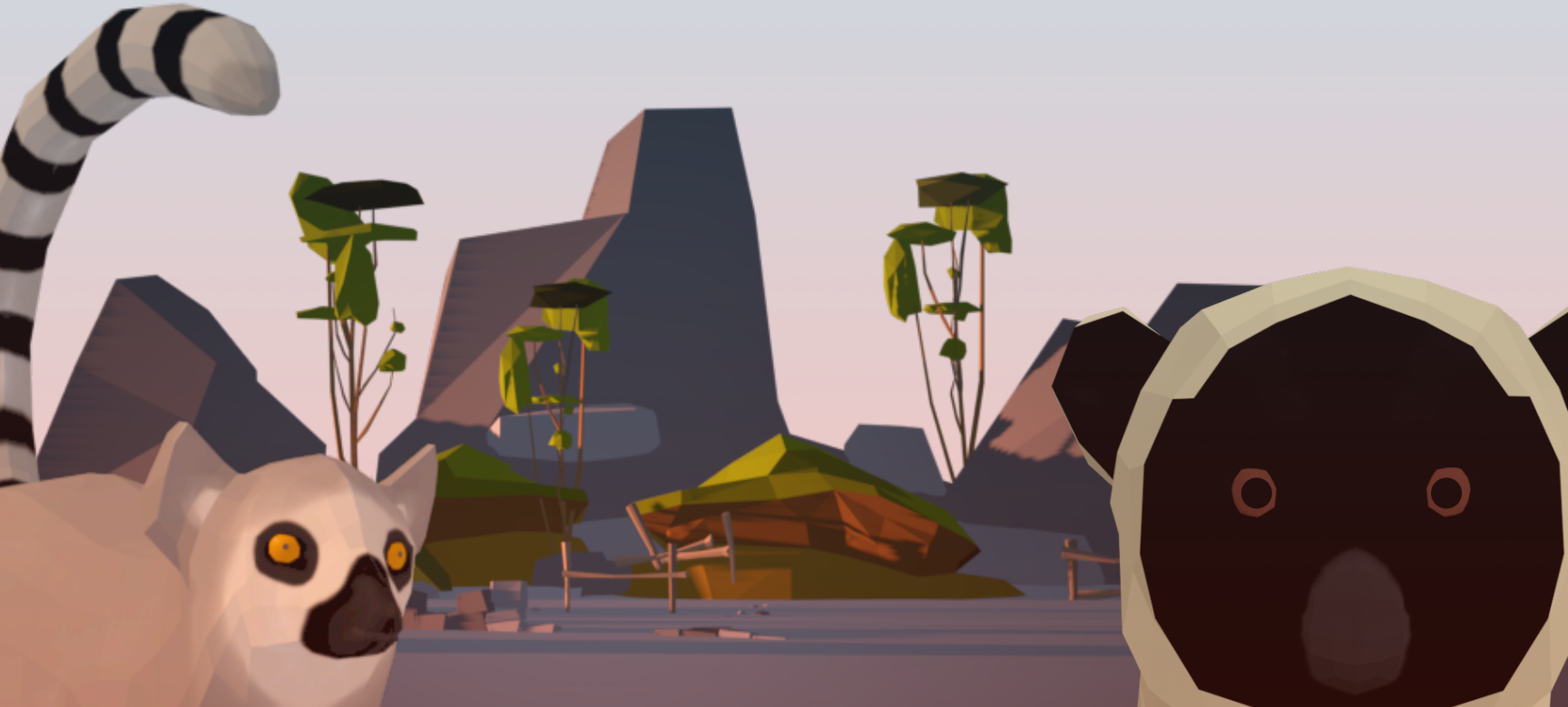
Probability for Lemurs



Probability for Lemurs

ISBN: XXX-XXXXXXXX (ISBN To Be Determined)

This book is for sale at https://leanpub.com/probabilityforlemurs

This is version V.20220314 of the book, date stamp equal to the publish date in CST.

This is a Leanpub book. Leanpub empowers authors and publishers with the Lean Publishing process. Lean Publishing is the act of publishing an in-progress ebook using lightweight tools and many iterations to get reader feedback, pivot until you have the right book and build traction once you do.

Special Request from the Author:

If You Came Across This Book for Free and Found It Useful Please Buy It!

I get it - on the Internet it's always easier to just take things for free and not think twice about it, that's just how it goes. Times are often tough for many of us, particularly for those of us starting out in our careers who perhaps don't have as much disposable income. I totally understand and that's fine.

But here's the thing - I spent many hours writing this book, and I don't have any nefarious designs toward creating some income scheme through trying to sell my readers on some bag of goods. This is literally just math education. I want to make it just a little bit easier for people around the world to learn about probability (and other math concepts in the future).

So, if you want to pitch in I would like to humbly request that you please buy this book. This will help me pay illustrators and improve future books and help me convince my wife that this is a good idea.

You can buy it here on Leanpub, it's only $4.99 USD. http://leanpub.com/probabilityforlemurs.

The upside is that you will get updates as the book improves. If you're in a really tough spot and still really need the book, that's fine, I trust you. You can buy it for free at this link here. This at least helps me track how many people are following along.

You could also help by promoting my book on Twitter by clicking here if you wish. Thanks!

I would like to give a special thanks to those who helped make this book a success, either by reviewing scribbles and drawings, buying LeanPub copies early on, supporting me through the process, giving me tips on what to do or other ways I may not have even realized.

Wout Mertens • Andrea Salazar • Daniel Feldman • Vance Trendov • Justin Grammens • Daniel Acosta •

This book is Dedicated to my son,

Sebastian Salazar Delaney.

00. Contents

Table of Contents Key:

Completed * Incomplete * Not Started

Exploring Data

00. Contents
01. Numbers
02. Populations
03. Samples
04. Mean Value
05. Variance
06. Random Variables
07. Probability Density Function
08. Central Limit Theorem
09. Standard Deviation
10. Cumulative Distribution Function
11. Confidence Intervals
12. Other Distributions
13. Significance Level, P-Value
14. Probability Space

Finding Relationships

15. Inference
16. Hypothesis Testing
17. Chi Squared Independence Test
18. t-Test
19. ANOVA
20. Regression Analysis
21. Pearson's R
22. Conditional Probability
23. Martingales
24. Correlation

Predictions

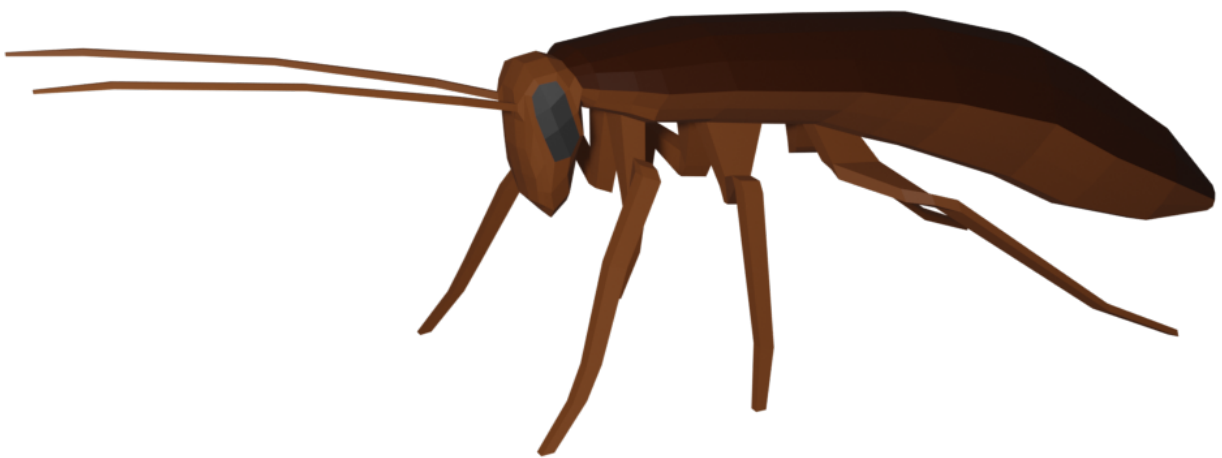
25. Frequentist Probability
26. Expected Value
27. Binomial
28. Bayes' Theorem
29. Poisson Process
30. Markov Chains
31. Statistical Process Control

Credits

32. Licenses to Graphics Used

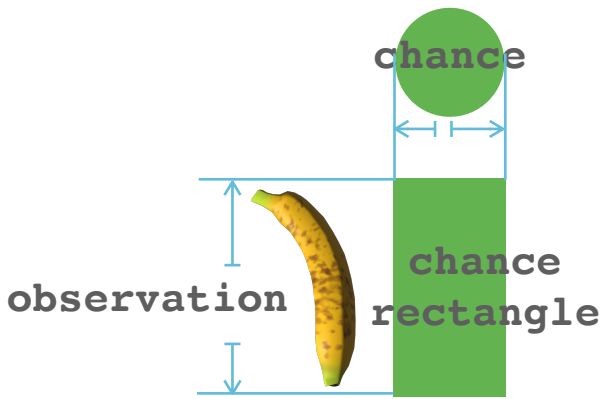
28. Bayes Theorem

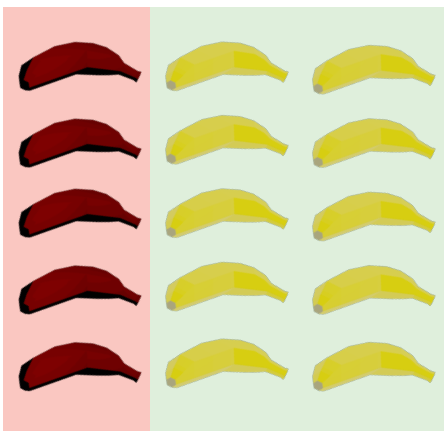
Of course we have already covered that Lemurs like bananas. Unfortunately for lemurs, there are bugs that also like bananas. This can be a big problem for lemurs, especially since it is impossible to tell for sure whether a banana has a bug in it until after the banana has been opened.



28. Bayes Theorem

We learned from Expected Values that the value of an Observation one can expect, or Probability, from a future output can possibly be based upon an input Chance value multiplied by measurements. When lemurs do this, they are creating a fictional idea about how the world works based upon frequencies of past Observations.

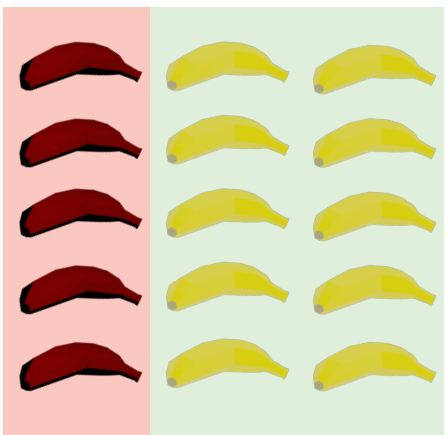




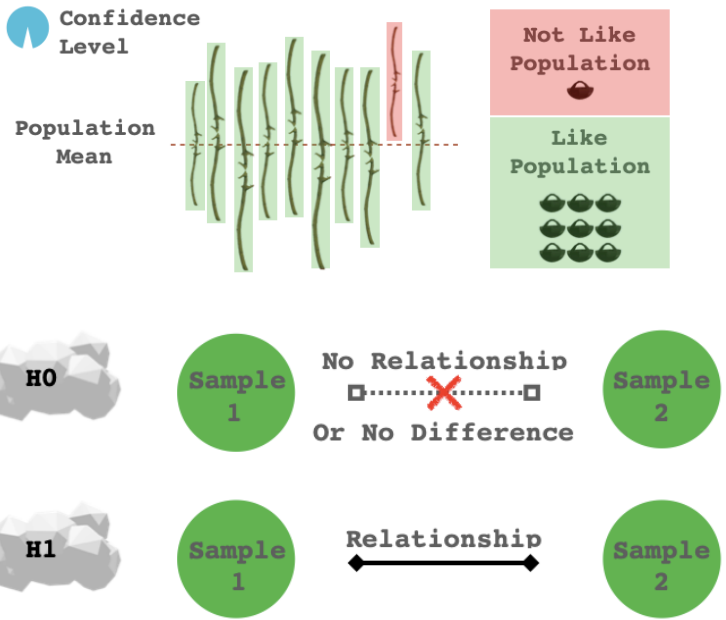
28. Bayes Theorem

Using Categorial Data, the Expected Value from past Observations of bug-ridden and clear bananas might be 2/3. There should always be another idea behind this type of thinking, that some kind of hypothesis testing and a confidence interval is included in this 2/3 Expected Value.

Categorical Data

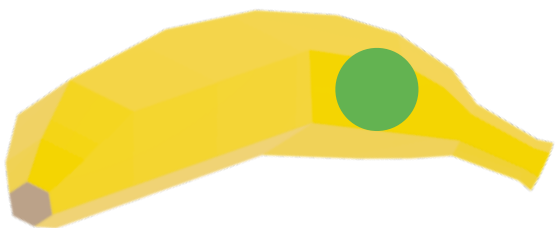


EV = 2/3 Good



28. Bayes Theorem

In the case of Bayesian Probability there is however, a test that can be performed which basically gives a clue as to what level of certainty, called a Bayes factor, there may be bug inside the banana even before opening it. Instead of using hypothesis testing and confidence intervals as our clues, the clues are within the Observations.



28. Bayes Theorem

In a whole forest of bananas, what are the chances that you pick a banana with a bug in it, before you open the banana. Imagining that we have a test for bugs that doesn't always work.

Chances of Something, Given Experience and a Clue

Base Rate = 1/36 bug

99% of time,

spot -> bug

no spot ->

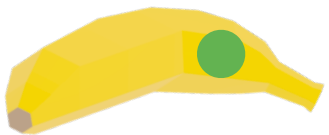
no bug

Bayes Factor

28. Bayes Theorem

In a whole forest of bananas, we pick one banana. There is a spot on it, or not. There are actually two directions the clue can go.

Direction 1



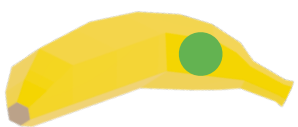
```
99% of time,
          spot -> bug
1% of time
        spot -> no bug
```

The Bayes factor, or percent value of the clue could be different in each direction. But in our example here, the clue works 99% of the time in both directions.

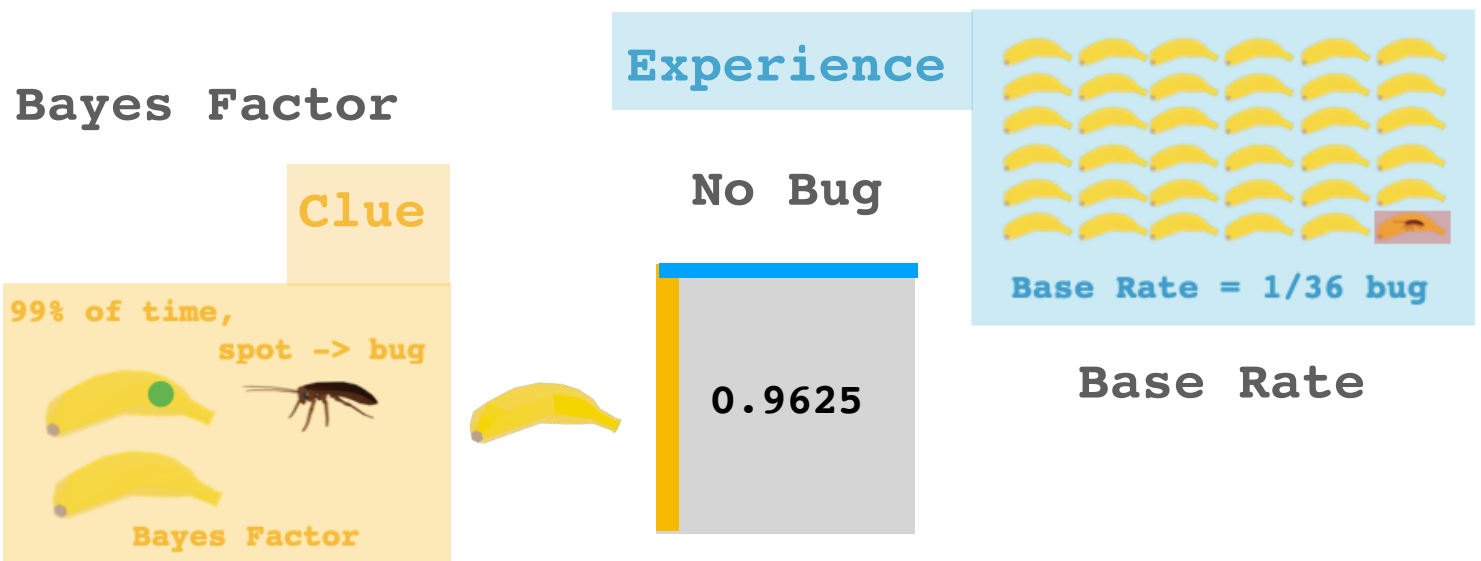
Direction 2

```
99% of time,
     no spot -> no bug
1% of time
     no spot -> bug
```

There are four possible events that can happen that we know ahead of time.

	Possible Test Outcomes	
Clue	✓ Clue was Correct	✗ Clue was Incorrect
	No Bug	
		No Bug

The Clue could actually have different values in different directions, but here it is 99% in both directions to keep things simple. So a banana with no spot means no bug 99% of the time, so the height of the rectangle is 99%. The width of the rectangle is 35/36 because this is the Base Rate from Experience out of 36 bananas, not taking into account any clue.

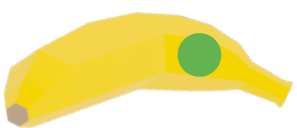


The concept of Types of Errors from Hypothesis Testing is kind of the same to Possible Test Outcomes in Bayesian thinking. But instead of the Error Type meaning incorrectly accepting or rejecting the Null Hypothesis, now it means the Clue was wrong.

Types of Errors - Bayesian Probability		
Part of Error	Type 1	Type 2
Test Outcome	Clue Said Yes, But Was Wrong	Clue Said No, But was Wrong
Our Prediction based upon Refutation	We predicted Yes because clue showed us most likely Yes.	We predicted No because clue showed us most likely No.
Reality	Clue was wrong - False Positive.	Clue was wrong - False Negative.

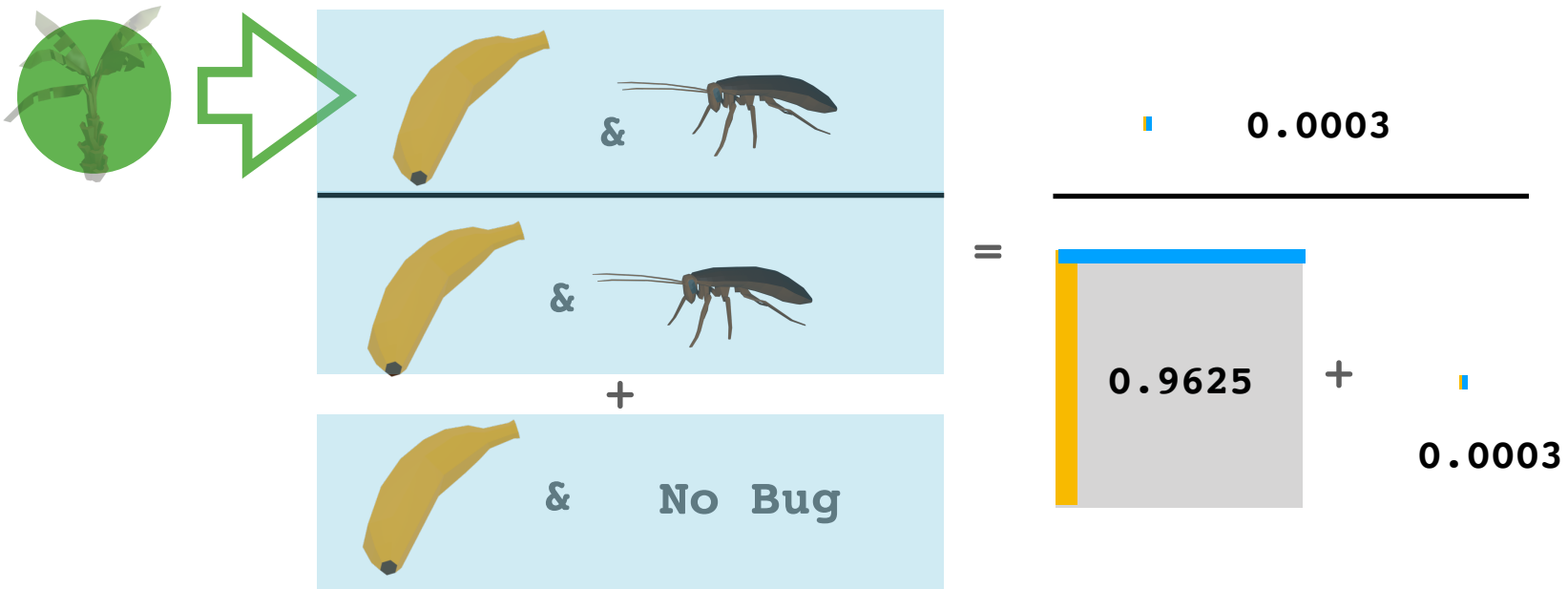
28. Bayes Theorem

To make things easier for lemurs, we can look at each of the Base Cases vs the Clues put near each other on a table.

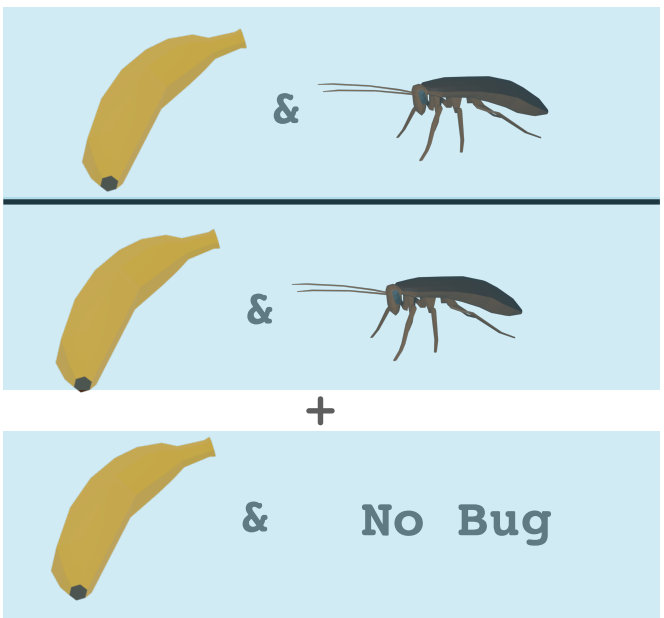
Clue	Experience:  1/36	No Bug 35/36
99% correct 1% not correct 	0.000278	0.9625
99% correct 1% not correct 	0.0257	0.0097

28. Bayes Theorem

Out of a whole forest of bananas, we pick one banana. Look at it and see there is no green spot. What are the chances it has a bug? It's not 1 in 36 anymore because we know there is no spot. It's not 0% because we know the test is not perfect.



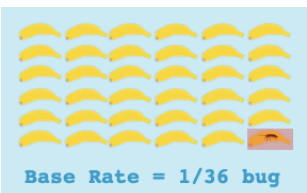
To find the accurate result, which is still small, we have to divide the chance of no spot and a bug by the sum of no spot and a bug plus no spot and no bug. These might seem like small numbers, but there is a 4% difference between 0.000278 and 0.000289!



$$= \frac{0.000278}{0.9625 + 0.000278} = 0.000289\ldots$$

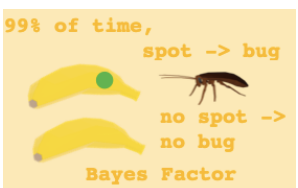
This final calculation including both the base case and the clue shows us the real risk. Using just experience or just the clue would be much too fearful.

0.0289% **is the risk in reality**



$1/36 = 2.78\%$

96 times too fearful

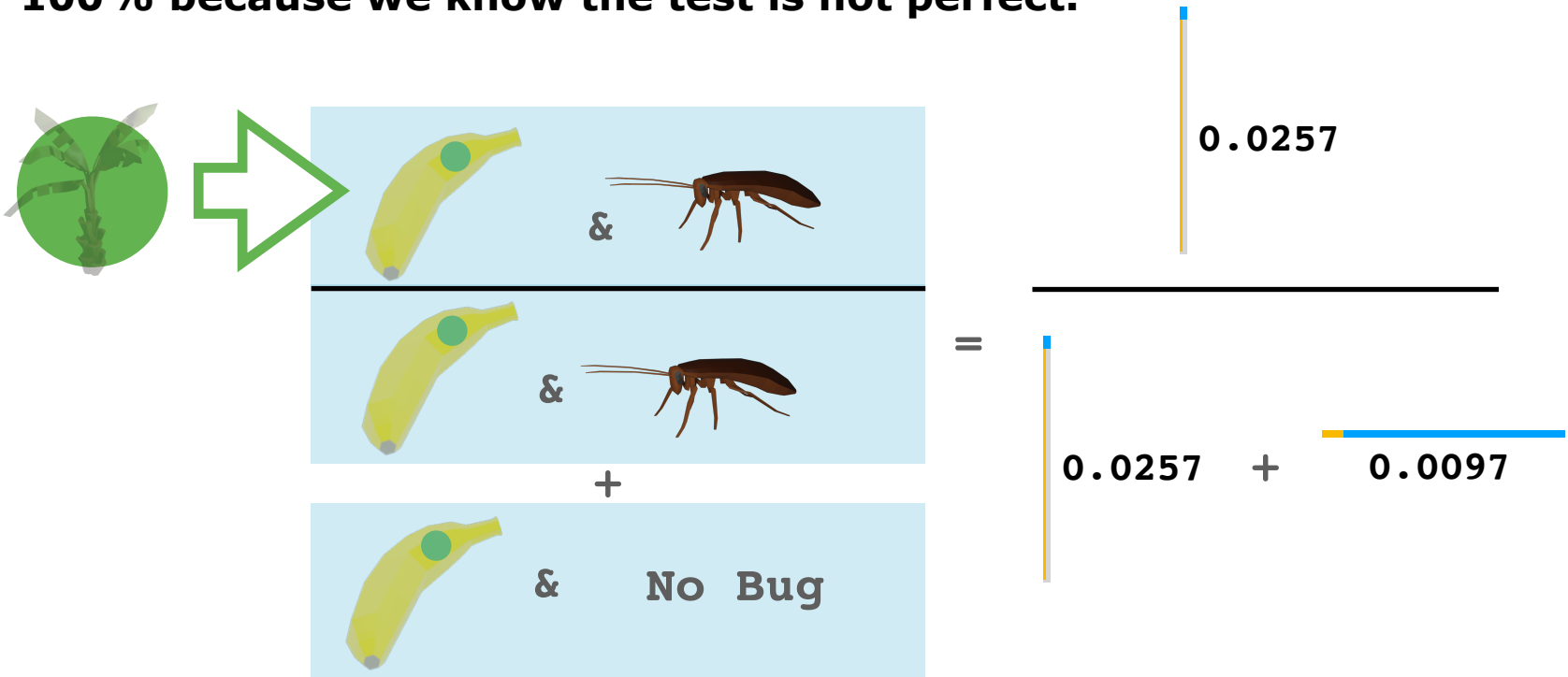


$1/100 = 1.00\%$

35 times too fearful

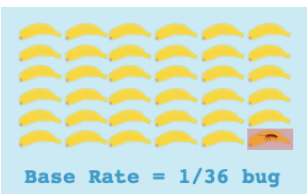
28. Bayes Theorem

Out of a whole forest of bananas, we pick one banana. Look at it and see there is a green spot. What are the chances it has a bug? It's not 1 in 36 anymore because we know there is a spot. It's not 100% because we know the test is not perfect.



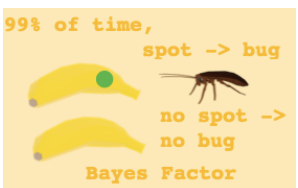
This final calculation including both the base case and the clue shows us the real risk. Using just experience or just the clue would be much too fearful.

72.60% **is the risk in reality**



$35/36 = 97.2\%$

1.3 times too confident



$99/100 = 99.0\%$

1.4 times too confident

28. Bayes Theorem

Crate of 1000 bananas marked to be thrown out because of spots.

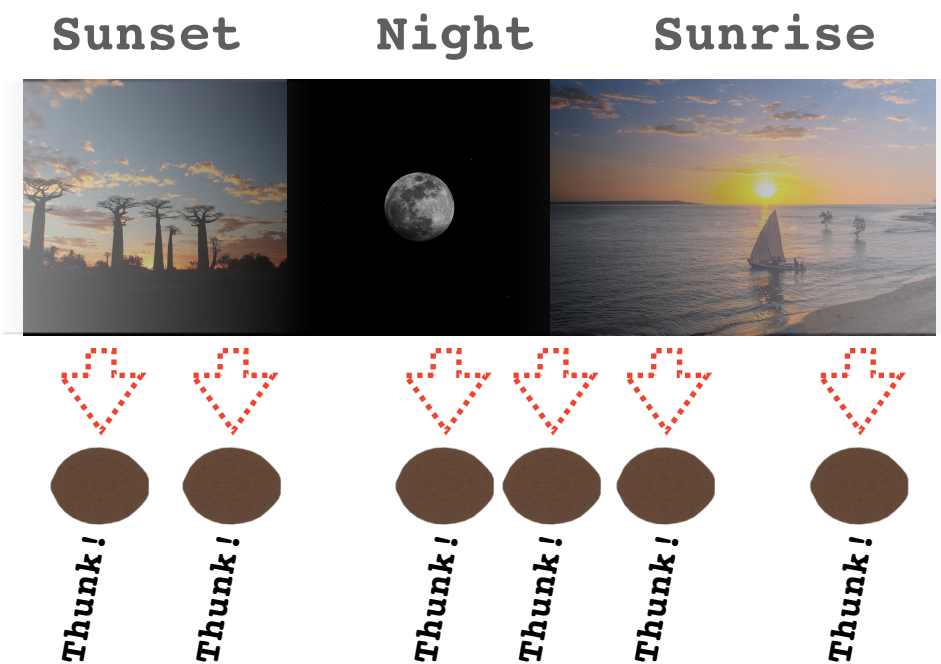
72.60% * 1000 = 726 Have Bugs

Bayesian Bug Probability

274 Would have been wasted!

29. Poisson Process

Some lemurs eat coconuts. These nocturnal lemurs, called Aye-ayes, can listen for the sound of coconuts falling and hitting the ground at night.



30. Markov Chains

Lemurs are good at leaping. But, they can't leap across a whole forest in one jump. They leap from tree to tree, using the trees or banana plants that are within leaping distance.