

Univerzita Pardubice
Fakulta elektrotechniky a informatiky

Statistické vyhodnocení objevených exoplanet

Michal Struna

Semestrální práce

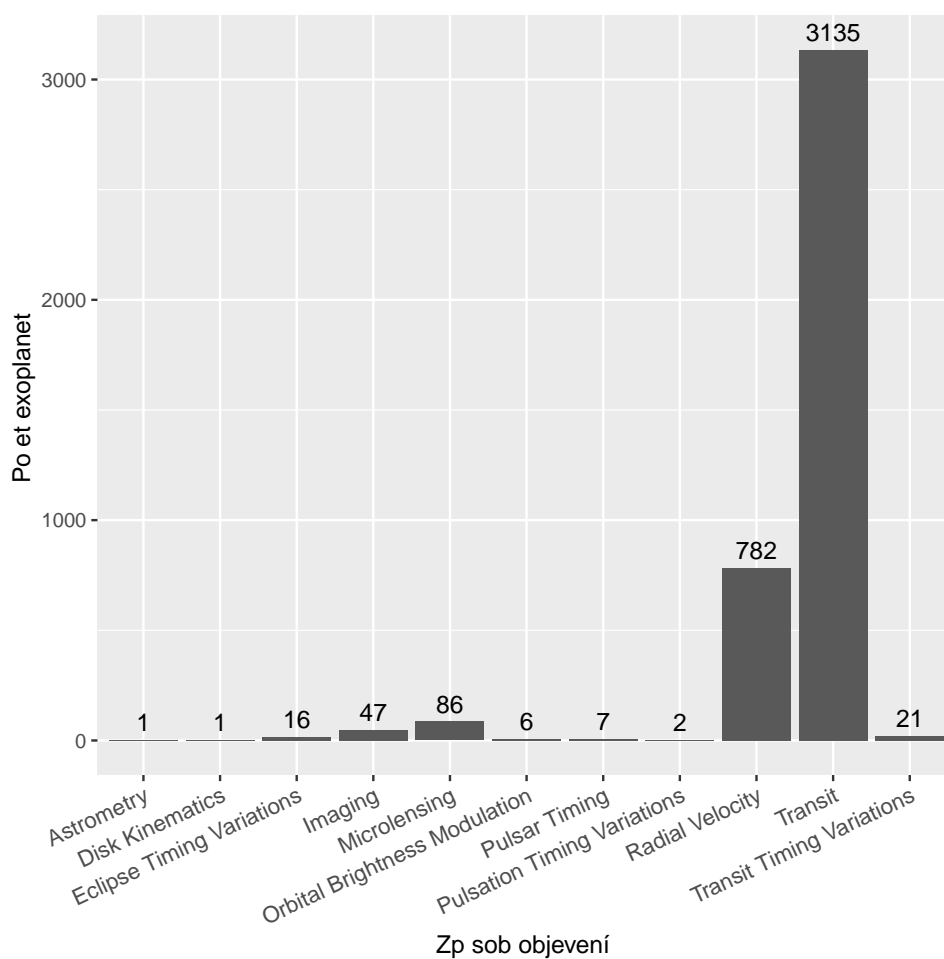
2019

1 Úvod

Tato semestrální práce zkoumá data a vztahy mezi daty u objevených exoplanet. Jedná se o planety mimo sluneční soustavu, kterých bylo k listopadu 2019 potvrzeno 4104. K práci byl využit archiv pocházející od NASA a obsahující všechny tyto planety. Je dostupný přes webové API nebo ke stažení jako CSV na adrese <https://exoplanetarchive.ipac.caltech.edu>.

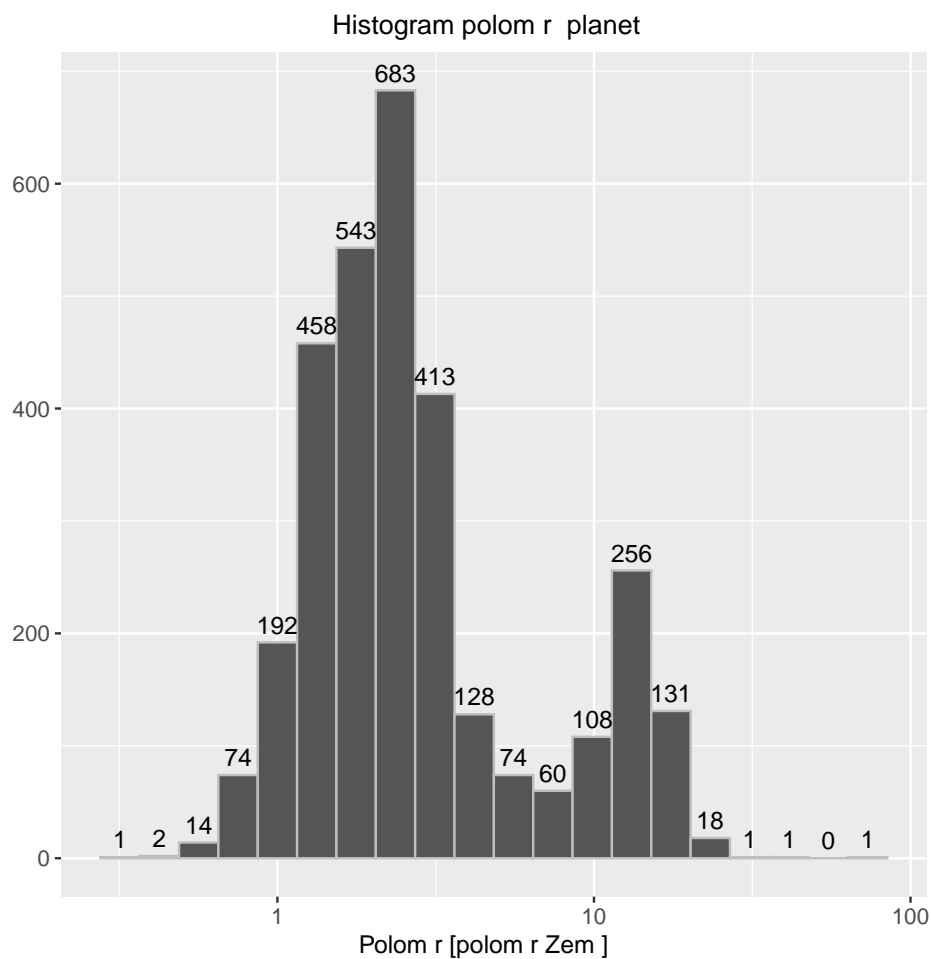
2 Zkoumané údaje

2.1 Způsob objevení



Protože exoplanety nelze pozorovat přímo, jsou k jejich objevování využívány různé techniky nepřímého pozorování. Počet planet objevených jednotlivými metodami je zobrazen v histogramu výše.

2.2 Poloměr planety



Poloměr planet je uváděn v poloměrech Země. Histogram četností planet v jednotlivých velikostních kategoriích má tvar bimodálního rozdělení. To je způsobeno tím, že terestrické (kamenné) planety podobné Zemi jsou často mnohem menší, než velcí plynní obři typu Jupiter.

Všechny planety byly následně rozděleny do všeobecně uznávaných velikostních kategorií dle následující tabulky:

Min	0.34
Qu1	1.57
Median	2.34
Mean	4.23
Qu3	3.60
Max	77.34

	Mercurian	Earth	Neptunian	Jovian
Poloměr [poloměry Země]	< 0,7	0,7-2	2-6	> 6
Počet	25	1224	1313	596

Table 1: Rozdělení planet do kategorií dle velikosti

2.3 Vzdálenost od Země

Min	4.24
Qu1	501.19
Median	1542.21
Mean	2069.64
Qu3	2816.64
Max	27710.00

Vzdálenost planet od Země je udávána ve světelných letech (zkratka ly). Světelný rok je vzdálenost, kterou světlo ve vakuu urazí za jeden rok, tedy přibližně $9.46 \cdot 10^{12}$ km.

Nejbližší exoplaneta vůbec, ve vzdálenosti 4,2 ly od Země, obíhá kolem nejbližší hvězdy Slunci – Proxima Centauri. Většina planet se ale nachází ve vzdálenosti 1 000 až 3 000 ly.

2.4 Hmotnost planety

Hmotnost planet je vyjádřena v hmotnostech Země. Nejvyhledávanějšími planety jsou ty, jejichž hmotnost je podobná Zemi.

Hmotnost malých planet je obtížné určit, a proto je u více jak poloviny (2 448) hmotnost neznámá. Z tohoto důvodu je 75 % změřených planet více jak 28krát hmotnějších, než Země.

Min	0.02
Qu1	28.38
Median	273.33
Mean	798.17
Qu3	807.26
Max	17668.17

2.5 Velká poloosa

Min	0.00
Qu1	0.06
Median	0.12
Mean	6.56
Qu3	0.69
Max	2500.00

Velká poloosa je průměrem minimální a maximální vzdálenosti planety od své hvězdy. Je udávána v astronomických jednotkách (AU), která je odvozena od velké poloosy Země (149 597 870 km).

Opět platí, že nejvyhledávanější hodnota je rovna 1, protože pak na planetě budou Zemi podobnější podmínky. Záleží však i na typu hvězdy.

2.6 Perioda oběhu planety

Perioda oběhu planety udává, za kolik pozemských dní planeta oběhne kolem své hvězdy. 75 % planet má dobu oběhu menší než 42 dní a pouze malé procento planet větší, než 1 000 dní.

Planety byly dále rozděleny do kategorií dle následující tabulky v závislosti na jejich periodě oběhu:

Min	0.09
Qu1	4.49
Median	11.87
Mean	2306.82
Qu3	42.60
Max	7300000.00

	Lava	Hot	Normal	Cold
Doba oběhu [dny]	< 5	5-50	50-500	> 500
Počet	1107	1958	540	381

Table 2: Rozdělení planet do kategorií dle periody oběhu

2.7 Počet planet v systému

Min	1.00
Qu1	1.00
Median	1.00
Mean	1.78
Qu3	2.00
Max	8.00

V jedné soustavě je zřídka pouze 1 planeta, avšak z důvodu omezených podmínek pozorování je i přesto více jak polovina planet v systému samotných.

Z druhé strany existuje 1 systém se 7 planetami a 1 systém s 8 planetami.

2.8 Poloměr hvězdy

Poloměr rodičovských hvězd planet je určován v poloměrech Slunce.

50 % objevených planet obíhá kolem hvězdy, jejíž poloměr se pohybuje mezi 0,8 až 1,26 poloměry Slunce.

Min	0.01
Qu1	0.80
Median	0.98
Mean	1.54
Qu3	1.26
Max	71.23

2.9 Hmotnost hvězdy

Min	0.01
Qu1	0.81
Median	0.97
Mean	1.00
Qu3	1.13
Max	23.56

Hmotnost hvězd je udávána v násobcích hmotnosti Slunce.

Stejně jako u poloměru, i hmotnost velkého množství hvězd je podobná Slunci. Konkrétně 50 % hvězd, kolem kterých je objevena alespoň 1 exoplaneta, má hmotnost v rozmezí 0,81 až 1,13 hmotností Slunce.

2.10 Povrchová teplota hvězdy

Povrchová teplota hvězdy je udávána v Kelvinech. Nejchladnější hvězdy mají červenou barvu a čím je jejich teplota větší, tím více jdou do žluté, bílé a nakonec i modré.

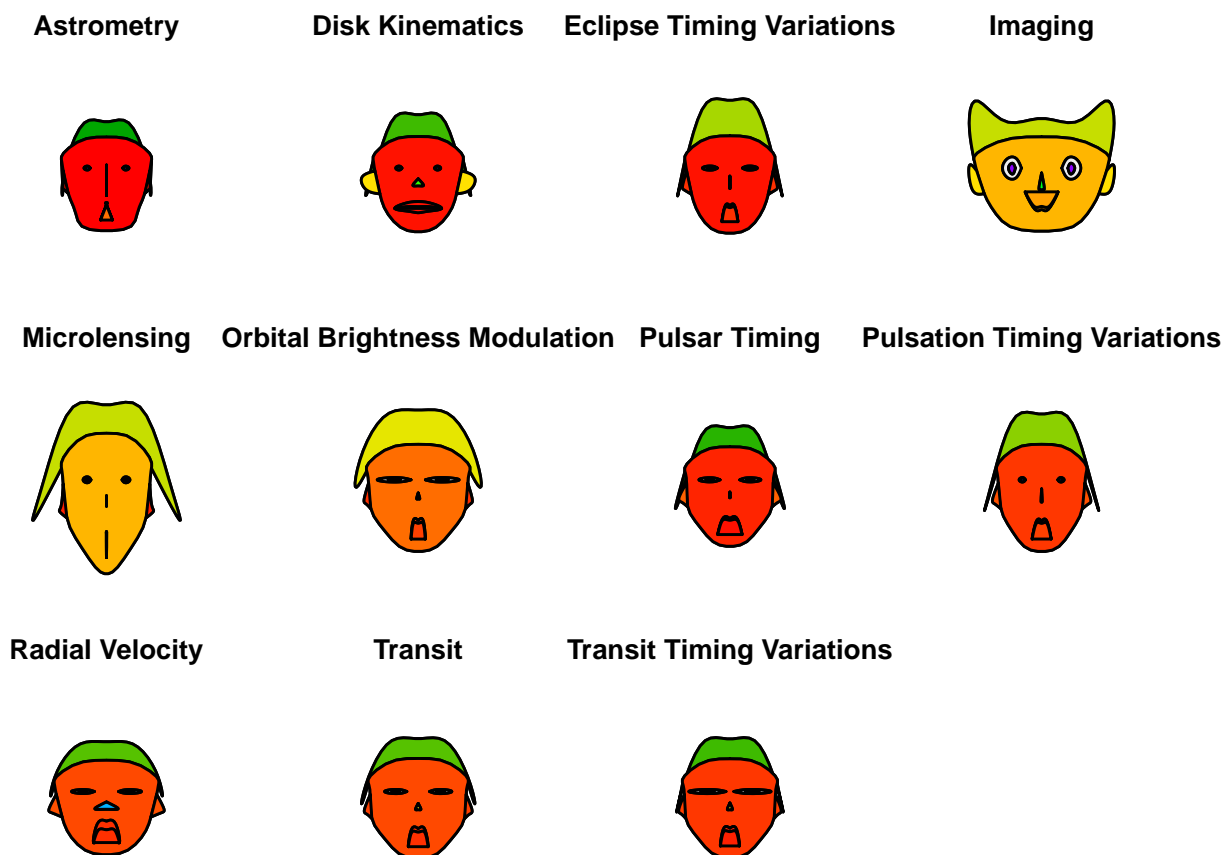
V tabulce níže je rozdělení hvězd do spektrálních kategorií dle jejich teploty. Je patrné, že nejvíce exoplanet bylo objeveno u žlutých hvězd (G) podobných Slunci.

Min	575.00
Qu1	5020.00
Median	5599.50
Mean	5492.44
Qu3	5923.00
Max	57000.00

	M	K	G	F	A	B	O
Teplota [tisíce K]	< 3,5	3,5-5	5-6	6-7,5	7,5-11	11-25	> 25
Počet	110	839	2134	763	16	1	9

Table 3: Rozdělení hvězd do spektrálních tříd dle povrchové teploty

3 Chernoffovy tváře



Na Chernoffových tvářích jsou ukázány rozdíly průměrných hodnot planet mezi planetami objevenými jednotlivými metodami.

modified.item	Var	modified.item	Var	modified.item	Var
height of face	dist	smiling	smax	style of hair	rad
width of face	rad	height of eyes	per	height of nose	mass
structure of face	mass	width of eyes	pnum	width of nose	srad
height of mouth	srad	height of hair	stem	width of ear	smass
width of mouth	smass	width of hair	dist	height of ear	smax

Table 4: Vlastnosti tváří popisující vlastnosti planet

4 Korelační matice

	dist	mass	rad	smax	per	pnum	smass	srad	stem
dist	1	0.17	0.2	-0.01	-0.04	-0.11	0.31	0.34	0.25
mass	0.17	1	0.25	0.16	0.15	-0.19	0.22	0.16	0.09
rad	0.2	0.25	1	0	0.03	-0.64	0.58	0.43	0.18
smax	-0.01	0.16	0	1	0.98	0.03	-0.05	0.04	-0.05
per	-0.04	0.15	0.03	0.98	1	0.01	-0.05	0.02	-0.05
pnum	-0.11	-0.19	-0.64	0.03	0.01	1	-0.36	-0.24	-0.18
smass	0.31	0.22	0.58	-0.05	-0.05	-0.36	1	0.72	0.36
srad	0.34	0.16	0.43	0.04	0.02	-0.24	0.72	1	0.17
stem	0.25	0.09	0.18	-0.05	-0.05	-0.18	0.36	0.17	1

Table 5: Korelační matice

V korelační matici lze vidět lineární závislost mezi velkou poloosou a dobou oběhu planety, která vychází z Keplerových zákonů. Slabá záporná korelace je i mezi počtem planet a poloměrem planety. Čím větší planeta je, tím spíše bude v soustavě méně planet.

Velice slabá korelace (0,58) je viditelná mezi hmotností hvězdy a poloměrem planety. U hmotnějších hvězd byly častěji pozorovány větší planety. Naopak vzdálenost planety od Země nezávisí na žádném jiném údaji – planety všech možných vlastností jsou objevovány v různých vzdálenostech od Země.

5 Chí-kvadrát test

Do kontingenční tabulky byla umístěna četnost planet v různých velikostních kategoriích v závislosti na periodě oběhu kolem hvězdy. Kvůli nedostatku objevených ledových planet se poslední sloupec nebude promítat do testu nezávislosti.

	Lava	Hot	Normal	Cold
Mercurian	14	11	0	0
Earth	507	669	48	0
Neptunian	133	975	200	5
Jovian	388	130	56	11

Table 6: Četnost planet dle velikosti a periody oběhu hvězdy

- H_0 : Vzdálenost planet od mateřské hvězdy nezávisí na jejich velikosti.
- H_1 : Existuje spojitost mezi vzdáleností planety od hvězdy a její velikostí.

Pearson's Chi-squared test

```
data: size.vs.period[, 1:3]
X-squared = 719.24, df = 6, p-value < 2.2e-16
```

Protože p-value je dostatečně malá, velikost planety a její perioda oběhu jsou závislé proměnné.

6 T-test

Cílem t-testu je demonstrovat na příkladu závěr chí-kvadrát testu v předchozí kapitole – velikost planety a perioda oběhu kolem hvězdy jsou závislé proměnné. Pro tento účel byl proveden náhodný výběr o velikosti 1 / 10 z velikostních kategorií planet Neptunian a Earth. T-test testuje doby periody planet v těchto náhodných výběrech.

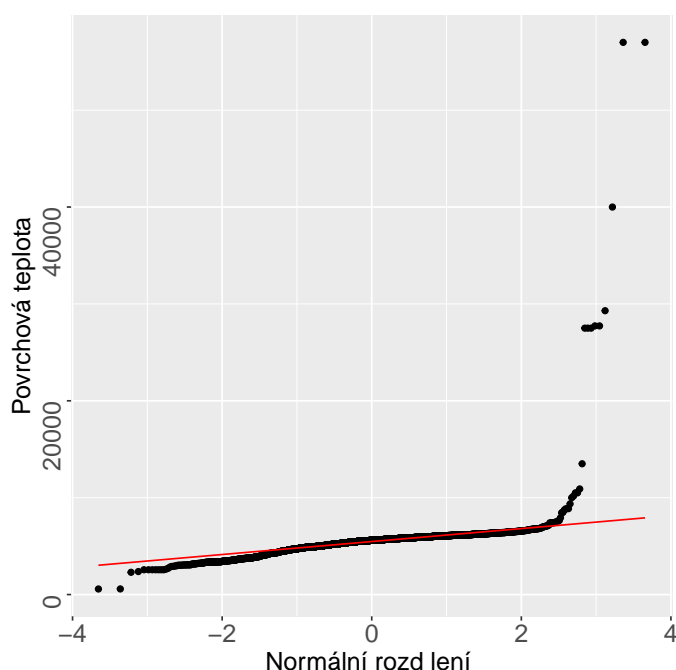
Two Sample t-test

```
data: earths$per and neptunians$per
t = -3.9297, df = 251, p-value = 0.00011
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -33.19149 -11.02941
sample estimates:
mean of x mean of y
 11.86790  33.97836
```

Protože p-value je menší než 0,05, můžeme nulovou hypotézu na 95% hladině významnosti zamítnout a říct, že střední hodnota period oběhů planet ve skupinách Earth a Neptunian není stejná. V konfidenčním intervalu je uvedeno, že střední hodnota periody oběhu planet typu Earth je o několik jednotek až desítek (dní) menší, než u planet ze skupiny Neptunian.

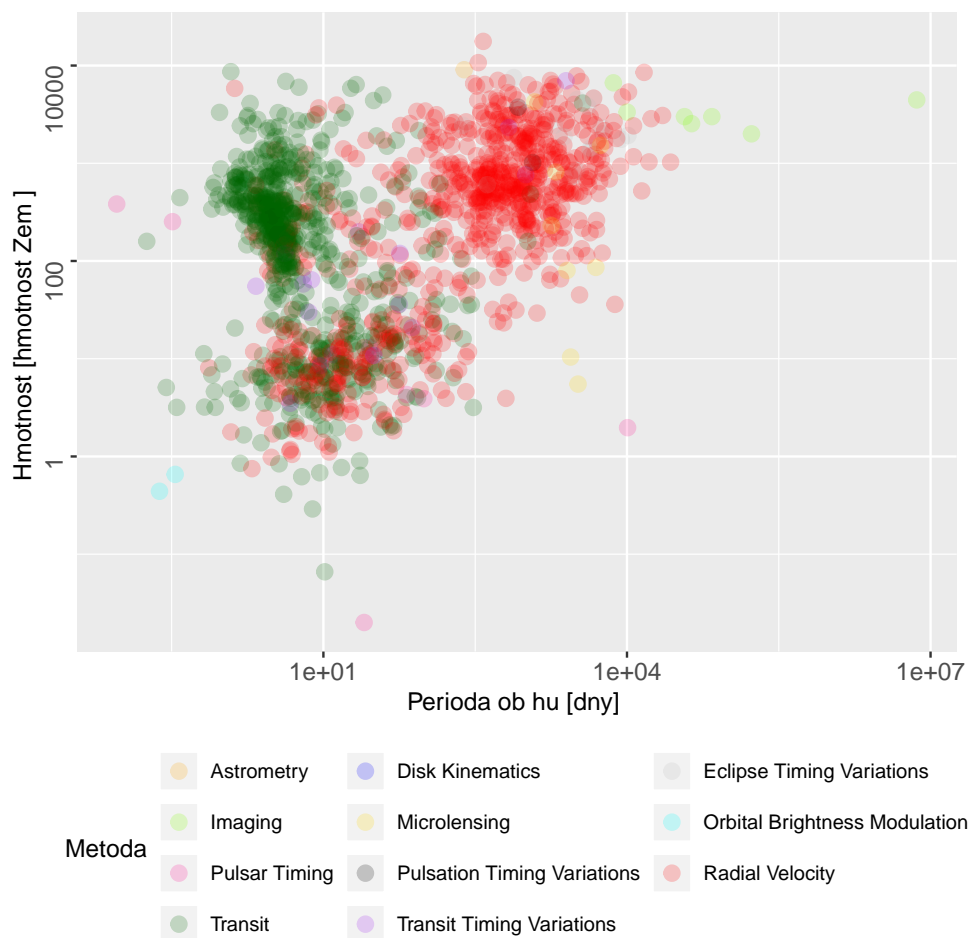
7 QQplot

V kapitole 2.10 bylo řečeno, že nejvíce objevených exoplanet obíhá kolem hvězd s podobnou teplotou jako naše Slunce. Do QQ grafu bude proto zanesena právě povrchová teplota hvězd, u níž se bude zjišťovat, zda vyhovuje normálnímu rozdělení.



Většina bodů leží na požadované přímce, a proto lze říci, že rozdělení teplot je blízké normálnímu rozdělení. Výjimkou jsou extrémy, které se od normálního rozdělení vzdalují.

8 Analýza rozptylu



Graf výše ukazuje rozdílné hmotnosti a periody planet při rozdílných metodách objevení. Hmotné hvězdy blízko své hvězdy byly objevovány převážně metodou Transit, kdežto hmotné hvězdy daleko od své hvězdy metodou Radial Velocity.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
stype	3	1.562e+08	5.207e+07	10.415	8.12e-07	***
meth	4	2.116e+10	5.289e+09	1058.009	< 2e-16	***
dtype	1	1.438e+05	1.438e+05	0.029	0.865	
stype:meth	4	3.371e+03	8.430e+02	0.000	1.000	
stype:dtype	3	2.319e+05	7.730e+04	0.015	0.997	
meth:dtype	1	4.200e+01	4.200e+01	0.000	0.998	
stype:meth:dtype	1	2.750e+02	2.750e+02	0.000	0.994	
Residuals	3128	1.564e+10	4.999e+06			

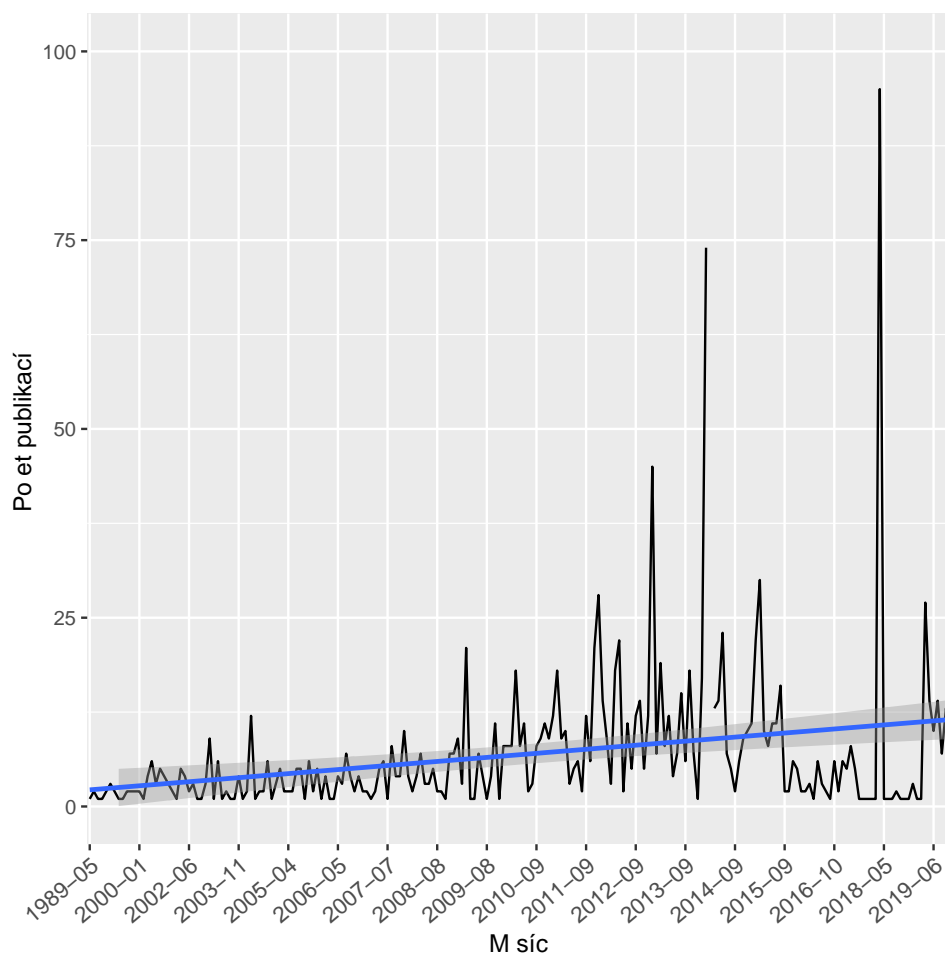
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

958 observations deleted due to missingness

Ve výsledcích analýzy rozptylu je uvedeno, že na dobu periody skutečně má vliv metoda objevení (meth) i velikost (stype). Naproti tomu vzdálenost od Země (dtype) nikoliv.

9 Regrese

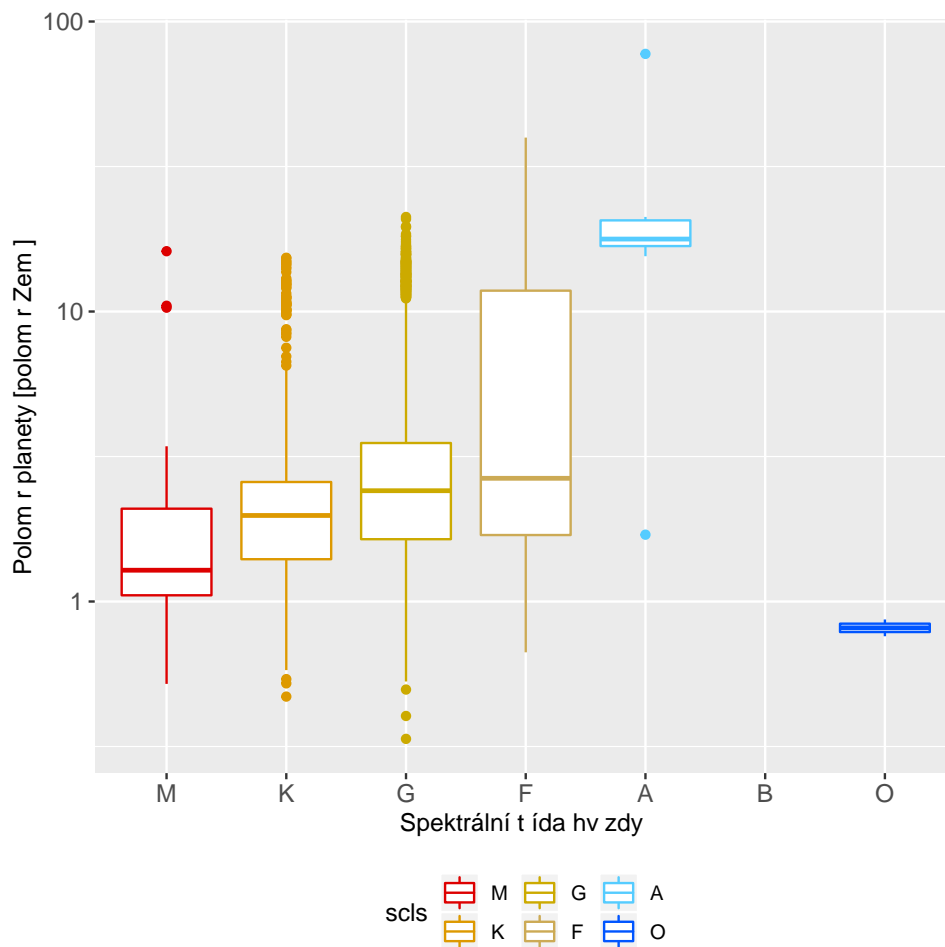
Cílem regrese je určit, zda počet publikací objevů exoplanet v průběhu posledních let klesá nebo stoupá.



Měsíční počet publikací exoplanet stále stoupá. V grafu je vyznačen i 95% interval spolehlivosti.

10 Boxplot

V závěrečné kapitole budou na boxplotu zobrazeny statistiky poloměrů planet v závislosti na spektrální třídě hvězdy, kolem které obíhají.



Největší planety obíhají kolem modrobílých hvězd (třída A). Nejmenší objevená planeta obíhá kolem žluté hvězdy (třída G). Kolem hvězd třídy B byla objevena pouze jedna planeta, a proto není box pro tuto skupinu zobrazen.