1) а) $p(X,m,\sigma^2|\alpha,\beta,m_0,\sigma_0^2) = \prod_{i=1}^{n} p(x_i|m,\sigma^2)\, p(m|m_0,\sigma^2)\, p(\sigma^2|\alpha,\beta) =$

$= \prod_{i=1}^{n} N(x_i|m,\sigma^2)\, N(m|m_0,\sigma_0^2)\, \Gamma(\frac{1}{\sigma^2}|\alpha,\beta)$

б) $p(m,\sigma^2|X,\alpha,\beta,m_0,\sigma_0) \propto p(X|m,\sigma^2)\, p(m|m_0,\sigma_0^2)\, p(\sigma|\alpha,\beta) =$

$= \prod_{i=1}^{n} N(x_i|m,\sigma^2)\, N(m|m_0,\sigma_0^2)\, \Gamma(\frac{1}{\sigma^2}|\alpha,\beta) \propto (\frac{1}{\sigma^2})^{n/2} \exp(-\frac{1}{2\sigma^2}\cdot$

$\sum_{i=1}^{n}(x_i-m)^2) \exp(-\frac{1}{2\sigma_0^2}(m-m_0)^2) (\frac{1}{\sigma^2})^{\alpha-1} \exp(-\beta\frac{1}{\sigma^2}) = (\frac{1}{\sigma^2})^{\frac{n}{2}+\alpha-1}\cdot$

$\exp(-\frac{n}{2\sigma^2}(S^2+\bar{x}^2-2m\bar{x}+m^2) - \frac{1}{2\sigma_0^2}(m^2-2mm_0+m_0^2) - \beta\frac{1}{\sigma^2}) \propto$

$\propto (\frac{1}{\sigma^2})^{\frac{n}{2}+\alpha-1} \exp(-\frac{n}{2\sigma^2}(S^2+\bar{x}^2-2m\bar{x}+m^2+\frac{2}{n}\beta) - \frac{1}{2\sigma_0}(m^2-$

$-2mm_0)) =$

Нормально-гамма распределение: $\propto \tau^{\alpha-\frac{1}{2}} \exp(-\tau(\frac{\lambda(x-\mu)^2}{2}+\beta))$

Следовательно $p(m,\sigma^2|X,\alpha,\beta,m_0,\sigma_0)$ не принадлежит нормально-гамма распределению, т.к. есть член: $\exp(-\frac{1}{2\sigma_0}(m^2-2mm_0))$.

в) Да, является

$= [\alpha' = \frac{n}{2}+\alpha] = (\frac{1}{\sigma^2})^{\alpha'-1} \exp(-\frac{1}{2}((\frac{n}{\sigma^2}+\frac{1}{\sigma_0^2})m^2 - 2(\frac{n\bar{x}}{\sigma^2}+\frac{m_0}{\sigma_0^2})m) -$

$-\frac{n}{2\sigma^2}(S^2+\bar{x}^2+\frac{2}{n}\beta)) = [\beta' = \beta+\frac{n}{2}S^2+\frac{n}{2}\bar{x}^2,\ \sigma'^2 = (\frac{n}{\sigma^2}+\frac{1}{\sigma_0^2})^{-1},$

$\mu = (\frac{n\bar{x}}{\sigma^2}+\frac{m_0}{\sigma_0^2})(\frac{n}{\sigma^2}+\frac{1}{\sigma_0^2})] = \Gamma(\frac{1}{\sigma^2}|\frac{n}{2}+\alpha, \beta+\frac{n}{2}S^2+\frac{n}{2}\bar{x}^2)\cdot$

$N(m|(\frac{n\bar{x}}{\sigma^2}+\frac{m_0}{\sigma_0^2})(\frac{n}{\sigma^2}+\frac{1}{\sigma_0^2})^{-1}, (\frac{n}{\sigma^2}+\frac{1}{\sigma_0^2})^{-1})$

2)

② а) $p(t|\lambda) = \prod_{i=1}^{n} p(t_i|\lambda) = \prod_{i=1}^{n} \lambda \exp(-\lambda t_i) = \lambda^n \exp(-\lambda \sum_{i=1}^{n} t_i) =$

$= \underline{\lambda^n \exp(-\lambda n \bar{t})}$

Пусть $p(\lambda|\alpha,\beta) = \Gamma(\lambda|\alpha,\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} \exp(-\lambda\beta)$

$p(\lambda|t,\alpha,\beta) \propto p(t|\lambda)p(\lambda|\alpha,\beta) \propto \lambda^{(\alpha+n)-1} \exp(-\lambda(\beta+n\bar{t})) =$

$= \Gamma(\alpha|\alpha+n, \beta+n\bar{t})$

Тогда $\underline{p(\lambda|\alpha,\beta) = \Gamma(\lambda|\alpha,\beta)}$

С таким правдоподобием сопряжено <u>семейство гамма-распределений</u>

б) $p(t|\alpha,\beta) = \int p(t,\lambda|\alpha,\beta)d\lambda = \int p(t|\lambda)p(\lambda|\alpha,\beta)d\lambda =$

$= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^{+\infty} \lambda^{n+\alpha-1} \exp(-\lambda(\beta+n\bar{t}))d\lambda = \frac{\beta^\alpha}{\Gamma(\alpha)} \cdot \frac{\Gamma(n+\alpha)}{(\beta+n\bar{t})^{n+\alpha}} \cdot$

$\int_0^{+\infty} \frac{(\beta+n\bar{t})^{n+\alpha}}{\Gamma(n+\alpha)} \lambda^{n+\alpha-1} \exp(-\lambda(\beta+n\bar{t}))d\lambda = \frac{\Gamma(n+\alpha)}{\Gamma(\alpha)} \cdot \frac{\beta^\alpha}{(\beta+n\bar{t})^{n+\alpha}}$

$\frac{\partial p(t|\alpha,\beta)}{\partial\beta} = \frac{\Gamma(n+\alpha)}{\Gamma(\alpha)} \left(\frac{\alpha\beta^{\alpha-1}}{(\beta+n\bar{t})^{n+\alpha}} - \frac{(n+\alpha)\beta^\alpha}{(\beta+n\bar{t})^{n+1+\alpha}} \right) =$

$= \frac{\Gamma(n+\alpha)}{\Gamma(\alpha)} \cdot \frac{\beta^{\alpha-1}}{(\beta+n\bar{t})^{n+\alpha+1}} \left(\alpha(\beta+n\bar{t}) - (n+\alpha)\beta \right) =$

$= \frac{\Gamma(n+\alpha)}{\Gamma(\alpha)} \cdot \frac{\beta^{\alpha-1}}{(\beta+n\bar{t})^{n+\alpha+1}} n(\alpha\bar{t} - \beta) = 0 \Rightarrow \hat{\beta} = \hat{\alpha}\bar{t}$

$\frac{\partial p(t|\alpha,\beta)}{\partial\alpha} = \frac{1}{(\beta+n\bar{t})^n} \Big(\frac{\Gamma'(n+\alpha)\Gamma(\alpha) - \Gamma(n+\alpha)\Gamma'(\alpha)}{\Gamma^2(\alpha)} \cdot \exp(\alpha(\ln\beta -$

$- \ln(\beta+n\bar{t}))) + \frac{\Gamma(n+\alpha)}{\Gamma(\alpha)} \cdot (\ln\beta - \ln(\beta+n\bar{t}))\exp(\alpha(\ln\beta -$

$- \ln(\beta+n\bar{t})))\Big) = \frac{1}{\Gamma(\alpha)} \cdot \frac{\beta^\alpha}{(\beta+n\bar{t})^{n+\alpha}} \Big(\Gamma'(n+\alpha) - \frac{\Gamma(n+\alpha)\Gamma'(\alpha)}{\Gamma(\alpha)} +$

$+ \Gamma(n+\alpha)(\ln\beta - \ln(\beta+n\bar{t}))\Big) = \frac{\Gamma(n+\alpha)}{\Gamma(\alpha)} \cdot \frac{\beta^\alpha}{(\beta+n\bar{t})^{n+\alpha}} \Big(\psi(n+\alpha) -$

$\psi(\alpha) + \ln\beta - \ln(\beta+n\bar{t})\Big) = 0 \Rightarrow$

$$\Rightarrow \frac{1}{\hat{\alpha}+(n-1)} + \ldots + \frac{1}{\hat{\alpha}} = \ln\left(1 + \frac{n\bar{t}}{\hat{\beta}}\right)$$

т.е. $$\begin{cases} \hat{\beta} = \hat{\alpha}\bar{t} \\ \frac{1}{\hat{\alpha}+(n-1)} + \ldots + \frac{1}{\hat{\alpha}} = \ln\left(1 + \frac{\bar{t}}{\hat{\alpha}}\right) \end{cases}$$

$$\underline{p(t|\alpha,\beta) = \frac{\Gamma(n+\hat{\alpha})}{\Gamma(\hat{\alpha})} \cdot \frac{\hat{\beta}^{\hat{\alpha}}}{(\hat{\beta}+n\bar{t})^{n+\hat{\alpha}}}}$$

В частном случае: $p(\lambda|\alpha,\beta) = \Gamma(\lambda|1,\alpha) = Exp(\alpha)$:

$$p(t|\alpha) = \frac{\bar{t}\,\Gamma(n+1)}{(\bar{t}+n\bar{t})^{n+1}} = \underline{\frac{\Gamma(n+1)}{\bar{t}^{n}(n+1)^{n+1}}}$$

$p(\sigma^2|\alpha,\beta) = \mathcal{N}(m|m_0, \sigma_0^2)\prod^{K}\Gamma(\frac{1}{\sigma^2}|\alpha, \beta)\prod^{N}(\ldots)$

④ Будем считать, что $y \in [-1, 1]$. Результаты будут такими же.

а) $p(y, w | X, A) = \dfrac{\sqrt{\det A}}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2} w^T A w\right) \prod\limits_{i=1}^{m} \sigma(w^T x_i y_i)$

б) Признак незначим, если $Dw_i \to 0 \iff \alpha_i^{-1} \to 0 \iff \alpha_i \to \infty$

в) $\sigma(x) \geqslant \sigma(\xi) \exp\left(-\frac{1}{4\xi}(2\sigma(\xi)-1)(x^2-\xi^2) + \frac{x-\xi}{2}\right)$

$$p(y, w | x, A) \geqslant \frac{\sqrt{\det A}}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2} w^T A w\right) \prod_{i=1}^{m} \sigma(\xi_i) \exp\Big(-\frac{2\sigma(\xi_i)-1}{4\xi_i} \cdot (w^T x_i x_i w - \xi_i^2) + \frac{y_i w^T x_i - \xi_i}{2}\Big) = \Big[A' = A + \sum_{i=1}^{m} \frac{2\sigma(\xi_i)-1}{2\xi_i} x_i x_i^T,$$
$$v = \frac{1}{2}\sum_{i=1}^{m} y_i x_i\Big] = \frac{\sqrt{\det A}}{(2\pi)^{n/2}} \prod_{i=1}^{m} \sigma(\xi_i) \exp\left(\frac{2\sigma(\xi_i)-3}{4}\xi_i\right) \cdot$$
$$\cdot \exp\left(-\frac{1}{2} w^T A' w + w^T v\right) = L(w, A, \xi)$$

г) $p(y|X, A) = \int p(y, w | X, A)\, dw \geqslant \int L(w, A, \xi)\, dw = \hat{L}(A, \xi) \to$

$\to \max\limits_{A,\xi}$

$F(q, A, \xi) = \log \hat{L}(A, \xi) - D_{KL}(q(w) \| p(w|A, \xi)) \to \max\limits_{q, A, \xi}$

а) E-шаг: $F(q, A, \xi) \to \max\limits_{q}$

$q(w) = p(w | A, \xi) \propto L(w, A, \xi) \propto \mathcal{N}(w | v, (A')^{-1})$

б) M-шаг: $\tilde{F}(A, \xi) = \mathbb{E}_{q(w)} \log L(w, A, \xi) \to \max\limits_{A,\xi}$

$$\hat{F}(A, \xi) \propto \mathbb{E}_{q(w)}\Big(\frac{1}{2}\sum_{i=1}^{n} \log \alpha_i + \sum_{i=1}^{m} \log \sigma(\xi_i) + \frac{2\sigma(\xi_i)-3}{4}\xi_i - \frac{1}{2} w^T A' w +$$
$$+ w^T v\Big)$$

$\dfrac{\partial L}{\partial \alpha_i} = \dfrac{1}{2\alpha_i} - \dfrac{1}{2}\mathbb{E}_{q(w)} w_i^T w_i \Rightarrow \alpha_i = (\mathbb{E}_{q(w)} w_i^T w_i)^{-1}$

$$\frac{\partial L}{\partial \xi_i} = \sigma(-\xi_i) + \frac{1}{2}\sigma(\xi_i)\sigma(-\xi_i)\xi_i + \frac{\sigma(\xi_i)}{2} - \frac{3}{4} - \frac{1}{2}\Big(\frac{\sigma(\xi_i)\sigma(-\xi_i)}{\xi_i} -$$
$$- \frac{\sigma(\xi_i)}{\xi_i^2} + \frac{1}{2\xi_i^2}\Big) x_i^T \mathbb{E}_{q(w)} w w^T x_i = 0$$

5 а) $p(X, m, \sigma^2 | \alpha, \beta, m_0, \sigma_0, \bar{\pi}) = p(X | m, \sigma^2, \bar{\pi})\, p(m | m_0, \sigma_0^2)$

$$p(\sigma^2 | \alpha, \beta) = \mathcal{N}(m | m_0, \sigma_0^2) \prod_{i=1}^{K} \Gamma\left(\frac{1}{\sigma_i^2} \Big| \alpha_i, \beta_i\right) \prod_{n=1}^{N} \left(\sum_{i=1}^{K} \pi_i \mathcal{N}(x_n | m, \sigma_i^2) \right)$$

б) $p(m, \sigma^2 | X, \alpha, \beta, m_0, \sigma_0, \bar{\pi}) \propto p(m, \sigma^2, X | \alpha, \beta, m_0, \sigma_0, \bar{\pi}) =$

$= p(X | m, \sigma^2, \alpha, \beta, m_0, \sigma_0, \bar{\pi})\, p(m, \sigma^2 | \alpha, \beta, m_0, \sigma_0, \bar{\pi}) =$

$$= \prod_{n=1}^{N} \left(\sum_{i=1}^{K} \pi_i \mathcal{N}(x_n | m, \sigma_i^2) \right) \mathcal{N}(m | m_0, \sigma_0^2) \prod_{i=1}^{K} \Gamma\left(\frac{1}{\sigma_i^2} \Big| \alpha_i, \beta_i\right)$$

в) $p(X, m, \sigma^2, Z | \alpha, \beta, m_0, \sigma_0, \bar{\pi}) = \prod_{n=1}^{N} \prod_{j=1}^{K} \left(\pi_j \mathcal{N}(x_n | m, \sigma_j^2)\right)^{z_{nj}}$

$$\cdot\, \mathcal{N}(m | m_0, \sigma_0^2) \prod_{i=1}^{K} \Gamma\left(\frac{1}{\sigma_i^2} \Big| \alpha_i, \beta_i\right)$$

⑥ а) $p(t,\vec{\pi},\lambda|\alpha,\beta,\mu) = p(t|\pi,\lambda)\,p(\vec{\pi}|\mu)\,p(\lambda|\alpha,\beta) =$

$= \prod_{n=1}^{N}\left(\sum_{k=1}^{K}\pi_k\lambda_k\exp(-\lambda_k t_n)\right)\mathrm{Dir}(\vec{\pi}|\mu\vec{e})\prod_{k=1}^{K}\Gamma(\lambda_k|\alpha,\beta)$

б) $p(t,\vec{\pi},\lambda,z|\alpha,\beta,\mu) = \prod_{n=1}^{N}\prod_{k=1}^{K}\left(\pi_k\lambda_k\exp(-\lambda_k t_n)\right)^{z_{nk}}\cdot$

$\mathrm{Dir}(\vec{\pi}|\mu\vec{e})\prod_{k=1}^{K}\Gamma(\lambda_k|\alpha,\beta)$

в) $q(\vec{\pi},\lambda,z) = q(\vec{\pi})q(\lambda)q(z)$

E-шаг:

1. $\log q(z) \propto \mathbb{E}_{q/z}\log p(t,\vec{\pi},\lambda,z|\alpha,\beta,\mu) = \mathbb{E}_{q\backslash z}\sum_{n=1}^{N}\sum_{k=1}^{K}$ $z_{nk}(\log\pi_k + \log\lambda_k - \lambda_k t_n) = \sum_{n=1}^{N}\sum_{k=1}^{K}z_{nk}(\mathbb{E}\log\pi_k + \mathbb{E}\log\lambda_k -$ $- t_n\mathbb{E}\lambda_k)$

2. $\log q(\vec{\pi}) \propto \mathbb{E}_{q\backslash\pi}\log p(t,\vec{\pi},\lambda,z|\alpha,\beta,\mu) = \mathbb{E}_{q\backslash\pi}\sum_{n=1}^{N}\Big(\sum_{k=1}^{K}$ $z_{nk}(\log\pi_k + \log\lambda_k - \lambda_k t_n) + (\mu e - 1)\log\pi_k\Big) \propto$

$\propto \sum_{k=1}^{K}\left(\mu e_k - 1 + \sum_{n=1}^{N}\mathbb{E}z_{nk}\right)\log\pi_k \Rightarrow q(\vec{\pi}) \propto \prod_{k=1}^{K}\pi_k^{\mu e_k - 1 + \theta_k} \Rightarrow$

$q(\vec{\pi}) \sim \mathrm{Dir}(\vec{\pi}|\mu\vec{e} + \vec{\theta}),\ \theta_k = \sum_{n=1}^{N}z_{nk}$

3. $\log q(\lambda) \propto \mathbb{E}_{q\backslash\lambda}\log p(t,\pi,\lambda,z|\alpha,\beta,\mu) = \mathbb{E}_{q\backslash\lambda}\sum_{k=1}^{K}\Big(\sum_{n=1}^{N}$ $z_{nk}(\log\pi_k + \log\lambda_k - \lambda_k t_n) + (\alpha-1)\log\lambda_k - \beta\lambda_k)\Big) \propto$

$\propto \sum_{k=1}^{K}\left(\alpha\log\lambda_k - (\beta + \sum_{n=1}^{N}t_n)\lambda_k\right) \Rightarrow q(\lambda) \propto = \prod_{k=1}^{K}\lambda_k^{(\alpha+1)-1}\exp($ $-\lambda_k(\beta + \sum_{n=1}^{N}t_n)) \Rightarrow q(\lambda_k) = \Gamma(\lambda_k|\alpha+1, \beta + \sum_{n=1}^{N}t_n)$

M-шаг:

$\mathbb{E}_{q(\pi,\lambda,z)}\log p(t,\vec{\pi},\lambda,z|\alpha,\beta,\mu) \propto \sum_{k=1}^{K}\mathbb{E}\Big(\sum_{n=1}^{N}z_{nk}(\log\pi_k + \log\lambda_k -$

$$-\lambda_k t_n) + (\mu e_k - 1 + \sum_{n=1}^{N} z_{nk}) \log \pi_k + \sum_{n=1}^{N} (z_{nk} ((\log \lambda_k -$$

$$-\lambda_k t_n) + (\alpha - 1) \log \lambda_k - \beta \lambda_k) = \widehat{F}(\alpha, \beta)$$