Exercises

E1. You are going to play poker at a casino and you consider the chances of winning in each game to be about .4. This value is based on 12 earlier games you consider to represent exchangeable random quantities with respect to the dichotomous result (win/loose). After the night at the casino you have won 17 out of 25 games that you continue to consider exchangeable in a similar sense. What is your renewed opinion about chances of winning if you describe the situation using a Beta-Binomial model?

E2. Consider a dichotomous property of individuals in a large finite population. Let the population size be N and the size of a sample taken without replacement be n. This situation corresponds to an urn containing N balls, which are either black or white. Let θ denote the number of black balls in the urn. The probability of a sample of n balls contains x black ones is given by the hypergeometric expression

$$p(x|\theta) = \frac{\binom{\theta}{x}\binom{N-\theta}{n-x}}{\binom{N}{n}}.$$

Further, if $p(\theta = r), r = 0, ..., N$, specifies the prior probabilities for θ, we get the posterior probability of $\theta = N$ as

$$p(\theta|x = n) = \frac{p(x = n|\theta = N)p(\theta = N)}{\sum_{r=n}^{N} p(x = n|\theta = r)p(\theta = r)}.$$

Assume a uniform prior distribution over the set of N possible values for the parameter θ and calculate explicitly the above posterior probability. Analyze the situation from a general scientific perspective. Hint: you can utilize the following general result

$$\sum_{r=0}^{N} \binom{r}{l}\binom{N-r}{m} = \binom{N+1}{l+m+1},$$

where $l + m = n$ is the sample size for the hypergeometric model and l stands for the number of picked individuals having the characteristic of interest (*e.g.* black colour).

E3. You observe $X = x$ to learn about parameter θ, for which you have the prior $p(\theta)$. You are uncertain about the shape of a suitable distribution, so you define the likelihood of x as

$$p(x|\theta) = \sum_{i=1}^{k} \pi_i p_i(x|\theta),$$

where $0 < \pi_i < 1$ and $\sum_{i=1}^{k} \pi_i = 1$. What is the shape of your posterior for θ? Consider how the information contained in the observation is reflected by the posterior.

E4. Consider a case of suspected discrimination, where 48 women and 259 men with equal educational background and working history, have taken a test required for promotion in their organization. Of women 26 passed the test, and the corresponding figure for men was 206. What would you say about the suspected discrimination on the basis of this evidence? As a guideline, use the following two ways of comparing statistically the hypothesis of no discrimination versus discrimination: 1) frequentist hypothesis test, 2) Bayesian model comparison, e.g. Bayes factor, which compares the marginal likelihoods of the data under the two alternative models.

E5. You work in an immigration office and have to determine the kinship of two persons. They claim to be siblings, however, officials in their country of departure provide you with the information that they might be first cousins. To get more evidence these persons are subjected to a genetic test, where the identity of their alleles in determined for 7 marker loci (each located in a different chromosome). Let us denote the identity of alleles at one locus with 1 and inequality by 0. You observe the values 0,0,1,0,0,0,0. Assume that siblings have probability 1/2 for allele identity, and correspondingly for cousins this figure is 1/4, independently for loci in different chromosomes. What can you conclude about the kinship by making a statistical comparison of the two hypotheses? How would your conclusions be affected by the fact that the two individuals would not be related at all?