

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY



ŠTRUKTÚRA SLOV V PÍSANOM TEXTE

DIPLOMOVÁ PRÁCA

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

ŠTRUKTÚRA SLOV V PÍSANOM TEXTE

DIPLOMOVÁ PRÁCA

Študijný program: Ekonomicko-finančná matematika a modelovanie
Študijný odbor: 1114 Aplikovaná matematika
Školiace pracovisko: Katedra aplikovanej matematiky a štatistiky
Vedúci práce: doc. Mgr. Richard Kollár, PhD.



Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Bc. Anna Košútová
Študijný program: ekonomicko-finančná matematika a modelovanie
(Jednoodborové štúdium, magisterský II. st., denná forma)
Študijný odbor: aplikovaná matematika
Typ záverečnej práce: diplomová
Jazyk záverečnej práce: slovenský
Sekundárny jazyk: anglický

Názov: Štruktúra slov v písanom texte
The word structure in written text

Anotácia: V bakalárskej práci A. Košútovej v roku 2018 bola pozorovaná dvojmodálna distribúcia dĺžky slov v literárnych textoch v slovenskom jazyku a taktiež v ďalších jazykoch. Výsledky sú dôležitými novými poznatkami v kvantitatívnej lingvistike a diplomová práca by ich mala prehĺbiť.

Vedúci: doc. Mgr. Richard Kollár, PhD.
Katedra: FMFI.KAMŠ - Katedra aplikovanej matematiky a štatistiky
Vedúci katedry: prof. RNDr. Daniel Ševčovič, DrSc.
Dátum zadania: 08.01.2019

Dátum schválenia: 08.01.2019
prof. RNDr. Daniel Ševčovič, DrSc.
garant študijného programu

študent

vedúci práce

Podakovanie Moja vďaka patrí predovšetkým školiteľovi doc. Mgr. Richardovi Kolárovi, PhD. za cenné a odborné rady, venovaný čas, prejavenu ochotu, motiváciu, podporu a pomoc, ktoré mi veľmi pomohli pri tvorbe tejto diplomovej práce. Veľké ďakujem patrí aj môjmu priateľovi za trpezlivosť, podporu a obetavosť. V neposlednom rade sa chcem poďakovať aj mojej rodine za podporu a psychickú oporu.

Abstrakt

KOŠÚTOVÁ, Anna: Štruktúra slov v písanom texte [Diplomová práca], Univerzita Komenského v Bratislave, Fakulta matematiky, fyziky a informatiky, Katedra aplikovanej matematiky a štatistiky, škooliteľ: doc. Mgr. Richard Kollár, PhD., Bratislava, 2020, 103 s.

V tejto diplomovej práci sa zaoberáme analýzou štruktúry slov v písanom texte. Popisujeme teóriu dĺžky slov, princípy a metódy spracovania a transformácie textov na dáta. Štruktúru slov analyzujeme na základe významnosti a obsahu slov. Využívame na to model aproximujúci rozdelenia dĺžok slov použitý v práci [11], ale aj jeho odvodenú diskretnú verziu. Taktiež sa venujeme správnosti modelov a bimodalite rozdelenia dĺžok slov v písaných textoch.

Kľúčové slová: dĺžka slova, modelovanie, diskretné gama rozdelenie, štruktúra slov, bimodalita

Abstract

KOŠÚTOVÁ, Anna: The word structure in written text [Diploma Thesis], Comenius University in Bratislava, Faculty of Mathematics, Physics and Informatics, Department of Applied Mathematics and Statistics, Supervisor: doc. Mgr. Richard Kollár, PhD., Bratislava, 2020, 103 p.

This diploma thesis analyses the subject of word structure in written text. We describe some word length theories, principles and methods of processing as well as transforming texts into data. Our analysis is based on word meaning and the contents of word. For this analysis we use models of word length distribution from [11] and also we use its discrete versions. Also we deal with the correctness of this models and bimodality of word length distribution in written texts.

Keywords: word length, modelling, discrete gamma distribution, word structure, bimodality

Obsah

Úvod	9
1 Teória dĺžky slova	10
1.1 História	10
1.2 Princípy na získavanie dát a ich úprava	11
1.3 Konštrukcia modelov	13
1.4 Zadefinovanie princípov pre ďalšie kapitoly	16
2 Texty a ich spracovanie do dát	17
2.1 Zdroje, rozsah a typy textov	17
2.2 Úprava textov a transformácia na dáta	18
3 Modelovanie dĺžky slov	20
3.1 Model využívajúci zmes gama rozdelení	20
3.2 Diskretizácia modelu	21
3.3 Vhodnosť modelov	25
4 Analýza štruktúr v písanom texte	26
4.1 Plnovýznamové a neplnovýznamové slová	26
4.2 Obsahové a neobsahové slová	29
5 Bimodalita rozdelenia dĺžok slov v textoch	35
5.1 Štandardné výpočty miery bimodality	35
5.2 Aplikácia na dáta	36
Záver	40
Literatúra	42

Úvod

V každodennom živote prichádzame do styku s jazykom, nielen v ústnej podobe. Písomná podoba jazyka je špecifická v tom, že sa riadi striktnejšími pravidlami a je konzistentnejšia, zatiaľ čo hovorený jazyk zaobchádza s gramatikou či syntaxou veľmi autonómne. Písané texty majú oveľa väčšiu výpovednú hodnotu ako hovorené slovo. Väčšina z nás berie písané texty ako prirodzenú súčasť svojho života a nezamýšľame sa nad tým, ako zložito sú písané, pretože stačí, že im rozumieme.

No dokážeme si na pozadí písaných textov predstaviť nebadateľný mechanizmus, ktorý vie štruktúru slov jednoducho popísať a interpretovať? Cieľom tejto diplomovej práce je nájsť taký mechanizmus, ktorý zároveň bude nadväzovať na prácu [11].

Hlavným ťažiskom našej práce je analýza štruktúry slov v textoch. Na konci práce [11] boli stručne naznačené tri hypotézy vychádzajúce zo zostaveného modelu, ktorých obsahom bolo, že písaný text môže byť rozdelený na dva samostatné celky.

Prvá hypotéza sa venovala plnovýznamovým a neplnovýznamovým slovám, druhá slovám po obsahovej stránke a tretia štýlu autora a hlavnej informácie textu. Obsahové a neobsahové slová úzko súvisia s hypotézou o štýle autora a hlavnej informácie textu, pretože slovník, štýl a dĺžky viet sa odrazia vo frekvencii obsahových a neobsahových slov autorovho textu. Preto skúmame hypotézy o významnosti a obsahovosti slov v písanom jazyku.

V prvej časti práce si približujeme teóriu dĺžky slova, pozeráme sa na jej historické pozadie a modelovanie a stanovujeme si princípy, ktorých sa budeme pridŕžiavať počas celej práce. Následne si približujeme získavanie textov do databázy a ich transformáciu na dáta. Tieto dáta tvoríme ako počty písmen slov jednotlivých textov, ktorých rozdelenie modelujeme pomocou dvoch modelov.

Prvý model, ktorý používame bol vytvorený v práci [11] ako zmes dvoch spojitých gama rozdelení. Nakoľko dĺžky slov spracované do dát sú diskrétné navrhujeme model využívajúci diskrétna rozdelenia.

Pretože dĺžky slov v písaných textoch vykazujú bimodálne rozdelenie, v závere práce sa pozeráme na štandardné miery testovania bimodality, ktoré aplikujeme na naše dáta.

1 Teória dĺžky slova

Dĺžka slova je asi najprirodzenejšou fyzickou vlastnosťou slova. Je to vlastnosť, ktorá sleduje nejaký distribučný zákon a tak sa považuje za náhodnú premennú. Dĺžkou slov sa zaoberá mnoho štúdií a vo väčšine prípadov ju merajú v počte slabík. Dĺžka slova sa však dá merať aj v počte písmen, morféme, dokonca aj v počte mór či taktov, a všetky tieto dĺžky sú vzájomne korelované [15, str.88]. V našej diplomovej práci sme dĺžku slova nemerali zaužívaným spôsobom a to v počte slabík, ale v počte písmen.

V tejto kapitole si priblížime historické korene skúmania dĺžky slova, odporúčané princípy na získavanie a úpravu dát pre jeho rozdelenie a následné modelovanie. V poslednej podkapitole si vymedzíme princípy, ktorých sa budeme pridržovať pri ďalšej práci.

Budeme vychádzať z [2], [7], [11] a [15].

1.1 História

Štúdium dĺžky slova podľa [7] môžeme datovať do roku 1851, kedy známy anglický matematik a logik Augustus de Morgan, v liste svojmu priateľovi, popísal myšlienku, že dĺžka slova môže byť dobrým indikátorom individuálneho štýlu a môže byť možným faktorom pri určovaní autorstva. De Morgan sa v liste zameral na počet písmen v jednotlivých slovách v Epištolách Svätého Pavla a prosil priateľa o zamyslenie, či priemerná dĺžka slov v rôznych Epištolách Svätého Pavla môže objasniť otázku ich autorstva. Ak by sme zovšeobecniť jeho myšlienky, tak de Morgan predpokladal, že priemerné dĺžky slov v dvoch textoch od toho istého autora, ktorých obsah je úplne odlišný, by sa mali na seba podobáť viac ako v dvoch textoch s rovnakou témou avšak napísané dvoma rôznymi autormi.

Prvý empirický dôkaz v prospech de Morganových predpokladov bol vykonaný o pár desaťročí neskôr americkým fyzikom a meteorológom Thomasom Corwinom Mendenhallom. Vo svojich štúdiách sa Mendenhall nezameral len na de Morganove myšlienky, ktoré boli založené na základe priemernej dĺžky slova. Využíval aj grafické znázornenia, v ktorých zobrazoval usporiadanie slov podľa ich dĺžky v počte písmen a relatívne frekvencie ich výskytu. Takto neanalyzoval celé texty, ale len ukážky po 1000 slov

od rôznych autorov, medzi ktorými boli aj Francis Bacon, William Shakespeare, John Stuart Mill či Charles Dickens.

Mendenhall spomenul aj možnosti ďalších aplikácií, ktoré môžu nadväzovať na jeho prístup, ako napríklad meranie dĺžok slov v slabikách alebo počtov slov vo vetách. Napriek tomu sa Mendenhall vo prácach sústredil len na dĺžku slova v počte písmen a do svojho výskumu zahrňal aj cudzojazyčné texty.

Mendenhallova práca sa stretla v tej dobe s nepochopením, no aj tak znamenala v štúdiu dĺžok slov enormný krok vpred a položila dobrý základ pre budúci výskum, ktorý sa systematicky začal rozvíjať až v 50. rokoch minulého storočia, kedy začali vznikať prvé modely slovnej dĺžky, následne ich zovšeobecnenia, metódy na budovanie modelov a pod.

1.2 Princípy na získavanie dát a ich úprava

Pri kvantitatívnej analýze textov sa odporúča najprv formulovať hypotézu a až po jej formulácii urobiť výber dát [15, str. 16]. To znamená, že dáta nie sú presne dané, ale si ich musíme skonštruovať tak, aby boli relevantné pre testovanie týchto hypotéz. Treba si tiež dávať pozor a nemiešať dáta s pozorovaniami. Aj pozorovania sa najprv zostrojujú, no dáta sú oveľa abstraktnejšie. Pri analýze textu za pozorovania môžeme považovať hlásky, slová, vety . . . , ktoré potom premeníme na dáta.

Prvý spôsob transformácie pozorovaní na dáta je postavený na základe formálnych vlastností textu, ktorými sú napríklad dĺžky, výskyty, kombinácie jednotiek a iné. Tieto jednotky sú stabilnejšie a každému poskytujú rovnaké výsledky.

Druhý spôsob transformácie je založený na dojmach o povahe textu, napríklad či je text smutného alebo veselého charakteru, či je písaný jednoducho alebo zložito a pod. Takéto kvalitatívne vlastnosti textu protikladného charakteru sú veľmi nepresné a na získanie dát potrebujeme veľa subjektívnych dojmov od jednotlivcov, ktoré škálovaním premeníme na kvantitatívne dáta ako napríklad priemery, stupne a pod., poprípade sa pokúsime nájsť také entity, ktoré vyvolávajú daný dojem a vyčíslime ich výskyt. Čiže aj subjektívne dojmy vieme premeniť na užitočné dáta [15, str. 20].

Pri získavaní dát pre rozdelenia dĺžok slov sa podľa [15, str. 89] odporúča pridržovať nasledovných princípov:

1. Text by sa mal vyhodnotiť ako celok. Neodporúča sa vyhodnocovať len časti textu alebo robiť náhodné výbery slov.
2. Ak text obsahuje kompaktné časti, napr. kapitoly, v takom prípade sa rozdeliť môže, pretože sa dá predpokladať, že text nebol napísaný tzv. jedným ťahom, ale s prestávkami, ktoré mohli viesť k zmene rytmu.
3. Texty by sa nemali miešať a nemali by sa hľadať rovnaké rozdelenia u dvoch rôznych autorov, ani v skupinách textov podľa žánru. Pri zmiešavaní príbuzných textov totiž môže dôjsť k zaniknutiu vhodnosti modelu.
4. Ak je text príliš dlhý bez akéhokoľvek naznačenia prirodzeného členenia textu, odporúča sa urobiť test vhodnosti modelu podľa nasledovného kritéria

$$C = \frac{1}{N} \cdot \chi^2, \quad (1)$$

kde N predstavuje rozsah výberu a χ^2 je testovacia štatistika z Pearsonovho chí-kvadrát testu dobrej zhody.

Testovacia štatistika χ^2 sa počíta podľa vzorca

$$\chi^2 = \sum_{i=1}^k \frac{(f_i - NP_i)^2}{E_i}. \quad (2)$$

Vo vzorci (2) k predstavuje celkový počet tried, f_i je experimentálna frekvencia i -tej triedy a $E_i = NP_i$ je teoretická frekvencia i -tej triedy.

Ak vypočítaný koeficient $C < 0.02$ model možno akceptovať a ak $C < 0.01$ tak ide o veľmi dobrý fit [7, str. 23].

Takto počítaný koeficient s využitím testovacej štatistiky χ^2 má výhodu v tom, že nezávisí od stupňov voľnosti chí-kvadrát rozdelenia.

Ešte pred modelovaním rozdelenia dĺžok slov je potrebné sa pozrieť na text a upraviť nasledujúce súčasti textov [2, str. 122]:

1. Skratky je potrebné rozpísať do tvaru podľa príslušných gramatických a syntaktických pravidiel. Napríklad skratka *atď.* sa rozpíše na *a tak ďalej*, čím vzniknú tri samostatne stojace slová. Skratka *UK* sa nahradí slovným spojením *Univerzita Komenského* alebo jeho príslušným vyskloňovaným tvarom podľa kontextu.

2. Čísla, číslice a dátumy písané arabskými alebo rímskymi číslicami je potrebné taktiež rozpísať do príslušnej gramatickej formy.
3. Cudzojazyčné slová by sa mali nahradiť podľa výslovnosti v jazyku, v ktorom je text primárne písaný.
4. Slovné spojenia písané cez spojovník je potrebné brať ako dve samostatne stojace slová. Napríklad *česko-slovenský* sa rozloží na *česko slovenský*.

1.3 Konštrukcia modelov

V tejto podkapitole sme vychádzali z [15, str. 89], ktorá sa zameriava na modelovanie rozdelenia dĺžok slov meraných v počte slabík.

Pri každom jednom texte môžeme predpokladať, že autor sa pri jeho tvorení držal istého rytmu, ktorý sa premenil do dĺžky vety, dĺžky slova a pozície prízvukov. Tento rytmus môže do veľkej miery súvisieť s obmedzeniami a predpismi žánru, momentálnym fyzickým stavom autora, ale aj s príslušnosťou autora k nejakej škole. Toto všetko sú faktory, ktoré podliehajú istým zákonitostiam, čo pri modelovaní dĺžky slov musí viesť k variabilite modelov.

Existencia rytmu na základe slovnej dĺžky nás môže v najjednoduchšom prípade priviesť k predpokladu, že pravdepodobnosť slov dĺžky x je proporcionálna pravdepodobnosti dĺžky $x - 1$, t.j.

$$P_x \sim P_{x-1}. \quad (3)$$

Ďalej sa predpokladá, že proporcionalita nie je konštantná a môže byť závislá od dosiahnutej dĺžky a aj iných faktorov. Môžeme ju označiť ako $f(x)$ a dosadiť do (3), čím dostaneme

$$P_x = f(x) \cdot P_{x-1}. \quad (4)$$

Funkciou $f(x)$ sa dajú zachytiť všetky faktory, ktoré pôsobia na rozdelenie dĺžok. Tieto faktory môžeme rozdeliť na dve skupiny a to na kreatívne a konzervatívne. Ich vzájomné pôsobenie sa dá vyjadriť pomocou racionálnej lomenej funkcie, pričom prvá skupina je v čitateli a druhá v menovateli:

$$f(x) = \frac{g(x)}{h(x)}.$$

Vo všeobecnosti môžeme predpokladať, že $g(x)$ obsahuje jazykové konštanty a všetky aktuálne faktory textu, zatiaľ čo $h(x)$ obsahuje vplyv jazykového spoločenstva, ktorý kontroluje odklon od daných noriem a zabezpečuje stabilitu a súvislosť prenosu informácie.

Na nasledujúcom jednoduchom príklade uvedenom v [15, str. 90] si stručne priblížime postup modelovania dĺžok slov v počte slabík v básni *Večer po práci* od Milana Rúfusa.

x	f_x	x	f_x
0	7	3	27
1	33	4	6
2	49		

Tabuľka 1: Rozdelenie dĺžok slov v básni *Večer po práci*. Zdroj: [15, str. 90].

Tabuľka 1 obsahuje frekvenčne rozdelené dĺžky slov z básne, kde x predstavuje slovnú dĺžku v slabikách a f_x príslušnú frekvenciu.

Na skonštruovanie modelu predpokladáme, že $f(x) = a/x^b$ a dosadením do (4) dostaneme:

$$P_x = \frac{a}{x^b} \cdot P_{x-1}, \quad x = 1, 2, \dots$$

Dostali sme jednoduchú diferenčnú rovnicu, ktorú riešime

$$\begin{aligned}
 P_1 &= \frac{a}{1^b} \cdot P_0 \\
 P_2 &= \frac{a}{2^b} \cdot P_1 = \frac{a^2}{(2!)^b} \cdot P_0 \\
 P_3 &= \frac{a}{3^b} \cdot P_2 = \frac{a^3}{(3!)^b} \cdot P_0 \\
 &\vdots \\
 P_x &= \frac{a^x}{(x!)^b} \cdot P_0.
 \end{aligned} \tag{5}$$

Nakoľko $\sum_{x \geq 0} P_x = 1$, dá sa zistiť, že $P_0^{-1} = \sum_{j=0}^{\infty} \frac{a^j}{(j!)^b}$.

Rozdelenie (5) sa nazýva Conwayovo-Maxwellovo-Poissonovo rozdelenie. Na nafitovanie tohto rozdelenia na dáta z tabuľky 1 potrebujeme ešte odhadnúť parametre a a

b , na čo sa odporúča použiť softvér, ktorý automaticky nafituje rozdelenie na dáta a odhadne parametre.

V nasledujúcom zozname sú vymenované rozdelenia získané na základe prístupu $P_x = f(x) \cdot P_{x-1}$, ktoré sa vyskytli pri jednoduchom modelovaní dĺžok slov v rôznych literatúrach. Sú to Poissonovo, pozitívne Poissonovo, binomické, pozitívne binomické, negatívne binomické, pozitívne negatívne binomické, (pozitívne) hyperpoissonovské, (pozitívne) Conwayovo-Maxwellovo-Poissonovo, (pozitívne) hyperpascalovské, Palmovo-Poissonovo, Pólyovo, ...

Nakoľko musíme rátať s tým, že autori, ich kreativita a s nimi aj texty sa postupom času vyvíjajú, tak sa tieto zmeny musia prejaviť aj v modeloch a to pomocou modifikácií.

Existujú tri druhy modifikácií [15, 93]:

1. Prvou modifikáciou je rozšírenie k -tého rádu. V niektorých prípadoch môžeme predpokladať, že pravdepodobnosť výskytu slov v triede dĺžky x nie je závislá len od predošlej triedy dĺžky $x - 1$, ale od viacerých, prípadne od všetkých. To sa dá vyjadriť nasledovným vzorcom:

$$P_x = f(x) \cdot \sum_{j=1}^x h(j) \cdot P_{x-j}, \quad x = 1, 2, \dots,$$

kde $f(x)$ je proporčná funkcia a $h(j)$ je váha triedy $(x-j)$. Volením funkcie $f(x)$ a váhy $h(j)$ môžeme dostať celé triedy tzv. zovšeobecnených rozdelení.

2. Druhou modifikáciou sú lokálne modifikácie. Lokálna modifikácia sa v modeli prejaví tak, že sa vyzdvihne alebo potlačí frekvencia v niektorej triede a to buď v prospech/neprospech všetkých ostatných tried alebo len v prospech/neprospech susedných tried.
3. Treťou modifikáciou sú zmiešané rozdelenia. Môžu vzniknúť tým, že daný jazyk obsahuje dve vrstvy, napríklad pri angličtine je to románska a germánska vrstva, alebo tým, že text je napísaný viacerými autormi, poprípade si prešiel silnou dodatočnou korektúrou.

1.4 Zadefinovanie princípov pre ďalšie kapitoly

V nasledujúcich kapitolách sa budeme venovať modelovaniu dĺžky slova s ním súvisiacimi analýzami. Preto sme si zdefinovali princípy inšpirované podkapitolami 1.2 a 1.3, ktorých sme sa pridržovali, ale zároveň sme sa pokúsili priniest iný pohľad na danú problematiku:

1. Na začiatku sme si sformulovali hypotézy a až potom sme získavali texty, ktoré sme transformovali na dáta.
2. Pred transformáciou textov na dáta sme rozpísali skratky, čísla, dátumy a odstránili interpunkčné znamienka. Ak bolo v texte uvedené meno ako iniciála a v celom texte sa nenachádzalo rozpísané, túto iniciálu sme odstránili. Cudzie slová sme do slovenskej výslovnosti neprepisovali.
3. Dáta boli vytvorené na základe dĺžok slov, ktoré sme merali v počte písmen.
4. Texty sme vyhodnocovali po celkoch, t. j. nerobili sme z nich náhodné výbery slov ani sme ich nevyhodnocovali po častiach.
5. Ak boli texty členené na kompaktné časti ako napríklad kapitoly, tak každú jednu časť sme vyhodnotili zvlášť.
6. V niektorých prípadoch sme texty zmiešavali. Zvyčajne sa jednalo o texty s malým počtom slov, na ktorých sa nemohla urobiť potrebná analýza. Podmienkou na zmiešavanie textu bola rovnaká téma textu a autor, prípadne iba téma.
7. Pri modelovaní dát sme počítali chybovosť modelu a robili aj test vhodnosti modelu podľa kritéria (1).
8. Na modelovanie rozdelenia dĺžok slov všetkých textov sme používali rovnaký model.

2 Texty a ich spracovanie do dát

Skúmanie štruktúry slov, ktorému sa budeme venovať v nasledujúcich kapitolách si vyžaduje dáta spracované z textov v písanej forme. Tieto texty musia byť upravené do jednotnej podoby, ktorá je prehľadná a s ktorou sa dobre pracuje. Preto si v tejto kapitole priblížime priebeh získavania a spracovania textov do našej databázy, z ktorých sa tvoria dáta. Tiež sa pozrieme aj na typy textov, ich rozsah a zdroje, z ktorých pochádzajú.

V tejto kapitole sme čerpali poznatky z [11] a [13].

2.1 Zdroje, rozsah a typy textov

Prvým zdrojom textov, bola databáza vytvorená v práci [11], ktorá obsahovala 2 skupiny. Bolo v nich spolu 173 textových súborov, ktoré vznikli zdigitalizovaním a spracovaním textov v slovenskom jazyku publikovaných od roku 1845 po rok 2018.

V prvej skupine bolo 112 textových súborov a každý jeden z nich obsahoval niekoľko krátkych správ z novín vydaných v jednom roku. Všetky tieto novinové správy boli rovnakého charakteru. Druhou skupinou bolo 61 textových súborov a v každom bola jedna rozprávka.

Pre našu štúdiu sme si zobrali iba textové súbory obsahujúce rozprávky. Nakoľko táto databáza obsahovala aj texty veľmi starého dáta, v ktorých slovenčina nespĺňa dnešnú spisovnú normu, vybrali sme z nej iba také textové súbory s rozprávkami, ktoré obsahovali texty napísané po roku 1900. Týmto obmedzením sa nám výrazne zredukoval počet textových súborov a to na 28.

Ďalej sme si digitalizovali všetky rozprávky z knihy [5] pomocou metodiky uvedenej v [11, str. 13], čím sme dostali 13 textových súborov.

Nakoľko ručné digitalizovanie kníh je veľmi prácne a časovo náročné, tak sme sa snažili nájsť voľne dostupné a stiahnuteľné e-knihy. Poslúžili nám na to internetové stránky [9] a [4]. Z nich sme získali 20 kníh s rozprávkami, z ktorých vzniklo 325 textových súborov a v každom bola jedna rozprávka. Ďalšie rozprávky sme získali z internetovej stránky [3], kde sa nachádzalo 7 rozprávkových celkov s 36 rozprávkami.

Rozprávky tvorili hlavnú časť našej databázy. Okrem nich sme si spracovali 5 eko-

nomických článkov a recepty na špagety, torty, polievky a šaláty. Každý jeden textový súbor sme si uložili pod špeciálnym názvom, t.j. kódom, do priečinka podľa žánru resp. podľa témy.

2.2 Úprava textov a transformácia na dáta

Získané texty uložené v textových súboroch sme sa rozhodli upraviť pomocou matematického softvéru *Matlab* do formy, s ktorou by sa nám pri ďalších analýzach lepšie pracovalo.

Najprv sme si vytvorili tabuľku v programe *Excel*, ktorá obsahovala hypertextové prepojenia na tieto súbory. Pomocou nej sme jednotlivé textové súbory do *Matlabu* načítali.

Vytvorili sme skript, ktorý nám očistil texty od rôznych znakov, ktoré nie sú súčasťou slov ako je napríklad bodka, čiarka, výkričník, otáznik, pomlčka, spojovník, percento, zátvorky, úvodzovky a iné. Ponechal nám iba samostatne stojace slová oddelené medzerou, ktoré uložil do nového textového súboru po 15 slov v riadku s predponou v názve a zároveň vytvoril v novom stĺpci tabuľky hypertextové prepojenie na tieto upravené textové súbory.

Jednotlivé slová v textových súboroch sme považovali za pozorovania, z ktorých sme chceli vhodnou transformáciou získať dáta. Transformáciu sme urobili na základe dĺžky slova. Pre každé slovo sme pomocou ďalšieho skriptu spočítali počet písmen, z ktorého sa skladá. Tieto vzniknuté dáta sme uložili do nového textového súboru s opäť novou predponou a v tabuľke sa na ne vytvorilo hypertextové prepojenie v ďalšom stĺpci, čo máme možnosť vidieť na nasledujúcom obrázku 1.

Rozprávky				
	Kód	Textový súbor - originál	Textový súbor - upravený	Textový súbor - dáta
1	01zb1	Rozprávky\Originál\01zb1.txt	Rozprávky\Úprava\u01zb1.txt	Rozprávky\Dáta\d01zb1.txt
2	01zb2	Rozprávky\Originál\01zb2.txt	Rozprávky\Úprava\u01zb2.txt	Rozprávky\Dáta\d01zb2.txt
3	01zb3	Rozprávky\Originál\01zb3.txt	Rozprávky\Úprava\u01zb3.txt	Rozprávky\Dáta\d01zb3.txt
4	01zb4	Rozprávky\Originál\01zb4.txt	Rozprávky\Úprava\u01zb4.txt	Rozprávky\Dáta\d01zb4.txt

Obr. 1: Tabuľka s hypertextovými prepojeniami na textové súbory.

Priebeh úpravy textov a transformáciu na dáta môžeme vidieť na úryvku rozprávky z knihy [5] na nasledujúcom Obr. 2.

ORIGINAL TEXT:
O divotvornom hrnčíčku Boli raz dve deti, ktoré nemali ani otca, ani matku. Bývali v malom domčeku a veru im tam nebolo veľmi dobre, lebo nemali z čoho žiť. Zobral sa raz chlapček do hory niečo nazbierať. Ako tak ide cestou, vidí v jednej barine starú babku s vozíčkom. Nijako ho nemohla vytiahnuť z bariny. Janko nelenivý skočil a pomohol jej. Babka mu poďakovala a vraví:
UPRAVENÝ TEXT:
O divotvornom hrnčíčku Boli raz dve deti ktoré nemali ani otca ani matku Bývali v malom domčeku a veru im tam nebolo veľmi dobre lebo nemali z čoho žiť Zobral sa raz chlapček do hory niečo nazbierať Ako tak ide cestou vidí v jednej barine starú babku s vozíčkom Nijako ho nemohla vytiahnuť z bariny Janko nelenivý skočil a pomohol jej Babka mu poďakovala a vraví Vezmi si tento hrnček a choď domov Nebudete viac
TEXT TRANSFORMOVANÝ NA DÁTA:
1 10 8 4 3 3 4 5 6 3 4 3 5 6 1 5 7 1 4 2 3 6 5 5 4 6 1 4 3 6 2 3 8 2 4 5 9 3 3 3 6 4 1 6 6 5 5 1 8 6 2 7 9 1 6 5 8 6 1 7 3 5 2 10 1 5 5 2 5 6 1 4 5 8 4

Obr. 2: Priebeh úpravy textu a jeho transformácia na dáta.

3 Modelovanie dĺžky slov

V práci [11] bola pozorovaná bimodálna distribúcia dĺžky slov meraná v počte písmen v literárnych textoch v slovenskom jazyku. Modelovanie pomocou jedného unimodálneho rozdelenia nebolo vhodné, preto bol zostavený zmiešaný model pozostávajúci z dvoch gama rozdelení, ktorý uspokojivo dobre aproximoval rozdelenie dĺžky slov v jednotlivých textoch. Tento model si priblížime v prvej časti kapitoly.

Nakoľko dĺžky slov spracované do dát sú diskkrétne a sústredené na hodnotách zhruba od 1 do 10, pokúsime sme sa v nasledujúcich častiach kapitoly daný model zdiskretizovať. Na záver si zostrojené modely porovnáme.

Vychádzať budeme z [7], [8], [11] a [12].

3.1 Model využívajúci zmes gama rozdelení

Nech dĺžky slov merané v počte písmen sú náhodná premenná X . Model zostrojený ako zmes dvoch gama rozdelení [11, str. 24] aproximujúci rozdelenie dĺžky slov je funkciou s piatimi parametrami $p, k_1, \theta_1, k_2, \theta_2$, ktorá vyzerá nasledovne

$$f_{modelG}(x|p, k_1, \theta_1, k_2, \theta_2) = p \cdot f_G(x|k_1, \theta_1) + (1 - p) \cdot f_G(x|k_2, \theta_2), \quad (6)$$

kde $f_G(x|k, \theta)$ je funkcia hustoty gama rozdelenia. Funkcia hustoty gama rozdelenia pre náhodnú premennú X má tvar

$$f_G(x|k, \theta) = \begin{cases} \frac{x^{k-1} e^{-\frac{x}{\theta}}}{\theta^k \Gamma(k)}, & \text{ak } x > 0 \\ 0, & \text{ak } x \leq 0 \end{cases} \quad \text{pre } k > 0, \theta > 0, \quad (7)$$

kde $\Gamma(k)$ je gama funkcia s parametrom k daná ako

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

Optimálne parametre, ktoré sa dosadzovali do (6) sa získali vyriešením nasledovnej minimalizačnej úlohy

$$\begin{aligned} \min_{p, k_1, \theta_1, k_2, \theta_2} \quad & f_{CH}(p, k_1, \theta_1, k_2, \theta_2) \\ & - f_{modelG}(x|p, k_1, \theta_1, k_2, \theta_2) \leq 0 \\ & 0 \leq p \leq 1 \\ & k_1, \theta_1, k_2, \theta_2 \geq 0, \end{aligned} \quad (8)$$

v ktorej sa minimalizovala funkcia chyby

$$f_{CH} = \sum_{i=1}^n (y_i - f_{modelG}(x_i|p, k_1, \theta_1, k_2, \theta_2))^2, \quad (9)$$

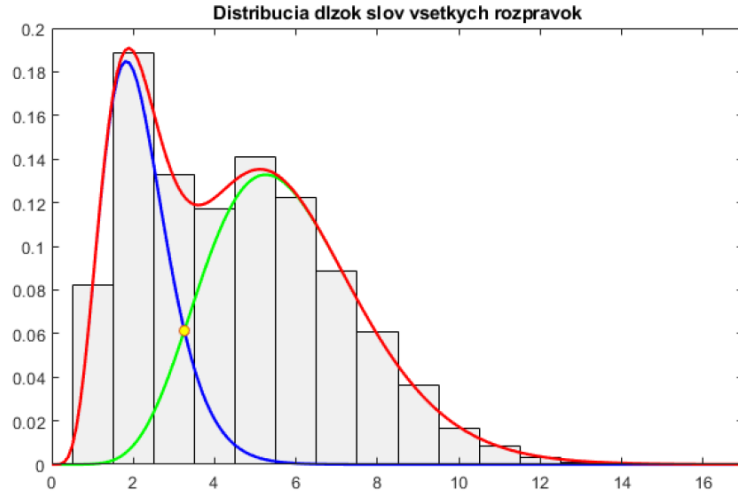
kde y_i predstavuje relatívnu početnosť dĺžky slov x_i a n je maximálna dĺžka slov.

Začiatkový bod x_0 pre optimalizáciu, ktorou sa získavali optimálne parametre sa volil nasledovne

$$x_0 = (0.5, k_{11}, \theta_{11}, k_{22}, \theta_{22}),$$

kde 0.5 je začiatková hodnota pre parameter p a θ_{11} , k_{11} , θ_{22} a k_{22} sú krajné body z 95% intervalov spoľahlivosti, do ktorých spadajú parametre pri nabitovaní jedného gama rozdelenia pomocou softvéru *Matlab*.

Na nasledujúcom Obr. 3 môžeme červenou farbou vidieť nabitovaný model na dátach rozprávkového charakteru z práce [11]. Modrou a zelenou sú vykreslené jednotlivé hustoty gama rozdelení uložené v modeli.



Obr. 3: Nabitovaný model (6) na rozdelení dĺžok slov u textov rozprávkového charakteru. Zdroj: [11, str. 33].

3.2 Diskretizácia modelu

Spôsobov diskretizácie spojitého rozdelenia existuje niekoľko. Jedným z nich je nasledujúci spôsob uvedený v [8, str. 3302]:

Ak spojitá náhodná premenná X má funkciu spoľahlivosti $S_X(x) = P(X \geq x)$, potom náhodná premenná $Y = \lfloor X \rfloor$ bude mať pravdepodobnostnú funkciu

$$\begin{aligned}
P(Y = y) &= P(y \leq X < y + 1) \\
&= P(X \geq y) - P(X \geq y + 1) \\
&= S_X(y) - S_X(y + 1), \quad \text{pre } y = 0, 1, 2, \dots
\end{aligned} \tag{10}$$

So spojitým gama rozdelením je úzko späté diskkrétne negatívne binomické rozdelenie, ktorého definícia je

Definícia 3.1 (Negatívne binomické rozdelenie $Ne(p, k)$, [12]). *Hovoríme, že náhodná premenná X má negatívne binomické rozdelenie s parametrami $p \in (0, 1)$ a $k \in N$, ak nadobúda hodnoty $x = 0, 1, \dots$ s pravdepodobnosťami*

$$P(X = x) = \binom{k+x-1}{x} \cdot p^k \cdot (1-p)^x \quad \text{pre } x = 0, 1, \dots$$

Aby negatívne binomické rozdelenie mohlo byť považované za diskrétnu verziu gama rozdelenia, tak parameter p musí mať tvar $e^{-\lambda}$ [8, str. 3304].

Na základe tohto poznatku sme si zostrojili nový model využívajúci negatívne binomické rozdelenie, ktorý mal tvar

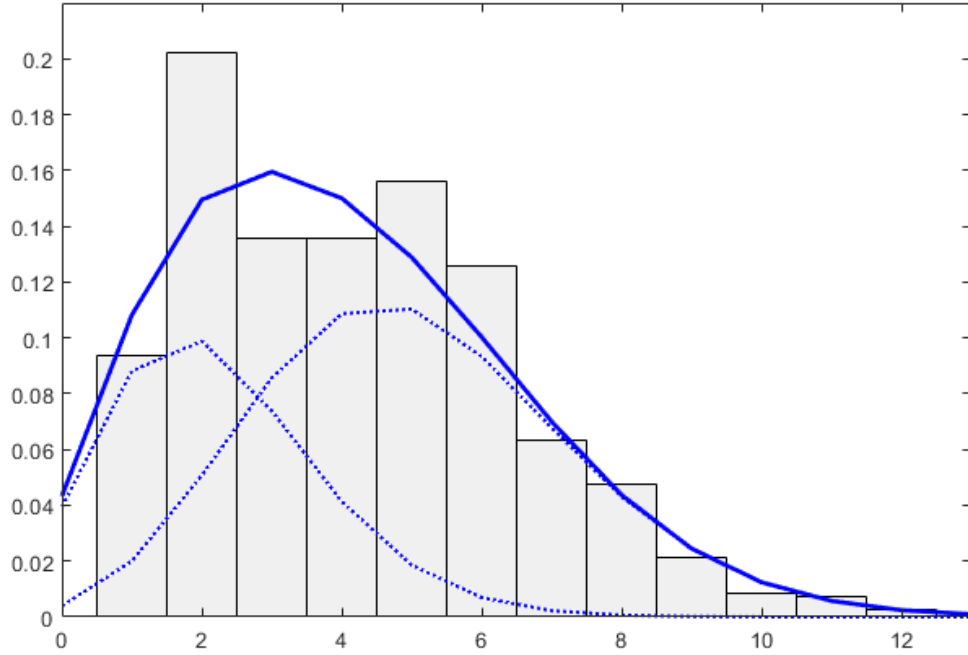
$$f_{modelNe}(x|p, e^{-\lambda_1}, k_1, e^{-\lambda_2}, k_2) = p \cdot f_{Ne}(x|e^{-\lambda_1}, k_1) + (1-p) \cdot f_{Ne}(x|e^{-\lambda_2}, k_2), \tag{11}$$

kde $f_{Ne}(x|e^{-\lambda}, k)$ je pravdepodobnostná funkcia negatívne binomického rozdelenia $Ne(e^{-\lambda}, k)$.

Na získanie optimálnych parametrov sme postupovali analogicky ako pri modeli (6) využívajúc (8) a (9) s tým rozdielom, že sme skúšali nastaviť rôzne verzie začiatočného bodu x_0 .

Tento model (11) nebol uspokojujúci, pretože koeficient vhodnosti modelu (1) vyšiel na všetkých nafitovaných dátach väčší ako 0.02.

Ukážku nafitovaného modelu (11) na dátach z rozprávky o Červenej Čiapočke z knihy [5] s príslušnými negatívne binomickými rozdeleniami uloženými v modeli (bodkované krivky) môžeme vidieť na Obr. 4.



Obr. 4: Nafitovaný model (11) na rozdelení dĺžok slov rozprávky o Červenej Čiapočke.

Nakoľko model (11) nebol vhodný, rozhodli sme sa zostaviť model využívajúci pravdepodobnostné funkcie diskretných gama rozdelení, ktorých tvar je uvedený v nasledujúcej definícii.

Definícia 3.2 (Pravdepodobnostná funkcia diskretného gama rozdelenia, [8]). *Nech X je spojitá náhodná premenná majúca gama rozdelenie $\gamma(k, \theta)$ s dvoma parametrami k a θ , s funkciou hustoty (γ) a funkciou spoľahlivosti*

$$S(x) = \frac{1}{\theta^k \cdot \Gamma(k)} \cdot \int_x^\infty u^{k-1} \cdot e^{-u/\theta} du.$$

Využitím prístupu diskretizácie (10) bude mať diskretné gama rozdelenie $d\gamma(k, \theta)$ s parametrami k a θ pravdepodobnostnú funkciu

$$\begin{aligned} P(Y = y) &= P(y \leq Y < y + 1) \\ &= S(y) - S(y + 1) \\ &= \frac{1}{\theta^k \cdot \Gamma(k)} \cdot \int_y^\infty x^{k-1} \cdot e^{-x/\theta} dx - \frac{1}{\theta^k \cdot \Gamma(k)} \cdot \int_{y+1}^\infty x^{k-1} \cdot e^{-x/\theta} dx \\ &= \frac{1}{\theta^k \cdot \Gamma(k)} \cdot \int_y^{y+1} x^{k-1} \cdot e^{-x/\theta} dx, \quad \text{pre } y = 0, 1, \dots; \quad k > 0; \quad \theta > 0. \end{aligned} \tag{12}$$

Zostavený model ako zmes dvoch diskretných gama funkcií mal tvar

$$f_{modelDG}(x|p, k_1, \theta_1, k_2, \theta_2) = p \cdot f_{DG}(x|k_1, \theta_1) + (1 - p) \cdot f_{DG}(x|k_2, \theta_2), \quad (13)$$

kde $f_{NG}(x|k, \theta)$ je pravdepodobnostná funkcia diskretného gama rozdelenia $d\gamma(k, \theta)$ (12).

Na získanie optimálnych parametrov sme postupovali opäť analogicky ako pri modeli (6) využívajúc (8) a (9). Začiatočný bod sme volili ako jednu z nasledujúcich možností:

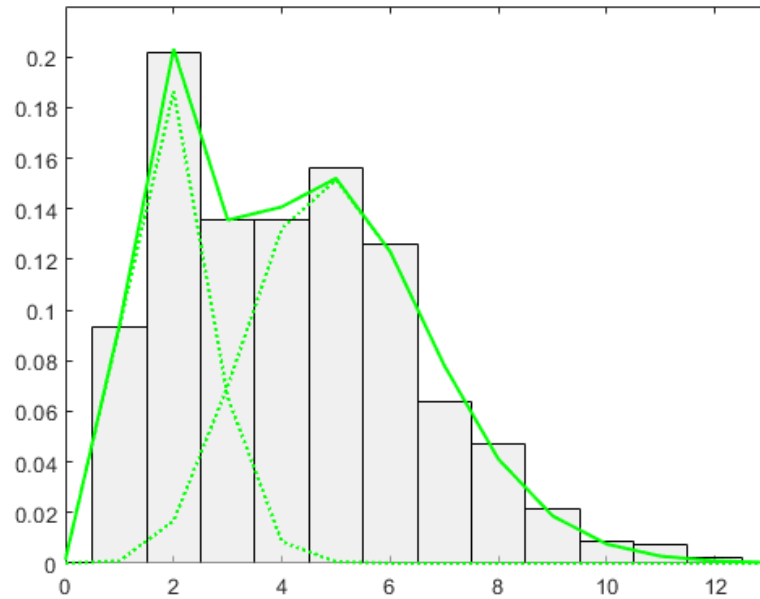
$$x_{0DG1} = (0.2, k_{11}, \theta_{11}, k_{22}, \theta_{22}),$$

$$x_{0DG2} = (p_G, k_{G1}, \theta_{G1}, k_{G2}, \theta_{G2}),$$

$$x_{0DG3} = (0.2, k_{G1}, \theta_{G1}, k_{G2}, \theta_{G2}),$$

kde x_{0DG1} je upravený začiatočný bod x_0 z modelu (6) a $p_G, k_{G1}, \theta_{G1}, k_{G2}, \theta_{G2}$ sú optimálne parametre získané nafitovaním modelu (6) na dátach.

Grafické znázornenie modelu (13) na dátach z rozprávky o Červenej Čiapočke z knihy [5] s príslušnými pravdepodobnostnými funkciami uloženými v modeli (bodkované krivky) môžeme vidieť na Obr. 5.



Obr. 5: Nafitovaný model (13) na rozdelení dĺžok slov rozprávky o Červenej Čiapočke.

3.3 Vhodnosť modelov

Pre modeli (6) a (13) sme urobili test vhodnosti modelu podľa kritéria (1) a taktiež sme sa pozreli na chybovosť, ktorú sme si definovali ako

$$CH = \sum_{i=1}^n (y_i - f_{model})^2, \quad (14)$$

kde y_i predstavuje relatívnu početnosť dĺžky slov i a n je maximálna dĺžka slov v texte.

Tieto údaje sme vypočítali pre všetky texty v databáze a výsledky sme uviedli v nasledujúcej Tabuľke 2.

		minimum	maximum	priemer	medián
model (6)	C	0.0011	58.2904	0.2792	0.0147
	CH	0.00002	0.0121	0.0010	0.00062
model (13)	C	0.0013	61.6506	0.2726	0.0152
	CH	0.00002	0.0127	0.0012	0.00064

Tabuľka 2: Ukazovatele vhodnosti modelov (6) a (13). C -kritérium vhodnosti modelu (1), CH -chybovosť modelu (14).

V Tabuľke 2 si v poslednom stĺpci môžeme všimnúť, že medián koeficientu vhodnosti modelu (1) vyšiel pod kritickou hodnotou 0.02 u oboch modelov. Pri dátach, v ktorých vyšiel vyšší ako 0.02 bol problém v krátkosti textu. Chybovosť bola postačujúco nízka. Preto v nasledujúcich analýzach používame oba modely.

4 Analýza štruktúr v písanom texte

V závere práce [11, str. 54] bola spomenutá otázka, čo presne môžu predstavovať jednotlivé gama rozdelenia v modeli (6). Na túto otázku boli načrtnuté tri hypotézy a to, že sa môže jednať o plnovýznamové a neplnovýznamové slová, o obsahové a neobsahové slová, alebo môže ísť o štýl autora a hlavnú informáciu v celého textu [11, str. 54]. Pretože štýl autora a hlavná informácia textu sa prejavujú na obsahových a neobsahových slovách, tak sa v tejto kapitole zameriame len na hypotézy o významnosti a obsahovosti textov, ktoré budeme analyzovať na našich dátach.

4.1 Plnovýznamové a neplnovýznamové slová

Prvá hypotéza, ktorou sme sa zaoberali bola o významnosti slov. Táto hypotéza tvrdí, že jedna funkcia hustoty gama rozdelenia z modelu (6), prípadne pravdepodobnostná funkcia diskrétného gama rozdelenia z modelu (13) kopíruje rozdelenie dĺžok slov kratších neplnovýznamových slov a druhá zase dlhšie plnovýznamové slová [11, str. 54].

Podľa významu sa slová delia na plnovýznamové a neplnovýznamové. Plnovýznamové slová majú svoj vecný význam a niečo pomenúvajú. Do tejto skupiny patria podstatné mená, prídavné mená, zámená, číslovky, slovesá a príslovky. Naopak neplnovýznamové slová nemajú samy o sebe vecný význam ani nič konkrétne nepomenúvajú, ale slúžia na vyjadrenie vzťahov. Sú to predložky, spojky, častice a citoslovčia.

Na to, aby sme vedeli automaticky rozlišovať plnovýznamové a neplnovýznamové slová v textoch, sme si potrebovali vytvoriť jednoduchý ľahko kontrolujúci mechanizmus.

Vytvorili sme si ho v programe *Excel* a fungoval na nasledujúcom princípe. V jednom hárku sme si vytvorili pomocné slovníky najčastejšie používaných neplnovýznamových slov. Do ďalšieho hárku sme si načítali textový súbor z databázy, ktorý bol upravený a očistený od interpunkčných znamienok a tak sa každé slovo zobrazilo do samostatnej bunky. Pomocou programovacieho jazyka *VBA* (Visual Basic for Applications) sme si vytvorili skript, ktorý po spustení prechádzal každým slovom z textu a porovnával ho zo slovami v slovníkoch. Ak sa našla zhoda, tak zhodujúce sa slová boli v texte farebne rozlíšené na základe príslušnosti k nejakému slovníku.

V prípade, že sa slovo z textu považovalo za plnovýznamové, tak danú bunku podfarbilo. Následne sme si prešli text a dodatočne vyfarbili/odfarbili tie slová, ktoré patrili do druhej skupiny. Keď sme mali text farebne rozlíšený podľa významovej príslušnosti, tak druhým skriptom sme si ho rozdelili do dvoch nových hárkov, z ktorých sme neplnovýznamové a plnovýznamové slová uložili do samostatných textových súborov.

Výber zo slovníkov neplnovýznamových slov môžeme vidieť na Obr. 6 a priebeh mechanizmu na ukážke textu z knihy [5] na Obr. 7.

PREDLOŽKY				ČASTICE		
z	Z	zo		nuž	Nuž	kiež
Poza	ponad	Ponad		Reku	vraj	Vraj
od	Od	okrem		just	Just	najmä
SPOJKY				CITOSLOVCA		
a	A	i		au	Au	oj
Ale	no	No		Ó	ach	Ach
že	Že	aby		fuj	Fuj	haha

Obr. 6: Výber zo slovníkov s neplnovýznamovými slovami.

Pôvodný upravený text:

O	divotvornom	hrnčičku	Boli	raz	dve	deti	ktoré	nemali	ani	otca	ani	matku	Bývali	v
malom	domčeku	a	veru	im	tam	nebolo	veľmi	dobre	lebo	nemali	z	čoho	žiť	Zobral
sa	raz	chlapček	do	hory	niečo	nazbierať	Ako	tak	ide	cestou	vidí	v	jednej	barine
starú	babku	s	vozičkom	Nijako	ho	nemohla	vytiahnuť	z	bariny	Janko	nelenivý	skočil	a	pomohol

Farebne rozlíšený text podľa príslušnosti k slovným druhom na základe významu:

O	divotvornom	hrnčičku	Boli	raz	dve	deti	ktoré	nemali	ani	otca	ani	matku	Bývali	v
malom	domčeku	a	veru	im	tam	nebolo	veľmi	dobre	lebo	nemali	z	čoho	žiť	Zobral
sa	raz	chlapček	do	hory	niečo	nazbierať	Ako	tak	ide	cestou	vidí	v	jednej	barine
starú	babku	s	vozičkom	Nijako	ho	nemohla	vytiahnuť	z	bariny	Janko	nelenivý	skočil	a	pomohol

Oddelené plnovýznamové slová:

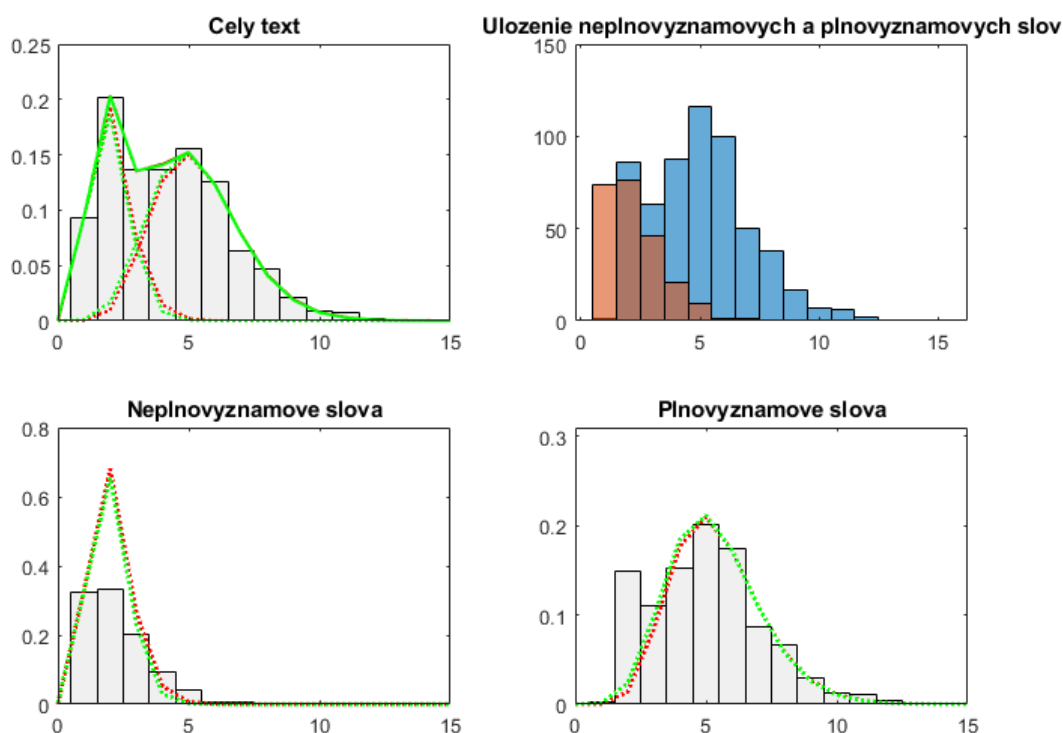
	divotvornom	hrnčičku	Boli		dve	deti		nemali		otca		matku	Bývali	
malom	domčeku			im	tam	nebolo	veľmi	dobre		nemali		čoho	žiť	Zobral
sa		chlapček		hory	niečo	nazbierať			ide	cestou	vidí		jednej	barine
starú	babku		vozičkom	Nijako	ho	nemohla	vytiahnuť		bariny	Janko	nelenivý	skočil		pomohol

Oddelené neplnovýznamové slová:

O			raz			ktoré		ani		ani				v
		a	veru					lebo		z				
	raz		do			Ako	tak				v			
		s					z					a		

Obr. 7: Priebeh mechanizmu na oddelenie slov podľa významu. Zdroj textu: [5].

Takto sme si spracovali 12 rozprávok z knihy [5]. Na dátach transformovaných z celého textu rozprávky sme nafitovali modely (6) a (13) a pozreli sme sa, či jednotlivé funkcie gama rozdelení kopírujú rozdelenia plnovýznamových a neplnovýznamových slov. Ukážku jednej rozprávky môžeme vidieť na nasledujúcom Obr. 8.



Obr. 8: Celkové rozdelenie dĺžok slov z rozprávky o Červenej Čiapočke z knihy [5] a jeho rozdelenie na neplnovýznamové a plnovýznamové slová s modelmi (6) a (13).

Na Obr. 8 môžeme vidieť, že plnovýznamové slová obsahujú veľkú časť slov, ktorá je v celkovom rozdelení popísaná prvou funkciou hustoty gama rozdelenia (červená krivka), resp. pravdepodobnostnou funkciou diskrétného gama rozdelenia (zelená krivka) z modelov.

Na všetkých rozprávkach z knihy [5] zaradenie plnovýznamových a neplnovýznamových slov vyšlo obdobne, ž čoho sme usúdili, že jednotlivé gama funkcie z modelov nepopisujú rozdelenia dĺžok slov podľa významu.

Túto hypotézu sme preto zamietli.

4.2 Obsahové a neobsahové slová

V tejto podkapitole sme overovali hypotézu, ktorá hovorí, že slovenský jazyk v písaných textoch možno rozdeliť na dve vrstvy, ktoré sú tvorené obsahovými a neobsahovými slovami a zároveň sú rozdelenia dĺžok týchto slov popísané gama funkciami uloženými v modeloch (6) a (13) [11, str. 54].

Za neobsahové slová sme pokladali také, bez ktorých by dával text zmysel a pochopili by sme jeho hlavnú informáciu. Patria sem všetky neplnovýznamové slová a takmer celá trieda zámen. K obsahovým slovám sme radili podstatné mená, prídavné mená, číslovky a takmer všetky slovesá a príslovky. Slová, ktoré sme na začiatku analýzy nezaradili ani do jednej skupiny sme nazývali strednou triedou a boli to nasledovné podskupiny slov:

1. Opakujúce sa slová - boli to slová, ktoré sa vyskytovali v texte viackrát za sebou a ich opakovanie nedávalo obsahu textu pridanú hodnotu. Napríklad vo vete *Ide, ide, stretne vlka* [5]. bolo druhé *ide* bez výpovednej hodnoty. Rovnako to boli aj tri za sebou idúce slová *var* na konci vety *Var, hrnčiček, var, var, var* [5].
2. Pomocné slovesá - boli to slovesá, ktoré pomáhali pri tvorení časov, ale z kontextu textu sme ich mohli vynechať, pretože sme rozumeli, čo sa udialo, deje alebo sa bude diať. Napr. vo vete *Čiapočka položí, košík na trávku a ide trhať kvety*[5]. bolo sloveso *ide* nepotrebné.
3. Príslovky - boli to príslovky, ktoré niečo umocňovali (*veľmi, viac,...*), nepriamo určovali čas (*potom, hneď, teraz,...*) alebo to boli zámenné príslovky (*tam, akože...*). Napr. vo vete *Veľmi vďačne jablonka*[5]. príslovku *veľmi* sme mohli vynechať, lebo v rámci kontextu sme rozumeli, že daná osoba niečo vďačne urobí.
4. Iné slová - do tejto skupiny sme radili slová, ktoré sa v rámci kontextu radili k rôznym slovným druhom, napr. *čosi, raz, práve, hádam,...*

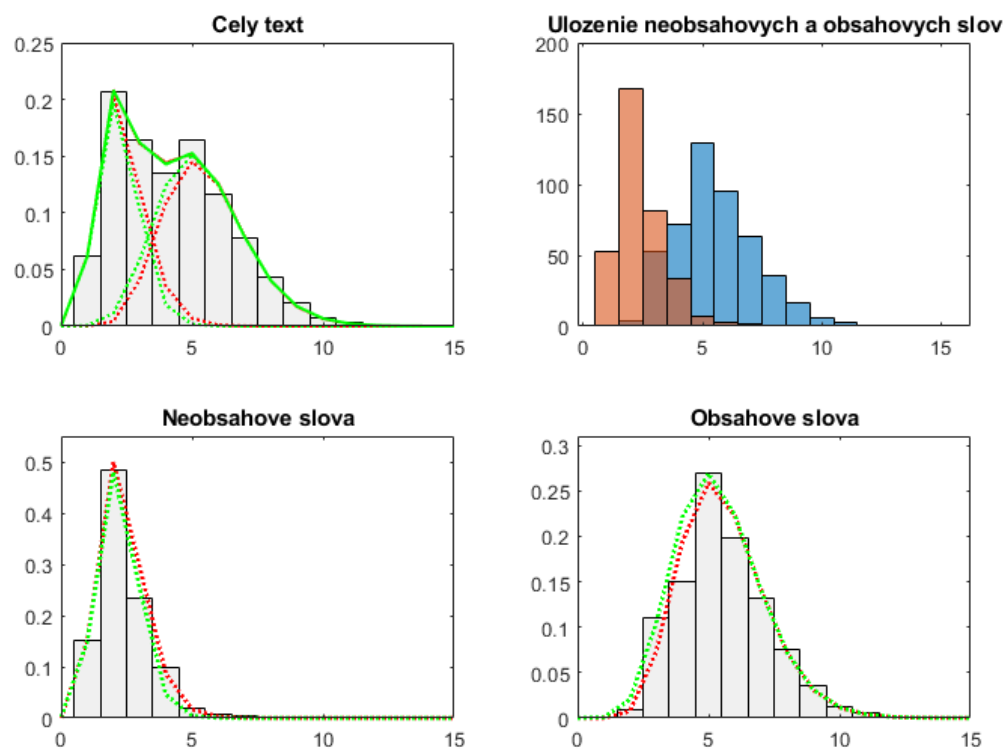
Pri rozdelení textu na obsahové a neobsahové slová sme postupovali analogicky ako v podkapitole 4.1 s tým rozdielom, že sme si zadefinovali slovník najčastejšie používaných zámen a dodatočným prechádzaním textu sme farebne odlíšili 4 podskupiny tvoriace

strednú triedu. Takto nám z jedného textu vznikli dve hlavné skupiny slov - obsahové a neobsahové a štyri podskupiny strednej triedy, ktoré sme oddelili do samostatných textových súborov a transformovali sme ich na dáta.

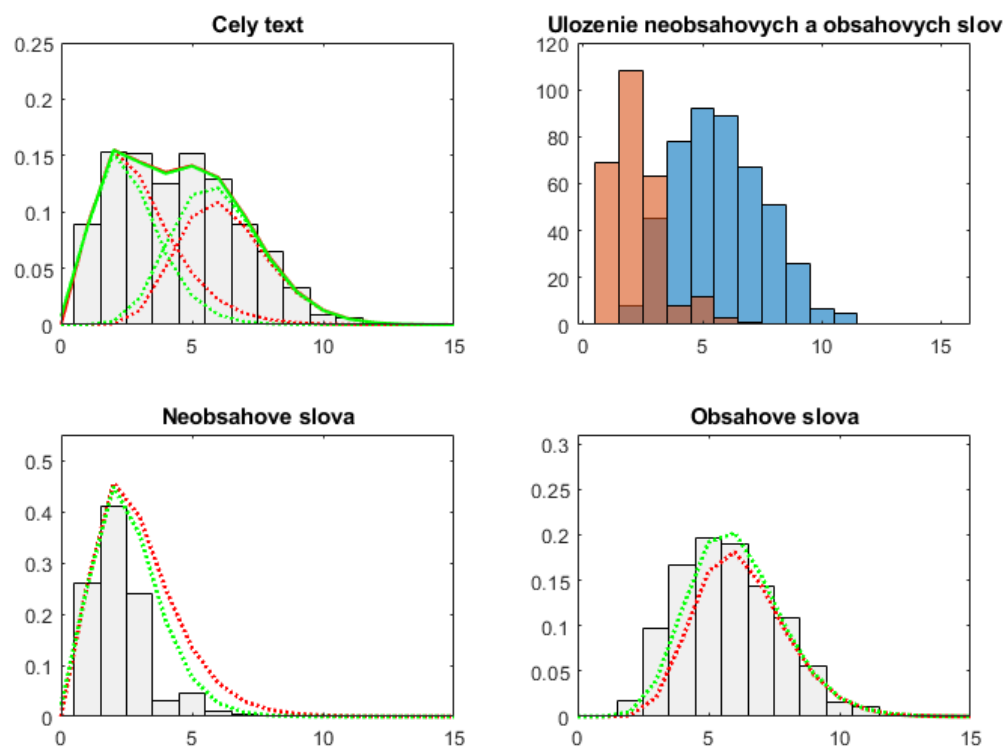
Následne sme pre každý analyzovaný text na základe kritéria vhodnosti modelu (1) a chybovosti (14) zisťovali, či stredná trieda patrí do skupiny obsahových alebo neobsahových slov tak, že sme postupne k hlavným skupinám neobsahových a obsahových slov pridávali kombinácie z podskupín strednej triedy.

Zároveň sme si do grafov vykreslili rozdelenie dĺžky slov s naftovanými modelmi (6) a (13), uložením obsahových a neobsahových slov v rozdelení dĺžky slov s ich prekryvom bez strednej triedy a nakoniec obsahové a neobsahové slová s jednotlivými gama funkciami z modelov taktiež bez strednej triedy.

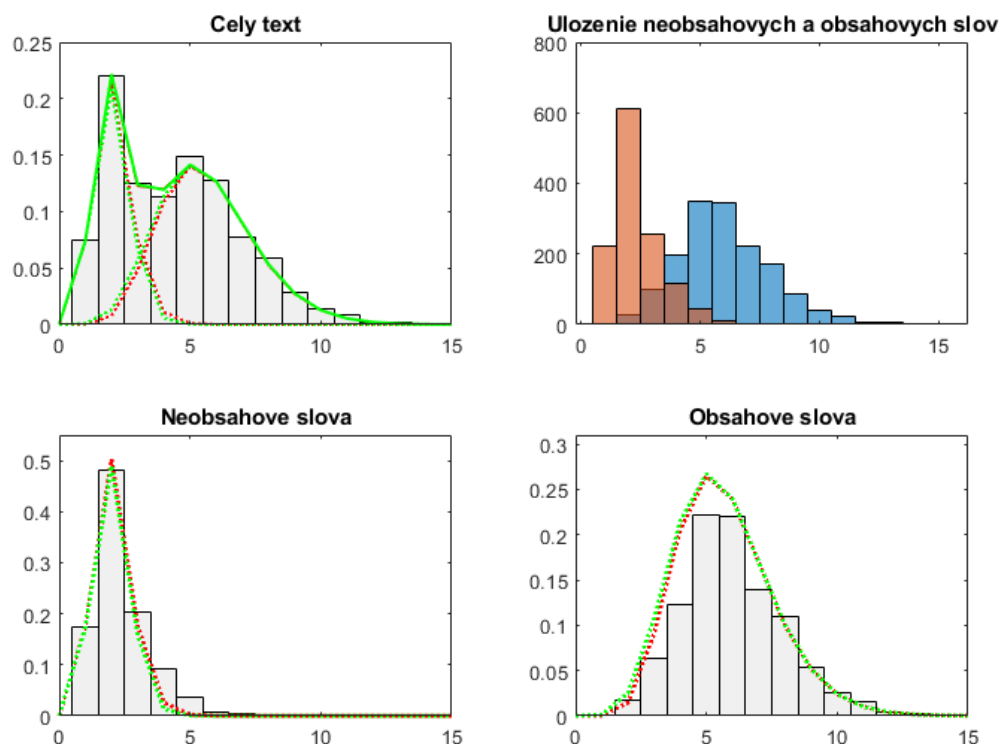
Tieto výpočty a grafické znázornenia sme urobili pre 12 rozprávok z knihy [5] a zistili sme, že nie je jednoznačné kam zaradiť strednú triedu slov. Pri niektorých rozprávkach sa stredná trieda radila k obsahovým slová, pri iných jasne k neobsahovým a u niektorých rozprávok nebolo jasné kam ju zaradiť, čo môžeme vidieť v nasledujúcich troch tabuľkách.



Obr. 9: r01zb1



Obr. 10: r01zb2



Obr. 11: r01zb7

Z Tabuliek v prílohách k daným rozprávkam na Obr. 9, 10 a 11 môžeme vyčítať nasledujúce informácie:

- kód rozprávky (01zb1 a pod.);
- koeficient vhodnosti modelu C podľa kritéria (1) pre model (6) a Cdg pre model (13), ktoré sme nafitovali na celkové rozdelenie dĺžok slov daného textu (druhý riadok);
- chybovosť CH podľa (14) pre model (6) a CHdg pre model (13), ktoré sme nafitovali na celkové rozdelenie dĺžok slov daného textu (druhý riadok);
- počet slov nachádzajúcich sa v celom texte (druhý riadok a šiesty stĺpec);
- koeficienty C, CH rátané s príslušnou funkciou hustoty gama rozdelenia z modelu (6) zodpovedajúcou hlavnej skupine neobsahových alebo obsahových slov a ich kombináciám so strednou triedou (2., 3., 7. a 8. stĺpec);

- koeficienty Cdg, CHdg rátané s príslušnou pravdepodobnostnou funkciou diskrétného gama rozdelenia z modelu (13) zodpovedajúcou hlavnej skupine neobsahových alebo obsahových slov a ich kombináciám so strednou triedou (4., 5., 9. a 10. stĺpec);
- kombinácie písané v skratkách, ktoré znamenajú neob = neobsahové slová, ob = obsahové slová, op = opakujúce sa slová, pr = príslovky, sl = pomocné slovesá, ine = iné slová;
- farebne sú rozlíšené tri najnižšie hodnoty koeficientov C, Cdg, CH, CHdg, podľa ktorých sme sa pozerali na zaradenie strednej k hlavným skupinám.

Na ukázaných Obr. 9, Obr. 10 a Obr. 11 sme si všimli pre našu prácu zaujímavé skutočnosti. Uložené gama funkcie v modeloch (6) a (13) samostatne dobre popisujú skupinu obsahových, či neobsahových slov.

Nakoľko táto analýza bola robená na textoch rozprávkového charakteru, pozreli sme sa na celok ekonomických textov, v ktorom bolo 5 článkov s témou o 10 rokoch eura na Slovensku a na štyri celky s receptami špagiet, polievok, nápojov a celok o tortách. Aj na týchto celkoch sme spozorovali rovnakú zákonitosť ako u rozprávok.

To znamená, že v slovenčine existujú dve vrstvy jazyka, ktoré sa vzájomne dopĺňajú, no na pochopenie textu stačí človeku čítať len obsahové slová s prípadnou strednou triedou.

Bližšie sme sa pozreli na jednotlivé texty daných troch rozprávok a nezistili sme žiadnu zákonitosť, či črtu textu, aby sme mohli povedať kvôli čomu sa stredná trieda radí, buď k obsahovým alebo k neobsahovým slovám, prípadne ako v Tabuľke ?? k obom skupinám, no môže úzko súvisieť so štýlom autora.

5 Bimodalita rozdelenia dĺžok slov v textoch

Testovanie miery bimodality v dátach je dôležité, pretože nám môže odhaliť ich rôznorodosť ako napríklad dve rozdelenia skombinované do jedného. V predchádzajúcej kapitole sme si ukázali, že celkové rozdelenie dĺžok slov v textoch pozostáva z dvoch samostatných rozdelení. Jedno rozdelenie kopíruje tzv. obsahové slová a druhé neobsahové slová.

Existuje niekoľko výpočtov na zistenie bimodality v dátach. V [10] Thomas R. Knapp zhrnul väčšinu z nich a aj s nimi spojené výhody a nevýhody.

V tejto kapitole si priblížime tri výpočty, aplikujeme ich na naše dáta a pozrieme sa, či odhalili bimodálne rozdelenie dĺžok slov v textoch. Dopĺňujúce poznatky sme čerpali z [14].

5.1 Štandardné výpočty miery bimodality

Prvý výpočet na zistenie miery bimodality uvedený v [10] bol od **Karla Pearsona**. Ten vymyslel zložitý postup, ktorým zisťoval, či rozdelenie početností môže byť rozdelené na dve samostatné normálne rozdelenia. Dospel k záveru, že ak dané rozdelenie má dva vrcholy ležiace dostatočne ďaleko od seba, môže sa označovať ako bimodálne.

Neskôr Karl Pearson zdefinoval nasledujúci výpočet

$$B = b_2 - b_1, \quad (15)$$

kde b_2 je štandardizovaný štvrtý štatistický moment - špicatosť a b_1 je štvorec štandardizovaného tretieho štatistického momentu - šikmosti.

B musí byť väčšie alebo rovné 1. Ak $B = 1$, tak ide o najextrémnejší prípad bimodality, ktorý má alternatívne (Bernoulliho) rozdelenie.

Ďalší vzorec, ktorý sa zaoberá bimodalitou rozdelenia, zdefinoval **Brad S. Chissom** nasledovne

$$\alpha_4 = b_2 - 3. \quad (16)$$

Chissom ukázal, že pre normálne rozdelenie je α_4 rovná 0 a ak $\alpha_4 = -2$, tak ide o perfektné bimodálne rozdelenie.

Posledný koeficient, ktorý sme si z [10] vybrali, zdefinoval **Warren Sarle** nasledovne

$$BC_1 = \frac{m_3^2 + 1}{m_4 + 3 \cdot \frac{(n-1)^2}{(n-2) \cdot (n-3)}}, \quad (17)$$

kde m_3 je šikmosť rozdelenia, m_4 je špicatosť rozdelenia a n je veľkosť dát.

Kritická hodnota pre BC_1 je $5/9 \approx 0.555$, ktorú by malo nadobúdať rovnomerné rozdelenie. Vyššie hodnoty môžu indikovať bimodálne alebo multimodálne rozdelenie. Maximálna hodnota pre BC_1 je 1 a tú dosahuje alternatívne (Bernoulliho) rozdelenie.

Vo vzorci (17) sú použité neštandardizované štatistické momenty šikmosť a špicatosť, ktoré sa dajú nahradiť pomocou b_1 a b_2 . Vďaka tomu sa zápis vzorca zjednoduší do tvaru

$$BC_2 = \frac{b_1 + 1}{b_2}. \quad (18)$$

5.2 Aplikácia na dáta

Vzorci (15), (16), (17) a (18) sme aplikovali v softvéri *Matlab* na každý dátový súbor získaný z textov rozprávok, ktorých bolo 401. Taktiež sme sa pozreli pomocou funkcie *findpeaks*, koľko lokálnych maxím, resp. vrcholov sa v dátach nachádza. Zo získaných výsledkov sme pre každý koeficient vypočítali minimálnu, maximálnu, priemernú a mediánovú hodnotu. Tieto hodnoty môžeme vidieť na nasledujúcej tabuľke 3.

	minimum	maximum	priemer	medián
počet vrcholov	1	5	2.25	2
B	1.8461	4.3163	2.4540	2.4354
α_4	-1.0562	2.2233	-0.1990	-0.2183
BC_1	0.3395	0.6210	0.4802	0.4848
BC_2	0.3403	0.6486	0.4829	0.4879

Tabuľka 3: Výsledky koeficientov bimodality na dátach získaných z textov rozprávok.

V Tabuľke 3 si môžeme všimnúť, že počet lokálnych maxím, teda vrcholov v rozdelení dát bol vo väčšine rozprávok 2. Iba 7% rozprávok malo 1 vrchol a vo väčšine prípadov

šlo o kratšie rozprávky. 68% rozprávok malo počet vrcholov rovný 2 a zvyšných 25% ich malo viac, čo bolo spôsobené prevažne tým, že napr. počet slov s desiatimi písmenami bol nula, ale s jedenástimi písmenami boli tri.

Koeficienty hovoriace o miere bimodality B a α_4 vyšli dosť vysoké na to, aby sme mohli tvrdiť, že rozdelenia počtov písmen v rozprávkach sú bimodálne. Iba 2 rozprávky mali α_4 menšiu ako -1 a 99% textov malo B väčšie ako 2.

Koeficienty BC_1 a BC_2 , ktoré majú určenú kritickú hodnotu 0.555, iba na 1% indikovali možnú bimodalitu.

Odpoveďou na otázku, prečo jednotlivé koeficienty nevykazovali bimodalitu môže byť fakt, že boli konštruované pre zmes normálnych rozdelení a v našich dátach sa nachádza zmes dvoch gama rozdelení [10]. Ďalším dôvodom môže byť to, že jednotlivé vrcholy sa nachádzajú veľmi blízko seba (prvý vrchol je zvyčajne na hodnote 2 a druhý vrchol na hodnote 5) a pokles medzi nimi nie je dostatočne hlboký.

Problém môže byť aj v chybovosti koeficientu BC_2 , čo ukázal autor článku [10] na 10 dátových súboroch.

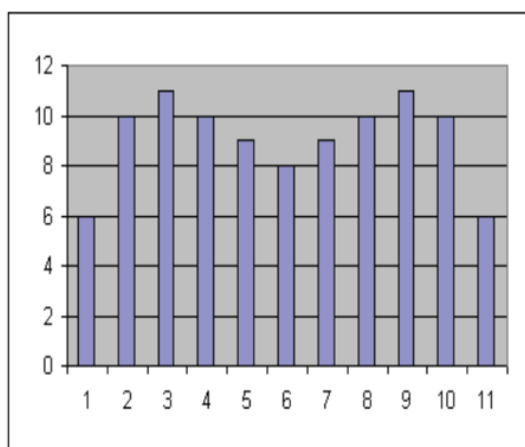


Figure 1

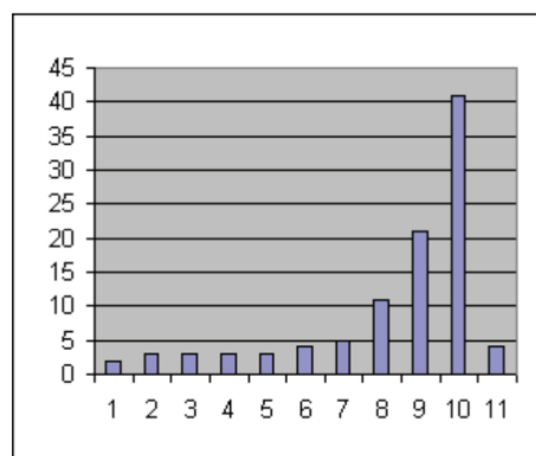


Figure 7

Figure	b_1	b_2	$b_2 - b_1$	$(b_1 + 1)/b_2$	$b_2 - 3$
1	0.000	1.757	1.757	0.569	-1.243
2	2.443	4.414	1.971	0.780	1.414

Obr. 12: Zdroj: [10, str. 11, 14, 15].

Na Obr. 12 sú vykreslené dve rozdelenia dát z [10] a vypočítané ich príslušné koeficienty bimodality. Na prvom rozdelení dát (Figure 1) môžeme voľným okom pozorovať bimodalitu na rozdiel od druhého rozdelenia dát (Figure 7), ktoré by malo byť unimodálne. Koeficienty $B = b_2 - b_1$ a $\alpha_4 = b_2 - 3$ ukázali mieru bimodality v prospech prvého rozdelenia. No koeficient $BC_2 = (b_1 + 1)/b_2$ bol výrazne vyšší pre druhé rozdelenie.

V nasledujúcej analýze sme sa pozreli, či tieto koeficienty zlyhajú aj keď zoberieme dve presné diskkrétne gama rozdelenia z modelu (??) nafitovanom na rozprávke o Červenej čiapočke z knihy [5]. Následne sme jedno gama rozdelenie postupne posúvali o jednotku doprava, čím sme zväčšovali vzdialenosť a pokles medzi vrcholmi, čo môžeme vidieť na nasledujúcom obrázku.

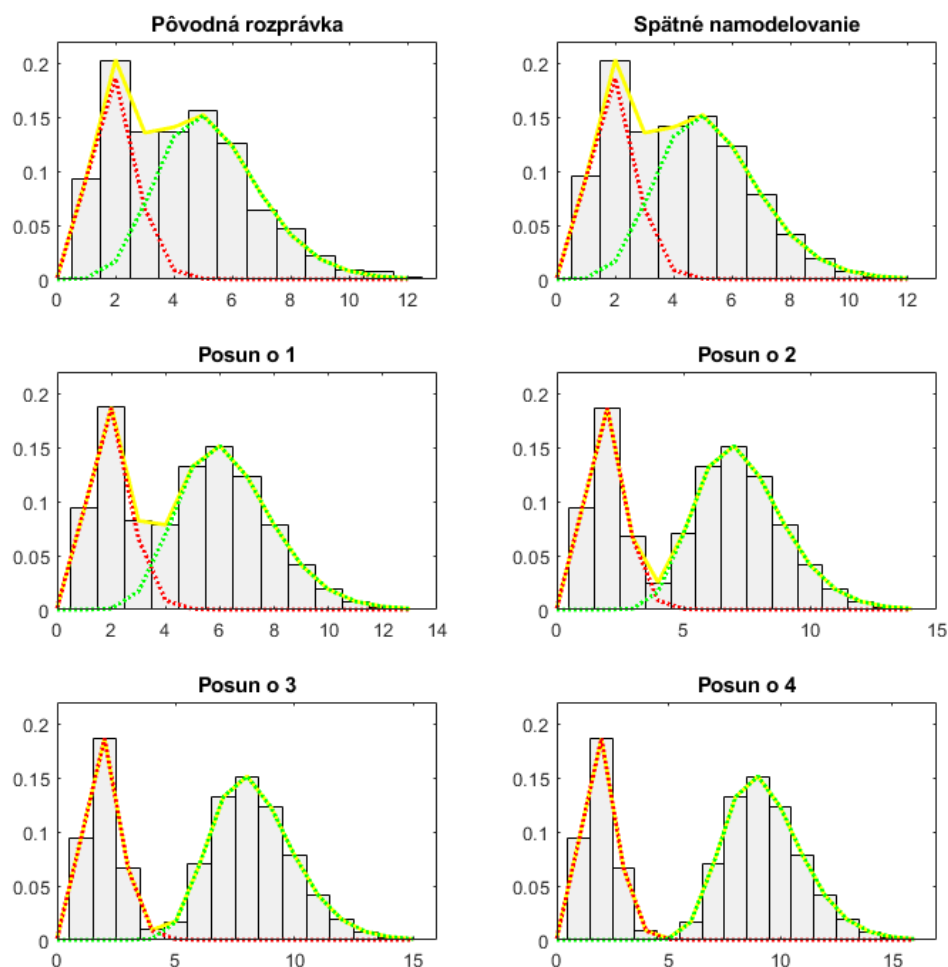
Na Obr. 13 je na prvom grafe vykreslené rozdelenie dĺžok slov rozprávky o Červenej Čiapočke s nafitovaným modelom (žltej farby) a s príslušnými gama rozdeleniami uloženými v modeli (červená a zelená farba). Druhý graf obsahuje presné rozdelenie z modelu, vďaka ktorému sme si spätne namodelovali dáta potrebné na výpočet koeficientov bimodality. Na zvyšných štyroch grafoch sú hodnoty v druhom gama rozdelení postupne zvyšované o jednotku.

	pôvodné	spätne namodelovanie	posun o 1	posun o 2	posun o 3	posun o 4
B	2.5650	2.4021	2.1607	1.9456	1.7670	1.6250
α_4	-0.0703	-0.3577	-0.7876	-1.0535	-1.2200	-1.3276
BC_1	0.4634	0.4668	0.4725	0.5110	0.5656	0.6225
BC_2	0.4658	0.4694	0.4754	0.5142	0.5691	0.6263

Tabuľka 4: Koeficienty miery bimodality.

Pre všetkých 6 prípadov sme si vypočítali hodnoty koeficientov, ktoré sú uvedené v Tabuľke 4. Najmenej bimodálne je rozdelenie pôvodnej rozprávky. Spätne namodelovanie dát, v ktorých sa nachádzali explicitné diskkrétne gama rozdelenia veľmi nepomohlo k odhaleniu bimodality. Výsledky koeficientov sa začali zlepšovať až keď sme jedno gama rozdelenie posúvali. Kritická hodnota pri koeficientoch BC_1 a BC_2 sa prekročila až pri posune dát jedného z gama rozdelení o 3.

Na záver možno skonštatovať, že koeficienty bimodality (15), (16), (17) a (18) neboli



Obr. 13: Postupné posúvanie pravdepodobnostnej funkcie diskrétného gama rozdelenia z modelu (13).

vhodné na naše rozdelenia dĺžok slov, pretože bimodalitu nepotvrdili. Zároveň však môžeme povedať, že bimodalita sa potvrdila tým, že sme dĺžky slov v jednotlivých textoch namodelovali zmesou dvoch gama rozdelení a následne sme jednotlivé gama rozdelenia identifikovali.

Záver

Cieľom tejto diplomovej práce bolo nájsť takú štruktúru slov v písaných textoch, ktorá by bola jasne interpretovateľná a dala by sa ľahko modelovať.

V prvej kapitole sme si priblížili historické korene skúmania dĺžky slova, odporúčané princípy na získanie dát a ich úpravu, pozreli sme sa na zaužívané modelovanie rozdelenia dĺžky slova a vymedzili sme si princípy, ktorých sme sa pridržovali v nasledujúcich kapitolách.

Následne sa druhá kapitola venovala získavaniu textov, ktoré sme transformovali na dáta, ktoré boli tvorené počtami písmen v slovách jednotlivých textoch. Transformáciu textov na dáta sme si ukázali na príklade v Obr. 2.

Pre vytvorené dáta sme si v tretej kapitole priblížili zmiešaný model dvoch gama rozdelení (6) z práce [11]. Nakoľko dĺžky slov spracované do dát boli diskkrétne a model využíval zmes spojitých gama rozdelení, ukázali sme odvodenie pravdepodobnostnej funkcie diskrétného gama rozdelenia. Túto pravdepodobnostnú funkciu sme využili pri zostavení modelu (13). Oba modely sme nafitovali na všetky dáta z databázy a pomocou testu vhodnosti modelu a chybovosti sme demonštrovali správnosť zostavených modelov, ktoré sme použili pri analýze štruktúr slov v písaných textoch v nasledujúcej kapitole.

Pomocou zmiešaného modelu dvoch gama rozdelení (6) a zmiešaného modelu dvoch diskrétnych gama rozdelení (13) sme modelovali dĺžku slov v počte písmen jednotlivých textov. Zistili sme, že jednotlivé gama rozdelenia uložené v modeli kopírujú dve vrstvy jazyka. Týmito dvoma vrstvami boli tzv. obsahové a neobsahové slová. Pri dostatočnej znalosti slovenského jazyka človek dokáže pochopiť základnú informáciu z textu, ktorý je bez neobsahových slov. Sú to také slová, ktoré sami o sebe nemajú význam a to neplnovýznamové slová, ale aj také, ktoré nenesú hlavnú informáciu. Môžeme do nich zaradiť všetky zámená a niektoré príslovky, pomocné slovesá či iné slová, ktoré však závisia od kontextu.

Toto zistenie považujeme za dôležité, pretože nám hovorí akým spôsobom používame jazyk. Jeho aplikáciu si vieme predstaviť pri chatbotoch, ktoré sa často využívajú na komunikáciu s klientami, ktorá by sa dala výrazne zlepšiť.

V záverečnej kapitole sme sa venovali bimodalite rozdelenia dĺžok slov v textoch.

Pozreli sme sa na štandardné výpočty miery bimodality, ktoré však neboli vhodné pre naše dáta, čo sme ukázali.

Literatúra

- [1] Ameijeiras-Alonso J., et al.: *Package ‘multimode’*, 2018, dostupné na internete (1.4.2020): <https://cran.r-project.org/web/packages/multimode/multimode.pdf>
- [2] Antić, G., Kelih, E., Grzybek, P.: Zero-Syllable Words in Determining Word Length. In: Grzybek, P. *Contributions to the Science of Text and Language*, Springer, Dordrecht, 2007, 117-156
- [3] Blog šikovná mamina: *Rozprávky na dobrú noc*, dostupné na internete (21.3.2020): <https://www.sikovnamamina.sk/blog/rozpravka-na-dobru-noc/>
- [4] Denník SME: *Zlatý fond*, dostupné na internete (21.10.2019): <https://zlatyfond.sme.sk/>
- [5] Ďuričková, M.: *Zlatá brána*, Mladé letá, Bratislava, 1975
- [6] Eknížky, dostupné na internete (10.1.2020): <https://eknizky.sk>
- [7] Grzybek, P.: History and Methodology of Word Length Studies. In: Grzybek, P. *Contributions to the Science of Text and Language*, Springer, Dordrecht, 2007, 15-90
- [8] Chakraborty, S., Chakravarty, D.: Discrete Gamma Distribution: Properties and Parameter Estimations. In: Balakrishnan, N., et al eds. *Communications in Statistics - Theory and Methods, Volume 41*, Taylor and Francis, 2012, 3301-3323
- [9] Knihy Václava Šuplatu, dostupné na internete (10.1.2020): <http://www.knihysuplata.sk/knihy/>
- [10] Knapp, T.R.: Bimodality Revised. In: *Journal of Modern Applied Statistical Methods, Vol. 6*, 2007, 8-20, dostupné na internete (20.4.2020): <https://digitalcommons.wayne.edu/cgi/viewcontent.cgi?article=1120&context=jmasm>
- [11] Košútová, A.: *Kvantitatívna analýza vývoja komplexity textov v slovenskom jazyku*, bakalárska práca, FMFI UK, Bratislava, 2018

- [12] Markechová, D., Stehlíková, B., Tirpáková, A.: *štatistické metódy a ich aplikácie*, Nitra, 2011
- [13] MathWorks, dostupné na internete (10.1.2020): <https://mathworks.com/products/matlab.html>
- [14] Pfister, R., Schwarz, K. A., Janczyk, M., Dale, R., Freeman, J. B.: Good things peak in pairs: a note on the bimodality coefficient. In: *Frontiers in psychology*, 2013, dostupné na internete (1.4.2020): <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3791391/pdf/fpsyg-04-00700.pdf>
- [15] Wimmer, G., Altmann, L., Hřebíček, L., Ondrejovič, S., Wimmerová, S.: *Úvod do analýzy textov*, VEDA, Bratislava, 2003

Prílohy

Zoznam rozprávok (Kód, Názov rozprávky, Názov knihy, Autor):

1. r01zb1, Starý Bodrík a vlk, Zlatá brána, Mária Rázusová - Martáková
2. r01zb2, Červená Čiapočka, Zlatá brána, Mária Rázusová - Martáková
3. r01zb3, Oráč a obri, Zlatá brána, Mária Ďuríčková
4. r01zb4, O divotvornom hrnčíčku, Zlatá brána, bez autora
5. r01zb5, Ako šlo vajce na vandrovkú, Zlatá brána, Mária Rázusová - Martáková
6. r01zb6, Dve sestry, Zlatá brána, Mária Ďuríčková
7. r01zb7, Janko Hraško, Zlatá brána, Mária Ďuríčková
8. r01zb8, Koza odratá a jež, Zlatá brána, Mária Rázusová - Martáková
9. r01zb9, Kozliatka, Zlatá brána, Mária Rázusová - Martáková
10. r01zb10, Krátka rozprávka, Zlatá brána, Pavol Dobšinský
11. r01zb11, O Jankovi Polienkovi, Zlatá brána, Mária Ďuríčková
12. r01zb12, O medovníkovom domčeku, Zlatá brána, Mária Rázusová - Martáková
13. r01vs1, Aj duchovia sa vedia báť, S duchmi je vždy veselo, Václav Šuplata
14. r01vs2, Aj duchovia sa radi hrajú, S duchmi je vždy veselo, Václav Šuplata
15. r01vs3, Aj duchom môže byť zima, S duchmi je vždy veselo, Václav Šuplata
16. r01vs4, Aj duchovia potrebujú okuliare, S duchmi je vždy veselo, Václav Šuplata
17. r01vs5, Aj duchovia chcú byť pekní, S duchmi je vždy veselo, Václav Šuplata
18. r01vs6, Aj duchovia majú svoje tajomstvá, S duchmi je vždy veselo, Václav Šuplata
19. r02vs1, Prečo pavúky pletú siete, Prečo..., Václav Šuplata

20. r02vs2, Prečo sú vrany čierne, Prečo..., Václav Šuplata
21. r02vs3, Prečo komáre štípu, Prečo..., Václav Šuplata
22. r02vs4, Prečo chodia slimáky pomaly, Prečo..., Václav Šuplata
23. r02vs5, Prečo majú datle na hlave červené perie, Prečo..., Václav Šuplata
24. r02vs6, Prečo majú ježkovia pichliače, Prečo..., Václav Šuplata
25. r02vs7, Prečo svätojánske mušky svietia, Prečo..., Václav Šuplata
26. r02vs8, Prečo majú labute dlhý krk, Prečo..., Václav Šuplata
27. r02vs9, Prečo včely robia med, Prečo..., Václav Šuplata
28. r02vs10, Prečo spia netopiere dolu hlavou, Prečo..., Václav Šuplata
29. r02vs11, Prečo sú ryby nemé, Prečo..., Václav Šuplata
30. r02vs12, Prečo majú lienky na chrbte bodky, Prečo..., Václav Šuplata
31. r02vs13, Prečo majú stonožky sto nôh, Prečo..., Václav Šuplata
32. r03vs1, 1, Nafukovacia rozprávka, Václav Šuplata
33. r03vs2, 2, Nafukovacia rozprávka, Václav Šuplata
34. r03vs3, 3, Nafukovacia rozprávka, Václav Šuplata
35. r03vs4, 4, Nafukovacia rozprávka, Václav Šuplata
36. r03vs5, 5, Nafukovacia rozprávka, Václav Šuplata
37. r03vs6, 6, Nafukovacia rozprávka, Václav Šuplata
38. r03vs7, 7, Nafukovacia rozprávka, Václav Šuplata
39. r03vs8, 8, Nafukovacia rozprávka, Václav Šuplata
40. r03vs9, 9, Nafukovacia rozprávka, Václav Šuplata

41. r04vs1, Ako sa kráľ Tukey Prvý zoznámil s bádateľom Cypriánom, TUKY PRVÝ kráľ tučniakov, Václav Šuplata
42. r04vs2, Ako kráľ Tukey Prvý dostal najkrajší darček k narodeninám, TUKY PRVÝ kráľ tučniakov, Václav Šuplata
43. r04vs3, Ako bol kráľ Tukey Prvý prvý a posledný krát v kine, TUKY PRVÝ kráľ tučniakov, Václav Šuplata
44. r04vs4, Ako sa kráľ Tukey Prvý stal hradným pánom, TUKY PRVÝ kráľ tučniakov, Václav Šuplata
45. r04vs5, Ako sa kráľ Tukey Prvý nestal cirkusovým umelcom, TUKY PRVÝ kráľ tučniakov, Václav Šuplata
46. r04vs6, Ako sa kráľ Tukey Prvý zmenil na hračku, TUKY PRVÝ kráľ tučniakov, Václav Šuplata
47. r04vs7, Ako chcel kráľ Tukey Prvý zachraňovať priateľov v ZOO, TUKY PRVÝ kráľ tučniakov, Václav Šuplata
48. r04vs8, Ako kráľ Tukey Prvý spieval v opere, TUKY PRVÝ kráľ tučniakov, Václav Šuplata
49. r04vs9, Ako sa kráľ Tukey Prvý vybral na huby, TUKY PRVÝ kráľ tučniakov, Václav Šuplata
50. r04vs10, Ako kráľ Tukey Prvý vymyslel morskú penu, TUKY PRVÝ kráľ tučniakov, Václav Šuplata
51. r05vs1, Moje kráľovstvo, Kráľovský detektív Sebastián, Václav Šuplata
52. r05vs2, Deň veľkej záhady, Kráľovský detektív Sebastián, Václav Šuplata
53. r05vs3, Keď bolo zajtra, Kráľovský detektív Sebastián, Václav Šuplata
54. r05vs4, Čakanie na detektíva, Kráľovský detektív Sebastián, Václav Šuplata
55. r05vs5, Prichádza detektív, Kráľovský detektív Sebastián, Václav Šuplata

56. r05vs6, Nenápadné sledovanie detektíva, Kráľovský detektív Sebastián, Václav Šuplata
57. r05vs7, Záhadný zlodej, Kráľovský detektív Sebastián, Václav Šuplata
58. r05vs8, Vypočúvanie plyšového medveda, Kráľovský detektív Sebastián, Václav Šuplata
59. r06vs1, 1, Päť zázračníkov, Václav Šuplata
60. r07vs1, Zjedená chalúpka, Ježibabovská detektívka, Václav Šuplata
61. r07vs2, Výborný nápad, Ježibabovská detektívka, Václav Šuplata
62. r07vs3, Veľká knihová lúpež, Ježibabovská detektívka, Václav Šuplata
63. r07vs4, Odhalené tajomstvo, Ježibabovská detektívka, Václav Šuplata
64. r08vs1, Namyslený kohút, Jullien a jeho noví kamaráti, Václav Šuplata
65. r08vs2, Smutný pes, Jullien a jeho noví kamaráti, Václav Šuplata
66. r08vs3, Lenivý králik, Jullien a jeho noví kamaráti, Václav Šuplata
67. r08vs4, Nespokojná hus, Jullien a jeho noví kamaráti, Václav Šuplata
68. r08vs5, Netrpezlivé kačiatko, Jullien a jeho noví kamaráti, Václav Šuplata
69. r08vs6, Čistotné prasiatko, Jullien a jeho noví kamaráti, Václav Šuplata
70. r08vs7, Urazené teliatko, Jullien a jeho noví kamaráti, Václav Šuplata
71. r08vs8, Krava kuchárka, Jullien a jeho noví kamaráti, Václav Šuplata
72. r08vs9, Závistlivý kôň, Jullien a jeho noví kamaráti, Václav Šuplata
73. r08vs10, Závodný kôň, Jullien a jeho noví kamaráti, Václav Šuplata
74. r08vs11, Zvedavá žirafa, Jullien a jeho noví kamaráti, Václav Šuplata
75. r08vs12, Veľké trápenie malého slona, Jullien a jeho noví kamaráti, Václav Šuplata

76. r08vs13, Nahnevaný tiger, Jullien a jeho noví kamaráti, Václav Šuplata
77. r08vs14, Lev - kráľ zvierat, Jullien a jeho noví kamaráti, Václav Šuplata
78. r09vs1, Maco Paco horolezcom, Maco Paco, Václav Šuplata
79. r09vs2, Maco Paco záhradníkom, Maco Paco, Václav Šuplata
80. r09vs3, Maco Paco pri vode, Maco Paco, Václav Šuplata
81. r10vs1, Diery v syre, Rozprávky zo šuplíka, Václav Šuplata
82. r10vs2, Myší kráľ sveta, Rozprávky zo šuplíka, Václav Šuplata
83. r10vs3, Najväčší z leňochov, Rozprávky zo šuplíka, Václav Šuplata
84. r10vs4, Namyslené vajce, Rozprávky zo šuplíka, Václav Šuplata
85. r10vs5, Šťastný vôl, Rozprávky zo šuplíka, Václav Šuplata
86. r10vs6, Zaľúbený bocian, Rozprávky zo šuplíka, Václav Šuplata
87. r10vs7, Žubrienky, Rozprávky zo šuplíka, Václav Šuplata
88. r10vs8, Ostrov nesplnených snov, Rozprávky zo šuplíka, Václav Šuplata
89. r10vs9, Cmúľanie palca, Rozprávky zo šuplíka, Václav Šuplata
90. r10vs10, Pravdivá rozprávka o Jari, Rozprávky zo šuplíka, Václav Šuplata
91. r01ha1, Snehová kráľovná, Pohádky a poviedky I, Hans Christian Andersen
92. r01ha2, Anjel, Pohádky a poviedky I, Hans Christian Andersen
93. r01ha3, Chrabrý cínový vojačik, Pohádky a poviedky I, Hans Christian Andersen
94. r01ha4, Kresadlo, Pohádky a poviedky I, Hans Christian Andersen
95. r01ha5, O hlúpom Jankovi, Pohádky a poviedky I, Hans Christian Andersen
96. r01ha6, Lietajúci kufor, Pohádky a poviedky I, Hans Christian Andersen
97. r01ha7, Jedlička, Pohádky a poviedky I, Hans Christian Andersen

98. r01ha8, Kovové prasiatko, Pohádky a povídky I, Hans Christian Andersen
99. r01ha9, Cisárov slávik, Pohádky a povídky I, Hans Christian Andersen
100. r01ha10, Ružový škriatok, Pohádky a povídky I, Hans Christian Andersen
101. r01ha11, Pod vrbou, Pohádky a povídky I, Hans Christian Andersen
102. r01ha12, O dieťati v hrobe, Pohádky a povídky I, Hans Christian Andersen
103. r01ha13, Divé labuti, Pohádky a povídky I, Hans Christian Andersen
104. r01ha14, Čo otec robí, vždy je dobré, Pohádky a povídky I, Hans Christian Andersen
105. r01ha15, O korenárovej nočnej čiapke, Pohádky a povídky I, Hans Christian Andersen
106. r01ha16, Dievčatko so zápalkami, Pohádky a povídky I, Hans Christian Andersen
107. r01ha17, O lístku s nebies, Pohádky a povídky I, Hans Christian Andersen
108. r02ha1, O dcére kráľa vodníka, Pohádky a povídky II, Hans Christian Andersen
109. r02ha2, O dievčati, ktoré na chlieb stúplo, Pohádky a povídky II, Hans Christian Andersen
110. r02ha3, Poviedka, Pohádky a povídky II, Hans Christian Andersen
111. r02ha4, Kameň mudrcov, Pohádky a povídky II, Hans Christian Andersen
112. r02ha5, Súdruh z ciest, Pohádky a povídky II, Hans Christian Andersen
113. r02ha6, Sviniar, Pohádky a povídky II, Hans Christian Andersen
114. r02ha7, Jakub a Krystýnka, Pohádky a povídky II, Hans Christian Andersen
115. r02ha8, Biskup börglumský a jeho rodina, Pohádky a povídky II, Hans Christian Andersen
116. r02ha9, Židovské dievča, Pohádky a povídky II, Hans Christian Andersen

117. r02ha10, Poviedka o roku, Pohádky a poviedky II, Hans Christian Andersen
118. r02ha11, Posledný sen starého dubu, Pohádky a poviedky II, Hans Christian Andersen
119. r02ha12, O škriatkovi a paničke, Pohádky a poviedky II, Hans Christian Andersen
120. r02ha13, Starý dom, Pohádky a poviedky II, Hans Christian Andersen
121. r02ha14, Kvety malej Idy, Pohádky a poviedky II, Hans Christian Andersen
122. r02ha15, Ruža s hrobu Homérovho, Pohádky a poviedky II, Hans Christian Andersen
123. r02ha16, O zlom kniežati, Pohádky a poviedky II, Hans Christian Andersen
124. r02ha17, Detská pletka, Pohádky a poviedky II, Hans Christian Andersen
125. r02ha18, Malý a veľký Kaloš, Pohádky a poviedky II, Hans Christian Andersen
126. r02ha19, Pastierka a kominár, Pohádky a poviedky II, Hans Christian Andersen
127. r02ha20, Všetko kam patrí, Pohádky a poviedky II, Hans Christian Andersen
128. r03ha1, Fľaškové hrdlo, Pohádky a poviedky III, Hans Christian Andersen
129. r03ha2, Pobratimstvo, Pohádky a poviedky III, Hans Christian Andersen
130. r03ha3, Čo v srdci skryté, nie je zabudnuté, Pohádky a poviedky III, Hans Christian Andersen
131. r03ha4, Zlato, Pohádky a poviedky III, Hans Christian Andersen
132. r03ha5, Krása, Pohádky a poviedky III, Hans Christian Andersen
133. r03ha6, O tetuške, Pohádky a poviedky III, Hans Christian Andersen
134. r03ha7, Hlásny Ole, Pohádky a poviedky III, Hans Christian Andersen
135. r03ha8, V deň stahovania, Pohádky a poviedky III, Hans Christian Andersen

136. r03ha9, Peter, Petrík a Petříček, Pohádky a poviedky III, Hans Christian Andersen
137. r03ha10, Majiteľ bábkového divadla, Pohádky a poviedky III, Hans Christian Andersen
138. r03ha11, Súsedia, Pohádky a poviedky III, Hans Christian Andersen
139. r03ha12, Rudé črievce, Pohádky a poviedky III, Hans Christian Andersen
140. r03ha13, Sedmikráska, Pohádky a poviedky III, Hans Christian Andersen
141. r03ha14, Slimák a ružička, Pohádky a poviedky III, Hans Christian Andersen
142. r03ha15, Pohanka, Pohádky a poviedky III, Hans Christian Andersen
143. r03ha16, Nemá kniha, Pohádky a poviedky III, Hans Christian Andersen
144. r03ha17, O starom náhrobku, Pohádky a poviedky III, Hans Christian Andersen
145. r03ha18, Kalamár a pero, Pohádky a poviedky III, Hans Christian Andersen
146. r03ha19, Motýľ, Pohádky a poviedky III, Hans Christian Andersen
147. r03ha20, Posledná perla, Pohádky a poviedky III, Hans Christian Andersen
148. r03ha21, O piatich v jednom, Pohádky a poviedky III, Hans Christian Andersen
149. r03ha22, Lan, Pohádky a poviedky III, Hans Christian Andersen
150. r03ha23, Dvaja kohúti, Pohádky a poviedky III, Hans Christian Andersen
151. r03ha24, Čajová konvička, Pohádky a poviedky III, Hans Christian Andersen
152. r03ha25, Na dvore kačiek, Pohádky a poviedky III, Hans Christian Andersen
153. r03ha26, Bolesť zubov, Pohádky a poviedky III, Hans Christian Andersen
154. r03ha27, Škriatok v skliepku, Pohádky a poviedky III, Hans Christian Andersen
155. r03ha28, Bazová matička, Pohádky a poviedky III, Hans Christian Andersen
156. r03ha29, V detskej izbe, Pohádky a poviedky III, Hans Christian Andersen

157. r03ha30, O sneženke, Pohádky a poviedky III, Hans Christian Andersen
158. r03ha31, Búrka štíty premiestňuje, Pohádky a poviedky III, Hans Christian Andersen
159. r03ha32, V deň posledný, Pohádky a poviedky III, Hans Christian Andersen
160. r03ha33, O najkrajšej ruži sveta, Pohádky a poviedky III, Hans Christian Andersen
161. r03ha34, Obrázok s hradieb, Pohádky a poviedky III, Hans Christian Andersen
162. r03ha35, V okne vo Vartone, Pohádky a poviedky III, Hans Christian Andersen
163. r03ha36, Zelenáčik, Pohádky a poviedky III, Hans Christian Andersen
164. r03ha37, Zvoniaca tíšina na potoku, Pohádky a poviedky III, Hans Christian Andersen
165. r01pd1, Zakliata hora, Prostonárodné slovenské povesti (Prvý zväzok), Pavol Dobšinský
166. r01pd2, Laktibrada, Prostonárodné slovenské povesti (Prvý zväzok), Pavol Dobšinský
167. r01pd3, Panna z rosy počatá a z deväť matiek splodená, Prostonárodné slovenské povesti (Prvý zväzok), Pavol Dobšinský
168. r01pd4, Matej veľký kráľ a Uliana veľká kráľovná, Prostonárodné slovenské povesti (Prvý zväzok), Pavol Dobšinský
169. r01pd5, Radúz a Ludmila, Prostonárodné slovenské povesti (Prvý zväzok), Pavol Dobšinský
170. r01pd6, Myšacia bundička, Prostonárodné slovenské povesti (Prvý zväzok), Pavol Dobšinský
171. r01pd7, Hadogašpar, Prostonárodné slovenské povesti (Prvý zväzok), Pavol Dobšinský

172. r01pd8, Dvanásť bratia a trinásť sestra, Prostonárodné slovenské povesti (Prvý zväzok), Pavol Dobšinský
173. r01pd9, Tri stromy, Prostonárodné slovenské povesti (Prvý zväzok), Pavol Dobšinský
174. r01pd10, Za zlatým jabĺčkom, Prostonárodné slovenské povesti (Prvý zväzok), Pavol Dobšinský
175. r01pd11, Zlatá podkova, zlatô pero, zlatý vlas, Prostonárodné slovenské povesti (Prvý zväzok), Pavol Dobšinský
176. r01pd12, Lomidrevo alebo Valibuk, Prostonárodné slovenské povesti (Prvý zväzok), Pavol Dobšinský
177. r01pd13, Tri perá draka alebo hľadanie zlého a dobrého, Prostonárodné slovenské povesti (Prvý zväzok), Pavol Dobšinský
178. r01pd14, Chorý kráľ, Prostonárodné slovenské povesti (Prvý zväzok), Pavol Dobšinský
179. r01pd15, Zlatovláska, Prostonárodné slovenské povesti (Prvý zväzok), Pavol Dobšinský
180. r01pd16, O kráľovičovi, čo si mladú hľadal, Prostonárodné slovenské povesti (Prvý zväzok), Pavol Dobšinský
181. r01pd17, Černokňažník, Prostonárodné slovenské povesti (Prvý zväzok), Pavol Dobšinský
182. r01pd18, Dvã šibalci, Prostonárodné slovenské povesti (Prvý zväzok), Pavol Dobšinský
183. r01pd19, Mahulienka zlatá panna alebo slncová panička, Prostonárodné slovenské povesti (Prvý zväzok), Pavol Dobšinský
184. r01pd20, Vlkolak, Prostonárodné slovenské povesti (Prvý zväzok), Pavol Dobšinský

185. r01pd21, Zlatovlasé dvojčatá, Prostonárodné slovenské povesti (Prvý zväzok), Pavol Dobšinský
186. r01pd22, Berona, Prostonárodné slovenské povesti (Prvý zväzok), Pavol Dobšinský
187. r01pd23, Jelenček, Prostonárodné slovenské povesti (Prvý zväzok), Pavol Dobšinský
188. r01pd24, Dalajláma, Prostonárodné slovenské povesti (Prvý zväzok), Pavol Dobšinský
189. r01pd25, Čarodejná lampa, Prostonárodné slovenské povesti (Prvý zväzok), Pavol Dobšinský
190. r01pd26, Pani Mačička, Prostonárodné slovenské povesti (Prvý zväzok), Pavol Dobšinský
191. r01pd27, Zlí bratia, Prostonárodné slovenské povesti (Prvý zväzok), Pavol Dobšinský
192. r01pd28, Víťaz Kuchta, Prostonárodné slovenské povesti (Prvý zväzok), Pavol Dobšinský
193. r01pd29, Či še hnevace?, Prostonárodné slovenské povesti (Prvý zväzok), Pavol Dobšinský
194. r01pd30, Svetovládny rytier, Prostonárodné slovenské povesti (Prvý zväzok), Pavol Dobšinský
195. r01pd31, Otcov hrob, Prostonárodné slovenské povesti (Prvý zväzok), Pavol Dobšinský
196. r01pd32, Vintalko, Prostonárodné slovenské povesti (Prvý zväzok), Pavol Dobšinský
197. r01pd33, Ludojedi, Prostonárodné slovenské povesti (Prvý zväzok), Pavol Dobšinský

198. r01pd34, Zlatá priadka, Prostonárodné slovenské povesti (Prvý zväzok), Pavol Dobšinský
199. r01pd35, Škrupinový zámok, Prostonárodné slovenské povesti (Prvý zväzok), Pavol Dobšinský
200. r01pd36, O kocúrikovi, Prostonárodné slovenské povesti (Prvý zväzok), Pavol Dobšinský
201. r02pd1, Janko a Macko, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšinský
202. r02pd2, Sirota, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšinský
203. r02pd3, Zlatná krajina, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšinský
204. r02pd4, Dlhý nos, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšinský
205. r02pd5, Zlatý kľúč, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšinský
206. r02pd6, Čarodejná kráľovná, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšinský
207. r02pd7, Zlatý zúbok, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšinský
208. r02pd8, O veternom kráľovi, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšinský
209. r02pd9, Šurienska a Atalienka, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšinský
210. r02pd10, Tri holúbky, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšinský

211. r02pd11, Neverná žena, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšinský
212. r02pd12, My trã brarã, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšinský
213. r02pd13, Had, mačička a psík, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšinský
214. r02pd14, Fundži palica, cepy von z vreca, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšinský
215. r02pd15, Tri píšťalky, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšinský
216. r02pd16, O Mechúrikovi, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšinský
217. r02pd17, Kozičky, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšinský
218. r02pd18, Mať a macocha alebo zlatá páva, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšinský
219. r02pd19, Rovnopekný pár, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšinský
220. r02pd20, Hrdá panička, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšinský
221. r02pd21, Kraľčík, Kucharčík, Popelčík, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšinský
222. r02pd22, Piatko a Pustaj, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšinský
223. r02pd23, Zlý brat, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšinský

224. r02pd24, Pijan, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšínský
225. r02pd25, Tri zlaté hrušky, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšínský
226. r02pd26, Bračok vtáčok, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšínský
227. r02pd27, Baláž, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšínský
228. r02pd28, Zlatovláska, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšínský
229. r02pd29, Mlynček, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšínský
230. r02pd30, Či jesto pravda na svete, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšínský
231. r02pd31, Janko Hraško, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšínský
232. r02pd32, Stará dievka a čert, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšínský
233. r02pd33, Červenkráľ a žltovláska, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšínský
234. r02pd34, Červenovlasý princ, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšínský
235. r02pd35, O krásnej Ibronke, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšínský
236. r02pd36, O sirote, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšínský

237. r02pd37, Divostvorný lovec, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšinský
238. r02pd38, Cesta k Slncu, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšinský
239. r02pd39, Mataj, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšinský
240. r02pd40, Lubka a Kovovlad, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšinský
241. r02pd41, O dvanástich mesiačikoch, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšinský
242. r02pd42, Plavčík a Vratko čo osúšajú slzy sveta, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšinský
243. r02pd43, O hlúpej žene, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšinský
244. r02pd44, Zlatí bratkovia, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšinský
245. r02pd45, Zakliaty zámok, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšinský
246. r02pd46, Drevená krava, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšinský
247. r02pd47, Čert slúži, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšinský
248. r02pd48, Žabina kmotra, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšinský
249. r02pd49, Hádač, strúhač, strelec, poprac, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšinský

250. r02pd50, Najmladší z dvanástich, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšinský
251. r02pd51, O troch grošoch, Prostonárodné slovenské povesti (Druhý zväzok), Pavol Dobšinský
252. r03pd1, Soľ nad zlato, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
253. r03pd2, Kubove príhody, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
254. r03pd3, Nebeská sláva, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
255. r03pd4, Chudobných rodičov syn, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
256. r03pd5, Na Boha s kyjom, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
257. r03pd6, Sitno, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
258. r03pd7, Jhopsa, hor sa - zem, otvor sa!, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
259. r03pd8, Ako šlo vajce na vandrovku, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
260. r03pd9, Tetka smrť, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
261. r03pd10, Svenči gazda, zvieracej reči vedomý, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
262. r03pd11, Stratený chlapec, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský

263. r03pd12, Peračina, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
264. r03pd13, Traja zhavranelí bratia, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
265. r03pd14, Prorok Rak, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
266. r03pd15, Tri práčky, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
267. r03pd16, Psota, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
268. r03pd17, Zabitá sestra, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
269. r03pd18, Panna za drakom, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
270. r03pd19, Ružová Anička, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
271. r03pd20, Petor a Otec nebeský v širokom údolí, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
272. r03pd21, Nebojsa, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
273. r03pd22, Klinko a Kompit kráľ, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
274. r03pd23, Bača a šarkan, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
275. r03pd24, Kráľ času, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský

276. r03pd25, Trojruža, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
277. r03pd26, O človeku, čo nikdy nehrešil, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
278. r03pd27, O pohanskom Kráľovi, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
279. r03pd28, Meveď a komár, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
280. r03pd29, Mechúrik - Koščúrik s kamarátmi, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
281. r03pd30, Drotári a ten špatný, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
282. r03pd31, Slncový kôň, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
283. r03pd32, O Víťazkovi, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
284. r03pd33, Smrť kmotra a zázračný lekár, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
285. r03pd34, Múdry chlapec, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
286. r03pd35, Sirôtky, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
287. r03pd36, Pamodaj šťastia, lavička, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
288. r03pd37, Jeleň zlatorohý, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský

289. r03pd38, Mŕtvy frajer, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
290. r03pd39, Popolvár najväčší na svete, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
291. r03pd40, ženský vtip, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
292. r03pd41, Dobrý strelec, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
293. r03pd42, Cigán čerta ošialil, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
294. r03pd43, Pes, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
295. r03pd44, O Peterkovi, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
296. r03pd45, Starý Bodrík a vlk, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
297. r03pd46, Starý človek a dvanásť oviec, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
298. r03pd47, Ďuro-truľo, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
299. r03pd48, Súdnej pám, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
300. r03pd49, Veštec, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
301. r03pd50, Cesta k Slncu a Mesiacu, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský

302. r03pd51, Popelvár - Hnusná tvár, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
303. r03pd52, Tri zakliate kniežatá, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
304. r03pd53, Tri citróny, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
305. r03pd54, O zaklietej sestre, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
306. r03pd55, Zelezník, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
307. r03pd56, Múdry Maťko a blázni, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
308. r03pd57, O vlku, čo si dal šiť boty, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
309. r03pd58, Svetská krása, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
310. r03pd59, Statočný valach, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
311. r03pd60, O ednom chudobnom mlinarevi, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
312. r03pd61, Pravda, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
313. r03pd62, Popeluša, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
314. r03pd63, Salaš na čertovici, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský

315. r03pd64, Pecko-sprostáčik, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
316. r03pd65, Koza odratá a jež, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
317. r03pd66, Kozliatka, Prostonárodné slovenské povesti (Tretí zväzok), Pavol Dobšinský
318. r01ot1, Ruženka a trpaslíci, Tri rozprávky, Olga Textorisová
319. r01ot2, Divotvorný kalich, Tri rozprávky, Olga Textorisová
320. r01ot3, Kráľevič Černohlav, Tri rozprávky, Olga Textorisová
321. r01rp1, Mária Morevna, Ruské národné povesti, Neznámy autor
322. r01rp2, Biela kačka, Ruské národné povesti, Neznámy autor
323. r01rp3, Sestrička Anička a braček Ivanuško, Ruské národné povesti, Neznámy autor
324. r01rp4, O Ivanovi-cárevičovi, vtákovi Ohnivákovi a sivom vlkovi, Ruské národné povesti, Neznámy autor
325. r01rp5, Krásna Vasilisa, Ruské národné povesti, Neznámy autor
326. r01rp6, Žabka-cárevna, Ruské národné povesti, Neznámy autor
327. r01rp7, Pierko Finista-Jasného Sokola, Ruské národné povesti, Neznámy autor
328. r01pb1, Pracháreň, Vrbovčanské povesti a báje, Pavel Beblavý
329. r01pb2, Nekrštenec, Vrbovčanské povesti a báje, Pavel Beblavý
330. r01pb3, Žalostivá, Vrbovčanské povesti a báje, Pavel Beblavý
331. r01pb4, Štinberk, Vrbovčanské povesti a báje, Pavel Beblavý
332. r01pb5, Žurkuv mlyn, Vrbovčanské povesti a báje, Pavel Beblavý
333. r01pb6, Ježuv mlyn, Vrbovčanské povesti a báje, Pavel Beblavý

334. r01pb7, Trafil majster na majstra, Vrbovčanské povesti a báje, Pavel Beblavý
335. r01pb8, Krajanek, Vrbovčanské povesti a báje, Pavel Beblavý
336. r01pb9, Kováčova žena, Vrbovčanské povesti a báje, Pavel Beblavý
337. r01pb10, Zlaté zrnká na Ostrom vrchu, Vrbovčanské povesti a báje, Pavel Beblavý
338. r01sm1, Dračia rozprávka, Príbehy na dobrú noc, kde dobro vždy víťazí, [3]
339. r01sm2, Rozprávka o drakovi Ferovi a jeho pahltom bratovi, Príbehy na dobrú noc, kde dobro vždy víťazí, [3]
340. r01sm3, Rozprávka o malej neposednej hviezdičke, Príbehy na dobrú noc, kde dobro vždy víťazí, [3]
341. r01sm4, Rozprávka o tom ako vznikla Coca Cola, Príbehy na dobrú noc, kde dobro vždy víťazí, [3]
342. r01sm5, Príbeh zachráneného lesa: O kráľovi stromov, ľudoch a lesnej víle, Príbehy na dobrú noc, kde dobro vždy víťazí, [3]
343. r01sm6, Príbeh zachráneného lesa: Ako Matej bojoval s lesným škriatkom a ľudskou hlúposťou, Príbehy na dobrú noc, kde dobro vždy víťazí, [3]
344. r01sm7, Rozprávka o Nezábudke, Príbehy na dobrú noc, kde dobro vždy víťazí, [3]
345. r01sm8, Rozprávka o počarbanej stene a tajomnom kráľovstve, Príbehy na dobrú noc, kde dobro vždy víťazí, [3]
346. r01sm9, Rozprávka o Linduške, Príbehy na dobrú noc, kde dobro vždy víťazí, [3]
347. r01sm10, Rozprávka o dievčatku v mydlovej bublinke, Príbehy na dobrú noc, kde dobro vždy víťazí, [3]
348. r01sm11, Rozprávka o vetre a vetrisku, Príbehy na dobrú noc, kde dobro vždy víťazí, [3]

349. r01sm12, Ďurimor a Zuzkin klobúčik, Príbehy na dobrú noc, kde dobro vždy víťazí, [3]
350. r01sm13, Odvážna Katarínka, Príbehy na dobrú noc, kde dobro vždy víťazí, [3]
351. r01sm14, Rozprávka vodníka Topolca, Príbehy na dobrú noc, kde dobro vždy víťazí, [3]
352. r01sm15, Vianočná rozprávka o Žofke a permoníkovi, Príbehy na dobrú noc, kde dobro vždy víťazí, [3]
353. r02sm1, Ako žabiak Čvachtoš naučil vílu Zvedavku plávať, Rozprávky zo Spevavého lesa, [3]
354. r02sm2, Bludička Cecília a čarovná studnička, Rozprávky zo Spevavého lesa, [3]
355. r02sm3, Chamtivý kominár a víla Zvedavka, Rozprávky zo Spevavého lesa, [3]
356. r03sm1, Ako malá myška opustila domček v automate na kávu, Krátke rozprávky na čítanie o malej myške, [3]
357. r03sm2, Ako malá myška spoznala nových priateľov, Krátke rozprávky na čítanie o malej myške, [3]
358. r03sm3, Ako malá myška hľadala pomoc pre mamu, Krátke rozprávky na čítanie o malej myške, [3]
359. r03sm4, Ako malá myška zvládla skúšku odvahy, Krátke rozprávky na čítanie o malej myške, [3]
360. r03sm5, Ako malá myška našla nový domov , Krátke rozprávky na čítanie o malej myške, [3]
361. r04sm1, Dúhový sen, Krátke rozprávky na dobrú noc: Lovci motýľov, [3]
362. r04sm2, Lovci motýľov, Krátke rozprávky na dobrú noc: Lovci motýľov, [3]
363. r04sm3, Ospalé kráľovstvo, Krátke rozprávky na dobrú noc: Lovci motýľov, [3]
364. r04sm4, Tam, kde končil svet, Krátke rozprávky na dobrú noc: Lovci motýľov, [3]

365. r05sm1, Zázračné okuliare, Príbehy pre deti - Moja šiši teta a ja, [3]
366. r06sm1, Fatra a starý pútnik, Slovenské povesti o Fatre, [3]
367. r06sm2, Prúnov dar, Slovenské povesti o Fatre, [3]
368. r06sm3, Krížna hora, Slovenské povesti o Fatre, [3]
369. r06sm4, Ako vznikol Rakytov, Slovenské povesti o Fatre, [3]
370. r06sm5, Perúnov hnev, Slovenské povesti o Fatre, [3]
371. r07sm1, Povest o Železnej studienke, Slovenské povesti od babky Elenky, [3]
372. r07sm2, Povest o Čertovej skale, Slovenské povesti od babky Elenky, [3]
373. r07sm3, Povest o perinovom škriatkovi, Slovenské povesti od babky Elenky, [3]
374. r01bp1, Lož, Klasy, K. Salva - Čebratský
375. r01bp2, Kto do teba kameňom, ty do neho chlebom, Tri rozprávočky pre dietky,
Z. Lojková
376. r01bp3, Zbojník a pocestný, Tisíc a jedna noc, Cyrill Galla
377. r01bp4, Divotvorný lovec, Prostonárodné Slovenské Povesti, Pavel Dobšinský
378. r01bp5, Pavúk, Bájky, Somolický
379. r01bp6, Hráč, Päťdesiat malých rozprávok, Strýčko Janko
380. r01bp7, Vlk a sedem kozliatok, Vlk a sedem kozliatok, -
381. r01bp8, Čarodejná kapsa, Čarodejná kapsa, -
382. r01bp9, Hostina vo Vranovciach, O vtáčkoch letáčkoch, Ivan Javor
383. r01bp10, Zápas, Rozprávky statočných detí, Ruža Jilemnická
384. r01bp11, Mamička bude mať dovolenku, Rozprávky statočných detí, Ruža Jilemnická
385. r01bp12, Africká rozprávka, Slnčné a iné rozprávky, Mária a Ján Dub

386. r01bp13, My traja bratia, Čo figeľ, to groš a iné rozprávky, Jozef Ludovít Holuby
387. r01bp14, Ako starú mater poklal býk, Rozprávky starej matere, Mária Jančová
388. r01bp15, O sladičovom koreni, Rozprávky z hôr, Ludo Ondrejov
389. r01bp16, O dievčatku, ktoré nemalo starkú, Rozprávky, Klára Jarunková
390. r01bp17, Ivanko, Rozprávky spod slnečníka, Zlatko Klátik
391. r01bp18, Malá rozprávka maličkého Malíčka o malom chlapcovi a malej farbičke, Rozprávky z červenej ponožky, Jaroslava Blažková
392. r01bp19, Vlnolapka, Nie je škola ako škola, Mária Ďuričková
393. r01bp20, O psíčkovi a mačičke, ako si umývali dlážku, Rozprávky o psíčkovi a mačičke, Josef Čapek
394. r01bp21, Ako zajko Nebojko vyplel žihľavu, O zajkovi Nebojkovi, -
395. r01bp22, Čo Lubke v Afrike závideli, Rozprávky štrbavej opičky, Jana Šrámková
396. r01bp23, O princeznej sýkorke, Rozprávky štrbavej opičky, Jana Šrámková
397. r01bp24, V najväčšej horúčave, Deduško Večerníček, Jozef Pavlovič
398. r01bp25, Láska k práci, Laskavé rozprávke, Branislav Jobus
399. r01bp26, Domček pre nosáčika, Rozprávky z trávy, Silvia Havelková
400. r01bp27, Najväčší poklad, Rozprávky utkané zo snov, Janka Valentovičová
401. r01bp28, O namyslenom kopci, O upratovacom robotovi a iné čudné rozprávky, Silvia Havelková

Zoznam ekonomických článkov (Kód, Názov článku, Zdroj):

1. ne0003, 10 rokov eura na Slovensku: Prinieslo viaceré pozitíva, hodnotí naša
Národná banka, Topky
2. ne0005, Euro čaká test života a smrti, SME
3. ne0017, Desatročie s eurom, Pravda
4. ne0026, Desť rokov eura u nás: Fakty nepustia, toto nám vzala spoločná mena!,
Pluska
5. ne0078, 10 rokov s eurom, Týždeň

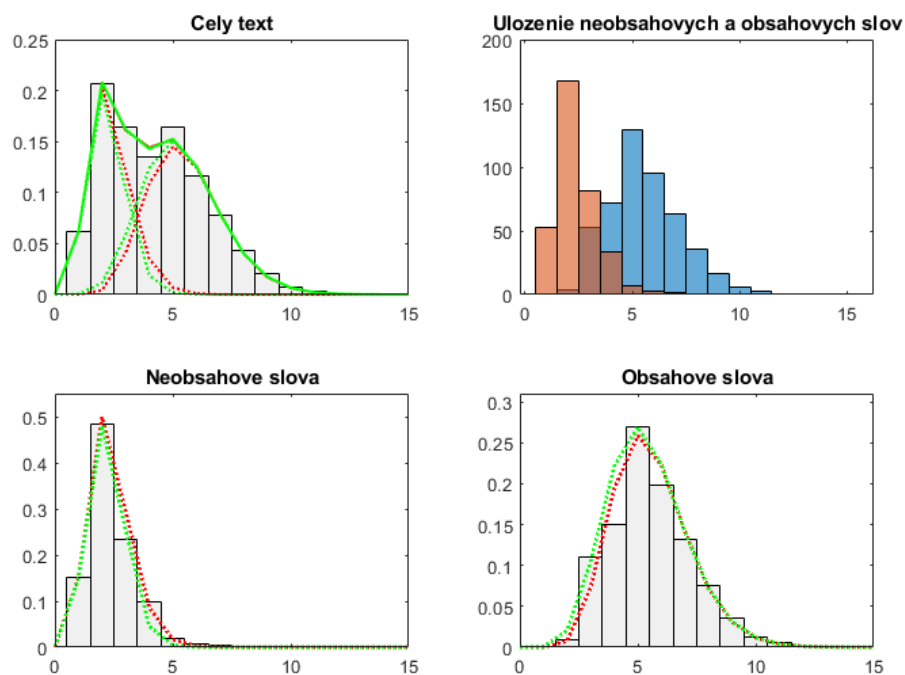
Zoznam receptov (Kód, Názov receptu, Zdroj):

1. s1, Makaróny so šalviovými čipsami, www.bonvivani.sk
2. s2, Cuketové špagety s rajčinovou omáčkou a orieškovo semienkovými gulôčkami, www.bonvivani.sk
3. s3, Špagety Carbonara, www.bonvivani.sk
4. s4, Špagety s tuniakom a zeleninou, www.bonvivani.sk
5. s5, Špagety s hubami hráškom a parmezánom , www.bonvivani.sk
6. s6, Špagety s údeným lososom cuketou cherry rajčinami a citrónovým tymiánom, www.bonvivani.sk
7. s7, Bolonské špagety , www.bonvivani.sk
8. s8, Špagety so špenátom a kuracím mäsom, www.bonvivani.sk
9. s9, Špagety so sušenými paradajkami, www.bonvivani.sk
10. s10, Špagety so šampiňónmi, www.bonvivani.sk
11. s11, Špagety s telacím mäsom a kuriatkami v smotanovej omáčke, www.bonvivani.sk
12. t1, Malinová torta, www.bonvivani.sk
13. t2, Nepečená kokosová torta, www.bonvivani.sk
14. t3, Višňová torta, www.bonvivani.sk
15. t4, Orechová torta bez múky, www.bonvivani.sk
16. t5, Čínska torta, www.bonvivani.sk
17. t6, Orechová torta s datlovo-figovou marmeládou a jogurtovou polevou, www.bonvivani.sk
18. p1, Paradajková polievka, www.bonvivani.sk
19. p2, Biela fazuľová polievka, www.bonvivani.sk
20. p3, Gulášová polievka, www.bonvivani.sk

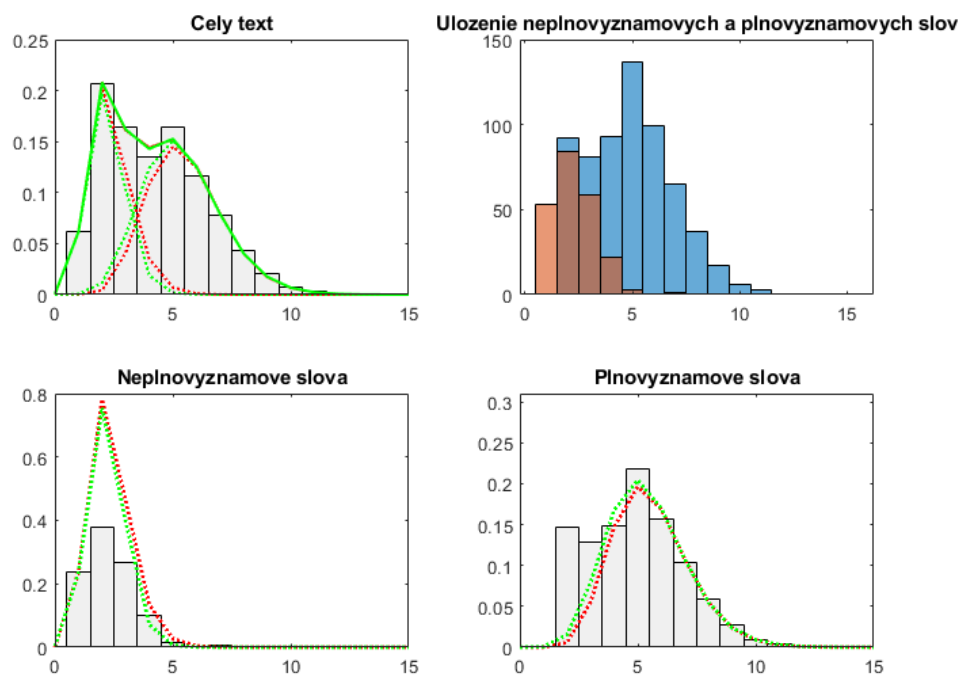
21. p4, Cesnačka - spravovačka, www.bonvivani.sk
22. p5, Brokolicový krém, www.bonvivani.sk
23. p6, Krémová polievka z batátov, www.bonvivani.sk
24. p7, Hovädzí vývar s pertžlenovými haluškami, www.bonvivani.sk
25. p8, Slepacia polievka so slížami, www.bonvivani.sk
26. p9, Hrášková polievka s karotkou a mrveničkou, www.bonvivani.sk
27. p10, Hrachová polievka, www.bonvivani.sk
28. p11, Frankfurtská polievka, www.bonvivani.sk
29. sa1, Cícerový šalát s restovaným lososom, www.kuchynalidla.sk
30. sa2, Šalát na grécky spôsob so syrom feta, www.kuchynalidla.sk
31. sa3, Lososový šalát, www.kuchynalidla.sk
32. sa4, Šalát s bulgurom, www.kuchynalidla.sk
33. sa5, Avokádový šalát, www.kuchynalidla.sk
34. sa6, Šalát „teriyaki“, www.kuchynalidla.sk
35. sa7, Kuskusový šalát s ananásom a lososom, www.kuchynalidla.sk

r01zb1	C	Cdg	CH	CHdg	pocet slov celok	C	Cdg	CH	CHdg
celok	0,0039	0,0033	0,0004	0,0003	852,0000				
neob	0,1134	3,3749	0,0007	0,0006	ob_pr_sl_op_ine	0,0500	0,0205	0,0013	0,0012
neob_pr	0,4185	27,9606	0,0007	0,0010	ob_sl_op_ine	0,0412	0,0316	0,0014	0,0017
neob_sl	0,1119	3,3648	0,0007	0,0006	ob_pr_op_ine	0,0475	0,0200	0,0012	0,0011
neob_op	0,1139	3,3718	0,0007	0,0006	ob_pr_sl_ine	0,0507	0,0215	0,0013	0,0012
neob_ine	0,1293	3,5793	0,0006	0,0006	ob_pr_sl_op	0,0418	0,0218	0,0011	0,0012
neob_pr_sl	0,4163	27,8837	0,0007	0,0010	ob_op_ine	0,0388	0,0312	0,0013	0,0017
neob_pr_op	0,4195	27,8927	0,0008	0,0011	ob_sl_ine	0,0421	0,0329	0,0014	0,0018
neob_pr_ine	0,4332	27,7095	0,0007	0,0011	ob_sl_op	0,0369	0,0350	0,0013	0,0018
neob_sl_op	0,1124	3,3617	0,0007	0,0006	ob_pr_ine	0,0482	0,0209	0,0012	0,0012
neob_sl_ine	0,1279	3,5690	0,0006	0,0006	ob_pr_op	0,0395	0,0213	0,0011	0,0012
neob_op_ine	0,1300	3,5761	0,0006	0,0007	ob_pr_sl	0,0425	0,0228	0,0012	0,0013
neob_pr_sl_op	0,4173	27,8162	0,0007	0,0011	ob_ine	0,0398	0,0325	0,0014	0,0017
neob_pr_sl_ine	0,4311	27,6348	0,0006	0,0011	ob_op	0,0347	0,0348	0,0013	0,0018
neob_pr_op_ine	0,4343	27,6439	0,0007	0,0012	ob_pr	0,0402	0,0224	0,0011	0,0013
neob_sl_op_ine	0,1286	3,5658	0,0006	0,0007	ob_sl	0,0379	0,0364	0,0014	0,0019
neob_pr_sl_op_ine	0,4322	27,5696	0,0007	0,0012	ob	0,0357	0,0362	0,0013	0,0019

Tabulka 5: r01zb1



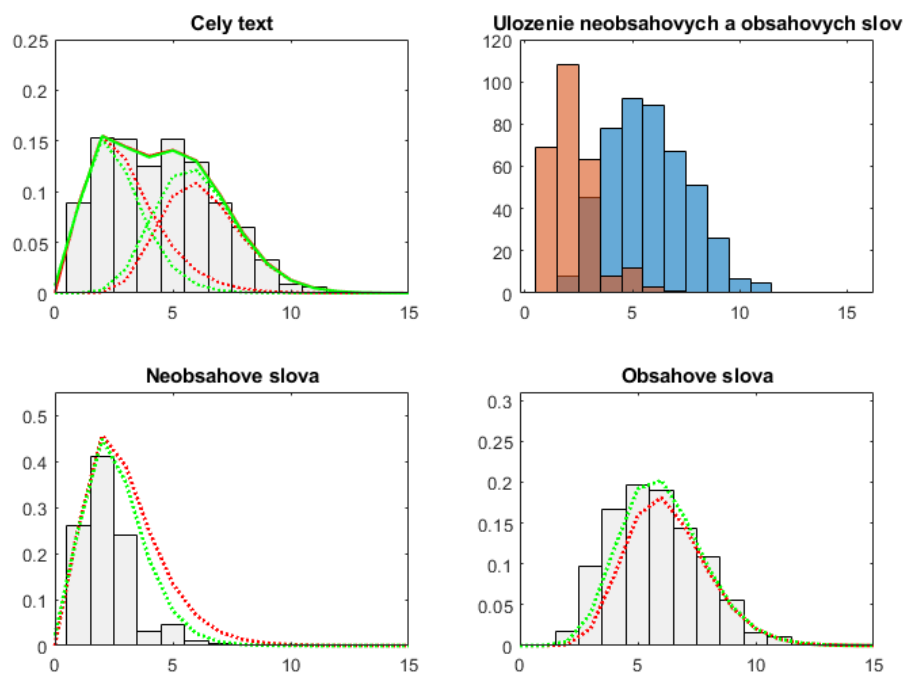
Obr. 14: r01zb1



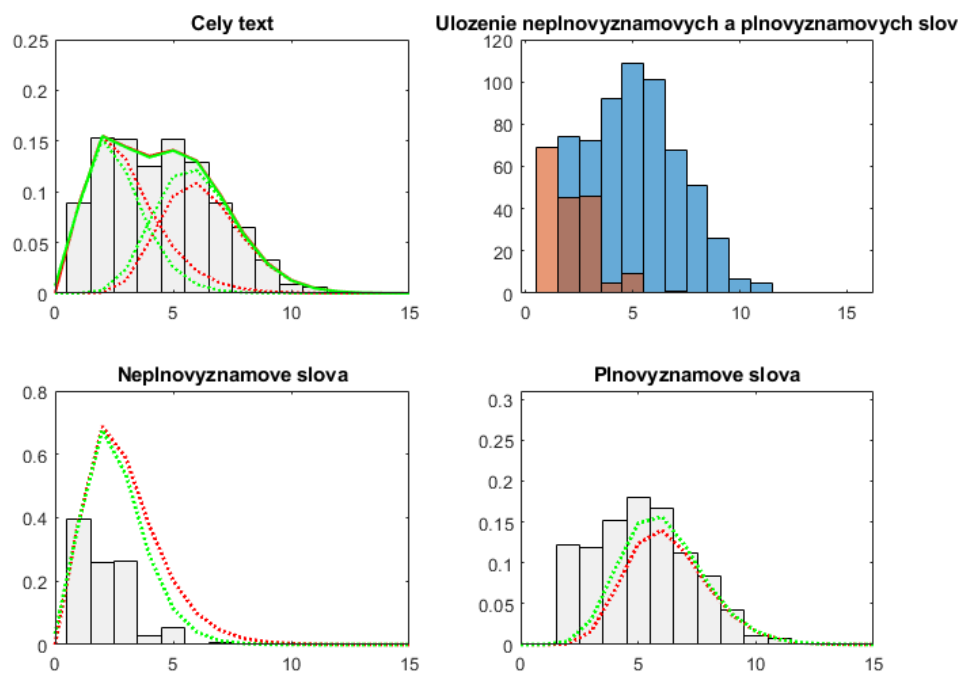
Obr. 15: r01zb1

r01zb2	C	Cdg	CH	CHdg	pocet slov celok	C	Cdg	CH	CHdg
celok	0,0056	0,0058	0,0004	0,0004	780				
neob	0,4008	0,1973	0,0096	0,0046	ob_pr_sl_op_ine	0,9358	0,2551	0,0095	0,0048
neob_pr	0,2446	0,1186	0,0070	0,0033	ob_sl_op_ine	0,6777	0,1828	0,0068	0,0032
neob_sl	0,3884	0,1871	0,0093	0,0044	ob_pr_op_ine	0,9169	0,2461	0,0092	0,0045
neob_op	0,3821	0,1836	0,0092	0,0044	ob_pr_sl_ine	0,9170	0,2459	0,0091	0,0045
neob_ine	0,3289	0,1517	0,0083	0,0038	ob_pr_sl_op	0,8698	0,2235	0,0083	0,0039
neob_pr_sl	0,2354	0,1111	0,0068	0,0031	ob_op_ine	0,6588	0,1744	0,0065	0,0030
neob_pr_op	0,2325	0,1123	0,0067	0,0031	ob_sl_ine	0,6591	0,1746	0,0065	0,0030
neob_pr_ine	0,1986	0,1071	0,0060	0,0027	ob_sl_op	0,6113	0,1539	0,0057	0,0026
neob_sl_op	0,3701	0,1740	0,0089	0,0042	ob_pr_ine	0,8984	0,2372	0,0088	0,0043
neob_sl_ine	0,3181	0,1431	0,0080	0,0036	ob_pr_op	0,8521	0,2153	0,0080	0,0037
neob_op_ine	0,3126	0,1403	0,0080	0,0036	ob_pr_sl	0,8522	0,2151	0,0079	0,0037
neob_pr_sl_op	0,2237	0,1052	0,0064	0,0029	ob_ine	0,6406	0,1663	0,0062	0,0028
neob_pr_sl_ine	0,1906	0,1007	0,0057	0,0025	ob_op	0,5937	0,1463	0,0055	0,0024
neob_pr_op_ine	0,1882	0,1022	0,0057	0,0026	ob_pr	0,8348	0,2071	0,0076	0,0035
neob_sl_op_ine	0,3022	0,1320	0,0077	0,0034	ob_sl	0,5940	0,1465	0,0054	0,0024
neob_pr_sl_op_ine	0,1805	0,0961	0,0054	0,0024	ob	0,5768	0,1390	0,0052	0,0023

Tabulka 6: r01zb2



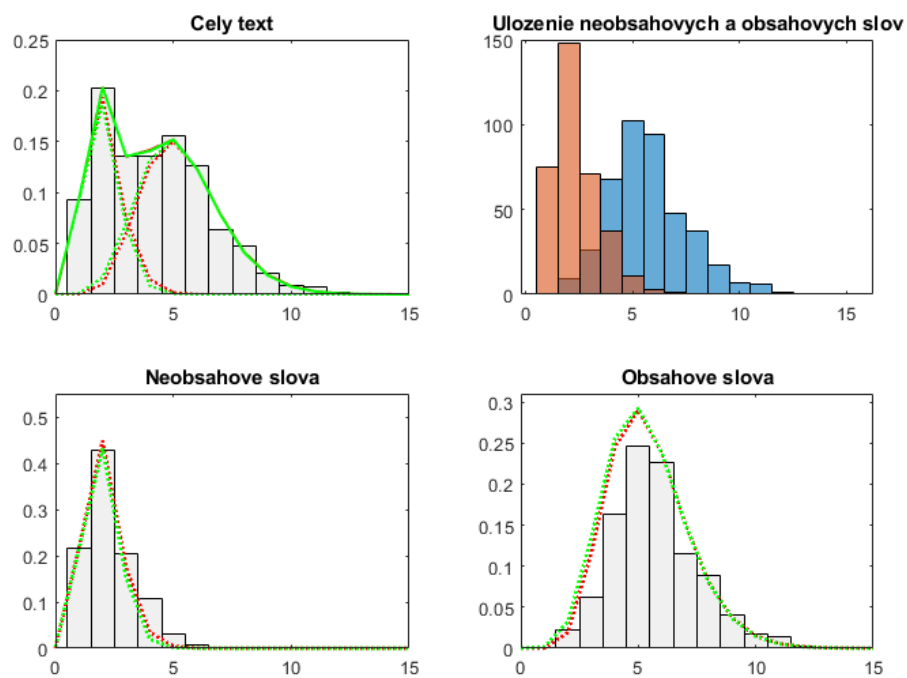
Obr. 16: r01zb2



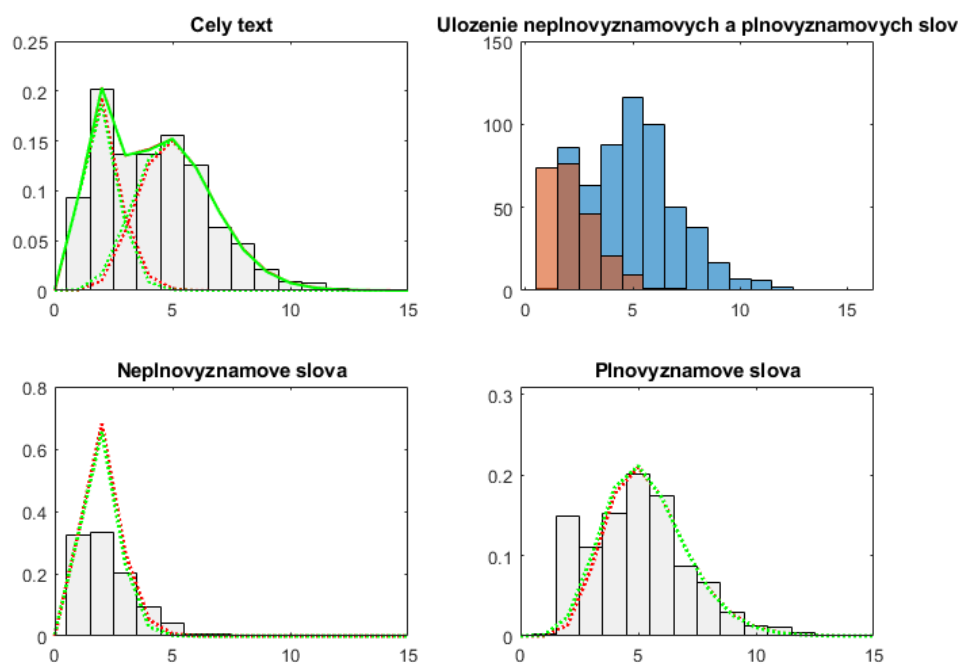
Obr. 17: r01zb2

r01zb3	C	Cdg	CH	CHdg	pocet slov celok	C	Cdg	CH	CHdg
celok	0,0134	0,0149	0,0003	0,0003	802				
neob	0,6375	6,1342	0,0014	0,0021	ob_pr_sl_op_ine	0,0586	0,0661	0,0020	0,0027
neob_pr	108469,8509	925624671,4682	0,0022	0,0033	ob_sl_op_ine	0,0639	0,0882	0,0028	0,0038
neob_sl	0,6354	6,0880	0,0015	0,0023	ob_pr_op_ine	0,0622	0,0712	0,0021	0,0029
neob_op	0,6818	6,2588	0,0015	0,0023	ob_pr_sl_ine	0,0604	0,0681	0,0021	0,0029
neob_ine	0,8962	8,1183	0,0017	0,0025	ob_pr_sl_op	0,0655	0,0751	0,0022	0,0031
neob_pr_sl	107608,9836	918278443,9244	0,0024	0,0035	ob_op_ine	0,0693	0,0950	0,0030	0,0041
neob_pr_op	107894,4957	920714196,0299	0,0024	0,0034	ob_sl_ine	0,0664	0,0910	0,0030	0,0040
neob_pr_ine	106482,5596	908662965,7404	0,0026	0,0037	ob_sl_op	0,0737	0,1002	0,0031	0,0043
neob_sl_op	0,6794	6,2118	0,0016	0,0025	ob_pr_ine	0,0640	0,0732	0,0022	0,0030
neob_sl_ine	0,8932	8,0575	0,0018	0,0027	ob_pr_op	0,0703	0,0813	0,0024	0,0033
neob_op_ine	0,9387	8,2301	0,0018	0,0026	ob_pr_sl	0,0674	0,0773	0,0024	0,0032
neob_pr_sl_op	107042,7015	913445399,7530	0,0025	0,0037	ob_ine	0,0718	0,0979	0,0031	0,0042
neob_pr_sl_ine	105652,8309	901582475,1069	0,0028	0,0040	ob_op	0,0804	0,1083	0,0033	0,0045
neob_pr_op_ine	105928,0436	903930346,3954	0,0027	0,0039	ob_pr	0,0722	0,0835	0,0025	0,0035
neob_sl_op_ine	0,9353	8,1687	0,0019	0,0029	ob_sl	0,0763	0,1031	0,0033	0,0044
neob_pr_sl_op_ine	105106,9014	896923134,4164	0,0029	0,0042	ob	0,0831	0,1113	0,0035	0,0047

Tabulka 7: r01zb3



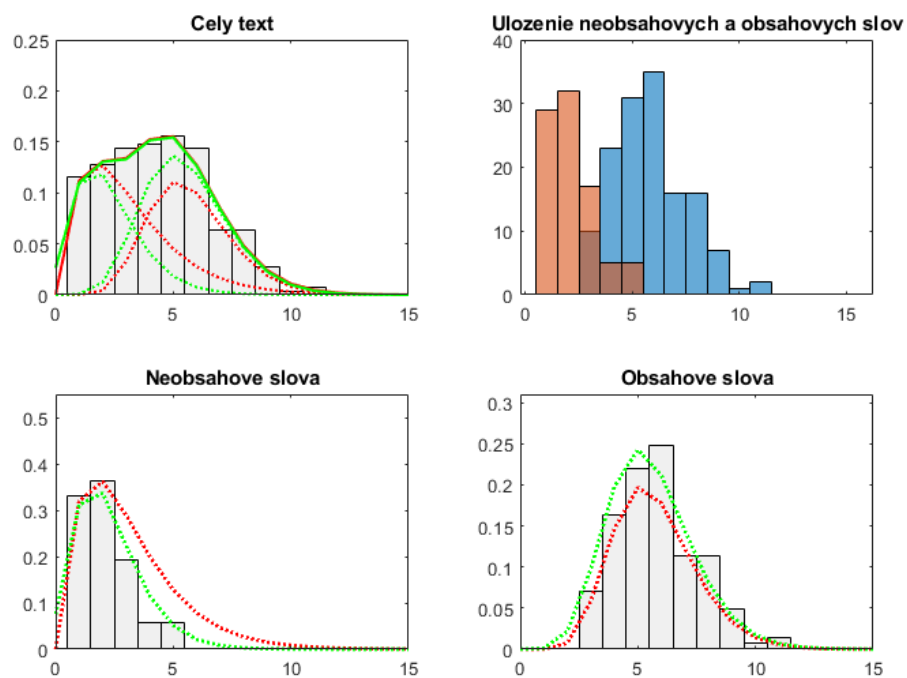
Obr. 18: r01zb3



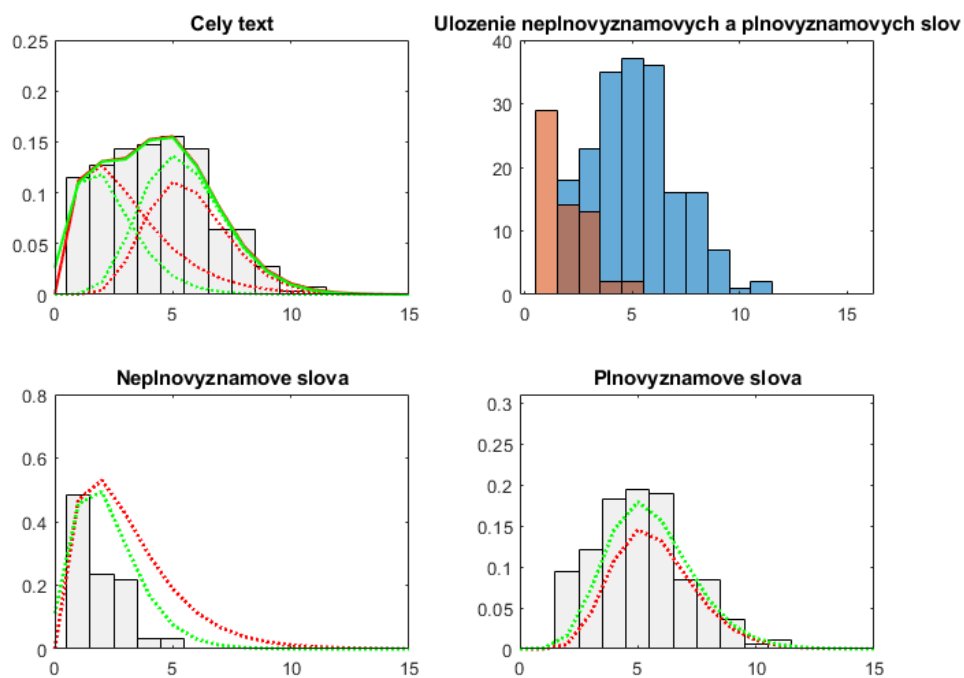
Obr. 19: r01zb3

r01zb4	C	Cdg	CH	CHdg	pocet slov celok	C	Cdg	CH	CHdg
celok	0,0209	0,0237	0,0012	0,0012	251				
neob	0,3526	0,0701	0,0054	0,0007	ob_pr_sl_op_ine	0,2175	0,0745	0,0072	0,0022
neob_pr	0,2283	0,0426	0,0030	0,0003	ob_sl_op_ine	0,1767	0,0727	0,0050	0,0019
neob_sl	0,3526	0,0701	0,0054	0,0007	ob_pr_op_ine	0,2175	0,0745	0,0072	0,0022
neob_op	0,2390	0,0357	0,0034	0,0004	ob_pr_sl_ine	0,1436	0,0609	0,0047	0,0014
neob_ine	0,3084	0,0681	0,0046	0,0006	ob_pr_sl_op	0,1777	0,0665	0,0062	0,0018
neob_pr_sl	0,2283	0,0426	0,0030	0,0003	ob_op_ine	0,1767	0,0727	0,0050	0,0019
neob_pr_op	0,1553	0,0486	0,0015	0,0007	ob_sl_ine	0,1139	0,0705	0,0031	0,0017
neob_pr_ine	0,1970	0,0493	0,0022	0,0004	ob_sl_op	0,1388	0,0674	0,0040	0,0017
neob_sl_op	0,2390	0,0357	0,0034	0,0004	ob_pr_ine	0,1436	0,0609	0,0047	0,0014
neob_sl_ine	0,3084	0,0681	0,0046	0,0006	ob_pr_op	0,1777	0,0665	0,0062	0,0018
neob_op_ine	0,2167	0,0548	0,0030	0,0009	ob_pr_sl	0,1244	0,0669	0,0042	0,0015
neob_pr_sl_op	0,1553	0,0486	0,0015	0,0007	ob_ine	0,1139	0,0705	0,0031	0,0017
neob_pr_sl_ine	0,1970	0,0493	0,0022	0,0004	ob_op	0,1388	0,0674	0,0040	0,0017
neob_pr_op_ine	0,1416	0,0725	0,0013	0,0012	ob_pr	0,1244	0,0669	0,0042	0,0015
neob_sl_op_ine	0,2167	0,0548	0,0030	0,0009	ob_sl	0,0983	0,0806	0,0027	0,0020
neob_pr_sl_op_ine	0,1416	0,0725	0,0013	0,0012	ob	0,0983	0,0806	0,0027	0,0020

Tabuľka 8: r01zb4



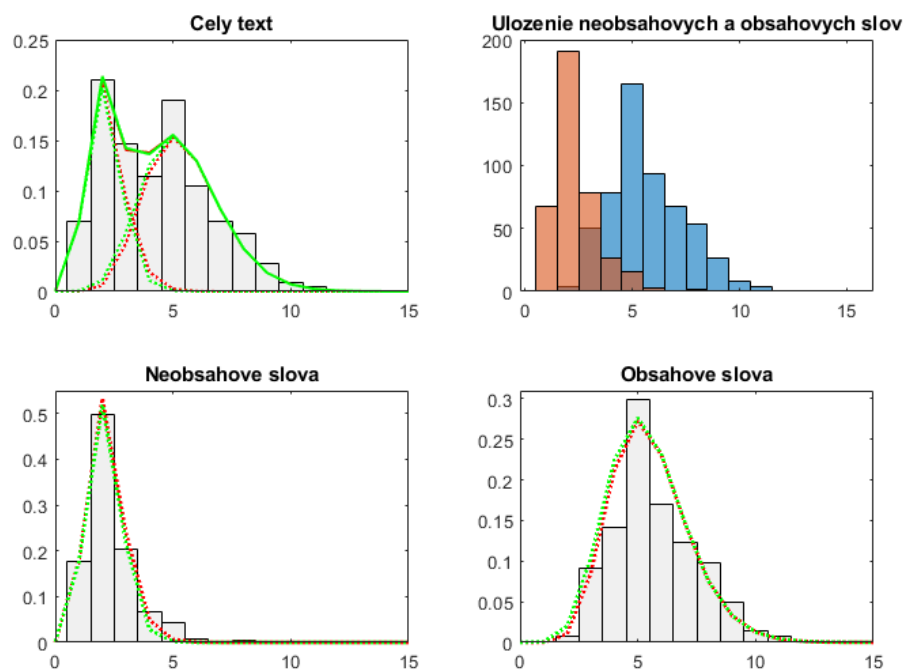
Obr. 20: r01zb4



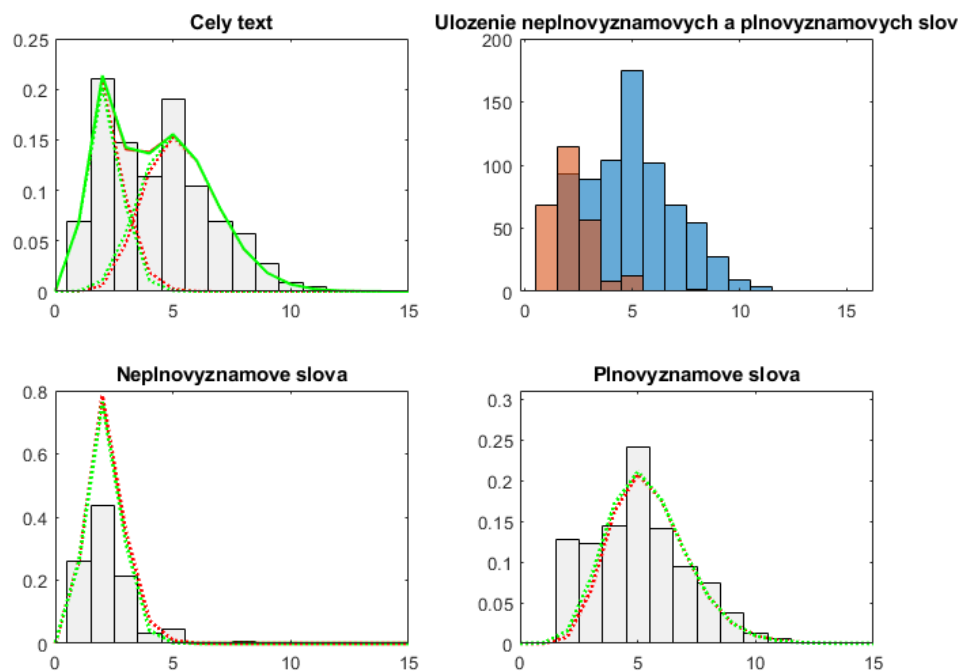
Obr. 21: r01zb4

r01zb5	C	Cdg	CH	CHdg	pocet slov celok	C	Cdg	CH	CHdg
celok	0,0302	0,0298	0,0029	0,0028	986				
neob	6,8286	810,9512	0,0006	0,0006	ob_pr_sl_op_ine	0,0934	0,0600	0,0031	0,0032
neob_pr	6,5272	760,5221	0,0005	0,0009	ob_sl_op_ine	0,0638	0,0655	0,0032	0,0036
neob_sl	6,7946	806,7568	0,0006	0,0006	ob_pr_op_ine	0,0932	0,0607	0,0032	0,0032
neob_op	401,1200	864749,9035	0,0005	0,0006	ob_pr_sl_ine	0,0858	0,0583	0,0029	0,0031
neob_ine	7,1204	791,0276	0,0007	0,0007	ob_pr_sl_op	0,0767	0,0592	0,0032	0,0033
neob_pr_sl	6,4990	756,8409	0,0006	0,0009	ob_op_ine	0,0643	0,0670	0,0033	0,0037
neob_pr_op	376,4393	811218,9386	0,0005	0,0010	ob_sl_ine	0,0602	0,0679	0,0032	0,0037
neob_pr_ine	6,8997	744,0950	0,0007	0,0011	ob_sl_op	0,0661	0,0755	0,0034	0,0038
neob_sl_op	399,0959	860382,4875	0,0005	0,0006	ob_pr_ine	0,0861	0,0595	0,0030	0,0032
neob_sl_ine	7,0859	787,0598	0,0007	0,0008	ob_pr_op	0,0764	0,0599	0,0032	0,0033
neob_op_ine	389,8279	839196,1359	0,0006	0,0007	ob_pr_sl	0,0687	0,0576	0,0030	0,0032
neob_pr_sl_op	374,6591	807374,3099	0,0006	0,0011	ob_ine	0,0613	0,0698	0,0032	0,0039
neob_pr_sl_ine	6,8705	740,5954	0,0007	0,0011	ob_op	0,0667	0,0770	0,0034	0,0039
neob_pr_op_ine	366,6003	788690,9161	0,0007	0,0012	ob_pr	0,0690	0,0588	0,0030	0,0033
neob_sl_op_ine	387,9187	835082,4369	0,0006	0,0008	ob_sl	0,0625	0,0782	0,0033	0,0039
neob_pr_sl_op_ine	364,9146	785056,4063	0,0007	0,0013	ob	0,0637	0,0802	0,0034	0,0040

Tabulka 9: r01zb5



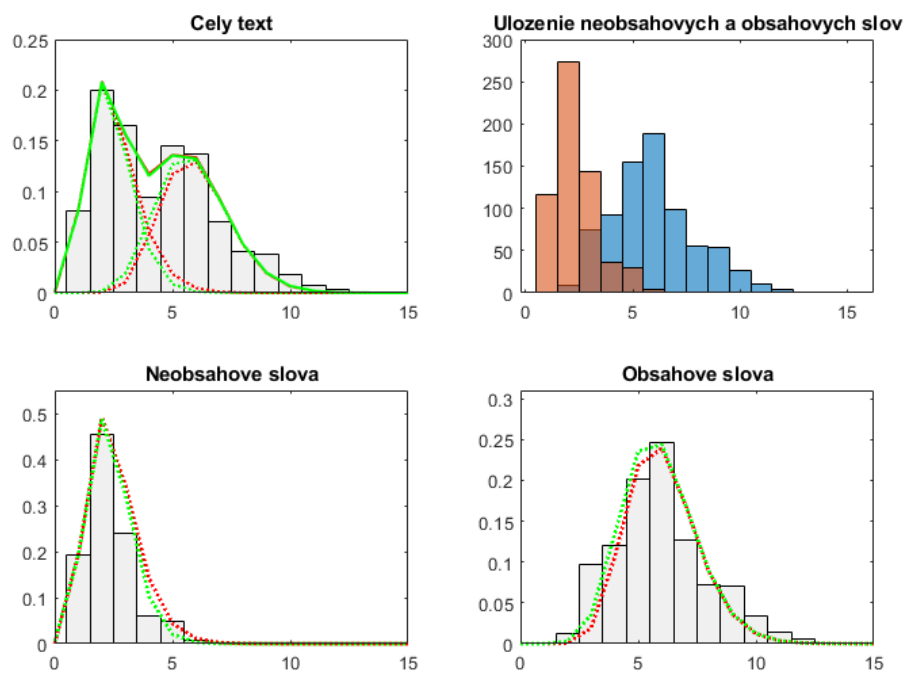
Obr. 22: r01zb5



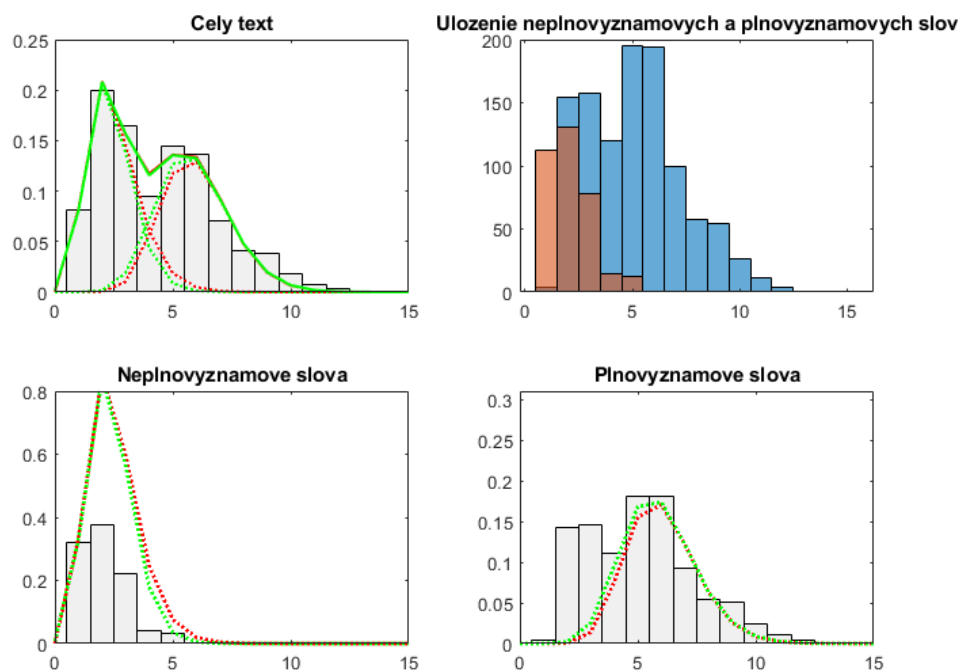
Obr. 23: r01zb5

r01zb6	C	Cdg	CH	CHdg	pocet slov celok	C	Cdg	CH	CHdg
celok	0,0817	0,0745	0,0018	0,0017	1426				
neob	0,0868	0,0900	0,0035	0,0020	ob_pr_sl_op_ine	0,9440	0,3482	0,0042	0,0031
neob_pr	0,1106	0,5077	0,0033	0,0022	ob_sl_op_ine	0,8791	0,3389	0,0040	0,0032
neob_sl	0,0804	0,0886	0,0031	0,0018	ob_pr_op_ine	0,8779	0,3181	0,0038	0,0028
neob_op	0,0814	0,0895	0,0032	0,0018	ob_pr_sl_ine	0,8887	0,3230	0,0039	0,0028
neob_ine	0,0761	0,1074	0,0032	0,0019	ob_pr_sl_op	0,8595	0,3273	0,0039	0,0030
neob_pr_sl	0,1067	0,5066	0,0030	0,0020	ob_op_ine	0,8133	0,3095	0,0036	0,0029
neob_pr_op	0,1077	0,5081	0,0030	0,0020	ob_sl_ine	0,8241	0,3143	0,0036	0,0030
neob_pr_ine	0,1070	0,5308	0,0030	0,0021	ob_sl_op	0,7992	0,3201	0,0037	0,0031
neob_sl_op	0,0757	0,0889	0,0029	0,0016	ob_pr_ine	0,8267	0,2952	0,0034	0,0025
neob_sl_ine	0,0706	0,1069	0,0028	0,0016	ob_pr_op	0,7947	0,2983	0,0035	0,0026
neob_op_ine	0,0716	0,1078	0,0029	0,0016	ob_pr_sl	0,8053	0,3030	0,0036	0,0027
neob_pr_sl_op	0,1044	0,5077	0,0028	0,0018	ob_ine	0,7626	0,2873	0,0032	0,0027
neob_pr_sl_ine	0,1037	0,5305	0,0027	0,0019	ob_op	0,7349	0,2919	0,0033	0,0028
neob_pr_op_ine	0,1046	0,5320	0,0028	0,0019	ob_pr	0,7448	0,2763	0,0032	0,0024
neob_sl_op_ine	0,0668	0,1080	0,0026	0,0014	ob_sl	0,7454	0,2965	0,0034	0,0029
neob_pr_sl_op_ine	0,1020	0,5324	0,0025	0,0018	ob	0,6855	0,2707	0,0030	0,0026

Tabuľka 10: r01zb6



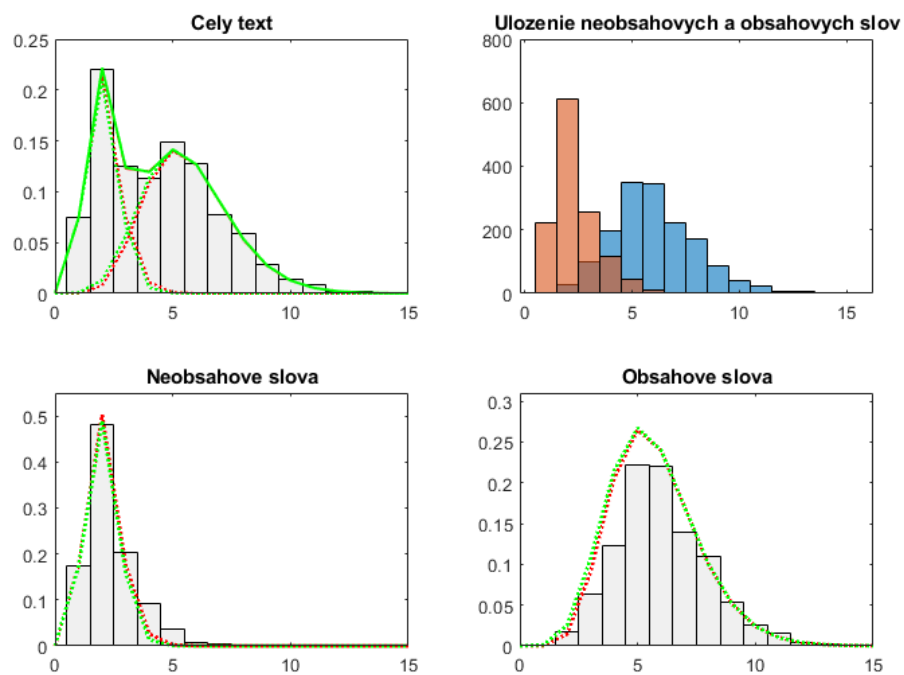
Obr. 24: r01zb6



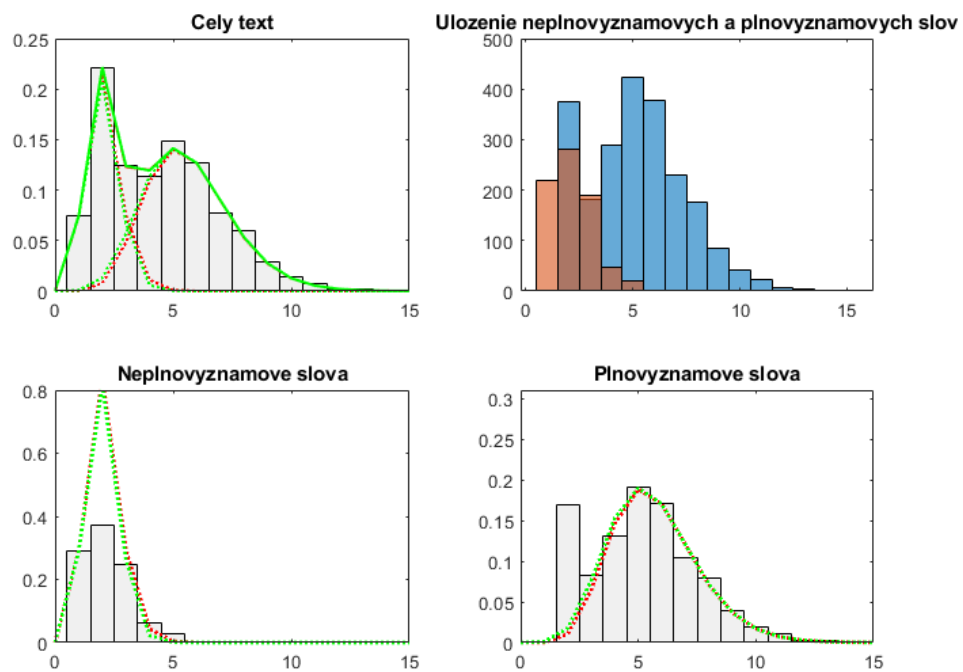
Obr. 25: r01zb6

01zb7	C	Cdg	CH	CHdg	pocet slov celok	C	Cdg	CH	CHdg
celok	0,0099	0,0101	0,0003	0,0003	2980				
neob	37,3008	49791,0686	0,0012	0,0017	ob_pr_sl_op_ine	0,0551	0,0547	0,0017	0,0022
neob_pr	38,3780	46837,9410	0,0020	0,0027	ob_sl_op_ine	0,0565	0,0749	0,0024	0,0032
neob_sl	37,8479	49711,7524	0,0012	0,0018	ob_pr_op_ine	0,0557	0,0555	0,0017	0,0023
neob_op	37,9597	49597,6510	0,0012	0,0018	ob_pr_sl_ine	0,0559	0,0557	0,0017	0,0023
neob_ine	59,4479	54146,2099	0,0016	0,0022	ob_pr_sl_op	0,0617	0,0623	0,0020	0,0026
neob_pr_sl	39,2892	46793,9810	0,0021	0,0028	ob_op_ine	0,0573	0,0758	0,0025	0,0032
neob_pr_op	39,4494	46694,1282	0,0021	0,0028	ob_sl_ine	0,0577	0,0762	0,0025	0,0032
neob_pr_ine	59,6827	51025,7430	0,0025	0,0033	ob_sl_op	0,0656	0,0854	0,0029	0,0037
neob_sl_op	38,6396	49527,5930	0,0013	0,0018	ob_pr_ine	0,0565	0,0565	0,0017	0,0023
neob_sl_ine	59,9329	54059,5091	0,0016	0,0022	ob_pr_op	0,0624	0,0631	0,0020	0,0026
neob_op_ine	60,2311	53942,8431	0,0016	0,0022	ob_pr_sl	0,0627	0,0635	0,0020	0,0027
neob_pr_sl_op	40,4827	46658,5898	0,0021	0,0028	ob_ine	0,0585	0,0772	0,0025	0,0033
neob_pr_sl_ine	60,5264	50974,1387	0,0026	0,0033	ob_op	0,0665	0,0865	0,0029	0,0037
neob_pr_op_ine	60,8622	50871,8070	0,0026	0,0034	ob_pr	0,0634	0,0644	0,0020	0,0027
neob_sl_op_ine	60,8448	53865,1619	0,0016	0,0022	ob_sl	0,0670	0,0870	0,0029	0,0038
neob_pr_sl_op_ine	61,8244	50828,4266	0,0026	0,0034	ob	0,0679	0,0881	0,0030	0,0038

Tabuľka 11: r01zb7



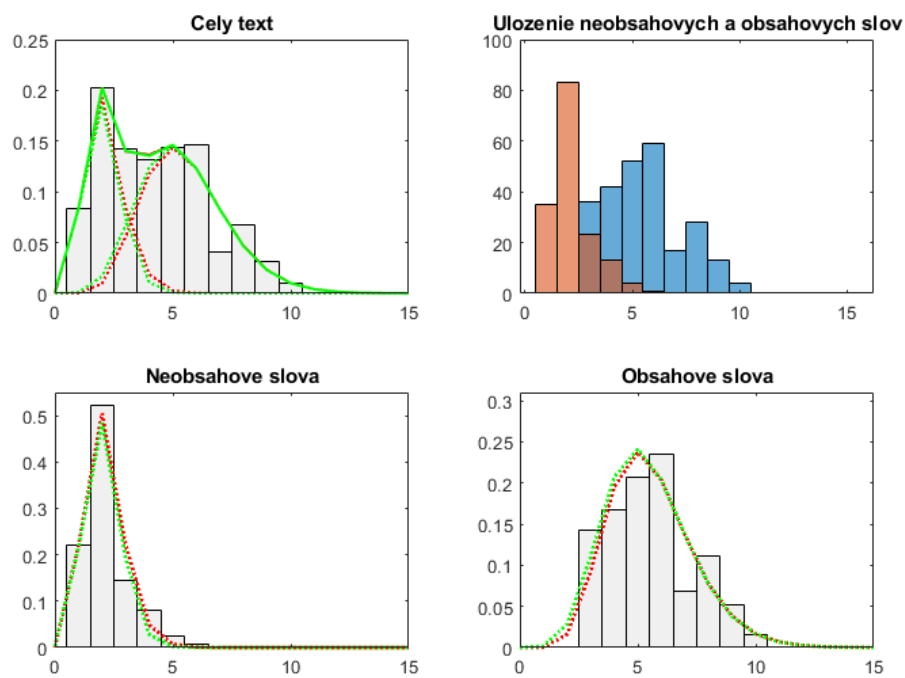
Obr. 26: r01zb7



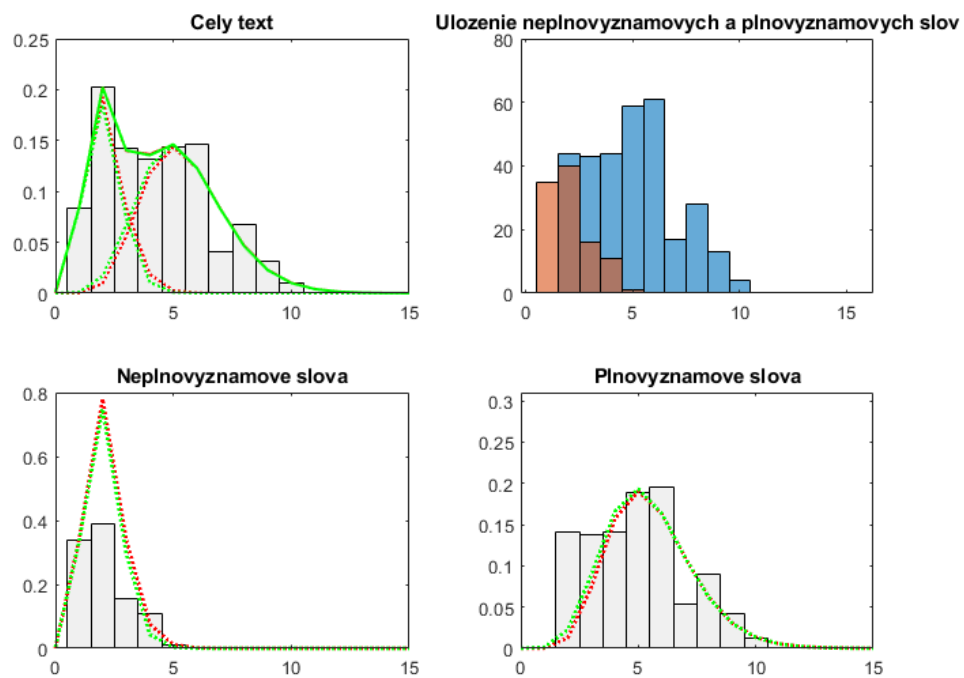
Obr. 27: r01zb7

r01zb8	C	Cdg	CH	CHdg	pocet slov celok	C	Cdg	CH	CHdg
celok	0,0381	0,0383	0,0028	0,0028	416				
neob	0,1251	0,6915	0,0011	0,0010	ob_pr_sl_op_ine	0,1013	0,1005	0,0041	0,0040
neob_pr	0,1248	0,6882	0,0011	0,0011	ob_sl_op_ine	0,1085	0,1081	0,0041	0,0041
neob_sl	0,1251	0,6915	0,0011	0,0010	ob_pr_op_ine	0,1013	0,1005	0,0041	0,0040
neob_op	0,1251	0,6915	0,0011	0,0010	ob_pr_sl_ine	0,1013	0,1005	0,0041	0,0040
neob_ine	0,4398	2,4555	0,0013	0,0013	ob_pr_sl_op	0,1050	0,1046	0,0042	0,0042
neob_pr_sl	0,1248	0,6882	0,0011	0,0011	ob_op_ine	0,1085	0,1081	0,0041	0,0041
neob_pr_op	0,1248	0,6882	0,0011	0,0011	ob_sl_ine	0,1085	0,1081	0,0041	0,0041
neob_pr_ine	0,4376	2,4416	0,0014	0,0014	ob_sl_op	0,1124	0,1124	0,0043	0,0043
neob_sl_op	0,1251	0,6915	0,0011	0,0010	ob_pr_ine	0,1013	0,1005	0,0041	0,0040
neob_sl_ine	0,4398	2,4555	0,0013	0,0013	ob_pr_op	0,1050	0,1046	0,0042	0,0042
neob_op_ine	0,4398	2,4555	0,0013	0,0013	ob_pr_sl	0,1050	0,1046	0,0042	0,0042
neob_pr_sl_op	0,1248	0,6882	0,0011	0,0011	ob_ine	0,1085	0,1081	0,0041	0,0041
neob_pr_sl_ine	0,4376	2,4416	0,0014	0,0014	ob_op	0,1124	0,1124	0,0043	0,0043
neob_pr_op_ine	0,4376	2,4416	0,0014	0,0014	ob_pr	0,1050	0,1046	0,0042	0,0042
neob_sl_op_ine	0,4398	2,4555	0,0013	0,0013	ob_sl	0,1124	0,1124	0,0043	0,0043
neob_pr_sl_op_ine	0,4376	2,4416	0,0014	0,0014	ob	0,1124	0,1124	0,0043	0,0043

Tabuľka 12: r01zb8



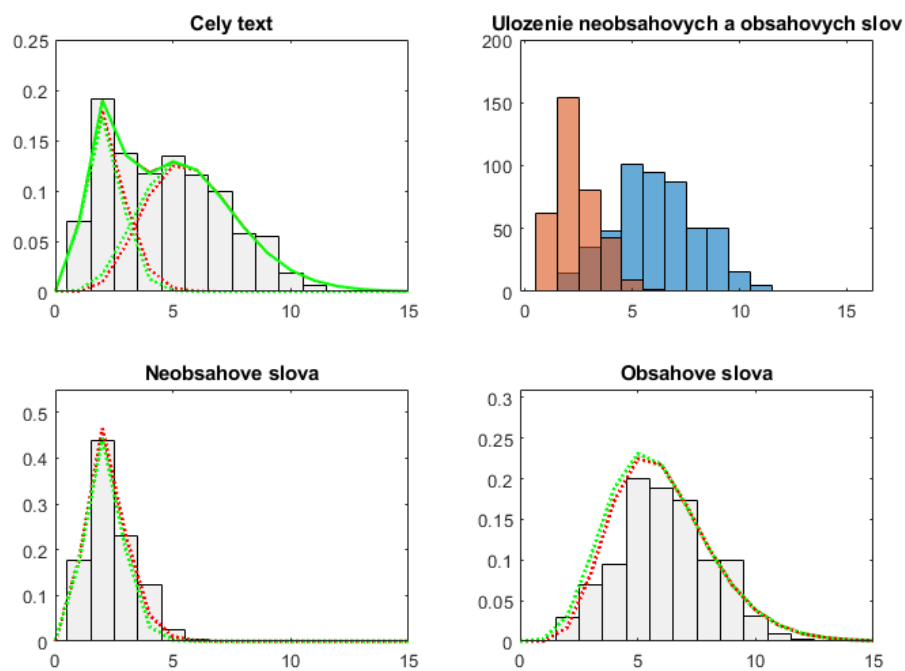
Obr. 28: r01zb8



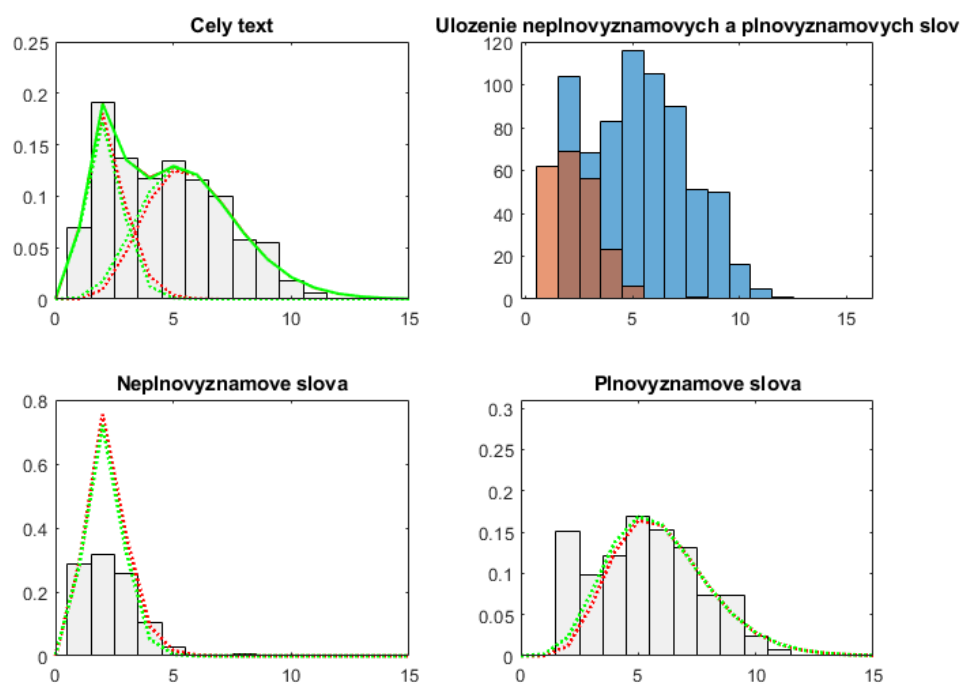
Obr. 29: r01zb8

r01zb9	C	Cdg	CH	CHdg	pocet slov celok	C	Cdg	CH	CHdg
celok	0,0150	0,0152	0,0005	0,0005	907				
neob	0,6375	38,6508	0,0008	0,0015	ob_pr_sl_op_ine	0,0608	0,0520	0,0013	0,0019
neob_pr	2,3376	145,2430	0,0016	0,0026	ob_sl_op_ine	0,0670	0,0715	0,0021	0,0030
neob_sl	0,6375	38,6508	0,0008	0,0015	ob_pr_op_ine	0,0608	0,0520	0,0013	0,0019
neob_op	0,6677	38,1844	0,0009	0,0017	ob_pr_sl_ine	0,0626	0,0558	0,0014	0,0020
neob_ine	1,1606	44,3178	0,0011	0,0020	ob_pr_sl_op	0,0648	0,0649	0,0016	0,0024
neob_pr_sl	2,3376	145,2430	0,0016	0,0026	ob_op_ine	0,0670	0,0715	0,0021	0,0030
neob_pr_op	2,3694	143,2334	0,0018	0,0029	ob_sl_ine	0,0701	0,0767	0,0023	0,0032
neob_pr_ine	3,0939	150,5908	0,0021	0,0034	ob_sl_op	0,0755	0,0890	0,0026	0,0037
neob_sl_op	0,6677	38,1844	0,0009	0,0017	ob_pr_ine	0,0626	0,0558	0,0014	0,0020
neob_sl_ine	1,1606	44,3178	0,0011	0,0020	ob_pr_op	0,0648	0,0649	0,0016	0,0024
neob_op_ine	1,1962	43,8386	0,0013	0,0023	ob_pr_sl	0,0682	0,0702	0,0017	0,0027
neob_pr_sl_op	2,3694	143,2334	0,0018	0,0029	ob_ine	0,0701	0,0767	0,0023	0,0032
neob_pr_sl_ine	3,0939	150,5908	0,0021	0,0034	ob_op	0,0755	0,0890	0,0026	0,0037
neob_pr_op_ine	3,1273	148,6542	0,0024	0,0037	ob_pr	0,0682	0,0702	0,0017	0,0027
neob_sl_op_ine	1,1962	43,8386	0,0013	0,0023	ob_sl	0,0804	0,0960	0,0028	0,0040
neob_pr_sl_op_ine	3,1273	148,6542	0,0024	0,0037	ob	0,0804	0,0960	0,0028	0,0040

Tabuľka 13: r01zb9



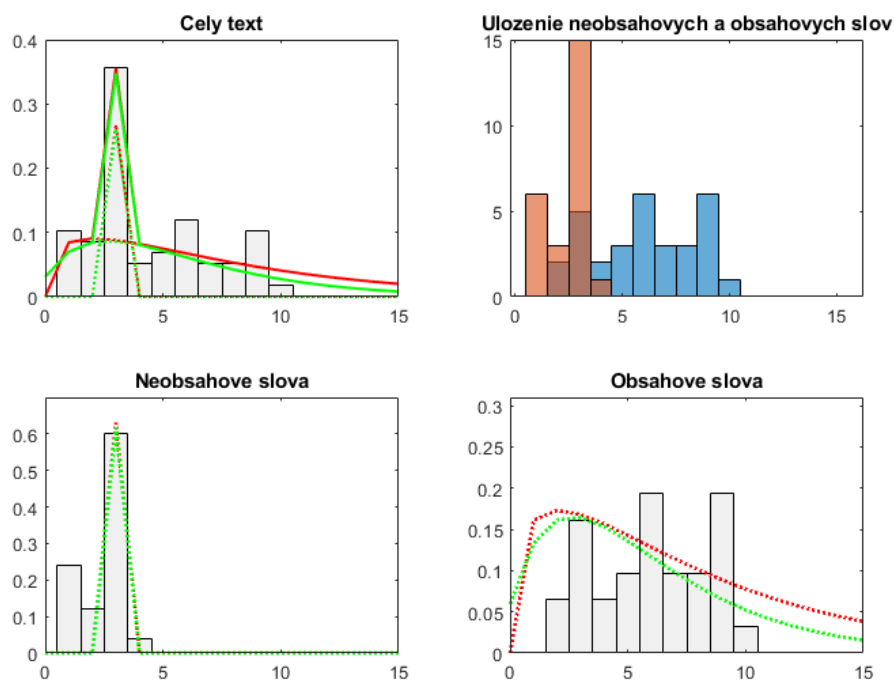
Obr. 30: r01zb9



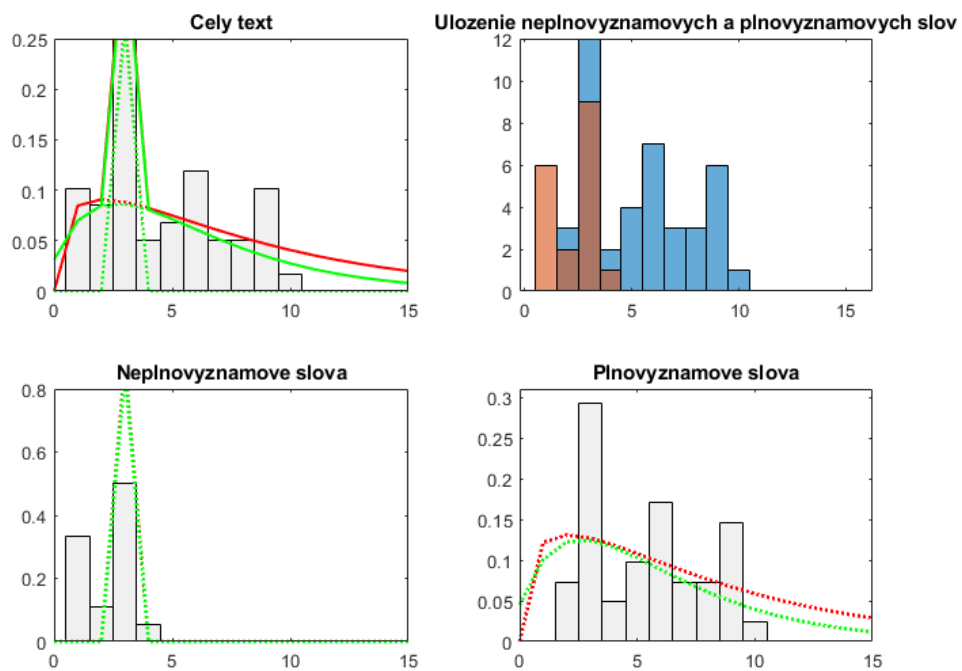
Obr. 31: r01zb9

01zb10	C	Cdg	CH	CHdg	pocet slov celok	C	Cdg	CH	CHdg
celok	0,1360	0,2184	0,0077	0,0100	59				
neob	65535	65535	0,0134	0,0133	ob_pr_sl_op_ine	0,4703	0,5609	0,0193	0,0180
neob_pr	65535	65535	0,0137	0,0136	ob_sl_op_ine	0,4972	0,5881	0,0198	0,0184
neob_sl	65535	65535	0,0134	0,0133	ob_pr_op_ine	0,4703	0,5609	0,0193	0,0180
neob_op	65535	65535	0,0134	0,0133	ob_pr_sl_ine	0,4703	0,5609	0,0193	0,0180
neob_ine	65535	65535	0,0135	0,0136	ob_pr_sl_op	0,4565	0,5410	0,0177	0,0161
neob_pr_sl	65535	65535	0,0137	0,0136	ob_op_ine	0,4972	0,5881	0,0198	0,0184
neob_pr_op	65535	65535	0,0137	0,0136	ob_sl_ine	0,4972	0,5881	0,0198	0,0184
neob_pr_ine	65535	65535	0,0138	0,0139	ob_sl_op	0,4847	0,5693	0,0182	0,0165
neob_sl_op	65535	65535	0,0134	0,0133	ob_pr_ine	0,4703	0,5609	0,0193	0,0180
neob_sl_ine	65535	65535	0,0135	0,0136	ob_pr_op	0,4565	0,5410	0,0177	0,0161
neob_op_ine	65535	65535	0,0135	0,0136	ob_pr_sl	0,4565	0,5410	0,0177	0,0161
neob_pr_sl_op	65535	65535	0,0137	0,0136	ob_ine	0,4972	0,5881	0,0198	0,0184
neob_pr_sl_ine	65535	65535	0,0138	0,0139	ob_op	0,4847	0,5693	0,0182	0,0165
neob_pr_op_ine	65535	65535	0,0138	0,0139	ob_pr	0,4565	0,5410	0,0177	0,0161
neob_sl_op_ine	65535	65535	0,0135	0,0136	ob_sl	0,4847	0,5693	0,0182	0,0165
neob_pr_sl_op_ine	65535	65535	0,0138	0,0139	ob	0,4847	0,5693	0,0182	0,0165

Tabulka 14: r01zb10



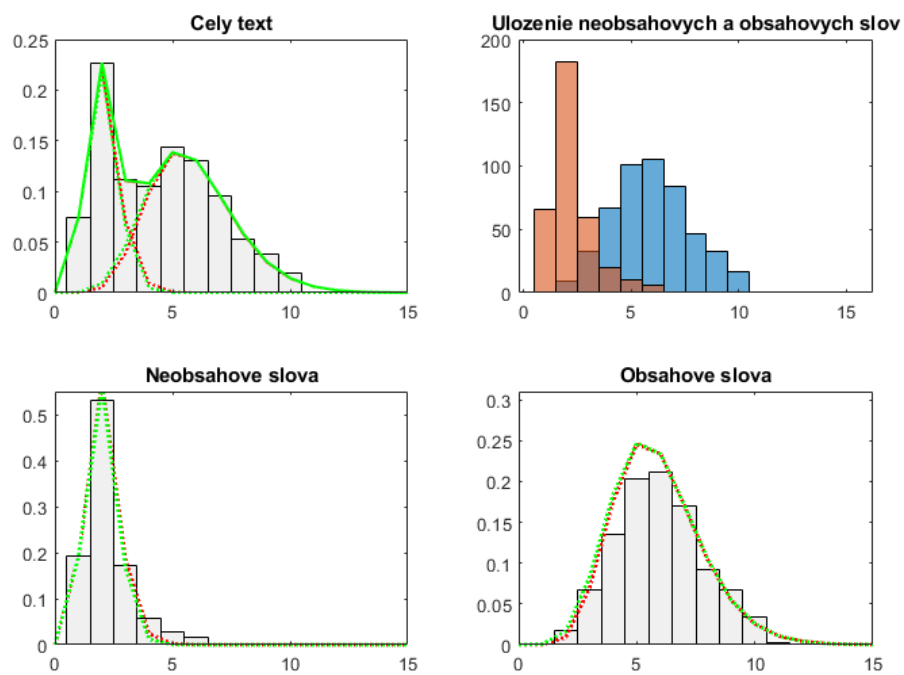
Obr. 32: r01zb10



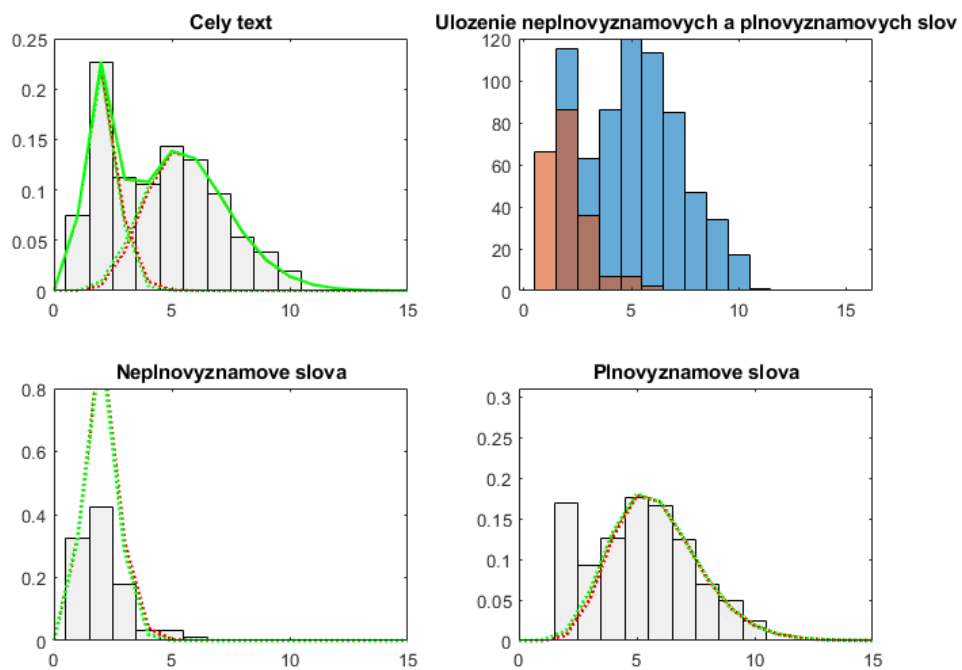
Obr. 33: r01zb10

01zb11	C	Cdg	CH	CHdg	pocet slov celok	C	Cdg	CH	CHdg
celok	0,0086	0,0084	0,0002	0,0002	885				
neob	3,2240	60,4136	0,0006	0,0006	ob_pr_sl_op_ine	0,1063	0,0479	0,0008	0,0008
neob_pr	45,1405	20438,4704	0,0009	0,0010	ob_sl_op_ine	0,0573	0,0358	0,0010	0,0012
neob_sl	3,2240	60,4136	0,0006	0,0006	ob_pr_op_ine	0,1063	0,0479	0,0008	0,0008
neob_op	3,4651	61,2330	0,0007	0,0007	ob_pr_sl_ine	0,0958	0,0445	0,0008	0,0009
neob_ine	1175,9583	3597042,7797	0,0008	0,0009	ob_pr_sl_op	0,0777	0,0404	0,0009	0,0010
neob_pr_sl	45,1405	20438,4704	0,0009	0,0010	ob_op_ine	0,0573	0,0358	0,0010	0,0012
neob_pr_op	44,9760	20164,8382	0,0010	0,0012	ob_sl_ine	0,0520	0,0358	0,0011	0,0013
neob_pr_ine	1149,3729	3410345,9279	0,0012	0,0015	ob_sl_op	0,0441	0,0387	0,0013	0,0016
neob_sl_op	3,4651	61,2330	0,0007	0,0007	ob_pr_ine	0,0958	0,0445	0,0008	0,0009
neob_sl_ine	1175,9583	3597042,7797	0,0008	0,0009	ob_pr_op	0,0777	0,0404	0,0009	0,0010
neob_op_ine	1160,3611	3548039,1911	0,0009	0,0010	ob_pr_sl	0,0701	0,0392	0,0009	0,0011
neob_pr_sl_op	44,9760	20164,8382	0,0010	0,0012	ob_ine	0,0520	0,0358	0,0011	0,0013
neob_pr_sl_ine	1149,3729	3410345,9279	0,0012	0,0015	ob_op	0,0441	0,0387	0,0013	0,0016
neob_pr_op_ine	1135,1661	3366514,3816	0,0014	0,0017	ob_pr	0,0701	0,0392	0,0009	0,0011
neob_sl_op_ine	1160,3611	3548039,1911	0,0009	0,0010	ob_sl	0,0421	0,0412	0,0014	0,0017
neob_pr_sl_op_ine	1135,1661	3366514,3816	0,0014	0,0017	ob	0,0421	0,0412	0,0014	0,0017

Tabulka 15: r01zb11



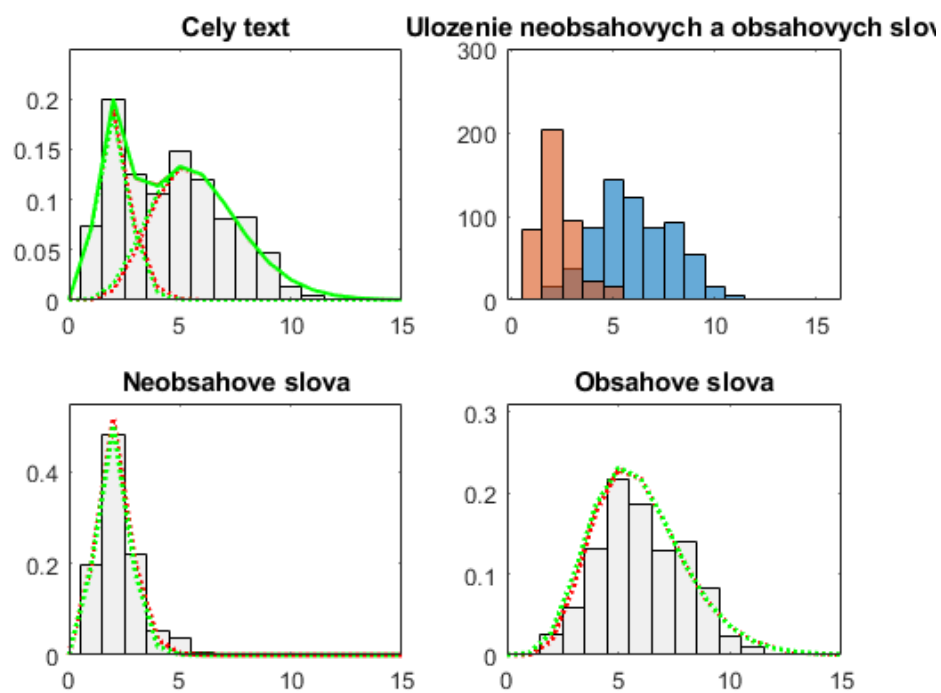
Obr. 34: r01zb11



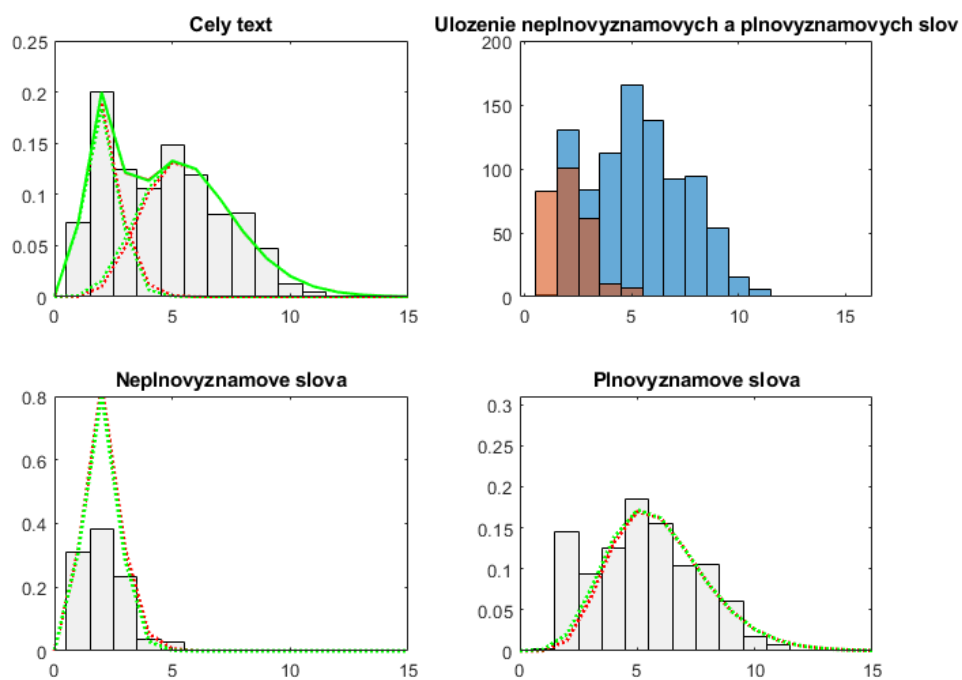
Obr. 35: r01zb11

r01zb12	C	Cdg	CH	CHdg	pocet slov celok	C	Cdg	CH	CHdg
celok	0,0213	0,0209	0,0012	0,0011	1155				
neob	4,8759	620,2219	0,0005	0,0007	ob_pr_sl_op_ine	0,0682	0,0492	0,0013	0,0015
neob_pr	18,3341	2270,9916	0,0009	0,0013	ob_sl_op_ine	0,0510	0,0566	0,0016	0,0021
neob_sl	4,8759	620,2219	0,0005	0,0007	ob_pr_op_ine	0,0682	0,0492	0,0013	0,0015
neob_op	10,8521	840,4666	0,0005	0,0008	ob_pr_sl_ine	0,0716	0,0530	0,0015	0,0017
neob_ine	6,6537	629,9103	0,0008	0,0011	ob_pr_sl_op	0,0715	0,0570	0,0015	0,0018
neob_pr_sl	18,3341	2270,9916	0,0009	0,0013	ob_op_ine	0,0510	0,0566	0,0016	0,0021
neob_pr_op	25,4641	2520,6081	0,0009	0,0014	ob_sl_ine	0,0551	0,0611	0,0018	0,0023
neob_pr_ine	21,2147	2237,1606	0,0014	0,0020	ob_sl_op	0,0605	0,0703	0,0021	0,0027
neob_sl_op	10,8521	840,4666	0,0005	0,0008	ob_pr_ine	0,0716	0,0530	0,0015	0,0017
neob_sl_ine	6,6537	629,9103	0,0008	0,0011	ob_pr_op	0,0715	0,0570	0,0015	0,0018
neob_op_ine	14,2370	911,8710	0,0009	0,0012	ob_pr_sl	0,0755	0,0613	0,0017	0,0020
neob_pr_sl_op	25,4641	2520,6081	0,0009	0,0014	ob_ine	0,0551	0,0611	0,0018	0,0023
neob_pr_sl_ine	21,2147	2237,1606	0,0014	0,0020	ob_op	0,0605	0,0703	0,0021	0,0027
neob_pr_op_ine	29,7529	2542,5673	0,0015	0,0021	ob_pr	0,0755	0,0613	0,0017	0,0020
neob_sl_op_ine	14,2370	911,8710	0,0009	0,0012	ob_sl	0,0653	0,0756	0,0023	0,0029
neob_pr_sl_op_ine	29,7529	2542,5673	0,0015	0,0021	ob	0,0653	0,0756	0,0023	0,0029

Tabulka 16: r01zb12



Obr. 36: r01zb12



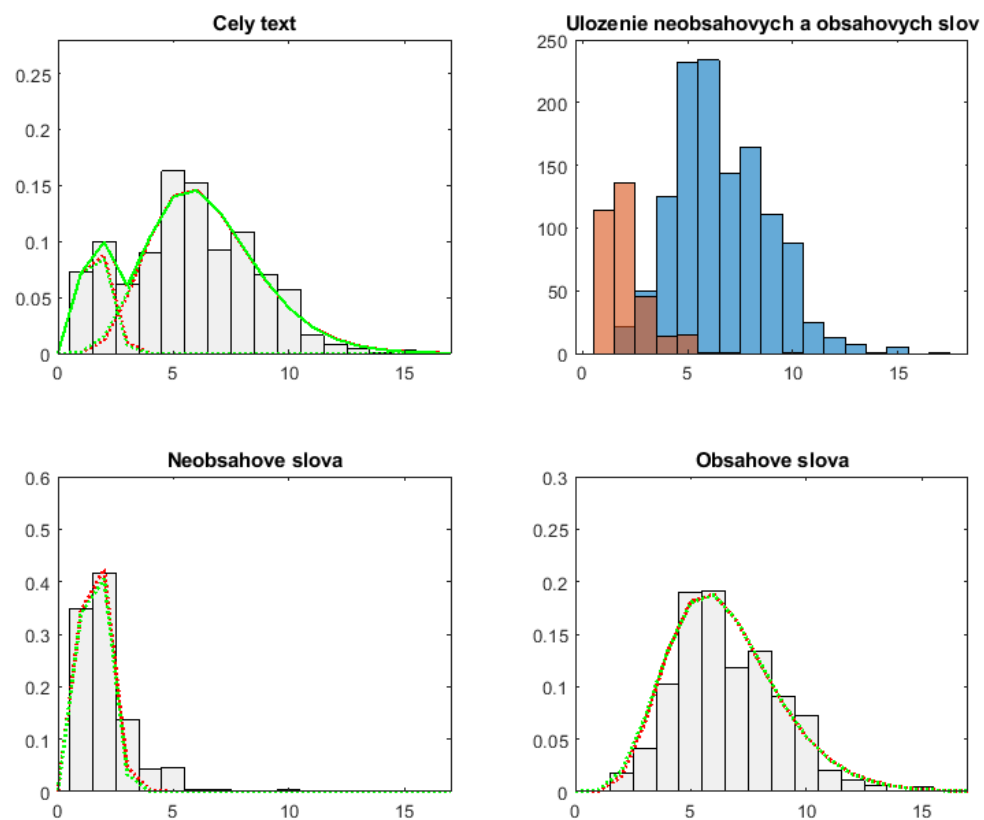
Obr. 37: r01zb12

s1-s11	C	Cdg	CH	CHdg	pocet slov celok	C	Cdg	CH	CHdg
celok	0,014383814	0,015104909	0,014383814	0,015104909	2298	0,000536333			
neob	0,206541501	1,084099529	0,001649323	0,001031099	ob_pr_sl_op_ine	0,529972484	0,210288141	0,002111922	0,001566417
neob_pr	0,250263004	2,288806689	0,001511595	0,001065888	ob_sl_op_ine	0,538043928	0,214301455	0,001964321	0,0015728
neob_sl	0,206541501	1,084099529	0,001649323	0,001031099	ob_pr_op_ine	0,529972484	0,210288141	0,002111922	0,001566417
neob_op	0,206541501	1,084099529	0,001649323	0,001031099	ob_pr_sl_ine	0,529972484	0,210288141	0,002111922	0,001566417
neob_ine	0,17981393	1,054950581	0,001511195	0,000935527	ob_pr_sl_op	0,519733659	0,203517868	0,001981123	0,001471976
neob_pr_sl	0,250263004	2,288806689	0,001511595	0,001065888	ob_op_ine	0,538043928	0,214301455	0,001964321	0,0015728
neob_pr_op	0,250263004	2,288806689	0,001511595	0,001065888	ob_sl_ine	0,538043928	0,214301455	0,001964321	0,0015728
neob_pr_ine	0,232262994	2,256846473	0,001386344	0,000983193	ob_sl_op	0,527812684	0,207572992	0,001846399	0,001491236
neob_sl_op	0,206541501	1,084099529	0,001649323	0,001031099	ob_pr_ine	0,529972484	0,210288141	0,002111922	0,001566417
neob_sl_ine	0,17981393	1,054950581	0,001511195	0,000935527	ob_pr_op	0,519733659	0,203517868	0,001981123	0,001471976
neob_op_ine	0,17981393	1,054950581	0,001511195	0,000935527	ob_pr_sl	0,519733659	0,203517868	0,001981123	0,001471976
neob_pr_sl_op	0,250263004	2,288806689	0,001511595	0,001065888	ob_ine	0,538043928	0,214301455	0,001964321	0,0015728
neob_pr_sl_ine	0,232262994	2,256846473	0,001386344	0,000983193	ob_op	0,527812684	0,207572992	0,001846399	0,001491236
neob_pr_op_ine	0,232262994	2,256846473	0,001386344	0,000983193	ob_pr	0,519733659	0,203517868	0,001981123	0,001471976
neob_sl_op_ine	0,17981393	1,054950581	0,001511195	0,000935527	ob_sl	0,527812684	0,207572992	0,001846399	0,001491236
neob_pr_sl_op_ine	0,232262994	2,256846473	0,001386344	0,000983193	ob	0,527812684	0,207572992	0,001846399	0,001491236

Tabuľka 17: s1-s11

t1-t6	C	Cdg	CH	CHdg	pocet slov celok	C	Cdg	CH	CHdg
celok	0,0337	0,0337	0,0025	0,0025	1572				
neob	659548789,2082	65535,0000	0,0005	0,0007	ob_pr_sl_op_ine	0,0515	0,0556	0,0028	0,0030
neob_pr	620896800,0590	65535,0000	0,0007	0,0009	ob_sl_op_ine	0,0506	0,0547	0,0026	0,0028
neob_sl	659548789,2082	65535,0000	0,0005	0,0007	ob_pr_op_ine	0,0515	0,0556	0,0028	0,0030
neob_op	659548789,2082	65535,0000	0,0005	0,0007	ob_pr_sl_ine	0,0515	0,0556	0,0028	0,0030
neob_ine	655539374,0998	65535,0000	0,0006	0,0007	ob_pr_sl_op	0,0528	0,0569	0,0028	0,0030
neob_pr_sl	620896800,0590	65535,0000	0,0007	0,0009	ob_op_ine	0,0506	0,0547	0,0026	0,0028
neob_pr_op	620896800,0590	65535,0000	0,0007	0,0009	ob_sl_ine	0,0506	0,0547	0,0026	0,0028
neob_pr_ine	617358926,5765	65535,0000	0,0008	0,0009	ob_sl_op	0,0519	0,0561	0,0026	0,0028
neob_sl_op	659548789,2082	65535,0000	0,0005	0,0007	ob_pr_ine	0,0515	0,0556	0,0028	0,0030
neob_sl_ine	655539374,0998	65535,0000	0,0006	0,0007	ob_pr_op	0,0528	0,0569	0,0028	0,0030
neob_op_ine	655539374,0998	65535,0000	0,0006	0,0007	ob_pr_sl	0,0528	0,0569	0,0028	0,0030
neob_pr_sl_op	620896800,0590	65535,0000	0,0007	0,0009	ob_ine	0,0506	0,0547	0,0026	0,0028
neob_pr_sl_ine	617358926,5765	65535,0000	0,0008	0,0009	ob_op	0,0519	0,0561	0,0026	0,0028
neob_pr_op_ine	617358926,5765	65535,0000	0,0008	0,0009	ob_pr	0,0528	0,0569	0,0028	0,0030
neob_sl_op_ine	655539374,0998	65535,0000	0,0006	0,0007	ob_sl	0,0519	0,0561	0,0026	0,0028
neob_pr_sl_op_ine	617358926,5765	65535,0000	0,0008	0,0009	ob	0,0519	0,0561	0,0026	0,0028

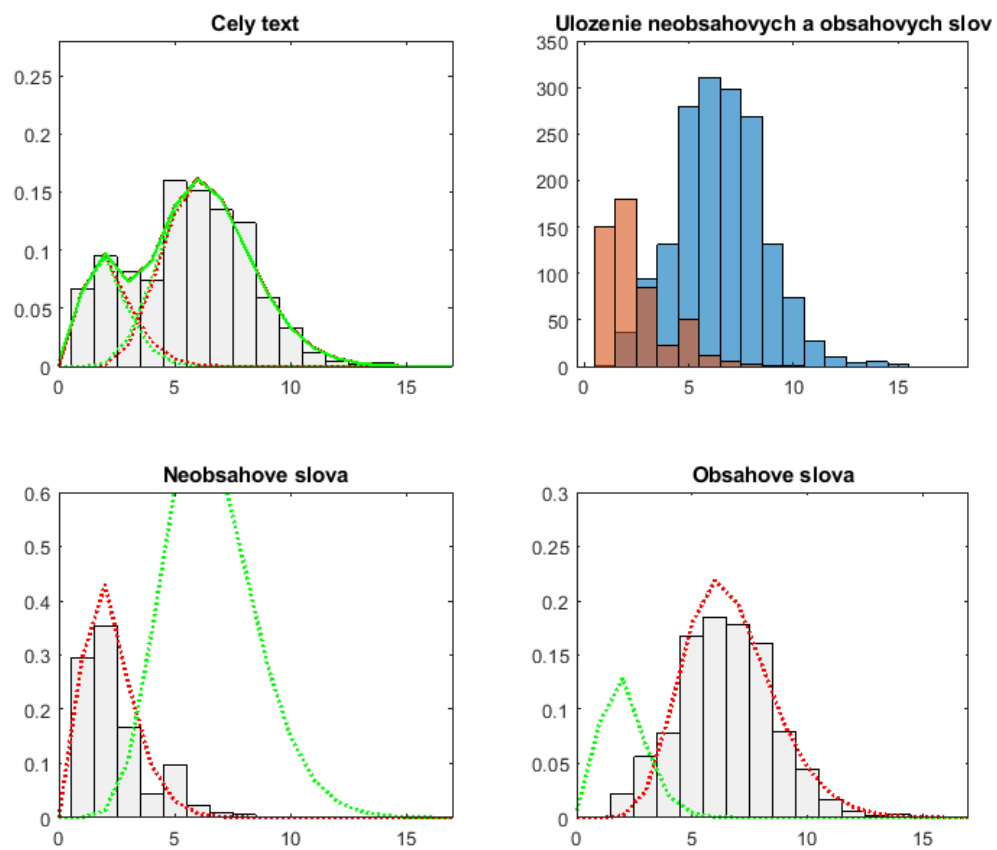
Tabulka 18: t1-t6



Obr. 38: t1-t6

p1-p11	C	Cdg	CH	CHdg	pocet slov celok	C	Cdg	CH	CHdg
celok	0,0185	0,0190	0,0015	0,0014	2286				
neob	0,5124	264,6575	0,0010	0,0875	ob_pr_sl_op_ine	0,3265	1249763,9284	0,0018	0,0864
neob_pr	2,5715	235,3280	0,0015	0,0820	ob_sl_op_ine	0,3371	1295982,5430	0,0019	0,0804
neob_sl	0,5299	260,5345	0,0010	0,0868	ob_pr_op_ine	0,3288	1255502,4477	0,0019	0,0857
neob_op	0,5124	264,6575	0,0010	0,0875	ob_pr_sl_ine	0,3265	1249763,9284	0,0018	0,0864
neob_ine	0,5100	262,0887	0,0009	0,0874	ob_pr_sl_op	0,3214	1253344,6264	0,0018	0,0863
neob_pr_sl	2,5915	232,0588	0,0015	0,0813	ob_op_ine	0,3401	1302156,6385	0,0020	0,0798
neob_pr_op	2,5715	235,3280	0,0015	0,0820	ob_sl_ine	0,3371	1295982,5430	0,0019	0,0804
neob_pr_ine	2,5597	233,2944	0,0014	0,0819	ob_sl_op	0,3321	1299834,7587	0,0019	0,0803
neob_sl_op	0,5299	260,5345	0,0010	0,0868	ob_pr_ine	0,3288	1255502,4477	0,0019	0,0857
neob_sl_ine	0,5300	258,0446	0,0009	0,0867	ob_pr_op	0,3237	1259116,1583	0,0018	0,0856
neob_op_ine	0,5100	262,0887	0,0009	0,0874	ob_pr_sl	0,3214	1253344,6264	0,0018	0,0863
neob_pr_sl_op	2,5915	232,0588	0,0015	0,0813	ob_ine	0,3401	1302156,6385	0,0020	0,0798
neob_pr_sl_ine	2,5821	230,0810	0,0015	0,0812	ob_op	0,3350	1306045,7078	0,0020	0,0797
neob_pr_op_ine	2,5597	233,2944	0,0014	0,0819	ob_pr	0,3237	1259116,1583	0,0018	0,0856
neob_sl_op_ine	0,5300	258,0446	0,0009	0,0867	ob_sl	0,3321	1299834,7587	0,0019	0,0803
neob_pr_sl_op_ine	2,5821	230,0810	0,0015	0,0812	ob	0,3350	1306045,7078	0,0020	0,0797

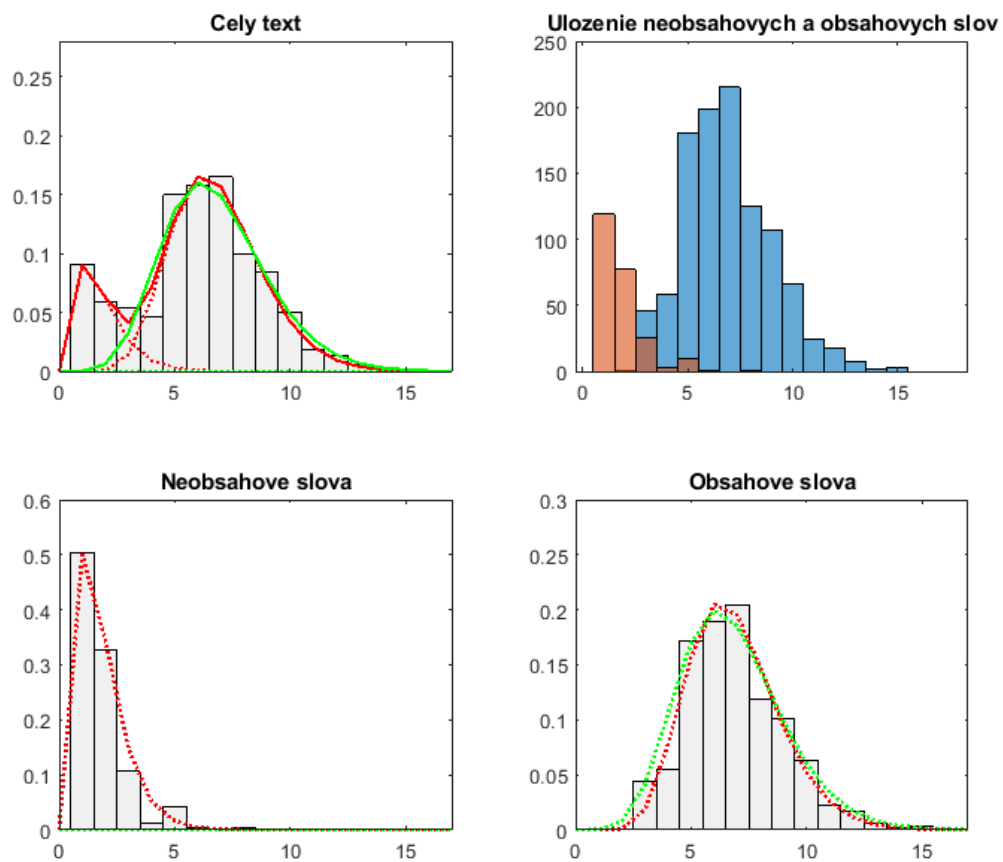
Tabuľka 19: p1-p11



Obr. 39: p1-p11

sa1-sa7	C	Cdg	CH	CHdg	pocet slov celok	C	Cdg	CH	CHdg
celok	0,0303	28,3986	0,0018	0,0138	1310				
neob	0,1241	65535,0000	0,0001	0,0121	ob_pr_sl_op_ine	0,0611	0,0413	0,0017	0,0024
neob_pr	3,2473	65535,0000	0,0002	0,0122	ob_sl_op_ine	0,0607	0,0417	0,0016	0,0024
neob_sl	0,1757	65535,0000	0,0001	0,0121	ob_pr_op_ine	0,0617	0,0416	0,0017	0,0024
neob_op	0,1241	65535,0000	0,0001	0,0121	ob_pr_sl_ine	0,0611	0,0413	0,0017	0,0024
neob_ine	0,8099	65535,0000	0,0001	0,0121	ob_pr_sl_op	0,0625	0,0423	0,0018	0,0024
neob_pr_sl	3,2745	65535,0000	0,0002	0,0122	ob_op_ine	0,0614	0,0421	0,0016	0,0024
neob_pr_op	3,2473	65535,0000	0,0002	0,0122	ob_sl_ine	0,0607	0,0417	0,0016	0,0024
neob_pr_ine	4,3719	65535,0000	0,0002	0,0122	ob_sl_op	0,0623	0,0429	0,0017	0,0024
neob_sl_op	0,1757	65535,0000	0,0001	0,0121	ob_pr_ine	0,0617	0,0416	0,0017	0,0024
neob_sl_ine	0,9057	65535,0000	0,0002	0,0122	ob_pr_op	0,0632	0,0427	0,0018	0,0025
neob_op_ine	0,8099	65535,0000	0,0001	0,0121	ob_pr_sl	0,0625	0,0423	0,0018	0,0024
neob_pr_sl_op	3,2745	65535,0000	0,0002	0,0122	ob_ine	0,0614	0,0421	0,0016	0,0024
neob_pr_sl_ine	4,4349	65535,0000	0,0002	0,0123	ob_op	0,0630	0,0433	0,0018	0,0025
neob_pr_op_ine	4,3719	65535,0000	0,0002	0,0122	ob_pr	0,0632	0,0427	0,0018	0,0025
neob_sl_op_ine	0,9057	65535,0000	0,0002	0,0122	ob_sl	0,0623	0,0429	0,0017	0,0024
neob_pr_sl_op_ine	4,4349	65535,0000	0,0002	0,0123	ob	0,0630	0,0433	0,0018	0,0025

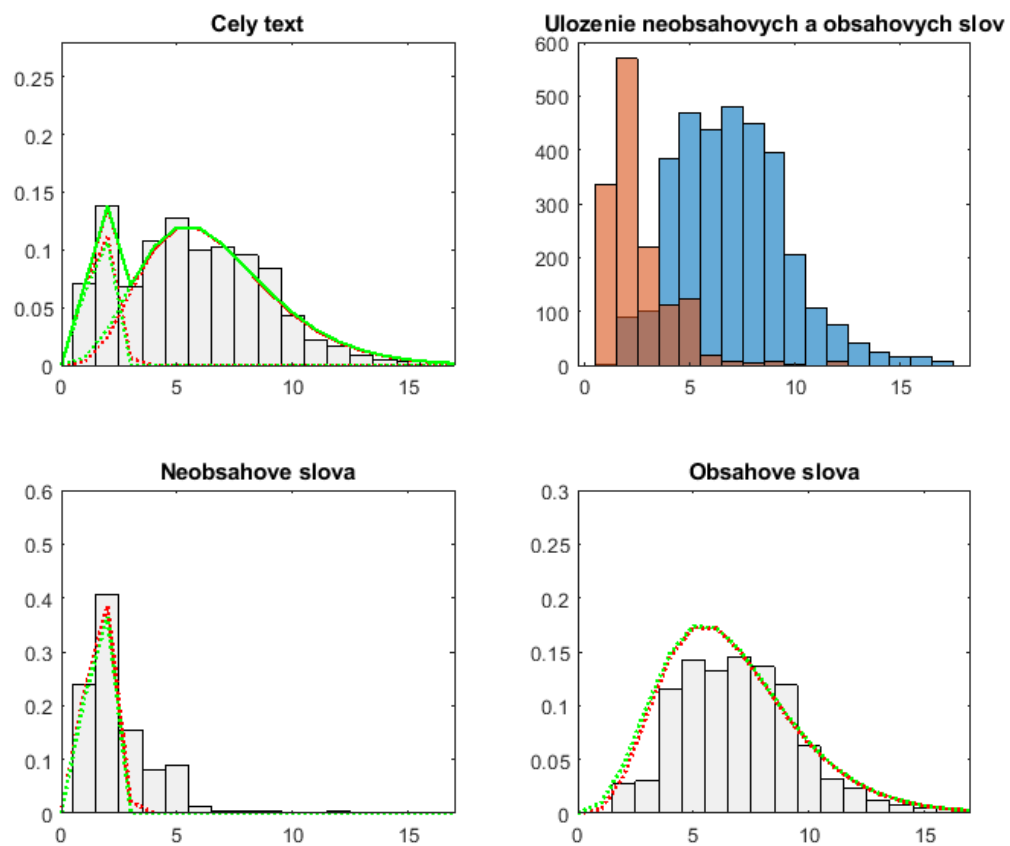
Tabulka 20: sa1-sa7



Obr. 40: sa1-sa7

ne0003-ne0078	C	Cdg	CH	CHdg	pocet slov celok	C	Cdg	CH	CHdg
celok	0,0248	0,0243	0,0012	0,0011	4804				
0,002802893	70277639386206700000,0000	65535,0000		0,0035	ob_pr_sl_op_ine	0,0831	0,1025	0,0034	0,0041
0,003231457	67347419954845200000,0000	65535,0000		0,0039	ob_sl_op_ine	0,0908	0,1114	0,0038	0,0046
0,002876235	69779922393386600000,0000	65535,0000		0,0036	ob_pr_op_ine	0,0846	0,1043	0,0035	0,0042
0,002802893	70277639386206700000,0000	65535,0000		0,0035	ob_pr_sl_ine	0,0831	0,1025	0,0034	0,0041
0,002944833	69337966516653900000,0000	65535,0000		0,0036	ob_pr_sl_op	0,0853	0,1049	0,0035	0,0043
0,003317625	66890207327860500000,0000	65535,0000		0,0040	ob_op_ine	0,0926	0,1134	0,0039	0,0047
0,003231457	67347419954845200000,0000	65535,0000		0,0039	ob_sl_ine	0,0908	0,1114	0,0038	0,0046
0,003387697	66483991494439400000,0000	65535,0000		0,0041	ob_sl_op	0,0933	0,1141	0,0040	0,0047
0,002876235	69779922393386600000,0000	65535,0000		0,0036	ob_pr_ine	0,0846	0,1043	0,0035	0,0042
0,003019389	68853424472512400000,0000	65535,0000		0,0037	ob_pr_op	0,0869	0,1068	0,0036	0,0044
0,002944833	69337966516653900000,0000	65535,0000		0,0036	ob_pr_sl	0,0853	0,1049	0,0035	0,0043
0,003317625	66890207327860500000,0000	65535,0000		0,0040	ob_ine	0,0926	0,1134	0,0039	0,0047
0,003475078	66038388334289000000,0000	65535,0000		0,0042	ob_op	0,0951	0,1161	0,0041	0,0048
0,003387697	66483991494439400000,0000	65535,0000		0,0041	ob_pr	0,0869	0,1068	0,0036	0,0044
0,003019389	68853424472512400000,0000	65535,0000		0,0037	ob_sl	0,0933	0,1141	0,0040	0,0047
0,003475078	66038388334289000000,0000	65535,0000		0,0042	ob	0,0951	0,1161	0,0041	0,0048

Tabuľka 21: ne0003-ne0078



Obr. 41: ne0003-ne0078