

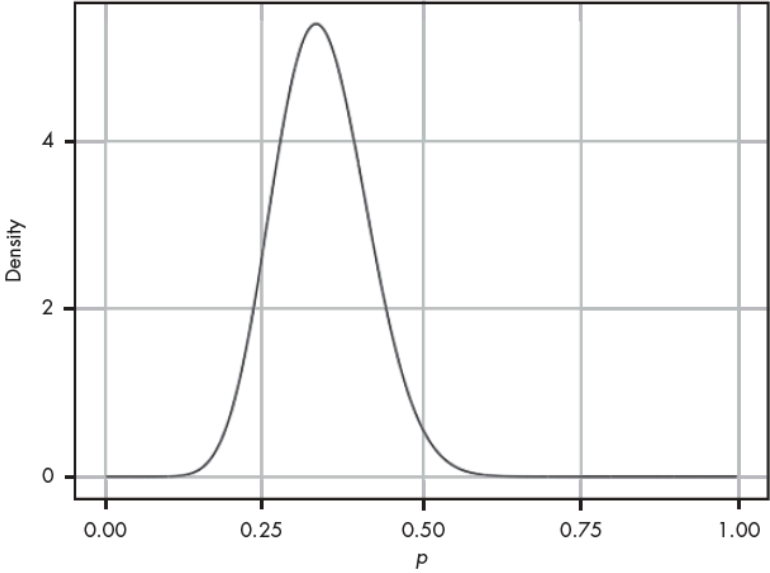
Bayesian Statistics the Fun Way

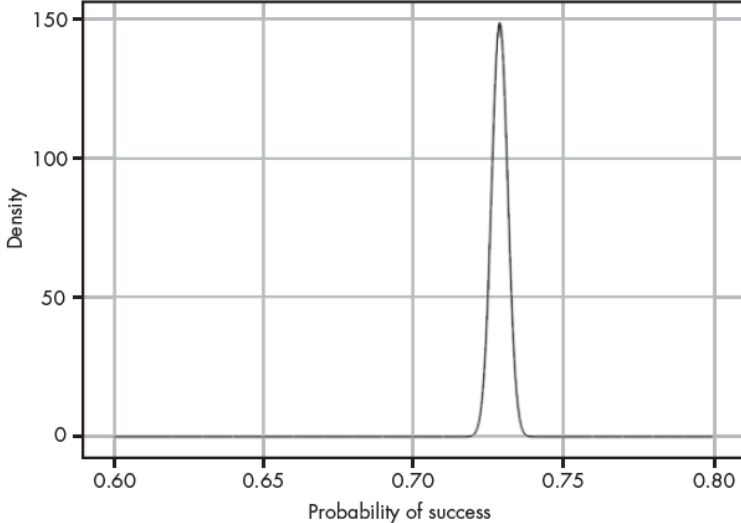
Understanding Statistics and Probability with Star Wars, LEGO, and Rubber Ducks

by Will Kurt

Errata updated to print 9

Page	Error	Correction	Print corrected																								
17	Equation replacement	$P(H_{\text{no article}}) = 20 \times (1 - P(H_{\text{no article}}))$	Print 2																								
29	So, using our die roll and coin toss example, the probability of rolling a number less than 6 or flipping a heads is:	So, using our die roll and coin toss example, the probability of rolling a number equal to 6 or flipping a heads is:	Print 3																								
40	Figure replacement	<p style="text-align: center;">Binomial Distribution for 10 Rolls of a Six-Sided Die</p> <table border="1"> <caption>Data for Figure 4-2: Binomial Distribution for 10 Rolls of a Six-Sided Die</caption> <thead> <tr> <th>k</th> <th>Probability B(k; 10, 1/6)</th> </tr> </thead> <tbody> <tr><td>0</td><td>0.0107</td></tr> <tr><td>1</td><td>0.3206</td></tr> <tr><td>2</td><td>0.2980</td></tr> <tr><td>3</td><td>0.1503</td></tr> <tr><td>4</td><td>0.0547</td></tr> <tr><td>5</td><td>0.0107</td></tr> <tr><td>6</td><td>0.0009</td></tr> <tr><td>7</td><td>0.0001</td></tr> <tr><td>8</td><td>0.0000</td></tr> <tr><td>9</td><td>0.0000</td></tr> <tr><td>10</td><td>0.0000</td></tr> </tbody> </table> <p style="text-align: center;">Figure 4-2: The probability of getting δ k times when rolling a six-sided die 10 times</p>	k	Probability B(k; 10, 1/6)	0	0.0107	1	0.3206	2	0.2980	3	0.1503	4	0.0547	5	0.0107	6	0.0009	7	0.0001	8	0.0000	9	0.0000	10	0.0000	Print 3
k	Probability B(k; 10, 1/6)																										
0	0.0107																										
1	0.3206																										
2	0.2980																										
3	0.1503																										
4	0.0547																										
5	0.0107																										
6	0.0009																										
7	0.0001																										
8	0.0000																										
9	0.0000																										
10	0.0000																										

Page	Error	Correction	Print corrected
51	Figure replacement	 <p data-bbox="1058 808 1860 834">Figure 5-3: Visualizing the beta distribution for our data collected about the black box</p>	Print 3
51	What we get in the end is a function that describes the probability of each possible hypothesis for our true belief in the probability of getting two heads from the box . . .	What we get in the end is a function that describes the probability of each possible hypothesis for our true belief in the probability of getting two coins from the box . . .	Print 5
53	Here we calculate the probability that the chance of getting two coins from the box is 0.5, given the data:	Here we calculate the probability that the chance of getting two coins from the box is less than or equal to 0.5, given the data:	Print 3
71	numberOfRedStuds = $P(\text{yellow} \mid \text{red}) \times \text{numberOfRedStuds} = 1/5 \times 20 = 4$	numberOfRedUnderYellow = $P(\text{yellow} \mid \text{red}) \times \text{numberOfRedStuds} = 1/5 \times 20 = 4$	Print 5
78	<ol style="list-style-type: none"> 1. A neighborhood kid hit a baseball through the front window. 2. You left your door unlocked. 2. You forgot that you brought your laptop to work and it's still there. 	<ol style="list-style-type: none"> 1. A neighborhood kid hit a baseball through the front window. 2. You left your door unlocked. 3. You forgot that you brought your laptop to work and it's still there. 	Pending
87	<p>We just add the alphas for our prior and posterior and the betas for our prior and posterior, and we arrive at a normalized posterior. Because this is so simple, working with the beta distribution is very convenient for Bayesian statistics. To determine our posterior for Han making it through the asteroid field, we can perform this simple calculation:</p> <p>Beta (20002, 7401) = Beta (2 + 20000, 7400 + 1)</p>	<p>We just add the alphas for our prior and posterior and the betas for our prior and likelihood and we arrive at a normalized posterior. Because this is so simple, working with the beta distribution is very convenient for Bayesian statistics. To determine our posterior for Han making it through the asteroid field, we can perform this simple calculation:</p> <p>Beta (20002, 7441) = Beta (2 + 20000, 7440 + 1)</p>	Print 5

Page	Error	Correction	Print corrected												
88	Figure replacement	<p style="text-align: center;">Distribution of our posterior belief Beta(2+20000,7440+1)</p>  <p style="text-align: center;">Figure 9-3: Combining our likelihood with our prior gives us a more intriguing posterior.</p>	Print 5												
94	You first instinct is probably to average these measurements.	Your first instinct is probably to average these measurements.	Print 8												
105	<table border="1" data-bbox="172 922 762 1175"> <thead> <tr> <th data-bbox="172 922 317 992">Observation</th> <th data-bbox="317 922 468 992">Difference from mean</th> </tr> </thead> <tbody> <tr> <td data-bbox="172 992 317 1081">Group <i>b</i></td> <td data-bbox="317 992 468 1081"></td> </tr> <tr> <td data-bbox="172 1081 317 1175">2.80</td> <td data-bbox="317 1081 468 1175">-0.16</td> </tr> </tbody> </table>	Observation	Difference from mean	Group <i>b</i>		2.80	-0.16	<table border="1" data-bbox="1039 922 1629 1175"> <thead> <tr> <th data-bbox="1039 922 1184 992">Observation</th> <th data-bbox="1184 922 1335 992">Difference from mean</th> </tr> </thead> <tbody> <tr> <td data-bbox="1039 992 1184 1081">Group <i>b</i></td> <td data-bbox="1184 992 1335 1081"></td> </tr> <tr> <td data-bbox="1039 1081 1184 1175">2.80</td> <td data-bbox="1184 1081 1335 1175">-0.2</td> </tr> </tbody> </table>	Observation	Difference from mean	Group <i>b</i>		2.80	-0.2	Print 5
Observation	Difference from mean														
Group <i>b</i>															
2.80	-0.16														
Observation	Difference from mean														
Group <i>b</i>															
2.80	-0.2														
105	Equation replacement	$\sum_{i=1}^5 a_i - \mu_a = 0 \quad \sum_{i=1}^5 b_i - \mu_b = 0$	Print 5												
106	Equation replacement	$\frac{1}{5} \times \sum_1^5 a_i - \mu_a = 0.04 \quad \frac{1}{5} \times \sum_1^5 b_i - \mu_b = 0.416$	Print 5												

Page	Error	Correction	Print corrected
116	Equation replacement	$N(\mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \times e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	Print 3
127	<pre>xs <- seq(0.005,0.01,by=0.00001) xs.all <- seq(0,1,by=0.0001)</pre>	<pre>xs <- seq(0.005,0.01,by=0.00001) xs.all <- seq(0,1,by=0.0001)</pre>	Print 5
130	As Figure 3-5 illustrates, the point where this line intersects the x-axis gives us our median!	As Figure 13-5 illustrates, the point where this line intersects the x-axis gives us our median!	Print 5
163	$P(D H_1) = 0.94 \times 0.89 = 0.78$	$P(D H_1) = 0.94 \times 0.83 = 0.78$	Print 3
164	The prior odds look like this:	The probabilities look like this:	Print 5
164	Equation replacement	$O(H_1) \times \frac{P(D H_1)}{P(D H_2)} = \frac{11}{37,000} \times 2.23 = \frac{245}{370,000}$	Print 5
178	Equation replacement	$PO = O(H_2)' \times \frac{P(D_{15} H_2)}{P(D_{15} H_3)} = \frac{1}{1,000} \times \frac{\left(\frac{9}{10}\right)^{14} \times \left(1 - \frac{9}{10}\right)^1}{\left(\frac{9}{10}\right)^{14} \times \left(1 - \frac{9}{10}\right)^1} = \frac{1}{1,000}$	Print 3
224	Since you've run half a mile , using this simple formula, we can figure out:	Since you've run half an hour , using this simple formula, we can figure out:	Print 5
234	A3. This is the same as B(5; 10, 1/23). As expected, the probability of this is extremely low: about 1/32,000 .	A3. This is the same as B(5; 10, 1/13). As expected, the probability of this is low: about 1/2,200 .	Print 6
236	Luckily we already did all this work earlier in the chapter, so we know that (A) = 4/1,000 and P(B) = 3/(100,000).	Luckily we already did all this work earlier in the chapter, so we know that (A) = 8/100 and P(B) = 3/(100,000).	Print 5
237	Plugging in our numbers, we get an answer of 100,747/25,000,000 or 0.00403 .	Plugging in our numbers, we get an answer of 800,276/10,000,000 or 0.0800276 .	Print 5
242	<pre>temp.sd <- my.sd(temp.data)</pre>	<pre>temp.sd <- sd(temp.data)</pre>	Print 4
250	$P(D H_2) = 0.63 \times 0.55 \times 0.49 = 0.170$	$P(D H_2) = 0.94 \times 0.83 \times 0.49 = 0.382$	Print 5

Page	Error	Correction	Print corrected
250	This means that given the Bayes factor alone, vestibular schwannoma is a roughly two times better explanation than labyrinthitis. Now we have to look at the odds ratio:	This means that given the Bayes factor alone, vestibular schwannoma is a roughly four times better explanation than labyrinthitis. Now we have to look at the prior odds ratio:	Print 5
251	The end result is that labyrinthitis is only a slightly better explanation than vestibular schwannoma .	The end result is that vestibular schwannoma is only a slightly better explanation than labyrinthitis .	Print 5
254	Equation replacement	$50 = \frac{9}{19} \times 950$ $\text{BF} = 950$	Print 5
254	<pre>dx <- 0.01 hypotheses <- seq(0,1,by=0.01)</pre>	<pre>dx <- 0.01 hypotheses <- seq(0,1,by=dx)</pre>	Print 5