Mixture distributions

Useful distributions

Francesco Corona

Department of Computer Science
Federal University of Ceará, Fortaleza

Mixture distributions

Useful distributions

Mixture distributions

Suppose that we have k distributions with their respective PDFs

$$f(x_1), f(x_2), \dots, f(x_n)$$

- $\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_n$, the supports
- $\mu_1, \mu_2, \dots, \mu_k$, the means
- $\sigma_1^2, \sigma_2^2, \dots, \sigma_k^2$, the variances

Mixture distributions (cont.)

Let $p_1, p_2, \ldots, p_n$ be positive mixing probabilities, $p_1 + p_2 + \cdots + p_k = 1$

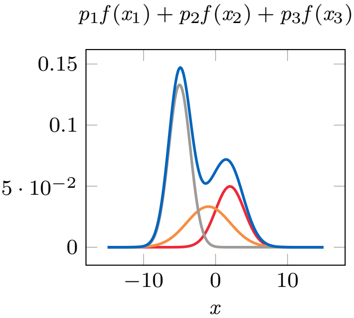
Let $\mathcal{S} = \cup_{i=1}^{k} \mathcal{S}_i$ and consider the function

$$\begin{aligned} f(x) &= p_1 f_1(x) + p_2 f_2(x) + \cdots + p_k f_k(x) \\ &= \sum_{i=1}^{k} p_i f_i(x), \quad x \in \mathcal{S} \end{aligned} \tag{1}$$

$f(x)$ is non-negative and it integrates to one over $(-\infty, \infty)$

$f(x)$ is a PDF for some RV X of the continuous-type

Mixture distributions (cont.)



Mixture distributions (cont.)

The mean of X

$$E(X) = \sum_{i=1}^{k} p_i \int_{-\infty}^{\infty} x f_i(x) \mathrm{d}x = \sum_{i=1}^{k} p_i \mu_i = \bar{\mu} \quad (2)$$

This is a weighted average of $\mu_1, \mu_2, \ldots, \mu_k$

Mixture distributions (cont.)

The variance of X

$$\begin{aligned}
\text{Var}(X) &= \sum_{i=1}^{k} p_i \int_{-\infty}^{\infty} (x-\overline{\mu})^2 f_i(x)\text{d}x \\
&= \sum_{i=1}^{k} p_i \int_{-\infty}^{\infty} \big[(x-\mu_i)+(\mu_i-\overline{\mu}_i)\big]^2 f_i(x)\text{d}x \\
&= \sum_{i=1}^{k} p_i \underbrace{\int_{-\infty}^{\infty} (x-\mu_i)^2 f_i(x)\text{d}x}_{\sigma_i^2} + \sum_{i=1}^{k} p_i(\mu_i-\overline{\mu})^2 \underbrace{\int_{-\infty}^{\infty} f_i(x)\text{d}x}_{1}
\end{aligned}$$

The cross-products terms integrate to zero

Mixture distributions (cont.)

$$\text{Var}(X) = \sum_{i=1}^{k} p_i \sigma_i^2 + \sum_{i=1}^{k} p_i (\mu_i - \overline{\mu})^2 \quad (3)$$

The variance is not simply the weighted average of the k variances

- ⇝ It includes a positive term, a weighted variance of the means

Mixture distributions (cont.)

Remark

These characteristics are associated with mixtures of k distributions

- Not related with linear combinations of k random variables

$$\rightsquigarrow \quad \sum a_i X_i$$

These two are different objects

Mixture distributions (cont.)

The process of mixing distributions is sometimes called **compounding**

$\rightsquigarrow$ It is not restricted to a limited number of distributions

We think of the original distribution of X as a conditional distribution

- It is the conditional distribution of X, given θ

$$\rightsquigarrow \quad f(x|\theta)$$

- The weighting function is treated as a PDF for θ

$$\rightsquigarrow \quad g(\theta)$$

Accordingly, the joint PDF of (X, θ)

$$\rightsquigarrow \quad f(x|\theta)g(\theta)$$

The compound PDF is understood as the marginal PDF of X

$$\rightsquigarrow \quad h(x) = \int_{\theta} g(\theta)f(x|\theta)\mathrm{d}\theta$$

Mixture distributions (cont.)

Example

Contaminated Gaussians

Suppose we are observing a random variable

- ⇝ Most of the times the RV follows a standard normal distribution
- ⇝ At times, the RV follows a normal distribution with larger variance

Some might say that most of the data are good, but there are outliers

We can make the statement more precise

Let $Z \sim N(0,1)$ and let $I_{1-\varepsilon}$ be a discrete random variane

$$I_{1-\varepsilon} = \begin{cases} 1, & \text{with probability } 1-\varepsilon \\ 0, & \text{with probability } \varepsilon \end{cases}$$

Mixture distributions (cont.)

Assume that Z and $I_{1-\varepsilon}$ are independent

Let $W = ZI_{1-\varepsilon} + \sigma_c Z(1 - I_{1-\varepsilon})$

By the independence of Z and $I_{1-\varepsilon}$, the CDF of W

$$
\begin{aligned}
& F_W(w) = \\
& = P(W \leq w) = P(W \leq w, I_{1-\varepsilon} = 1) + P(W \leq w, I_{1-\varepsilon} = 0) \\
& = P(W \leq w | I_{1-\varepsilon} = 1) P(I_{1-\varepsilon} = 1) + P(W \leq w | I_{1-\varepsilon} = 0) P(I_{1-\varepsilon} = 0) \\
& = P(Z \leq w)(1 - \varepsilon) + P(Z \leq w\sigma_c)\varepsilon \\
& = \Phi(w)(1 - \varepsilon) + \Phi(w/\sigma_c)\varepsilon
\end{aligned}
\tag{4}
$$

The distribution of W is a mixture of two normals

- $E(W) = 0$
- $\text{Var}(W) = 1 + \varepsilon(\sigma_c^2 - 1)$

The PDF of W

$$f_W(w) = \phi(w)(1 - \varepsilon) + \phi(w/\sigma_2^c)\frac{\varepsilon}{\sigma_c}$$

Mixture distributions (cont.)

Suppose, in general, that we are interested in the RV $X = a + bW$, $b > 0$

The mean and variance of X

- $E(X) = a$
- $\text{Var}(X) = b^2\big[1 + \varepsilon(\sigma_c^2 - 1)\big]$

The CDF of X

$$F_X(x) = \Phi\Big(\frac{x-a}{b}\Big)(1-\varepsilon) + \Phi\Big(\frac{x-a}{b\sigma_c}\Big) \tag{5}$$

This is a mixture of normal CDFs