2.4 Laplace's Rule and Priors

Laplace has got into trouble for this for a long time.
If you do a binary experiment once, and get one success, the Rule suggests that your best estimate for a success the next time is 1/3. This seems very conservative; a more reasonable value would presumably have been 1. We can trace the source of this to the very broad prior we put on ρ in the first place. The difficulty becomes clearer if we make the example more specific. In 1908, Onnes succeeded in liquifying helium at a temperature of 4 K. After he had done it once, do you suppose he thought the chance of doing it again was 1/3? As an experienced experimentalist, he may well have thought the chance was not as high as 1, but as low as 1/3? Of course all sorts of background information comes into play here. In the case of physical phenomena, where we believe in the "regularity of nature", our prior on ρ is biassed towards zero or one. This is the point of the Haldane prior. Using this, instead of the ordinary prior, it is easy to show that the rule of succession becomes $N/(N+M)$. Now, the probability of success after one successful trial is 1.
Laplace got into trouble because he used his Rule to predict the probability of the sun rising again tomorrow. He actually did this to illustrate the importance of background information, but this point was too subtle for his detractors. Jaynes (2003) §18.6 discusses various apparent paradoxes that result from uninformed use of the Rule.
Incidentally, although we have expressed the Rule as an average value (the posterior mean) it is more informative to see it as an answer to the question "What is the probability of success at this trial, given that we had N successes and M failures before?" This is

$$\text{prob}(\text{success} \mid \rho)\text{prob}(\rho \mid N, M).$$

The first term is just ρ. The second term is the posterior density distribution we got before, so is

$$\text{prob}(\rho|\text{data}) = \rho^n(1-\rho)^{m-n}/B[n+1, m-n+1].$$

(This is an example of taking the posterior from one experiment as the prior of the next.) Integrating over ρ we get the Rule of Succession again.