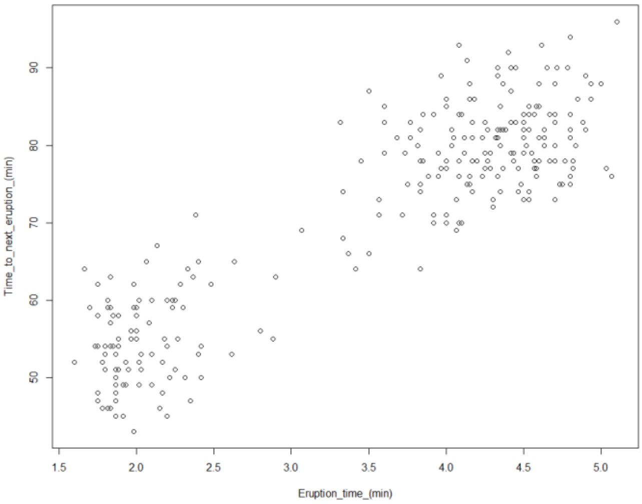
統計模擬作業 2　授課教授:翁久幸　學生:統計碩一　鄭昕東 106354023

Homework 2: (EM for old faithful data)

(1) Dataset:
272 obs. of 2 variables:
"Eruptions" and "Waiting"



We can find that they are roughly divided into two groups from the figure，K=2。

(2) EM for GMM

Let $X = (X_1, \dots, X_{272}) = ((E_1, W_1), \dots, (E_{272}, W_{272})), E_i\text{: Eruption}, W_i\text{: Waiting}$

$Z = (Z_1, \dots, Z_{272}), Z\text{: latent variable}$

And then

$X_i | Z_i = 1 \sim BN(\mu_{11}, \mu_{12}, \sigma_{11}^2, \sigma_{12}^2, \rho_1) \equiv BN(\mu_1, \sum_1)$

Where $\mu_1 = (\mu_{11}, \mu_{12}), \sum_1 = covariance\ \ matrix$

$X_i | Z_i = 2 \sim BN(\mu_{21}, \mu_{22}, \sigma_{21}^2, \sigma_{22}^2, \rho_2) \equiv BN(\mu_2, \sum_2)$

Where $\mu_2 = (\mu_{21}, \mu_{22}), \sum_2 = covariance\ \ matrix$

$\theta = (\mu_1, \sum_1, \mu_2, \sum_2, \alpha)$,where $\alpha = (\alpha_1, \alpha_2), \alpha_1 + \alpha_2 = 1$

(i)E-step: Calculates conditional probabilities given latent variables

Let ϕ $means\ \ pdf\ \ of\ \ bivariate\ \ normal$

$$Q(\theta|\theta^{(t)}) = E_z(logp(x,z|\theta)|x,\theta^{(t)})$$

$$= \sum_{i=1}^{272}\sum_{j=1}^{2}[\log(\alpha_j\phi(x_i|\mu_j,\sigma_j^2))p(z_i=j|x_i,\theta^{(t)})]$$

$$where \quad p(z_i=j|x_i,\theta^{(t)}) = \frac{\alpha_j^{(t)}\phi(x_i,\mu_j^{(t)},(\sigma_j^2)^{(t)})}{\sum_{j=1}^{2}\alpha_j^{(t)}\phi(x_i,\mu_j^{(t)},(\sigma_j^2)^{(t)})}$$

(ii)M-step:Calculates θ to maximize $Q(\theta|\theta^{(t)})$

$$\alpha_j^{(t+1)} = \frac{1}{272}\sum_{i=1}^{272}p(z_i=j|x_i,\theta^{(t)})$$

$$u_j^{(t+1)} = \frac{\sum_{i=1}^{272}x_i p(z_i=j|x_i,\theta^{(t)})}{\sum_{i=1}^{272}p(z_i=j|x_i,\theta^{(t)})}$$

$$\sum\nolimits_j^{(t+1)} = \frac{\sum_{i=1}^{272}(x_i-u_j^{(t)})(x_i-u_j^{(t)})^T \; p(z_i=j|x_i,\theta^{(t)})}{\sum_{i=1}^{272}p(z_i=j|x_i,\theta^{(t)})}$$

We can find that we only need $p(z_i=j|x_i,\theta^{(t)})$ to calculate M-step.
Therefor we just need to calculate $p(z_i=j|x_i,\theta^{(t)})$ in the E-step.

(iii)

Take the value that we get in the M-step returns into E-step and repeat (i),(ii),(iii) until convergence.

Iteration:11

Convergence criteria: $|Q(\theta^{(t)}|\theta^{(t+1)}) - Q(\theta^{(t-1)}|\theta^{(t)})| \le 10^{-6}$

Finally parameters we got:

α	[0.356 ,0.644]
μ_1	[2.036,54.48]
μ_2	[4.289,79.969]
$\sum_1$	$\begin{bmatrix}0.07 & 0.44\\0.44 & 34.06\end{bmatrix}$
$\sum_2$	$\begin{bmatrix}0.17 & 0.94\\0.94 & 36.23\end{bmatrix}$