Percentage Error: What Denominator?

FINDINGS OF A SURVEY CONDUCTED BY KESTEN GREEN AND LEN TASHMAN

THE ISSUE

This is our second survey on the measurement of forecast error. We reported the results of our first survey in the Summer 2008 issue of *Foresight* (Green & Tashman, 2008). The question we asked in that survey was whether to define forecast error as Actual minus Forecast (A-F) or Forecast minus Actual (F-A). Respondents made good arguments for both of the alternatives.

In the current survey, we asked how *percentage forecast error* should be measured. In particular: What should the denominator be when calculating percentage error?

We posed the question to the International Institute of Forecasters discussion list as well as to *Foresight* subscribers, in the following way:

> To calculate a percentage error, it is better to use…
>
> | | (*Check or write in*) |
> |---|---|
> | 1. The actual value (A) as the denominator | [ ] |
> | 2. The forecast (F) as the denominator | [ ] |
> | 3. Neither (A) nor (F) but some other value | [ ] |
>
> I recommend my choice of denominator, because: _____

The first two options in the questionnaire have each been used when calculating the mean absolute percentage error (MAPE) for multiple forecast periods. The first option is the more traditional form.

One popular alternative to using either A or F as the denominator is to take an average of the two: (A+F)/2. Calculated over multiple forecast periods, this measure is most commonly called the symmetric MAPE (sMAPE) and has been used in recent forecasting competitions to compare the accuracy of forecasts from different methods. See, for example, www.neural-forecasting-competition.com/index.htm.

SURVEY RESULTS

We received 61 usable responses. 34 of these (a majority of 56%) preferred option 1: using the **Actual** as the denominator for the percentage error. 15% preferred option 2, using the **Forecast** as the denominator, while 29% chose option 3, something other than the actual or the forecast.

One respondent wrote: "*For our company, this issue led to a very heated debate with many strong points of view. I would imagine that many other organizations will go through the same experience.*"

Option 1

Percentage Error = Error / Actual * 100

Of the 34 proponents of using the Actual value for the denominator, 31 gave us their reasons. We have organized their responses by theme.

A. The Actual is the forecaster's target.

Actual value is the forecast target and therefore should represent the baseline for measurement.

The measure of our success must be how close we came to "the truth."

Actual is the "stake in the ground" against which we should measure variance.

Since forecasting what actually happened is always our goal, we should be comparing how well we did to the actual value.

We should measure performance against reality.

B. The Actual is the only consistent basis for comparing forecast accuracy against a benchmark or for judging improvement over time.

Actual is the only acceptable denominator because

it represents the only objective benchmark for comparison.

Without a fixed point of reference quantity in the denominator, you will have trouble comparing the errors of one forecast to another.

You want to compare the forecast to actuals and not the other way around. The actuals are the most important factor. It drives safety stock calculations that are based on standard deviation of forecast error calculations that use actuals as the denominator.

Forecast error is measured here as (actual-forecast)/ actual, for comparability to other studies.

C. The Actuals serve as the weights for a weighted MAPE.

Using the Actuals is more consistent for calculating a weighted average percentage error (WAPE) for a group of SKUs or even for the full product portfolio. Using actual value as denominator is providing the weight for the different SKUs, which is more understandable – one is weighting different SKUs based on their actual contribution. If we use F (forecast), this means we will weigh them based on the forecast – but this can be challenged as subjective. Someone may calculate the single SKU accuracy based on F as denominator, and then weigh according to Actual sales of each SKU, but this unnecessarily complicates the formula.

D. The Actual is the customary and expected denominator of the MAPE.

I would argue that the standard definition of "percent error" uses the Actual. The Actual is used without any discussion of alternatives in the first three textbooks I opened, it is used in most forecasting software, and it is used on Wikipedia (at least until someone changes it).

If you are creating a display that reads "percent error" or "MAPE" for others to read without further explanation, you should use Actual – this is what is expected.

Actual is the generally used and accepted formula; if you use an alternative, such as the Forecast, you might need to give it a new name in order to avoid confusion.

E. Use of the Actual gives a more intuitive interpretation.

If the forecast value is > the actual value, then the percentage error with the forecast in the denominator cannot exceed 100%, which is misleading. For example, if the Actual is 100 and the Forecast is 1,000, the average percentage error with Actual is 900% but with Forecast is only 90%. (Ed. note: See Table 1a for an illustrative calculation.)

The reason is pragmatic. If Actual is, say, 10 and Forecast is 20, most people would say the percentage error is 100%, not 50%. Or they would say forecast is twice what it should have been, not that the actual is half the forecast.

By relating the magnitude of the forecast error to an Actual figure, the result can be easily communicated to non specialists.

From a retail perspective, explaining "over-forecasting" when Forecast is the denominator seems illogical to business audiences.

F. Using the Forecast in the denominator allows for manipulation of the forecast result.

Utilizing the Forecast as the benchmark is subjective and creates the opportunity for the forecaster to manipulate results.

Use of the Actual eliminates "denominator management."

Using Forecast encourages high forecasting.

G. Caveats: There are occasions when the Actual can't be used.

Use of Actual only works for non-0 values of the Actual.

Table 1. Illustrative Calculations

A	F	Absolute Error	% Error with A	% Error with F	Avg A+F	% Error with Avg of A&F
1a. If the Forecast exceeds the Actual, the % error cannot exceed 100%.						
100	200	100	100%	50%	150	0.667
100	1000	900	900%	90%	550	164%
100	10000	9900	9900%	99%	5050	196%
1b. Illustration of the Symmetry of the sMAPE.						
100	50	50	50%	100%	75	67%
50	100	50	100%	50%	75	67%
1c. When the Actual equals zero, use of sMAPE always yields 200%.						
0	50	50	#DIV/0!	100%	25	200%
0	100	100	#DIV/0!	100%	50	200%

If you are trying to overcome difficulties related to specific data sets (e.g., low volume, zeroes, etc.) or biases associated with using a percentage error, then you may want to create a statistic that uses a different denominator than the Actual. However, once you do so, you need to document your nonstandard definition of "percentage error" to anyone who will be using it.

For me, the Actual is the reference value. But in my job I deal with long-term (5-10 years+) forecasts, and the Actual is seldom "actually" seen. And since you're asking this question, my suspicion tells me the issue is more complicated than this.

Option 2

Percentage Error = Error / Forecast * 100

Eight of the 9 respondents who preferred to use the Forecast value for the denominator provided their reasons for doing so. Their responses fell into two groups.

A. Using Forecast in the denominator enables you to measure performance against forecast or plan.

For business assessment of forecast performance, the relevant benchmark is the plan – a forecast, whatever the business term. The relevant error is percent variation from plan, not from actual (nor from an average of the two).

For revenue forecasting, using the Forecast as the denominator is considered to be more appropriate since the forecast is the revenue estimate determining and constraining the state budget. Any future budget adjustments by the governor and legislature due to changing economic conditions are equal to the percentage deviations from the forecasted amounts initially used in the budget. Therefore, the error as a percent of the forecasted level is the true measure of the necessary adjustment, instead of the more commonly used ratio of (actual-forecast)/actual.

It has always made more sense to me that the forecasted value be used as the denominator, since it is the forecasted value on which you are basing your decisions.

The forecast is what drives manufacturing and is what is communicated to shareholders.

You are measuring the accuracy of a forecast, so you divide by the forecast. I thought this was a standard approach in science and statistics.

If we were to measure a purely statistical forecast (no qualitative adjustments), we would use Actual value (A) as the denominator because statistically this should be the most consistent number. However, once qualitative input (human judgment) from sales is included, there is an element that is not purely statistical in nature.

For this reason, we have chosen to rather divide by forecast value (F) such that we measure performance to our forecast.

B. The argument that the use of Forecast in the denominator opens the opportunity for manipulation is weak.

The politicizing argument is very weak, since the forecast is in the numerator in any case. It also implies being able to tamper with the forecast after the fact, and that an unbiased forecast is not a goal of the forecasting process.

Option 1 or 2
Percentage Error = Error / [Actual or Forecast: It Depends] * 100

Several respondents indicated that they would choose A *or* F, depending on the purpose of the forecast.

Actual, if measuring deviation of forecast from actual values. Forecast, if measuring actual events deviated from the forecast.

If the data are always positive and if the zero is meaningful, then use Actual. This gives the MAPE and is easy to understand and explain. Otherwise we need an alternative to Actual in the denominator.

The actual value must be used as a denominator whenever comparing forecast performance over time and/or between groups. Evaluating performance is an assessment of how close the forecasters come to the actual or "true" value. If forecast is used in the denominator, then performance assessment is sullied by the magnitude of the forecasted quantity.

If Sales and Marketing are being measured and provided incentives based on how well they forecast, then we measure the variance of the forecast of each from the actual value. If Sales forecast 150 and Marketing forecast 60 and actual is 100, then Sales forecast error is (150-100)/150=33% while Marketing forecast error is (70-100)/70=43%. When Forecast is the denominator, then Sales appears to be the better forecaster – even though their forecast had a greater difference to actual.

When assessing the impact of forecast error on deployment and/or production, then forecast error should be calculated with Forecast in the denominator because inventory planning has been done assuming the forecast is the true value.

Option 3
Percentage Error = Error / [Something Other Than Actual or Forecast] * 100

One respondent indicated use of Actual or Forecast, whichever had the higher value. No explanation was given.

Three respondents use the average of the Actual and the Forecast.

Averaging actual and forecast to get the denominator results in a symmetrical percent-error measure. (Ed. note: See Table 1b for an illustration, and the article by Goodwin and Lawton (1999) for a deeper analysis of the symmetry of the sMAPE.)

There likely is no "silver bullet" here, but it might be worthwhile to throw into the mix using the average of F and A – this helps solve the division-by-zero issues and helps take out the bias. Using F alone encourages high forecasting; using A alone does not deal with zero actuals. (Ed. note: Unfortunately, the averaging of A and F does *not* deal with the zero problem. When A is zero, the division of the forecast error by the average of A and F always results in a percentage error equal to 200%, as shown in Table 1c below and discussed by Boylan and Syntetos [2006].)

I find the corrected sMAPE adequate for most empirical applications without implying any cost structure, although it is slightly downward biased. In company scenarios, I have switched to suggesting a weighted MAPE (by turnover, etc.) if it is used for decision making and tracking.

Four respondents suggest use of some "average of Actual values" in the denominator.

Use the mean of the series. Handles the case of intermittent data, is symmetrical, and works for cross section. (Ed. note: This recommendation leads to use of the MAD/Mean, as recommended by Kolassa and Schutz [2007].)

My personal favorite is MAD/Mean. It is stable, even for slow-moving items, it can be easily explained, and it has a straightforward percentage interpretation.

A median baseline, or trimmed average, using recent periods, provides a stable and meaningful denominator.

I prefer a "local level" as the denominator in all the error % calculations. (Ed. note: The local level can be thought of as a weighted average of the historical data.) *When using Holt-Winters, I use the level directly, as it is a highly reliable indication of the current trading level of the time series. In addition, it isn't affected by outliers and seasonality. The latter factors may skew readings (hence interpretations) dramatically and lead to incorrect decisions.*

With other types of forecasting – such as multivariate – there's always some "local constant" that can be used. Even a median of the last 6 months would do. The main problem that arises here is what to do when this level approaches zero. This – hopefully – does not happen often in any set of data to be measured. It would rather point, as a diagnostic, to issues other than forecasting that need dire attention.

Two respondents recommend that the denominator be the absolute average of the period-over-period differences in the data, yielding a MASE (Mean Absolute Scaled Error).

The denominator should be equal to the mean of the absolute differences in the historical data. This is better, for example, than the mean of the historical data, because that mean could be close to zero. And, if the data are nonstationary (e.g., trended), then the mean of the historical data will change systematically as more data are collected. However, the mean of the absolute differences will be well behaved, even if the data are nonstationary, and it will always be positive. It has the added advantage of providing a neat, interpretable statistic: the MASE. Values less than 1 mean that the forecasts are more accurate than the in-sample, naïve, one-step forecasts. (See Hyndman, 2006.)

Mean absolute scaled error, which uses the average absolute error for the random walk forecast (i.e., the absolute differences in the data).

FOLLOW-UP

We welcome your reactions to these results. Have they clarified the issue? Have they provided new food for thought? Have they changed your mind? See our contact information at bottom.

REFERENCES

Boylan, J. & Syntetos, A. (2006). Accuracy and accuracy-implication metrics for intermittent demand, *Foresight: The International Journal of Applied Forecasting*, Issue 4, 39-42.

Goodwin, P. & Lawton R. (1999). On the asymmetry of the symmetric MAPE, *International Journal of Forecasting*, 15, 405-408.

Green, K.C. & Tashman, L. (2008). Should we define forecast error as e= F-A or e= A-F? *Foresight*: *The International Journal of Applied Forecasting*, Issue 10, 38-40.

Hyndman, R. (2006). Another look at forecast-accuracy metrics for intermittent demand, *Foresight: The International Journal of Applied Forecasting*, Issue 4, 43-46.

Kolassa, S. & Schutz, W. (2007). Advantages of the MAD/MEAN ratio over the MAPE, *Foresight: The International Journal of Applied Forecasting*, Issue 6, 40-43.

CONTACT
Kesten@ForPrin.com
LenTashman@forecasters.org