Risk Comparison and Planning for Bayesian Assurance Tests

Hyoshin Kim, Kate Sanborn, Alyson Wilson
NC State University
Department of Statistics

Motivation

- We are interested in assessing the reliability of a system by doing a series of pass/fail tests.

- How many tests should we do?

Motivation

- Formally, we select a ***binomial test plan***, (n,c):
  - Number of test units *n*
  - Maximum number of failures *c*

- If we observe ***fewer than c*** failures in n tests, the system "passes."

- There are many ways to choose a test plan (n,c) so that we can balance our risks and resources.

Questions of Interest

- ***Questions***
  - What are the risks that we want to control?
  - What are some of the commonly used methods for choosing a test plan (n,c)?
  - How do we compare across test plans (n,c) found using different methods?

Risk Metrics

- We consider 7 risk metrics that we may control to choose a test plan (n,c).
- ***[Classical Risks]*** 2 risk metrics that make inference based on the pass/fail information we will observe in our *n* tests.
  - Consumer's Risk (CR), Producer's Risk (PR)
- ***[Bayesian Risks]*** 5 risk metrics that make inference based on the pass/fail information we will observe ***and*** relevant pass/fail data we have observed previously.
  - Assurance, Posterior Consumer's Risk (PCR), Posterior Producer's Risk (PPR), Effect Size, Average Length Criteria (ALC)

Methods for Selecting a Test Plan

- We compare 5 methods for choosing a test plan (n,c).
- 2 methods that control Classical Risks.
  - ***OC Curve*** and ***Sequential Probability Ratio Test (SPRT)***: controls Consumer's Risk (CR), Producer's Risk (PR)
- 3 methods that control Bayesian Risks.
  - ***Assurance Reliability Demonstration Testing (ARDT)***: controls Assurance
  - ***Bayesian Assurance (BA)***: controls Posterior Consumer's Risk (PCR), Posterior Producer's Risk (PPR)
  - ***Bayesian Power (BP)***: controls Effect Size or Average Length Criteria (ALC)

Results Example: Bayesian Assurance

Method	n	c	CR	PR	Assurance	PCR	PPR	Effect Size	ALC
BA	108	0	0.012	0.338	0.440	**0.050**	**0.002**	0.950	0.045

- The ***Bayesian Assurance*** method chooses test plans that control risks PCR at 5% and PPR at 10%.
- Out of all possible test plans (n,c) that control the risks, we choose the test plan with smallest n and biggest c: ***(108, 0)***.
- There is no universally "best" method that controls all 7 risks.
- Methods that control Bayesian risks tend to give smaller test plans and incorporate additional relevant information into the planning.

References

- **OC Curves**: MIL-HDBK-781, Revision A, April 1, 1996: Reliability Test Methods, Plans, and Environments for Engineering Development, Qualification, and Production.
- **SPRT:** A Wald (1945). Sequential Tests of Statistical Hypotheses, The Annals of Mathematical Statistics 16(2): 117–86.
- **Bayesian Power:** AE Gelfand and F Wang (2002). A simulation-based approach to Bayesian sample size determination for performance under a given model and for separating models, Statistical Science 17(2): 193-208.
- **Bayesian Assurance:** MS Hamada, AG Wilson, CS Reese, HF Martz. *Bayesian Reliability*, Springer, New York, 2011.
- **ARDT:** KJ Wilson and M Farrow (2021). Assurance for Sample Size Determination in Reliability Demonstration Testing, Technometrics, 63(4): 523-535.