**Autoregressive Model**<https://www.statisticshowto.datasciencecentral.com/autoregressive-model/>

An autoregressive (AR) model **predicts future behavior based on past behavior**. It’s used for forecasting when there is some correlation between values in a time series and the values that precede and succeed them. You *only*use past data to model the behavior, hence the name*auto*regressive (the Greek prefix *auto*– means “self.” ). The process is basically a [linear regression](https://www.statisticshowto.datasciencecentral.com/probability-and-statistics/regression-analysis/find-a-linear-regression-equation/#definition) of the data in the current series against one or more past values in the same series.

In an AR model, the value of the [outcome variable (Y)](https://www.statisticshowto.datasciencecentral.com/dependent-variable-definition/#DVOutcome) at some point *t* in time is — like “regular” linear regression — directly related to the [predictor variable](https://www.statisticshowto.datasciencecentral.com/independent-variable-definition/#Predictor) (X). Where simple linear regression and AR models differ is that Y is dependent on X **and previous values**for Y.

The AR process is an example of a [stochastic](https://www.statisticshowto.datasciencecentral.com/stochastic-model/)process, which have degrees of uncertainty or randomness built in. The randomness means that you might be able to predict future trends pretty well with past data, but you’re never going to get 100 percent accuracy. Usually, the process gets “close enough” for it to be useful in most scenarios.

AR models are also called conditional models, [Markov models](https://www.statisticshowto.datasciencecentral.com/hidden-markov-model/), or transition models.

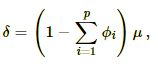
## AR(p) Models

An AR(p) model is an autoregressive model where specific lagged values of yt are used as [predictor variables](https://www.statisticshowto.datasciencecentral.com/independent-variable-definition/#Predictor). Lags are where results from one time period affect following periods.

**The value for “p” is called the order.** For example, an AR(1) would be a “first order autoregressive process.” The outcome variable in a first order AR process at some point in time t is related only to time periods that are one period apart (i.e. the value of the variable at t – 1). A second or third order AR process would be related to data two or three periods apart.

**The AR(p) model is defined by the equation:**  
yt = δ + φ1yt-1 + φ2yt-2 + … + φpyt-1 + At  
Where:

* yt-1, yt-2…yt-p are the past series values (lags),
* At is white noise (i.e. randomness),
* and δ is defined by the following equation:

[](https://www.statisticshowto.datasciencecentral.com/wp-content/uploads/2015/08/autoregressive-model.png)

*where μ is the process mean*