1，不同数据频率下，周期差分之阶数选取

根据数据频率的不同，周期阶数选取也不同。根据经验总结，可以将周期性数据分为以下三种：一是单一周期，且数据频率与周期一致，如季度数据（周期为4）、月度数据（周期为12）、周数据（周期为7）；二是单一周期，但数据频率与周期不一致，如；三是混合周期，如客运量数据包含月度效应、星期效应，周期为12、7.

如何将日数据转换为月度数据、季度数据？再周期差分？

理解差分公式及含义？

<http://robjhyndman.com/hyndsight/detecting-seasonality/>

混合周期、长周期 推荐傅立叶方法或状态空间方法

强烈参考：

<http://robjhyndman.com/hyndsight/longseasonality/>

<http://robjhyndman.com/hyndsight/tbats-with-regressors/>

I have a question regarding setting annual seasonality to 365.25. Sometimes I get an error message when using tbats() as follows:

Error in optim(par = param.vector$vect, fn = calcLikelihoodTBATS, method = "Nelder-Mead", : function cannot be evaluated at initial parameters

However, when I switch it to 365.25001 things work just fine. Any insight will be greatly appreciated! Thanks again.

7 is correct for weekly patterns. But months are of unequal lengths so it is not possible to handle with Fourier like terms. It is rare to see daily data with a monthly pattern. It is much more common to have an annual pattern where the seasonality is 365.25,

#### complex seasonal

bats() and tbats() in the forecast package

<http://robjhyndman.com/hyndsight/seasonal-periods/>

De Livera, A.M., Hyndman, R.J., & Snyder, R. D. (2011), Forecasting time series with complex seasonal patterns using exponential smoothing, Journal of the American Statistical Association, **106**(496), 1513-1527.

Forecasting time series with complex seasonal patterns using exponential smoothing

<http://robjhyndman.com/papers/ComplexSeasonality.pdf>

#### multiple seasonality

There are no R packages that handle multiple seasonality for ARIMA models as far as I know. You could try the forecast package which implements multiple seasonality using models based on exponential smoothing. The dshw, bats and tbats functions will all handle data with multiple seasonal periods.

参考：

<http://stats.stackexchange.com/questions/47729/two-seasonal-periods-in-arima-using-r?rq=1>

#### Seasonal ARIMA Modelling

##### 1 季节差分与1阶季节差分

**如果季节差分后，仍有季节性（检验非平稳）、再次1阶差分。**

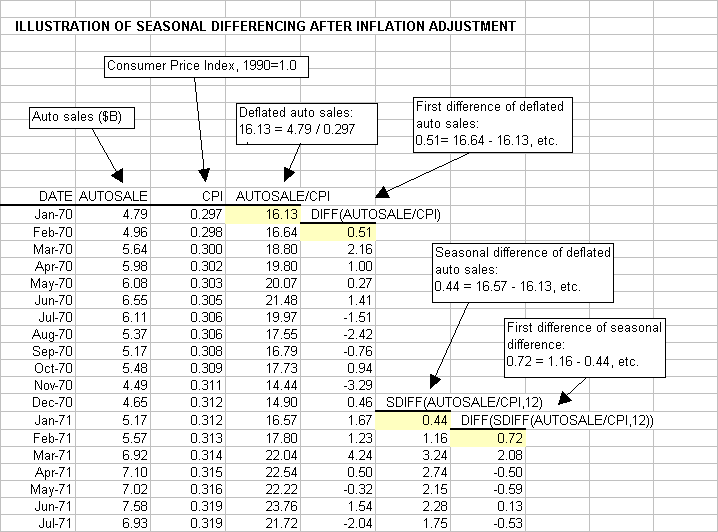
If the seasonal difference of Y is stationary white noise (independently and identically distributed values with no autocorrelation), then Y is described by a [seasonal random walk](http://people.duke.edu/~rnau/411searw.htm) model: each value is a random step away from the value that occurred exactly one *season* ago.

If the first difference of the seasonal difference of Y is pure noise, then Y is described by a [seasonal random trend](http://people.duke.edu/~rnau/411seart.htm) model.

For monthly data, in which there are 12 periods in a season, the seasonal difference of Y at period t is Yt - Yt-12.

The **first difference of the seasonal difference** of a monthly time series Y at period t is equal to (Yt - Yt-12) - (Yt-1 - Yt-13). Equivalently, it is equal to (Yt - Yt-1) - (Yt-12 - Yt-13). This is the amount by which the change from the previous period to the current period is different from the change that was observed exactly one year earlier. Thus, for example, the first difference of the seasonal difference in September 1995 is equal to the August-to-September change in 1995 minus the August-to-September change in 1994.

示例：The following spreadsheet shows how the seasonal difference and first difference of the seasonal difference are calculated in this example:



##### 2 季节ARIMA

The seasonal part of an ARIMA model has the same structure as the non-seasonal part: it may have an AR factor, an MA factor, and/or an order of differencing. In the seasonal part of the model, all of these factors operate across multiples of lag s (the number of periods in a season).

A seasonal ARIMA model is classified as an **ARIMA(p,d,q)x(P,D,Q)** model, where P=number of seasonal autoregressive (SAR) terms, D=number of seasonal differences, Q=number of seasonal moving average (SMA) terms

In identifying a seasonal model, the *first* step is to determine whether or not a seasonal *difference* is needed, in addition to or perhaps instead of a non-seasonal difference. You should look at time series plots and ACF and PACF plots for all possible combinations of 0 or 1 non-seasonal difference and 0 or 1 seasonal difference.

**Rule 12: If the series has a strong and consistent seasonal pattern, then you should use an order of seasonal differencing--but never use more than one order of seasonal differencing or more than 2 orders of total differencing (seasonal+nonseasonal).**

**Rule 13: If the autocorrelation at the seasonal period is positive, consider adding an SAR term to the model. If the autocorrelation at the seasonal period is negative, consider adding an SMA term to the model. Try to avoid mixing SAR and SMA terms in the same model, and avoid using more than one of either kind.**

Usually an SAR(1) or SMA(1) term is sufficient. You will rarely encounter a genuine SAR(2) or SMA(2) process, and even more rarely have enough data to estimate 2 or more seasonal coefficients without the estimation algorithm getting into a "feedback loop."

**Probably the most commonly used seasonal ARIMA model is the (0,1,1)x(0,1,1) model**--i.e., an MA(1)xSMA(1) model with both a seasonal and a non-seasonal difference. This is essentially a "seasonal exponential smoothing" model.

**重点参考：**

[**http://people.duke.edu/~rnau/arimrule.htm**](http://people.duke.edu/~rnau/arimrule.htm)

[**http://people.duke.edu/~rnau/seasarim.htm**](http://people.duke.edu/~rnau/seasarim.htm)

[**http://people.duke.edu/~rnau/411seas.htm**](http://people.duke.edu/~rnau/411seas.htm)

**书籍：**

**Forecasting principles and practice**

案例：

<http://stats.stackexchange.com/questions/14742/auto-arima-with-daily-data-how-to-capture-seasonality-periodicity>

https://rpubs.com/ryankelly/tsa5

<http://www.r-bloggers.com/seasonal-or-periodic-time-series/>

<http://www.r-bloggers.com/seasonal-unit-roots/>

#### **差分函数**

**diff()不带参数名的参数指滞后阶数，也就是与滞后第几阶的数据进行差分。如果要指定差分的阶数，则一定要使用带名称的参数:diff=2。**

1、diff(sample,2)表示是对滞后2阶的数据进行差分，一阶差分，等同于：

diff(sample,lag=2)

2、diff(sample,diff=2)才是表示二阶差分