



## Lineárna regresná priamka programom MS EXCEL<sup>®</sup>

Použitie funkcie LINEST() sa v programe Excel ponúka ako optimálne. Pomocník programu ani manuály však na túto možnosť neupozorňujú. Výhodou je výpočet najvýznamnejších štatistických hodnôt a jednoduchosť aplikácie. Aplikácia tejto funkcie má dve obmedzenia:

- ▶ polia definované ako argumenty musia obsahovať číselné dáta (nie prázdne bunky alebo napr. text), počet datových bodov je teda už pre aplikovanú funkciu nemenný,
- ▶ funkcia je nepoužiteľná pre dáta, ktoré sú štrukturované inak ako vektory rovnakých dĺžok.

Funkcia volá tieto parametre: LINEST(pole\_y;pole\_x;B;Štat)

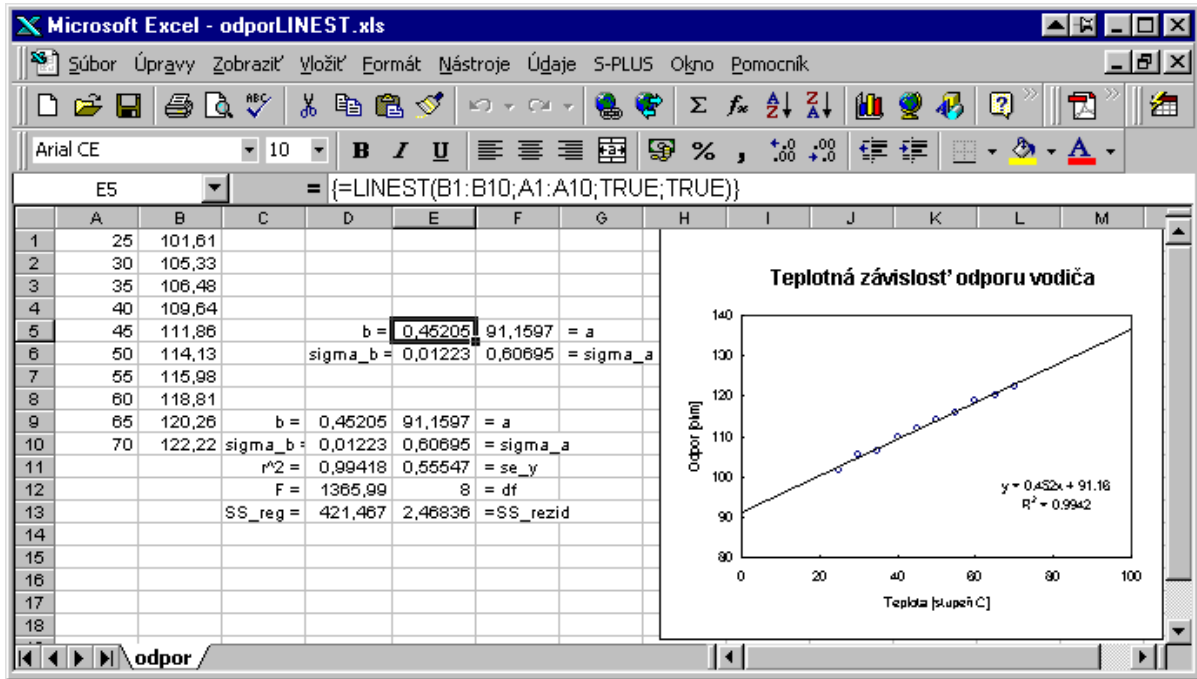
vypočíta regresnú priamku zo zadanej matice hodnôt podľa vzorca  $y = a + bx$ , kde  $b$  je parameter hodnoty  $x$  a  $a$  je konštanta. Výsledkom funkcie LINEST je matica v tvare  $\{b_n, b_{n-1}, \dots, b_1, a\}$ . Pri zadaní do vzorca tvorí pole\_y už vypočítané hodnoty, pole\_x je možná známa množina hodnôt  $x$  (dvojice  $x_i, y_i$  teda reprezentujú experimentálne body) a B je logická hodnota (keď je B *logická nula*, počíta vzorec  $a = 0$ ).

Štat je logický argument, ktorý umožňuje zistenie ďalších hodnôt regresnej štatistiky. Keď je uvedený ako *logická nula*, výsledkom výpočtu budú len hodnoty  $a$  a  $b$ . V prípade uvedenia *logickej jednotky* dostaneme výslednú maticu v tvare

$$\{b_n, b_{n-1}, \dots, b_1, a; se_n, se_{n-1}, \dots, se_1, se_a; r^2, se_y; F, df; SSreg, SSrezid\},$$

kde jednotlivé symboly majú takýto význam:

- ▶  $se_n, se_{n-1}, \dots, se_1, se_a$  sú smerodajné odchýlky koeficientov  $b_n, b_{n-1}, \dots, b_1, a$  (v pomocníkovi je nesprávne pomenovanie štandardná chyba)
- ▶  $r^2$  je koeficient determinovanosti (Pearsonov koef. korelácie), porovnáva skutočné hodnoty  $y$  a ich odhady, pričom nadobúda hodnoty od 0 do 1 (keď je  $r^2 = 1$  potom je dokonalá korelácia medzi odhadom a zistenými hodnotami)
- ▶  $se_y$  je reziduálna smerodajná odchýlka odhadu  $y$  (v Exceli netradične štandardná chyba odhadu)
- ▶ F je hodnota Fischerovho-Snedecovho rozdelenia (tzv.  $F$ -štatistika, vzťah medzi nezávislými a závislými)  $F(\alpha, v_1, v_2)$ , kde  $\alpha$  je hladina významnosti,  $v_1 = p - 1$  je počet stupňov voľnosti,  $p$  je počet koeficientov modelu,  $v_2 = n - p$  je počet stupňov voľnosti regresie. Kritická hodnota  $F_{krit}(\alpha, v_1, v_2)$  sa vypočíta funkciou FINV( $\alpha; p-1; n-p$ ). Splnenie nerovnosti  $F \geq F_{krit}$  znamená pozitívny výsledok.
- ▶ df (degree of freedom) je počet stupňov voľnosti ( $n - 2$ )
- ▶ SSreg (Sum Square of Regression) je regresná suma štvorcov  $\sum_{i=1}^n (y_i - \bar{y})^2$
- ▶ SSrezid je reziduálna suma štvorcov  $\sum_{i=1}^n (y_i - a - bx_i)^2$



## Literatúra

WLSfit.xls Program written by IAN COOPER

School of Physics, The University of Sydney, NSW 2006, Australia

Email: cooper@physics.usyd.edu.au

[http://www.physics.usyd.edu.au/teach\\_res/projects/projects.htm](http://www.physics.usyd.edu.au/teach_res/projects/projects.htm)

Funkcie pre Excel:

<http://www.netrax.net/~jdavita/XlXtrFun/XlXtrFun.htm>

URBÁNEK, T. – ŠKÁRKA, J. 1998. *Microsoft Excel 97. Pro vědce a inženýry*. Praha : Computer Press, 1998, ISBN 80-7226-099-5

ŠŤASTNÝ, Z. 1999. *Microsoft Excel. Matematické a statistické výpočty*. Praha : Computer Press, 1999, ISBN 80-7226-141-X

LIENGME, B. V. 2002. *A Guide to Microsoft Excel 2002. For Scientists and Engineers*. Oxford : Butterworth-Heinemann, Third Edition, 2002, ISBN 0-7506-5613-1

CHAJDIÁK, J. 2002. *Štatistika v EXCELi*. Bratislava : Statis, 2002, ISBN 80-85659-27-1