

# COMBINING SIGNALS

## 1. MULTIPLE PREDICTORS

We commonly think of strategies as relying on just one model, but of course any single-model strategy will compete poorly against more sophisticated rivals with multiple models driving revenue. Just as a single good model is worth testing and applying to multiple related securities, a security with good trading characteristics is worth spending the effort to derive several models. If we have spent that effort, we need to combine the models to form trading decisions.

## 2. SIMPLE CASE: GAUSSIAN RANGES

In the simplest case, we have a set of a set of independent models providing return predictions  $p_1, \dots, p_K$  for our security, each with gaussian error bars  $\sigma_1, \dots, \sigma_K$ . The model independence lets us then combine into a single prediction

$$p \sim N \left( \frac{\sum p_k \sigma_k^2}{\sum \sigma_k^2}, \left( \sum \sigma_k^{-2} \right)^{-\frac{1}{2}} \right)$$

However, this equation is quite limited, unrealistic due to underlying assumptions, and so rather misleading as well.

Most importantly, the errors models make are far from independent. In finance, essentially all models are very weak predictors, and so any impossible-to-model phenomenon will fool them all equally. On the other hand, their uniform weakness gives a useful approximation, namely that all their errors are equal to each other. Therefore all the  $\sigma_k$  are the same. This gives rise to the new, simpler equation

$$p \sim N \left( \text{Mean}(\{p_k\}), \sigma \right)$$

**2.1. Accumulation Criteria.** We might take on a position if  $|p|$  is large enough for us to have a high expectation of tolerable return on capital. For example, we can set a *hurdle rate*  $h$ , and demand that

$$L \cdot (p - \sigma) > h$$

where  $L$  is our leverage ratio.

**2.2. Calibration And Weighting.** If we wish to retain some sense of certain models being “better” than others on an *ad hoc* basis, we can maintain it by using a weighted average

$$p \sim N\left(\sum w_k p_k, \sigma\right)$$

One reason to do this is that many models are far from gaussian in errors, and tricky to calibrate. Post calibration, where the tails have been dealt with by excess caution, they often underpredict, with  $|p_k|$  biased toward zero in order to avoid their rare but expensive mistakes.

### 3. VOTING AND CENSORSHIP

If we use weighted sums, then even in cases where our models work in very different ways, they need to somehow be calibrated for compatibility with our averaging formula for combining them. We may instead wish to give each model a vote in some system designed to combine discrete outputs.

A common form of this approach is to force each of our  $K$  models to make a vote  $v_k$  of  $-1, 0$ , or  $1$  (corresponding to *short*, *neutral*, and *buy*), and then demand that some supermajority of models agree on a position to take. For example, we might demand that

$$\left| \text{Mean}(\{p_k\}) \right| > b$$

for some  $b \approx 0.75$ . Larger values of  $b$  will involve greater agreement among models, but fewer cases when we actually take on positions, providing a clear tradeoff between per-trade profit and strategy capacity.

We can include *vetos* as well, demanding for example that we take a position only if no models, or only one or two models, disagree with it. Vetos are a powerful limit on trading, of course. In the independent case, 8 models each with a veto will reduce our position-taking to less than 1% of its invulnerable level. For models with veto power, we sometimes say that they are *censoring* the trades.

**3.1. Compatibility With Return Predictions.** If we have some models with return predictions, and other models suitable only for voting or censorship, we can see that the two concepts can be combined. The main idea is to convert our combination of return models into a single (or multiple) voting models, by choosing a reasonable hurdle

rate. We can then proceed with voting as before. naturally since our combination of predictive return models is one we are likely to consider a “good” voter, we probably want to give it veto power.

#### 4. MODEL TIME HORIZONS

So far, we have ignored that different models operate at different time frequencies. A model that is expected to operate monthly is quite different from one with a time horizon of a few minutes or hours. Our returns are predicted to be *realized* in very different patterns.

A good starting assumption for a model prediction with input data of age about  $\tau$  is that its predictions will decay in about  $\tau/4$ . Obviously, this decay is measurable, so we do not need to worry too much about choosing a value arbitrarily. Once we have a decay number, we prefer a simple model for prices over time, namely, starting with present price  $P_0$  at time  $t = 0$ , we can assume exponential decay of returns

$$P_t = P_0 \exp \left( 2r \left( 1 - 2^{-\frac{t}{\tau/4}} \right) \right)$$

Choosing some arbitrary canonical horizon  $\tau_c$  we then achieve a prediction for each of the models with characteristic time horizon  $\tau_k$  as

$$p_k^{(c)} = 2p_k \left( 1 - 2^{-\frac{\tau_c}{\tau_k/4}} \right)$$

This equation is a useful way to help combine models, but of course when time horizons differ wildly we may find that longer-term models fail to provide sizable predictions on the time scales at which short-term models make sense.

**4.1. Predictions Versus Trading Costs.** If we are subject to punitive trading costs, short term models begin to look less attractive. The need to quickly enter a position involves high trading costs that will counterbalance any predicted returns.

Given our equations for decay above, we can perform fairly precise estimates of the tradeoff between quick position accumulation and trading costs. We simply take some trading cost model with configurable accumulation rate  $a_t$ , and optimize our profit (or Sharpe ratio) over the decaying return expectations.