3.5. **Distribuciones de Probabilidad.** Una distrbución de probabilidad es una función matemática que relaciona los estados de una variable de un *experimento aleatorio* con la probabilidad de que estos ocurran (Ash, 2008). Es importante distinguir entre las variables aleatorias clasificandolas como *continuas* o *discretas*, pues la forma de tratarlas a través de distribuciones de probabilidad es distinta.

3.5.1. *Variables Discretas.* En este caso basta con utilizar una función de masa de probabilidad, cuya salida al evaluar sobre un valor (para la variable aleatoria) es directamente su probabilidad de ocurrencia. Algunos ejemplos comunes de distribuciones discretas son:

- Distribución binomial: Expresa la probabilidad de obtener cierto número de éxitos en una serie de experimentos aleatorios e independientes entre sí, dada su probabilidad de éxito individual.
- Distribución de Poisson: Expresa la probabilidad de que ocurra un determinado número de eventos independientes en un intervalo de tiempo, dada la frecuencia media de los eventos.
- Distribución hipergeométrica: Expresa la probabilidad de obtener cierta cantidad de elementos de un tipo al realizar una muestra aleatoria y sin reemplazo de una población cuya distribución se conoce.

3.5.2. *Variables Continuas.* Por otro lado, dada la naturaleza de estas variables, sería absurdo asignar directamente una probabilidad a cada estado (pues son infinitos incontables), por lo que se utilizan funciones de densidad de probabilidad a través de las cuales puede calcularse la *masa de probabilidad* integrando sobre el intervalo cuya probabilidad desea obtenerse. Algunos ejemplos comunes de distribuciones continuas son:

- Distribución exponencial: Modela la probabilidad asociada al tiempo de espera para la ocurrencia de un evento, conociendo su frecuencia media de ocurrencia.
- Distribución normal: Modela datos que se distribuyen simétricamente al rededor de su media, volviéndose progresivamente menos comunes conforme se alejan de ella.

4. Definición del problema

Uno de los mayores desperdicios derivados de los procesos de restauración ecológica es el desaprovechamiento de las plantas durante la reforestación. La causa del desaprovechamiento de recursos radica en que se determina en base a estimaciones la cantidad de plantas necesarias para restaurar un área dada. Sin embargo, es común sobrestimar el número de plantas debido a la topografía del terreno y las plantas y elementos ya presentes en el área. Además, la sobrevivencia de las plantas disminuye si entre las plantas vecinas existen relaciones negativas. Considerando esto, el presente reto busca propooner un modelo matemático y computacional que permita plantear una mejor distribución de las especies a plantar en un polígono minimizando la competencia entre ellas y tomando en cuenta los elementos ya presentes en el mismo.

5. Modelo Matemático

En adelante se define la notación $[n] = \{i \mid i \in \mathbb{N}, i \leq n\}$ para facilitar la lectura.

5.1. **Conjuntos.** Sea $G = (N, M)$ un grafo no dirigido y completo, de nodos N y arcos M; donde $N = [n]$ es el conjunto de *espacios de plantación*. Se define el grafo $H = (S, A)$ como un grafo no dirigido y completo, de nodos S y arcos A, donde $S = [s]$ es el conjunto de las especies de plantas a considerar. Sea $L = \{l_i\} \forall i \in S$ el conjunto de requerimentos por especie.

5.2. **Parámetros.** Sean $n \in \mathbb{N}$ la cantidad de *espacios de plantación*. Sea $s \in \mathbb{N}$ el número de especies de plantas a considerar. Los parámetros $l_i \in L$ corresponden a la cantidad requerida de especímenes de la especie $i \in S$ por hectárea. El grafo G tiene ponderaciones binarias $g_m : m \in M$ que toman valor de 1 si los *espacios de plantación* representados por los nodos unidos por el arco m mantienen una relación de *adyacencia espacial* en la plantación. El grafo H tiene ponderaciones no-negativas $h_a : a \in A$ que corresponden a una penalización por la competencia entre las especies que representan los nodos conectados por el arco a.

5.3. **Variables.** Sea $p_{ia} : i \in N, a \in S$ una binaria que toma como valor 1 si en el *espacio de plantación* i está plantado un especimen de la especie k, de lo contrario toma el valor 0. Sean z_{ijab} tal que $i, j \in N$ y $a, b \in S$ variables auxiliares que toman como valor el resultado de la operación lógica p_{ia} AND p_{jb}.

5.4. **Problema de Programación Lineal.** El planteamiento es el siguiente:

5.4.1. *Función Objetivo.*

$$\text{Minimizar} \quad \sum_{i \in N} \sum_{j \in N} \sum_{a \in S} \sum_{b \in S} (g_{(i,j)} * z_{ijab}) * h_{(a,b)} \tag{1}$$

5.4.2. *Restricciones.*

$$\sum_{a \in S} p_{ia} = 1 \qquad \forall i \in N \tag{2}$$

$$\sum_{i \in N} p_{ia} = l_a \qquad \forall a \in S \tag{3}$$

$$z_{ijab} \geq p_{ia} + p_{jb} - 1 \qquad \forall i, j \in N \quad a, b \in S \tag{4}$$

$$z_{ijab} \leq p_{ia} \qquad \forall i, j \in N \quad a, b \in S \tag{5}$$

$$z_{ijab} \leq p_{jb} \qquad \forall i, j \in N \quad a, b \in S \tag{6}$$

$$z_{ijab} \in \{0, 1\} \qquad \forall i, j \in N \quad a, b \in S \tag{7}$$

$$p_{ia} \in \{0, 1\} \qquad \forall i \in N, \ a \in S \tag{8}$$

5.4.3. *Lógica tras las Ecuaciones.* La función objetivo (Ecuación 1) puede entenderse pensando en el primer producto como una operación que solo toma el valor 1 cuando se reunen a la vez las siguientes condiciones:

- Los *espacios de plantación* i y j son *espacialmente adyacentes*, por lo que compiten entre sí.
- La planta en el *espacio de plantación* i es, en efecto, de la especie a.
- La planta en el *espacio de plantación* j es, en efecto, de la especie b.

Y toma el valor 0 en cualquier otro caso. Es decir, que solo vale 1 cuando la competencia entre el par de plantas p_{ia} y p_{jb} debe ser considerada. Luego, este valor se multiplica por $h_{(a,b)}$ que es la penalización por la competencia entre dicho par de plantas.

La primera restricción (Ecuación 2) asegura que en cada *espacio de plantación* haya exactamente una planta; no espacios vacíos, no plantas superpuestas. La siguiente restricción (Ecuación 3) asegura que se cumplan los requerimentos de INECOL sobre cantidad de especímenes por especie en toda la plantación. Las siguientes tres restricciones (Ecuaciones 4, 5, 6) se utilizan para linealizar la operación lógica p_{ia} AND p_{jb} asegurando que z_{ijab} tome su valor. Las últimas dos restricciones (Ecuaciones 7 y 8) describen la naturaleza de las variables como binarias.

5.5. **Nivel de competencia.** De acuerdo a la información proporcionada por el INECOL, en cuanto a la diversidad de especies de plantas a utilizar, se investigaron datos acorde a las necesidades principales de cada especie, como su región natural, la cantidad de agua que requieren, sus usos y datos adicionales, para crear los parámetros que determinaron el nivel de competencia entre especies, siendo el siguiente.