

## Téma 5 – Příklad 3

### Zadání příkladu

Data v tabulce (cena; poptávka) udávají poptávku a cenu určitého spotřebního zboží. Zvolte vhodný regresní model popisující závislost poptávky na ceně.

Cena	100	110	140	160	200
Množství	120	89	56	41	22

### Vypracování příkladu

Jelikož postupujeme (jak je v realitě nejběžnější) bez předem zadaného regresního modelu, musíme jej nejprve vybrat. K tomu nás mohou dovést:

1. Věcně-ekonomická kritéria.
2. Grafické znázornění.
3. Matematicko-statistická kritéria.

Vztah mezi poptávaným množstvím a cenou je známý ekonomický problém, můžeme vycházet z této diskuze. Z didaktických důvodů ovšem budeme uvažovat i o ostatních, teoriích nepopsaných modelech běžných regresních funkcí. Až teprve poté provedeme testy o statistické významnosti zvoleného modelu. Tyto testy jsou popsány v předešlých příkladech.

### Řešení v SGP

Zadání do programu není v tomto případě vůbec problematické. Stačí přepsat obě proměnné do samostatných sloupců.

**Procedura v SGP: Relate – One Factor - Simple Regression (Analysis Summary, Comparison of Alternative Models)**

Při vstupním dialogu zadáme jako závislou proměnnou y *mnozstvi* a jako proměnnou nezávislou x *cena*.

V druhém kroku – Analysis options – vybíráme regresní funkci. Můžeme začít od známé regresní přímky, je označena jako model *Linear*.

Okno Analysis Options shrnuje výsledky regresní analýzy. V tabulce Coefficients jsou odhady jednotlivých parametrů a jejich dílčí t-testy. V tomto případě parametr průsečíku (Intercept) a směrnice (Slope). T-Statistic a P-Value ukazuje příslušnou hodnotu kritéria a jeho významnosti. Druhá tabulka s názvem Analysis of Variance zobrazuje provedený F-test. Rozklad variability na teoretickou (Model) a residuální (Residual). Hodnota kritéria F (F-Ratio) a jeho významnost (P-Value) nám vypoví, zda je model na zvolené hladině významnosti vhodný k vystižení závislosti.

Pod tabulkami jsou hodnoty ukazatelů, které slouží k charakteristice a posouzení vhodnosti modelu. Index determinace odpovídá hodnotě R-squared. Okno Comparison of Alternative Models nám napoví, jaké jsou hodnoty indexu determinace u ostatních modelů (pozor, nejedná se o upravenou verzi!).

**Comparison of Alternative Models**

<i>Model</i>	<i>Correlation</i>	<i>R-Squared</i>
Logarithmic-Y square root-X	-0,9982	99,64%
Square root-Y reciprocal-X	0,9972	99,45%
Multiplicative	-0,9970	99,41%
Exponential	-0,9970	99,39%
Reciprocal-Y squared-X	0,9941	98,82%
Square root-Y logarithmic-X	-0,9939	98,79%
Reciprocal-X	0,9903	98,07%
Double square root	-0,9887	97,75%
S-curve model	0,9879	97,60%
Logarithmic-Y squared-X	-0,9873	97,48%
Square root-Y	-0,9811	96,25%
Reciprocal-Y	0,9790	95,85%
Logarithmic-X	-0,9758	95,21%
Reciprocal-Y square root-X	0,9678	93,66%
Square root-X	-0,9649	93,10%
Square root-Y squared-X	-0,9592	92,00%
Reciprocal-Y logarithmic-X	0,9542	91,05%
Linear	-0,9517	90,57%
Squared-Y reciprocal-X	0,9478	89,84%
Double reciprocal	-0,9218	84,97%
Squared-X	-0,9193	84,51%
Squared-Y logarithmic-X	-0,9169	84,07%
Squared-Y square root-X	-0,8980	80,64%
Squared-Y	-0,8770	76,92%
Double squared	-0,8306	68,99%
Logistic	<no fit>	
Log probit	<no fit>	

Z daných výsledků můžeme vyčíst, že poměrně významně vyšší hodnota indexu determinace je u exponenciálního modelu. Zvolíme jej a podíváme se, zda splňuje dílčí t-testy a celkový F-test.

### Simple Regression - poptavka vs. cena

Dependent variable: poptavka  
 Independent variable: cena  
 Exponential model:  $Y = \exp(a + b \cdot X)$   
 Number of observations: 5

#### Coefficients

	Least Squares	Standard	T	
Parameter	Estimate	Error	Statistic	P-Value
Intercept	6,35802	0,108647	58,5202	0,0000
Slope	-0,0164563	0,000741653	-22,1887	0,0002

NOTE: intercept =  $\ln(a)$

#### Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Model	1,75485	1	1,75485	492,34	0,0002
Residual	0,010693	3	0,00356432		
Total (Corr.)	1,76555	4			

Correlation Coefficient = -0,996967  
 R-squared = 99,3944 percent  
 R-squared (adjusted for d.f.) = 99,1925 percent  
 Standard Error of Est. = 0,0597019  
 Mean absolute error = 0,039759  
 Durbin-Watson statistic = 1,92051 (P=0,1619)  
 Lag 1 residual autocorrelation = -0,251619

Model regresní exponenciály je vhodným modelem. K posouzení vhodnosti je možné ještě aplikovat porovnání testové statistiky F, nebo reziduálního součtu čtverců (pouze u modelů se stejným počtem parametrů).

### Interpretace

Na 5% hladině významnosti bylo prokázáno, že model regresní exponenciály je statisticky významný k vystižení závislosti poptávaného množství zboží na jeho ceně. Tento model vysvětluje více než 99 % celkové variability.

Pozn.: Lze uvažovat i o modelech některých polynomů.