时间序列也称动态序列，是指将某种现象的指标数值按照时间顺序排列而成的数值序列。时间序列分析大致可分成三大部分，分别是描述过去、分析规律和预测未来，本讲将主要介绍时间序列分析中常用的三种模型：季节分解、指数平滑方法和ARIMA模型，并将结合Spss软件对时间序列数据进行建模。

时间序列数据

时间序列数据：对同一对象在不同时间连续观察所取得的数据。

Time Series Data

例如：

（1）从出生到现在，你的体重的数据（每年生日称一次）。

（2）中国历年来GDP的数据。

（3）在某地方每隔一小时测得的温度数据。

时间序列的基本概念

时间序列也称动态序列，是指将某种现象的指标数值按照时间顺序排列而成的数值序列。

时间序列由两个组成要素构成：

1、第一个要素是**时间要素**；

年、季度、月、周、日、小时、分钟、秒

2、第二个要素是**数值要素**。

时间序列根据时间和数值性质的不同，可以分为时期时间序列和时点时间序列。

时期序列中，数值要素反映现象在一定时期内发展的结果；

时点序列中，数值要素反映现象在一定时点上的瞬间水平。

区分时期和时点时间序列

例如：

（1）从出生到现在，你的体重的数据（每年生日称一次）。

（2）中国历年来GDP的数据。

（3）在某地方每隔一小时测得的温度数据。

（1）和（3）是时点时间序列；（2）是时期时间序列

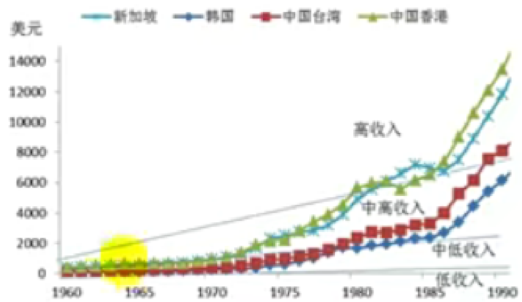
时期序列可加，时点序列不可加。

时期序列中的观测值反映现象在一段时期内发展过程的总量，不同时期的观测值可以相加，相加结果表明现象在更长一段时间内的活动总量；而时点序列中的观测值反映现象在某一瞬间上所达到的水平，不同时期的观测值不能相加，相加结果没有实际意义。

（灰色预测模型里面有一个累加的过程）

长期趋势：T

长期趋势(Secular trend,T)指的是统计指标在相当长的一段时间内，受到长期趋势影响因素的影响，表现出持续上升或持续下降的趋势，通常用字母T表示。例如，随着国家经济的发展，人均收入将逐渐提升；随着医学水平的提高，新生儿死亡率在不断下降。



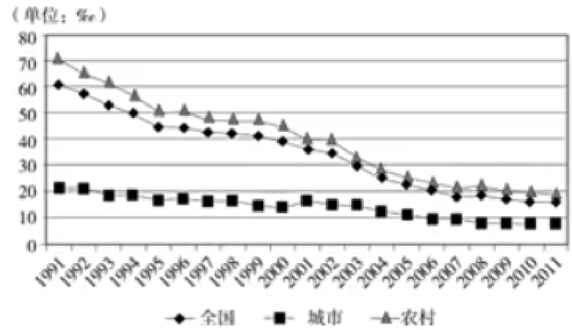


图4 中国5岁以下儿童死亡率

季节趋势：S

季节趋势(Seasonal Variation,S)是指由于季节的转变使得指标数值发生周期性变动。这里的季节是广义的，一般以月、季、周为时间单位，不能以年作单位。例如雪糕和棉衣的销量都会随着季节气温的变化而周期变化；每年的长假（五一、十一、春节）都会引起出行人数的大量增加。

时间序列分解

以上四种变动就是时间序列数值变化的分解结果。有时这些变动会同时出现在一个时间序列里面，有时也可能只出现一种或几种，这是由引起各种变动的影响因素决定的。正是由于变动组合的不确定性，时间序列的数值变化才那么千变万化。

四种变动与指标数值最终变动的关系可能是叠加关系，也可能是乘积

叠加模型和乘积模型

(1) 如果四种变动之间是相互独立的关系，那么叠加模型可以表示为：

$$Y = T + S + C + I$$

(2) 如果四种变动之间存在相互影响关系，那么应该使用乘积模型：

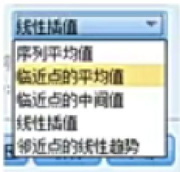
$$Y = T \times S \times C \times I$$

Y: 指标数值的最终变动;
T: 长期趋势变动;
S: 季节变动;
C: 循环变动;
I: 不规则变动。

（1）数据具有周期性时才能使用时间序列分解，例如数据是月份数据(周期为12)、季度数据(周期为4)，如果是年份数据则不行。

（2）在具体的时间序列图上，如果随着时间的推移，序列的季节波动变得越来越大，则反映各种变动之间的关系发生变化，建议使用乘积模型；反之，如果时间序列图的波动保持恒定，则可以直接使用叠加模型；当然，如果不存在季节波动，则两种分解均可以。

替换缺失值的五种方法



序列平均值
- 用整个序列的平均数代替缺失值

临近点的平均值
- 用相邻若干个点的平均数来替换缺失值（默认为两个点）

临近点的中位数
- 用相邻若干个点的中位数来替换缺失值（默认为两个点）

线性插值
- 用相邻两个点的平均数来替换缺失值

邻近点的线性趋势
- 将时期数作为x，时间序列值作为y进行回归，求缺失点的预测值

时间序列分析

描述过去	分析规律	预测未来
描述时间序列的动态变化	揭示时间序列数值变化背后的规律	依据数值变化规律预测未来数值趋势

时间序列分析的具体步骤

- 作时间序列图；
- 判断时间序列包含的变动成分；
- 时间序列分解（有周期性且包含长期趋势、季节变动或循环变动）；
- 建立时间序列分析模型；
- 预测未来的指标数值。

SPASS中七种指数平滑方法

Simple模型

名称	适用条件	与之类似的ARIMA模型
简单指数平滑法	不含趋势和季节成分	ARIMA（0,1,1）

令x_t为t时刻的观测数据，S_t为第t期的平滑值，且令$S_t=\hat{x}_{t+1}$，即第$t+1$期的预测值，且满足：$\hat{x}_{t+1}=\alpha x_t+(1-\alpha)\hat{x}_t$

可以证明：$\hat{x}_{t+1}=\alpha x_t+\alpha(1-\alpha)x_{t-1}+\alpha(1-\alpha)^2x_{t-2}+\cdots+\alpha(1-\alpha)^{t-1}x_1+(1-\alpha)^t l_0$ 其中$l_0=\hat{x}_1$视为初始值

例如：当$\alpha=0.5$时，$S_t=\hat{x}_{t+1}=0.5x_t+0.25x_{t-1}+0.125x_{t-2}+0.0625x_{t-3}+\cdots$，其中$\alpha$被称为平滑系数（$0\le\alpha\le1$）

显然：每一个平滑后的数据都是由过去的数据加权求和后所得，越接近当期的数据，其权重越大

这说明距离当期越接近的数据，对当期的影响也越大；反之，越早期的数据，对当期影响越小。

关于平滑系数α的选取原则：

1、如果时间序列具有不规则的起伏变化，但长期趋势接近一个稳定常数，α值一般较小(取0.05-0.02之间)

2、如果时间序列具有迅速明显的变化倾向，则α应该取较大值（取0.3-0.5）

3、如果时间序列变化缓慢，亦应选较小的值（一般在0.1-0.4之间）

实际上，Spss的专家建模如果选择了Simple模型用来估计，那么软件会帮我们自动选取一个适合的平滑系数使得预测误差最小。

线性趋势模型(linear trend)

名称	适用条件	与之类似的ARIMA模型
霍特线性趋势模型	线性趋势、不含季节成分	ARIMA（0,2,2）

Holt 在1957年把简单的指数平滑模型进行了延伸，能够预测包含趋势的数据，该方法包含一个预测方程和两个平滑方程（一个用于水平，另一个用于趋势）：

$$\begin{cases} l_t = \alpha x_t + (1-\alpha)(l_{t-1} + b_{t-1}) & \text{（水平平滑方程）} \\ b_t = \beta(l_t - l_{t-1}) + (1-\beta) b_{t-1} & \text{（趋势平滑方程）} \\ \hat{x}_{t+h} = l_t + h b_t,\ h = 1, 2, \cdots & \text{（预测方程）} \end{cases}$$

t: 当前期;

h: 预测超前期数，也称之为预测步长;

x_t: 第t期的实际观测值

T_t: 时刻t的预估水平

b_t: 时刻t的预测趋势（或坡度）

α: 水平的平滑参数

β: 趋势的平滑参数

布朗(Brown)线性趋势模型

假定$\alpha = \beta$，即认为水平平滑参数和趋势平滑参数相等。（是Holt线性趋势模型的特例）

阻尼趋势模型(Damped trend)

名称	适用条件	与之类似的ARIMA模型
阻尼趋势模型	线性趋势逐渐减弱且不含季节成分	ARIMA（1,1,2）

经验表明，Holt的线性模型和指数模型倾向于对未来预测值过高，特别是对于长期预测。Gardner 和 McKenzie (1985)在霍特的模型基础上引入了一种阻尼效应，用来缓解较高的线性趋势。

$$\begin{cases} l_t = \alpha x_t + (1-\alpha)(l_{t-1} + \phi b_{t-1}) & \text{（水平平滑方程）} \\ b_t = \beta(l_t - l_{t-1}) + (1-\beta) \phi b_{t-1} & \text{（趋势平滑方程）} \\ \hat{x}_{t+h} = l_t + (\phi + \phi^2 + \cdots + \phi^h) b_t & \text{（预测方程）} \end{cases}$$

α: 水平的平滑参数

β: 趋势的平滑参数

ϕ: 阻尼参数（读作$phi, 0 < \phi \leq 1$）

如果$\phi = 1$，则阻尼趋势模型就是霍特线性趋势模型

对于在0到1的值，ϕ会对趋势产生阻尼效应

简单季节性(Simple seasonal)

名称	适用条件	与之类似的ARIMA模型
简单季节性	含有稳定的季节成分、不含趋势	SARIMA $(0,1,1)\times(0,1,1)_S$

$$
\begin{cases}
l_t=\alpha(x_t-s_{t-m})+(1-\alpha)l_{t-1} & \text{（水平平滑方程）}\\
s_t=\gamma(x_t-l_{t-1})+(1-\gamma)s_{t-m} & \text{（季节平滑方程）}\\
\hat{x}_{t+h}=l_t+s_{t+h-m(k+1)}\text{ , }k=[\dfrac{h-1}{m}] & \text{（预测方程）}
\end{cases}
$$

m: 周期长度（月度数据取12，季度数据取4）

α: 水平的平滑参数

γ: 季节的平滑参数

h: 预测超前期数

$\hat{x}_{t+h}$: 第h期的预测值

https://otexts.com/fpp2/taxonomy.html

温特加法模型(Winters' additive)

名称	适用条件	与之类似的ARIMA模型
温特加法模型	含有线性趋势和稳定的季节成分	SARIMA $(0,1,0)\times(0,1,1)_S$

Holt (1957) and Winters (1960) extended Holt's method to capture seasonality.

$$
\begin{cases}
l_t=\alpha(x_t-s_{t-m})+(1-\alpha)(l_{t-1}+b_{t-1}) & \text{（水平平滑方程）}\\
b_t=\beta(l_t-l_{t-1})+(1-\beta)b_{t-1} & \text{（趋势平滑方程）}\\
s_t=\gamma(x_t-l_{t-1}-b_{t-1})+(1-\gamma)s_{t-m} & \text{（季节平滑方程）}\\
\hat{x}_{t+h}=l_t+hb_t+s_{t+h-m(k+1)}\text{ , }k=[\dfrac{h-1}{m}] & \text{（预测方程）}
\end{cases}
$$

m: 周期长度（月度数据取12，季度数据取4）

α: 水平的平滑参数

β: 趋势的平滑参数

γ: 季节的平滑参数

$\hat{x}_{t+h}$: 第h期的预测值

温特乘法模型(Winters' multiplicative)

名称	适用条件	与之类似的ARIMA模型
温特乘法模型	含有线性趋势和不稳定的季节成分	不存在

$$
\begin{cases}
l_t = \alpha \dfrac{x_t}{s_{t-m}} + (1-\alpha)(l_{t-1} + b_{t-1}) & \text{（水平平滑方程）} \\
b_t = \beta(l_t - l_{t-1}) + (1-\beta) b_{t-1} & \text{（趋势平滑方程）} \\
s_t = \gamma \dfrac{x_t}{l_{t-1} + b_{t-1}} + (1-\gamma) s_{t-m} & \text{（季节平滑方程）} \\
\hat{x}_{t+h} = (l_t + h b_t) s_{t+h-m(k+1)}, \ k = [\dfrac{h-1}{m}] & \text{（预测方程）}
\end{cases}
$$

m: 周期长度（月度数据取12，季度数据取4）

α: 水平的平滑参数

β: 趋势的平滑参数

γ: 季节的平滑参数

$\hat{x}_{t+h}$: 第h期的预测值