8-a) $var(y_{i,j} \mid \mu, \tau^2)$이 $Var(y_{i,j} \mid \theta_j, \sigma^2)$보다 더 크다.

그 이유는 $var(y_{i,j} \mid \theta_j, \sigma^2)$은 fixed group에서 sampling 하는 반면,
$var(y_{i,j} \mid \mu, \tau^2)$은 θ_j를 선택하고 그 안에서 sampling 하기 때문에
그룹간 변동성이 더해지기 때문이다.

8-b) $cov[y_{i1,j}, y_{i2,j} \mid \theta_j, \sigma^2] = 0 \quad (\because IID)$

$cov[y_{i1,j}, y_{i2,j} \mid \mu, \tau] > 0$ (같은 그룹 j에서 sampling 한 애들)

8-c) $var(y_{ij} \mid \theta_i, \sigma^2) = \sigma^2$

$var(\bar{y}_{\cdot,j} \mid \theta_i, \sigma^2) = \sigma^2 / n_j$

$$
\begin{aligned}
Var(y_{ij} \mid \mu, \tau^2) &= var(E[y_{ij} \mid \theta_j, \sigma^2] \mid \mu, \tau^2) + E[var(y_{ij} \mid \theta_j, \sigma^2) \mid \mu, \tau^2] \\
&= var(\theta_j \mid \mu, \tau^2) + E[\sigma^2 \mid \mu, \tau^2] \\
&= \tau^2 + \sigma^2
\end{aligned}
$$

$\Rightarrow$ (a)의 예상과 같이 $var(y_{ij} \mid \mu, \tau^2)$이 더 크다

$$
\begin{aligned}
var(\bar{y}_{\cdot,j} \mid \mu, \tau^2) &= var(E[\bar{y}_{\cdot j} \mid \theta_j, \sigma^2] \mid \mu, \tau^2) + E[var(\bar{y}_{\cdot j} \mid \theta_j, \sigma^2) \mid \mu, \tau^2] \\
&= var(\theta_j \mid \mu, \tau^2) + E[\sigma^2 / n_j \mid \mu, \tau^2] \\
&= \tau^2 + \sigma^2 / n_j
\end{aligned}
$$

$$
\begin{aligned}
cov(y_{i1,j}, y_{i2,j} \mid \theta_j, \sigma^2) &= E[y_{i1j}, y_{i2j} \mid \theta_j, \sigma^2] - E[y_{i1,j} \mid \theta_j, \sigma^2] E[y_{i2,j} \mid \theta_j, \sigma^2] \\
&= 0
\end{aligned}
$$

$$
\begin{aligned}
cov(y_{i1,j}, y_{i2,j} \mid \mu, \tau^2) &= \cancel{E[cov(y_{i1,j}, y_{i2,j} \mid \theta_j, \sigma^2) \mid \mu, \tau^2]} \to 0 \\
&\quad + cov(E[y_{i1,j} \mid \theta_j, \sigma^2], E[y_{i2,j} \mid \theta_j, \sigma^2]) \\
&= cov(\theta_j, \theta_j) = \tau^2
\end{aligned}
$$

$\Rightarrow$ 마찬가지로 (b)에서의 예상과 같이

$cov(y_{i1,j}, y_{i2,j} \mid \theta_j, \sigma^2) = 0$ & $cov(y_{i1,j}, y_{i2,j} \mid \mu, \tau^2) > 0$

8-d) $p(\mu \mid \theta_1 \cdots \theta_m, \sigma^2, \tau^2, y_1 \cdots y_m) = p(\mu \mid \theta_1 \cdots \theta_m, \sigma^2, \tau^2, y_1 \cdots y_m)$

$$
= \frac{p(\mu, \theta_1 \cdots \theta_m, \sigma^2, \tau^2, y_1 \cdots y_m)}{\int p(\mu, \theta_1 \cdots \theta_m, \sigma^2, \tau^2, y_1 \cdots y_m) d\mu} = \frac{p(\mu) p(\tau^2) p(\sigma^2) p(Y \mid \theta_1 \cdots \theta_m, \sigma^2) p(\theta_1 \ldots \theta_m \mid \mu, \tau^2)}{\int p(\mu) p(\tau^2) p(\sigma^2) p(Y \mid \theta_1 \cdots \theta_m, \sigma^2) p(\theta_1 \ldots \theta_m \mid \mu, \tau^2) d\mu}
$$

$$
= \frac{p(\mu) p(\tau^2) p(\sigma^2) p(Y \mid \theta_1 \ldots \theta_m, \sigma^2) p(\theta_1 \cdots \theta_m \mid \mu, \tau^2)}{p(\tau^2) p(\sigma^2) p(Y \mid \theta_1 \ldots \theta_m \mid \sigma^2) \int p(\mu) p(\theta_1 \cdots \theta_m \mid \mu, \tau^2) d\mu}
$$

$= p(\mu \mid \theta_1 \ldots \theta_m, \tau^2)$