

# KOMPARÁCIA PROGNÓZY PREPRAVNÝCH VÝKONOV ŽELEZNIČNEJ OSOBNEJ DOPRAVY S VYUŽITÍM KAUZÁLNYCH MODELOV

## THE PROGNOSIS OF RAIL PASSENGER TRANSPORT PERFORMANCES THROUGH CAUSUAL MODELS

Anna Dolinayová,<sup>1</sup> Darina Chlebíková<sup>2</sup>

---

*Anotácia: Prognóza vývoja prepravných výkonov osobnej dopravy sa v súčasnosti využíva predovšetkým v dopravných podnikoch pri plánovacom procese. Prognózu je možné zostaviť na základe matematicko-štatistických metód použitím rôznych modelov, pričom výsledné hodnoty môžu byť veľmi odlišné, v závislosti od použitého modelu. V príspevku sú na základe vybraných makroekonomických veličín vytvorené krátkodobé prognózy vývoja prepravných výkonov osobnej železničnej dopravy pomocou jednoduchých regresných modelov a viacnásobného lineárneho regresného modelu. Príspevok sa ďalej zaoberá porovnaním výsledkov týchto modelov a ich interpretáciou pre dopravnú prax.*

*Kľúčové slová: prognóza, osobná železničná preprava, regresný model, verifikácia modelu*

*Summary: Nowadays the transport companies use the prognosis of rail passenger transport performances especially in planning processes. The prognosis is possible to formate on the basis of mathematic–statistical models whereas their results can be very different in dependence on applied model. In the paper there is constructed prognosis of rail passenger transport performances for a short term with the aid of single regressive models and multiply linear regressive model. The paper deals with the comparison results of these models and their interpretation for transport practice.*

*Key words: prognosis, passenger railway transport, regressive model, verification of model*

### 1. ÚVOD

Prognóza vývoja prepravných výkonov železničnej osobnej dopravy je dôležitým vstupným predpokladom napr. pri zostavení rozpočtu podniku na budúce obdobie, ale aj pri plánovaní dotácií zo štátneho rozpočtu, keďže ide o službu vo verejnom záujme. Prognózu je možné uskutočniť rôznymi metódami od objektívnych matematicko-štatistických metód až po využitie niektorých zo subjektívnych, t.j. expertných metód prognózovania. Na zostavenie krátkodobej prognózy vývoja prepravných výkonov sú vhodné kauzálne prognostické modely, ktoré využívajú nástroje regresnej analýzy. Aj keď ide o objektívne metódy

---

<sup>1</sup> Ing. Anna Dolinayová, PhD., Žilinská univerzita v Žiline, Fakulta PEDAS, Katedra železničnej dopravy, Univerzitná 8215/1, 010 26 Žilina, Tel.: +421415133424, Fax: +421415655816, E-mail: [anna.dolinayova@fpedas.uniza.sk](mailto:anna.dolinayova@fpedas.uniza.sk)

<sup>2</sup> Ing. Darina Chlebíková, PhD., Žilinská univerzita v Žiline, Fakulta PEDAS, Katedra ekonomiky, Univerzitná 8215/1, 010 26 Žilina, Tel.: +421415133222, E-mail: [darina.chlebikova@fpedas.uniza.sk](mailto:darina.chlebikova@fpedas.uniza.sk)

prognózovania, samotná prognóza je závislá od príčinných faktorov, ktoré sú brané do úvahy a od použitia konkrétneho regresného modelu.

## 2. KAUZÁLNY PROGNOTICKÝ MODEL

Zostavenie prognózy na základe kauzálneho prognostického modelu sa uskutočňuje modelovaním časového radu prognózovaného javu pomocou nezávislých faktorov. Ak je do úvahy braný len jeden nezávislý faktor, ide o jednoduchý regresný model, v prípade použitia viacerých nezávislých faktorov, hovoríme o viacnásobnom regresnom modeli.

### 2.1 Jednoduchý regresný model

Jednoduchý regresný model umožňuje zostavenie prognózy závislej premennej na základe vývoja jednej nezávislej premennej, pričom môže byť použitý lineárny, kvadratický, exponenciálny, resp. iný druh regresného modelu. Samotný výber modelu často závisí od skúseností spracovateľa prognózy, a to predovšetkým vtedy, keď je potvrdená verifikácia viacerých jednoduchých regresných modelov. V praxi je možné použiť ten model, ktorý vykazuje najvyššiu hodnotu koeficientu determinácie.

Postup zostavenia prognózy možno zhrnúť do nasledovných krokov:

- získanie časového radu vývoja závisle a nezávisle premennej,
- kvantifikácia regresného modelu,
- verifikácia regresného modelu,
- získanie alebo odhad vývoja nezávisle premennej na prognózované obdobie,
- vytvorenie bodovej a intervalovej predpovede vývoja nezávisle premennej.

Časový rad vývoja závisle premenných je dostupný z vnútropodnikových údajov. Pri získavaní časového radu údajov nezávisle premenných je potrebné zamerať sa na tie, ktoré sú v príčinnom vzťahu k nezávisle premennej, ale zároveň, pri ktorých je možné zistiť z externých zdrojov prognózy ich vývoja, resp. je možné zostaviť prognózu na základe modelu trendu danej nezávisle premennej.

Kvantifikáciu jednoduchého regresného modelu je možné uskutočniť výpočtom alebo využitím niektorého zo štatistických programov. Niektoré jednoduchšie regresné modely, ktoré sú však na krátkodobé prognózovanie vývoja prepravných výkonov železničnej osobnej dopravy postačujúce, je možné zostaviť využitím štandardného prostredia tabuľkového procesora MS EXCEL, ktorý zároveň pri každom type regresnej funkcie určí aj koeficient determinácie ( $R^2$ ).

Verifikáciou modelu sa zisťuje, či daný matematicko-štatistický model dostatočne popisuje skúmaný jav, t.j.: [4]

- štatistická významnosť parametrov modelu, t.j. či parametre veličín vystihujú významnosť premenných,
- štatistická významnosť modelu ako celku, t.j. či model ako celok vystihuje interpretáciu všetkých modelových veličín.

Testovanie štatistickej významnosti parametrov modelu sa uskutočňuje pomocou *t*-štatistiky:

$$t_i = \frac{\hat{b}_i}{s_{b_i}} \quad (1)$$

kde:  $\hat{b}_i$  - regresný parameter

$s_{b_i}$  - štandardná chyba parametra:

$$s_{b_i} = \frac{s}{\sqrt{\sum_{i=1}^n (x_{ti} - \bar{x}_{ti})^2}} \quad (2)$$

Ak je hodnota *t*-štatistiky vyššia ako tabuľková hodnota na zvolenej hladine významnosti  $\alpha$  a pri  $n - (k+1)$  stupňoch voľnosti, parameter  $b_i$  je štatisticky významný.

Verifikácia modelu ako celku sa uskutočňuje pomocou *F*-štatistiky, ktorá má tvar:

$$F_r = \frac{R^2 / k}{(1 - R^2) / [n - (k + 1)]} \quad (3)$$

Ak je hodnota  $F_r$  vyššia ako tabuľková hodnota *F*-štatistiky na hladine významnosti  $\alpha$  a pri  $k$  a  $n - (k+1)$  stupňoch voľnosti, model je štatisticky významný a je možné použiť ho na tvorbu prognózy.

Bodová predpoveď na nasledujúce obdobie sa vypočíta dosadením prognózy vývoja nezávisle premennej na nasledujúce obdobie do rovnice regresnej krivky. Pri konštrukcii intervalovej predpovede je potrebné najskôr vypočítať štandardnú chybu predpovede podľa vzorca [4]:

$$s_{n+i} = \sqrt{s^2 + s^2 \left( \frac{1}{n} + \frac{(x_{n+i} - \bar{x}_t)^2}{\sum_{t=1}^n (x_t - \bar{x}_t)^2} \right)} \quad (4)$$

Intervalová predpoveď nezávisle premennej na zvolenej hladine významnosti  $\alpha$  má tvar:

$$\hat{y}_{n+i} \pm t_{\alpha; [n-(k+1)]} s_{n+i} \quad (5)$$

## 2.2 Viacnásobný regresný model

Pomocou viacnásobného regresného modelu sa modeluje vývoj nezávisle premennej vzhľadom na vývoj viacerých závisle premenných. V praxi sa najčastejšie používa viacnásobný lineárny regresný model, pretože interpretácia vplyvov jednotlivých nezávisle premenných na výslednú hodnotu závisle premennej je pri použití iného ako lineárneho modelu zložitá. Konštrukciu viacnásobného lineárneho regresného modelu je možné uskutočniť výpočtom alebo využitím štatistického softvéru. Postup tvorby predpovede je analogický ako pri jednoduchom regresnom modeli, ale pri verifikácii modelu je potrebné, okrem štatistickej významnosti parametrov modelu a modelu ako celku, zistiť, či medzi nezávislými premennými nie je závislosť, t.j. multikolarita. Diagnostiku multikolarity je možné uskutočniť viacerými nástrojmi. Jedným z nich je matica koeficientov korelácie pre všetky dvojice nezávisle premenných. V modeli je významná multikolarita, ak niektorý korelačný koeficient v matici nadobúda hodnoty vyššie ako 0,8.

Bodová predpoveď vývoja nezávisle premennej sa získa podobne ako v prípade jednoduchej regresie dosadením predpokladaných budúcich hodnôt nezávisle premenných do regresného modelu. Intervalová predpoveď sa konštruuje podľa vzťahu [3]:

$$\hat{y}_{n+i} \pm t_{\alpha;[n-(k+1)]} s_{n+i} \quad (6)$$

kde:  $t_{\alpha;[n-(k+1)]}$  – tabuľková hodnota  $t$ -štatistiky

$s_{n+i}$  – očakávaná stredná chyba prognózy:

$$s_{n+i} = s \sqrt{1 + X'_{n+i} (X'X)^{-1} X_{n+i}} \quad (7)$$

kde:  $X_{n+i}$  – vektor vysvetľujúcich premenných modelu

$X$  – matica hodnôt vysvetľujúcich premenných

$(X'X)^{-1}$  – inverzná matica k súčinu pôvodnej a transponovanej matice nezávislých premenných modelu

Predpokladané hodnoty nezávisle premenných je možné získať z prognóz spracovávaných relevantnými inštitúciami, príp. zistiť ich pravdepodobnú hodnotu na nasledujúce obdobie zostavením modelu časového radu danej nezávisle premennej a jej extrapolácie na budúce obdobie.

### 3. KONŠTRUKCIA KRÁTKODOBEJ PROGNÓZY JEDNODUCHÝM REGRESNÝM MODELOM

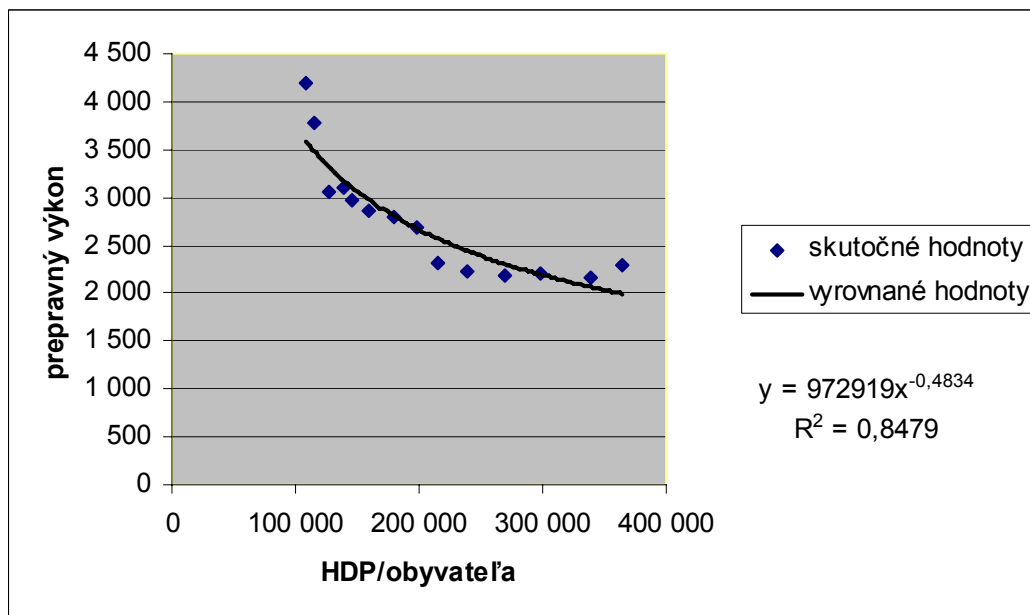
Krátkodobú prognózu prepravných výkonov jednoduchým regresným modelom uskutočníme na základe modelovania prepravných výkonov železničnej osobnej dopravy parciálne v závislosti od HDP na osobu, reálnej hrubej mesačnej mzdy a počtu osobných automobilov. Potrebné štatistické údaje dokumentuje nasledujúca tabuľka.

Tab. 1 - Vývoj HDP na osobu, priemernej reálnej mzdy, počtu os. automobilov a prepravných výkonov železničnej osobnej dopravy v SR v rokoch 1995 až 2008

Rok	HDP/obyvateľ/a [Sk v stálych cenách predchádz. roka]	priemerná hrubá reálna mesačná mzda [Sk]	počet osobných automobilov [ks]	prepravné výkony železničnej osobnej dopravy [mil. oskm]
1995	108 406	6 483	1 015 794	4202
1996	115 236	7 681	1 058 425	3769
1997	127 119	8 663	1 135 914	3057
1998	139 038	9 333	1 196 109	3092
1999	145 967	9 591	1 236 396	2968
2000	158 760	10 058	1 274 244	2870
2001	180 270	11 462	1 292 843	2805
2002	198 360	13 065	1 326 891	2682
2003	215 827	13 144	1 356 185	2316
2004	238 948	14 638	1 197 030	2228
2005	269 352	16 808	1 303 704	2182
2006	299 024	17 917	1 333 749	2213
2007	339 334	19 582	1 433 926	2165
2008	364 566	20 780	1 544 888	2296

Zdroj: Spracované z literatúry [5], [6], [7], [8]

Jednoduché regresné modely boli vytvorené využitím štandardného prostredia tabuľkového procesora MS EXCEL. V nasledujúcich grafoch sú zobrazené tie modely, ktoré najlepšie vystihujú danú závislosť.



Zdroj: Autori

Obr. 1 - Regresný model závislosti prepravných výkonov železničnej osobnej dopravy od HDP na obyvateľa

Overenie modelu sme uskutočnili pomocou *F-štatistiky*, podľa vzorca (3), ktorej hodnota je 66,9, čo je viac ako tabuľková hodnota na hladine významnosti  $\alpha = 0,05$  a pri 1 a 13 stupňoch voľnosti ( $F=4,7472$ ). Na základe *t-štatistiky* podľa vzorca (1) a (2) sme zistili, že parameter  $b_1$  je štatisticky významný. Z uvedeného vyplýva, že vytvorený regresný model je možné použiť na zostavenie krátkodobej prognózy.

Predpokladaná hodnota HDP/obyvateľa je podľa prognóz Štatistického úradu SR ako aj INFOSTATU na úrovni 96,5 % HDP z predchádzajúceho obdobia, t.j. približne 351 800 Sk/obyvateľa. Dosadením tejto hodnoty do rovnice regresnej krivky zobrazenej na obr. 1 získame bodovú predpoveď prepravných výkonov vo výške 2 027,7 mil. oskm. Intervalovú predpoveď sme zostavili na základe vzorcov (4) a (5). Potrebné výpočty sú uvedené v nasledujúcej tabuľke.

Tab. 2 - Výpočty na zostavenie intervalovej prognózy prepravných výkonov

Rok	x	y	$x - \bar{x}$	$(x - \bar{x})^2$	$\hat{y}$	$y - \hat{y}$	$(y - \hat{y})^2$
1995	108 406	4 202	-98 752	9751958631	3 582,06	620	384325,5
1996	115 236	3 769	-91 921	8449475343	3 477,60	291	84799,87
1997	127 119	3 057	-80 039	6406228327	3 316,67	-260	67427,54
1998	139 038	3 092	-68 120	4640323880	3 176,04	-84	7063,466
1999	145 967	2 968	-61 190	3744238970	3 102,24	-134	18020,6
2000	158 760	2 870	-48 398	2342351170	2 978,78	-109	11833,76
2001	180 270	2 805	-26 888	722957575,4	2 801,33	4	13,49577
2002	198 360	2 682	-8 797	77388173,23	2 674,77	7	52,23301
2003	215 827	2 316	8 669	75154770,43	2 567,85	-252	63430,21
2004	238 948	2 228	31 791	1010643635	2 444,58	-217	46908
2005	269 352	2 182	62 195	3868177961	2 307,06	-125	15641,15
2007	299 024	2 213	91 866	8439404178	2 193,41	20	383,6206
2007	339 334	2 165	132 176	17470500790	2 063,34	102	10333,88
2008	364 566	2 296	157 408	24777307397	1 993,03	303	91789,48
Celkom	x	x	x	91776110802	x	x	802022,8

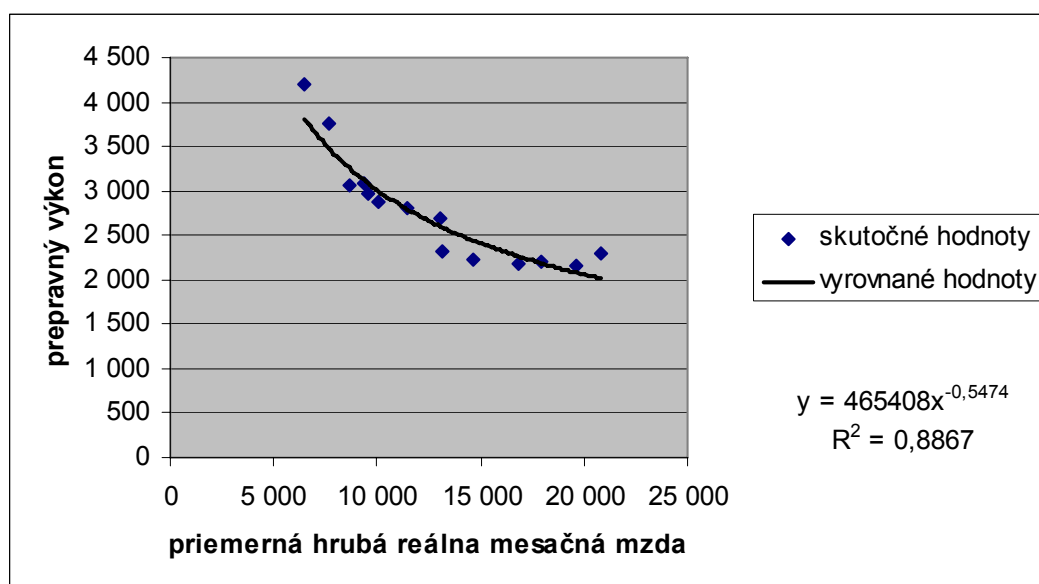
Zdroj: Vlastný výpočet

Hodnota štandardnej chyby predpovede je:

$$s_{n+1} = \sqrt{66\,835,2 + 66\,835,2 \left( \frac{1}{14} + \frac{351\,800 - 207\,158}{91\,776\,110\,802} \right)} = 294,67$$

Tabuľková hodnota *t*-štatistiky na hladine významnosti  $\alpha = 0,05$  a pri 13 stupňoch voľnosti je 1,7823. Na základe bodovej predpovede, štandardnej chyby predpovede a tabuľkovej hodnoty *t*-štatistiky sme zostavili intervalovú predpoveď na nasledujúce obdobie podľa vzorca (5), t.j. s pravdepodobnosťou 95 % budú prepravné výkony železničnej osobnej dopravy v roku 2009 podľa zostaveného regresného modelu v intervale (1 502,4; 2 552,9).

Pri ostatných nezávislých premenných, vzhľadom na rovnakú metodiku výpočtu, uvádzame len grafické zobrazenie modelu, výsledné hodnoty verifikácie modelu a prognózu na nasledujúce obdobie.

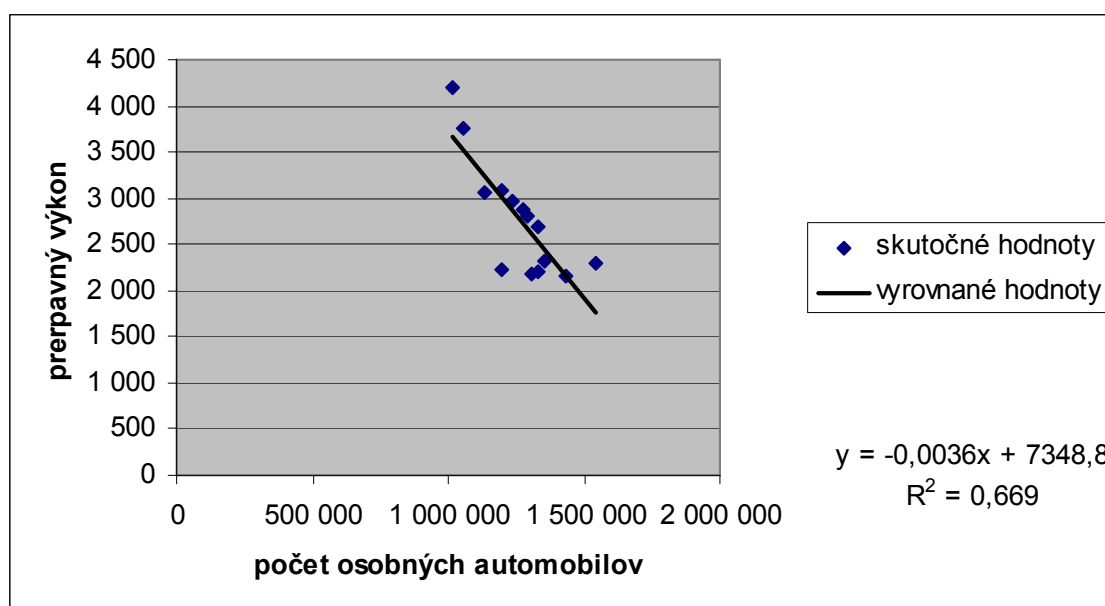


Zdroj: Autori

Obr. 2 - Regresný model závislosti prepravných výkonov železničnej osobnej dopravy od hrubej reálnej mesačnej mzdy

Ako vidieť z obr. 2, koeficient determinácie nadobúda vysokú hodnotu, to znamená, že model je štatisticky významný a taktiež je štatisticky významný parameter  $b_1$  a model je možné použiť na zostavenie krátkodobej prognózy. Predpoveď vývoja reálnej hrubej mesačnej mzdy podľa prognóz Štatistického úradu SR, ako aj INFOSTATU je 102,6 % priemernej hrubej reálnej mzdy predchádzajúceho obdobia.

Bodová predpoveď (vytvorená dosadením hodnoty 21 320 do rovnice regresnej krivky z obr. 2) výšky prepravných výkonov je 1 987,3 mil. oskm. Použitím štandardnej chyby predpovede, ktorej hodnota je 242,4 a tabuľkovej hodnoty *t*-štatistiky sme zostrojili intervalovú predpoveď prepravných výkonov. S pravdepodobnosťou 95 % budú prepravné výkony železničnej osobnej dopravy vzhľadom na zmenu priemernej reálnej hrubej v intervale (1 555,2; 2 419,3).



Zdroj: Autori

Obr. 3 - Regresný model závislosti prepravných výkonov železničnej osobnej dopravy od počtu osobných automobilov

V prípade skúmania závislosti prepravných výkonov železničnej osobnej dopravy od počtu osobných automobilov je výška koeficientu determinácie podstatne nižšia ako pri predchádzajúcich dvoch nezávisle premenných. Kritická hodnota  $F_r$  je však stále dostatočne vysoká ( $F_r=23,9$ ) a regresný model ako celok je štatisticky významný. Podobne aj štandardná chyba parametra  $b_1$  je vyššia ako tabuľková hodnota *t*-štatistiky.

Na zostavenie krátkodobej prognózy sme využili odhad počtu osobných automobilov na nasledujúce obdobie pomocou modelu trendu vo výške 1 449 190. Bodový odhad prepravných výkonov železničnej osobnej dopravy na základe zostaveného jednoduchého regresného modelu je 2 131,7 mil. oskm, intervalový s pravdepodobnosťou 95 % (1 401,4; 2 862,0).

#### 4. KONŠTRUKCIA KRÁTKODOBEJ PROGNÓZY VIACNÁSOBNÝM LINEÁRNÝM REGRESNÝM MODELOM

Pri zostavení prognózy pomocou viacnásobného lineárneho regresného modelu sme využili údaje v tabuľke 1. Keďže v matici korelačných koeficientov bol vysoký korelačný koeficient HDP/osobu a priemerných hrubých reálnych príjmov (jeho hodnota bola 0,92), nezávislý parameter HDP/osobu sme z regresného modelu vylúčili a zostavili viacnásobný lineárny regresný model závislosti vývoja prepravných výkonov železničnej osobnej dopravy od priemerných hrubých reálnych príjmov a počtu osobných automobilov.

Použitím metódy najmenších štvorcov sme dostali nasledujúce tri normálové rovnice:



$$38\,845 = 14b_0 + 179\,205b_1 + 17\,706\,098b_2$$

$$465\,541\,429 = 179\,205b_0 + 2\,564\,220\,628b_1 + 233\,886\,945\,495b_2$$

$$48\,195\,966\,827 = 17\,706\,098b_0 + 233\,886\,945\,495b_1 + 22\,651\,005\,670\,982b_2$$

Riešením uvedených troch normálových rovníc sme získali odhady parametrov  $b_0$ ,  $b_1$ , a  $b_2$  a zostavili nasledujúci viacnásobný lineárny regresný model:

$$y = 5\,478,7 - 0,082x_1 - 0,0013x_2$$

Hodnota koeficientu determinácie  $R^2$  pre uvedený regresný model je 0,76, kritická hodnota Fisherovho kritéria  $F_r = 17,29$ , čo je vyššie ako tabuľková hodnota  $F$ -štatistiky (3,9823) a model je možné použiť pre krátkodobú prognózu. Bodovú predpoveď sme zostavili dosadením pravdepodobných budúcich hodnôt nezávisle premenných, uvedených v predchádzajúcej kapitole, do rovnice modelu:

$$y = 5\,478,7 - 0,082 \cdot 21\,320 - 0,0013 \cdot 1\,449\,190 = 1\,833,14$$

Pred konštrukciou intervalovej predpovede sme najskôr vypočítali očakávanú strednú chybu prognózy podľa vzorca (7):

$$X'_{n+1} = [1 \quad 21\,320 \quad 149\,190]$$

$$X = \begin{bmatrix} 14 & 179\,204,897 & 17\,706\,098 \\ 179\,204,897 & 2\,564\,220\,628 & 2,33887E \cdot 10^{11} \\ 17\,706\,098 & 2,33887 \cdot 10^{11} & 2,2651 \cdot 10^{13} \end{bmatrix}$$

$$(X'X)^{-1} = \begin{bmatrix} 14,02 & 3,4 \cdot 10^{-3} & -1,45 \cdot 10^{-5} \\ 3,4 \cdot 10^{-3} & 1,49 \cdot 10^{-8} & -4,21 \cdot 10^{-10} \\ -1,45 \cdot 10^{-5} & -4,21 \cdot 10^{-10} & 1,57 \cdot 10^{-11} \end{bmatrix}$$

$$X'_{n+1} (X'X)^{-1} X_{n+1} = 0,34072$$

$$s_{15} = 332,5 \cdot \sqrt{1 + 0,34072} = 113,29$$

Intervalová predpoveď na hladine významnosti  $\alpha = 0,05$  je:

$$1833,14 \pm 332,5 \cdot 0,34072 \cdot 1,7959 \approx 1\,833,14 \pm 203,4555$$

S pravdepodobnosťou 95 % budú prepravné výkony železničnej osobnej dopravy na základe zostaveného viacnásobného lineárneho regresného modelu v intervale (1 629,7; 2 336,6) mil. oskm.

## 5. ZÁVER

Manažéri podnikov zaoberajúcich sa osobnou dopravou by mali vedieť relevantne využiť spracované prognózy vývoja prepravných výkonov, pretože na základe ich pravdepodobného vývoja zostavujú operatívne, príp. aj strednodobé plány a rozpočty. V praxi sa často používajú na zostavenie prognóz budúceho vývoja jednoduché modely časových radov alebo jednoduché regresné modely. Z uvedených krátkodobých prognóz vývoja prepravných výkonov železničnej osobnej dopravy vyplýva, že aj pri použití objektívnych matematicko–štatistických metód je veľmi dôležité, aký model sa na vytvorenie prognózy použije. Vyššiu vypovedaciu schopnosť má viacnásobný regresný model ako väčší počet jednoduchých regresných modelov. Na zvýšenie spoľahlivosti prognózy je potrebné matematicko–štatistický model doplniť ešte niektorou z expertných metód prognózovania, ktoré môžu poskytnúť predovšetkým kvalitatívne výpovede o prognózovanom jave.

## POUŽITÁ LITERATÚRA

- [1] Aktualizovaná prognóza vývoja vybraných ukazovateľov na rok 2009. Online. Dostupné na: <http://portal.statistics.sk/showdoc.do?docid=18788>
- [2] Chlebíková, D.: Ekonomická štatistika, Žilinská univerzita v Žiline, EDIS 2009, ISBN 978-80-554-0058-7
- [3] Kvasnička, M., Vašíček, O.: Úvod do analýzy časových řad. Masarykova univerzita v Brne. 2001, sylaby predmetu EMM. Dostupné na: [www.is.muni.cz](http://www.is.muni.cz), 173 s.
- [4] Mikolaj, J., Kľučka, J., Vančo, B.: Plánovanie a prognostika, Multiprint Košice 2005, 242 str., ISBN 80-969148-3-9
- [5] Revidované údaje HDP za roky 1996-2005. Online. Dostupné na: <http://portal.statistics.sk/showdoc.do?docid=9617>
- [6] Štatistická ročenka Slovenskej republiky 2000. VEDA, vydavateľstvo SAV, Bratislava, 2000. ISBN 80-224-0742-9
- [7] Štatistická ročenka Slovenskej republiky, 2005. VEDA, vydavateľstvo SAV, Bratislava, 2005. ISBN 80-224-0882-4
- [8] Štatistická ročenka Slovenskej republiky, 2008. VEDA, vydavateľstvo SAV, Bratislava, 2008. ISBN 978-80-224-1053-3

Recenzenti: doc. Ing. Viera Bartošová, PhD.  
Žilinská univerzita v Žiline, FPEDaS, Katedra ekonomiky  
doc. Ing. Rudolf Kampf, Ph.D.  
Univerzita Pardubice, DFJP, Katedra dopravného managementu, marketingu  
a logistiky