

Moving averages: what the number computes

MOVING AVERAGE

A moving average is the mean of the last n closing prices, recalculated each period. A simple moving average weights every price in the window equally; an exponential one weights recent prices more heavily. Both are summaries of prices that have already printed, so both necessarily lag - the lag is arithmetic, not a defect to be tuned away.

You can compute a moving average by hand, and doing it once removes most of the mystique.

The arithmetic

A 20-period simple moving average is the sum of the last 20 closes divided by 20. Next period, drop the oldest close, add the newest, divide again. That is the entire calculation.

An exponential moving average weights recent prices more. Each new value is the previous EMA plus a fraction of the gap between it and the newest close. A 20-period EMA uses a fraction of about 0.095. Recent prices count more, older ones fade rather than dropping out abruptly.

That is the only difference. An EMA responds faster to new prices and gives up some smoothness for it. Neither is more advanced.

Lag is not a flaw

An average of the last 20 closes cannot tell you anything the last 20 closes did not already contain. It is a summary, so it necessarily sits behind.

Traders often chase this with shorter periods. A 5-period average lags less and is noisier, which is the same trade in different clothes. There is no setting that removes lag, because removing lag would mean removing the averaging.

Anyone selling a moving average that anticipates price is selling a curve fitted to history.

What a crossover mechanically is

A 50-period average crossing above a 200-period one means the mean of the last 50 closes has risen above the mean of the last 200. That is the whole event.

It is commonly read as a shift from a shorter-term downtrend to an uptrend, and the crossing of those two particular periods is widely enough followed to have its own name. What it describes is arithmetic about closes that already happened.

Why averages act like levels

Price often turns near a well-watched moving average, and the explanation is the one from the previous lesson. Enough participants watch the 200-day that orders cluster around it. The average is not exerting anything; the orders placed because of it are.

This also predicts where averages will not work. An average nobody watches - a 37-period on a five-minute chart - has no such effect, because there is no crowd placing orders around it.

The honest position

Moving averages are good at one thing: making a trend definition mechanical, so you can state a rule and test it. Lesson two argued that a trend definition is only useful once it is explicit, and an average is one way of being explicit.

Treat crossovers as a description of what prices have done, not a signal about what they will do. And be wary of any period chosen because it worked on the chart in front of you - that is fitting, and it does not travel.

Check your understanding

1. Why does a moving average lag price?

- a. Because platforms calculate it on a delay
- b. Because it is an average of prices that have already printed
- c. Because it uses closes rather than intraday highs and lows

Answer: b. The lag is arithmetic. A summary of the last n closes cannot contain anything those closes did not, and shortening the period trades lag for noise rather than removing it.

2. A 50-period average crosses above a 200-period one. What has happened, precisely?

- a. The mean of the last 50 closes has risen above the mean of the last 200
- b. Buying pressure has exceeded selling pressure
- c. A new uptrend has started

Answer: a. That is the event in full. Traders commonly read it as a trend shift, and it is followed widely enough to attract orders, but the thing that occurred is arithmetic about closes already on the chart.