Sandya Subramanian
Stat Conn HW2

1) **SAMPLE SPACE**

Graph space of fixed size $\{V, E, Y\}$

$V \in \{V_1 \dots V_n\}$ (vertices)

$E \in \{E_{11} \dots E_{nn}\}$ (edges)

$Y \in \{0\ 1\}$ (class labels)

OR

$A \in \{0,1\}^{nxn}$ (all possible adjacency matrices)

$Y \in \{0,1\}^n$ (class labels)

2) **MODEL**

$ER(p)$ – Erdos-Renyi model with p being the probability of an edge

OR

$SBM(k)$ – Stochastic Block Model with k being the number of clusters

If $k = 2$

$\beta \in \begin{pmatrix} p & q \\ q & r \end{pmatrix}$ $where\ p, q, r \in [0,1]$ (probabilities of an edge in each block)

3) **ACTION SPACE**

Vector of all cluster assignments possible

$A = \{y \in \{0,1\}^n\}$

4) **LOSS FUNCTION**

Loss function: total number of incorrect labels (after checking all possible permutations)

OR

ARI – Adjusted Rand Index – a way to compare two clusterings based on number of pairs of matrices that in same and different clusters between clusterings

5) **RISK FUNCTIONAL**

$E(L)$ – Expected value of Loss Function