

Correctness and Power Guarantees for Statistical Model Checking

Daniël Reijsbergen Pieter-Tjerk de Boer
Werner Scheinhardt Boudewijn Haverkort

ROCKS meeting, April 26, 2012

1 Introduction

- Statistics
- Hypothesis Testing
- Example: Cola Cans

2 Statistical Model Checking

- Model Checking
- Hypothesis Testing for Model Checking

3 Conclusion

1 Introduction

- Statistics
- Hypothesis Testing
- Example: Cola Cans

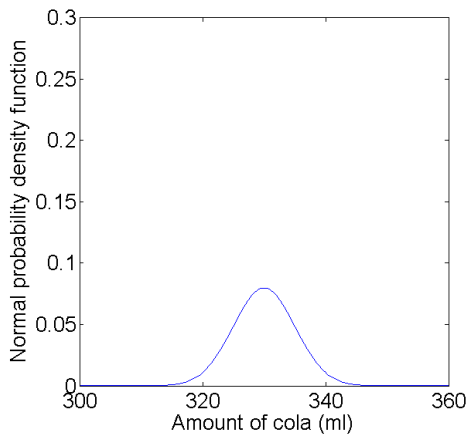
2 Statistical Model Checking

- Model Checking
- Hypothesis Testing for Model Checking

3 Conclusion

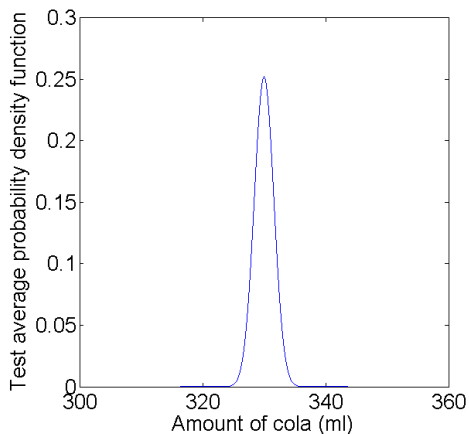
*“A statistical hypothesis test is a method of **making decisions** using **experimental data**.”*

Hypothesis Test I



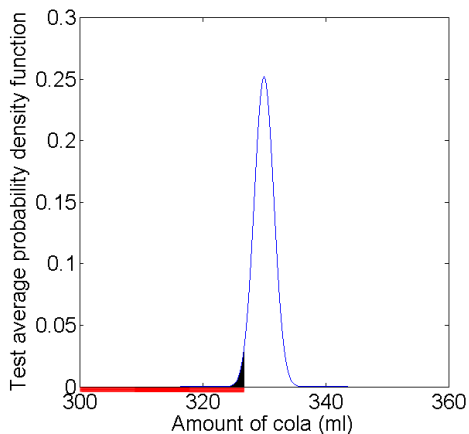
$$\mu = 330, \sigma = 5$$

Hypothesis Test II



$$\mu = 330, \sigma = 5/\sqrt{10}$$

Hypothesis Test III



Critical area $\approx (-\infty, 327.4)$

1 Introduction

- Statistics
- Hypothesis Testing
- Example: Cola Cans

2 Statistical Model Checking

- Model Checking
- Hypothesis Testing for Model Checking

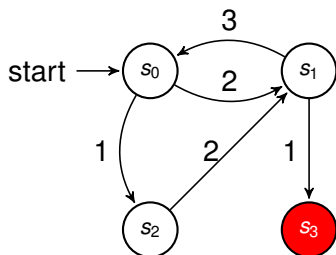
3 Conclusion

Hypothesis Testing for Model Checking I

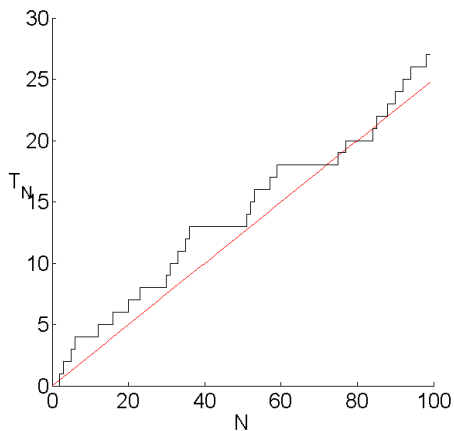
Is the probability that I reach s_3 before time t smaller than 25%?

Computer program:

- reach $s_3 \Rightarrow$ return 1; quit;
- reach $t \Rightarrow$ return 0; quit;

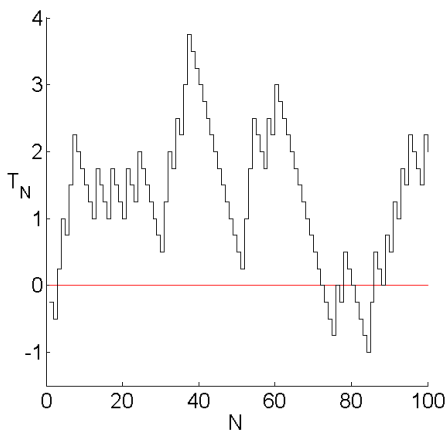


Hypothesis Testing for Model Checking II



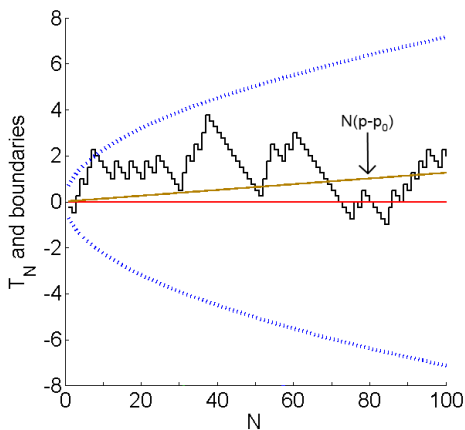
$$T_N = \sum_{i=1}^N X_i, p_0 = 0.25$$

Hypothesis Testing for Model Checking III



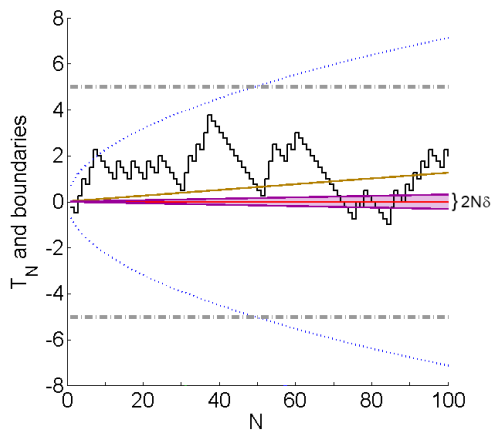
$$T_N = \sum_{i=1}^N (X_i - p_0), \quad p_0 = 0.25$$

Hypothesis Testing for Model Checking IV



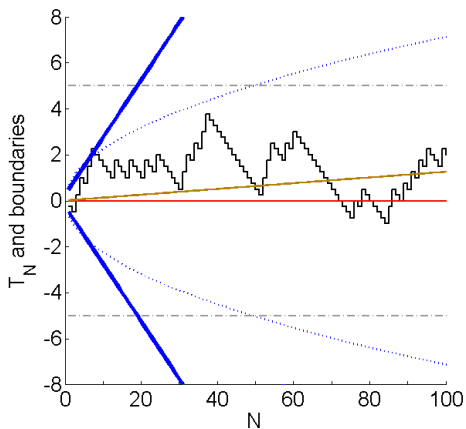
Fixed sample size test ('confidence intervals').

Hypothesis Testing for Model Checking V



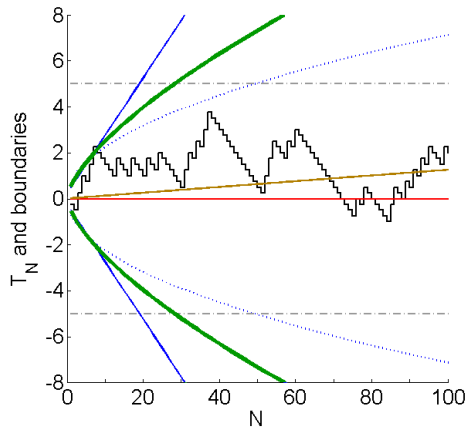
Sequential test (Younes).

Hypothesis Testing for Model Checking VI



Linearly diverging boundaries.

Hypothesis Testing for Model Checking VII



Better.

Outline

1 Introduction

- Statistics
- Hypothesis Testing
- Example: Cola Cans

2 Statistical Model Checking

- Model Checking
- Hypothesis Testing for Model Checking

3 Conclusion

Conclusion

- Existing tests have shortcomings:
 - 'Confidence intervals': depend on fixed N .
 - Younes: depends on δ .
- Better test \rightarrow not have shortcomings.

Future Work:

- Improve bounds.
- Generalise results: steady-state, importance sampling.

Thank you for your attention.