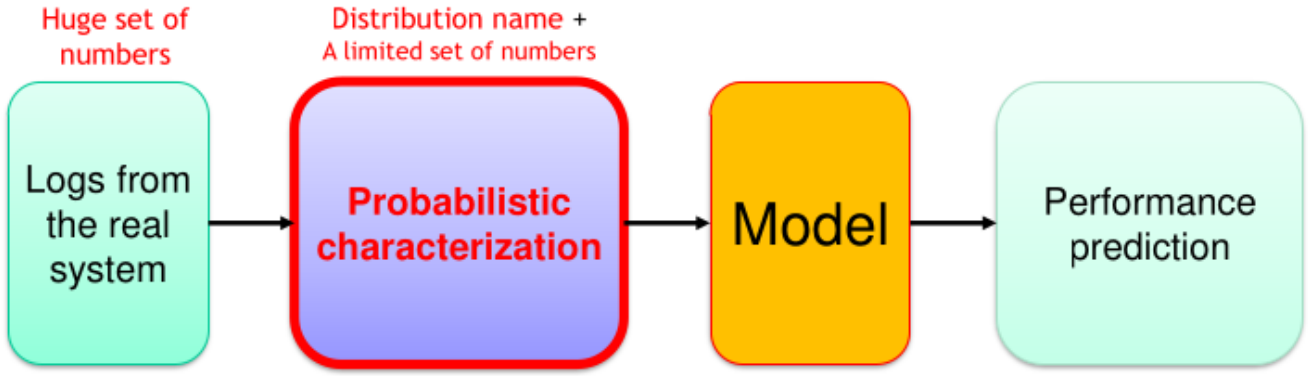
06.Fitting



In Performance Evaluation, the distributions allow us to reduce a real trace to a limited set of parameters (values) that can completely describe the corresponding trace. Moreover, from a probability distribution, traces of arbitrary length can be easily generated, allowing to compute performance metrics with a desired accuracy

For example, the distributions presented in the previous lesson require:

Distribution	N. parameters	Meaning
Discrete (*k* values)	*k-1*	*k* choices
Geometric	*1*	Success probability
Poisson	*1*	Rate
Deterministic	*1*	Value
Uniform	*2*	Minimum, maximum
Exponential	*1*	Rate
Hyper-exponential (*k* branches)	*2k - 1*	*k - 1* choices, *k* rates
Hypo-exponential (*k* stages)	*k*	k rates
Erlang	*2*	Rate, stages
Hyper-Erlang (k branches)	*3k - 1*	*K - 1* choices, *k* rates, *k* stages

The coefficient of variation

We can check which distribution using the coefficient of variation. Before determining the parameters, we must choose an appropriate distribution. The coefficient of variation of a set of data can guide the choice of the most appropriate distribution. Hypo-exponentials are generally used to model process where the composing stages can be measured separately. Usually

Erlang are preferred to model c.v. less than one since they are easier to manage than Hypo-exponentials.

Parameter estimation and fitting

When a distribution has been selected, its parameters must be estimated from measurements acquired from the corresponding real system. The process of determining the best parameters that would make a distribution produce samples with characteristics similar to the ones measured in the real trace is called parameter estimation or fitting.
The two main methods available are:

- Method of moments
- Maximum likelihood method