

Mendelova univerzita v Brně
Provozně ekonomická fakulta

Webová aplikace pro zpracování statistických dat

Bakalářská práce

Vedoucí práce:
Ing. Pavel Haluza, Ph.D.

Josef Jelínek

Brno 2017

Velmi rád bych poděkoval panu Ing. Pavlovi Haluzovi Ph.D. za cenné připomínky a odborné rady, které mi poskytl během zpracování bakalářské práce.

Čestné prohlášení

Prohlašuji, že jsem tuto práci: **Webová aplikace pro zpracování statistických dat**

vypracoval samostatně a veškeré použité prameny a informace jsou uvedeny v seznamu použité literatury. Souhlasím, aby moje práce byla zveřejněna v souladu s § 47b zákona č. 111/1998 Sb., o vysokých školách ve znění pozdějších předpisů, a v souladu s platnou *Směrnicí o zveřejňování vysokoškolských závěrečných prací*.

Jsem si vědom, že se na moji práci vztahuje zákon č. 121/2000 Sb., autorský zákon, a že Mendelova univerzita v Brně má právo na uzavření licenční smlouvy a užití této práce jako školního díla podle § 60 odst. 1 Autorského zákona.

Dále se zavazuji, že před sepsáním licenční smlouvy o využití díla jinou osobou (subjektem) si vyžádám písemné stanovisko univerzity o tom, že předmětná licenční smlouva není v rozporu s oprávněnými zájmy univerzity, a zavazuji se uhradit případný příspěvek na úhradu nákladů spojených se vznikem díla, a to až do jejich skutečné výše.

V Brně dne 4. ledna 2017

.....

Abstract

Jelínek, J. *Web application for processing of statistical data*. Bachelor thesis. Brno, 2017

The bachelor thesis deals with statistics, statistical software and evaluation of statistical variables. The first part describes the current statistical applications. Another part deals with statistical variables and their interpretation. The third passage is focused on the technology used. The last part deals with the implementation of site through the use of technology.

Keywords:

Web application, statistics, Nette Framework, statistical software

Abstrakt

Jelínek, J. *Webová aplikace pro zpracování statistických dat*. Bakalářská práce. Brno, 2017

Bakalářská práce se zabývá problematikou statistiky, statistického softwaru a vyhodnocováním statistických veličin. První část popisuje současné statistické aplikace. Další část se zabývá statistickými veličinami a jejich interpretaci. Třetí pasáž je zaměřená na použité technologie. Poslední část se věnuje implementaci stránky pomocí použitých technologií.

Klíčová slova:

Webová aplikace, statistika, Nette Framework, statistický software

Obsah

1	Úvod a cíl práce	11
1.1	Úvod	11
1.2	Cíl práce	11
2	Porovnání současných aplikací	13
2.1	STATISTICA	13
2.2	UNISTAT	14
2.3	R (programovací jazyk)	15
2.4	Gretl	15
2.5	Statistické webové aplikace	16
3	Statistické veličiny	18
3.1	Charakteristiky polohy	18
3.1.1	Aritmetický průměr (ozn. \bar{x})	18
3.1.2	Geometrický průměr (ozn. x_g)	19
3.1.3	Kvantil (ozn. x_p), Medián (ozn. \hat{x})	19
3.1.4	Modus (ozn. \tilde{x})	19
3.2	Charakteristiky variability	20
3.2.1	Průměrná absolutní odchylka (ozn. d)	20
3.2.2	Rozptyl (ozn. s^2) nebo σ^2	20
3.2.3	Směrodatná odchylka (ozn. s nebo σ)	20
3.2.4	Variační koeficient (ozn. v)	21
3.3	Charakteristiky asymetrie	21
3.3.1	Šikmost	21
3.3.2	Špičatost	21
3.4	Typy rozdělení	22
3.4.1	Normální rozdělení	22
3.4.2	Studentovo rozdělení	23
3.4.3	Rozdělení χ^2	25
4	Použité technologie	27
4.1	Framework	27
4.2	MVC	27
4.2.1	Model	28
4.2.2	View	28
4.2.3	Controller (Presenter)	28
4.3	NetBeans	29
4.4	php5.net	29
4.5	XAMPP	29
4.6	Adminer	29
4.7	Knihovna JpGraph	30
4.7.1	Použité objekty v třídě Graph	30

4.7.2	Třída LinePlot	30
5	Implementace	31
5.1	Příprava aplikace	31
5.2	Doporučená adresářová struktura Nette Frameworku	31
5.3	Popis funkcí aplikace	32
5.3.1	Registrace	32
5.3.2	Přihlašování	32
5.3.3	Manuální vkládání hodnot	32
5.3.4	Výpis hodnot	32
5.3.5	Editace hodnot	32
5.3.6	Upload souboru	32
5.3.7	Vygenerování náhodných čísel	33
5.3.8	Výpis charakteristik polohy	33
5.3.9	Výpis charakteristik variability	33
5.3.10	Výpis charakteristik asymetrie	33
5.3.11	Shapirův-Wilkův test	33
5.3.12	χ^2 test	33
5.3.13	Statistické tabulky	34
5.3.14	Grafy rozdělení	34
6	Diskuse	36
7	Závěr	37
8	Reference	38
	Přílohy	40
A	Zdrojový kód	41

1 Úvod a cíl práce

1.1 Úvod

V dnešní době informačních technologií je informací a dat nesmírně mnoho. Proto je nutné je nějakým způsobem sumarizovat a interpretovat, aby bylo dosaženo určité výpovědní hodnoty. K tomu slouží právě vědní obor statistika.

Statistika jako věda může poskytnout informace, které nejsou v daném souboru dat na první pohled vidět. Pomocí výpočtů statistických ukazatelů lze např. optimalizovat výrobní procesy ve firmě, sledovat výkonnost lidí, ale i strojů, nebo vyhodnocovat ekonomickou aktivitu dané země. Ve statistice jsou výpočty tím přesnější, čím jsou přesnější data. To znamená, že kdyby byl počítán průměr známek žáků 6.C v Chrudimi, je lepší zahrnout všechny žáky této třídy a všechny jejich známky. Kdyby byl někdo vynechán, sníží to vypovídací hodnotu souboru dat, v tomto případě průměru známek žáků 6.C v Chrudimi.

Je také vhodné položit si otázku, co se vlastně bude počítat. S tím souvisí vhodnost použití té, či oné statistické veličiny.

Například průměrná mzda v České republice ve 2. čtvrtletí roku 2016 činila něco málo přes 27 tisíc korun, jenže toto číslo nevypovídá nic o tom, kolik lidé skutečně vydělávají. Mohla by nastat situace, že bude na jedné straně část populace, která vydělává málo a na straně druhé ta co vydělává často mnohokrát více. Potom jsou lidé překvapeni, proč nedosahují na průměrný plat. V tomto případě by bylo vhodnější použít medián. Má mnohem lepší vypovídací schopnost co se týče extrémních hodnot.

Statistika se využívá také na vyhodnocování různých průzkumů, ověřování hypotéz nebo zobecnění souboru dat.

Právě statistice a statistickým výpočtům se bude věnovat tato práce, a to vytvořením webové aplikace pro nenáročného uživatele, kteří si budou chtít vypočítat základní statistické veličiny.

1.2 Cíl práce

Cílem práce je naprogramovat webovou aplikaci, která dokáže zpracovat datový soubor a vyhodnotit statistické výpočty z daného souboru. Aplikace poskytne uživateli přehledné statistické výsledky vložených dat.

Webová aplikace by měla umět:

- vkládání hodnot, jak prostřednictvím souboru .csv nebo manuálně, popřípadě je měnit
- vyhodnocení základních statistických veličin z uložených statistických dat jako je aritmetický průměr, modus, medián, rozptyl, směrodatná odchylka, nebo šikmost či špičatost statistických dat

- výpočet kvantilů 3 nejznámějších rozdělení (Studentovo rozdělení, normální rozdělení a rozdělení chí-kvadrát) ze zadaných údajů
- zobrazení grafů funkcí jednotlivých rozdělení a jejich distribuční funkci ze zadaných parametrů
- testování hypotézy, jestli daný soubor dat má normální rozdělení, či nikoliv.

2 Porovnání současných aplikací

Současné aplikace jsou především desktopové. Důvodem je to, že obsahují velké množství funkcí, které vyžadují velký výpočetní výkon a většinou i velkou paměť. Právě proto je webových aplikací tak málo. Desktopových aplikací je celá řada, od těch základních, až po velmi robustních. V této kapitole zhodnotím některé z nich.

Statistické programové systémy jsou komplexním nástrojem pro analýzu dat. Lze uvést jejich následující přednosti:

- umožňují zpracovávat velké soubory dat,
- umožňují analyzovat datové soubory obsahující chybějící údaje,
- zahrnují větší rozsah statistických metod,
- nabízejí rozsáhlejší pomoc uživateli ze statistického hlediska (systémy radí při výběru metod, interpretují výsledky, jsou dodávány vzorové soubory statistických dat),
- některé umožňují lepší manipulaci s výstupy. (Statistické programové systémy)

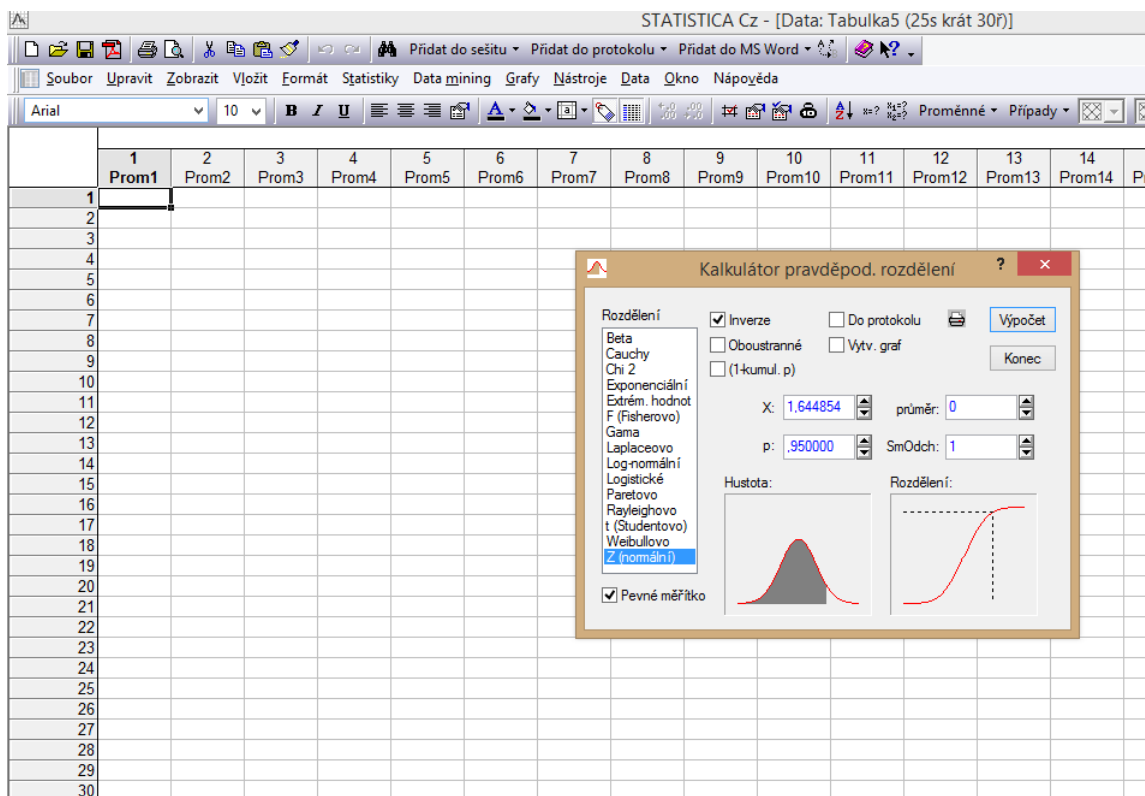
2.1 STATISTICA

STATISTICA je komplexní statistický software pro analýzu a správu dat. Je určený pro užití v oblasti bankovníctví, pojišťovnictví, financí, energetiky, marketingu a také v oblasti akademického výzkumu.

Výrobcem softwaru a také dodavatelem pro Českou republiku je společnost StatSoft. Program je chráněn licenci a v nabídce firmy není žádná studentská verze této licence. Společnost StatSoft však umožňuje stažení Trial verze programu po dobu 30ti dní zdarma. StatSoft nabízí jako zkušební vždy nejnovější verzi programu. (Statistický software STATISTICA 10, 2012)

Software STATISTICA nabízí lokalizaci češtiny avšak některé nápisy jsou v angličtině. Má příjemné uživatelské prostředí a umožňuje import velkého množství formátů, od .CSV přes Excelovské tabulky až po xml dokumenty. Program má velké množství funkcí např. popisné statistiky, korelační koeficienty, kontingenční tabulky nebo metody vícerozměrné regrese. Avšak pro laika, který chce vypočítat kolik průměrně nasbírá jablek za den, by to mohl být problém. Společnost Statsoft pravidelně pořádá kurzy ovládání softwaru poněvadž je složitý. Také tato společnost vydává návody na internetu aby byli uživatelé efektivnější při práci s tímto softwarem.

Na druhou stranu zkušený uživatel nemá problém pracovat se softwarem. Návody na ovládání jsou na vysoké úrovni a když má uživatel čas si návod přečíst, velice si usnadní práci. Program STATISTICA ze všech dále zkoumaných programů má nejlépe udělaný kalkulátor pravděpodobnostního rozdělení. Na obrázku 1 lze vidět jak početní výsledek, tak i grafické znázornění daného rozdělení. V pozadí je vidět prostředí tohoto programu.



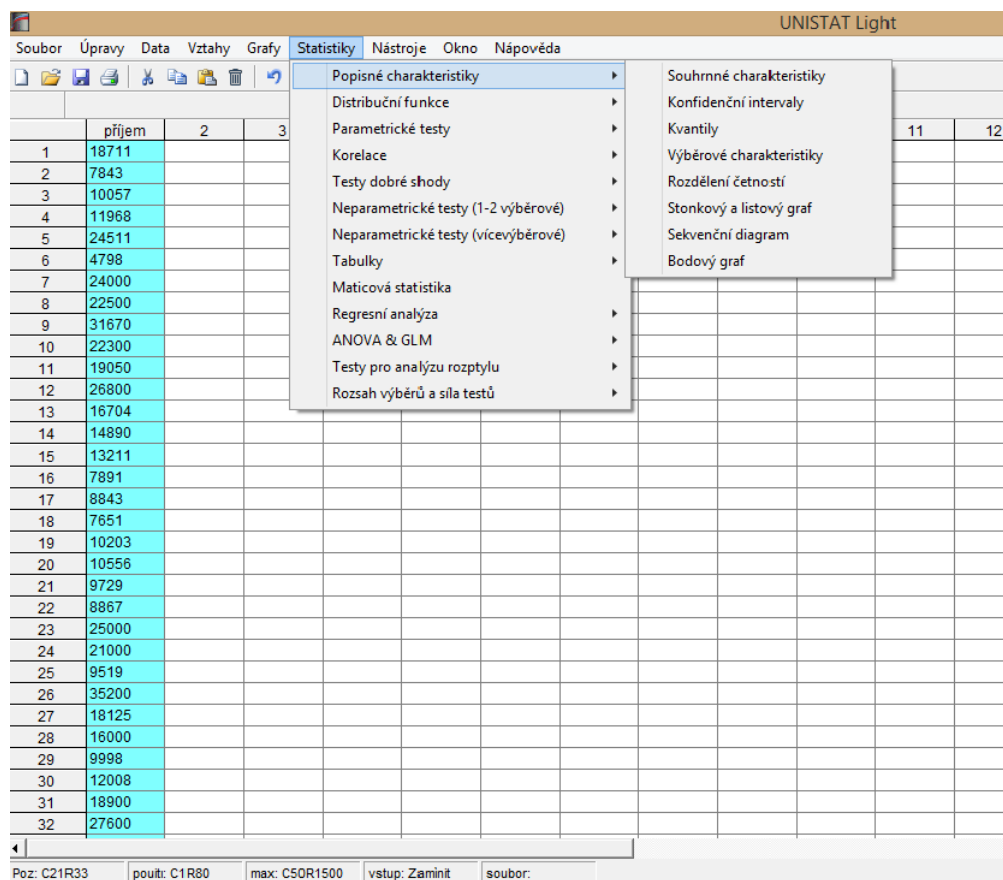
Obrázek 1: Kalkulátor v programu STATISTICA

2.2 UNISTAT

UNISTAT na svých stránkách píše toto: UNISTAT je kompletním statistickým balíkem, který nabízí všechny základní a pokročilejší statistické procedury

- Popisná statistika, kontingenční tabulky
- Testy dobré shody
- Parametrické i neparametrické testy
- Regrese a ANOVA + Testy
- Korelační koeficienty
- Rozsah výběrů a síla testu (UNISTAT)

UNISTAT není tak komplexní jak STATISTICA ovšem o to přehlednější. Na obrázku 2 vidíme, že neobsahuje tolik funkcí jako u předchozího softwaru. UNISTAT dbá na jednoduchost a dobrou orientaci v programu. Program taktéž podporuje češtinu a taktéž umožňuje import dat různých formátů. Nevýhoda demoverze spočívá v tom, že se musí pracovat pouze se vzorovými daty. Program je levnější než STATISTICA a má nižší hardwarové nároky.



Obrázek 2: Prostředí v programu UNISTAT

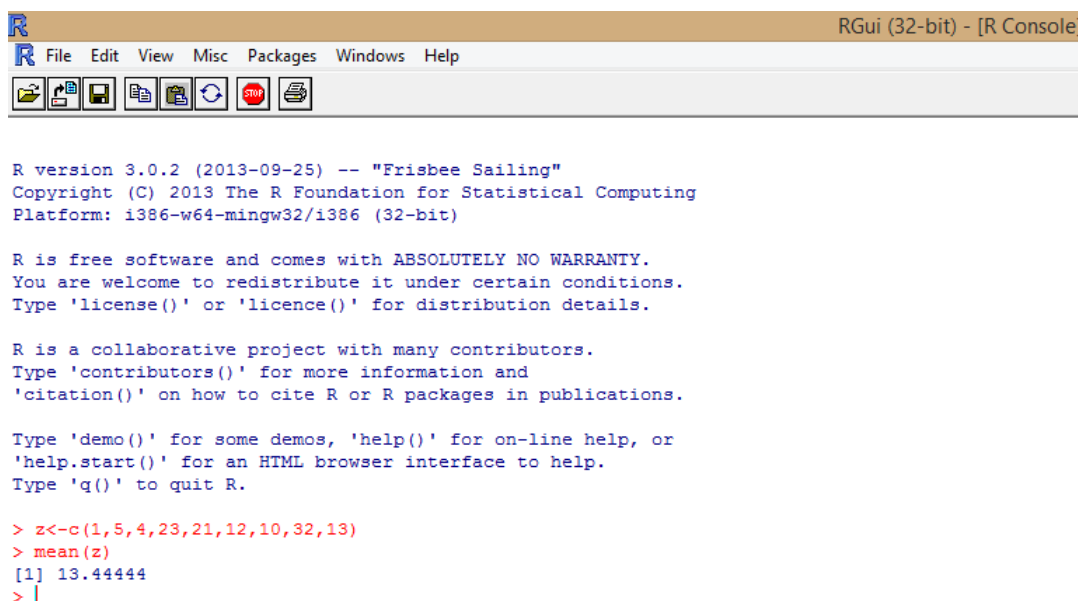
2.3 R (programovací jazyk)

R je spíše výpočetní prostředí než statistický software, a milují ho lidé, kteří musí vidět pod pokličku toho, co dělají. Ti, kdo s ním umí, říkají, že v něm lze udělat vše. Výhodou je samozřejmě, že je zadarmo, a lze proto velmi přizpůsobit vašim potřebám. Nevýhodou je pomalá „křivka učení“. V poslední době R respektují kromě vývojářů open-source řešení i komerční firmy, které nabízejí statistický software, a mnoho výrobců software se snaží najít s R společnou řeč, neboli rozhraní. (Statistický software, 2010)

Jazyk R je interpretovaný a ukrývá v sobě spoustu zajímavých funkcí. Funkce na popisnou statistiku, analýzu časových řad, nebo vykreslení grafů. Na obrázku 3 je ukázán aritmetický průměr, který se počítá funkcí `mean()`.

2.4 Gretl

Název programu Gretl je zkratkou vycházející z Gnu Regression, Econometrics and Time-series Library. Jedná se o softwarový balíček, který obsahuje užitečné a jednoduše aplikovatelné nástroje ekonometrické analýzy. Potěšující vlastností je jeho



Obrázek 3: Prostředí v programu R

volná dostupnost, díky které si program můžete stáhnout zdarma z internetové adresy gretl.sourceforge.net.

Gretl je možné rozšířit o množství vzorových datových zdrojů a databází makroekonomických časových řad. Program využívá při výpočtech plejádu účelových odhadových technik. Samozřejmostí je schopnost vykreslovat data do přehledných grafů, případně generovat textový výstup do standardních formátů (TXT, RTF), včetně populárního \LaTeX . (gretl – uživatelská příručka 2009)

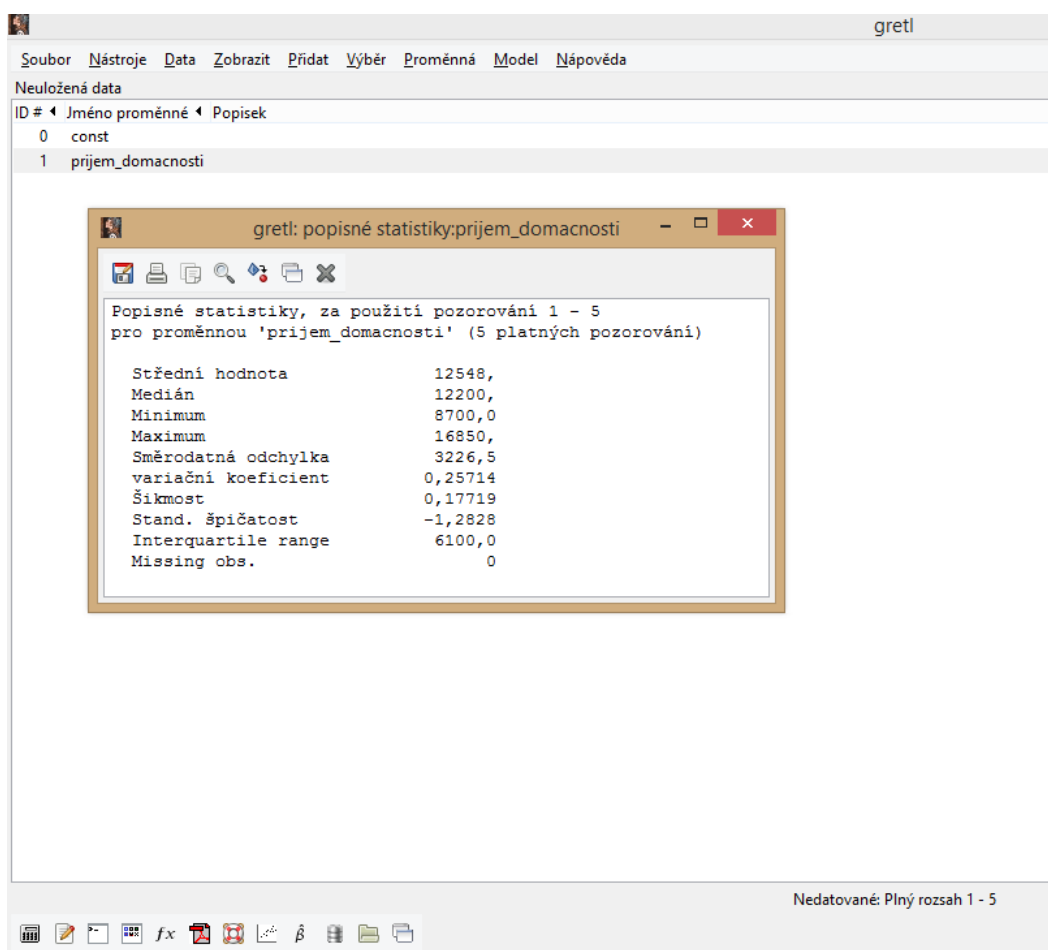
Gretl se od ostatního softwaru v lecčem liší. Hodnoty se zadávají a ukládají do proměnné, která reprezentuje statistický soubor. Pomocí přehledného menu lze počítat výsledky, vykreslovat grafy nebo řadit podle velikosti. Program, ačkoliv nemá tolik funkcí jako STATISTICA nebo UNISTAT, má intuitivní ovládání a není třeba zvláštní znalosti této aplikace. Ovšem největší výhodou tohoto programu je skutečnost, že je zdarma.

Na obrázku 4 lze vidět výčet popisné statistiky proměnné `prijem_domacnosti`. Jednoduché, přesné a rychlé.

2.5 Statistické webové aplikace

Webových aplikací na výpočet statistických dat je málo. Hezká aplikace je např. <http://www.hackmath.net/cz/kalkulacka/statistika> kde jednoduše zadáte hodnoty a stiskem tlačítka Vypočítej, se vypočítají hodnoty statistických výpočtů jako je aritmetický průměr, rozptyl, kvartily, medián, atd.

Poněkud skromnější je aplikace na adrese <http://www.kalkulacka.info/statisticke-vypocty>, kde jsou jen základní statistické veličiny.



Obrázek 4: Prostředí v programu Gretl

3 Statistické veličiny

V této kapitole budou popsány použité statistické veličiny. Tyto veličiny (charakteristiky) se dělí do 3 skupin. Charakteristiky polohy, charakteristiky variability a charakteristiky asymetrie.

- **Charakteristiky polohy** - poskytují představu o celém sledovaném statistickém souboru ve formě jedné nebo více statistických charakteristik. Mezi nejznámější charakteristiky polohy patří *průměry* (aritmetický, geometrický), *medián* a *modus*.
- **Charakteristiky variability** - popisují míru rozptýlení hodnot znaku okolo střední hodnoty. K měření variability se používá *variační rozpětí*, *průměrná odchylka*, *směrodatná odchylka*, *rozptyl* a další.
- **Charakteristiky asymetrie** - informují o odlišnosti od normálního rozdělení. K tomu slouží *šikmost* a *špičatost*.

3.1 Charakteristiky polohy

Charakteristiky polohy jsou čísla, která charakterizují úroveň hodnot znaku ve statistickém souboru.

3.1.1 Aritmetický průměr (ozn. \bar{x})

Aritmetický průměr je nejznámější a nejčastěji používaná charakteristika polohy. Podle Mgr. Veroniky Blaškové, Ph.D. (Blašková, 2012) aritmetický průměr definujeme jako součet všech hodnot znaku dělený jejich počtem, tj. rozsahem souboru. Jestliže statistický soubor má rozsah n a statistický znak X nabývá hodnoty x_1, x_2, \dots, x_n , potom aritmetický průměr je dán vztahem

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{1}{n} \sum_{i=1}^n x_i.$$

Jestliže je dána tabulka rozdělení četností, potom aritmetický průměr počítáme podle vzorce

$$\bar{x} = \frac{x_1 f_1 + x_2 f_2 + \dots + x_n f_n}{n} = \frac{1}{n} \sum_{i=1}^n x_i f_i.$$

Absolutní četnosti f_i jednotlivých hodnot znaku ve vzorci vystupují jako váhy příslušných hodnot x_i , a proto uvedený aritmetický průměr nazýváme také vážený aritmetický průměr.

Jak už bylo řečeno v úvodu, aritmetický průměr má jistou nevýhodu. Počítá totiž se všemi hodnotami, tedy i s extrémními hodnotami. Extrémní hodnoty mohou průměr vychýlit tak, že většina hodnot bude od průměru vzdálená. Tento problém řeší medián, který bude popsán níže.

3.1.2 Geometrický průměr (ozn. x_g)

Geometrický průměr lze používat pouze znaky s kladnými hodnotami. Je definován jako n -tá odmocnina součinu hodnot znaku. Je tedy definován pomocí následujícího vztahu:

$$x_g = \sqrt[n]{x_1 \cdot x_2 \cdot \dots \cdot x_n}.$$

Jestliže jsou data zapsaná v tabulce rozdělení četností, geometrický průměr počítáme podle vzorce:

$$x_g = \sqrt[n]{x_1^{f_1} \cdot x_2^{f_2} \cdot \dots \cdot x_n^{f_n}}.$$

Používá se jako charakteristika polohy znaku, jehož hodnoty narůstají geometricky. Vyjadřuje průměrnou velikost změny. Geometrický průměr je veličina, která udává v jistém smyslu typický koeficient v souboru koeficientů, nahrazuje hodnoty, co se týče jejich součinu. (Blašková, 2012)

3.1.3 Kvantil (ozn. x_p), Medián (ozn. \hat{x})

Medián označuje prostřední hodnotu statistického souboru. Je to 50% kvantil. Využívají se též 25% a 75%, souhrně se jim říká kvartily. Nejprve je nutné aby byly hodnoty uspořádány od nejmenšího po největší a hodnota, která leží uprostřed je hodnota mediánu. Pokud má soubor sudý počet hodnot, potom je medián aritmetickým průměrem dvou prostředních hodnot. Jeho označení je $x = x_{0,5}$.

Příkladem můžou být mzdy ve firmě. máme 5 zaměstnanců, kteří mají následující mzdu v českých korunách: 10000, 11500, 12000, 13000 a 35000. Když je vypočítán aritmetický průměr, vyjde číslo 16300 Kč. Je vidět, že jediná extrémní hodnota vychýlila aritmetický průměr. Drtivá většina zaměstnanců na tento aritmetický průměr nedosahuje. Proto je vhodnější použít medián.

V seřazených hodnotách se vybere prostřední hodnota, což je v tomto případě hodnota 12000 Kč. Tedy $x = 12000$. Jak je vidět, medián nezkrsluje danou skutečnost extrémními hodnotami.

3.1.4 Modus (ozn. \tilde{x})

Modus je hodnota znaku X , která má ve statistickém souboru největší četnost. Stejně jako medián i modus je charakteristika polohy, pro jejíž výpočet nepotřebujeme všechny naměřené hodnoty – na rozdíl od průměrů. Modus můžeme určit jen tehdy, když se četnosti f_i hodnot x_i znaku X odlišují. V případě kdy jsou četnosti stejné, nemůžeme modus určit.

Jestliže v souboru vystupují dvě nebo více navzájem sousedících hodnot stejně často a jejich četnost je větší jak četnost ostatních hodnot, potom modus vypočítáme jako aritmetický průměr nejfrekventovanějších hodnot.

Jestliže existují dvě navzájem nesousedící hodnoty s relativně největšími četnostmi, potom obě tyto hodnoty uvádíme jako modus. V takovém případě říkáme, že rozdělení je *bimodální* (dvojvrcholové).

V případě intervalového rozdělení četností budeme za modus považovat střed intervalu s největší četností. (Blašková, 2012)

Příkladem mohou být počty dětí v rodinách. Zkoumá se 5 rodin s následujícími počty dětí: 1, 3, 4, 3, 2. Je vidět, že nejčetnější hodnota je hodnota 3. Z těchto rodin mají 3 děti 2 rodiny. Tedy $x = 3$.

3.2 Charakteristiky variability

Statistické soubory se mohou lišit nejen v úrovni hodnot znaku (vyjádřené některou charakteristikou polohy), ale i variabilitou hodnot pozorovaného znaku. Míry variability charakterizují míru rozptýlení hodnot znaku. Některé míry variability se počítají vzhledem k vybrané charakteristice polohy. Takovými mírami jsou například průměrná odchylka, rozptyl a směrodatná odchylka. Zcela odlišně je koncipováno variační rozpětí a variační koeficient. (Blašková, 2012)

3.2.1 Průměrná absolutní odchylka (ozn. d)

Ve výpočtu průměrné absolutní odchylky její velikost závisí na každé naměřené hodnotě statistického souboru. Počítáme ji jako aritmetický průměr absolutních hodnot všech odchylek od nějaké konstanty, nejčastěji od aritmetického průměru nebo mediánu:

$$d = \frac{1}{n} \sum_{i=1}^n |x_i - a|,$$

3.2.2 Rozptyl (ozn. s^2) nebo σ^2

Rozptyl (variance, disperze) je nejčastěji používaná míra variability. Počítáme ho jako aritmetický průměr čtverců odchylek od aritmetického průměru podle vztahu:

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

Vlastnosti rozptylu:

- Rozptyl je vždy nezáporný, rozptyl z konstanty je 0
- Rozptyl se nemění s aditivní konstantou
- Rozptyl se mění se čtvercem multiplikativní konstanty (Blašková, 2012)

3.2.3 Směrodatná odchylka (ozn. s nebo σ)

Tato veličina je druhá odmocnina rozptylu: $s = \sqrt{s^2}$. Charakteristika má tu výhodu, že má stejný rozměr jako pozorovaný znak.

3.2.4 Variační koeficient (ozn. v)

Variační koeficient říká, jak moc jsou hodnoty rozptýleny kolem aritmetického průměru. Většinou se používá pro srovnání variability více souborů. Je to podíl směrodatné odchylky a aritmetického průměru:

$$v = \frac{s}{\bar{x}}.$$

Variační koeficient se většinou počítá v procentech:

$$v = \frac{s}{\bar{x}} \cdot 100.$$

3.3 Charakteristiky asymetrie

3.3.1 Šikmost

Šikmost je veličina, která udává, jak moc se liší od normálního rozdělení co se týče náklonu křivky.

Vztah mezi aritmetickým průměrem, modem, mediánem a šikmostí:

Jsou-li shodné – souměrné rozdělení

$\bar{x} \leq \tilde{x} \leq \hat{x}$ – pravotransné rozdělení = záporná asymetrie

$\hat{x} \leq \tilde{x} \leq \bar{x}$ – levostranné rozdělení = kladná asymetrie

$$k_3 = \frac{1}{n} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{s} \right)^3 \cdot f_i.$$

Je-li koeficient šikmosti větší než nula, hovoříme o levostranné šikmosti. V případě dokonalé symetrie hovoříme o nulové šikmosti.

Nulová šikmost značí, že hodnoty náhodné veličiny jsou rovnoměrně rozděleny vlevo a vpravo od střední hodnoty. Kladná šikmost značí, že vpravo od průměru se vyskytují odlehlejší hodnoty nežli vlevo (rozdělení má tzv. pravý ocas) a většina hodnot se nachází blízko vlevo od průměru. U záporné šikmosti je tomu naopak. (Blašková, 2012)

3.3.2 Špičatost

Špičatost souboru udává, jak jsou hodnoty nahuštěny okolo střední hodnoty souboru.

$$k_4 = \frac{1}{n} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{s} \right)^4 \cdot f_i - 3.$$

Je-li koeficient špičatosti roven nule, hovoříme o normální špičatosti, je-li větší než nula, hovoříme o nadnormální špičatosti.

Normální rozdělení má špičatost nula. Kladná špičatost značí, že většina hodnot náhodné veličiny leží blízko její střední hodnoty a hlavní vliv na rozptyl mají

málo pravděpodobné odlehle hodnoty. Křivka hustoty je špičatější, nežli u normálního rozdělení. Záporná špičatost značí, že rozdělení je rovnoměrnější a jeho křivka hustoty je plošší nežli u normálního rozdělení. (Blašková, 2012)

3.4 Typy rozdělení

Tato podkapitola se bude věnovat 3 nejznámějším rozdělením. Jsou to normální rozdělení, Studentovo rozdělení a rozdělení chí-kvadrát.

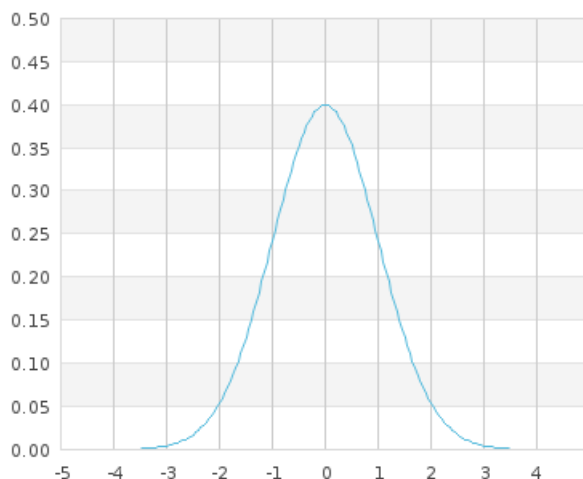
3.4.1 Normální rozdělení

Je to rozdělení, které se používá nejčastěji. Též se jí říká Gaussovo rozdělení, to proto, že vychází z Gaussovy funkce. Průběh funkce je definován podle vzorce:

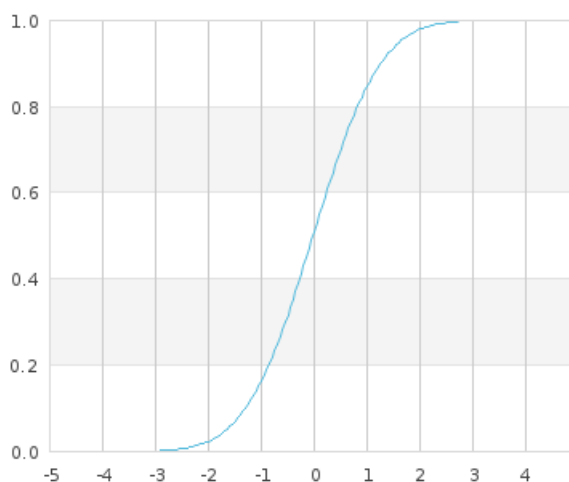
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

kde μ je střední hodnota souboru a σ je směrodatná odchylka.

Na obrázku 5 je zobrazen průběh funkce normálního rozdělení a na obrázku 6 je její distribuční funkce.



Obrázek 5: Funkce normálního rozdělení



Obrázek 6: Distribuční funkce normálního rozdělení

3.4.2 Studentovo rozdělení

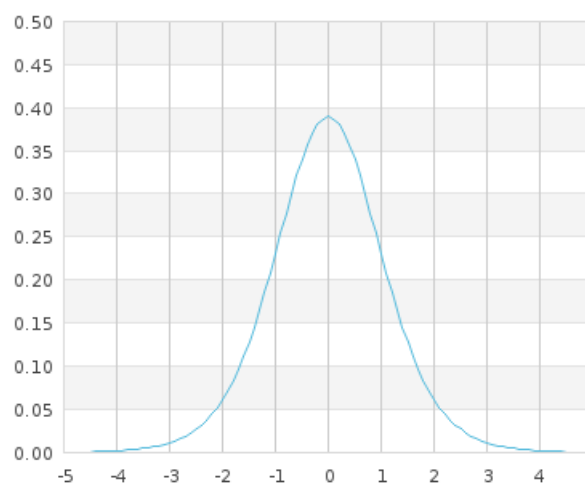
Studentovo rozdělení je podobné normálnímu rozdělení. Čím vyšší stupeň volnosti, tím více se blíží normálnímu rozdělení. Funkce Studentova rozdělení je dána:

$$f(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{\pi n}} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}$$

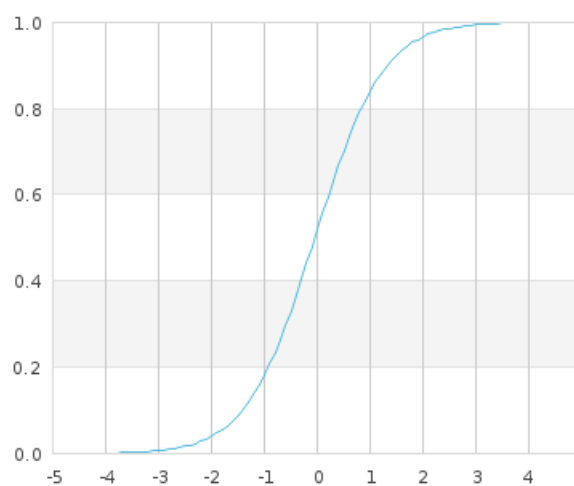
kde n je počet stupňů volnosti a

$$\Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t} dt$$

Toto rozdělení s počtem stupňů volnosti 10 je možné prohlédnout na obrázku 7 a průběh distribuční funkce na obrázku 8.



Obrázek 7: Funkce Studentova rozdělení



Obrázek 8: Distribuční funkce Studentova rozdělení

3.4.3 Rozdělení χ^2

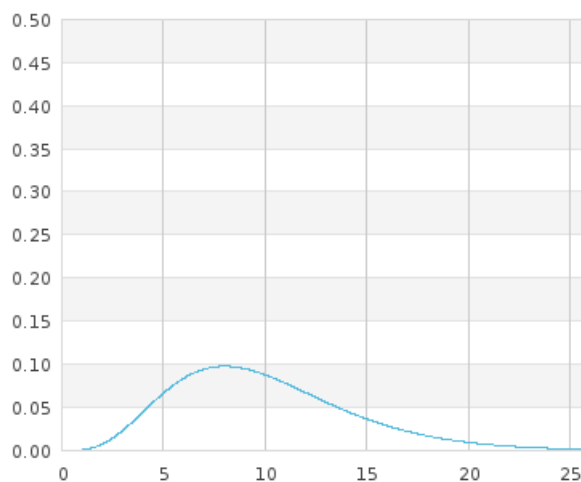
Stejně jako Studentovo rozdělení se i toto rozdělení s vyšším počtem stupňů volnosti blíží k normálnímu rozdělení. Proto je tzv. test dobré shody počítán přes toto rozdělení. Je ale nutné dbát na to, aby počet hodnot přesáhl hodnotu 30. Při nižším počtu hodnot klesá přesnost testování, zda soubor má či nemá normální rozdělení. Funkce je popsána následující rovnicí:

$$f(x) = \frac{1}{\Gamma\left(\frac{n}{2}\right) 2^{\frac{n}{2}}} e^{-\frac{x}{2}} x^{\frac{n}{2}-1}$$

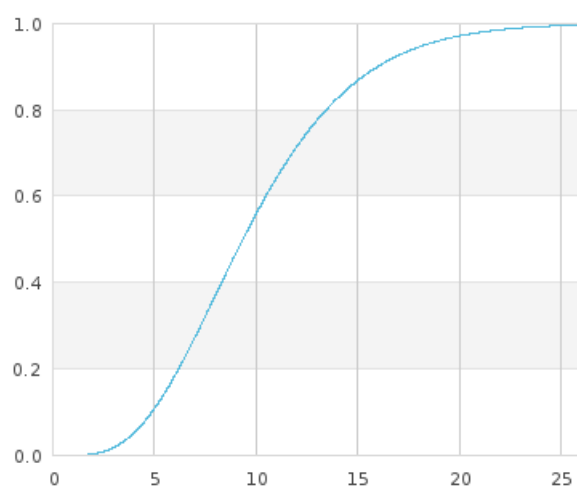
Pro $x \leq 0$ je výsledek 0.

Opět jsou přiloženy grafy funkcí při počtu stupňů volnosti 10, jak hustoty funkce obrázek 9, tak její distribuční funkce obrázek 10.

Jak je vidět na těchto obrázcích, při malém počtu stupňů volnosti je graf odlišný od grafu normálního rozdělení. Proto je vhodné testovat normalitu dat s více než 30 hodnotami.



Obrázek 9: Funkce rozdělení χ^2



Obrázek 10: Distribuční funkce rozdělení χ^2

4 Použité technologie

V této kapitole budou vysvětleny všechny použité technologie použité na tuto webovou aplikaci. Jedná se především o vývojové nástroje, programovací jazyky, knihovny nebo také webhosting.

4.1 Framework

Každý programátor ví, že obvykle jiné projekty řeší tu samou problematiku. Proto se napříč programátorským světem šíří různé knihovny, návrhové vzory nebo frameworky. Framework je obecně sada funkcí, kódů nebo tříd.

Výhoda frameworku tkví především v poskytnutí řešení drobných problémů, ať se jedná o bezpečnost aplikace nebo komunikaci s databází. Další velká výhoda je, že daný kód je většinou odladěný a funkční, takže programátor nemusí zasahovat do daného kódu. Přednost má také ve své stavbě. Framework je totiž modulární. To znamená, že je dělený na menší části, které se pak načtou při provádění kódu. Když chce programátor vytvořit formulář, stačí načíst soubor „Form“ a může pracovat s načtenou třídou.

Blog Pavla Železného popisuje framework takto: O frameworku je běžné říct, že je tvořen jen shlukem knihoven a proto nejlepší framework by jich měl obsahovat co nejvíce, ale není to pravda. Snad nejdůležitější částí frameworku by měla být vestavěná podpora pro řešení důležitých návrhových vzorů (MVC, ORM, Dependency Injection, ...), nebo alespoň snadné napojení komponent, které tuto otázku řeší. Pokud však framework neusnadňuje a neumožňuje její začlenění, nelze jej považovat za kvalitní. (Od špagety k objektům: Framework)

Jedním z mnoha PHP frameworků je Nette Framework, který byl použit na tvorbu této aplikace.

Nette Framework je napsaný v PHP 5 s plným využitím objektů (OOP). Vyrosla kolem něj jedna z nejaktivnějších komunit českých PHP vývojářů, ne-li nejaktivnější vůbec. Nette používají významné tuzemské společnosti. A podle testu uveřejněném na serveru Root je jedním z nejvýkonnějších frameworků. (Nette Framework, 2009)

Tento framework je postaven na architektuře MVC což je zkratka anglických slov model, view, controller. Nette používá místo controlleru slovo presenter, ale v podstatě v tom není rozdíl.

4.2 MVC

MVC je jeden z mnoha návrhových vzorů, který se uchytil zejména na webu, ačkoli původně vznikl na desktopech. Je součástí populárních webových frameworků, jakými jsou např. Zend nebo Nette pro PHP.

MVC architektura má za cíl oddělení logiky od výstupu. Problém nerozdělení aplikace je obrovská nepřehlednost. V třídě tak najdeme logické operace, včetně ren-

derování výstupu, databázové dotazy atd. Je velmi obtížné takový soubor udržovat natož rozšiřovat.

Proto je MVC architektura velice populární. Programátor přesně ví, kterou třídu otevřít, a kam se podívat, aby aplikaci rozšířil či opravil.

„Komponenty“ Celá aplikace je rozdělena na komponenty 3 typů, hovoříme o Modelech, View (pohledech) a Controllerech (Presenterech). Neexistuje žádná striktní definice architektury a tak je možné setkat se s více výklady.

Komponenty jsou samozřejmě třídy, oddělené z abstraktních tříd Model, View a Controller. (MVC architektura: návrhový vzor)

4.2.1 Model

Model je komponenta, která se stará o data. Nejen o to, aby data byla správná, ale také aby data byla ve správném formátu a validní. Také zajišťuje autorizaci uživatele, který se chce přihlásit nebo udělat nějaký úkon, který to vyžaduje.

Asi nejvhodnějším vodítkem pro pochopení otázky „Co do Modelu patří a co už ne?“ je právě snaha o snadnou oddělitelnost této části na dedikovaný server. Jakmile se tato část oddělí od zbytku aplikace, měla by zůstat zachována její bezpečnost, funkčnost a celková konzistence s minimem úsilí. Měla by tedy obsahovat všechny kontroly na formální i obsahovou validitu vstupních dat. A v případě, že dostane data v neočekávané podobě, měla by vrátit srozumitelnou výjimku. Výjimka by však neměla obsahovat citlivé informace jako názvy SQL tabulek, uživatelská hesla nebo úryvky zdrojového kódu. (Od špagety k objektům: MVC)

4.2.2 View

Tato část se stará o zobrazení výstupu aplikace. Blog Pavla Železného jej popisuje takto: část zvaná „View“ se stará o zobrazení výstupu uživateli. Pro větší pohodlí programátora a pro vyšší přehlednost se v této části obvykle využívá šablonovací systém, který se zároveň snaží odstínit kodéra od programovacího jazyku. Díky tomu by nemělo dojít k nechtěnému poškození „business logiky“ aplikace.

Část „View“ by měla pracovat výhradně s daty, které jsou jí předány, nebo si zažádat, aby jí byla předána data další, ale neměla by si je sama brát. Ačkoliv se to může v některých případech zdát jako přílišně dogmatické pravidlo, jeho porušení má fatální dopad na celkovou přehlednost a tedy i rozšiřitelnost a udržitelnost kódu. (Od špagety k objektům: MVC)

4.2.3 Controller (Presenter)

Controller se stará o běh aplikace. Je to takový prostředník mezi Modelem a View. Tato komponenta zpracovává uživatelské požadavky, jako kliknutí na odkaz nebo odeslání formuláře. Zpracování požadavku by mělo spočívat maximálně v sestavení příslušného objektu naplněného daty od uživatele.

Komponenta by ovšem neměla řešit vzhled, který se výhradně řeší v komponentě View. Je to z toho důvodu, aby programátor, který má za úkol vzhled aplikace nezasahoval do funkčních částí aplikace.

4.3 NetBeans

NetBeans je velmi užitečný nástroj pro programátora, který nabízí mnoho funkcí, nástrojů, pluginů apod. Oficiální stránka Netbeans (netbeans.org 2013) uvádí toto: NetBeans je úspěšný Open Source projekt s velmi rozsáhlou uživatelskou základnou, rostoucí komunitou vývojářů a téměř 100 partnery po celém světě. Firma Sun Microsystems založila Open Source projekt NetBeans v červnu 2000 a je zároveň hlavním sponzorem celého projektu.

Dnes existují dva produkty: vývojové prostředí NetBeans (NetBeans IDE) a vývojová platforma NetBeans (The NetBeans Platform).

Vývojové prostředí NetBeans IDE je nástroj, pomocí kterého programátoři mohou psát, překládat, ladit a distribuovat aplikace. Samotné vývojové prostředí je vytvářeno v jazyce Java - ovšem podporuje prakticky jakýkoliv programovací jazyk. Existuje rovněž velké množství modulů, které toto vývojové prostředí rozšiřují. Vývojové prostředí NetBeans je bezplatně šířený produkt a jeho užívání není nijak omezeno.

4.4 php5.net

Server php5.net je freehosting, který nabízí doménu 3. řádu, 50 MB úložiště, podporu PHP 5.5 a MySQL ve verzi 5.6 a připojení přes protokol FTP. Zřízení domény je velmi jednoduché a rychlé. Ačkoliv doména 2. řádu je vždy php5. Do databáze je přístup přes program phpMyAdmin, který je konkurentem programu Adminer. Aplikace je pod doménou xjeline4.php5.net/www.

4.5 XAMPP

XAMPP je podle (itbiz.cz) free a open-source multiplatformní balík, který umožňuje rychlou a snadnou instalaci Apache, MySQL, PHP, Perlu a dalších doplňků. Samotný název je akronym tvořený z: X – značí, že je multiplatformní, A – jako Apache, M – MySQL, P – PHP, P – Perl.

Tento balíček posloužil jako nástroj pro vývoj této aplikace, kde jeho úloha spočívala jako náhrada za server, aby nebylo potřeba pořád dokola nahrávat soubory na server. Tento způsob je více efektivní.

4.6 Adminer

Adminer je plně vybavený nástroj pro správu databáze napsaný v PHP. Na rozdíl od nástroje phpMyAdmin obsahuje jediný soubor, který je připraven pro nahrání

na cílový server. Adminer je k dispozici pro MySQL, PostgreSQL, SQLite, MS SQL, Oracle, SimpleDB, Elasticsearch a MongoDB. (adminer)

Mimo jiné program Adminer je součástí Nette Frameworku v doporučené struktuře.

4.7 Knihovna JpGraph

Jak píše web interval.cz (Profesionální grafy v PHP snadno a rychle) knihovna JpGraph je přídatná knihovna do PHP, která není standardní součástí distribuce PHP. Je napsaná v prostředí UNIXu, ale odzkoušená a plně funkční též pod Windows. Je plně objektově orientovaná a naprogramována v PHP. Její použití je snadné a vyžaduje pouze základní znalosti programování v PHP.

Právě díky jednoduchosti použití byla implementována do této webové aplikace. Využity byly pouze 2 soubory a to `jpgraph.php` a `jpgraph_line.php`. V následující části budou popsány jednotlivé metody tříd z knihovny Jpgraph.

4.7.1 Použité objekty v třídě Graph

Objektů je zde celá řada, použito bylo jen pár z nich. Prvním je objekt `title`. Jak je z názvu patrné jedná se o popis grafu. Dalšími jsou `xaxis` a `yaxis`. Jsou to objekty, které reprezentují osy X Y. Posledním použitým objektem je objekt `xgrid`. V grafu leckdy pomáhá v přehlednosti mřížka. Tento objekt tuto mřížku definuje. Tímto vytvoříme souřadnicový systém, neboli obrázek, ve kterém nebude vykreslen graf. Do tohoto obrázku je nutné vložit průběh funkce. K tomu slouží objekt třídy `LinePlot`.

4.7.2 Třída LinePlot

Objekt třídy `LinePlot` musí být vytvořen se dvěma parametry. Polem osy X a polem osy Y. Objekt z těchto polí vytvoří průběh funkce. Poté se vzniklý objekt vloží do souřadnicového systému popsaného výše pomocí metody `Add()` objektem třídy `Graph`.

5 Implementace

5.1 Příprava aplikace

V nainstalovaném vývojovém prostředí je možnost si vygenerovat doporučenou adresářovou strukturu. V aplikaci Adminer byly vytvořeny tabulky pro uživatele a jejich data.

5.2 Doporučená adresářová struktura Nette Frameworku

sandbox/

app/ ← adresář s aplikací

config/ ← konfigurační soubory

config.neon ← konfigurační soubor

config.local.neon

model/ ← modelová vrstva a její třídy

presenters/ ← třídy presenterů

HomepagePresenter.php ← třída presenteru Homepage

templates/ ← adresář se šablonami

@layout.latte ← šablona společného layoutu

Homepage/ ← šablony presenteru Homepage

default.latte ← šablona akce default

bootstrap.php ← zaváděcí soubor aplikace

log/ ← obsahuje logy, error logy atd.

temp/ ← pro dočasné soubory, cache, ...

vendor/ ← adresář na knihovny (např. třetích stran)

nette/ ← všechny knihovny Nette Frameworku

nette/Nette ← oblíbený framework nainstalovaný Composerem

...

autoload.php ← soubor, který načítá třídy nainstalovaných balíčků

www/ veřejný adresář, document root projektu

.htaccess ← pravidla pro mod_rewrite

index.php ← který spouští aplikaci

images/ ← další adresáře, třeba pro obrázky (MVC aplikace & Presentery)

5.3 Popis funkcí aplikace

5.3.1 Registrace

Funkce registrace je důležitá pro identifikaci uživatele a jeho dat. V této aplikaci musí uživatel zadat jméno, email a heslo aby se po registraci mohl přihlásit. Je zde ošetřeno, aby uživatel zadal platnou emailovou adresu a heslo musí mít alespoň 6 znaků. Při zaregistrování se automaticky vytvoří tabulka, do které může uživatel zadávat hodnoty.

5.3.2 Přihlašování

Aby mohl uživatel pracovat se svými daty, je potřeba se přihlásit. Přihlašuje se pomocí emailu a hesla. Heslo je v databázi osolené a šifrované pomocí metody sha1. Je zde i možnost zůstat přihlášen, která při zaškrtnutí políčka prodlouží session na 14 dní.

5.3.3 Manuální vkládání hodnot

V menu pod názvem Data je odkaz na vytvoření nového seznamu. Po kliknutí se zobrazí formulář s 10 políčky na vkládání hodnot. Je zde ošetřené, aby uživatel vkládal hodnoty a ne žádné nečíselné znaky. Také je jedno, jestli uživatel zadá desetinnou čárku či tečku. Také je nutností, aby byla zadána alespoň 1 hodnota. po odeslání formuláře je uživatel přesměrován na výpis těchto hodnot.

5.3.4 Výpis hodnot

Tato funkce se skrývá pod odkazem v menu data s názvem vypsát hodnoty. Vypíše tabulku s pořadovým číslem hodnoty a hodnotu samotnou. pod výpisem jsou odkazy na editaci, upload souboru a případné vygenerování náhodných hodnot.

5.3.5 Editace hodnot

Editovat data může uživatel pod odkazem data s názvem editovat hodnoty. Data se vypíší do tabulky, která obsahuje textfield a odkaz na smazání hodnoty. Pod tabulkou je checkbox se smazáním všech hodnot.

5.3.6 Upload souboru

Uživateli je také dovoleno uploadovat hodnoty pomocí souboru .csv. Funkce se vyskytuje pod menu data s názvem upload souboru. Tento soubor musí splňovat podmínku, že data budou jednotlivě na řádcích. Funkce zpracuje jak desetinné tečky, tak čárky.

5.3.7 Vygenerování náhodných čísel

Tato funkce se nachází pod menu data s názvem vygenerovat náhodná čísla. Po kliknutí na tento odkaz se zobrazí formulář, který se ptá na počet generovaných hodnot, na minimální a maximální hodnotu. Je zde také možnost smazat všechna předchozí data.

5.3.8 Výpis charakteristik polohy

Výsledky charakteristik polohy uživatelských dat se zobrazí po kliknutí na odkaz charakteristiky polohy v menu výpočty. Jsou zde všechny jmenované charakteristiky včetně dolního a horního kvartilu. Výpočty se provádí podle vzorců uvedených v kapitole statistické veličiny.

5.3.9 Výpis charakteristik variability

Výpis je proveden po kliknutí na odkaz charakteristiky variability v menu výpočty. Je vysvána přehledná tabulka s jednotlivými statistickými veličinami.

5.3.10 Výpis charakteristik asymetrie

Veličiny šikmost a špičatost se vypíší po kliknutí na odkaz charakteristiky asymetrie v menu výpočty. Je zde uvedeno, jak uživatelská data vypadají, a co s těchto hodnot plyne.

5.3.11 Shapirův-Wilkův test

V této funkci si uživatel může otestovat, jestli mají jeho data normální rozdělení. Testuje se pomocí Shapirovým-Wilkovým testem. Tento test je možné provést jen když je počet hodnot větší než 6. Je zde možnost zvolit hladinu významnosti. Většinou se testují hypotézy s hladinou významnosti 0,01, 0,05 nebo 0,1. Po zvolení hladiny významnosti a odkliknutí tlačítka vypočítat se nad menu zobrazí hláška, jestli se daná hypotéza zamítá, či nikoliv.

5.3.12 χ^2 test

Tento test je obdobný jako Shapirův-Wilkův test, jen se počítá jinak. Také testujeme hypotézu, o normalitě souboru. Jak bylo řečeno v kapitole statistické veličiny, tento test je vhodné používat, když je počet hodnot vyšší než 30. Opět po zadání hladiny významnosti a odkliknutí tlačítka vypočítat se nad menu zobrazí výsledek testování normality.

5.3.13 Statistické tabulky

V menu nástroje pod názvem statistické tabulky jsou funkce na zjištění kvantilů jednotlivých rozdělení. U Studentova rozdělení a rozdělení χ^2 je nutné zadat počet stupňů volnosti a pravděpodobnost alfa.

U Normálního rozdělení se zadávají střední hodnota, směrodatná odchylka a pravděpodobnost alfa.

Po odeslání formuláře se nad menu zobrazí výsledný koeficient vypočítaný pomocí vzorců zmíněných v kapitole statistické veličiny.

Níže vypsaná metoda zajišťuje výpočet Kvantilu normálního rozdělení.

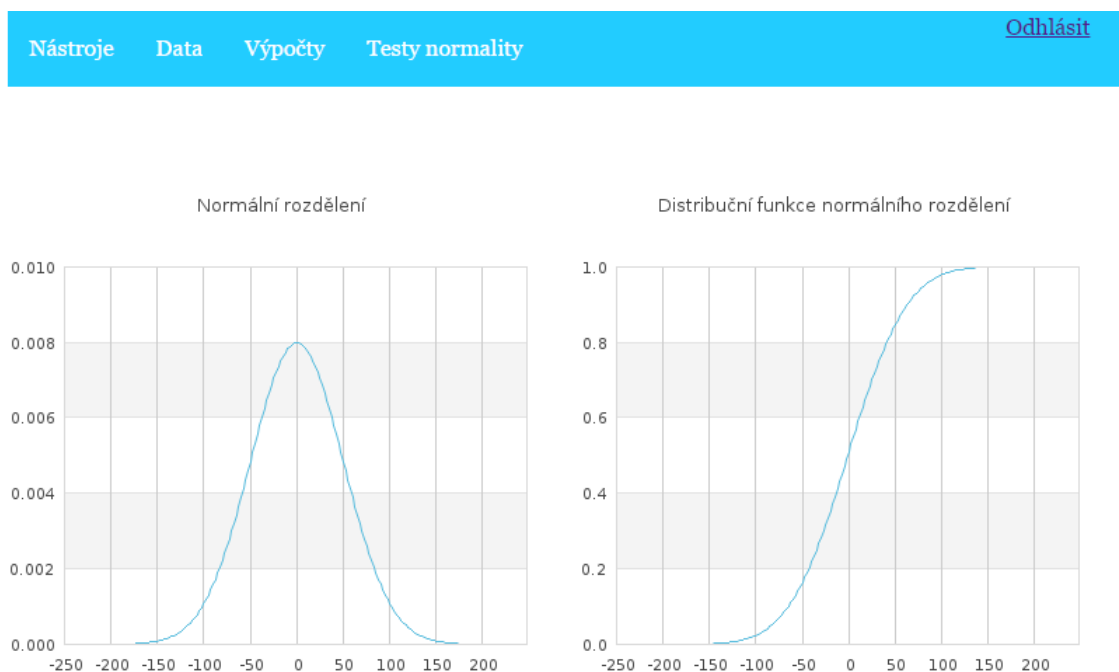
```
public function tablesNormalFormSucceeded($form)
{
    //Normalni rozdeleni
    $values = $form->getValues();
    $values->alfa=str_replace(",", ".", $values->alfa);
    if($values->alfa<=0 || $values->alfa>=1){
        $this->flashMessage('Hodnota musí být '
            . 'v~rozmezí 0 až 1', 'error');
    }
    else{
        $values->alfa=doubleval($values->alfa);
        $values->hodnota=doubleval($values->hodnota);
        $values->odchylka=doubleval($values->odchylka);
        $obsah = 0;
        for ($x = -(10*$values->odchylka)+$values->hodnota;
            $x <=(10*$values->odchylka)+$values->hodnota
                && $obsah < $values->alfa;
            $x=$x+$values->odchylka*0.0001)
        {
            $obsah = $obsah+((1/($values->odchylka*sqrt(2*pi())))
                *exp(-(pow(($x-$values->hodnota),2)/
                    (2*pow($values->odchylka,2)))))
                *$values->odchylka*0.0001);
        }
        $this->flashMessage(round($x,3), 'success');
    }
}
```

5.3.14 Grafy rozdělení

Pomocí této funkce je možné se zadáním příslušných parametrů zobrazit průběh funkcí daných rozdělení. Je zde opět formulář pro zadání nezbytných údajů. Stu-

dentovo rozdělení a rozdělení χ^2 se spokojí s parametrem počet stupňů volnosti a u normálního rozdělení je nutné zadat střední hodnotu a směrodatnou odchylku.

Po vyplnění jednoho z formulářů a odeslání se zobrazí 2 grafy. První graf je průběh funkce zvoleného rozdělení a druhý graf je jeho distribuční funkce. Výsledek je vidět na obrázku 11.



Obrázek 11: Výsledná stránka zobrazující průběh normálního rozdělení a její distribuční funkci pro střední hodnotu 0 a směrodatnou odchylku 50.

6 Diskuse

Na základě stanovených požadavků byla vytvořena webová aplikace, která je určena pro výpočty statistických ukazatelů z uživatelových dat. Aplikace je dostupná na adrese <http://xjeline4.php5.cz/www/>.

Všechny funkce, které byly naimplementovány fungují tak, jak by měly. Uživatel má funkce přehledně rozděleny v menu aplikace. Další výhodou je, že formuláře jsou ošetřeny tak, aby uživatel nezasílal chybné údaje, či hodnoty.

Nevýhodou aplikace je malé množství funkcí. Statistika je mnohem rozmanitější. Přes testování hypotézy, zda mají 2 soubory stejný rozptyl (když není znám), až po korelační a regresní analýzu. Další nevýhodou je, že uživatel může nahrát či vytvořit pouze 1 statistický soubor.

Možnosti zlepšení aplikace jsou následující: naimplementování dalších výpočtů statistických veličin a možnost vkládání či vytváření více statistických souborů dat.

7 Závěr

Práce se zabývala návrhem a implementací webové aplikace pro zpracování statistických dat. Webová aplikace obsahuje funkce pro vytvoření či editaci dat a také funkce na výpočty statistických veličin. Také jsou v aplikaci výpočty kvantilů nejznámějších rozdělení a zobrazení grafů funkcí těchto rozdělení.

V první části práce je popsáno, proč je statistika důležitá věda a je zde uveden cíl práce.

Další část se zabývá konkurenčními statistickými aplikacemi, a to jak desktopovými tak webovými. Jsou zde uvedeny výhody a nevýhody každé z nich.

Následující část se věnuje použitým statistickým veličinám a jejich vzorcům.

Čtvrtá část obsahuje výčet a popis všech použitých technologií na tvorbu této webové aplikace.

V další části je podrobný popis naimplementovaných funkcí a způsob obsluhy funkcí.

V diskusi jsou popsány výhody a nevýhody aplikace, popřípadě návrhy na zlepšení.

8 Reference

- Adminer* [online]. [cit. 2016-12-05]. Dostupné z: <http://www.adminer.org/>.
- BLAŠKOVÁ, V. a kol. *Statistika I*. Vyd. 2., přeprac. Brno: Mendelova univerzita v Brně, 2012, 204 s. ISBN 978-80-7375-648-2..
- Co je NetBeans?* [online]. [cit. 2016-12-05]. Dostupné z: https://netbeans.org/index_cs.html.
- gretl – uživatelská příručka* [online]. [cit. 2016-12-05]. Dostupné z: http://www.thunova.cz/wp-content/uploads/CZU/Manual_gretl.pdf.
- MVC aplikace & Presentery* [online]. [cit. 2016-12-05]. Dostupné z: <http://doc.nette.org/cs/2.1/presenters>.
- MVC architektura: návrhový vzor* [online]. [cit. 2016-12-05]. Dostupné z: <http://www.itnetwork.cz/mvc-architektura-navrhovy-vzor>.
- Nette Framework* [online]. [cit. 2016-12-05]. Dostupné z: <http://www.zdrojak.cz/clanky/nette-framework-zvyste-svoji-produktivitu/>.
- Od špagety k objektům: Framework* [online]. [cit. 2016-12-05]. Dostupné z: [http://blog.pavelzelezny.cz/od-spagety-k-objektum/\(4\)-framework.html](http://blog.pavelzelezny.cz/od-spagety-k-objektum/(4)-framework.html).
- Od špagety k objektům: MVC* [online]. [cit. 2016-12-05]. Dostupné z: [http://blog.pavelzelezny.cz/od-spagety-k-objektum/\(3\)-mvc.html](http://blog.pavelzelezny.cz/od-spagety-k-objektum/(3)-mvc.html).
- Profesionální grafy v PHP snadno a rychle* [online]. [cit. 2016-12-05]. Dostupné z: <https://www.interval.cz/clanky/professionalni-grafy-v-php-snadno-a-rychle/>.
- Statistické programové systémy* [online]. [cit. 2016-12-05]. Dostupné z: <http://badame.vse.cz/clanky/statsoft.php>.
- Statistický software* [online]. [cit. 2016-12-05]. Dostupné z: <http://www.datamind.cz/cz/blog/Statisticky-software-jak-na-vyber-statistickeho-software>.
- Statistický software STATISTICA 10* [online]. [cit. 2016-12-05]. Dostupné z: http://samba.fsv.cuni.cz/~soukp6as/OLD/AKD/old/zaverecneprace/Tumova_Software%20Statistica.doc.
- UNISTAT* [online]. [cit. 2016-12-05]. Dostupné z: <http://www.unistat.cz/>.
- XAMPP* [online]. [cit. 2016-12-05]. Dostupné z: <http://www.itbiz.cz/slovník/informacni-technologie-it/xampp>.

Seznam obrázků

Obrázek 1: Kalkulátor v programu STATISTICA	14
Obrázek 2: Prostředí v programu UNISTAT	15
Obrázek 3: Prostředí v programu R	16
Obrázek 4: Prostředí v programu Gretl	17
Obrázek 5: Funkce normálního rozdělení	22
Obrázek 6: Distribuční funkce normálního rozdělení	23
Obrázek 7: Funkce Studentova rozdělení	24
Obrázek 8: Distribuční funkce Studentova rozdělení	24
Obrázek 9: Funkce rozdělení χ^2	25
Obrázek 10: Distribuční funkce rozdělení χ^2	26
Obrázek 11: Výsledná stránka zobrazující průběh normálního rozdělení a její distribuční funkci pro střední hodnotu 0 a směrodatnou odchylku 50.	35

Přílohy

A Zdrojový kód

Zdrojové kódy aplikace jsou k dispozici na přiloženém CD.