Discrete Mixtures

30M

Sometimes, we may need to combine several basic distributions when modeling an event. For example, a randomly selected homeowners insurance claim may be from a fire, burglary, or liability incident. To model a randomly selected claim, we could use a ***mixture*** that is made up of the basic distributions used to individually model fire, burglary, and liability claims.

A random variable Y is a ***discrete mixture***, or ***finite mixture***, of random variables X_1, X_2, ..., X_n if its probability function can be expressed as a weighted average of the probability functions of these n random variables. Y is also called a ***n-point mixture***.

$$\begin{aligned} f_Y(y) &= w_1 \cdot f_{X_1}(y) + w_2 \cdot f_{X_2}(y) + \ldots + w_n \cdot f_{X_n}(y) \\ &= \sum_{i=1}^{n} w_i \cdot f_{X_i}(y) \end{aligned} \quad \text{(S2.2.}$$

In the equation above, w_i represents the ***weight*** of random variable X_i. Often, the weights can be treated as probabilities because they have the following properties:

- $\sum_{i=1}^{n} w_i = 1$
- $0 \le w_i \le 1$

Coach's Remarks

The $f_{X_i}(y)$'s in $(\mathrm{S2.2.8.1})$ do not necessarily need to be continuous PDFs. They are generic probability functions because mixture distributions can be a combination of discrete and continuous distributions.

The word "discrete" in "discrete mixtures" refers to combining a **finite** number of distributions.

The following are a direct consequence of (S2.2.8.1).

$$F_Y(y) = \sum_{i=1}^{n} w_i \cdot F_{X_i}(y) \qquad (\text{S2.2.8.2})$$

$$S_Y(y) = \sum_{i=1}^{n} w_i \cdot S_{X_i}(y) \qquad (\text{S2.2.8.3})$$

$$\mathrm{E}\left[Y^k\right] = \sum_{i=1}^{n} w_i \cdot \mathrm{E}\left[X_i^k\right] \qquad (\text{S2.2.8.4})$$

Keep in mind that a mixture is **not** the same as a *linear combination* of random variables. Consider the two examples below and note their differences:

- Mixture

$$f_Y(y) = w_1 \cdot f_{X_1}(y) + w_2 \cdot f_{X_2}(y)$$

  Y follows the distribution of X_1 for $100w_1\%$ of the time, and follows the distribution of X_2 for $100w_2\%$ of the time.

- Linear combination of random variables

$$Y = w_1 X_1 + w_2 X_2$$

  Y follows neither the distribution of X_1 nor X_2. Instead, its exact value is dependent on the values of both X_1 and X_2. Unlike mixtures, w_1 and w_2 are not weights. Therefore, they can be any real number and do not have to add to 1.

As an example, consider the scenario below and identify if it is a mixture or a linear combination.

A class consists of 17 boys and 13 girls. For boys, the probability of being late for class is 0.2; for girls, the probability of being late for class is 0.15.

What is the probability that a randomly chosen student is late for class?

The selected student can be either a boy **or** a girl; thus, the probability that the student is late for class is a **mixture**. Since there are 30 students in total, the probability that the randomly chosen student is a boy is 17/30, and the probability that the randomly chosen student is a girl is 13/30.

- $\mathrm{Pr}(\mathrm{Late} \mid \mathrm{Boy}) = 0.2$
- $\mathrm{Pr}(\mathrm{Late} \mid \mathrm{Girl}) = 0.15$

The desired probability is:

$$\begin{aligned}\mathrm{Pr}(\mathrm{Late}) &= \mathrm{Pr}(\mathrm{Boy}) \cdot \mathrm{Pr}(\mathrm{Late} \mid \mathrm{Boy}) + \mathrm{Pr}(\mathrm{Girl}) \cdot \mathrm{Pr}(\mathrm{Late} \mid \mathrm{Girl}) \\ &= \frac{17}{30}(0.2) + \frac{13}{30}(0.15) \\ &= \mathbf{0.17833}\end{aligned}$$

This is not a linear combination because it is **not true** that the selected student is "17 parts boy and 13 parts girl".

In the scenario above, notice that 0.2 and 0.15 are conditional probabilities. Therefore, $(\mathrm{S}2.2.8.1)$ can be interpreted as the weighted average of various conditional distributions. This fact allows us to apply the three important theorems introduced in Section S2.1.5:

- $\mathrm{Pr}(Y = y) = \mathrm{E}_X[\mathrm{Pr}(Y = y \mid X)]$
- $\mathrm{E}_Y[Y] = \mathrm{E}_X[\mathrm{E}_Y[Y \mid X]]$
- $\mathrm{Var}_Y[Y] = \mathrm{E}_X[\mathrm{Var}_Y[Y \mid X]] + \mathrm{Var}_X[\mathrm{E}_Y[Y \mid X]]$

In short, when dealing with mixtures, think in terms of conditional distributions.

Let's see these formulas in action.

Example S2.2.8.1

For an automobile insurance coverage issued by company XYZ, policyholders can be categorized into three types of risks: low, medium, and high. The annual loss amount has a mixed exponential distribution.

The average loss per year for each risk class is:

Risk	Average Loss per Year
Low	300
Medium	540
High	800

Suppose 40% of the insureds are low risk and 35% of the insureds are high risk.

For a randomly selected policyholder, calculate:

1. the probability that the annual loss is less than 500.
2. the expected annual loss.
3. the standard deviation of the annual loss.

Solution to (1)

Let X be the annual loss size. It is a mixture of 3 exponential distributions, one for each risk class.

Let L, M, and H be low, medium and high risk indicators, respectively.

Therefore,

- $(X \mid L) \sim \text{Exponential }(300)$, where $\Pr(L) = 0.4$
- $(X \mid M) \sim \text{Exponential }(540)$, where $\Pr(M) = 0.25$
- $(X \mid H) \sim \text{Exponential }(800)$, where $\Pr(H) = 0.35$

The probability that the annual loss is less than 500 is:

$$\begin{aligned}\Pr(X < 500) &= \mathrm{E}[\Pr(X < 500 \mid \mathrm{Risk})] \\ &= 0.4\Pr(X < 500 \mid L) + 0.25\Pr(X < 500 \mid M) + 0.35 \\ &= 0.4\left(1 - e^{-500/300}\right) + 0.25\left(1 - e^{-500/540}\right) + 0.35\left(1 - \right. \\ &= \mathbf{0.6381}\end{aligned}$$

Solution to (2)

The expected annual loss is

$$\begin{aligned}\mathrm{E}[X] &= \mathrm{E}[\mathrm{E}[X \mid \mathrm{Risk}]] \\ &= 0.4\mathrm{E}[X \mid L] + 0.25\mathrm{E}[X \mid M] + 0.35\mathrm{E}[X \mid H] \\ &= 0.4(300) + 0.25(540) + 0.35(800) \\ &= \mathbf{535}\end{aligned}$$

Solution to (3)

To calculate the standard deviation of annual loss, the Law of Total Variance can be used. However, since $\mathrm{E}[X]$ has already been calculated above, it is more efficient to calculate the second moment of X and compute $\mathrm{E}\left[X^2\right] - \mathrm{E}[X]^2$. Both methods should produce the same answer.

Refer to the exam tables for the second moment of the exponential distribution.

$$\begin{aligned}\mathrm{E}\left[X^2\right] &= \mathrm{E}\left[\mathrm{E}\left[X^2 \mid \mathrm{Risk}\right]\right] \\ &= 0.4\mathrm{E}\left[X^2 \mid L\right] + 0.25\mathrm{E}\left[X^2 \mid M\right] + 0.35\mathrm{E}\left[X^2 \mid H\right] \\ &= 0.4\left(2 \cdot 300^2\right) + 0.25\left(2 \cdot 540^2\right) + 0.35\left(2 \cdot 800^2\right) \\ &= 665{,}800\end{aligned}$$

$$\begin{aligned}\mathrm{Var}[X] &= \mathrm{E}\left[X^2\right] - \mathrm{E}[X]^2 \\ &= 665{,}800 - 535^2 \\ &= 379{,}575\end{aligned}$$

$$\sigma_X = \sqrt{\mathrm{Var}[X]} = \mathbf{616.097}$$

Coach's Remarks

In the calculation of $\mathrm{E}\left[X^2\right]$ above, notice that it is the weighted average of the second moment of the conditional random variables. This is not the same as the weighted average of the first moment squared for the conditional random variables.

$$\begin{aligned}\mathrm{E}\left[X^2\right] &= \sum_{\text{all risks}} \Pr(\mathrm{Risk}) \cdot \mathrm{E}\left[X^2 \mid \mathrm{Risk}\right] \\ &\neq \sum_{\text{all risks}} \Pr(\mathrm{Risk}) \cdot \mathrm{E}[X \mid \mathrm{Risk}]^2\end{aligned}$$

Example S2.2.8.2

Fred owns a new car and an old car. Every morning, he flips a fair coin to decide which car to drive to work.

- With his new car, the time to travel to his office, in minutes, is uniformly distributed on [20, 30].
- With his old car, the time to travel to his office, in minutes, is uniformly distributed on [35, 50].

Determine the variance of Fred's travel time to his office on a particular day.

Solution

Let T be the travel time to Fred's office.

- $(T \mid \text{New}) \sim \text{Uniform}\ (20,\ 30)$
- $(T \mid \text{Old}) \sim \text{Uniform}\ (35,\ 50)$

Apply the Law of Total Variance:

$$\text{Var}[T] = \text{E}[\text{Var}[T \mid \text{Car}]] + \text{Var}[\text{E}[T \mid \text{Car}]]$$

Car	New	Old
$\text{E}[T \mid \text{Car}]$	$\dfrac{20+30}{2} = 25$	$\dfrac{35+50}{2} = 42.5$
$\text{Var}[T \mid \text{Car}]$	$\dfrac{(30-20)^2}{12} = 8.3333$	$\dfrac{(50-35)^2}{12} = 18.75$
$\Pr(\text{Car})$	0.5	0.5

- Calculate $\text{E}[\text{Var}[T \mid \text{Car}]]$.

$$\begin{aligned}\text{E}[\text{Var}[T \mid \text{Car}]] &= 8.3333(0.5) + 18.75(0.5) \\ &= 13.5417\end{aligned}$$

- Calculate $\text{Var}[\text{E}[T \mid \text{Car}]]$.

$$\begin{aligned}\mathrm{E}\left[\mathrm{E}[T \mid \mathrm{Car}]^2\right] &= 25^2(0.5) + 42.5^2(0.5) \\ &= 1{,}215.625\end{aligned}$$

$$\begin{aligned}\mathrm{E}[\mathrm{E}[T \mid \mathrm{Car}]] &= 25(0.5) + 42.5(0.5) \\ &= 33.75\end{aligned}$$

$$\begin{aligned}\mathrm{Var}[\mathrm{E}[T \mid \mathrm{Car}]] &= 1{,}215.625 - 33.75^2 \\ &= 76.5625\end{aligned}$$

Therefore,

$$\begin{aligned}\mathrm{Var}[T] &= \mathrm{E}[\mathrm{Var}[T \mid \mathrm{Car}]] + \mathrm{Var}[\mathrm{E}[T \mid \mathrm{Car}]] \\ &= 13.5417 + 76.5625 \\ &= \mathbf{90.1042}\end{aligned}$$

■

Most discrete mixed distributions that appear on the exam are only a mixture of two distributions (or three, occasionally), as shown in Example S2.2.8.2. In these cases, the ***Bernoulli shortcut*** can expedite part of the Law of Total Variance calculation.

Bernoulli Shortcut

The Bernoulli shortcut is related to the Bernoulli distribution. It is a technique to quickly calculate the variance of a random variable that has only **two** possible values.

In general, for a random variable

$$X = \begin{cases} a, & \text{Probability} = q \\ b, & \text{Probability} = 1-q \end{cases}$$

the variance of X can be calculated as

$$\text{Var}[X] = (a-b)^2 q\,(1-q) \qquad \text{(S2.2.8.5)}$$

This shortcut is extremely useful because the variance can be determined without calculating the first and second moments.

Let's apply this technique to recalculate the $\text{Var}[\text{E}[T \mid \text{Car}]]$ in the previous example. Recall that there are two possible values for $\text{E}[T \mid \text{Car}]$:

$$\text{E}[T \mid \text{Car}] = \begin{cases} 25, & \text{Probability} = 0.5 \\ 42.5, & \text{Probability} = 0.5 \end{cases}$$

Therefore,

$$\begin{aligned} \text{Var}[\text{E}[T \mid \text{Car}]] &= (25-42.5)^2 \cdot 0.5 \cdot 0.5 \\ &= 76.5625 \end{aligned}$$

which matches the result calculated using the first and second moments.

The Bernoulli shortcut will be used repeatedly hereafter.

Let's work on another example.

Example S2.2.8.3

You are given two independent models:

- X is Pareto with $\alpha = 4$ and $\theta = 240$.

- Y is Pareto with $\alpha = 3$ and $\theta = 300$.

Evaluate:

1. the variance of Z, such that $f_Z(z) = 0.4 f_X(z) + 0.6 f_Y(z)$.
2. the variance of T, such that $T = 0.4X + 0.6Y$.

Solution to (1)

To calculate the variance of Z, notice that its density function is expressed as a weighted average of the density functions of X and Y. This implies Z is a mixture.

$$(Z \mid X) \sim \text{Pareto } (4,\ 240)$$

$$(Z \mid Y) \sim \text{Pareto } (3,\ 300)$$

To calculate its variance, apply the Law of Total Variance.

$$\text{Var}[Z] = \text{E}[\text{Var}[Z \mid \text{Model}]] + \text{Var}[\text{E}[Z \mid \text{Model}]]$$

Model	$\boldsymbol{X}$	$\boldsymbol{Y}$
$\text{E}[Z \mid \text{Model}]$	$\dfrac{240}{4-1} = 80$	$\dfrac{300}{3-1} = 150$
$\text{Var}[Z \mid \text{Model}]$	$80^2\left(\dfrac{4}{4-2}\right) = 12{,}800$	$150^2\left(\dfrac{3}{3-2}\right) = 67{,}500$
$\Pr(\text{Model})$	0.4	0.6

In the table above, the variances of $(Z \mid \text{Model})$ are calculated using the Pareto variance formula covered in Section S2.2.3. Alternatively, calculate the Pareto variance from the first and second moments. The moments formula is included in the exam tables.

- Calculate $\text{E}[\text{Var}[Z \mid \text{Model}]]$.

$$\begin{aligned}\mathrm{E}[\mathrm{Var}[Z \mid \mathrm{Model}]] &= 12{,}800(0.4) + 67{,}500(0.6) \\ &= 45{,}620\end{aligned}$$

- Calculate $\mathrm{Var}[\mathrm{E}[Z \mid \mathrm{Model}]]$ using the Bernoulli shortcut.

$$\mathrm{E}[Z \mid \mathrm{Model}] = \begin{cases} 80, & \text{Probability} = 0.4 \\ 150, & \text{Probability} = 0.6 \end{cases}$$

$$\begin{aligned}\mathrm{Var}[\mathrm{E}[Z \mid \mathrm{Model}]] &= (80 - 150)^2 \cdot 0.4 \cdot 0.6 \\ &= 1{,}176\end{aligned}$$

Therefore,

$$\begin{aligned}\mathrm{Var}[Z] &= 45{,}620 + 1{,}176 \\ &= \mathbf{46{,}796}\end{aligned}$$

Solution to (2)

To calculate the variance of T, notice that it is a linear combination of X and Y. Therefore,

$$\begin{aligned}\mathrm{Var}[T] &= \mathrm{Var}[0.4X + 0.6Y] \\ &= 0.4^2\mathrm{Var}[X] + 0.6^2\mathrm{Var}[Y] + 2(0.4)(0.6)\mathrm{Cov}[X,\, Y] \\ &= 0.4^2(12{,}800) + 0.6^2(67{,}500) + 2(0.4)(0.6)(0) \\ &= \mathbf{26{,}348}\end{aligned}$$

Note that $\mathrm{Cov}[X, Y] = 0$ because X and Y are independent.