Statistical Tests in SReach

Qinsi Wang

Computer Science Department, Carnegie Mellon University, USA

1 Statistical tests

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 [3]. 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. [3] 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 [2]. 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 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 [5] for details). A solution is to use an indifference region. The indifference region indicates the distance between two hypotheses, which is set to separate two hypotheses.

Sequential probability ratio test (SPRT) [4]. As for he SPRT, we consider a simple hypothesis $H_0 :\ \theta = \theta_0$ against a simple alternative $H_1 :\ \theta = \theta_1$. With the critical region Λ_n and two thresholds A, and B, SPRT decides that H_0 is true and stops when $\Lambda_n < A$. It decides that H_1 is true and terminates if $\Lambda_n > B$. If $A\ < \Lambda_n < B$, it will collect another observation to obtain a new critical region Λ_{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 [1]. To estimate the probability p for H to satisfy ϕ, given a precision δ', the Chernoff-Hoeffding bound can given a value p' such that $|p' - p| \leq \delta'$ with confidence $1 - \alpha$.

Bayesian Interval Estimation with Beta prior [6]. 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, congaing p with an arbitrary high probability. [6] assumes that the unknown p is given by a random variable, whose density is called the prior density, and focuses on Beta priors. It has been showed that, with this Bayesian interval estimation method, the probability of giving a wrong answer is arbitrarily small, and speed of obtaining an answer is higher than the sequential hypothesis testing.

References

1. W. Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963.
2. R. E. Kass and A. E. Raftery. Bayes factors. *Journal of the American Statistical Association*, 90(430):773–795, 1995.
3. T. L. Lai. Nearly optimal sequential tests of composite hypotheses. *The Annals of Statistics*, pages 856–886, 1988.
4. A. Wald. Sequential tests of statistical hypotheses. *Annals of Mathematical Statistics*, 16(2):117–186, 1945.
5. H. L. Younes. Verification and planning for stochastic processes with asynchronous events. Technical report, DTIC Document, 2005.
6. P. Zuliani, A. Platzer, and E. M. Clarke. Bayesian statistical model checking with application to Simulink/Stateflow verification. In *HSCC*, pages 243–252, 2010.