

Statistical Tests Implemented in SReach

Qinsi Wang

Computer Science Department, Carnegie Mellon University, USA

In this section we briefly describe the statistical techniques implemented in **SReach**. To deal with qualitative questions, **SReach** supports the following hypothesis testing methods.

Lai's test [4]. As a simple class of sequential tests, it tests the one-sided composite hypotheses $H_0 : \theta \leq \theta_0$ versus $H_1 : \theta \geq \theta_1$ for the natural parameter θ of an exponential family of distributions under the 0 – 1 loss and cost c per observation. [4] shows that these tests have nearly optimal frequentist properties and also provide approximate Bayes solutions with respect to a large class of priors.

Bayes factor test [3]. The use of Bayes factors is a Bayesian alternative to classical hypothesis testing. It is based on the Bayes theorem. Hypothesis testing with Bayes factors is more robust than frequentist hypothesis testing, as the Bayesian form avoids model selection bias, evaluates evidence in favor of the null hypothesis, includes model uncertainty, and allows non-nested models to be compared. Also, frequentist significance tests become biased in favor of rejecting the null hypothesis with sufficiently large sample size.

Bayes factor test with indifference region. A hypothesis test has ideal performance if the probability of the Type-I error (respectively, Type-II error) is exactly α (respectively, β). However, these requirements make it impossible to ensure a low probability for both types of errors simultaneously (see [6] for details). A solution is to use an indifference region. The indifference region indicates the distance between two hypotheses, which is set to separate the two hypotheses.

Sequential probability ratio test (SPRT) [5]. The SPRT considers a simple hypothesis $H_0 : \theta = \theta_0$ against a simple alternative $H_1 : \theta = \theta_1$. With the critical region A_n and two thresholds A , and B , SPRT decides that H_0 is true and stops when $A_n < A$. It decides that H_1 is true and terminates if $A_n > B$. If $A < A_n < B$, it will collect another observation to obtain a new critical region A_{n+1} . The SPRT is optimal, among all sequential tests, in the sense that it minimizes the average sample size.

To offer quantitative answers, **SReach** also supports estimation procedures as below.

Chernoff-Hoeffding bound [2]. To estimate the mean p of a (bounded) random variable, given a precision δ' and coverage probability α , the Chernoff-Hoeffding bound computes a value p' such that $|p' - p| \leq \delta'$ with probability at least α .

Bayesian Interval Estimation with Beta prior [7]. This method estimates p , the unknown probability that a random sampled model satisfies a specified reachability property. The estimate will be in the form of a confidence interval, containing p with an arbitrary high probability. [7] assumes that the unknown

p is given by a random variable, whose density is called the prior density, and focuses on Beta priors.

Direct sampling. Given N as the number of samples to be sampled, the direct sampling method estimates the mean of p of a (bounded) random variable. According to the central limit theorem [1], the error ϵ with a confidence c between the real probability p and the estimated \hat{p} is bounded:

$$\epsilon = \phi^{-1}\left(\frac{c+1}{2}\right) \sqrt{\frac{p(1-p)}{N}}$$

where $\phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-x}^x e^{-t^2/2} dt$. That is, as N goes to ∞ , the estimated probability approaches to the real one.

References

1. R. Durrett. *Probability: theory and examples*. Cambridge University Press, 2010.
2. W. Hoeffding. Probability inequalities for sums of bounded random variables. *J American Statistical Association*, 58(301):13–30, 1963.
3. R. E. Kass and A. E. Raftery. Bayes factors. *JASA*, 90(430):773–795, 1995.
4. T. L. Lai. Nearly optimal sequential tests of composite hypotheses. *AOS*, 16(2):856–886, 1988.
5. A. Wald. Sequential tests of statistical hypotheses. *The Annals of Mathematical Statistics*, 16(2):117–186, 1945.
6. H. L. Younes. Verification and planning for stochastic processes with asynchronous events. Technical report, DTIC Document, 2005.
7. P. Zuliani, A. Platzer, and E. M. Clarke. Bayesian statistical model checking with application to stateflow/simulink verification. *Formal Methods in System Design*, 43(2):338–367, 2013.