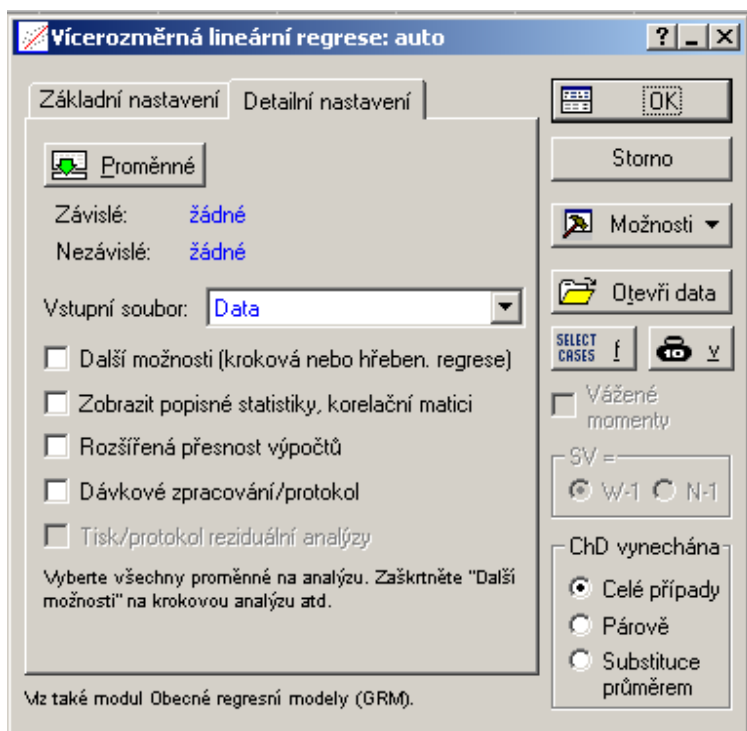


Téma 9: Vícenásobná regrese

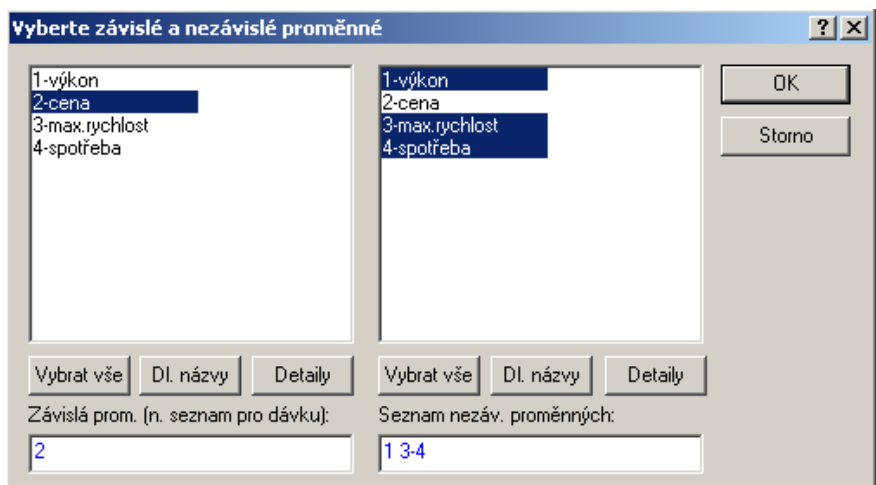
1) Vytvoření modelu

V menu *Statistika* zvolíme nabídku *Vícerozměrná regrese*. Aktivujeme kartu *Detailní nastavení* – viz obr.1. Nastavíme *Proměnné* tak, že v příslušném okně – viz. obr.2, v levém sloupci označíme závisle proměnnou a v pravém sloupci označíme nezávisle proměnné. Při nesouvhládnutém výběru nezávisle proměnných použijeme tlačítko CTRL Kdybychom chtěli vytvořit prostý lineární model, označíme v pravém sloupci pouze jedinou nezávisle proměnnou.

Obr. 1

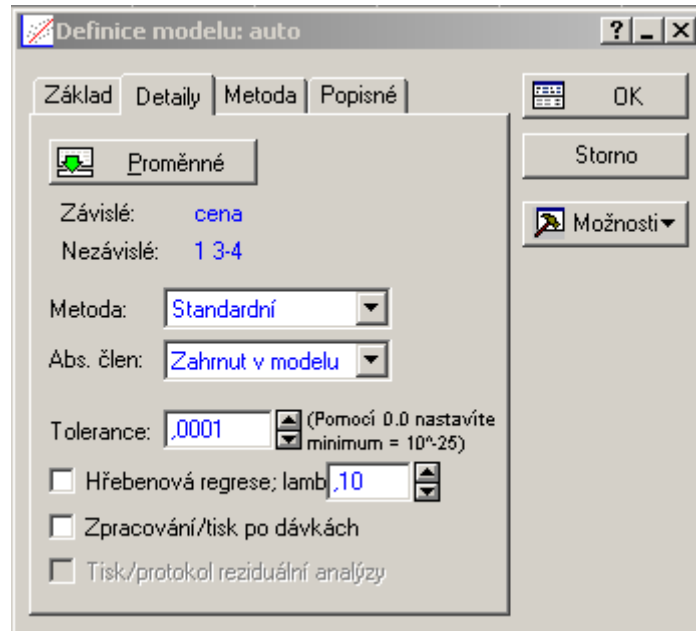


Obr. 2



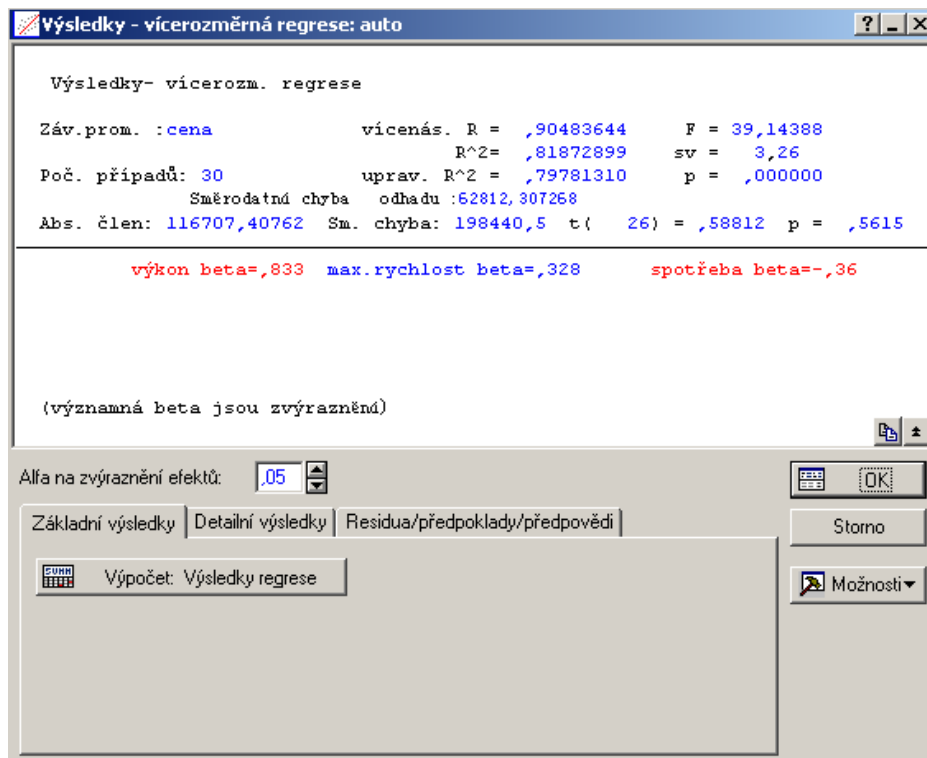
Výběr potvrdíme tlačítkem **OK**. Poté zajistíme, aby na kartě **Detailní nastavení** (obr.1) byla zaškrtnuta volba **Další možnosti (kroková nebo hřeben. regrese)**. Stiskneme – li opět tlačítko **OK**, otevře se okno **Definice modelu** – viz. obr.3. Zkontrolujeme, zda je nastavena **Standardní metoda** a zda je **absolutní člen zahrnut v modelu**.

Obr. 3



Stiskneme-li potřeť tlačítko tlačítko **OK**, otevře se okno **Výsledky – vícerozměrná regrese** –viz obr.4.

Obr.4



Vybereme kartu **Základní výsledky**. V ní stiskneme tlačítko **Výpočet: výsledky regrese**. Objeví se výstupní sestava - viz obr. 5.

Obr. 5

Výsledky regrese se závislou proměnnou : cena (auto)						
R= ,90483644 R^2= ,81872899 uprav. R^2= ,79781310						
F(3,26)=39,144 p<,00000 Směrod. chyba odhadu : 62812,						
N=30	Beta	Sm.chyba beta	B	Sm.chyba B	t(26)	Úroveň p
Abs.člen			116707,4	198440,4	0,58812	0,561524
výkon	0,833407	0,215908	4860,1	1259,1	3,86002	0,000673
max.rychlost	0,328018	0,200519	2225,5	1360,5	1,63584	0,113923
spotřeba	-0,356197	0,126022	-55252,4	19548,3	-2,82646	0,008930

V její horní polovině je řada užitečných informací. Pro nás jsou důležité hodnoty **R – totální (vícenásobný) koeficient korelace** a **R^2 totální (vícenásobný) koeficient determinace**.

Chceme-li porovnat totální koeficienty determinace v modelech s různým počtem proměnných, používáme k tomuto účelu nezkrácený odhad totál. koeficientu determinace - **uprav. R^2**.

Ve výstupní sestavě –viz obr. 5 - jsou pro vytvoření modelu důležité hodnoty, uvedené ve sloupci **B**. Jedná se o koeficienty u proměnných, jejichž název je uveden v prvním sloupci –tj. na obr. 5 se jedná o *Abs.člen*, *výkon*, *max. rychlost*, *spotřeba*. Ve výstupní sestavě uvedené na obr. 5 se tedy jedná o model

$$y' = 116707,4 + 4860,1x_1 + 2225,5x_2 - 55252,4x_3$$

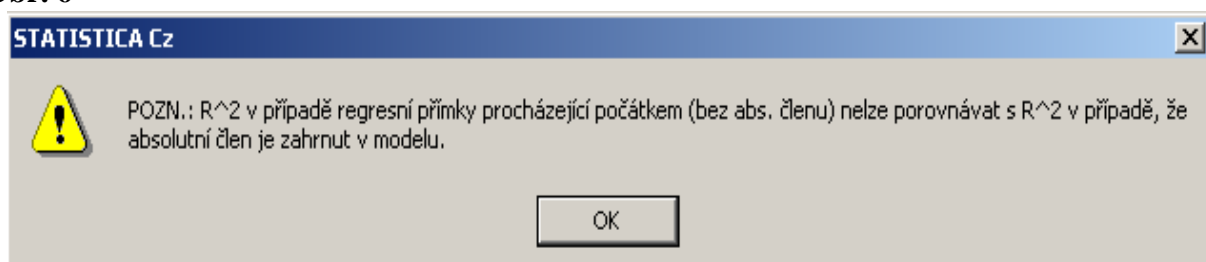
s vícenásobným koeficientem korelace $R_{y,x_1, x_2, x_3} = 0,905$ a s vícenásobným koeficientem determinace $R^2_{y,x_1, x_2, x_3} = 0,819$.

Čísla ve sloupci **Úroveň p** značí nejmenší hladiny významnosti, pro něž lze zamítnout hypotézy o nulových hodnotách regresních koeficientů, tj. výsledky testů $H_0: \beta_i = 0$ proti $H_A: \text{non } H_0, i = 0, 1, 2, 3$. (Interpretujeme-li tedy hodnoty *p* v obr. 5, hypotézu o nulové hodnotě β_0 a β_2 zamítnout nelze a hypotézu o nulové hodnotě β_1 a β_3 zamítnout lze.)

Čísla ve sloupci **t(B)** jsou pak hodnoty testovacích kritérií.

Pro nás z toho plyne, že model lze zjednodušit tím, že vypustíme absolutní člen. Aktivujeme lištu **Výsledky-vícerozměrné**. Ta je zobrazena **v levé spodní části obrazovky**. Objeví se okno, uvedené na obr. 4. (V něm můžeme v kartách **Detailní výsledky** či **Rezidua/předpoklady/předpovědi** zvolit provedení dalších výpočtů.) V okně, uvedeném na obr. 4, stiskneme tlačítko **Storno**. Vrátime se tak do okna **Definice modelu** (obr. 2). V něm v kartě **Detaily** rozbalíme Položku **Abs. člen** a vybereme nabídku **Nastaven na 0**. Klikneme na **OK**. Provede se nový výpočet. Před tím však budeme upozorněni, že nelze srovnávat R^2 původní výstupní sestavy s hodnotou R^2 na sestavě zjednodušené –viz obr. 6. Vícenásobný

Obr. 6



koeficient determinace je v modelu bez absolutního členu počítán podle jiného vzorce, než v modelu s absolutním členem.

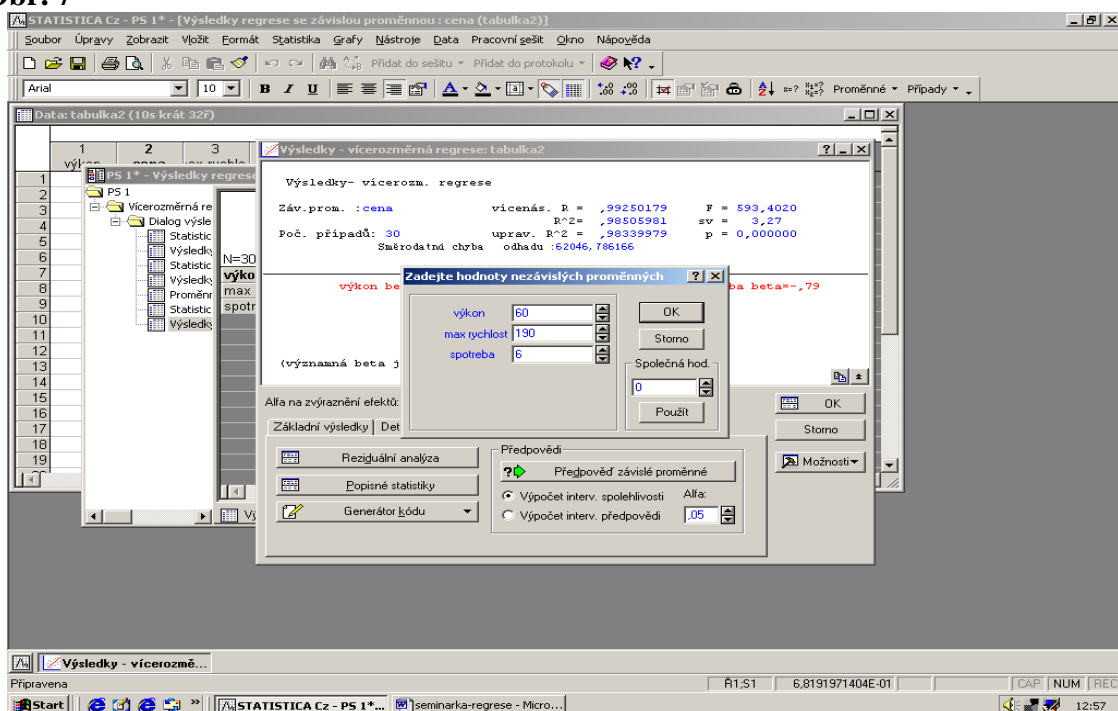
Nová výstupní sestava již absolutní člen nebude mít. Interpretace výsledků je stejná jako interpretace výstupní sestavy uvedené na obr.5.

2) Predikce

Predikci umožní provést nastavení karty **Residua/předpoklady/předpovědi** v okně **Výsledky – vícerozměrná regrese** - viz obr. 4. Do tohoto okna se nejrychleji vrátíme stiskem tlačítka **Výsledky – vícerozmě.**, umístěného ve spodní části obrazovky.

Zvolíme-li nabídku **Výpočet interv. spolehlivosti**, aktualizujeme-li hladinu významnosti **Alfa** a zadáme-li hodnotu **nezávisle proměnné**, pro niž chceme predikci provést – viz obr.7, pak vlastní predikci provedeme stiskem tlačítka **OK**. Výstupní sestava je uvedena na obr. 8.

Obr. 7



Obr. 8

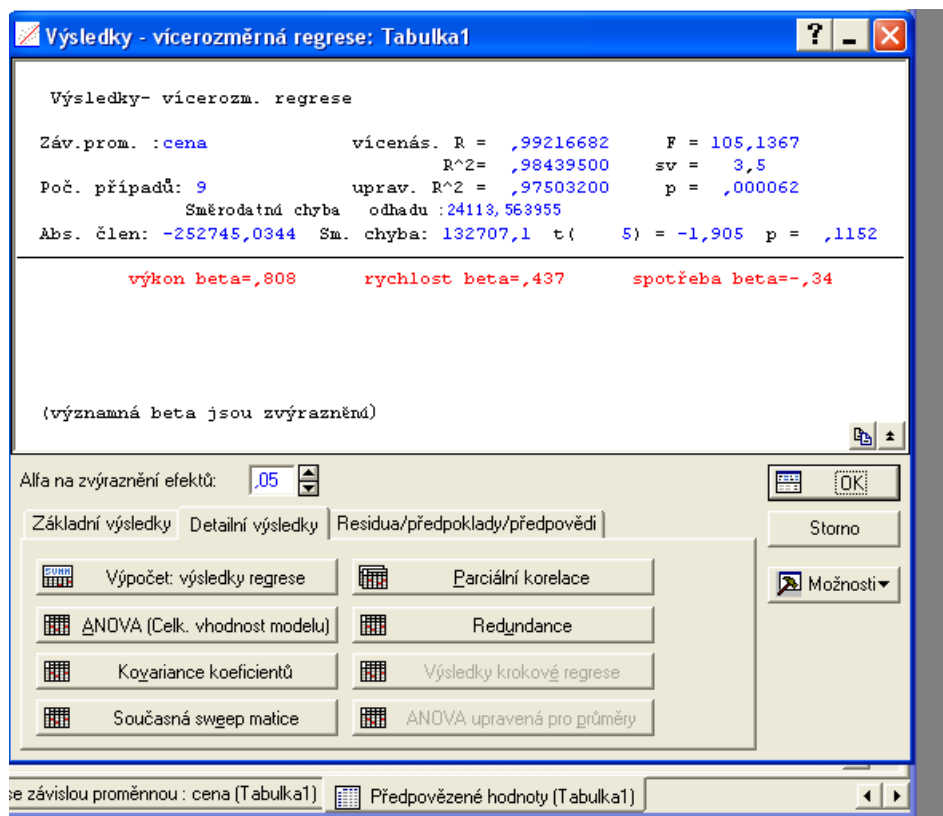
Proměnná	Předpovězené hodnoty (Tabulka1) proměnné: cena		
	B-váž.	Hodnota	B-váž. * Hodnot
výkon	5065,7	60,0000	303940
rychlost	3788,2	190,0000	719751
spotřeba	-52486,6	6,0000	-314920
Abs. člen			-252745
Předpověď			456026
-95,0%LS			390126
+95,0%LS			521926

Obsahuje informace o příspěvcích nezávisle proměnných k odhadu závisle proměnné (*sloupec B-váž.*), bodový a intervalový odhad hodnot závisle proměnné (*sloupec B-váž.*Hodnot*).

3) *Parciální korelace*

Potřebujeme-li znát charakteristiky parciální korelace, vrátíme se do okna *Výsledky – vícerozměrná regrese*- viz obr. 9. (Do tohoto okna se nejrychleji vrátíme stiskem tlačítka *Výsledky – vícerozmě.*, umístěného ve spodní části obrazovky.) Na kartě *Detailní výsledky* zvolíme tlačítko *Parciální korelace*. V sestavě, která se objeví - viz obr. 10, najdeme potřebné informace.

Obr. 9



Obr. 10

Proměnná	Proměnné obsažené v rovnici ; ZP: cena (auta)						
	Beta	Parciál. korelace	Semipar. korelace	Tolerance	R^2	t(27)	Úroveň p
výkon	0,681920	0,763478	0,144495	0,044899	0,955101	6,14267	0,000001
max. rychlost	1,098386	0,582953	0,087697	0,006375	0,993625	3,72811	0,000904
spotřeba	-0,788903	-0,492409	-0,069152	0,007683	0,992317	-2,93972	0,006657