

Základní charakteristiky časových řad

Zadání příkladu:

Určete základní charakteristiky časové řady středního počtu obyvatelstva České republiky (ozn. $P(t)$) od roku 1989 do roku 2011. Hodnoty jsou uvedeny v tabulce 1. Dále vytvořte vhodný graf této veličiny.

Tab. 1 - Počet obyvatel ČR letech 1989-2011 v tisících - ozn. P_t

Rok	1989	1990	1991	1992	1993	1994	1995	1996
P_t	10362	10363	10309	10318	10331	10336	10331	10315
Rok	1997	1998	1999	2000	2001	2002	2003	2004
P_t	10304	10295	10283	10272	10224	10201	10202	10207
Rok	2005	2006	2007	2008	2009	2010	2011	
P_t	10234	10267	10313	10430	10491	10517	10497	

Zdroj dat: Český statistický úřad (ČSÚ)

Vypracování příkladu

Ve statistice dělíme sledované veličiny podle různých hledisek. Například na slovní a číselné, diskrétní a spojité, alternativní a množné apod. Mimo jiné známe i dělení na prostorové a časové. U prostorových veličin není okamžik, ke kterému napozorovaná hodnota veličiny přísluší, důležitý. U časových veličin je ale každá hodnota pevně svázána s časovým údajem svého vzniku, takže soubor, se kterým pracujeme, je množinou uspořádaných dvojic $[t, X_t]$, kde čas $t = 1, \dots, T$ a veličina X má hodnoty X_1, \dots, X_T . Hodnota času t je u časových řad nezbytná pro jejich analýzu, neboť zde zkoumáme například trendovou nebo sezónní složku, neboli chování veličiny X jako funkci času t .

Program STATGRAPHICS Centurion XVII

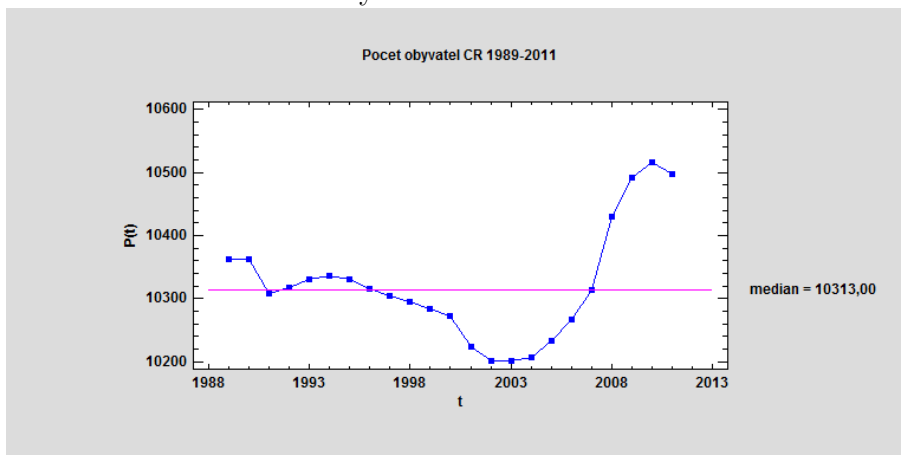
V programu STATGRAPHICS je pro grafy časových řad (ČŘ) vytvořena speciální funkcionálita. Najdeme ji v menu *Plot/Time Sequence Plots/Run Charts/Individuals*. Do povinného pole **Observations** je třeba zadat název proměnné, kde jsou hodnoty sledované veličiny X_t . Do pole (**Date/Time/Labels:**) je možné zadat odpovídající hodnoty času.

Vytvořený graf je možné upravovat přes menu, které se otevírá pravým tlačítkem myši, když je ukazatel myši na grafu, volba *Graphics Options*. V tomtéž menu je volba *Save Graph*, kterou lze graf vyexportovat do samostatného souboru. Grafické znázornění je ČŘ velmi důležité a slouží především k předběžné vizuální identifikaci trendů, periodicity a vůbec vývoje sledované veličiny v čase.

Nyní určíme základní číselné charakteristiky zkoumané řady. Těmi jsou popisné statistiky jako je **minimum, maximum, aritmetický průměr, medián, rozpětí a další**, ale také první diference, průměrný přírůstek a průměrný koeficient růstu.

Popisné statistiky lze získat přes *Describe/Numeric Data/One-variable Analysis*, kde do pole **Data** opět vložíme název sledované proměnné.

Obrázek 1: Počet obyvatel ČR v letech 1989-2011 v tis.



Zdroj: Údaje ČSÚ, vlastní zpracování

Výsledky:

$X_{min} = 10201$, $X_{MAX} = 10517$, rozpětí $R = 316$, $X_{0,25} = 10267$, medián $\tilde{X} = 10313$, $X_{0,75} = 10362$.

Průměr \bar{X} nelze v tomto případě spočítat podle klasického vzorce $\bar{X} = \frac{\sum_{t=1}^T X_t}{T}$, protože se jedná o **okamžikovou časovou řadu**, kde součet hodnot nedává reálně příliš smysl. Použijeme proto vztah pro **prostý chronologický průměr**, neboť vzdálenosti mezi časovými okamžiky jsou stejné (za střední stav obyvatelstva v kalendářním roce je v ČR považován počet obyvatel daného území vždy o půlnoci z 30. 6. na 1. 7. sledovaného roku). Platí

$$\bar{X} = \frac{\frac{X_1}{2} + \sum_{t=2}^{T-1} X_t + \frac{X_T}{2}}{T - 1}$$

Takže prostý chronologický průměr $\bar{X} = 10316,93$.

Nyní přikročme k výpočtu dalších charakteristik. **První diference** $\Delta_t^{(1)}$ zkoumané ČŘ vypočteme jako $\Delta_t^{(1)} = X_t - X_{t-1}$, $t = 2, \dots, T$. Jejich význam se projeví v trendové analýze ČŘ, která je dalším tématem přednášky. V programu STATGRAPHICS je pro výpočet prvních diferencí přímo vytvořena funkce **Diff (proměnná)**. Postup výpočtu je následující: označit myší volný sloupec datového listu kliknutím na záhlaví; kliknout na záhlaví sloupce pravým tlačítkem myši a v menu, které se otevře, vybrat položku *Generate Data*; do pole *Expression* napsat příkaz **Diff (proměnná)**, přičemž proměnná je ten sloupec datového listu, kde jsou hodnoty analyzované ČŘ.

Tab. 2 - První diference počtu obyvatel ČR letech 1989-2011 v tisících - ozn. $\Delta_t^{(1)}$

Rok	1989	1990	1991	1992	1993	1994	1995	1996
$\Delta_t^{(1)}$		1	-54	9	13	5	-5	-16
Rok	1997	1998	1999	2000	2001	2002	2003	2004
$\Delta_t^{(1)}$	-11	-9	-12	-11	-48	-23	1	5
Rok	2005	2006	2007	2008	2009	2010	2011	
$\Delta_t^{(1)}$	27	33	46	117	61	26	-20	

Zdroj dat: Český statistický úřad (ČSÚ) + vlastní zpracování

Průměrný přírůstek ČR $\bar{\Delta}$ se spočítá jako aritmetický průměr jejích prvních diferencí. Po zjednodušení vzorce pro tento průměr vyjde vztah

$$\bar{\Delta} = \frac{X_T - X_1}{T - 1}$$

Zde konkrétně $\bar{\Delta} = 135/22 \doteq 6,13$.

Poslední charakteristikou je **průměrný koeficient růstu** \bar{k} . Je to geometrický průměr všech $T - 1$ koeficientů růstu $k_2, k_3, \dots, k_{T-1}, k_T$, které se spočtou jako podíl hodnot ČR dvou sousedních období:

$$k_t = \frac{X_t}{X_{t-1}}, \quad t = 2, \dots, T$$

Hodnota k_t udává, kolikrát vzrostla hodnota ČR ve sledovaném období oproti období předchozímu. Geometrický průměr má potom tvar

$$\bar{k} = \sqrt[T-1]{k_2 \cdot k_3 \cdot \dots \cdot k_{T-1} \cdot k_T} = \sqrt[T-1]{\frac{X_T}{X_1}}$$

Zde je $\bar{k} = \sqrt[22]{\frac{10497}{10362}} \doteq 1,00059$. Pokud je hodnota $\bar{k} > 1$, potom řada v průměru meziročně roste. Pokud je menší než 1, jedná se o pokles. Percentuelní vyjádření získáme tak, že vypočteme výraz $100 \cdot (\bar{k} - 1)$. Pokud je kladný, jedná se opět o růst, v opačném případě o pokles. Pro sledovanou ČR je $100 \cdot (\bar{k} - 1) = 0,059\%$.

Interpretace výsledků

Při interpretaci výsledků je nutné vedle sebe mít jak grafické znázornění, tak číselné charakteristiky. Závěry založené pouze na popisných statistikách mohou často vést k chybným závěrům, zejména v trendové analýze!

Z grafu je patrné, že počet obyvatel během prvních 13 let klesal. V roce 2011 dosáhl svého minima (10201), aby poté začal znovu stoupat (maximum 10517 v roce 2010). S tím souvisí i znaménko prvních diferencí, které je v prvním jmenovaném období vesměs záporné, ve zbytku naopak kladné. Rozpětí činí 316. Vzhledem k tomu, že průměr je větší než medián, je většina hodnot podprůměrná. Průměrný přírůstek činí 6,13. Lze proto říci, že od roku 1989 do roku 2011 se počet obyvatel každý rok v průměru zvýšil o 6130. Relativně vyjádřeno, každým rokem se počet obyvatel zvýšil o 0,59 ‰.