A simple gibbs sampling example

lleia@qq.com

August 9, 2014

1 The Change Point Model

This example can be found here. Since its simplicity, we can focus on the most important sampling part of the model instead of being lost in the mathematical details. The quotation below is the problem:

> Suppose that we observe a sequence of counts $x_1, x_2, \cdots, x_N$ where the average of the counts has some value for time steps from 1 to n, and a different value for time steps $n+1$ to N.

We use poisson distributions to model the generating of the sequence and leverage a Γ distribution for its priors.So for the first part of the sequence, we have a parameter λ_1 and every $x_i \sim poisson(\lambda_1)$ and so are the remains. Finally the model has been crafted:

$$n \sim Uniform(1, 2, \cdots, N) \tag{1}$$

$$\lambda_i \sim \Gamma(\lambda; a, b) \tag{2}$$

$$x_i \sim \begin{cases} Poisson(x; \lambda_1), i \leq n \\ Poisson(x; \lambda_2), i > n \end{cases} \tag{3}$$

$$p(X, n, \lambda_1, \lambda_2) = p(n)p(\lambda_1)p(\lambda_2)p(X_{1,2,\cdots,n}|\lambda_1)p(X_{n+1,n+2,\cdots,N}|\lambda_2) \tag{4}$$

2 Poisson Distribution and Gamma Distribution

Suppose $p(\lambda) \sim \Gamma(\lambda; a, b)$, $p(x|\lambda) \sim Poisson(x; \lambda)$,now we calculate the $p(\lambda|x)$.

$$
\begin{aligned}
p(\lambda|X) &= \frac{p(X|\lambda)p(\lambda)}{p(X)} \\
&\propto p(X|\lambda)p(\lambda) \\
&= \prod_{i=1}^{N} p(x_i|\lambda)p(\lambda) \\
&= \prod_{i=1}^{N} e^{-\lambda}\frac{\lambda^{x_i}}{x_i!}\frac{1}{\Gamma(a)}b^a\lambda^{a-1}e^{-b\lambda} \\
&= e^{-N\lambda}\frac{\lambda^{\sum_{i=1}^{N}x_i}}{\prod_{i=1}^{N}x_i!}\frac{1}{\Gamma(a)}b^a\lambda^{a-1}e^{-b\lambda} \\
&= \frac{b^a}{\Gamma(a)\prod_{n=1}^{N}x_i!}\lambda^{a+\sum_{i=1}^{N}x_i-1}e^{-(b+N)\lambda} \\
&\propto \Gamma(\lambda; a+\sum_{i=1}^{N}x_i, b+N)
\end{aligned} \tag{5}
$$

So the Gamma and Poisson are conjugate distributions.

3 The Gibbs Sampler

We now derive a gibbs sampler for the change point model. Note that we want to kown the posterior:

$$p(n, \lambda_1, \lambda_2|X)$$

We need all the conditional distributions in order to run the gibbs sampler. For $p(n|\lambda_1, \lambda_2, X)$,

$$
\begin{aligned}
p(n|\lambda_1, \lambda_2, X) &= \frac{p(n, \lambda_1, \lambda_2, X)}{p(\lambda_1, \lambda_2, X)} \\
&\propto p(n, \lambda_1, \lambda_2, X) \\
&= p(n)p(\lambda_1)p(\lambda_2)p(X_{1,2,\cdots,n}|\lambda_1)p(X_{n+1,n+2,\cdots,N}|\lambda_2) \\
&\propto p(X_{1,2,\cdots,n}|\lambda_1)p(X_{n+1,n+2,\cdots,N}|\lambda_2) \\
&= \prod_{i=1}^{n} e^{-\lambda_1}\frac{\lambda_1^{x_i}}{x_i!}\prod_{i=n+1}^{N} e^{-\lambda_2}\frac{\lambda_2^{x_i}}{x_i!} \\
&\propto \prod_{i=1}^{n} e^{-\lambda_1}\lambda_1^{x_i}\prod_{i=n+1}^{N} e^{-\lambda_2}\lambda_2^{x_i} \\
&= exp(-n\lambda_1-(N-n)\lambda_2)\lambda_1^{\sum_{i=1}^{n}x_i}\lambda_2^{\sum_{i=n+1}^{N}x_i} \\
&= exp(\sum_{i=1}^{n} log\lambda_1 + \sum_{i=n+1}^{N} log\lambda_2 - n\lambda_1 - (N-n)\lambda_2)
\end{aligned} \tag{6}
$$

So we sample n from a multinomial distribution which depends on λ_1, λ_2. Since $p(\lambda_1|\lambda_2, n, X) \propto p(\lambda_1|X_{1,2,\cdots,n})$,just as in section 2, we have:

$$\begin{aligned} p(\lambda_1|\lambda_2, n, X) &\propto p(\lambda_1, \lambda_2, n, X) \\ &\propto p(\lambda_1|X_{1,2,\cdots,n}) \\ &= \Gamma(\lambda_1; a + \sum_{i=1}^{n}, b + n) \end{aligned} \tag{7}$$

λ_2 has the same story:

$$p(\lambda_2|\lambda_1, n, X) \propto \Gamma(\lambda_2; a + \sum_{i=n+1}^{N}, b + (N - n)) \tag{8}$$

Put all these things together, we have the gibbs sampler for our problem as below:

Algorithm 1: Gibbs Sampler

Data: observed sequence $x_1, x_2, \cdots, x_n$
Result: n, λ_1, λ_2
$i = 1$;
while the sampler has not converged do
    sample $n_i \sim multinomial$ as in formula 6;
    sample $\lambda_{i1} \sim \Gamma(a + \sum_{i=1}^{n} x_i, b + n_i)$;
    sample $\lambda_{i2} \sim \Gamma(a + \sum_{i=n+1}^{N} x_i, b + N - n_i)$;
    $i = i + 1$;
end

---

4 Experiments

Here is the r code. We use two binomial distributions (with 10 and 30 as expection respectively) to generate a sequence of length 100. After 10000 iterations,we get results as below,as we can see from the pictures, the sampling results work well to fit our test sequence.

Figure 1: Test Sequence

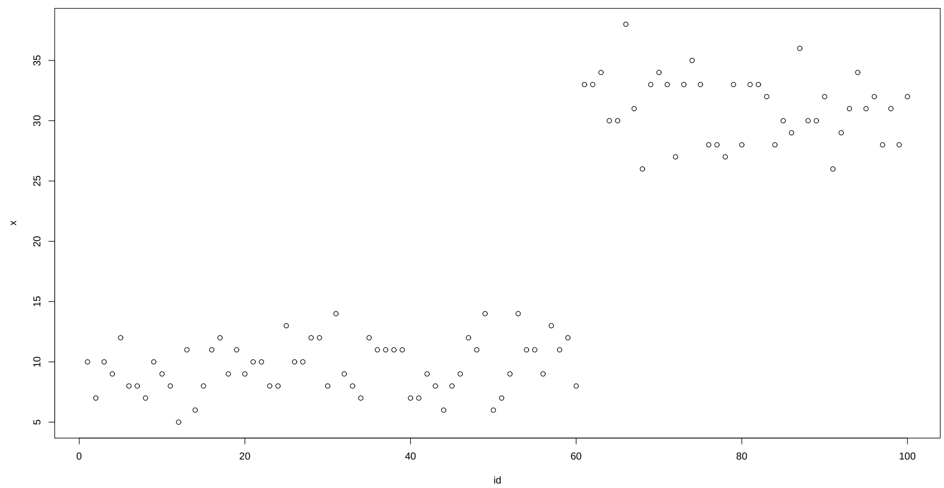

Figure 2: Sampling of n

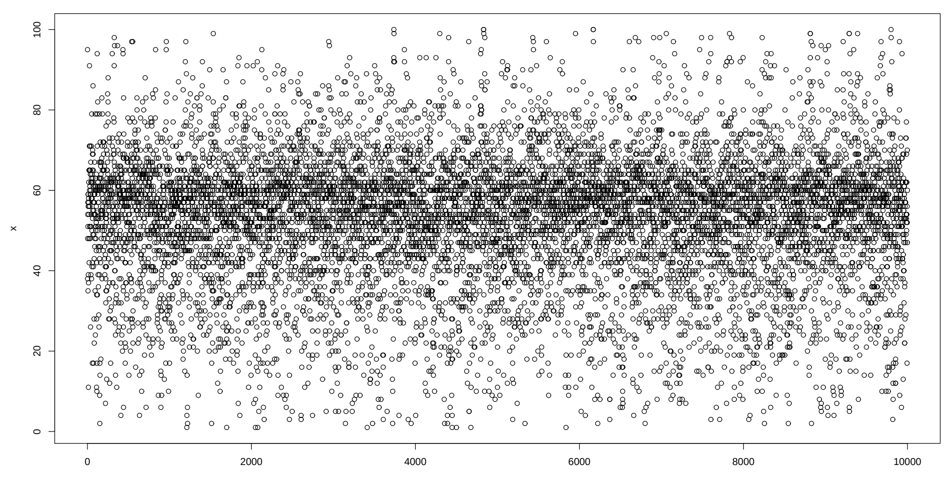

Figure 3: Sampling of λ_1

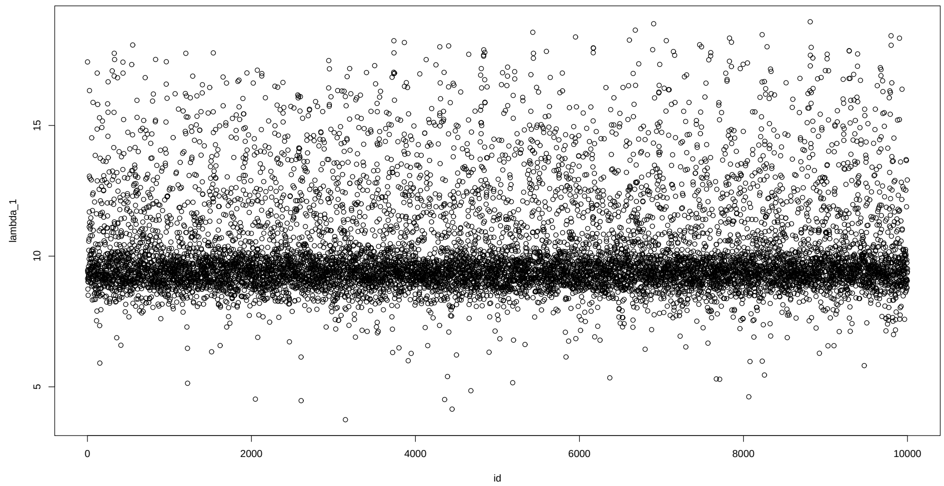

Figure 4: Sampling of λ_2