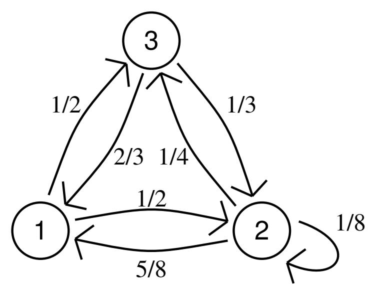
Assignment # 2 MCMC and Bayesian inference

1. Explain what are the essential ingredients to specify a the probability law of a Markov chain on a general state space and how you can use them to derive any finite dimensional distribution

2. Explain what ingredients are sufficient to derive an MCMC approximation of the (finite) quantity
$$I = E_{\pi}[h(X)]$$
and what are the key elements which the approximation error depends on.

3. Let us consider a Markov chain $(X_t)_{t \geq 0}$ defined on the state space $\mathcal{S} = \{1, 2, 3\}$ with the following transition



   (a) Starting at time $t = 0$ in the state $X_0 = 1$ simulate the Markov chain with distribution assignes as above for $t = 1000$ consecutive times

   (b) compute the empirical relative frequency of the two states in your simulation

   (c) repeat the simulation for 500 times and record only the final state at time $t = 1000$ for each of the 500 simulated chains. Compute the relative frequency of the 500 final states. What distribution are you approximating in this way? Try to formalize the difference between this point and the previous point.

   (d) compute the theoretical stationary distribution π and explain how you have obtained it

   (e) Is it well approximated by the simulated empirical relative frequencies computed in (b) and (c)?

   (f) what happens if we start at $t = 0$ from state $X_0 = 1$ instead of $X_0 = 2$?

4. *Coal mining disasters* - Poisson counts with a change point.

From 1851 to 1962 the number of accident in the UK coal minings. Carlin, Gelfand and Smith, 1992 *Applied Statistics*, **41**, 389-405) have used the following statistical model to see whether there has been a change in the rate of occurrence of fatal events possibly due to more severe safety regulations: $(Y_1, ..., Y_{m-1}, Y_m, Y_{m+1}, ...Y_n)$ are the annual counts of accidents

$$\begin{aligned} Y_i &\sim Poi(\lambda) \;\; i = 1, 2, ..., m \\ Y_j &\sim Poi(\phi) \;\; j = m+1, m+2, ..., n \end{aligned}$$

Unknown model parameters are (λ, ϕ, m) while n is known ($n = 112$). Use the following priors

$$\begin{aligned} \lambda &\sim Gamma(\alpha, \beta) \\ \phi &\sim Gamma(a, b) \\ m &\sim Unif\{1, 2, ..., n-1\} \end{aligned}$$

Observed values are:

```
y=c(4,5,4,1,0,4,3,4,0,6,
+ 3,3,4,0,2,6,3,3,5,4,5,3,1,4,4,1,5,5,3,4,2,5,2,2,3,4,2,1,3,2,
+ 1,1,1,1,1,3,0,0,1,0,1,1,0,0,3,1,0,3,2,2,0,1,1,1,0,1,0,1,0,0,
+ 0,2,1,0,0,0,1,1,0,2,2,3,1,1,2,1,1,1,1,2,4,2,0,0,0,1,4,0,0,0,
+ 1,0,0,0,0,0,1,0,0,1,0,0)
```

(a) Write the likelihood

(b) Derive all *full-conditionals*

(c) Implement a Gibbs Sampling to approximate the posterior distribution using $T = 10000$ and discarding the first $T_0 = 1000$ simulations as (*burn-in*).

(d) Provide comments on the simulations and the corresponding approximations of marginal posterior distributions of the parameter of interest.

(e) Can you say which in which year there has been a change? How would you justify your choice?

(f) What is the expected reduction (in percentage) of the rate of accidents in the two periods?

5. Ages (Y_i) and lengths (x_i) of 27 Dugongs have been recorded and the following (non linear) regression model is considered:

$$\begin{aligned} Y_i &\sim N(\mu_i, \tau^2) \\ \mu_i = f(x_i) &= \alpha - \beta\gamma^{x_i} \end{aligned}$$

Model parameters are $\alpha \in (1, \infty)$, $\beta \in (1, \infty)$, $\gamma \in (0, 1)$, $\tau^2 \in (0, \infty)$. Let us consider the following prior distributions:

$$\begin{aligned} \alpha &\sim N(0, \sigma_\alpha^2) \\ \beta &\sim N(0, \sigma_\beta^2) \\ \gamma &\sim Unif(0, 1) \\ \tau^2 &\sim IG(a, b))(InverseGamma) \end{aligned}$$

(a) Derive the functional form (up to proportionality constants) of all *full-conditionals*

(b) Which distribution can you recognize within standard parametric families so that direct simulation from full conditional can be easily implemented ?

(c) Using a suitable Metropolis-within-Gibbs algorithm simulate a Markov chain ($T = 10000$) to approximate the posterior distribution for the above model

(d) Show the 4 univariate trace-plots of the simulations of each parameter

(e) Evaluate graphically the behaviour of the empirical averages $\hat{I}_t$ with growing $t = 1, ..., T$

(f) Provide estimates for each parameter together with the approximation error and explain how you have evaluated such error

(g) Which parameter has the largest posterior uncertainty? How did you measure it?

(h) Which couple of parameters has the largest correlation (in absolute value)?

(i) Use the Markov chain to approximate the posterior predictive distribution of the length of a dugong with age of 20 years.

(j) Provide the prediction of another dugong with age 30

(k) Which prediction is less precise?