Optimisation and Statistical Data Analysis
Exercise Set 9 (Binomial proportion)

Problem 1

According to the Finnish Statistics Bureau (`http://www.stat.fi`), $30\,546$ boys and $29\,415$ girls were born in Finland in 2011.

(a) Find the posterior 95% confidence interval for θ, the chance of a Finnish birth being a boy. Use a uniform prior.

(b) What is the posterior probability of $\theta > 0.5$?

(c) Suppose 4000 children are born in Finland this month. Using the 2011 data, plot the posterior distribution of the number of male births. What is the posterior probability that this number is > 2000?

(d) Find the number of boys and girls born in Finland in 2012 and update the answer to (a). Use the recursive update formula.

Problem 2

You are consulting for a non-profit organisation (NGO) that is running an advertisement on a website. The NGO is considering changing the image that is displayed with the ad, so you suggest an A/B test: during one week, some of the visitors to the website will see the advertisement with image A (a kitten), and the remaining visitors will see image B (an endangered spider) instead.

(a) Based on their previous experience with this ad, people at the NGO say that the proportion of visitors who click on the ad (click-through rate, CTR) is between 10% and 50%, so assume CTR $\sim$ beta$(3, 7)$. Plot the density function of this prior distribution. Plot the distribution of the number of click-throughs in 100 visits.

(b) The results on day 1 of the test are as follows.

	# visits	# clicks
A	7706	2654
B	6821	2214

Plot the posterior density for $\mathrm{CTR}_A - \mathrm{CTR}_B$ and find the posterior probability that $\mathrm{CTR}_A > \mathrm{CTR}_B$.

Answers 1. (a) $(0.5054, 0.5134)$ (b) 0.9999981 (c) 0.87 (d) $(0.5066, 0.5123)$
2. (b) 0.995