Solutions for the exam in Statistical Reasoning

Date: Friday, November 11, 2016
Time: 09.00-12.00
Place: 5419.0013 (Landleven 12 (Kapteynborg))
Progress code: WISR-11

1. **Posterior Distribution of Binomial-Beta Model.** $\boxed{\mathbf{15}}$

   Compute the density of the joint distribution:

$$p(y_1,\ldots,y_n|\theta,N) = \prod_{i=1}^{n} p(y_i|\theta,N) = \prod_{i=1}^{n} \binom{N}{y_i} \cdot \theta^{y_i} \cdot (1-\theta)^{N-y_i} \quad \boxed{2/15}$$

$$= \left(\prod_{i=1}^{n} \binom{N}{y_i}\right) \cdot \left(\theta^{\sum_{i=1}^{n} y_i} \cdot (1-\theta)^{n\cdot N - \sum_{i=1}^{n} y_i}\right) \quad \boxed{5/15}$$

   Compute the posterior distribution of θ:

$$p(\theta|y_1,\ldots,y_n) \propto p(y_1,\ldots,y_n|\theta,N) \cdot p(\theta) \quad \boxed{6/15}$$

$$= \left(\prod_{i=1}^{n} \binom{N}{y_i}\right) \cdot \left(\theta^{\sum_{i=1}^{n} y_i} \cdot (1-\theta)^{n\cdot N - \sum_{i=1}^{n} y_i}\right) \cdot \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \cdot \theta^{a-1} \cdot (1-\theta)^{b-1} \quad \boxed{8/15}$$

$$\propto \theta^{a+(\sum_{i=1}^{n} y_i)-1} \cdot (1-\theta)^{b+n\cdot N-(\sum_{i=1}^{n} y_i)-1} \quad \boxed{10/15}$$

   From the last line it follows that the PDF of the posterior is proportional to the PDF of a Beta distribution with $\tilde{a} = a + \sum_{i=1}^{n} y_i$ and $\tilde{b} = b + n \cdot N - \sum_{i=1}^{n} y_i$.

$$\theta|(Y_1 = y_1,\ldots,Y_n = y_n) \sim \text{Beta}(a + \sum_{i=1}^{n} y_i, b + n \cdot N - \sum_{i=1}^{n} y_i) \quad \boxed{15/15}$$

2. **Marginal Likelihood of Poisson-Exponential Model.** $\boxed{\mathbf{15}}$
   For the PDF of the marginal distribution we have:

$$p(y) = \int p(y,\theta)d\theta = \int p(y|\theta)p(\theta)d\theta \quad \boxed{3/15}$$

$$= \int \frac{\theta^y \cdot e^{-\theta}}{y!} \cdot \lambda \cdot e^{-\lambda\theta} d\theta = \frac{\lambda}{y!} \cdot \int \theta^y \cdot e^{-(1+\lambda)\theta} d\theta \quad \boxed{5/15}$$

$$= \frac{\lambda}{y!} \cdot \frac{\Gamma(y+1)}{(1+\lambda)^{y+1}} \cdot \int \frac{(1+\lambda)^{y+1}}{\Gamma(y+1)} \cdot \theta^y \cdot e^{-(1+\lambda)\theta} d\theta \quad \boxed{10/15}$$

$$= \frac{\lambda}{y!} \cdot \frac{\Gamma(y+1)}{(1+\lambda)^{y+1}} \cdot 1 = \frac{\lambda}{(1+\lambda)^{y+1}} = \left(\frac{1}{1+\lambda}\right)^y \cdot \frac{\lambda}{1+\lambda} \quad \boxed{15/15}$$

   This is the PDF of a geometric distribution with parameter $p := \frac{\lambda}{1+\lambda}$.

3. **Predictive Distribution of Multinomial-Dirichlet Model.** **15**

(a) PDF of posterior distribution:

$$p(\theta_1, \ldots, \theta_K | n_1, \ldots, n_K) = \frac{\Gamma(\sum_{k=1}^K \alpha_k + n_k)}{\prod_{k=1}^K \Gamma(\alpha_k + n_k)} \cdot \prod_{k=1}^K \theta_k^{\alpha_k + n_k - 1} \quad \boxed{5/5}$$

(b) For such a new observation (see exercise) the PDF is:

$$p(\tilde{n}_1, \ldots, \tilde{n}_K | \theta_1, \ldots, \theta_K) = \frac{1!}{0! \cdot \ldots 0! \cdot 1! \cdot 0! \cdot \ldots \cdot 0!} \cdot \prod_{k=1}^K \theta_k^{\tilde{n}_k} = \prod_{k=1}^K \theta_k^{\tilde{n}_k} \quad \boxed{2/10}$$

Compute the PDF of the corresponding predictive distribution:

$$p(\tilde{n}_1, \ldots, \tilde{n}_K | n_1, \ldots, n_K)$$

$$= \int p(\tilde{n}_1, \ldots, \tilde{n}_K | \theta_1, \ldots, \theta_K) \cdot p(\theta_1, \ldots, \theta_K | n_1, \ldots, n_K) d\vec{\theta} \quad \boxed{3/10}$$

where $\vec{\theta} = (\theta_1, \ldots, \theta_K)^T$

$$= \int \left(\prod_{k=1}^K \theta_k^{\tilde{n}_k} \right) \cdot \frac{\Gamma(\sum_{k=1}^K (\alpha_k + n_k))}{\prod_{k=1}^K \Gamma(\alpha_k + n_k)} \cdot \prod_{k=1}^K \theta_k^{\alpha_k + n_k - 1} d\vec{\theta} \quad \boxed{4/10}$$

$$= \frac{\Gamma(\sum_{k=1}^K (\alpha_k + n_k))}{\prod_{k=1}^K \Gamma(\alpha_k + n_k)} \cdot \int \left(\prod_{k=1}^K \theta_k^{\tilde{n}_k} \right) \cdot \prod_{k=1}^K \theta_k^{\alpha_k + n_k - 1} d\vec{\theta} \quad \boxed{5/10}$$

$$= \frac{\Gamma(\sum_{k=1}^K (\alpha_k + n_k))}{\prod_{k=1}^K \Gamma(\alpha_k + n_k)} \cdot \int \prod_{k=1}^K \theta_k^{\alpha_k + n_k + \tilde{n}_k - 1} d\vec{\theta} \quad \boxed{6/10}$$

$$= \frac{\Gamma(\sum_{k=1}^K (\alpha_k + n_k))}{\prod_{k=1}^K \Gamma(\alpha_k + n_k)} \cdot \frac{\prod_{k=1}^K \Gamma(\alpha_k + n_k + \tilde{n}_k)}{\Gamma(\sum_{k=1}^K (\alpha_k + n_k + \tilde{n}_k))} \cdot 1 \quad \boxed{7/10}$$

Since $\tilde{n}_k = 0$ for all $k \neq j$ and $\tilde{n}_j = 1$:

$$= \frac{\Gamma(\sum_{k=1}^K (\alpha_k + n_k))}{\Gamma(\alpha_j + n_j)} \cdot \frac{\Gamma(\alpha_j + n_j + 1)}{\Gamma(1 + \sum_{k=1}^K (\alpha_k + n_k))} \quad \boxed{8/10}$$

$$= \frac{\alpha_j + n_j}{\sum_{k=1}^K (\alpha_k + n_k)} \quad \boxed{9/10}$$

With $\alpha := \sum_{k=1}^K \alpha_k$, and $n = \sum_{k=1}^K n_k$ we have:

$$p(\tilde{n}_1, \ldots, \tilde{n}_K | n_1, \ldots, n_K) = \frac{\alpha_j + n_j}{\alpha + n} \quad \boxed{10/10}$$

4. **Gibbs Sampling - Generic Pseudo Code.** **10**

Initialisation: Set $\theta_1^{(0)} = \theta_1 \in \Theta_1$, $\theta_2^{(0)} = \theta_2 \in \Theta_2$, and $\theta_3^{(0)} = \theta_3 \in \Theta_3$. $\boxed{3/10}$

Iterations: For $t = 1, \ldots, T$:

- Sample $\theta_1^{(t)} \sim \text{FCD}(\theta_1|\theta_2^{(t-1)}, \theta_3^{(t-1)}, \text{D})$ $\boxed{5/10}$
- Sample $\theta_2^{(t)} \sim \text{FCD}(\theta_2|\theta_1^{(t)}, \theta_3^{(t-1)}, \text{D})$ $\boxed{7/10}$
- Sample $\theta_3^{(t)} \sim \text{FCD}(\theta_3|\theta_1^{(t)}, \theta_2^{(t)}, \text{D})$ $\boxed{9/10}$

Output: $(\theta_1^{(t)}, \theta_2^{(t)}, \theta_3^{(t)})_{t=0,\dots,T}$. $\boxed{10/10}$

NOTE: Normally we would now have to remove the first samples (e.g. the first 50% of the samples) to take the burn-in phase into account and we would have to thin-out the remaining samples (e.g. by a factor of 10) to avoid auto-correlations.

5. **Hierarchical Bayesian Model - Coupled Variances.** $\boxed{\mathbf{20}}$

No digital solution for the graphical model. $\boxed{5/5}$ $\boxed{\text{-1 point per mistake}}$

(1) Compute the FCD of σ_1^2:

$$p(\sigma_1^{-2}|\sigma_2^{-2}, b, y_1, \dots, y_{n_1}, z_1, \dots, z_{n_2}) \propto$$

$$p(y_1, \dots, y_{n_1}|\sigma_1^2) \cdot p(z_1, \dots, z_{n_2}|\sigma_2^2) \cdot p(\sigma_1^{-2}|b) \cdot p(\sigma_2^{-2}|b) \cdot p(b) \propto p(y_1, \dots, y_{n_1}|\sigma_1^2) \cdot p(\sigma_1^{-2}|b)$$

$$= \prod_{i=1}^{n_1} \left(\frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sigma_1} \cdot \exp\{-0.5 \cdot \frac{(y_i - \mu)^2}{\sigma_1^2}\} \right) \cdot \frac{b^1}{\Gamma(1)} \cdot \left(\sigma_1^{-2}\right)^{1-1} \exp\{-b \cdot \sigma_1^{-2}\} \quad \boxed{1/5}$$

$$= \left(\frac{1}{\sqrt{2\pi}} \right)^{n_1} \cdot \left(\sigma_1^{-2}\right)^{n_1/2} \cdot \exp\{-0.5 \cdot \sum_{i=1}^{n_1} \frac{(y_i - \mu)^2}{\sigma_1^2}\} \cdot b \cdot \exp\{-b \cdot \sigma_1^{-2}\} \quad \boxed{2/5}$$

$$\propto \left(\sigma_1^{-2}\right)^{n_1/2} \cdot \exp\{-0.5 \cdot \sum_{i=1}^{n_1} (y_i - \mu)^2 \cdot \sigma_1^{-2}\} \cdot \exp\{-b \cdot \sigma_1^{-2}\} \quad \boxed{3/5}$$

$$\propto \left(\sigma_1^{-2}\right)^{(n_1/2+1)-1} \cdot \exp\{-\sigma_1^{-2} \left(b + 0.5 \cdot \sum_{i=1}^{n_1} (y_i - \mu)^2 \right)\} \quad \boxed{4/5}$$

From the last line it can be seen that the PDF of the FCD is proportional to the PDF of a Gamma distribution. Thus, the FCD of σ_1^{-2} is a Gamma distribution with parameters $\alpha_1 = n_1/2 + 1$ and $\beta_1 = b + 0.5 \cdot \sum_{i=1}^{n_1} (y_i - \mu)^2$. $\boxed{5/5}$

(2) Compute the FCD of σ_2^2:

$$p(\sigma_2^{-2}|\sigma_1^{-2}, b, y_1, \dots, y_{n_1}, z_1, \dots, z_{n_2}) \propto$$

$$p(y_1, \dots, y_{n_1}|\sigma_1^2) \cdot p(z_1, \dots, z_{n_2}|\sigma_2^2) \cdot p(\sigma_1^{-2}|b) \cdot p(\sigma_2^{-2}|b) \cdot p(b) \propto p(z_1, \dots, z_{n_2}|\sigma_2^2) \cdot p(\sigma_2^{-2}|b)$$

$$= \prod_{i=1}^{n_2} \left(\frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sigma_2} \cdot \exp\{-0.5 \cdot \frac{(z_i - \mu)^2}{\sigma_2^2}\} \right) \cdot \frac{b^1}{\Gamma(1)} \cdot \left(\sigma_2^{-2}\right)^{1-1} \exp\{-b \cdot \sigma_2^{-2}\} \quad \boxed{1/5}$$

and in analogy to the above computation it follows: 2/5 3/5

$$\cdots \propto \left(\sigma_2^{-2}\right)^{(n_2/2+1)-1} \cdot \exp\{-\sigma_2^{-2}\left(b + 0.5 \cdot \sum_{i=1}^{n_2}(z_i - \mu)^2\right)\} \quad \boxed{4/5}$$

so that the FCD of σ_2^{-2} is a Gamma distribution with parameters $\alpha_2 = n_2/2 + 1$ and $\beta_2 = b + 0.5 \cdot \sum_{i=1}^{n_2}(z_i - \mu)^2$. 5/5

(3) Compute the FCD of b:

$$p(b|\sigma_1^{-2}, \sigma_2^{-2}, y_1, \ldots, y_{n_1}, z_1, \ldots, z_{n_2}) \propto$$

$$p(y_1, \ldots, y_{n_1}|\sigma_1^2) \cdot p(z_1, \ldots, z_{n_2}|\sigma_2^2) \cdot p(\sigma_1^{-2}|b) \cdot p(\sigma_2^{-2}|b) \cdot p(b) \propto \cdot p(\sigma_1^{-2}|b) \cdot p(\sigma_2^{-2}|b) \cdot p(b)$$

1/5

$$= \frac{b^1}{\Gamma(1)}\left(\sigma_1^{-2}\right)^{1-1} \exp\{-b \cdot \sigma_1^{-2}\} \cdot \frac{b^1}{\Gamma(1)}\left(\sigma_2^{-2}\right)^{1-1} \exp\{-b \cdot \sigma_2^{-2}\} \cdot \frac{\beta^\alpha}{\Gamma(\alpha)} b^{\alpha-1} \exp\{-\beta \cdot b\}$$

2/5

$$= b \cdot \exp\{-b \cdot \sigma_1^{-2}\} \cdot b \cdot \exp\{-b \cdot \sigma_2^{-2}\} \cdot \frac{\beta^\alpha}{\Gamma(\alpha)} b^{\alpha-1} \exp\{-\beta \cdot b\} \quad \boxed{3/5}$$

$$\propto b^{\alpha+2-1} \cdot \exp\{-b \cdot (\sigma_1^{-2} + \sigma_2^{-2} + \beta)\} \quad \boxed{4/5}$$

From the last line it can be seen that the PDF of the FCD is proportional to the PDF of a Gamma distribution. Thus, the FCD of b is a Gamma distribution with parameters $\alpha_\star = \alpha + 2$ and $\beta_\star = \beta + \sigma_1^{-2} + \sigma_2^{-2}$. 5/5

6. **Metropolis-Hastings Sampling and Monte Carlo Approximations.** **15**

(a)

$$A\left((\theta_1, \theta_2), (\theta_1^\star, \theta_2)\right) = \min\{1, \frac{p(y_1, \ldots, y_n|\theta_1^\star)}{p(y_1, \ldots, y_n|\theta_1)} \cdot \frac{p(\theta_1^\star|\theta_2)}{p(\theta_1|\theta_2)} \cdot \frac{p(\theta_2)}{p(\theta_2)} \cdot \text{HR}\} \quad \boxed{1/5}$$

$$= \min\{1, \frac{p(y_1, \ldots, y_n|\theta_1^\star)}{p(y_1, \ldots, y_n|\theta_1)} \cdot \frac{p(\theta_1^\star|\theta_2)}{p(\theta_1|\theta_2)} \cdot 1 \cdot \frac{p(\theta_1|\theta_2)}{p(\theta_1^\star|\theta_2)}\} \quad \boxed{3/5}$$

$$= \min\{1, \frac{p(y_1, \ldots, y_n|\theta_1^\star)}{p(y_1, \ldots, y_n|\theta_1)}\} = \min\{1, \frac{2}{3}\} = \frac{2}{3} \quad \boxed{5/5}$$

(b)

$$A\left((\theta_1, \theta_2), (\theta_1, \theta_2^\star)\right) = \min\{1, \frac{p(y_1, \ldots, y_n|\theta_1)}{p(y_1, \ldots, y_n|\theta_1)} \cdot \frac{p(\theta_1|\theta_2^\star)}{p(\theta_1|\theta_2)} \cdot \frac{p(\theta_2^\star)}{p(\theta_2)} \cdot \text{HR}\} \quad \boxed{1/5}$$

$$= \min\{1, 1 \cdot \frac{p(\theta_1|\theta_2^\star)}{p(\theta_1|\theta_2)} \cdot 1 \cdot \frac{0.2}{0.2}\} \quad \boxed{3/5}$$

$$= \min\{1, \frac{p(\theta_1|\theta_2^\star)}{p(\theta_1|\theta_2)}\} = \min\{1, \frac{0.4}{0.5}\} = \frac{8}{10} = 0.8 \quad \boxed{5/5}$$

(c) $p(\tilde{y}|y_1, \ldots, y_n)$ can be approximated by $\frac{1}{T}\sum_{t=1}^{T} p(\tilde{y}|\theta_1^{(t)})$ 5/5

END