5.1 a) $p(\Theta, \alpha \mid y) = p(\alpha) \prod_{i=1}^{n} p(\Theta_i \mid \alpha)\, p(y_i \mid \Theta_i)$

```
data:
    int n
    real y[n]
parameters:
    real a
    real theta
model:
    a ~ p()
    for i in 1...n:
        theta[i] = p(α)
        y[i] ~ p(theta[i])
```

b) $p(\Theta, \mu, \alpha \mid y) = p(\alpha) \prod_{i=1}^{m} p(\mu_i \mid \alpha) \prod_{k=1}^{n} p(\Theta_{ik} \mid \mu_{ik})\, p(y_{ik} \mid \Theta_{ik})$

```
data:
    int n
    int m
    real y[n·m]
parameters:
    real a
    real mu[m]
    real theta[m·n]
model:
    a ~ p()
    for i in 1...m:
        mu[i] = p(a)
        for k in 1...n:
            theta[n·(i-1)+k] = p(μi)
            y[n·(i-1)+k] = p(theta[n·(i-1)+k])
```

c) $p(\gamma, \alpha, \sigma, \mu \mid y, x) = p(\alpha)p(\gamma)\prod_{i=1}^{n} p(\sigma_i \mid x_i, \gamma)\, p(\mu_i \mid x_i, \alpha)\, p(y_i \mid \sigma_i, \mu_i)$

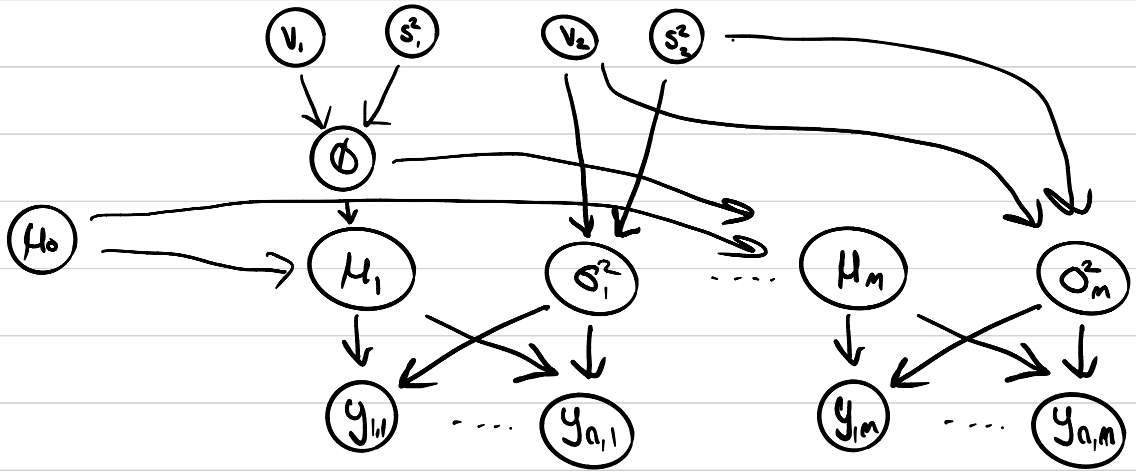
```
data:
    int n
    real y[n]
    real x[n)

parameters:
    real gamma
    real a
    real mu[n]
    real si[n

model:
    gamma ~ p()
    a ~ p()
    for i in 1...n:
        si[i] ~ p(gamma, x[i])
        mu[i] ~ p(a, x[i])
        y[i] ~ p(si[i], mu[i])
```

Values for a and gamma may not actually require sampling from some probability distribution.

d)



```
data:
    int n
    int m
    real y[n*m]
parameters:
    real v_1
    real v_2
    real s_1^2
    real s_2^2
    real mu[m]
    real sigma[m]
    real phi
model:
    phi ~ inv_chi_square(v_1, s_1)
    mu_o ~ normal(0, 10^c)
    for i in 1...m:
        mu[i] ~ normal(mu_o, phi)
        sigma[i] ~ inv_chi_square(v_2, s_2)
        for j in 1...n:
            y[(i-1)n+j] ~ normal(mu[i], sigma[i])
```