UNIVERSITY OF ALBERTA

Markov Chain Monte Carlo (MCMC) Methods

PHYS 574

April 5th, 2019

Prepared by: Mario Ivanov

Metropolis-Hastings Algorithm

Most common/Widely used

Metropolis-Hastings Algorithm

Most common/Widely used

Let’s explore it with an example!

Metropolis-Hastings Algorithm

Most common/Widely used

Let’s explore it with an example!

Python module: “emcee”

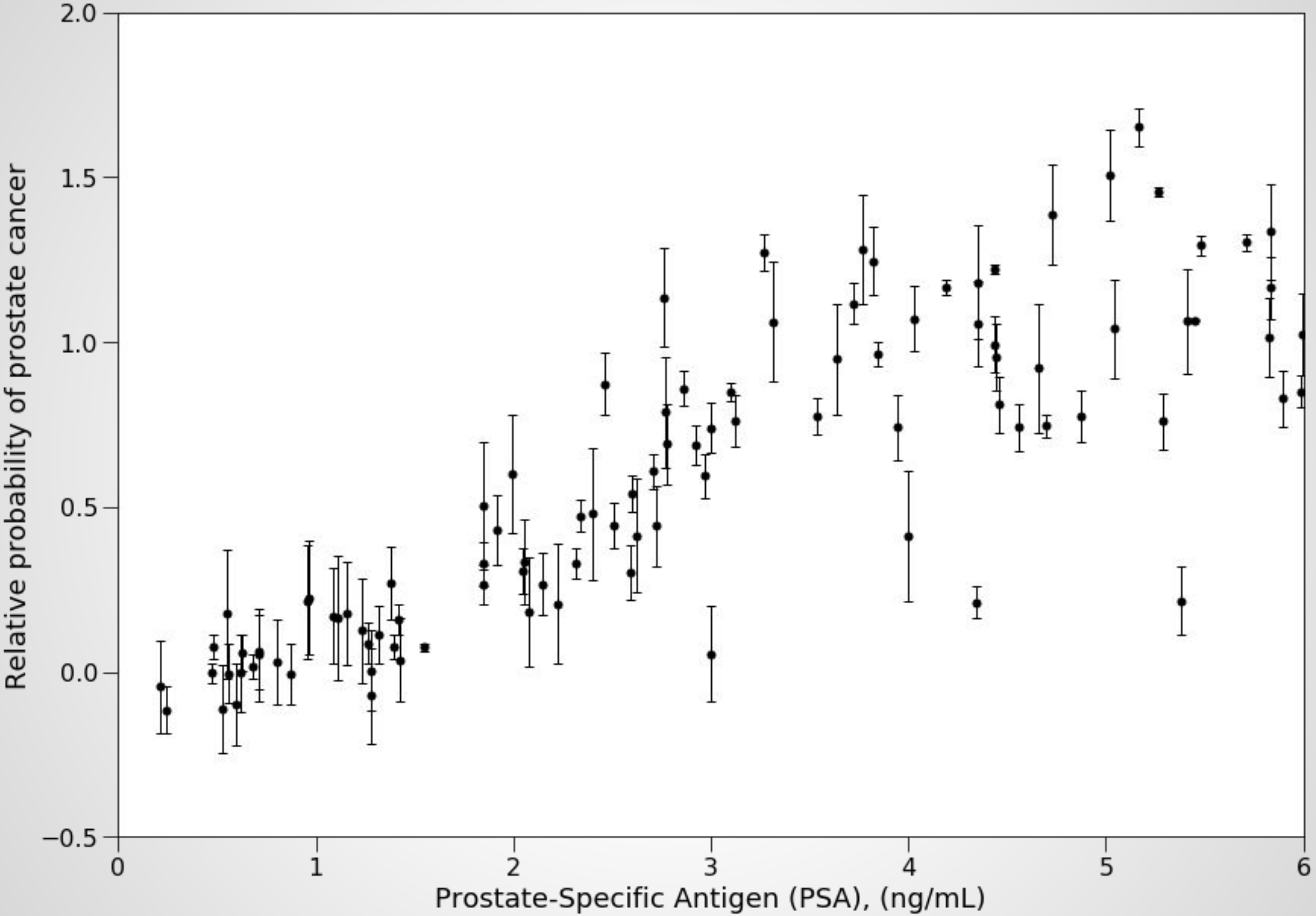
Prostate Cancer Test

- Blood test: Prostate-Specific Antigen (PSA)
  - PSA concentration in (ng/mL)

Prostate Cancer Test

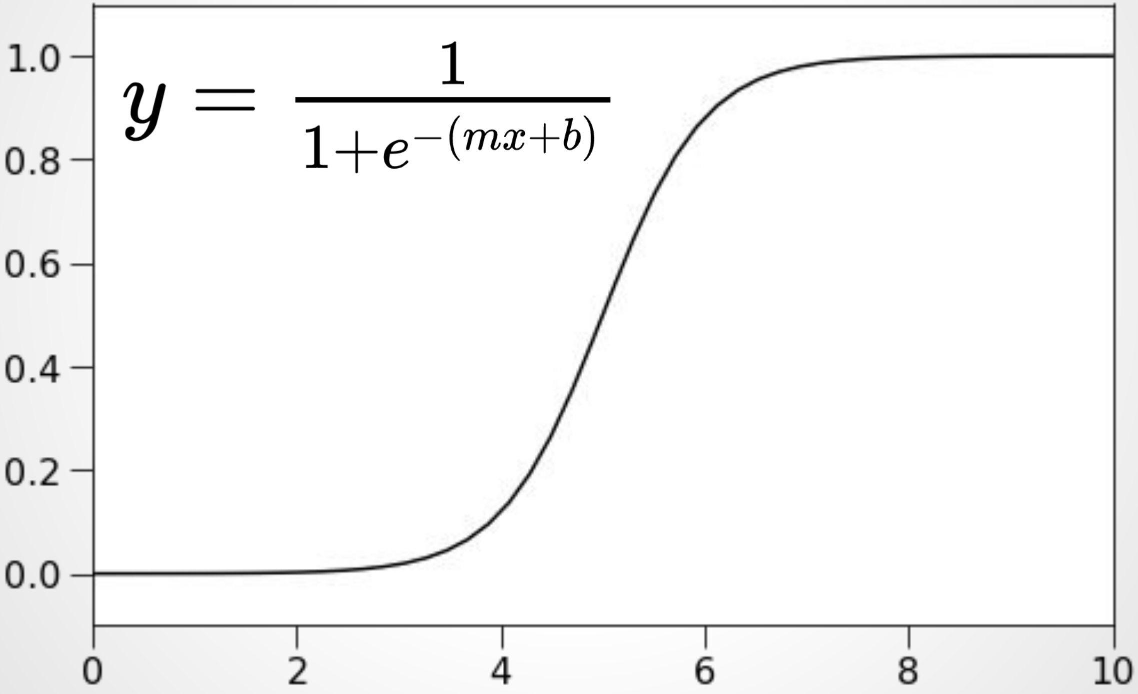
- Blood test: Prostate-Specific Antigen (PSA)
  - PSA concentration in (ng/mL)

- Problems:
  - Ethnicity/Age/Gender
  - PSA velocity
  - False positives and biopsy complications



Relative probability of prostate cancer
2.0
1.5
1.0
0.5
0.0
−0.5
0
1
2
3
4
5
6
Prostate-Specific Antigen (PSA), (ng/mL)

Sigmoid Model



Bayes' Theorem

$$p(B \mid A) = \frac{p(A \mid B)\, p(B)}{p(A)}$$

Bayes' Theorem

$$p(m, b, f \mid x, y, \sigma) \propto p(m, b, f)\, p(y \mid x, \sigma, m, b, f)$$

Bayes' Theorem

$$p(m, b, f \mid x, y, \sigma) \propto p(m, b, f)\, p(y \mid x, \sigma, m, b, f)$$

POSTERIOR

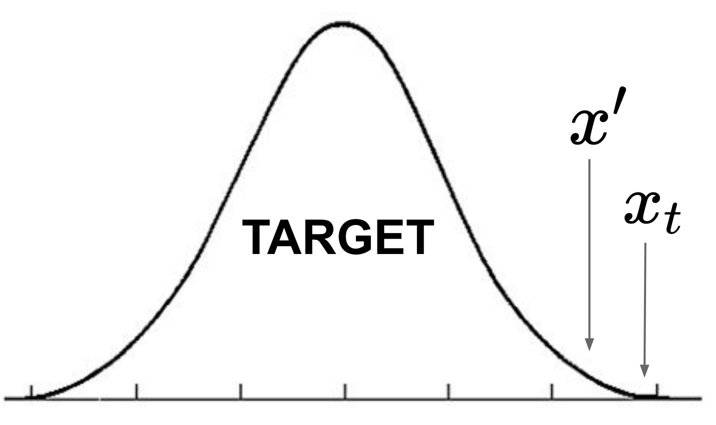
LIKELIHOOD

PRIOR

Metropolis-Hastings Algorithm

Draw a sample state from proposed distribution

$$x' \quad \text{from} \quad p(m, b, f)$$



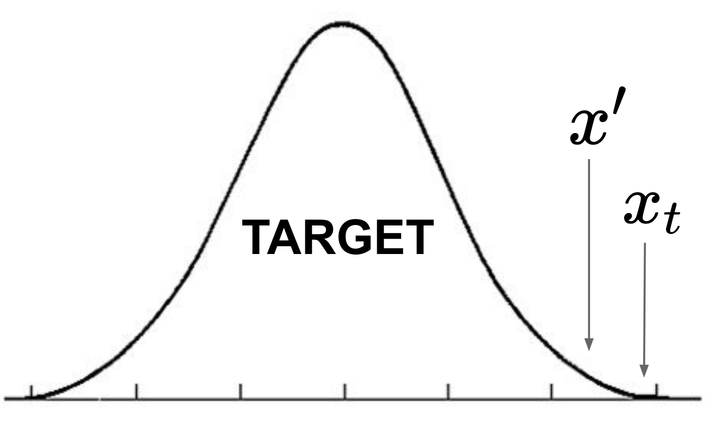
where $x_t = \{m^0, b^0, f^0\}$

where $x' = \{m', b', f'\}$

Metropolis-Hastings Algorithm

Calculate the acceptance probability ratio

$$\alpha = \frac{p(x')}{p(x_t)}$$

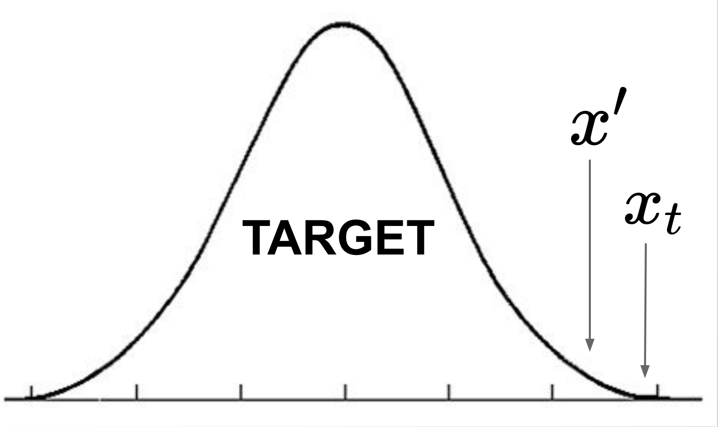


Metropolis-Hastings Algorithm

Accept or Reject:

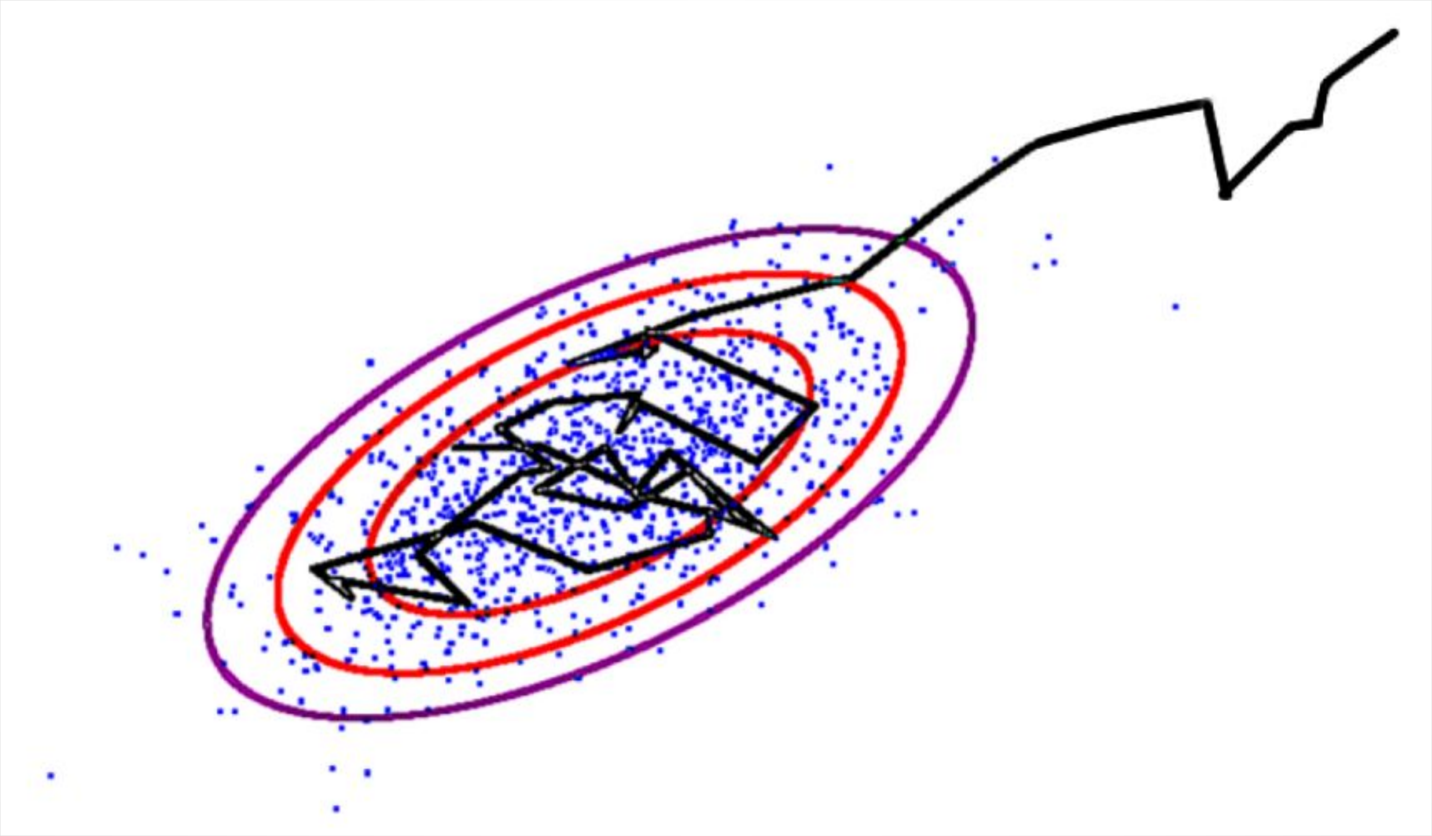
$$u \in [0, 1]$$

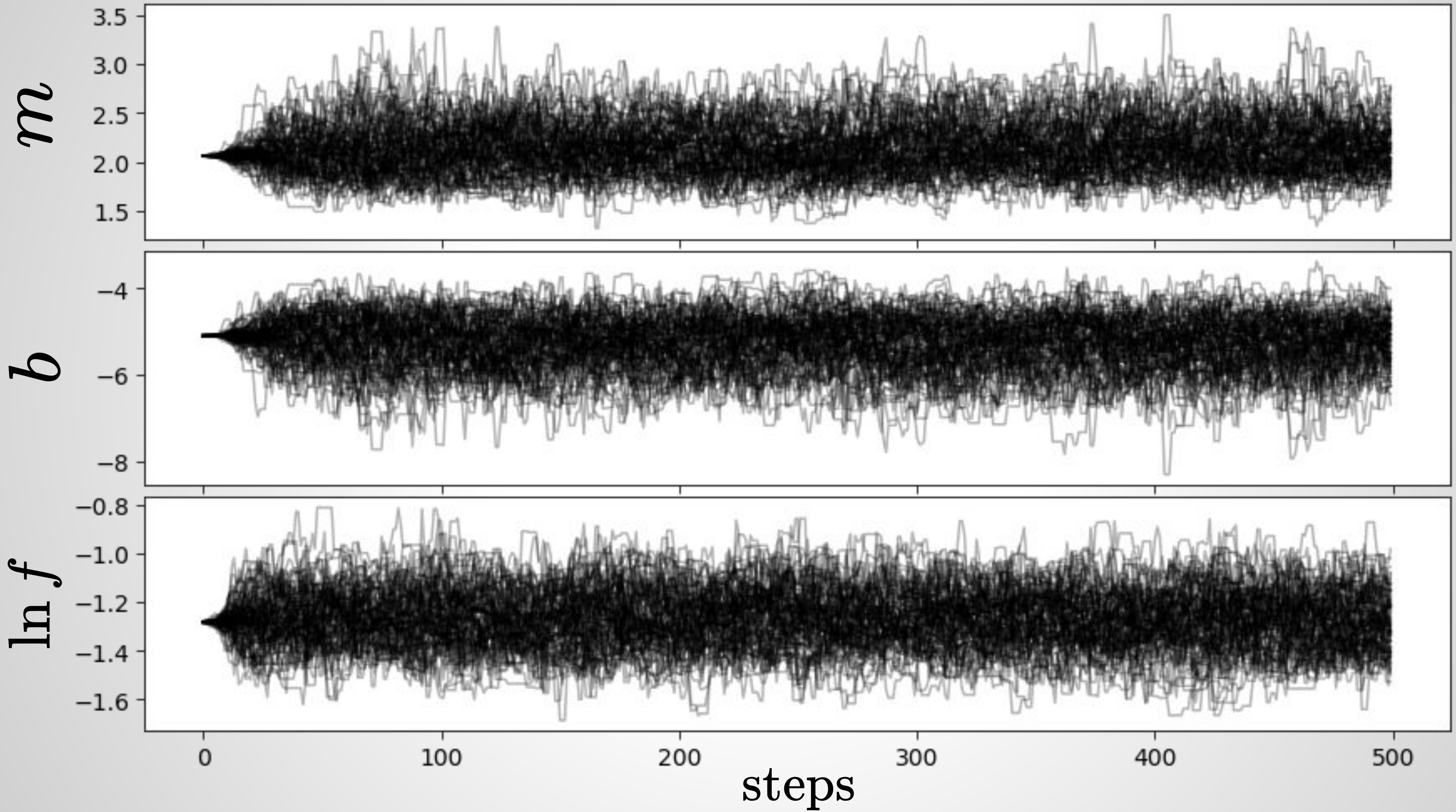
$$\alpha = \frac{p(x')}{p(x_t)}$$



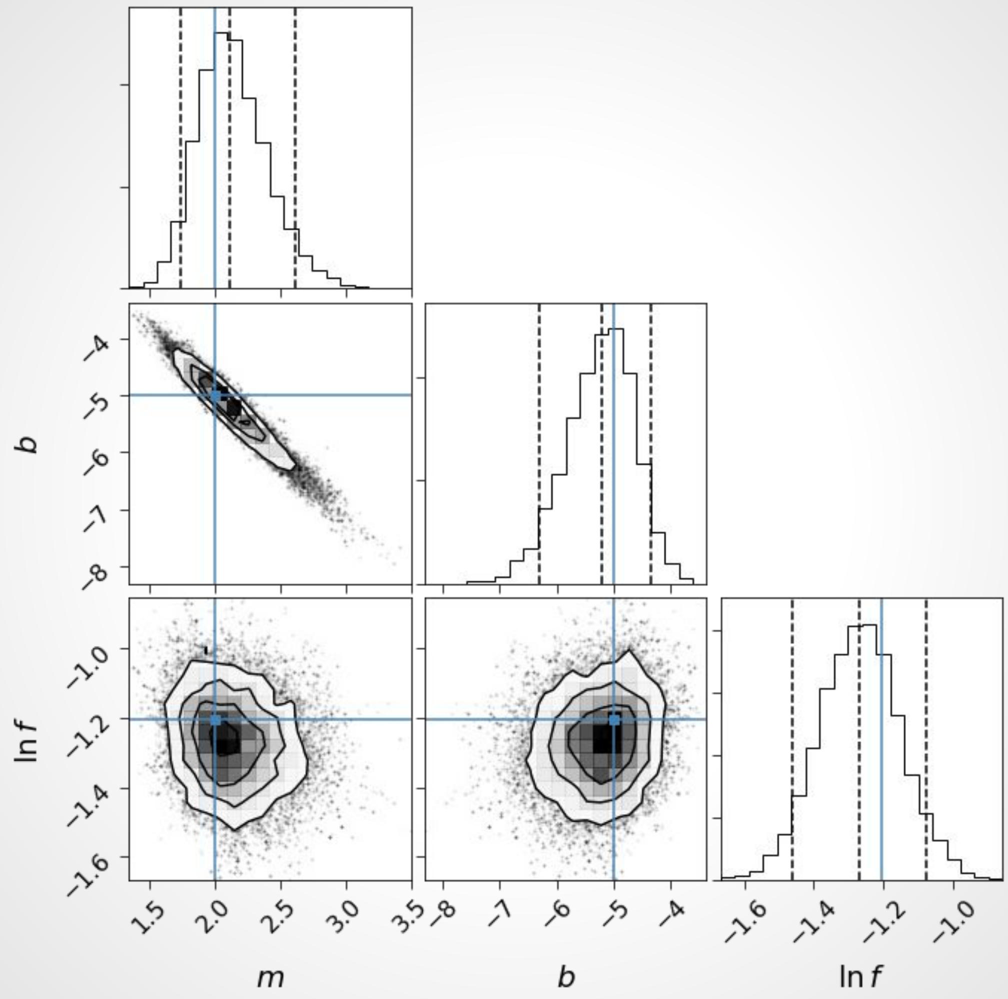
$$u \leq \alpha, \text{ accept}: \; x_{t+1} = x'$$

$$u > \alpha, \text{ reject}: \; x_{t+1} = x_t$$

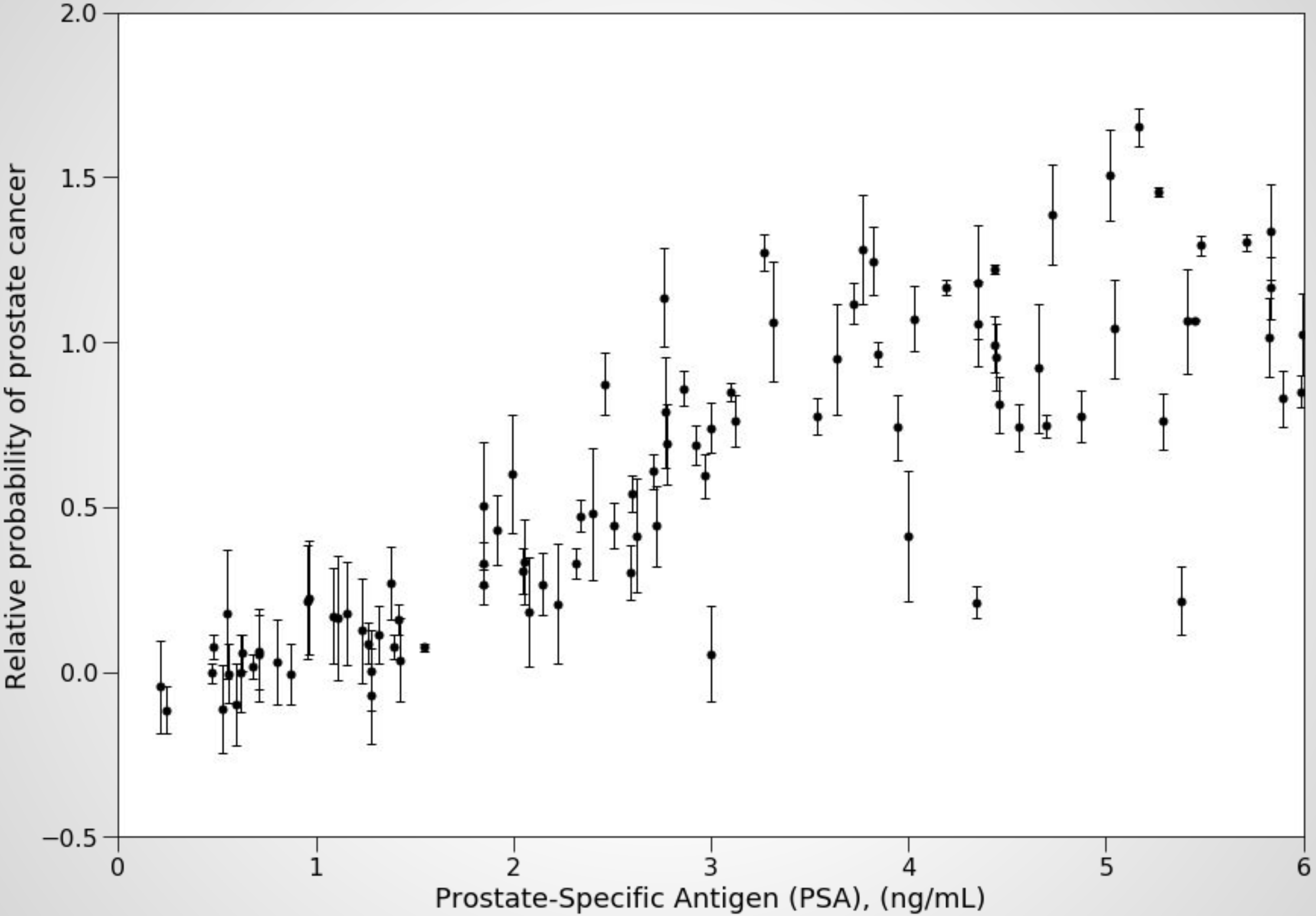


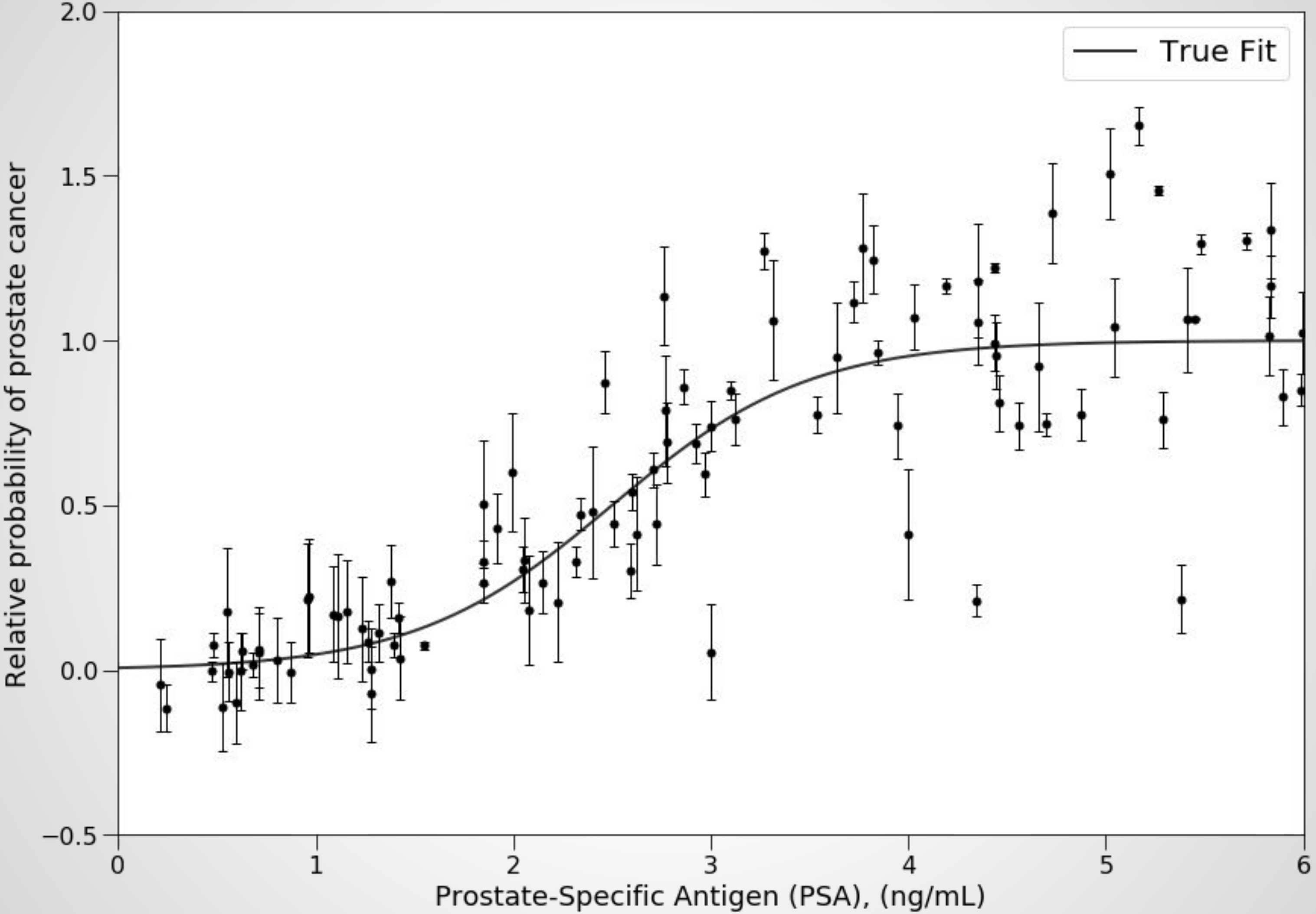


m
3.5
3.0
2.5
2.0
1.5
b
−4
−6
−8
$\ln f$
−0.8
−1.0
−1.2
−1.4
−1.6
0
100
200
300
400
500
steps

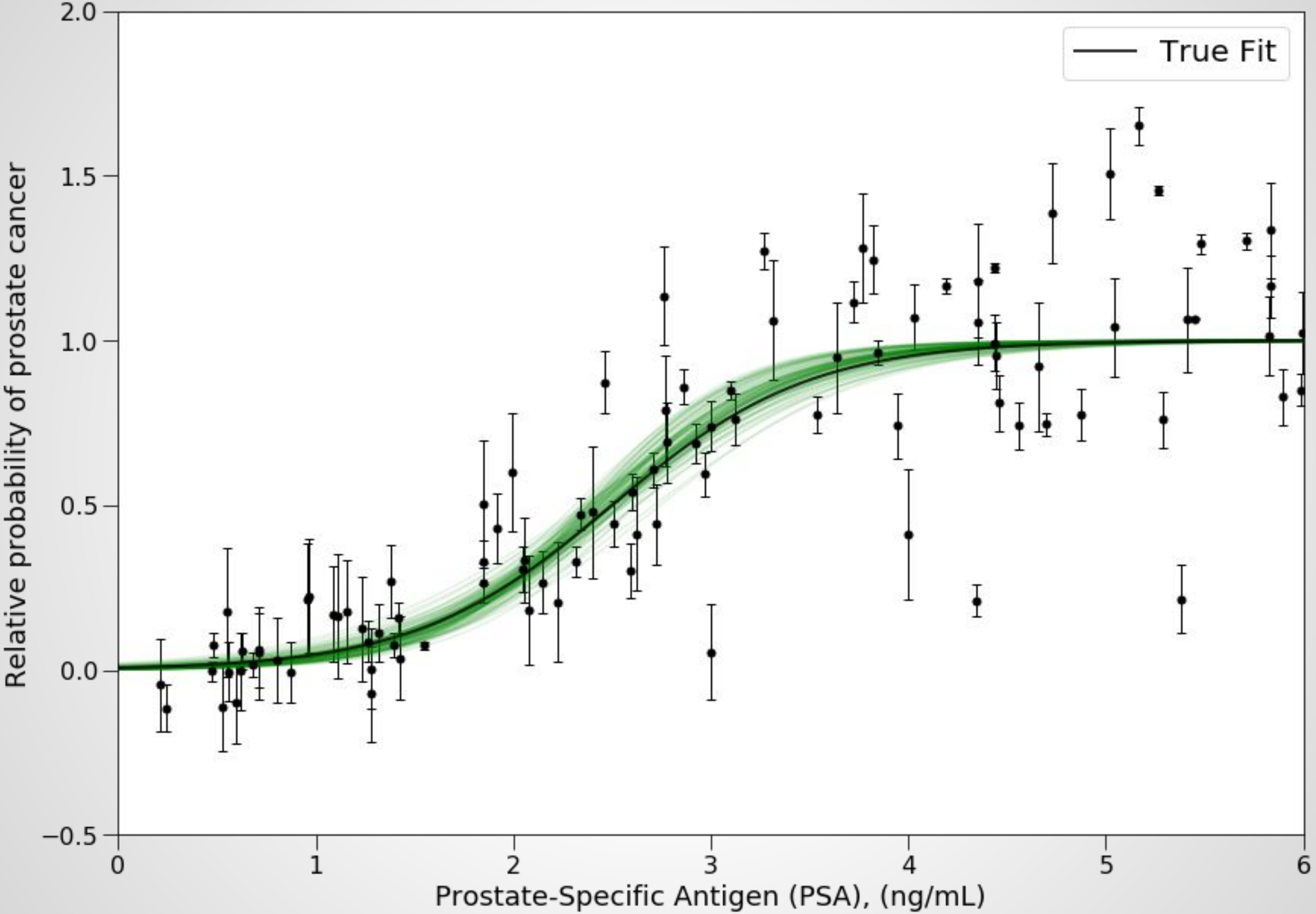


b
ln f
m
b
ln f

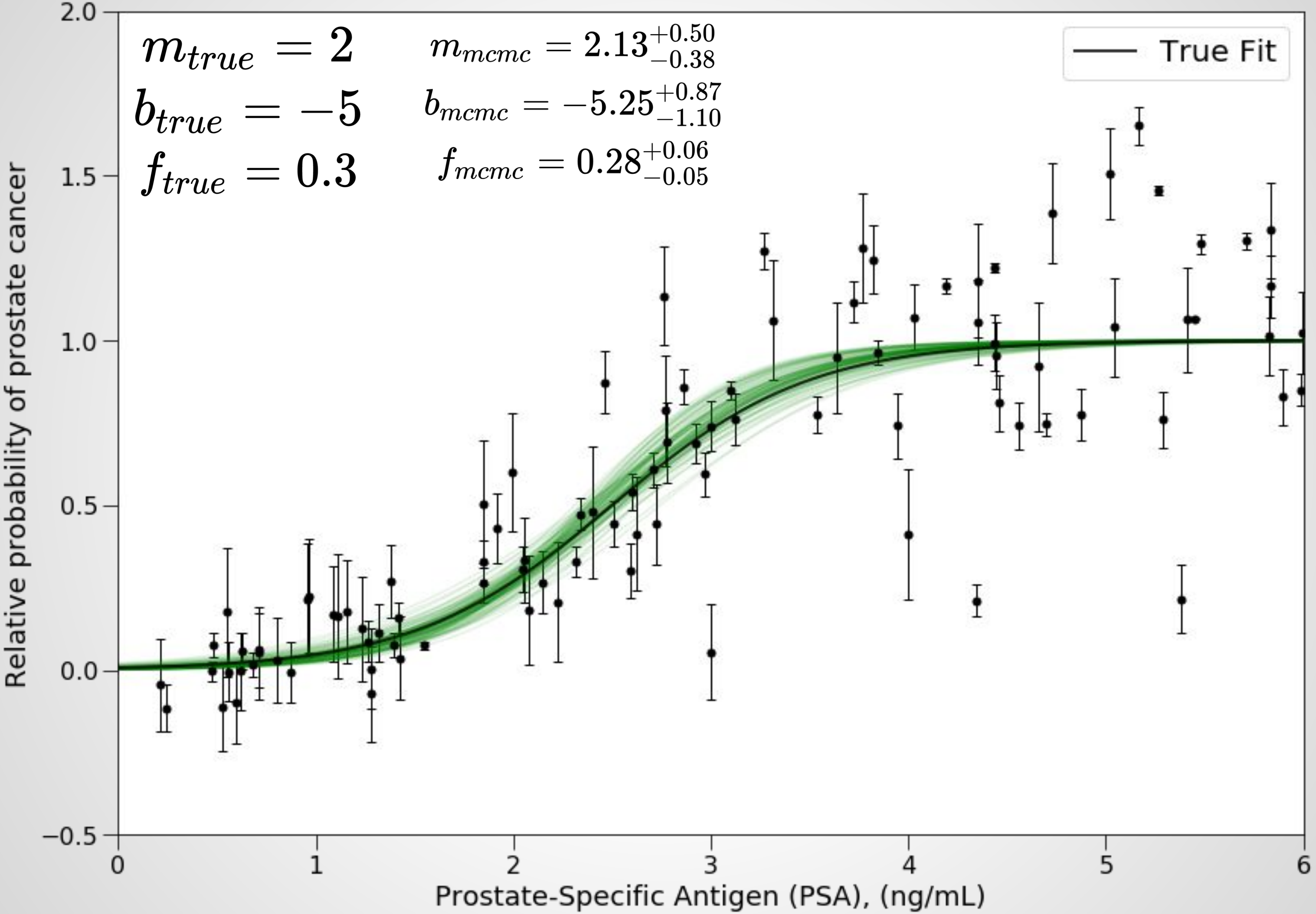

Relative probability of prostate cancer
2.0
1.5
1.0
0.5
0.0
−0.5
0
1
2
3
4
5
6
Prostate-Specific Antigen (PSA), (ng/mL)



True Fit
Relative probability of prostate cancer
Prostate-Specific Antigen (PSA), (ng/mL)
2.0
1.5
1.0
0.5
0.0
−0.5
0
1
2
3
4
5
6



True Fit
Relative probability of prostate cancer
Prostate-Specific Antigen (PSA), (ng/mL)
2.0
1.5
1.0
0.5
0.0
−0.5
0
1
2
3
4
5
6


$m_{true} = 2$
$b_{true} = -5$
$f_{true} = 0.3$
$m_{mcmc} = 2.13^{+0.50}_{-0.38}$
$b_{mcmc} = -5.25^{+0.87}_{-1.10}$
$f_{mcmc} = 0.28^{+0.06}_{-0.05}$
True Fit
Relative probability of prostate cancer
Prostate-Specific Antigen (PSA), (ng/mL)
2.0
1.5
1.0
0.5
0.0
−0.5
0
1
2
3
4
5
6

Gibbs Algorithm

Draw a sample state from proposed distribution

$$\{m^0, b^0, f^0\} \text{ from } p(m, b, f)$$

Gibbs Algorithm

Draw a sample state from proposed distribution

$$\{m^0, b^0, f^0\} \text{ from } p(m, b, f)$$

$$\text{for } t \text{ in } [1:T]:$$

$$m^t \;\sim\; p(m \mid b^{t-1}, f^{t-1})$$

$$b^t \;\sim\; p(b \mid m^t, f^{t-1})$$

$$f^t \;\sim\; p(f \mid m^t, b^t)$$

Gibbs Algorithm

Simplest of MCMC algorithms

Gibbs Algorithm

Simplest of MCMC algorithms

Use if conditional probabilities can be sampled

Gibbs Algorithm

Simplest of MCMC algorithms

Use if conditional probabilities can be sampled

Slow for correlated parameters

Metropolis-adjusted Langevin Rule

Similar to MH algorithm, with a target probability distribution function term

Metropolis-adjusted Langevin Rule

Similar to MH algorithm, with a target probability distribution function term

$$\alpha = \underbrace{\frac{p(x')}{p(x_t)}}_{\text{MH}}$$

Metropolis-adjusted Langevin Rule

Similar to MH algorithm, with a target probability distribution function term

$$\alpha = \frac{p(x')\, p(y \mid x')}{p(x_t)\, p(y \mid x_t)}$$

Hamiltonian Monte Carlo

MH’s and Gibbs’ (and a little bit of Langevin’s) weakness is their reliance on random steps

Hamiltonian Monte Carlo

MH’s and Gibbs’ (and a little bit of Langevin’s) weakness is their reliance on random steps

Add a momentum term to each variable

Hamiltonian Monte Carlo

MH’s and Gibbs’ (and a little bit of Langevin’s) weakness is their reliance on random steps

Add a momentum term to each variable

Gives the expected distance and direction of jump of variable based on the last few jumps

Hamiltonian Monte Carlo

Draw a sample state from proposed distribution

$$\{m^0, b^0, f^0\} \text{ from } p(m, b, f)\, p(\mathbf{v})$$

Hamiltonian Monte Carlo

Draw a sample state from proposed distribution

$$\{m^0, b^0, f^0\} \text{ from } p(m, b, f)\, p(\mathbf{v})$$

$$H(m, b, f, \mathbf{v}) = E(m, b, f) + K(\mathbf{v}) = \text{constant}$$

Hamiltonian Monte Carlo

Draw a sample state from proposed distribution

$$\{m^0, b^0, f^0\} \text{ from } p(m, b, f)\, p(\mathbf{v})$$

$$H(m, b, f, \mathbf{v}) = E(m, b, f) + K(\mathbf{v}) = \text{constant}$$

...followed by same accept/reject steps as MH

Slice Sampling

Sample the region under proposed distribution

Slice Sampling

Sample the region under proposed distribution

1. Sample a random initial value x_0

Slice Sampling

Sample the region under proposed distribution

1. Sample a random initial value x_0
2. At each iteration i
   a. Sample uniformly (vertical slice) for an auxiliary variable in the region $[0, p(x_{i-1} \mid y)]$
   b. Sample uniformly (horizontal slice) for x_i

Summary

Metropolis-Hastings:	drunk man walking
Gibbs:	drunk man hopscotch
Langevin:	drunk man with a map
Hamiltonian:	drunk man running
Slice Sampling:	drunk man in a room

Summary

While easy-to-use, they are easy to mess up

There are a lot of resources online, especially for the most common types of algorithms https://chi-feng.github.io/mcmc-demo/app.html

Prostate Cancer Research:
https://www.cancer.gov/types/prostate/psa-fact-sheet
https://www.ncbi.nlm.nih.gov/pubmed/11333995
https://www.ncbi.nlm.nih.gov/pubmed/15163773
https://www.ncbi.nlm.nih.gov/pubmed/29772218
https://www.ices.on.ca/Publications/Atlases-and-Reports/2002/Prostate-specific-antigen-PSA-screening
http://www.topalbertadoctors.org/download/276/Prostate%2BCancer%2BGuideline%2BEval%2B%2526%2BReferr.pdf

MH and Slice Sampling :
https://www.sheffield.ac.uk/polopoly_fs/1.60510!/file/MCMC.pdf

Gibbs Sampling:
http://www.mit.edu/~ilkery/papers/GibbsSampling.pdf

Metropolis-adjusted Langevin:
https://warwick.ac.uk/fac/sci/statistics/crism/visitors/vats/misc/malavsrwm.pdf

Hamiltonian Monte Carlo Reference:
https://www.cs.utoronto.ca/~radford/ftp/ham-mcmc.pdf

EXTRA SLIDES

Bayes’ Theorem (once again)

<u>LOG LIKELIHOOD FUNCTION</u>

$$\ln\, p(y\,|\,x,\sigma,m,b,f) = -\frac{1}{2}\sum_n \left[\frac{(y_n - \mathrm{logit}(x_n,m,b))^2}{s_n^2} + \ln(2\pi s_n^2)\right]$$

$$\mathrm{logit}(x,m,b) = \frac{1}{1+e^{-(mx+b)}}$$

$$s_n^2 = \sigma_n^2 + f^2\,(m x_n + b)^2$$

Doctor’s Visit

Doctor's Visit

Test Result	Prostate Cancer: Yes	Prostate Cancer: No
+	0.9	
-		

Doctor’s Visit

Prostate Cancer

Test Result	Yes	No
+	0.9	
-	0.1	

Doctor's Visit

Prostate Cancer

Test Result	Yes	No
+	0.9	0.2
-	0.1	

Doctor's Visit

Prostate Cancer

Test Result	Yes	No
+	0.9	0.2
-	0.1	0.8

Doctor's Visit

Test is positive!

Prostate Cancer

Test Result		Yes	No
	+	0.9	0.2
	-	0.1	0.8

Doctor's Visit

Test is positive!

$$p(\text{cancer} \mid +)?$$

Prostate Cancer

Test Result	Yes	No
+	0.9	0.2
-	0.1	0.8

Doctor’s Visit

$$p(\text{cancer} \mid +) = \frac{p(+ \mid \text{cancer})\, p(\text{cancer})}{p(+)}$$

Doctor’s Visit

$$p(\text{cancer} \mid +) = \frac{\overset{0.9}{\overbrace{p(+ \mid \text{cancer})}}\, p(\text{cancer})}{p(+)}$$

Doctor’s Visit

$$p(\text{cancer} \mid +) = \frac{\overset{0.9}{p(+ \mid \text{cancer})}\; \overset{0.00001}{p(\text{cancer})}}{p(+)}$$

Doctor's Visit

$$p(+) = p(+ \mid \text{cancer})\, p(\text{cancer}) + p(+ \mid \text{no cancer})\, p(\text{no cancer})$$

Doctor’s Visit

$$p(+) = \underbrace{p(+ \mid \text{cancer})}_{0.9}\, \overbrace{p(\text{cancer})}^{0.00001} + \underbrace{p(+ \mid \text{no cancer})}_{0.2}\, \overbrace{p(\text{no cancer})}^{0.99999}$$

Doctor's Visit

$$p(\text{cancer} \mid +) = 0.000045$$

Doctor's Visit

$$p(\text{cancer} \mid +) = 0.000045$$

Low chance of prostate cancer (0.00001)

High chance of false positive (0.2)

Doctor’s Visit

What if the prior was larger?

Doctor’s Visit

What if the prior was larger?

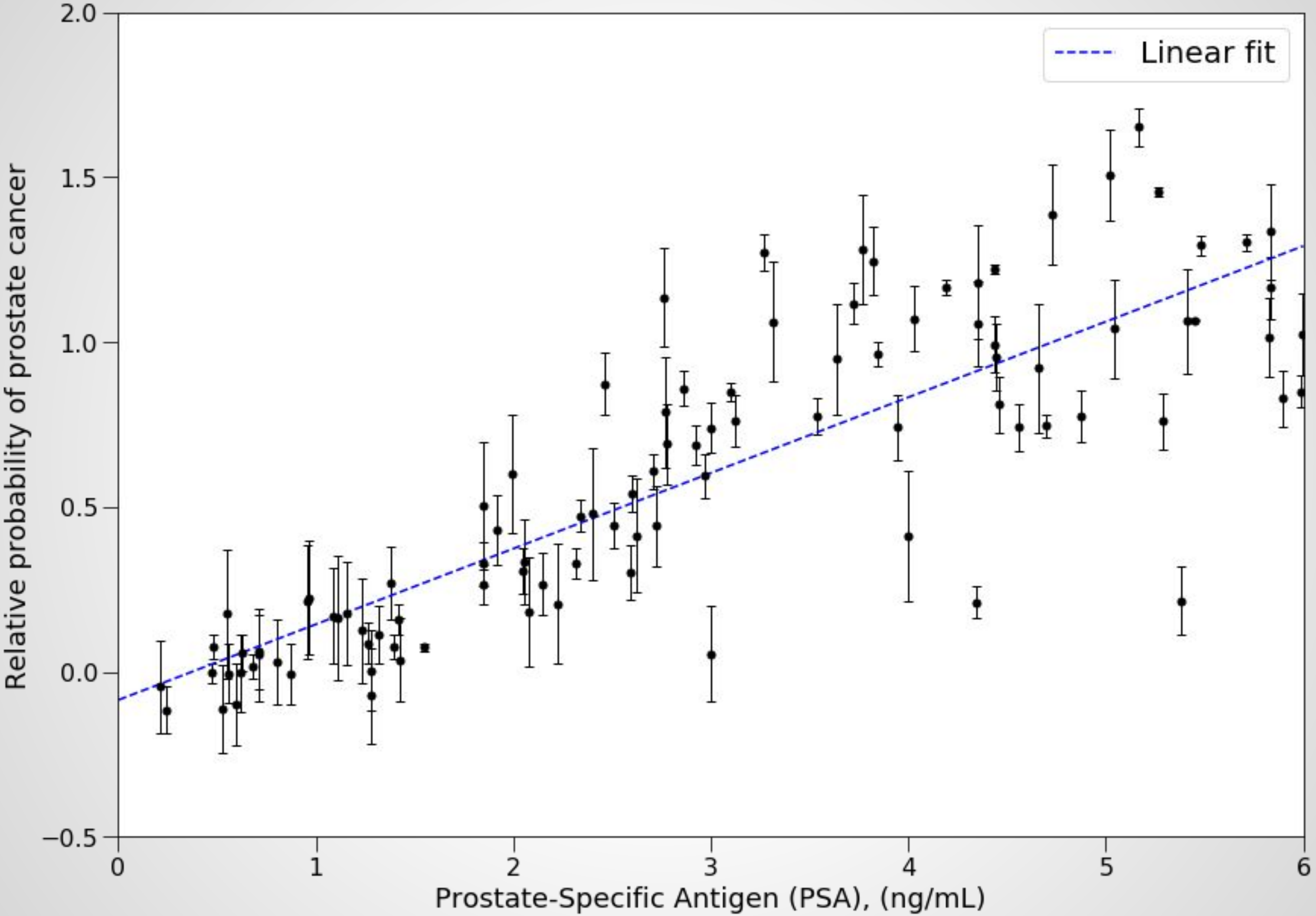
$$p(\text{cancer}) = 0.25$$

Doctor's Visit

What if the prior was larger?

$$p(\text{cancer}) = 0.25$$

$$p(\text{cancer} \mid +) \approx 0.58$$


Linear fit
Relative probability of prostate cancer
Prostate-Specific Antigen (PSA), (ng/mL)
2.0
1.5
1.0
0.5
0.0
−0.5
0
1
2
3
4
5
6