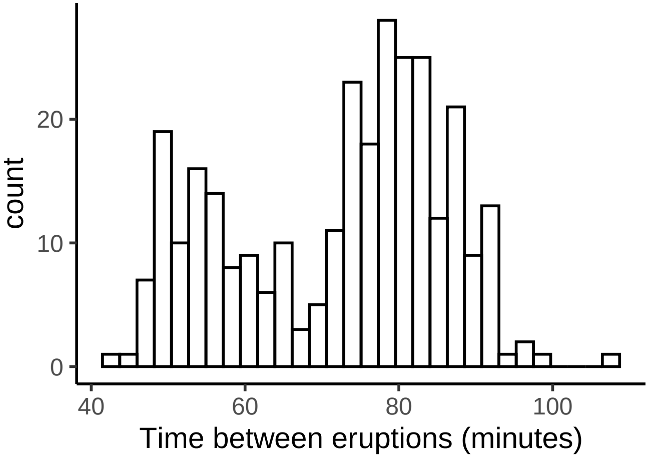
Lecture 30: Fitting Gaussian mixture models

Ciaran Evans

Last time



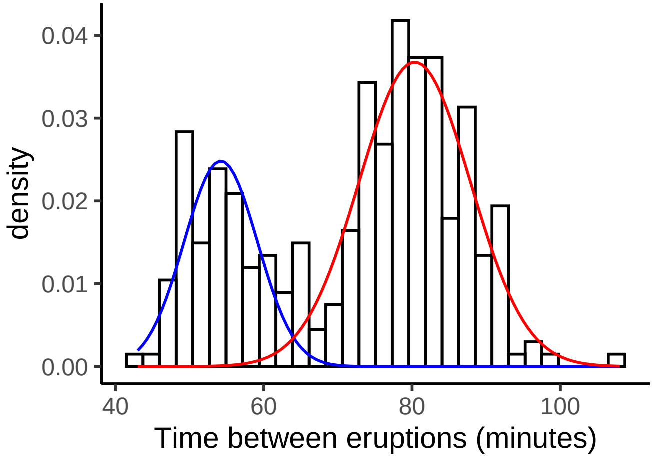
Last time: Gaussian mixture model

- Observe data $X_1, ..., X_n$
- Assume each observation i comes from one of k groups. Let $Z_i \in \{1, ..., k\}$ denote the group assignment
  - The group Z is an unobserved (**latent**) variable

Model:

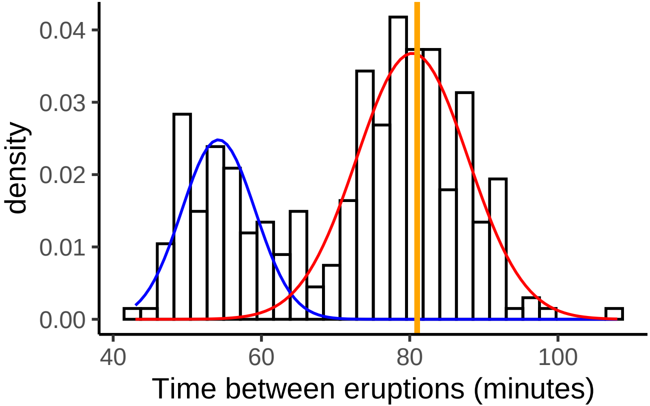
- $P(Z_i = j) = \lambda_j$
- $X_i | (Z_i = j) \sim N(\mu_j, \sigma_j^2)$

Last time: Gaussian mixture model



Group probabilities

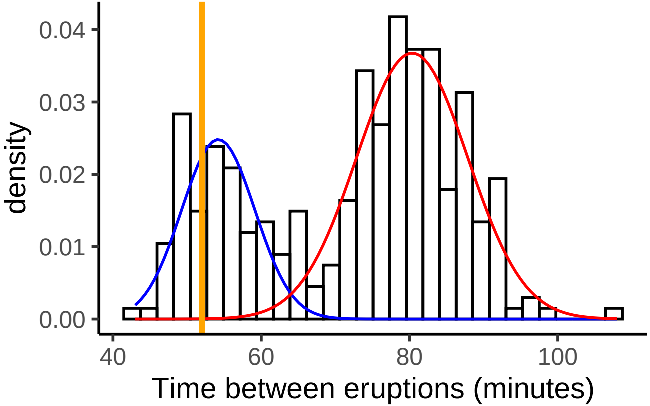
We never get to see the true group labels Z_i. Instead, we estimate the **probability** of belonging to each group.



Question: If $X_i = 81$, do you think it is more likely that $Z_i = 1$ or $Z_i = 2$?

Group probabilities

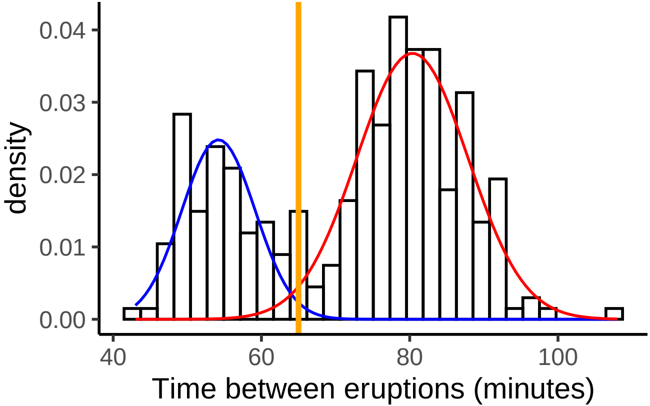
We never get to see the true group labels Z_i. Instead, we estimate the **probability** of belonging to each group.



Question: If $X_i = 52$, do you think it is more likely that $Z_i = 1$ or $Z_i = 2$?

Group probabilities

We never get to see the true group labels Z_i. Instead, we estimate the **probability** of belonging to each group.



Question: If $X_i = 65$, do you think it is more likely that $Z_i = 1$ or $Z_i = 2$?

Gaussian mixture model: posterior probabilities

Model: $P(Z_i = j) = \lambda_j \qquad X_i | (Z_i = j) \sim N(\mu_j, \sigma_j^2)$

Parameters: $\lambda = (\lambda_1, ..., \lambda_k)$, $\mu = (\mu_1, ..., \mu_k)$, $\sigma = (\sigma_1, ..., \sigma_k)$

Given λ, μ, and σ, we would like to calculate the **posterior probability** $P(Z_i = j | X_i)$

Estimating model parameters

- If we know **true** group labels Z_i, it is easy to estimate λ, μ, σ
- We don't have the true Z_i, but we **do** have $P(Z_i = j|X_i)$
- How do we estimate λ, μ, σ using the $P(Z_i = j|X_i)$?

Example: Suppose we observe 6 points:

X	45	50	65	66	80	90
$P(Z = 2\|X)$	1	1	0	0	0	0

Question: What is your estimate $\widehat{\lambda}_2 = \widehat{P}(Z = 2)$?

Estimating model parameters

- If we know **true** group labels Z_i, it is easy to estimate λ, μ, σ
- We don't have the true Z_i, but we **do** have $P(Z_i = j|X_i)$
- How do we estimate λ, μ, σ using the $P(Z_i = j|X_i)$?

Example: Suppose we observe 6 points:

X	45	50	65	66	80	90
$P(Z = 2\|X)$	1	0.75	0.25	0	0	0

Question: What is your estimate $\widehat{\lambda}_2 = \widehat{P}(Z = 2)$?

Estimating model parameters

- If we know **true** group labels Z_i, it is easy to estimate λ, μ, σ
- We don't have the true Z_i, but we **do** have $P(Z_i = j|X_i)$
- How do we estimate λ, μ, σ using the $P(Z_i = j|X_i)$?

Example: Suppose we observe 6 points:

X	45	50	65	66	80	90
$P(Z = 2\|X)$	1	1	0	0	0	0

Question: What is your estimate $\widehat{\mu}_2$?

Estimating model parameters

- If we know **true** group labels Z_i, it is easy to estimate λ, μ, σ
- We don't have the true Z_i, but we **do** have $P(Z_i = j|X_i)$
- How do we estimate λ, μ, σ using the $P(Z_i = j|X_i)$?

Example: Suppose we observe 6 points:

X	45	50	65	66	80	90
$P(Z = 2\|X)$	1	0.75	0.25	0	0	0

Question: What is your estimate $\widehat{\mu}_2$?

Posterior probabilities and parameter estimation

- **If** we know the parameters λ, μ, σ, we can calculate posterior probabilities:

$$P(Z_i = j|X_i) = \frac{\lambda_j f(X_i|Z_i = j)}{\lambda_1 f(X_i|Z_i = 1) + \cdots + \lambda_k f(X_i|Z_i = k)}$$

- **If** we know the posterior probabilities, we can estimate the model parameters λ, μ, and σ:
  - $\widehat{\lambda}_j = \frac{1}{n}\sum_{i=1}^{n} P(Z_i = j|X_i)$
  - $\widehat{\mu}_j = \dfrac{\sum_{i=1}^{n} X_i P(Z_i = j|X_i)}{\sum_{i=1}^{n} P(Z_i = j|X_i)}$
  - $\widehat{\sigma}_j = \sqrt{\dfrac{\sum_{i=1}^{n} (X_i - \widehat{\mu}_j)^2 P(Z_i = j|X_i)}{\sum_{i=1}^{n} P(Z_i = j|X_i)}}$

Putting everything together

Model: $P(Z_i = j) = \lambda_j \qquad X_i | (Z_i = j) \sim N(\mu_j, \sigma_j^2)$

Parameters: $\lambda = (\lambda_1, ..., \lambda_k)$, $\mu = (\mu_1, ..., \mu_k)$, $\sigma = (\sigma_1, ..., \sigma_k)$

Estimation:

1. Initialize parameter guesses $\lambda^{(0)}$, $\mu^{(0)}$, $\sigma^{(0)}$
2. Given current parameter estimates, compute $P^{(0)}(Z_i = j | X_i)$ for all i, j
3. Given current posterior probabilities $P^{(0)}(Z_i = j | X_i)$, update parameter estimates to $\lambda^{(1)}$, $\mu^{(1)}$, $\sigma^{(1)}$
4. Iterate: repeat steps 2–3 until convergence

Your turn

Implement the algorithm to fit a Gaussian mixture model:

https://sta379-s25.github.io/practice_questions/pq_30.html

- Start in class
- Welcome to work with a neighbor
- Solutions will be posted later on the course website