1. Simulate R.V.s

- Inverse CDF method
- Accept - Rejection method
  ↳ generate realisations from a pdf f using another pdf g
  ↳ find c?

$$\frac{f(x)}{g(x)} \leq c \quad \forall x$$

Find the maximum of the f/g using derivatives, i.e.

$$\frac{d}{dx}\left(\frac{f(x)}{g(x)}\right) = 0 \implies x^*.$$

Then verify that it is a maximum with 2nd order derivative

$$\frac{d^2}{dx^2}\left(\frac{f(x)}{g(x)}\right)\bigg|_{x=x^*}.$$

Finally plug the maximum x^* in

$$\frac{f(x^*)}{g(x^*)}.$$

2. Classification

- Given features and observed classifications

$$(X_1, Y_1), \ldots, (X_n, Y_n),$$

find a rule $h: \mathcal{X} \to \mathcal{Y}$ that minimizes a certain criterion.

- Generative or discriminative classifiers.

- Bayes rule with

$$\mathbb{P}(Y=y \mid X=x) = \frac{f_{X|Y=y}(x)\, \mathbb{P}(Y=y)}{\sum_{k=1}^{K} f_{X|Y=k}(x)\, \mathbb{P}(Y=k)}.$$

- Multi-class Logistic Regression with

$$\mathbb{P}(Y=y \mid X=x) = \frac{e^{w_y^T x}}{\sum_{k=1}^{K} e^{w_k^T x}}.$$

- Gaussian Bayes Classifier with

$$X|_{Y=y} \sim N(\mu_y, \Sigma_y).$$

3. Expectation - Maximization Algorithm

- (E-step) estimate latent variables given observations.
- (M-step) maximize likelihood given latent variables.

↳ Example: $K=3$

$$X \mid_{Y=k} \sim N\left(\begin{bmatrix} \mu_{1k} \\ \mu_{2k} \end{bmatrix}, \begin{bmatrix} \Sigma_{11k} & \Sigma_{12k} \\ \Sigma_{12k} & \Sigma_{22k} \end{bmatrix} \right).$$