

# Think Bayes

*Allen B. Downey*

# Table of Contents

Preface	ix
1. Bayes's Theorem	1
Conditional probability	1
Conjoint probability	2
The cookie problem	3
Bayes's theorem	3
The diachronic interpretation	5
The M&M problem	6
The Monty Hall problem	7
Discussion	9
2. Computational Statistics	11
Distributions	11
The cookie problem	12
The Bayesian framework	13
The Monty Hall problem	14
Encapsulating the framework	15
The M&M problem	16
Discussion	17
Exercises	18
3. Estimation	19
The dice problem	19
The locomotive problem	20
What about that prior?	22
An alternative prior	23
Credible intervals	25
Cumulative distribution functions	26

The German tank problem	27
Discussion	27
Exercises	28
<u>More Estimation</u>	<u>29</u>
The Euro problem	29
Summarizing the posterior	31
Swamping the priors	31
Optimization	33
The beta distribution	35
Discussion	36
Exercises	37
<u>Odds and Addends</u>	<u>39</u>
Odds	39
The odds form of Bayes's theorem	40
Oliver's blood	41
Addends	42
Maxima	45
Mixtures	47
Discussion	49
<u>Decision Analysis</u>	<u>51</u>
The Price is Right problem	51
The prior	52
Probability density functions	53
Representing PDFs	53
Modeling the contestants	55
Likelihood	58
Update	58
Optimal bidding	60
Discussion	62
<u>Prediction</u>	<u>65</u>
The Boston Bruins problem	65
Poisson processes	66
The posteriors	67
The distribution of goals	68
The probability of winning	70
Sudden death	71
Discussion	73
Exercises	74

8. Observer Bias	77
The Red Line problem	77
The model	77
Wait times	79
Predicting wait times	82
Estimating the arrival rate	84
Incorporating uncertainty	87
Decision analysis	88
Discussion	90
Exercises	91
9. Two Dimensions	93
Paintball	93
The suite	93
Trigonometry	95
Likelihood	96
Joint distributions	97
Conditional distributions	98
Credible intervals	99
Discussion	101
Exercises	103
10. Approximate Bayesian Computation	105
The Variability Hypothesis	105
Mean and standard deviation	106
Update	108
The posterior distribution of CV	108
Underflow	109
Log-likelihood	111
A little optimization	111
ABC	113
Robust estimation	114
Who is more variable?	116
Discussion	118
Exercises	119
11. Hypothesis Testing	121
Back to the Euro problem	121
Making a fair comparison	122
The triangle prior	123
Discussion	124
Exercises	125

12. Evidence	127
Interpreting SAT scores	127
The scale	128
The prior	128
Posterior	130
A better model	132
Calibration	134
Posterior distribution of efficacy	136
Predictive distribution	137
Discussion	138
13. Simulation	141
The Kidney Tumor problem	141
A simple model	143
A more general model	144
Implementation	146
Caching the joint distribution	147
Conditional distributions	148
Serial Correlation	150
Discussion	153
14. A Hierarchical Model	155
The Geiger counter problem	155
Start simple	156
Make it hierarchical	157
A little optimization	158
Extracting the posteriors	159
Discussion	159
Exercises	160
15. Dealing with Dimensions	163
Belly button bacteria	163
Lions and tigers and bears	164
The hierarchical version	166
Random sampling	168
Optimization	169
Collapsing the hierarchy	170
One more problem	173
We're not done yet	174
The belly button data	175
Predictive distributions	179
Joint posterior	182