Adaptive Thresholds FAQ

1. Why are my significance level thresholds missing?

Significance level thresholds are based on certain statistics computed over baseline metric data. Separate statistical computations are performed and saved for each of the so-called "Time Groups" associated with the baseline. The Time Groups are used to capture and profile several common database workload cycles, e.g. online during day vs. batch at night.

The statistics on which significance level thresholds depend are themselves dependent on the cardinality of the data within Time Groups. For example, it is not possible to make a useful statement about the 99.9^{th} percentile when the cardinality of data in a Time Group is 60 values.

Cardinality requirements

Significance level thresholds currently impose the following cardinality requirements:

- 0.95 = 100 values
- 0.99 = 500 values
- 0.999 = 700 values
- 0.9999 = 700 values

Cardinality Table

The following table shows Time Group cardinalities for several different System Moving Window baseline sizes. Note that the Weekday / Weekend time grouping introduces asymmetry in the size of groups (5 weekdays vs. 2 weekend days) so both values are listed. Note also that these values assume database is running and producing metric values over the entire moving window period, downtime will reduce data cardinalities.

Time Grouping (Day : Week)	Groups per week	Moving Window Size		
		1-week	3-weeks	5-weeks
X : X	1	10080		
X : W	2	7200 / 2880		
X : D	7	1440		
N : X	2	5040		
N : W	4	3600 / 1440		
N : D	14	720		
H : X	24	420	1260	
H : W	48	300 / 120	900 / 360	1500 / 600
H : D	168	60	180	300

Symbols:
X = All hours of day OR All days of week
W = Weekday / Weekend
D = Day of Week
N = Night / Day
H = Hour of Day

Significance Level Threshold Table

Combining the threshold cardinality requirements with Time Group cardinalities yields a table of maximum available significance level thresholds. If significance level is set above the maximum listed for by moving window size then thresholds will not be set because cardinality requirements are not satisfied.

Time Grouping (Day : Week)	Groups per week	Moving Window Size		
		1-week	3-weeks	5-weeks
X : X	1	0.9999		
X : W	2	0.9999		
X : D	7	0.9999		
N : X	2	0.9999		
N : W	4	0.9999		
N : D	14	0.9999		
H : X	24	0.95	0.9999	
H : W	48	0.95 / 0.95	0.9999 / 0.95	0.9999 / 0.99
H : D*	168	NA	0.95	0.95

* System Moving Window must be at least 12 weeks to allow for 0.9999 significance level thresholds.

Goodness of fit

There is an additional requirement in the case of the two highest significance level settings, 0.999 and 0.9999. The thresholds for these significance levels are determined by fitting a model to the data observations for each time group, and extrapolating from the model to get the threshold values. When the model is determined to be a bad fit for the data observed then thresholds from that model are deemed not good enough to set and will be suppressed.

Conclusion

Significance level thresholds depend on data cardinalities and goodness of fit which themselves depend on time grouping, moving window size, and the data itself. In particular, time grouping that includes Hour of Day partitions data such that quite large moving windows may be required to generate enough samples for the highest significance level thresholds to be reliably determined. Thus there is no guarantee in the significance level case that thresholds will be set.