Rather than worrying about whether a time series is stationary, or doing any Ljung-Box-Cox-Kwiatkowski-Phillips-Schmidt-Shin-[insert name here] tests, or fitting any distributions, **you can handle everything with features**.  
  
I.e.:

* you could just break out **seasonality**(even multiple layers of it) using dummy variables -  day of week, weekends, public holidays, seasons, time of day, etc.
* you can use with **cross sectional** data along with the time series data.
* you can introduce **lags**as features
* you can use the same **transformations**of features (e.g. polynomials) as if you were doing statistical modelling

Note that you will still have to do some form of feature selection. But after that you can just run it through a robust machine learning algorithm like an implementation of decision trees.

<https://www.quora.com/Data-Science-Can-machine-learning-be-used-for-time-series-analysis>

#### distribution

实际上，对时序的分布并没有要求，只要残差是白噪声即可.

所以在时序建模中，如果y为时序，则自变量x中应加入y的周期性、趋势性特征（此时，相当于对去除y的周期性、趋势性后剩余的残差建模，分析其影响因素）。

Box-Cox转换

优点：响应变量不满足正态分布的情况下，可在一定程度上减少不可观测的误差和预测变量的相关性

更多参考：

Forecasting time series using R-ppt.pdf

#### order (iid-independent)

As long as you are passing the features required to capture the order, the trend and seasonality effect (many more such features).

Many models assume that the observations are independent (among other things); if you were to **shuffle** the observations around the general properties of the system would remain the same.

One simple way to see how well your model is reproducing the structure of the data is checking whether your **residuals** are approximately white noise.

For xgboost example, the main reason why (by changing order of columns or rows) you get different results is because of subsample and colsample\_by\_tree parameter. they choose randomly a subset of rows or columns if subsample <1 & colsample\_bytree < 1. For example, if you try with colsample\_by\_tree =1 (subsample ratio of columns=1 which means using all the columns) and change the order of the columns,you will get the same result ! (the same process with subsample and order of rows)

<https://www.kaggle.com/c/rossmann-store-sales/discussion/17014>