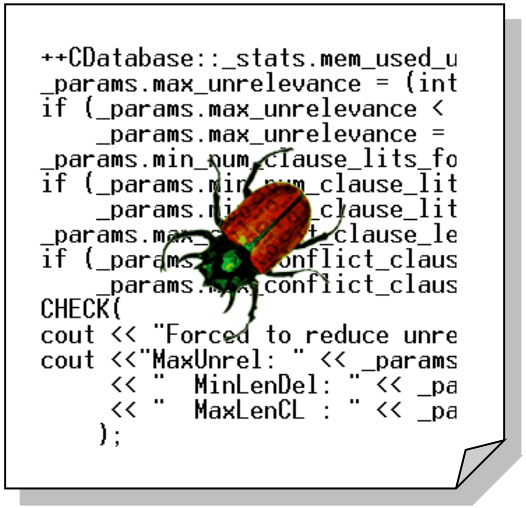
Bayesian Statistical Model Checking with Application to Stateflow/Simulink Verification

Paolo Zuliani

André Platzer

Edmund M. Clarke



Computer Science Department

Carnegie Mellon University

Problem

Verification of Stochastic Systems

- Uncertainties in the system environment, modeling a fault, stochastic processors, biological signaling pathways ...
  - Modeling uncertainty with a distribution → Stochastic systems
- Models:
  - for example, Discrete, Continuous Time Markov Chains
- Property specification:
  - "does the system fulfill a request within 1.2 ms with probability at least .99"?
- If Φ = "system fulfills request within 1.2 ms", decide between:

$$P_{\geq .99}(\Phi) \quad \text{or} \quad P_{<.99}(\Phi)$$

Equivalently

- A biased coin (Bernoulli random variable):
  - Prob (Head) = p Prob (Tail) = $1-p$
  - p is unknown
- Question: Is $p \geq \theta$? (for a fixed $0<\theta<1$)
- A solution: flip the coin a number of times, collect the outcomes, and use:
  - Statistical hypothesis testing: returns *yes/no*
  - Statistical estimation: returns "p in (a,b)" (and compare a with θ)

Motivation

- State Space Exploration infeasible for large systems
  - Symbolic MC with OBDDs scales to 10^{300} states
  - Scalability depends on the structure of the system
- Pros: Simulation is feasible for **many more** systems
  - Often easier to simulate a complex system than to build the transition relation for it
  - Easier to parallelize
- Cons: answers may be wrong
  - But error probability can be bounded

Towards verification

Property Φ

+

*Stochastic system **M***

=

Biased coin

Key: define a probability measure on the set of traces (simulations) of ***M***.
The set of traces satisfying *Φ* is measurable.

Statistical Model Checking

Key idea

- Suppose system behavior w.r.t. a (fixed) property Φ can be modeled by a Bernoulli random variable of parameter p:
  - System satisfies Φ with (unknown) probability p

- Question: $P_{\geq\theta}(\Phi)$? (for a fixed $0<\theta<1$)
- Draw a sample of system simulations and use:
  - Statistical hypothesis testing: Null vs. Alternative hypothesis

$$H_0 : \mathcal{M} \models P_{\geqslant\theta}(\phi) \qquad H_1 : \mathcal{M} \models P_{<\theta}(\phi)$$

  - Statistical estimation: returns "p in (a,b)" (and compare a with θ)

Bayesian Statistical Model Checking

- MC chooses between two mutually exclusive hypotheses

Null Hypothesis	$H_0 : \mathcal{M} \models P_{\geqslant\theta}(\phi)$
vs	
Alternate Hypothesis	$H_1 : \mathcal{M} \models P_{<\theta}(\phi)$

- We have developed a new statistical MC algorithm
  - Sequential sampling
  - Performs Composite Hypothesis Testing and Estimation
  - Based on Bayes Theorem and the Bayes Factor.

Bayesian Statistics

Three ingredients:

1. Prior probability
   - Models our initial (a priori) uncertainty/belief about parameters (what is Prob($p \geq \theta$) ?)

2. Likelihood function
   - Describes the distribution of data (*e.g.*, a sequence of heads/tails), given a specific parameter value

3. Bayes Theorem
   - Revises uncertainty upon experimental data - compute Prob($p \geq \theta \mid data$)

Sequential Bayesian Statistical MC - I

- Model Checking $H_0 : \mathcal{M} \models P_{\geqslant\theta}(\phi)$ $H_1 : \mathcal{M} \models P_{<\theta}(\phi)$
- Suppose $\mathcal{M}$ satisfies ϕ with (unknown) probability p
  - p is given by a random variable (defined on [0,1]) with density g
  - g represents the prior belief that $\mathcal{M}$ satisfies ϕ
- Generate independent and identically distributed (iid) sample traces.
- x_i: the i^{th} sample trace σ satisfies ϕ
  - $x_i = 1$ iff $\sigma_i \models \phi$
  - $x_i = 0$ iff $\sigma_i \not\models \phi$
- Then, x_i will be a Bernoulli trial with conditional density (likelihood function)

$$f(x_i|u) = u^{x_i}(1 - u)^{1-x_i}$$

Sequential Bayesian Statistical MC - II

- $X = (x_1, \ldots, x_n)$ a sample of Bernoulli random variables
- Prior probabilities $P(H_0)$, $P(H_1)$ strictly positive, sum to 1
- Posterior probability (Bayes Theorem [1763])

$$\boxed{P(H_0|X) = \frac{P(X|H_0)P(H_0)}{P(X)}}$$

for $P(X) > 0$

- Ratio of Posterior Probabilities:

$$\boxed{\frac{P(H_0|X)}{P(H_1|X)} = \frac{P(X|H_0)}{P(X|H_1)} \cdot \frac{P(H_0)}{P(H_1)}}$$

Bayes Factor

Sequential Bayesian Statistical MC - III

- Recall the Bayes factor $B = \frac{P(X|H_0)}{P(X|H_1)}$

- Jeffreys' [1960s] suggested the Bayes factor as a statistic:
  - For fixed sample sizes
  - For example, a Bayes factor greater than 100 "strongly supports" H_0

- We introduce a sequential version of Jeffrey's test

- Fix threshold $T \geq 1$ and prior probability.
  Continue sampling until
  - Bayes Factor > T: Accept H_0
  - Bayes Factor < $1/T$: Reject H_0

Sequential Bayesian Statistical MC - IV

Require: ***Property*** $P_{\geq\theta}(\Phi)$, ***Threshold*** $T \geq 1$, ***Prior density*** g

```
n := 0              {number of traces drawn so far}
x := 0              {number of traces satisfying Φ so far}
repeat
        σ := draw a sample trace of the system (iid)
        n := n + 1
        if  σ ⊨ Φ  then
          x := x + 1
        endif
        𝓑 := BayesFactor(n, x, θ, g)
until (𝓑 > T  ∨ 𝓑 < 1/T )
if (𝓑 > T ) then
          return "H_0 accepted"
else
          return "H_0 rejected"
endif
```

Correctness

Theorem (Error bounds). When the Bayesian algorithm – using threshold T – stops, the following holds:

$$\boxed{\begin{array}{l} \text{Prob}(\text{“accept } H_0\text{”} \mid H_1) \leq 1/T \\ \text{Prob}(\text{“reject } H_0\text{”} \mid H_0) \leq 1/T \end{array}}$$

Note: bounds independent from the prior distribution.

Computing the Bayes Factor - I

Definition: Bayes Factor of sample *X* and hypotheses H_0, H_1 *is*

$$\frac{P(H_0|X)}{P(H_1|X)} \cdot \frac{P(H_1)}{P(H_0)} = \frac{\int_\theta^1 \overbrace{f(x_1|u)\cdots f(x_n|u)}^{\text{joint (conditional) density of independent samples}} \cdot g(u)\ du}{\int_0^\theta f(x_1|u)\cdots f(x_n|u)\cdot g(u)\ du} \cdot \frac{1-\pi_0}{\pi_0}$$

- $\pi_0 = P(H_0) = \int_\theta^1 g(u)du$ prior *g* is Beta of parameters *α>0, β>0*

$$g(u) = \frac{1}{B(\alpha,\beta)} u^{\alpha-1}(1-u)^{\beta-1}$$

$$B(\alpha,\beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1}\ dt$$

Computing the Bayes Factor - II

Proposition

The Bayes factor of $H_0{:}\mathcal{M} \models P_{\geq\theta}(\Phi)$ vs $H_1{:}\mathcal{M} \models P_{<\theta}(\Phi)$ for n Bernoulli samples (with $x \leq n$ successes) and prior Beta(α,β)

$$B = \frac{1-\pi_0}{\pi_0} \cdot \left(\frac{1}{F_{(x+\alpha,n-x+\beta)}(\theta)} - 1 \right)$$

where $F_{(\cdot,\cdot)}(\cdot)$ is the Beta distribution function.

$$F_{(x+\alpha,n-x+\beta)}(\theta) = \frac{1}{B(x+\alpha,n-x+\beta)} \int_0^\theta u^{x+\alpha-1}(1-u)^{n-x+\beta-1}\, du$$

- No need of integration when computing the Bayes factor

Bayesian Interval Estimation - I

- Estimating the (unknown) probability *p* that "system $\models \Phi$"
- Recall: system is modeled as a Bernoulli of parameter *p*
- *Bayes' Theorem* (for iid Bernoulli samples)

$$f(u \mid x_1, \ldots, x_n) = \frac{f(x_1 \mid u) \cdots f(x_n \mid u) g(u)}{\int_0^1 f(x_1 \mid v) \cdots f(x_n \mid v) g(v) \, dv}$$

- We thus have the posterior distribution
- So we can use the mean of the posterior to estimate *p*
  - mean is a posterior Bayes estimator for *p* (it minimizes the integrated risk over the parameter space, under a quadratic loss)

Bayesian Interval Estimation - II

- By integrating the posterior we get Bayesian intervals for *p*
- Fix a coverage ½ < *c* < 1. Any interval (t_0, t_1) such that

$$\int_{t_0}^{t_1} f(u \mid x_1, \ldots, x_n)\, du = c$$

  is called a 100c percent Bayesian Interval Estimate of *p*
- *An optimal interval* minimizes $t_1 - t_0$: difficult in general
- Our approach:
  - fix a half-interval width δ
  - Continue sampling until the posterior probability of an interval of width 2δ containing the posterior mean exceeds coverage c

Bayesian Interval Estimation - III

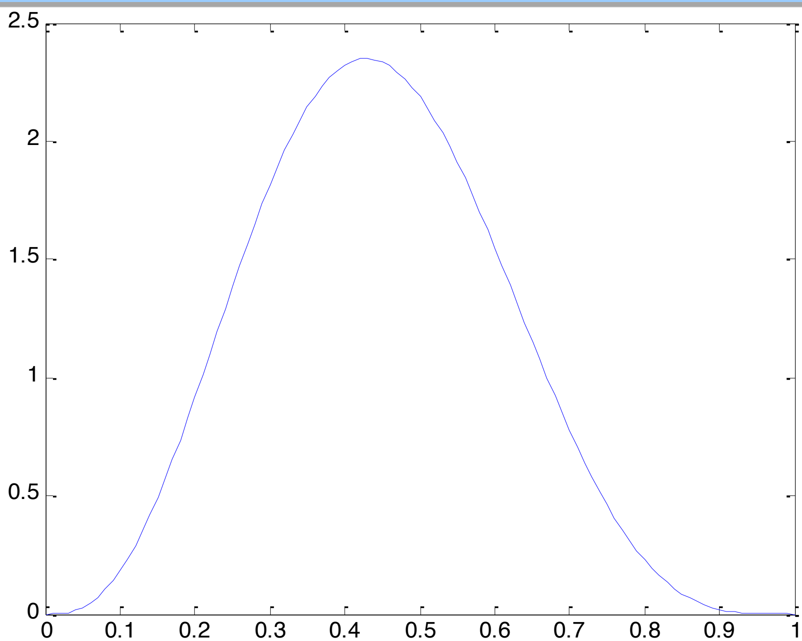
- Computing the posterior probability of an interval is easy

- Suppose *n* Bernoulli samples (with *x*≤*n* successes) and prior Beta(*α*,*β*)

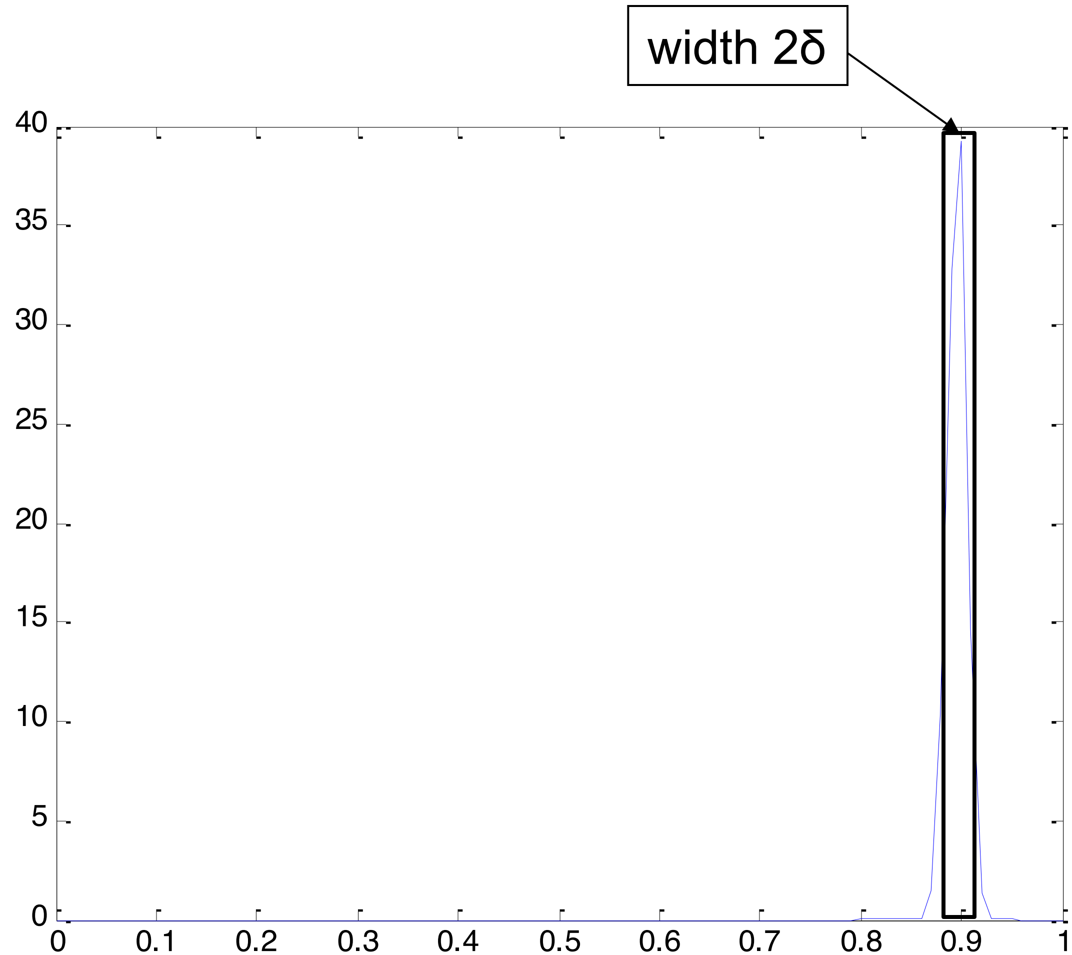
$$
\begin{aligned}
\mathrm{P}(t_0 < p < t_1 | x_1, \ldots, x_n) &= \int_{t_0}^{t_1} f(u \mid x_1, \ldots, x_n)\; du \\
&= \boxed{F_{(x+\alpha, n-x+\beta)}(t_1) - F_{(x+\alpha, n-x+\beta)}(t_0)}
\end{aligned}
$$

- No numerical integration

Bayesian Interval Estimation - IV



prior is beta(α=4,β=5)



posterior density after 1000 samples and 900 “successes” is beta(α=904,β=105)

posterior mean = 0.8959

Bayesian Interval Estimation - V

Require: **BLTL** ***property*** *Φ*, ***interval-width*** *δ*, ***coverage*** *c*, ***prior*** *beta parameters α,β*

$n := 0$ {*number of traces drawn so far*}

$x := 0$ {*number of traces satisfying so far*}

repeat

  $\sigma :=$ draw a sample trace of the system (iid)

  $n := n + 1$

  if $\sigma \models \Phi$ **then**

    $x := x + 1$

  endif

  mean = (x+α)/(n+α+β)

  (t_0,t_1) = (mean-δ, mean+δ)

  $\mathcal{I} := PosteriorProbability\ (t_0,t_1,n,x,\alpha,\beta)$

until $(\mathcal{I} > c)$

return (t_0,t_1), mean

Bayesian Interval Estimation - VI

- Recall the algorithm outputs the interval (t_0, t_1)
- Define the null hypothesis

$$H_0: t_0 < p < t_1$$

- We can use the previous results for hypothesis testing

Theorem (Error bound). When the Bayesian estimation algorithm (using coverage $½ < c < 1$) stops – we have

> Prob (“accept H_0” | H_1) $\leq (1/c - 1)\pi_0/(1-\pi_0)$
> Prob (“reject H_0” | H_0) $\leq (1/c - 1)\pi_0/(1-\pi_0)$

π_0 is the prior probability of H_0

Bounded Linear Temporal Logic

- Bounded Linear Temporal Logic (BLTL): Extension of LTL with time bounds on temporal operators.

- Let $\sigma = (s_0, t_0), (s_1, t_1), \ldots$ be an execution of the model
  - along states $s_0, s_1, \ldots$
  - the system stays in state s_i *for time* t_i
  - divergence of time: $\Sigma_i\, t_i$ diverges (i.e., non-zeno)

- σ^i: Execution trace starting at state i.
- A model for simulation traces (e.g. Simulink)

Semantics of BLTL

The semantics of BLTL for a trace σ^k:

- $\sigma^k \models ap$ iff atomic proposition *ap* true in state s_k
- $\sigma^k \models \Phi_1 \vee \Phi_2$ iff $\sigma^k \models \Phi_1$ or $\sigma^k \models \Phi_2$
- $\sigma^k \models \neg\Phi$ iff $\sigma^k \models \Phi$ does not hold
- $\sigma^k \models \Phi_1 \,\mathcal{U}^t\, \Phi_2$ iff there exists natural i such that
  1) $\sigma^{k+i} \models \Phi_2$
  2) $\Sigma_{j<i}\, t_{k+j} \leq t$
  3) for each $0 \leq j < i$, $\sigma^{k+j} \models \Phi_1$

  "within time t, Φ_2 will be true and Φ_1 will hold until then"

- In particular, $F^t \Phi = true \,\mathcal{U}^t\, \Phi$, $G^t \Phi = \neg F^t \neg\Phi$

Semantics of BLTL (cont'd)

- Simulation traces are finite: is $\sigma \models \Phi$ well defined?

- *Definition:* The time bound of Φ:
  - $\#(ap) = 0$
  - $\#(\neg\Phi) = \#(\Phi)$
  - $\#(\Phi_1 \vee \Phi_2) = \max(\#(\Phi_1), \#(\Phi_2))$
  - $\#(\Phi_1 \mathcal{U}^t \Phi_2) = t + \max(\#(\Phi_1), \#(\Phi_2))$

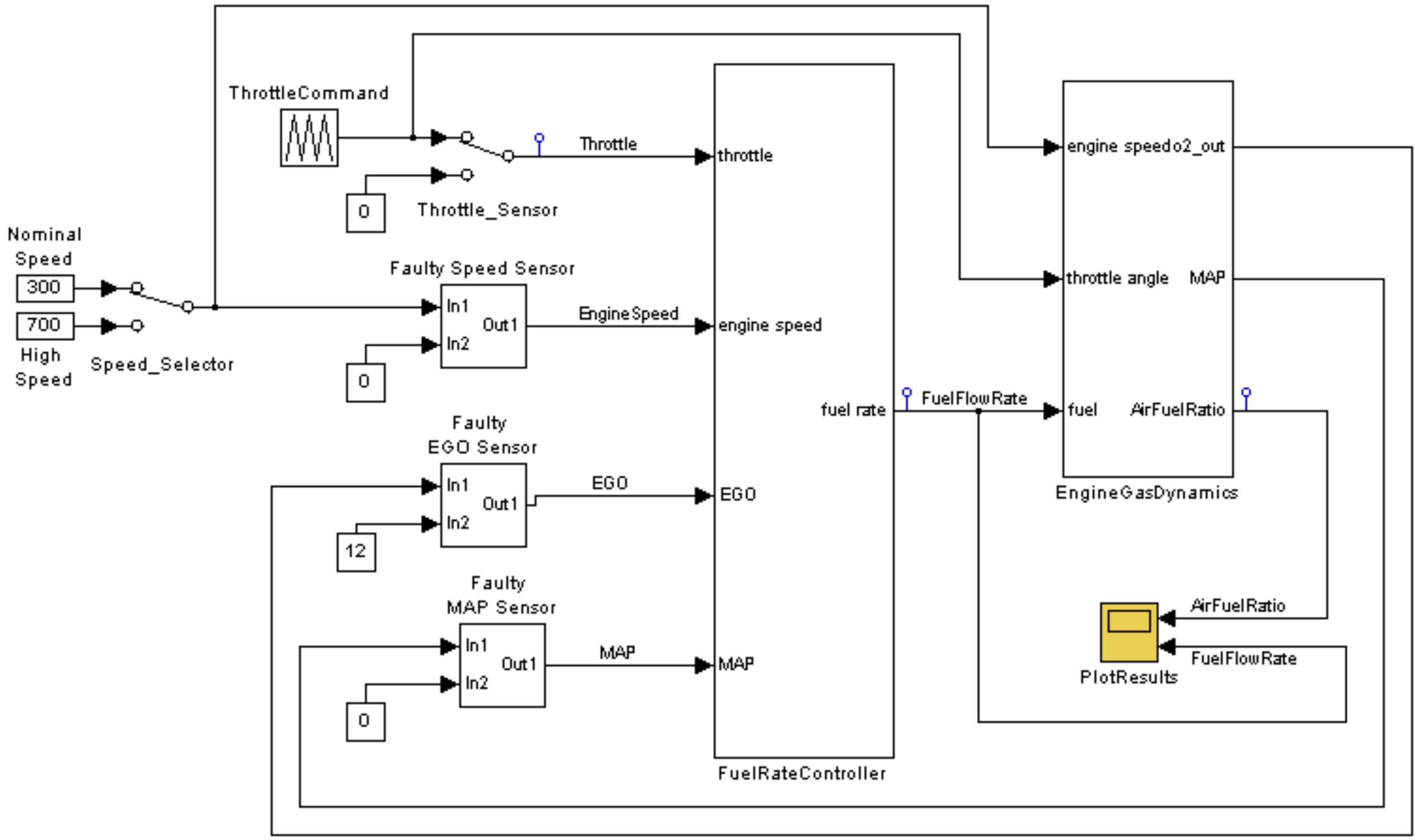
- *Lemma:* "Bounded simulations suffice"

  Let Φ be a BLTL *property, and $k \geq 0$. For any two infinite traces ρ, σ such that ρ^k* and *σ^k "equal up to time $\#(\Phi)$" we have*

$$\rho^k \models \Phi \quad \textit{iff} \quad \sigma^k \models \Phi$$

Fuel Control System - I

The Simulink model:



Fuel Control System - II

- Ratio between air mass flow rate and fuel mass flow rate
  - Stoichiometric ratio is 14.6
- Senses amount of oxygen in exhaust gas, pressure, engine speed and throttle to compute *correct* fuel rate.
  - Single sensor faults are compensated by switching to a higher oxygen content mixture
  - Multiple sensor faults force engine shutdown
- Probabilistic behavior because of random faults
  - In the EGO (oxygen), pressure and speed sensors
  - Faults modeled by three independent Poisson processes
  - We did not change the speed or throttle inputs

Fuel Control System - III

- We Model Check the formula (Null hypothesis)

  $\mathcal{M}, \mathit{FaultRate} \models P_{\geq\theta}(\neg \mathbf{F}^{100}\, \mathbf{G}^{1}(\mathit{FuelFlowRate} = 0))$

  for $\theta = .5, .7, .8, .9, .99$
- *"It is not the case that within 100 seconds, FuelFlowRate is zero for 1 second"*
- We use various values of *FaultRate for each of the* three sensors in the model
- We choose Bayes threshold $T = 1000$, i.e., stop when probability of error is < .001
- Uniform, equally likely priors

Fuel Control System:
Hypothesis testing

Recall the Null hypothesis:

$$\mathcal{M}, \mathit{FaultRate} \models P_{\geq\theta}(\neg \boldsymbol{F}^{100}\ \boldsymbol{G}^{1}(\mathit{FuelFlowRate} = 0))$$

Priors: *uniform, equally likely*.

Number of samples and test decision:

- red / blue number: reject / accept null hypothesis

		Probability threshold θ				
		.5	.7	.8	.9	.99
Fault rates	[3 7 8]	58	17	10	8	2
	[10 8 9]	32	95	394	710	8
	[20 10 20]	9	16	24	44	$\boxed{1{,}626}$
	[30 30 30]	9	16	24	44	239

Longest run: 1h 5’ on a 2.4GHz Pentium 4 computer

Fuel Control System results:
Interval estimation

- Bayesian estimation algorithm, uniform prior.
- Want to estimate the probability that

$$\mathcal{M}, \mathit{FaultRate} \models (\neg \boldsymbol{F}^{100}\, \boldsymbol{G}^{1}(\mathit{FuelFlowRate} = 0))$$

- For half-width δ=.01 and several values of coverage c
- Posterior mean: add/subtract δ to get the Bayesian interval

		Interval coverage c			
		.9	.95	.99	.999
Fault rates	[3 7 8]	.3603	.3559	.3558	.3563
	[10 8 9]	.8534	.8518	.8528	.8534
	[20 10 20]	.9764	.9784	.9840	.9779
	[30 30 30]	.9913	.9933	.9956	.9971

Fuel Control System results:
Interval estimation

- Number of samples
- Comparison with Chernoff-Hoeffding bound (Bernoulli r.v.'s)

$$\Pr(|X - p| \geq \delta) \leq exp(-2n\delta^2)$$

where $X = 1/n\ \Sigma_i X_i$, $E[X_i]=p$

		Interval coverage *c*			
		.9	.95	.99	.999
Fault rates	[3 7 8]	6,234	8,802	15,205	24,830
	[10 8 9]	3,381	4,844	8,331	13,569
	[20 10 20]	592	786	1,121	2,583
	[30 30 30]	113	148	227	341
Chernoff bound		119,829	147,555	211,933	304,036

Conclusions

- Use sequential sampling
- Bayesian Interval Estimation / Hypothesis Testing
- Statistical Model Checking is
    - Not the silver bullet
    - Another (useful) verification tool

The End

Thank you!

Bayes Estimators - I

- Quadratic loss function:

  u (unknown) parameter, *d(x)* estimator for *u*

$$L(u, d(x)) = |u - d(x)|^2$$

- Risk of estimator *d*: average loss over all possible data

$$R(u, d) = E_u[L(u, d)] = \int_X L(u, d(x)) f(x|u) \; dx$$

Bayes Estimators - II

- Integrated risk of estimator *d* with respect to prior *g*

$$r(g,d) = E[R(u,d)] = \int_U \int_X L(u,d(x)) f(x|u) dx \;\; g(u) \; du$$

- *U* is the parameter space ([0,1] for us).
- Using the posterior mean as estimator minimizes *r(g,d)*
- In our case the posterior mean is

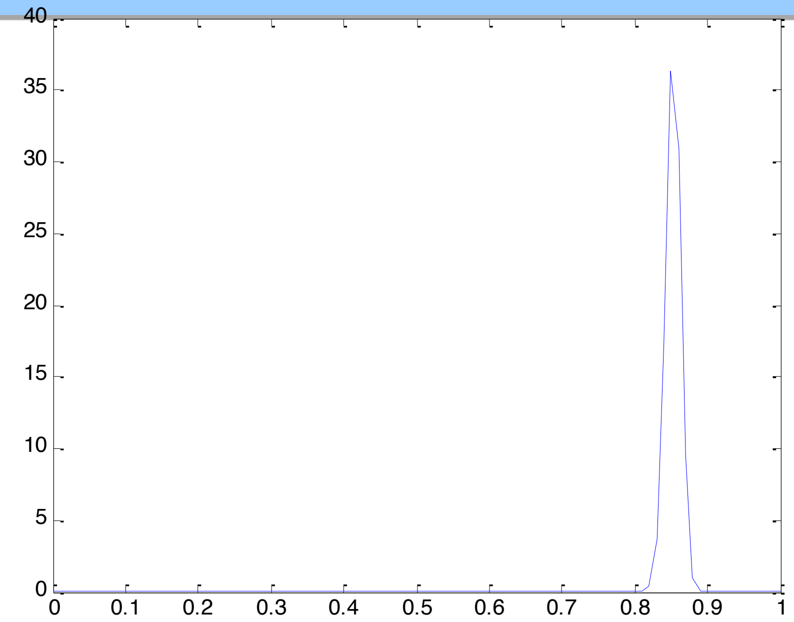
  $(x+\alpha)/(n+\alpha+\beta)$

  where $x \leq n$ number of successes, α,β Beta prior parameters.

Fuel Control System:
Hypothesis testing

Informative priors:
convex combinations of Betas

Example: for fault rates [10 8 9] we used
0.01 x beta(1,1) + 0.99 x beta(1000,172.6)



		Probability threshold θ				
		.5	.7	.8	.9	.99
Fault rates	[3 7 8]	55 (3)	12 (5)	10	8	2
	[10 8 9]	28 (4)	64 (31)	347 (47)	255 (455)	8
	[20 10 20]	8 (1)	13 (3)	20 (4)	39 (5)	1,463 (163)
	[30 30 30]	7 (2)	13 (3)	18 (6)	33 (11)	201 (38)

Computing the Bayes Factor - I

- The Bayes Factor uses posterior (and prior) probability

$$\boxed{\frac{P(X|H_0)}{P(X|H_1)} = \frac{P(H_0|X)}{P(H_1|X)} \cdot \frac{P(H_1)}{P(H_0)}}$$

- Posterior density (Bayes Theorem) (iid Bernoulli samples)

$$\boxed{f(u \mid x_1, \ldots, x_n) = \frac{\boxed{f(x_1 \mid u) \cdots f(x_n \mid u)} \cdot g(u)}{\int_0^1 f(x_1 \mid v) \cdots f(x_n \mid v) \cdot g(v)\; dv}}$$

Likelihood function

Why Beta priors?

- Defined over [0,1]
- Beta distributions are *conjugate* to Binomial distributions:
  - If prior g is Beta and likelihood function is Binomial then *posterior is Beta*
- Suppose likelihood Binomial(n,x), prior Beta(α,β): posterior

$$f(u \mid x_1, \ldots, x_n) \approx f(x_1|u) \cdot \cdot \cdot f(x_n|u) \cdot g(\mathrm{u})$$

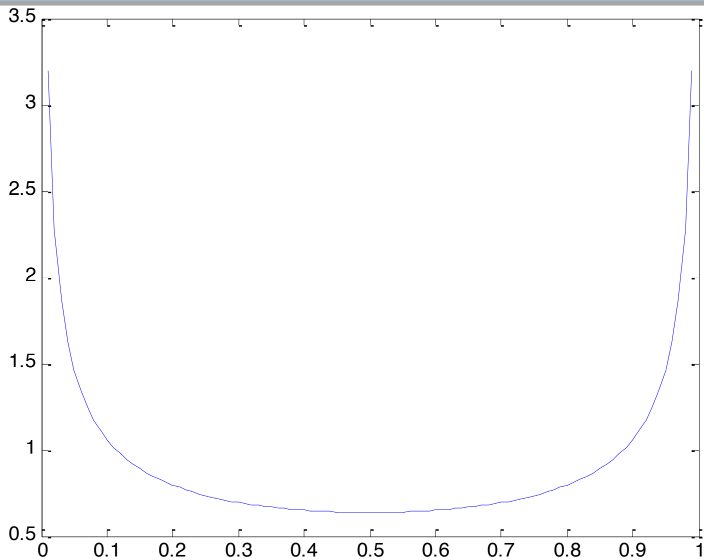
$$= u^x(1 - u)^{\mathrm{n-}x} \cdot u^{\alpha\text{-}1}(1 - u)^{\beta\text{-}1}$$

$$= u^{x+\alpha\,-1}(1 - u)^{\mathrm{n-x+}\beta\text{-}1}$$

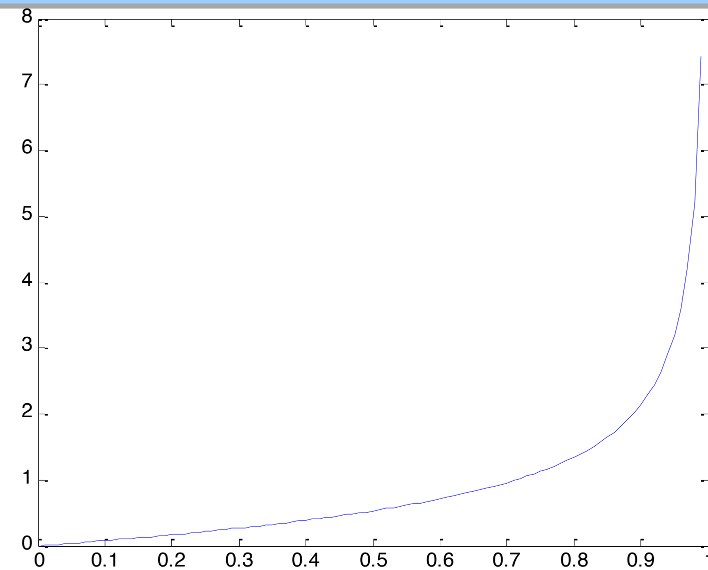
where $x = \Sigma_i x_i$

- Posterior is Beta of parameters $x+\alpha$ and n-x+β

Beta Density Shapes

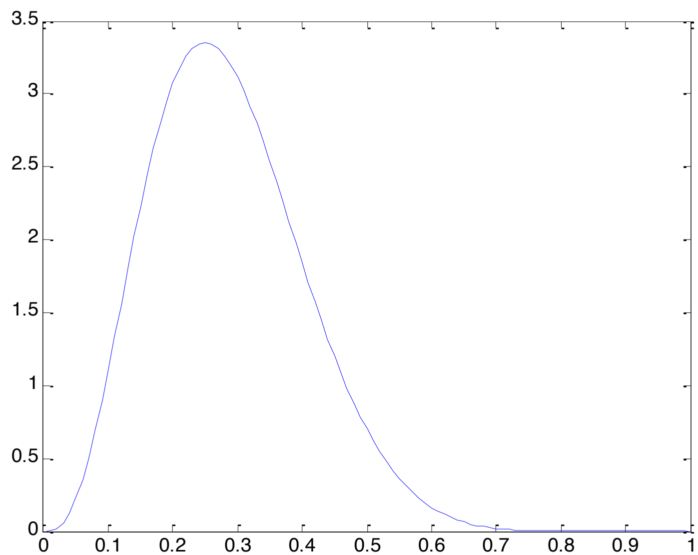


α=.5 β=.5



α=2 β=.5

Unimodal, but can form convex combinations …



α=4 β=10

Performance of Bayesian Estimation