

Tarea 4 Genómica Funcional 2025

Roberto Álvarez

07 April, 2025

Contents

Tarea 4	1
(20 %) Preguntas rápidas (simples)	1
(80 %) Scripts, código, texto y gráficas	1

Tarea 4

(20 %) Preguntas rápidas (simples)

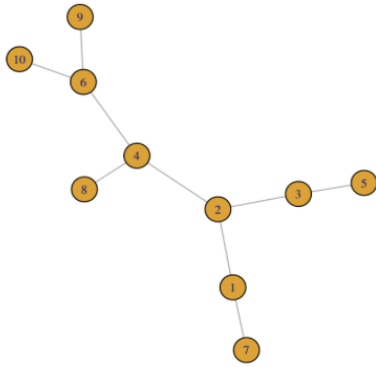
1. Considera una red de 15 nodos . Considera que tienes una sola conexión. ¿Cuál es el número máximo de nodos que puedes conectar? Dado un solo nodo , ¿cuál es el número máximo conexiones que puede tener tal nodo?
2. Considera una red dirigida de 10 nodos. Usa 25 conexiones y dibuja una red con ese número de nodos y conexiones dirigidas. Calcula la distribución de conectividades de salida. Calcula el número total de conexiones de salida ¿Cómo se compara el número total de de conectividades de entrada (son iguales, diferentes)? Explica tu respuesta. Calcula el promedio de entrada y de salida. Discute tu resultado.
3. Considera una red no dirigida de $n = 3, 4, 5, 6$ nodos respectivamente. ¿Cuál es el número máximo de conexiones que puedes poner? Dibuja cada caso. ¿Qué pasa si la red es dirigida?
4. Considera tu red de amigos reales. ¿Cuál es tu coeficiente de clusterización?

(80 %) Scripts, código, texto y gráficas

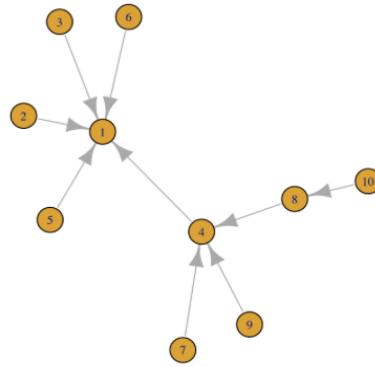
1. Propiedades básicas de redes

A partir de las siguientes redes calcula manualmente las siguientes propiedades de cada red

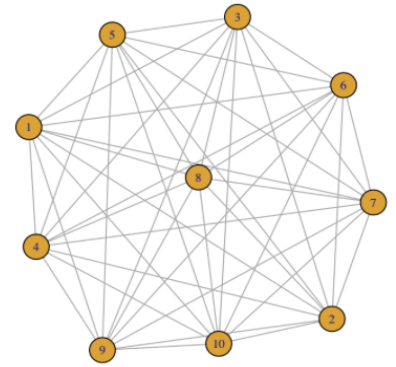
A



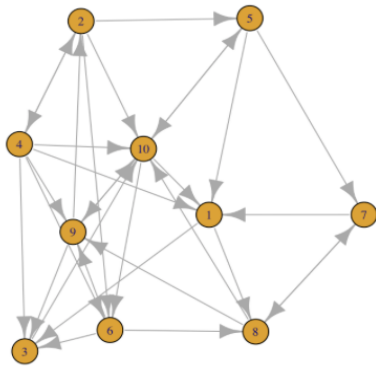
B



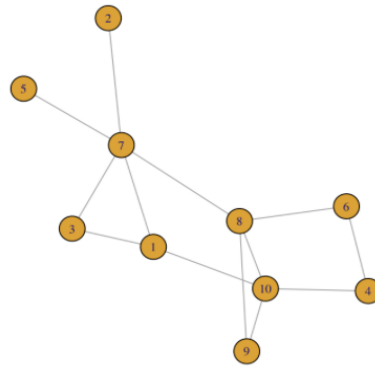
C



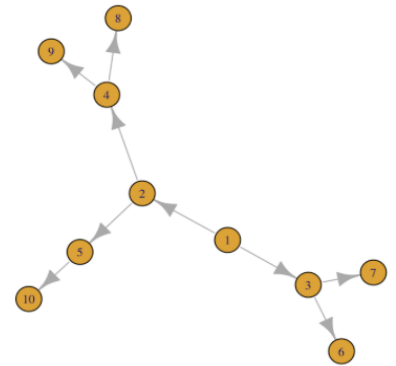
D



E



F



- Número de conexiones
- Número de nodos
- Degree
- Average degree
- Degree distribution
- Density

- Adjacency matrix
- Matriz de distancia
- Diámetro
- Nodos más distantes
- Coeficiente de clusterización

Después escribe código en R que genere las gráficas de las redes y que calcule las propiedades anteriores para cada red.

1. Karate

Considera la red del club de Karate de Zachary. Acá puedes leer sobre eso. En igraph la gráfica está precargada

```
library(igraph)
karate<-make_graph("Zachary")
```

- ¿Cuántos nodos y conexiones tiene?
- ¿Quiénes son los nodos y cuál es la regla de conexión?
- ¿Qué tan densa es la red?
- ¿Cómo obtienes la matriz de adyacencia?
- ¿Es una red dirigida, pesada?
- Calcula y gráfica la distribución de conectividades
- Calcula el diámetro, la matriz de distancias y la distancia promedio
- Encuentra la trayectoria de los nodos más alejados.
- Existen nodos con coeficiente de clusterización 1. ¿Qué significa?
- Mide , con al menos tres medidas de centralidad, los nodos más importantes.

Explica e interpreta tus resultados.

2. Amigues

A partir de la red de amigues que vimos en clase, en su versión no ponderada, contesta lo siguiente:

- Escribe una función que calcule el número de amigos que tiene cualquier persona arbitraria y lo compare con el número de los amigos de estos.
- Escribe una función que te de la trayectoria más larga entre dos nodos arbitrarios de la red.
- Encuentra las dos personas más populares.

3. Red PPI

A partir de la red de interacción proteína-proteína (PPI) de la levadura que se encuentra en la biblioteca `igraphdata` Elabora un script que conteste lo siguiente:

- Encuentre qué tipo de distribución de conectividades tiene. Haz un ajuste en log-log para ver que tipo de distribución podría ser.
- Encuentra las diez proteínas más conectadas
- Calcula el diámetro y promedio de las distancias
- Crea una función que, a partir de eliminar al azar un nodo de la red genere el promedio de las distancias después de eliminar $n = 1, 2, 3, \dots, 100$ nodos al azar
- Crea una función que elimine las proteínas más conectadas y calcule el promedio de las distancias cada vez que se remueve un nodo.
- Calcula el promedio del coeficiente de clusterización. ¿Hay proteínas que tengan un coeficiente de clusterización de 1? Eso qué significa.

Discute tus resultados.

4. Redes biológicas

Ve a al Network Repository y descarga la red en formato lista de adyacencia. Explica que representan los nodos y las conectividades.

Escribe código que resuelva lo siguiente:

- Cargue y genere una gráfica en **igraph**.
- Genera la gráfica con todos los layouts disponibles.
- Grafica la red con al menos tres medidas de centralidad.
- ¿Qué tan densa es la red?
- Clusteriza la red con al menos tres métodos de clusterización. Gráfica la red con esos métodos.

Explica tus resultados.

5. Red de coexpresión simulada

Simula una matriz de expresión con 20 genes en 6 condiciones diferentes.

- Calcula la correlación entre todos los pares de genes.
- Construye una red de coexpresión utilizando un umbral de correlación > 0.8 .
- Calcula la distribución de grados.
- Identifica si hay módulos o agrupamientos de genes altamente correlacionados.
- Visualiza la red y discute qué tipo de topología tiene.

6. Comparación de redes

Construye dos redes:

- Una red aleatoria tipo Erdos-Rényi.
- Una red tipo “scale-free” usando el modelo Barabási–Albert.

Para ambas redes:

- Compara su grado promedio, distribución de grados, coeficiente de clusterización y diámetro.
- Interpreta las diferencias en el contexto de redes biológicas.