M339J: March 10th, 2021.

Two-point mixtures.

We say that a r.v. Y is a **two-point mixture** of random variables X_1 and X_2 if its cdf is of the form:

$$F_Y(y) = a_1 \cdot F_{X_1}(y) + a_2 \cdot F_{X_2}(y)$$

w/ $a_1 > 0$, $a_2 > 0$, $a_1 + a_2 = 1$; constant.

Problem. [Sample STAM, Problem #169.]

The dist'n of a loss X is a two-pt mixture:

(i) w/ probability 0.8, X has a two-parameter Pareto dist'n w/ $\alpha = 2$ and $\theta = 100$;

(ii) w/ probability 0.2, X has a two-parameter Pareto dist'n w/ $\alpha = 4$ and $\theta = 3000$.

Find $\mathbb{P}[X \le 200]$.

$\longrightarrow$: By def'n: $\mathbb{P}[X \le 200] = F_X(200)$

$$X \sim \begin{cases} X_1 \sim \text{Pareto}(\alpha = 2, \theta = 100) & \text{w/ prob. } a_1 = 0.8 \\ X_2 \sim \text{Pareto}(\alpha = 4, \theta = 3000) & \text{w/ prob. } a_2 = 0.2 \end{cases}$$

$$F_X(x) = a_1 \cdot F_{X_1}(x) + a_2 \cdot F_{X_2}(x)$$

$$= 0.8\left(1 - \left(\frac{100}{x+100}\right)^2\right) + 0.2\left(1 - \left(\frac{3000}{x+3000}\right)^4\right)$$

$$= 1 - 0.8\left(\frac{100}{x+100}\right)^2 - 0.2\left(\frac{3000}{x+3000}\right)^4$$

answer: $F_X(200) = 1 - 0.8\left(\frac{100}{200+100}\right)^2 - 0.2\left(\frac{3000}{200+3000}\right)^4$

$$= 1 - \frac{4}{5}\left(\frac{1}{9}\right) - \frac{1}{5}\left(\frac{15}{16}\right)^4 = 0.7566$$

Problem. [Sample STAM Problem #288.]

The r.v. N has the following mixed dist'n:

(i) With probability p, N has a binomial dist'n w/ $q_1 = \frac{1}{2}$ and $m_1 = 2$.

(ii) With probability $1-p$, N has a binomial dist'n w/ $q_2 = \frac{1}{2}$ and $m_2 = 4$.

Find the expression for $\mathbb{P}[N=2]$ in terms of p.

$\rightarrow$:

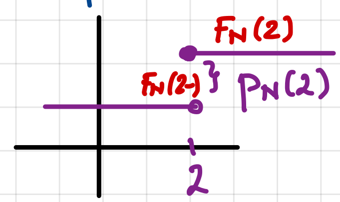
$$N \sim \begin{cases} \boxed{N_1 \sim bin(q_1 = 0.5, m_1 = 2)} & \text{w/ probab. } p \\ \boxed{N_2 \sim bin(q_2 = 0.5, m_2 = 4)} & \text{w/ probab. } 1-p \end{cases}$$

We seek: $\underline{\mathbb{P}[N=2]} = p_N(2) =$

$$= F_N(2) - F_N(2-)$$

$$= F_N(2) - F_N(1)$$

the support are integers



Because N is a 2-pt mixture, we have

$$p_N(2) = p \cdot \boxed{F_{N_1}(2)} + (1-p) \cdot \boxed{F_{N_2}(2)}$$
$$- \left(p \cdot \boxed{F_{N_1}(1)} + (1-p) \cdot \boxed{F_{N_2}(1)}\right)$$

$$p_N(2) = p\underbrace{\left(F_{N_1}(2) - F_{N_1}(1)\right)}_{p_{N_1}(2)} + (1-p)\underbrace{\left(F_{N_2}(2) - F_{N_2}(1)\right)}_{p_{N_2}(2)}$$

$$\boxed{p_N(2) = p \boxed{p_{N_1}(2)} + (1-p) \cdot \boxed{p_{N_2}(2)}}$$

Think about the generalization @ home?

Returning to our problem:

$$p_N(2) = p\left(\frac{1}{4}\right) + (1-p) \cdot \binom{4}{2} \cdot \left(\frac{1}{2}\right)^4$$

$$= p \cdot \frac{1}{4} + (1-p)\frac{\cancel{4}\cdot 3}{2} \cdot \frac{1}{\cancel{16}_4} = \frac{p}{4} + \frac{3}{8} - \frac{3}{8}p = \frac{3}{8} - \frac{1}{8}p$$ ■

k-pt mixture.

Let $X_1, X_2, \ldots, X_k$ be random variables w/ cdf F_{X_i}, $i=1..k$.
Let $a_1, a_2, \ldots, a_k$ be positive constants such that
$a_1 + a_2 + \cdots + a_k = 1$

Then, Y is a k-pt mixture if its cdf is of the form:

$$\boxed{F_Y(y) = a_1 \cdot F_{X_1}(y) + a_2 \cdot F_{X_2}(y) + \cdots + a_k \cdot F_{X_k}(y).}$$

Continuous Mixing.

- Start w/ a r.v.

  Λ which is going to play the role of the mixing parameter.

  Let f_Λ be the pdf of this continuous r.v.

- Suppose that, conditional on $\Lambda = \lambda$, the r.v. X has the pdf $f_{X|\Lambda}(x|\lambda)$

  and the cdf $F_{X|\Lambda}(x|\lambda)$.

Then, the unconditial pdf of X is

$$f_X(x) = \int f_{X|\Lambda}(x|\lambda) f_\Lambda(\lambda)\, d\lambda.$$

Also, the unconditional cdf of X is

$$F_X(x) = \int F_{X|\Lambda}(x|\lambda) \cdot f_\Lambda(\lambda)\, d\lambda$$

Consequences:

- $\mathbb{E}[X^k] = \mathbb{E}[\mathbb{E}[X^k | \Lambda]]$
- $\mathrm{Var}[X] = \mathbb{E}[\mathrm{Var}[X|\Lambda]] + \mathrm{Var}[\mathbb{E}[X|\Lambda]]$