Assignments III

T12. Consider a vector-valued parameter $\boldsymbol{\phi}$, satisfying $\boldsymbol{\phi}'\boldsymbol{\phi} = 1$. Assume *a priori* that the uncertainty about $\boldsymbol{\phi}$ is described by a uniform distribution in the q-dimensional unit hypersphere S_q. This corresponds to the circumference of the unit circle when $q = 2$, and to the surface of the unit ball when $q = 3$. Such a prior can be used in a statistical model for directional data (see e.g. http://en.wikipedia.org/wiki/Von_Mises_distribution) and it arises also in other contexts where probability model structure is such that the length of a certain parameter vector is not identifiable but only the direction of the vector (e.g. co-integration models in econometrics). The density function of the prior equals

$$\pi(\boldsymbol{\phi}) = \frac{\Gamma(q/2)}{(2\pi)^{q/2}}$$

Assume we wish to compare models in the set $I = \{2, ..., 10\}$, where each integer $j \in I$ corresponds to the dimension q of parameter $\boldsymbol{\phi}$. Consider how the prior behaves when we compare model dimension $q + 1$ to q for these values. What is the difference between this comparison and estimation problem, where q is fixed?

T13. Consider an analogous problem as in the previous case, but remove the length restriction to the unit hypersphere, and assume $\boldsymbol{\phi} \in \mathcal{R}^q$. Assume each element of $\boldsymbol{\phi}$ to follow independently the standard normal distribution $N(0, 1)$. How does this prior behave in the same model comparison problem as in the previous assignment?

T14. The New York Times has once upon a time published the following table:

Accused\Sentence	Death	Other
Black	59	2448
White	72	2185

presenting the results of 4764 murder cases in courts of the State of Florida durings years 1973-79 with respect to the ethnicity of the accused murderer and the sentence. Consider the evidence for independence of these two quantities on the basis of the observations using a multinomial model and a Dirichlet prior. Observe how the model structure is simplified under the independence assumption (you consider the marginal distributions). The marginal likelihood expression for the Multinomial-Dirichlet-model is (see also lecture material)

$$\frac{\Gamma(\sum_{i=1}^{k}\lambda_i)}{\Gamma(\sum_{i=1}^{k}\lambda_i + n_i)}\prod_{i=1}^{k}\frac{\Gamma(\lambda_i + n_i)}{\Gamma(\lambda_i)},$$

where k is the number of different outcomes in the multinomial distribution, λ_i is a hyperparameter of the Dirichlet distribution (the "prior relative importance" of the ith outcome) and n_i is the number of observed cases. Check how the inferences are affected by the choice of λ_i.

T15. We continue with the previous assignment, by splitting the table with respect to the ethnicity of the victim:

Black victim :

Accused\Sentence	Death	Other
Black	11	2209
White	0	111

White victim :

Accused\Sentence	Death	Other
Black	48	239
White	72	2074

This is an example of a so called Simpson's paradox, where absence of certain information (here victim ethnicity) affects the dependence between other things. For the three considered quantities (victim ethnicity (U), accused ethnicity (S) and penalty type (T)) there are eight possible models, possibly stating marginal or conditional independence between the U, S, T. Ordered with respect to the degree of simplicity, the models range from complete independence ($P(U, S, T) = P(U)P(S)P(T)$), to complete dependence (no restrictions on $P(U, S, T)$). If, for instance, a model states U to be independent from the others, we have $P(U, S, T) = P(U)P(S, T)$. Correspondingly, by assuming conditinal independence between S and T we get

$$P(U, S, T) = P(S, T|U)P(U) = P(S|U)P(T|U)P(U)$$

Show that we get for these three binary variables in total 8 different models stating either no or some independence. Use the same model family as in the previous assignment, and calculate the posterior probabilities for the models. Hint: By using the relationship $P(S, U) = P(S|U)P(U))$ between joint and conditional probabilities, you can see which marginal distributions need to be investigated for a model stating a conditional independence assumption.

T16. A typical test for the ESP sensitivity of an individual correcponds to the thumbtack tossing scenario. For instance, a test subject tries repearedly (n times) to guess from two choices the color of a covered card, while the the tester thinks about the color. Null hypothesis here is the symmetric probability (.5) between correct and wrong answers, meaning that the test subject does not have ESP sensitivity. In an alternative model one could use Beta(α, β) prior distribution for the probability of the correct answer. What would you conclude about ESP sensitivity on the basis of 7 correct and 3 wrong answers out 10 repeats? How does the choice of α and β affect your conclusions? What about the importance of the sample size?

T17. You have two putative models for a data set: Negative Binomial(θ_1)- and Poisson(θ_2)-distribution. Assume first that the expectations for the two models are equal and the parameter values are $\theta_1 = 1/3, \theta_2 = 2$. Compare models when you observe $x_1 = x_2 = 0$ or $x_1 = x_2 = 2$. How does the model comparison behave if you instead use a Beta prior distribution for θ_1 and Gamma prior distribution for θ_2 and utilize Bayes factor or posterior probabilities?