

VYUŽITÍ REGRESNÍ ANALÝZY V PŘESTUPNÍCH UZLECH NA ŽELEZNIČNÍ SÍTI

REGRESSION ANALYSIS RECOVERY FOR TRANSFER JUNCTIONS ON THE RAILWAY NETWORK

Pavel Drdla ¹

Anotace: Příspěvek se věnuje problematice využití regresní analýzy při zjišťování charakteristik trendů závislostí jednotlivých veličin v přestupních uzlech spolu s vymezením pravděpodobnosti dodržení přestupních vazeb.

Klíčová slova: regresní analýza, Beta rozdělení pravděpodobnosti, přestupy, doprava, statistika

Summary: The paper deals with a problematic of the use of regression analysis in determining the characteristics of trends of individual variables in the dependency transfer junctions, together with the likelihood of compliance with the definition of transfer connections.

Key words: regression analysis, Beta probability distribution, transfer, transport, statistics

1. ÚVOD

V běžném provozu v praktické rovině a v různých publikačních výstupech v teoretické rovině je řešena problematika přestupních vazeb v přestupních uzlech železniční osobní dopravy. Pozornost bývá věnována množině přestupů v různých uzlech a toto je pak dále analyzováno.

Méně často se ale věnuje pozornost otázce přestupních vazeb u periodicky se opakujících přestupů v daných přestupních uzlech osobní železniční dopravy. V tomto případě je možné provést lepší porovnání zjištěných hodnot u takto se opakujících procesů, než je tomu u relativně nesouvisejících případů.

Tento příspěvek se právě této oblasti věnuje a jeho cílem je pokusit se o zobecnění zjištěných informací a pro toto uvést příslušné závěry.

2. BETA ROZDĚLENÍ (4 PARAMETRICKÉ)

Pokud se věnuje pozornost studiu odborné literatury (například [1]), tak se zjistí, že přestupní vazby vykazují 4 parametrické Beta rozdělení pravděpodobnosti. Toto rozdělení pravděpodobnosti lze charakterizovat následovně [2]:

¹ doc. Ing. Pavel Drdla, Ph.D., Univerzita Pardubice, Dopravní fakulta Jana Pernera, Katedra technologie a řízení dopravy, Studentská 95, CZ-53210 Pardubice, tel.: +420 466036204, fax: +420 466306094, e-mail: pavel.drdla@upce.cz

Náhodná veličina X má Beta rozdělení s parametry a , b , α , β , jestliže její hustota pravděpodobnosti má tvar

$$f(x) = \frac{(x-a)^{\alpha-1} (b-x)^{\beta-1}}{B(\alpha, \beta)(b-a)^{\alpha+\beta-1}} \quad \text{pro } a < x < b, \alpha > 0, \beta > 0, b > a$$

$B(\alpha, \beta)$ je Beta funkce, definovaná jako

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1} (1-x)^{\beta-1} dx \quad \text{pro } \alpha > 0, \beta > 0$$

Střední hodnota a rozptyl mají tvar

$$E(X) = a + \frac{ba}{\alpha + \beta} \quad D(X) = \frac{\alpha\beta(b-a)^2}{(\alpha + \beta)^2 (\alpha + \beta + 1)^2}$$

Pro potřeby výpočtu hustoty pravděpodobnosti $f(x)$ je tedy potřeba zjistit pro každou přestupní vazbu výše uvedené 4 parametry (vše v minutách):

a – velikost minimálního zpoždění nebo maximálního náskoku prvního vlaku

b – velikost maximálního zpoždění prvního vlaku

α – rozdíl mezi průměrným zpožděním prvního vlaku a veličinou a

β – rozdíl mezi veličinou b a průměrným zpožděním prvního vlaku

Je tedy zřejmé, že pro každou dvojici přípojných vlaků (první vlak – druhý vlak) je kromě parametrů a a b třeba zjistit hodnotu průměrného zpoždění prvního vlaku.

Jako další nezbytný parametr bude sloužit i doba na přestup mezi prvním a druhým vlakem. Pro potřeby následující metody byla stanovena nezbytná podmínka, že jako zachovaný přípoj lze ještě považovat situaci, kde opožděný první vlak přijede nejpozději v čas pravidelného odjezdu druhého vlaku podle jízdního řádu (JŘ), tedy velikost jeho zpoždění nebude větší než je uvedená doba na přestup. Je samozřejmě možné si stanovit v rámci přestupních vazeb jinou podmínku, metodický postup ale zůstane zachován.

3. ZJIŠTĚNÍ VSTUPNÍCH HODNOT

Pro potřeby tohoto příspěvku byla provedena velmi podrobná analýza všech přestupních stanic na síti Českých drah s výjimkou uzlu Praha, který je zcela specifický [3]. V jednotlivých uzlech byly zjištěny [4] požadované vstupní hodnoty za předpokladu, že k přípojným vazbám dochází periodicky, nejčastěji s časovým odstupem 1 nebo 2 hodiny. Výpočet pravděpodobnosti přípoje (Beta rozdělení) byl proveden s využitím MS Excel. Výsledky analýzy byly zpracovány do velmi obsáhlé tabulky, kterou není možné z důvodu značného rozsahu zde uveřejnit (tabulka má kromě záhlaví 623 řádků) – v případě zájmu je k dispozici u autora. Jako příklad tohoto obsáhlého materiálu může posloužit tabulka č. 1 pro jeden vybraný přestupní uzel (5 řádků).

Tab. 1 – Výšek z tabulky s analýzou periodicky se opakujících přestupních vazeb

trať první	trať druhá	přestupní stanice	vlak první	vlak druhý	ujetá vzdálenost prvního vlaku [km]	ujetá vzdálenost od předchozího místa zastavení prvního vlaku [km]	doba jízdy prvního vlaku [min]	doba jízdy od předchozího místa zastavení prvního vlaku [min]
č. 340	č. 343	Veselí nad Moravou	Sp 1723	Os 2759	90	10	91	9
			Sp 1725	Os 2761				
			Sp 1727	Os 2765				
			Sp 1729	Os 2769				
			Sp 1731	Os 2707				

pokračování

příjezd prvního vlaku dle JŘ	odjezd druhého vlaku dle JŘ	minimální zpoždění (+) nebo maximální náskok (-) 1. vlaku [min]	průměrné zpoždění nebo náskok 1. vlaku [min]	maximální zpoždění 1. vlaku [min]	doba na přestup [min]	pravděpodobnost přípoje [%]
10:59	11:03	-1	1	5	4	97,25961
12:59	13:03	-2	0	8		99,66564
14:59	15:03	-2	1	27		99,61989
16:59	17:03	-2	1	18		94,83493
18:59	19:03	-2	5	123		95,37763

Zdroj: Autor, Knižní jízdní řád 2008/9

Protože jde o periodicky se opakující procesy ve stejném přestupním uzlu, lze pro další postup pracovat s průměrnou hodnotou vztahující se k tomuto uzlu. Z tohoto důvodu lze sestavit z původní velmi rozsáhlé tabulky sumarizovanou tabulku č. 2.

Tab. 2 – Sumarizovaná tabulka přestupních vazeb v jednotlivých uzlech

trať první	trať druhá	přestupní stanice	ujetá vzdálenost prvního vlaku [km]	ujetá vzdálenost od předchozího místa zastavení prvního vlaku [km]	doba jízdy prvního vlaku [min]	doba jízdy od předchozího místa zastavení prvního vlaku [min]	doba na přestup [min]	pravděpodobnost přípoje [%]
č. 020	č. 040	Chlumec nad Cidlinou	93	32	93	25	6	95,145
č. 040	č. 020	Chlumec nad Cidlinou	102	11	132	9	3	90,4037
č. 020	č. 231	Hradec Králové	120	27	120	21	14	100
č. 020	č. 231	Hradec Králové	62	4	79	5	11	100
č. 026	č. 047	Teplice nad Metují	32	7	46	8	3	55,1465
č. 010	č. 031	Pardubice	104	42	73	22	4	81,7974
č. 031	č. 010	Pardubice	161	2	182	3	4	95,8618
č. 040	č. 030	Stará Paka	84	3	100	4	5	77,02
č. 030	č. 064	Stará Paka	85	31	95	31	5	97,1474

trať první	trať druhá	přestupní stanice	ujetá vzdálenost prvního vlaku [km]	ujetá vzdálenost od předchozího místa zastavení prvního vlaku [km]	doba jízdy prvního vlaku [min]	doba jízdy od předchozího místa zastavení prvního vlaku [min]	doba na přestup [min]	pravděpodobnost připoje [%]
č. 040	č. 064	Stará Paka	84	4	100	4	9	99,2464
č. 040	č. 041	Ostroměř	75	10	103	10	5	98,1694
č. 040	č. 041	Ostroměř	59	5	72	5	4	89,1451
č. 070	č. 092	Neratovice	37	15	45	12	3	67,322
č. 070	č. 092	Neratovice	68	6	129	5	4	42,4881
č. 070	č. 072	Všetaty	43	6	52	6	3	44,6626
č. 071	č. 070	Mladá Boleslav	53	30	46	25	5	42,4747
č. 070	č. 071	Mladá Boleslav	61	10	86	11	7	67,5045
č. 071	č. 070	Mladá Boleslav	99	9	143	9	3	41,0038
č. 070	č. 080	Bakov nad Jizerou	84	9	93	9	2	2,4304
č. 231	č. 072	Lysá nad Labem	39	29	31	23	9	99,9383
č. 081	č. 080	Česká Lípa	54	20	54	21	3	82,2357
č. 080	č. 081	Česká Lípa	107	15	68	16	4	99,5223
č. 086	č. 036	Liberec	113	31	137	37	9	100
č. 090	č. 087	Lovosice	84	6	64	5	5	99,3072
č. 090	č. 114	Lovosice	45	22	35	16	5	93,2352
č. 090	č. 114	Lovosice	84	6	64	5	5	99,1539
č. 090	č. 081	Děčín	129	23	101	16	8	99,929
č. 090	č. 097	Lovosice	84	6	64	5	7	99,8717
č. 097	č. 090	Lovosice	45	22	35	16	7	96,8872
č. 120	č. 122	Hostivice	17	11	25	15	4	86,2398
č. 120	č. 124	Lužná u Rakovníka	64	13	79	13	3	27,4677
č. 120	č. 174	Rakovník	73	9	92	12	3	29,5538
č. 130	č. 123	Most	156	12	123	10	4	94,2446
č. 160	č. 170	Plzeň	157	23	169	29	11	99,9078
č. 160	č. 170	Plzeň	157	23	169	29	14	99,9937
č. 160	č. 190	Plzeň	157	23	169	29	9	99,5677
č. 010	č. 024	Ústí nad Orlicí	154	15	115	13	10	99,8629
č. 010	č. 024	Ústí nad Orlicí	154	15	121	13	4	99,9032
č. 010	č. 272	Česká Třebová	164	10	132	9	13	100
č. 170	č. 184	Planá u Marianských Lázní	184	31	178	34	4	66,7845
č. 190	č. 203	Strakonice	76	17	61	12	13	99,5535
č. 190	č. 201	Ražice	90	14	74	10	2	36,466
č. 190	č. 192	Čičenice	106	7	89	6	12	99,4965
č. 190	č. 197	Čičenice	106	7	89	6	13	99,8111
č. 200	č. 203	Březnice	100	18	95	16	5	67,2339
č. 220	č. 201	Tábor	103	39	96	31	14	99,9948

trať první	trať druhá	přestupní stanice	ujetá vzdálenost prvního vlaku [km]	ujetá vzdálenost od předchozího místa zastavení prvního vlaku [km]	doba jízdy prvního vlaku [min]	doba jízdy od předchozího místa zastavení prvního vlaku [min]	doba na přestup [min]	pravděpodobnost připoje [%]
č. 220	č. 201	Tábor	66	20	56	17	7	92,3915
č. 220	č. 202	Tábor	103	54	96	45	11	99,8655
č. 220	č. 202	Tábor	66	20	56	17	4	68,0024
č. 231	č. 230	Kolín	134	8	117	7	12	100
č. 230	č. 231	Kolín	195	11	164	9	7	90,1689
č. 230	č. 236	Čáslav	82	9	80	8	6	99,1223
č. 260	č. 240	Brno	255	22	194	22	9	99,9324
č. 240	č. 250	Brno	372	43	394	46	13	99,8435
č. 240	č. 260	Brno	372	43	394	46	7	95,0449
č. 250	č. 340	Brno	181	34	138	24	4	76,9522
č. 250	č. 251	Žďár nad Sázavou	169	17	152	13	13	100
č. 250	č. 254	Šakvice	34	34	22	22	4	83,4674
č. 250	č. 254	Šakvice	147	6	113	4	3	40,1479
č. 250	č. 255	Zaječí	141	8	108	6	14	99,9964
č. 250	č. 255	Zaječí	40	6	27	4	5	88,0174
č. 250	č. 246	Břeclav	59	11	45	10	10	99,8724
č. 260	č. 262	Skalice nad Svitavou	38	16	34	10	2	42,1697
č. 260	č. 261	Svitavy	181	15	148	11	7	99,8513
č. 270	č. 275	Olomouc	106	101	55	46	9	99,9992
č. 270	č. 290	Olomouc	106	101	55	46	9	99,9992
č. 301	č. 270	Olomouc	100	20	93	26	3	30,6048
č. 270	č. 301	Olomouc	106	101	55	46	8	99,9952
č. 270	č. 310	Olomouc	112	21	144	15	5	86,4861
č. 270	č. 310	Olomouc	106	101	55	46	11	100
č. 330	č. 300	Přerov	159	15	119	11	4	82,3961
č. 330	č. 300	Přerov	82	15	92	11	5	91,5738
č. 270	č. 280	Hranice na Moravě	117	29	102	17	4	59,2724
č. 270	č. 278	Suchdol nad Odrou	138	29	115	17	15	99,9933
č. 270	č. 278	Suchdol nad Odrou	42	12	37	7	12	99,9948
č. 270	č. 277	Suchdol nad Odrou	138	29	115	17	16	99,9979
č. 270	č. 277	Suchdol nad Odrou	42	12	37	7	13	99,9985
č. 270	č. 278	Suchdol nad Odrou	138	29	115	17	15	99,9966
č. 270	č. 278	Suchdol nad Odrou	42	12	37	7	12	99,9941
č. 270	č. 279	Studénka	150	12	46	8	4	76,1082
č. 321	č. 270	Ostrava- Svinov	117	28	150	22	8	99,1885
č. 321	č. 270	Ostrava- Svinov	28	15	23	10	4	75,8171
č. 321	č. 270	Ostrava-	28	15	23	10	8	96,2727

trať první	trať druhá	přestupní stanice	ujetá vzdálenost prvního vlaku [km]	ujetá vzdálenost od předchozího místa zastavení prvního vlaku [km]	doba jízdy prvního vlaku [min]	doba jízdy od předchozího místa zastavení prvního vlaku [min]	doba na přestup [min]	pravděpodobnost přípoje [%]
		Svinov						
č. 270	č. 321	Ostrava-Svinov	13	5	16	6	7	99,6764
č. 310	č. 311	Valšov	64	7	80	8	2	53,7585
č. 310	č. 315	Opava východ	89	4	121	5	5	43,7544
č. 310	č. 315	Opava východ	116	4	159	5	5	69,1077
č. 310	č. 317	Opava východ	89	4	121	5	4	25,7606
č. 310	č. 317	Opava východ	116	4	159	5	4	61,138
č. 330	č. 341	Staré Město u Uherského Hradiště	68	18	48	9	12	99,9125
č. 330	č. 341	Staré Město u Uherského Hradiště	113	14	85	8	5	88,3307
č. 330	č. 342	Moravský Písek	82	14	58	8	6	89,9823
č. 330	č. 342	Moravský Písek	99	20	76	10	18	100
č. 340	č. 342	Bzenec	80	81	3	4	16	99,9881
č. 340	č. 343	Veselí nad Moravou	90	91	10	9	4	97,3515
č. 340	č. 343	Veselí nad Moravou	90	91	10	9	14	99,954
č. 340	č. 343	Veselí nad Moravou	19	4	27	5	5	99,1655
č. 340	č. 343	Veselí nad Moravou	19	4	27	5	15	100

Zdroj: Autor, Knižní jízdní řád 2008/9

4. ZJIŠŤOVÁNÍ ZÁVISLOSTI PARAMETRŮ

Vlastním zjištěním vstupních parametrů pochopitelně metodický postup nekončí, protože je třeba zjistit závislostní vztah mezi jednotlivými parametry. Nabízí se následující varianty pro zjištění závislosti:

1. zjistit závislost 2 parametrů, zde doby na přestup a pravděpodobnosti přípoje (Varianta 1)
2. zjistit závislost 3 parametrů (Varianta 2), zde doby na přestup, pravděpodobnosti přípoje a
 - a) ujeté vzdálenosti prvního vlaku,
 - b) ujeté vzdálenosti od předchozího místa zastavení prvního vlaku,
 - c) doby jízdy prvního vlaku,
 - d) doby jízdy od předchozího místa zastavení prvního vlaku.

4.1 Varianta 1

Při sumarizaci hodnot bylo zjištěno, že některé z nich (vycházející logicky z reálného provozu) vybočují z dále uvedeného trendu vývoje závislosti (extrapolace funkce). Z uvedeného důvodu byly zjišťovány závislosti pro obě možnosti, tedy buď s využitím všech hodnot nebo s vyjmutím (redukováním) okrajových hodnot. Vše je patrné z tabulky č. 3.

Tab. 3 – Pravděpodobnosti přípojů při zohledňování všech nebo bez redukce přestupů

doba na přestup [min]	pravděpodobnost přípoje [%]	
	všechny přestupy	redukované přestupy
2	33,7062	44,1314
3	50,8549	56,4419
4	76,9607	79,8051
5	82,6785	88,7653
6	94,7499	95,1450
7	92,6746	96,2703
8	98,8464	99,9290
9	99,8119	99,9061
10	99,8677	99,8677
11	99,9433	99,9693
12	99,8796	99,9754
13	99,8678	99,9306
14	99,9878	99,9962
15	99,9966	99,9966
16	99,9930	99,9930
17	100,0000	100,0000
18	100,0000	100,0000

Zdroj: Autor

V obou případech, kdy byl zjišťován pomocí MS Excel typ spojnice trendu, vycházel nejlépe polynomický, a to 5. stupně. Z trendů je rovněž patrné, že u vyšších přestupních časů se pravděpodobnost přestupu limitně blíží 100%, tedy prakticky jistotě.

V příslušných rovnicích trendů (rovnících regrese) mají použité veličiny následující význam:

x – doba na přestup,

y – pravděpodobnost přípoje.

4.1.1 Zohlednění všech přestupů

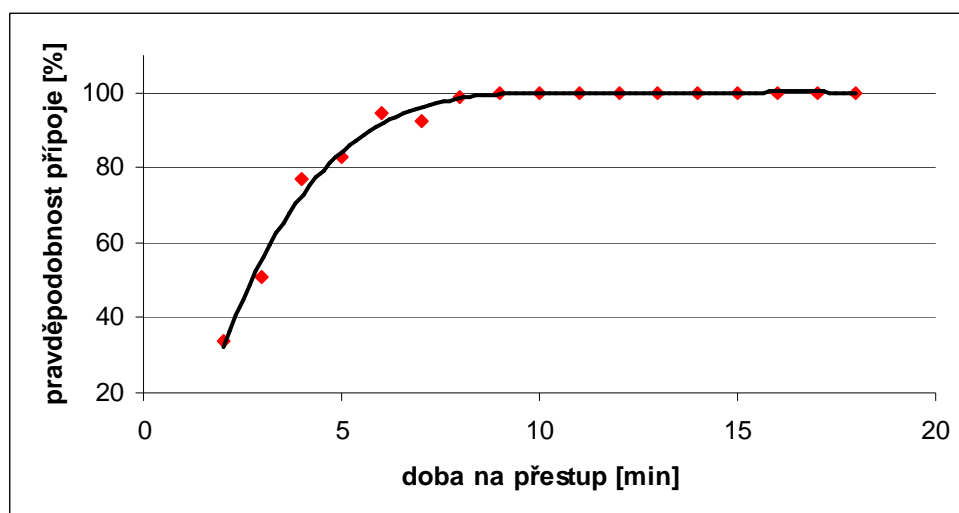
Rovnice spojnice trendu (rovnice regrese) má tento tvar:

$$y = 0,0001x^5 - 0,0114x^4 + 0,3963x^3 - 6,3979x^2 + 48,867x - 43,036$$

Hodnota spolehlivosti má tuto velikost (protože není rovna 1, může u grafického vyjádření a při aplikaci rovnice regrese vycházet v některých případech pravděpodobnost přípoje větší jak 100%, což samozřejmě v reálném případě není možné):

$$R^2 = 0,9884$$

Graficky lze rovnici regrese znázornit způsobem uvedeným na obrázku č. 1.



Zdroj: Autor

Obr. 1 – Grafické znázornění regrese při zohlednění všech přestupů

4.1.2 Redukce přestupů

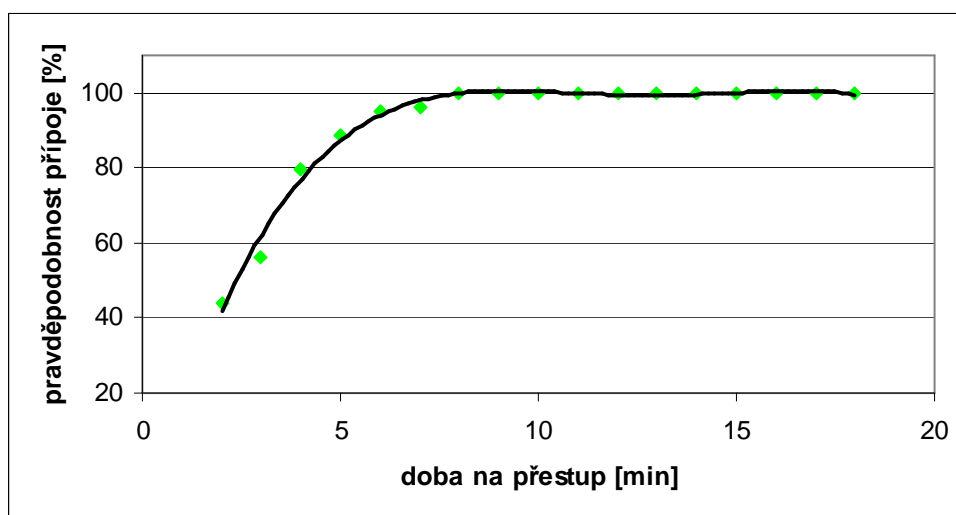
V tomto případě vyšla následující rovnice spojnice trendu (rovnice regrese):

$$y = -0,0001x^5 + 0,0001x^4 + 0,1643x^3 - 4,1482x^2 + 37,792x - 18,319$$

Hodnota spolehlivosti vychází následovně:

$$R^2 = 0,9878$$

Grafické znázornění rovnice regrese je na obrázku č. 2.

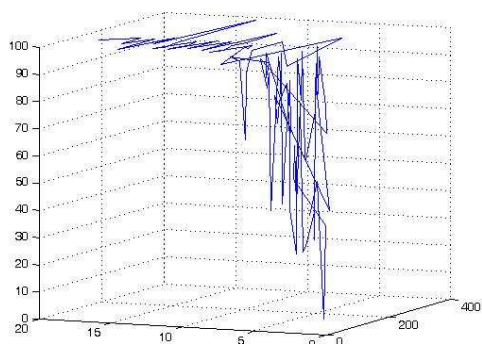


Zdroj: Autor

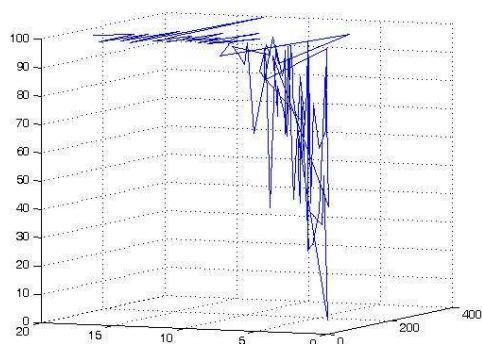
Obr. 2 – Grafické znázornění regrese při redukci některých přestupů

4.2 Varianta 2

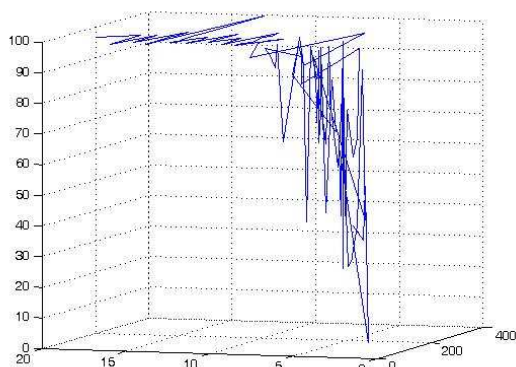
V tomto případě se pro extrapolaci hodnot využil MATLAB. Ve všech čtyřech případech se zohlednily všechny hodnoty bez redukce. Z následujících graficky provedených výstupů na obrázku č. 3 je patrné, že v žádném z případů není možné vyvodit závěr ohledně trendu vývoje, protože vstupní hodnoty jsou zcela neuspořádané. U jednotlivých grafických výstupů je na ose x (vodorovná osa) doba na přestup [min], na ose y (svislá osa) pravděpodobnost přestupu [%] a na ose z (hloubka grafu) variabilní veličina [min] uvedená pod grafickým výstupem obrázku č. 3.



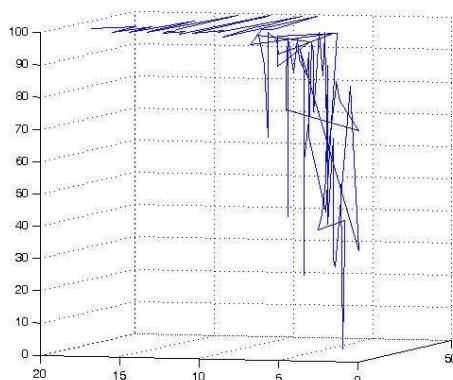
ujetá vzdálenost prvního vlaku



ujetá vzdálenost od předchozího místa zastavení prvního vlaku



doba jízdy prvního vlaku



doba jízdy od předchozího místa zastavení prvního vlaku

Zdroj: Autor + Ing. Strádal

Obr. 3 – Grafické znázornění závislosti doby na přestup a pravděpodobnosti přestupu na variabilní proměnné

Z uvedeného důvodu nemá smysl se výsledky Varianty 2 dále zabývat. Směrodatné budou výsledky uvedené ve Variantě 1.

5. VYHODNOCENÍ A ZÁVĚR

Pokud se výše uvedené má vyhodnotit, vyjde se z výsledků prezentovaných v rámci Varianty č. 1.

Z obrázků č. 1 a č. 2 je zcela jednoznačně patrný průběh regresních křivek pro obě možnosti, tedy využívání všech hodnot nebo při redukci některých velmi odchylných hodnot. Uvedené křivky mají regresivní průběh limitně se blížící 100% pravděpodobnosti přestupů.

Také je třeba vzít v potaz, že i v železniční dopravě se připouští určitá procentuální míra rizika. Obecně se používá zásada, že u železniční nákladní dopravy se operuje s 5% rizikem (tedy 95% spolehlivostí) a u železniční osobní dopravy se operuje s 1% rizikem (tedy 99% spolehlivostí).

Pokud se vyjde právě z tohoto doporučení, je třeba určit, při jaké přestupní době je splněna 99% spolehlivost přestupu. S využitím rovnic regrese vychází tyto výsledky:

- a) při zohlednění všech hodnot je tato spolehlivost dodržena u přestupních dob delších jak 8 minut,
- b) při redukci některých odchylných hodnot pak u přestupních dob větších jak 7,5 minuty.

Grafikon vlakové dopravy a tudíž i jízdní řády jsou sestavovány s přesností na 0,5 minuty, proto jsou uvedeny tyto zaokrouhlené hodnoty.

Samozřejmě je možno si stanovit i jiné podmínky pro zachování přípojných vazeb, ovšem metodicky postup zůstane zachován.

Na úplném konci příspěvku je třeba poděkovat panu Ing. Oktaviánu Strádalovi, Ph.D. za cenné rady k regresní analýze a za pomoc při vyhodnocování dat pomocí aplikace MATLAB.

POUŽITÁ LITERATURA

- [1] ČERNÝ, J., KLUVÁNEK, P.: *Základy matematickej teórie dopravy*. Bratislava: Veda, 1989. 279 s. ISBN 80-224-0099-8
- [2] *Pravděpodobnostní rozdělení v MS Excel* [online]. c2006 [cit. 2009-04-07]. Dostupné z <<http://panda.hyperlink.cz/cestapdf/pdf06c6/mares.pdf>>
- [3] *Železniční jízdní řád 2008/2009*. Praha: Správa železniční dopravní cesty, 2008.
- [4] *Babitrón – Zpoždění vlaků* [online]. c2007 [cit. 2009-02-26]. Dostupné z <<http://kam.mff.cuni.cz/~babilon/zponline>>

Příspěvek vznikl za podpory Institucionálního výzkumu „Teorie dopravních systémů“ (MSM 0021627505) Univerzity Pardubice.

Recenzenti: Ing. Karel Greiner, Ph.D.
Univerzita Pardubice, DFJP, Katedra informatiky v dopravě
doc. RNDr. Antonín Tuzar, CSc.
ČVUT v Praze, FD, Ústav řízení dopravních procesů a logistiky