kaplana a Meiera a Coxe (Kaplan a Meier, 1958; Cox, 1972). Student (1908) napsal článek o t-testu.

Prof. Anděl je autorem učebnice matematické statistiky (viz Anděl, 1998). Teorii odhadu se věnuje práce Lehmann a Casella (1998). V případě odkazů na specifickou informaci (definice, důkaz, ...) uvedenou v knize bývá užitečné uvést specificky číslo kapitoly, číslo věty atp. obsahující požadovanou informaci, např. viz Anděl (2007, Věta 4.22) nebo (viz Anděl, 2007, Věta 4.22).

Mnoho článků je výsledkem spolupráce celé řady osob. Při odkazování v textu na článek se třemi autory obvykle při prvním výskytu uvedeme plný seznam: Dempster, Laird a Rubin (1977) představili koncept EM algoritmu. Respektive: Koncept EM algoritmu byl představen v práci Dempstera, Lairdové a Rubina (Dempster, Laird a Rubin, 1977). Při každém dalším výskytu již používáme zkrácenou verzi: Dempster a kol. (1977) nabízejí též několik příkladů použití EM algoritmu. Respektive: Několik příkladů použití EM algoritmu lze nalézt též v práci Dempstera a kol. (Dempster a kol., 1977).

U článku s více než třemi autory odkazujeme vždy zkrácenou formou: První výsledky projektu ACCEPT jsou uvedeny v práci Genbergové a kol. (Genberg a kol., 2008). V textu *nenapíšeme*: První výsledky projektu ACCEPT jsou uvedeny v práci Genberg, Kulich, Kawichai, Modiba, Chingono, Kilonzo, Richter, Pettifor, Sweat a Celentano (2008).

Kapitola 3

Tabulky, obrázky, softwarový kód

Používání tabulek a grafů v odborném textu má některá společná pravidla a některá specifická. Tabulky a grafy neuvádíme přímo do textu, ale umístíme je buď na samostatné stránky nebo na vyhrazené místo v horní nebo dolní části běžných stránek. \LaTeX se o umístění plovoucích grafů a tabulek postará automaticky. Každý graf a tabulku očíslováme a umístíme pod ně legendu. Legenda má popisovat obsah grafu či tabulky tak podrobně, aby jim čtenář rozuměl bez důkladného studování textu práce. Na každou tabulku a graf musí být v textu odkaz pomocí jejich čísla. Na příslušném místě textu pak shrneme ty nejdůležitější závěry, které lze z tabulky či grafu učinit. Text by měl být čitelný a srozumitelný i bez prohlížení tabulek a grafů a tabulky a grafy by měly být srozumitelné i bez podrobné četby textu. Na tabulky a grafy odkazujeme pokud možno nepřímo v průběhu běžného toku textu; místo „*Tabulka 3.1 ukazuje, že muži jsou v průměru o 9,9 kg těžší než ženy*“ raději napíšeme „*Muži jsou o 9,9 kg těžší než ženy (viz Tabulka 3.1)*“.

3.1 Tabulky

U **tabulek** se doporučuje dodržovat následující pravidla:

- Vyhýbat se svislým linkám. Silnějšími vodorovnými linkami oddělit tabulku od okolního textu včetně legendy, slabšími vodorovnými linkami oddělovat záhlaví sloupců od těla tabulky a jednotlivé části tabulky mezi sebou. V \LaTeX u tuto podobu tabulek implementuje balík `booktabs`. Chceme-li výrazněji oddělit některé sloupce od jiných, vložíme mezi ně větší mezeru.
- Neměnit typ, formát a význam obsahu políček v tomtéž sloupci (není dobré do téhož sloupce zapisovat tu průměr onde procenta).

Efekt	Odhad	Směrod. chyba ^a	P-hodnota
Abs. člen	-10,01	1,01	—
Pohlaví (muž)	9,89	5,98	0,098
Výška (cm)	0,78	0,12	< 0,001

Pozn: ^a Směrodatná chyba odhadu metodou Monte Carlo.

Tabulka 3.1: Maximálně věrohodné odhady v modelu M.

- Neopakovat tentýž obsah políček mnohokrát za sebou. Máme-li sloupec *Rozptyl*, který v prvních deseti řádcích obsahuje hodnotu 0.5 a v druhých deseti řádcích hodnotu 1.5, pak tento sloupec raději zrušíme a vyřešíme to jinak. Například můžeme tabulku rozdělit na dvě nebo do ní vložit popisné řádky, které informují o nějaké proměnné hodnotě opakující se v následujícím oddíle tabulky (např. „*Rozptyl = 0.5*“ a níže „*Rozptyl = 1.5*“).
- Čísla v tabulce zarovnávat na desetinnou tečku.
- V tabulce je někdy potřebné používat zkratky, které se jinde nevyskytují. Tyto zkratky můžeme vysvětlit v legendě nebo v poznámkách pod tabulkou. Poznámky pod tabulkou můžeme využít i k podrobnějšímu vysvětlení významu některých sloupců nebo hodnot.

3.2 Obrázky

Několik rad týkajících se obrázků a grafů.

- Graf by měl být vytvořen ve velikosti, v níž bude použit v práci. Zmenšení příliš velkého grafu vede ke špatné čitelnosti popisků.
- Osy grafu musí být řádně popsány ve stejném jazyce, v jakém je psána práce (absenci diakritiky lze tolerovat). Kreslíme-li graf hmotnosti proti výšce, nenecháme na nich popisky *ht* a *wt*, ale osy popíšeme *Výška [cm]* a *Hmotnost [kg]*. Kreslíme-li graf funkce $h(x)$, popíšeme osy x a $h(x)$. Každá osa musí mít jasně určenou škálu.
- Chceme-li na dvourozměrném grafu vyznačit velké množství bodů, dáme pozor, aby se neslily do jednolité černé tmy. Je-li bodů mnoho, zmenšíme velikost symbolu, kterým je vykreslujeme, anebo vybereme jen malou část bodů, kterou do grafu zaneseme. Grafy, které obsahují tisíce bodů, dělají problémy hlavně v elektronických dokumentech, protože výrazně zvětšují velikost souborů.
- Budeme-li práci tisknout černobíle, vyhneme se používání barev. Čáry rozlišujeme typem (plná, tečkovaná, čerchovaná, . . .), plochy dostatečně rozdílnými intenzitami šedé nebo šrafováním. Význam jednotlivých typů čar a ploch vysvětlíme buď v textové legendě ke grafu anebo v grafické legendě, která je přímo součástí obrázku.

Pomocí příkazu `\psfrag` lze nahrazovat části *ps/eps* souborů (typicky popisky v obrázcích) libovolnou kombinací \LaTeX ových příkazů, jak ukazují následující příklady.

3.3 Softwarový kód

Softwarový kód, resp. výstupy z počítačových programů (je-li potřeba je v práci uvádět) je vhodné odlišit od ostatního textu. Jednou z možností je použití \LaTeX ového balíčku `fancyvrb` (fancy verbatim), pomocí něhož je v souboru `BcPrace.tex` nadefinováno prostředí `PCinout`. Pomocí něho lze vytvořit např. následující ukázky.

```
> mean(x)
[1] 158.90
```

```
> objekt$prumer  
[1] 158.90
```

Menší písmo:

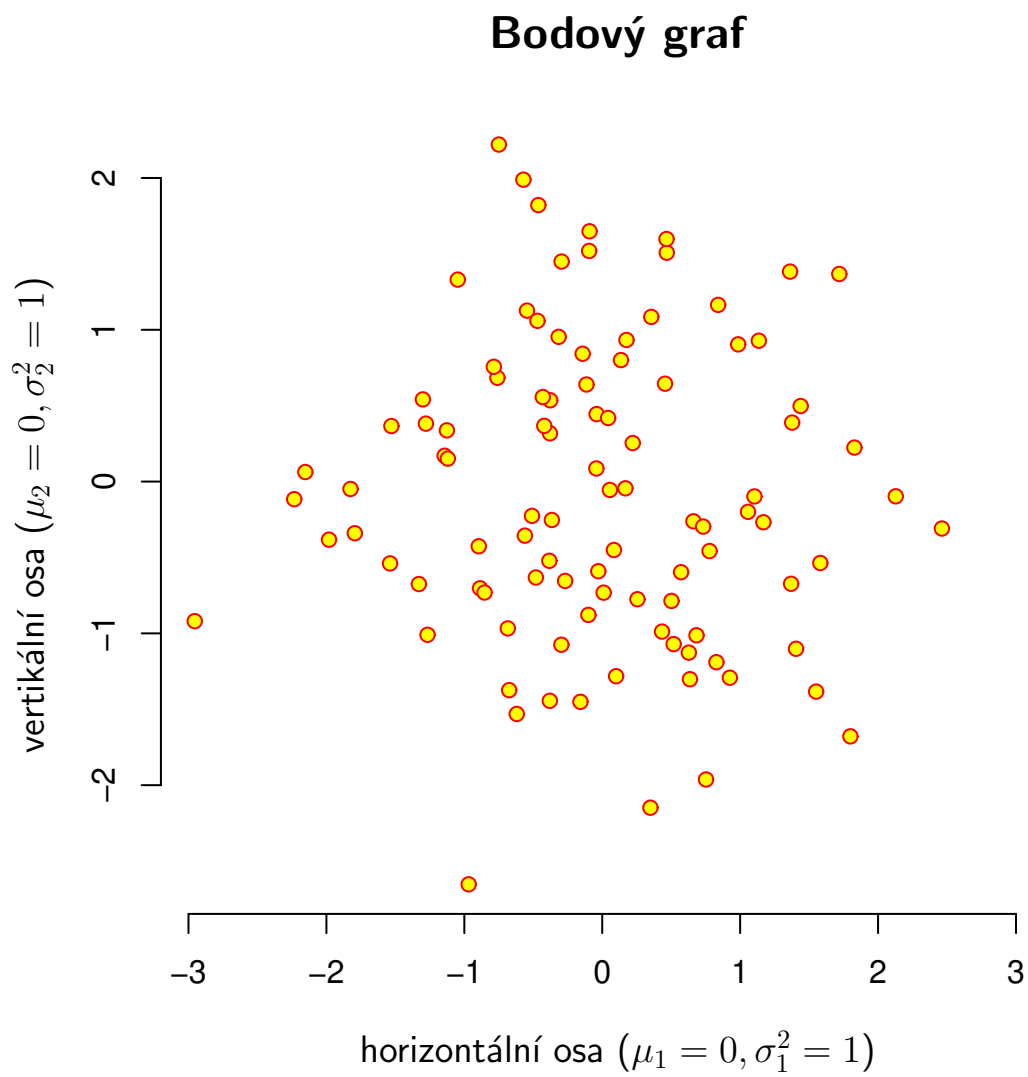
```
> mean(x)  
[1] 158.90  
> objekt$prumer  
[1] 158.90
```

Bez rámečku:

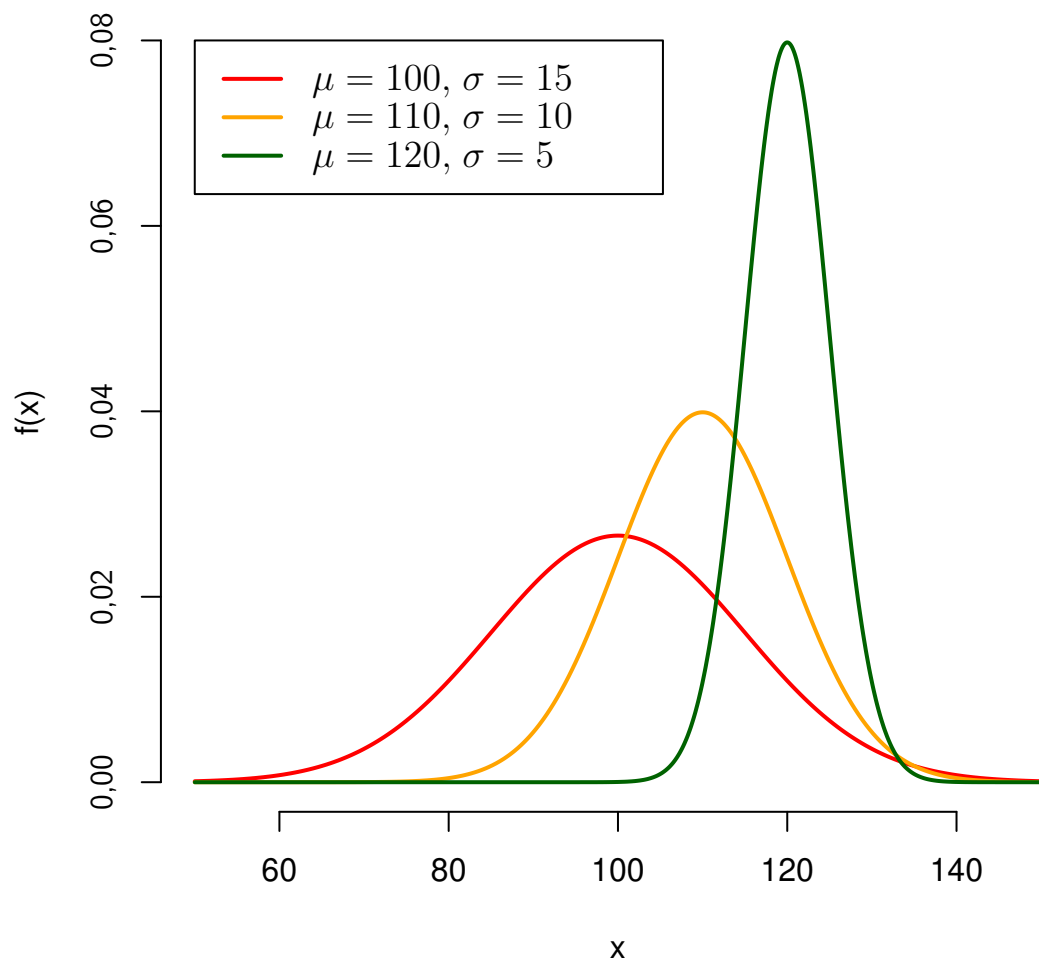
```
> mean(x)  
[1] 158.90  
> objekt$prumer  
[1] 158.90
```

Užší rámeček:

```
> mean(x)  
[1] 158.90  
> objekt$prumer  
[1] 158.90
```

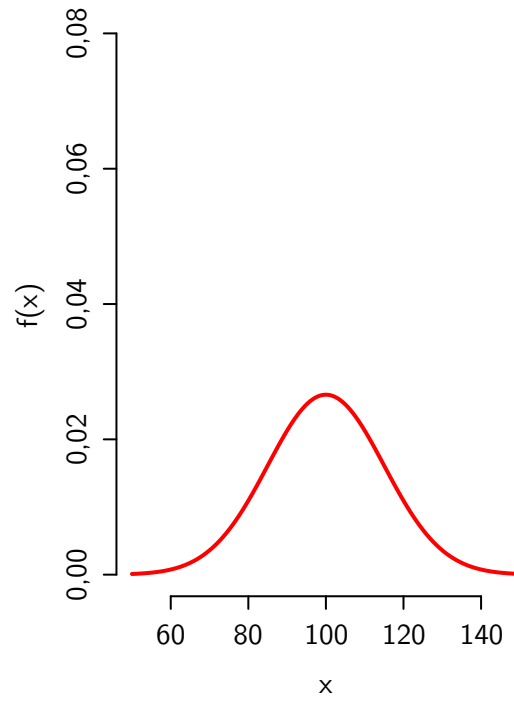


Obrázek 3.1: Náhodný výběr z rozdělení $\mathcal{N}_2(\mathbf{0}, I)$.

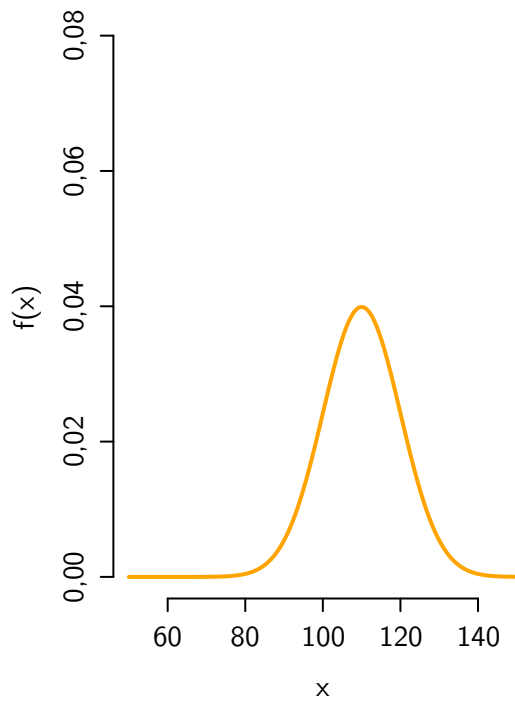


Obrázek 3.2: Hustoty několika normálních rozdělení.

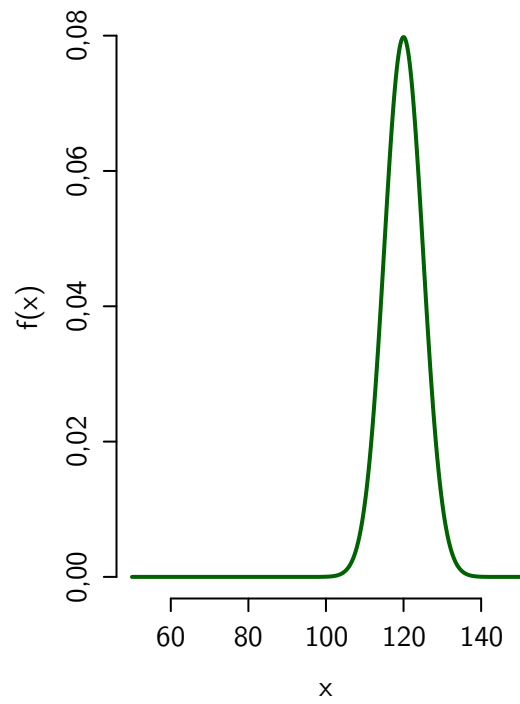
$$\mu = 100, \sigma = 15$$



$$\mu = 110, \sigma = 10$$



$$\mu = 120, \sigma = 5$$



Obrázek 3.3: Hustoty několika normálních rozdělení.

Literatura

- ANDĚL, J. (1998). *Statistické metody*. Druhé přepracované vydání. Matfyzpress, Praha. ISBN 80-85863-27-8.
- ANDĚL, J. (2007). *Základy matematické statistiky*. Druhé opravené vydání. Matfyzpress, Praha. ISBN 80-7378-001-1.
- COX, D. R. (1972). Regression models and life-tables (with Discussion). *Journal of the Royal Statistical Society, Series B*, **34**(2), 187–220.
- DEMPSTER, A. P., LAIRD, N. M. a RUBIN, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, **39**(1), 1–38.
- GENBERG, B. L., KULICH, M., KAWICHAJ, S., MODIBA, P., CHINGONO, A., KILONZO, G. P., RICHTER, L., PETTIFOR, A., SWEAT, M. a CELENTANO, D. D. (2008). HIV risk behaviors in sub-Saharan Africa and Northern Thailand: Baseline behavioral data from project Accept. *Journal of Acquired Immune Deficiency Syndrome*, **49**, 309–319.
- KAPLAN, E. L. a MEIER, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, **53**(282), 457–481.
- LEHMANN, E. L. a CASELLA, G. (1998). *Theory of Point Estimation*. Second Edition. Springer-Verlag, New York. ISBN 0-387-98502-6.
- STUDENT (1908). On the probable error of the mean. *Biometrika*, **6**, 1–25.

Seznam obrázků

3.1	Náhodný výběr z rozdělení $\mathcal{N}_2(\mathbf{0}, I)$	9
3.2	Hustoty několika normálních rozdělení.	10
3.3	Hustoty několika normálních rozdělení.	11

Seznam tabulek

3.1	Maximálně věrohodné odhady v modelu M.	6
-----	------------------------------------------------	---