Supplementary Materials to *Factorized Asymptotic Bayesian Inference for Factorial Hidden Markov Models*

March 18, 2014

Lemma 1. *Suppose y_n is the sum of n independent Bernoulli r.v. $z_1, \cdots, z_n$, i.e. $y_n = \sum_{i=1}^{n} z_i$, and $\bar{y}_n = E[y_n] = \sum_{i=1}^{n} \bar{z}_i$. Suppose further that $\bar{y}_n$ is $O(n)$, i.e., $\exists \epsilon \in (0,1), s.t. \forall n, \bar{y}_n > \epsilon n$.*

Then in the Taylor series of $y_n \log y_n$ expanded around $\bar{y}_n$: 1) the second order term $\frac{1}{2\bar{y}_n}(y_n - \bar{y}_n)^2$ has expectation $\frac{1}{2\bar{y}_n}\sum_{i=1}^{n}(\bar{z}_i - \bar{z}_i^2)$, which is in the interval $[0, 0.5]$; 2) when $n \to \infty$, the expectation of the residual $R_2(\bar{y}_n)$ of the above approximation will tend to 0.

Proof:

1) By straightforward computation, we can prove $\mathrm{E}\Big[\frac{1}{2\bar{y}_n}(y_n - \bar{y}_n)^2\Big] = \frac{1}{2\bar{y}_n}\sum_{i=1}^{n}(\bar{z}_i - \bar{z}_i^2)$. As $\forall i, 0 \le \bar{z}_i^2 \le \bar{z}_i \le 1$, then $\mathrm{E}\Big[\frac{1}{2\bar{y}_n}(y_n - \bar{y}_n)^2\Big] \le \frac{1}{2\bar{y}_n}\sum_{i=1}^{n}\bar{z}_i = 0.5$.

2) The absolute value of the residual

$$\Big|R_2(\bar{y}_n)\Big| = \Big|\sum_{m=3}^{\infty}\frac{(-1)^m}{m!\bar{y}_n^{m-1}}(y_n - \bar{y}_n)^m\Big| < \bar{y}_n\sum_{m=3}^{\infty}\frac{1}{m!}\frac{|y_n - \bar{y}_n|^m}{\bar{y}_n^m}. \tag{1}$$

Therefore

$$\Bigg|\mathrm{E}\Big[R_2(\bar{y}_n)\Big]\Bigg| \le \mathrm{E}\Bigg[\Big|R_2(\bar{y}_n)\Big|\Bigg] < \bar{y}_n\sum_{m=3}^{\infty}\frac{1}{m!}\frac{\mathrm{E}[|y_n - \bar{y}_n|^m]}{\bar{y}_n^m}. \tag{2}$$

Inside each term of the summation,

$$\begin{aligned}
&\mathrm{E}[|y_n - \bar{y}_n|^m] \\
=& \sum_{r_1+\cdots+r_n=m}\binom{m}{r_1, \cdots, r_n}\prod_{i=1}^{n}\mathrm{E}[(z_i - \bar{z}_i)^{r_i}] \\
\le& \sum_{r_1+\cdots+r_n=m}\binom{m}{r_1, \cdots, r_n}\prod_{i=1}^{n}\Big|\mathrm{E}[(z_i - \bar{z}_i)^{r_i}]\Big| \\
=& \sum_{u=1}^{\lfloor m/2\rfloor}\sum_{\substack{r_1+\cdots+r_u=m \\ \forall r_i \ge 2}}\binom{m}{r_1, \cdots, r_u}\sum_{(a_1,\cdots,a_u)}\prod_{i=1}^{u}\Big|\mathrm{E}[(z_{a_i} - \bar{z}_{a_i})^{r_i}]\Big|,
\end{aligned} \tag{3}$$

where u is the number of non-zero r_i, $(a_1, \cdots, a_u)$ are a tuple of distinct integers in the range $1 \le a_i \le n$, and $\lfloor m/2\rfloor$ is the floor of $m/2$.

The last equality in (**??**) holds since $\forall i, \mathrm{E}[z_i - \bar{z}_i] = 0$, and thus $\prod_{i=1}^{u} \Big|\mathrm{E}[(z_{a_i} - \bar{z}_{a_i})^{r_i}]\Big| > 0$ only when all $r_i \geq 2$. Consequently, $u \leq \lfloor m/2 \rfloor$.

For any i, $\Big|\mathrm{E}[(z_i - \bar{z}_i)^{r_i}]\Big| = \Big|(1-\bar{z}_i)\bar{z}_i^{r_i} + (-1)^{r_i}(1-\bar{z}_i)^{r_i}\bar{z}_i\Big| \leq (1-\bar{z}_i)\bar{z}_i^2 + (1-\bar{z}_i)^2\bar{z}_i = \bar{z}_i(1-\bar{z}_i) < 1$. Further, the number of all possible tuples $(a_1, \cdots, a_u)$ is $n \cdots (n-u+1)$. Thus,

$$
\begin{aligned}
&\mathrm{E}[|y_n - \bar{y}_n|^m] \\
&< \sum_{u=1}^{\lfloor m/2 \rfloor} \sum_{\substack{r_1+\cdots+r_u=m \\ \forall r_i \geq 2}} \binom{m}{r_1, \cdots, r_u} \frac{1}{u!} n \cdots (n-u+1) \\
&< n^{\lfloor m/2 \rfloor} \sum_{u=1}^{m} \sum_{\substack{r_1+\cdots+r_u=m \\ \forall r_i \geq 1}} \binom{m}{r_1, \cdots, r_u} \frac{1}{u!} \\
&< n^{\lfloor m/2 \rfloor} \sum_{u=1}^{m} \sum_{\substack{r_1+\cdots+r_u=m \\ \forall r_i \geq 1}} \binom{m}{r_1, \cdots, r_u} && (4) \\
&< n^{\lfloor m/2 \rfloor} \sum_{\substack{r_1+\cdots+r_m=m \\ \forall r_i \geq 0}} \binom{m}{r_1, \cdots, r_m} && (5) \\
&= n^{\lfloor m/2 \rfloor} m^m. && (6)
\end{aligned}
$$

The inequality from (4) to (5) is because each solution of the equation $r_1 + \cdots + r_u = m, 1 \leq u \leq m, \forall r_i \geq 1$, is a solution of $r_1 + \cdots + r_m = m, \forall r_i \geq 0$, by simply letting $r_{u+1} = \cdots = r_m = 0$.

Inside (5), $\binom{m}{r_1, \cdots, r_u}$ counts the number of ways of splitting m different balls into m groups, each having r_i balls. When the tuple $(r_1, \cdots, r_m)$ goes through all the solutions of $r_1 + \cdots + r_m = m$, the sum is the number of ways of splitting m different balls into m groups arbitrarily. Obviously the count is m^m.

It is known that

$$
m! > m^m e^{-m}, \forall m \in \mathcal{N}. \tag{7}
$$

Applying (6) and (7) onto (2), we obtain

$$
\begin{aligned}
&\Big|\mathrm{E}\Big[R_2(\bar{y}_n)\Big]\Big| < \bar{y}_n \sum_{m=3}^{\infty} \frac{1}{m!} \frac{n^{\lfloor m/2 \rfloor} m^m}{\bar{y}_n^m} \\
&< \bar{y}_n \sum_{m=3}^{\infty} \frac{n^{\lfloor m/2 \rfloor} e^m}{\bar{y}_n^m} < \bar{y}_n \sum_{m=3}^{\infty} \frac{n^{\lfloor m/2 \rfloor} e^m}{\epsilon^m n^m} \\
&= \bar{y}_n \sum_{m=3}^{\infty} \frac{e^m}{\epsilon^m n^{\lceil m/2 \rceil}} < n\Big(\frac{e^3}{\epsilon^3 n^2} + \sum_{m=4}^{\infty} (\frac{e}{\epsilon\sqrt{n}})^m\Big) \\
&= \frac{e^3}{\epsilon^3 n} + \frac{e^4}{\epsilon^4 n - \epsilon^3 e\sqrt{n}}. && (8)
\end{aligned}
$$

So when $n \to \infty$, $\Big|\mathrm{E}\Big[R_2(\bar{y}_n)\Big]\Big| \to 0$.