Take home final

STA701.01, Statistical Inference
Fall Semester, 2013

Due: Tues Dec 10th, 2013

Consider 2-way contingency table under the independence model, i.e., fixing the row and column sums ($\theta = \lambda + \lambda_X + \lambda_Y$).

1 Implement a MCMC procedure using "basic moves" in R with the hypotheses

$$H_0 : \lambda_{XY} = 0 \text{ vs } H_1 : \lambda_{XY} \neq 0,$$

where $\theta = \lambda + \lambda_X + \lambda_Y + \lambda_{AB}$ is the canonical parametor for each cell. Then consider the birthday-deathday data set:

Table 1: Relationship between birthday and death day

	Jan	Feb	March	April	May	June	July	Aug	Sep	Oct	Nov	Dec
Jan	1	0	0	0	1	2	0	0	1	0	1	0
Feb	1	0	0	1	0	0	0	0	0	1	0	2
March	1	0	0	0	2	1	0	0	0	0	0	1
April	3	0	2	0	0	0	1	0	1	3	1	1
May	2	1	1	1	1	1	1	1	1	1	1	0
June	2	0	0	0	1	0	0	0	0	0	0	0
July	2	0	2	1	0	0	0	0	1	1	1	2
Aug	0	0	0	3	0	0	1	0	0	1	0	2
Sep	0	0	0	1	1	0	0	0	0	0	1	0
Oct	1	1	0	2	0	0	1	0	0	1	1	0
Nov	0	1	1	1	2	0	0	2	0	1	1	0
Dec	0	1	1	0	0	0	1	0	0	0	0	0

Table 1 shows data gathered to test the hypothesis of association between birth day and death day. The table records the month of birth and death for 82 descendants of Queen Victoria. A widely stated claim is that birthday-death day pairs are associated. Columns represent the month of birth day and rows represent the month of death day.

Then test with the example on the birthday and deathday tables with the sample size 1,000. What is the estimated p-value?

2 Implement the SIS procedure for the two way tables under the independence model. What is the estimated number of tables for the birthday and deathday tables with the sample size 1,000?

3 Read the paper "Uniformly most powerful Bayesian tests" by Valen E. Johnson Ann. Statist. Volume 41, Number 4 (2013), 1716-1741. `http://projecteuclid.org/DPubS?service=UI&version=1.0&verb=Display&handle=euclid.aos/1378386237`. Read very carefully and

(a) summarize the paper,

(b) pick one theorem and write the proof very carefully, and

(c) discuss your thoughts on this paper. Any criticizm did you have? Any opinions? Do you think the methods in the paper is Bayesian to you? Why and why not?

4 [From Game Theory]
Assume $\Theta = \mathcal{A} = [0, 1]$. Find minimax strategy and the value of the game for each of the following losses:
(a) $\theta - 2\theta a + a/2$,
(b) $(\theta - a)^2$,
(c) $\theta^2 - \theta + a - a^2$,
(d) $a - 2\theta a + 1$,
(e) $|\theta - a| - (\theta - a)^2$.

5 Assume $X \sim Bin(10, \theta)$. It is desired to test

$$H_0 : \theta = 0.4 \text{ vs } H_1 : \theta = 0.6$$

under $0 - 1$ loss (i.e., if the hypothesis is true then 0 otherwise 1). Obtain minimax procedure and compute the least favorable prior distribution.