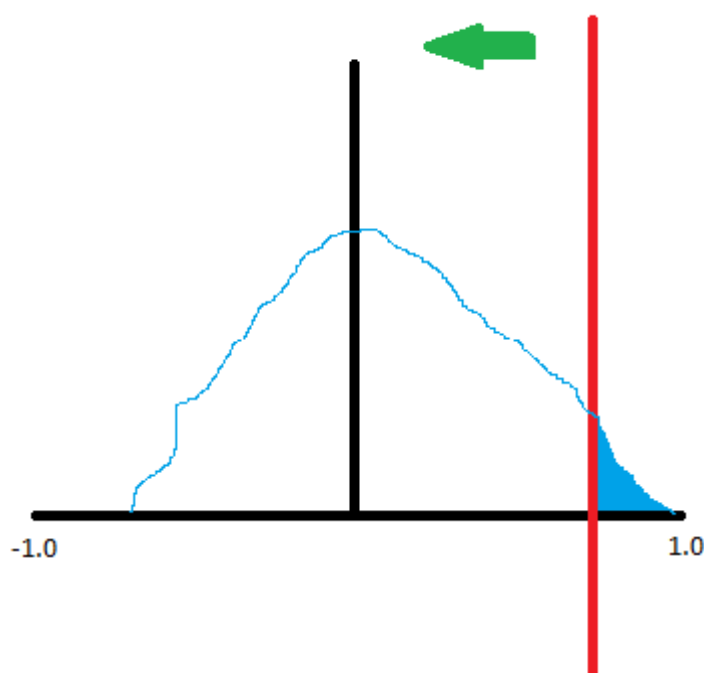


Análisis de P-values

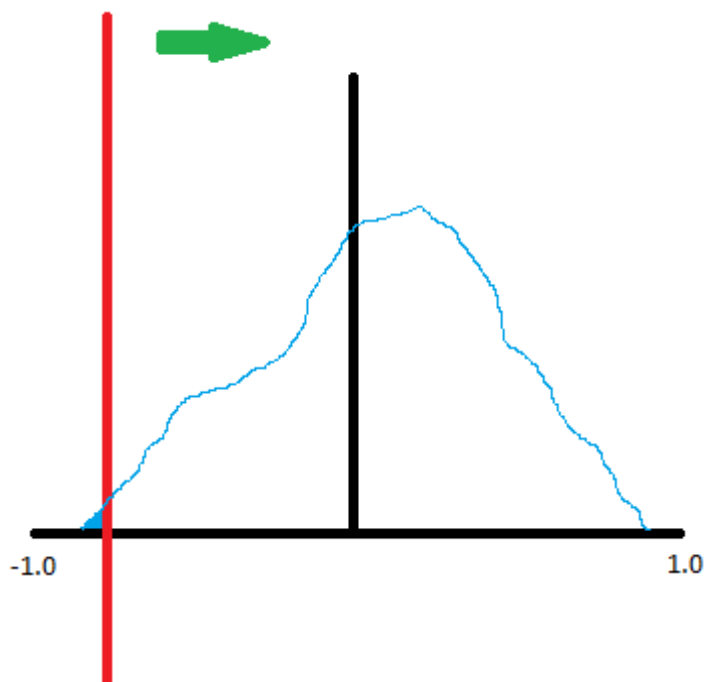
A continuación se muestran las gráficas de los P-values calculados para los cuatro coeficientes (Clustering Coefficient, Average Minimum Path Length, Global Efficiency y Small-World) en función del umbral seleccionado para binarizar el grafo.

Todos los grafos muestran P-value (eje Y) en función del umbral (eje X). La selección del umbral para realizar los barridos se ha hecho de 3 maneras distintas:

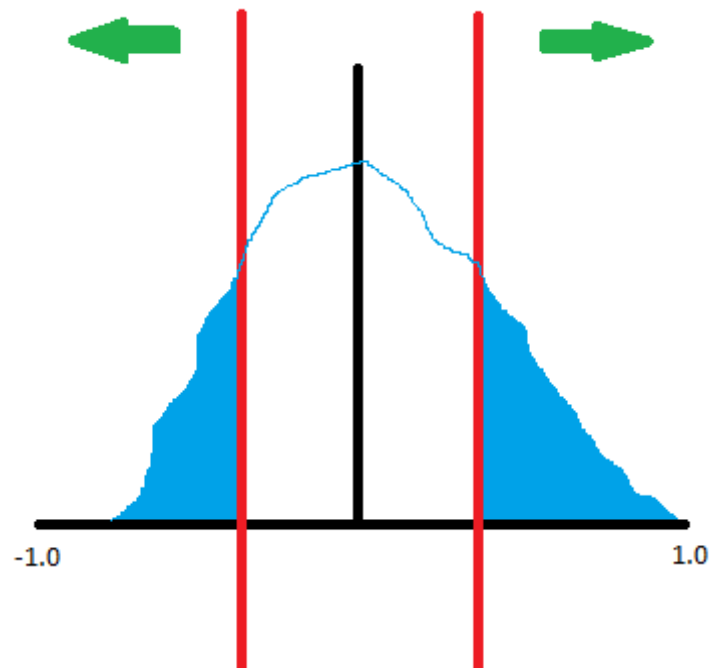
1. Umbral que mantiene únicamente el % especificado de aristas con valor más alto/positivo. Barrido entre 1% y 100%.



2. Umbral que mantiene únicamente el % especificado de aristas con valor más bajo/negativo. Barrido entre 1% y 100%.

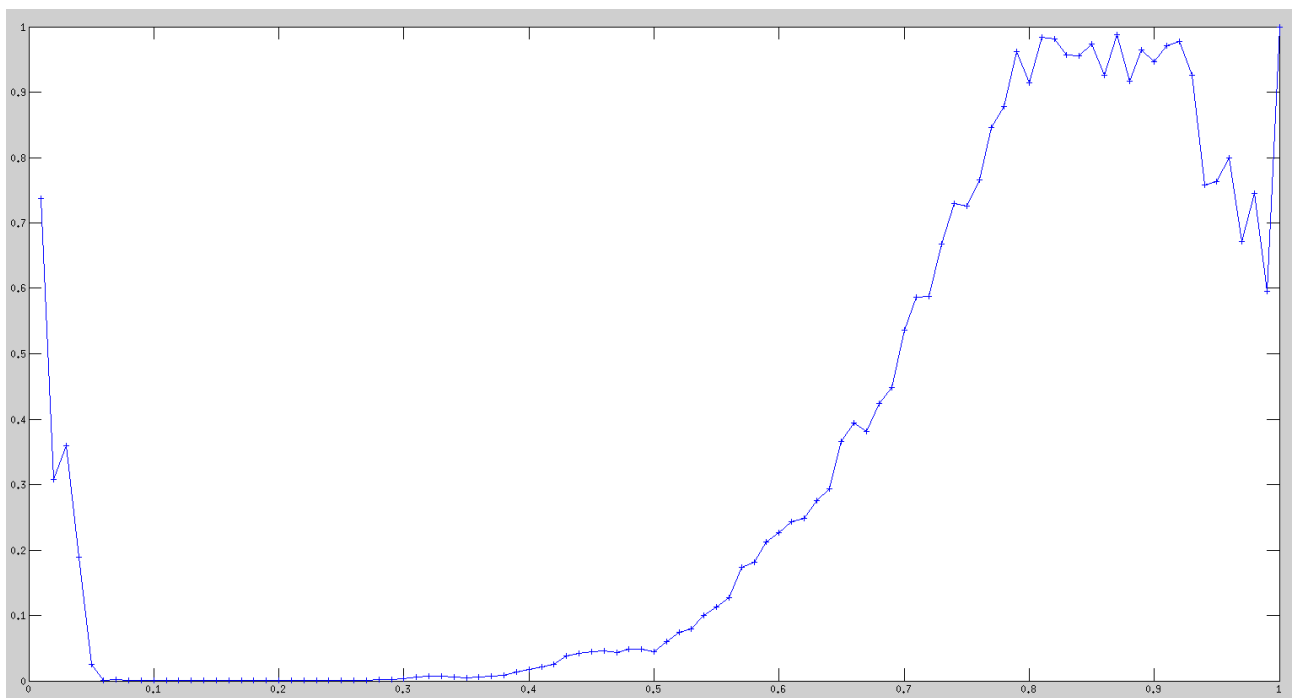


3. Umbral simétrico que binariza por valor absoluto del valor directo de las aristas. Este caso no se ha hecho mediante % para mantener la simetría del umbral. Barrido entre 0.01 y 1.0. Nótese que, a diferencia de los casos anteriores, el mayor número de aristas se encuentra para los valores más bajos del umbral. Por tanto, hay que “invertir” las gráficas para compararlas con las de los 2 casos anteriores.

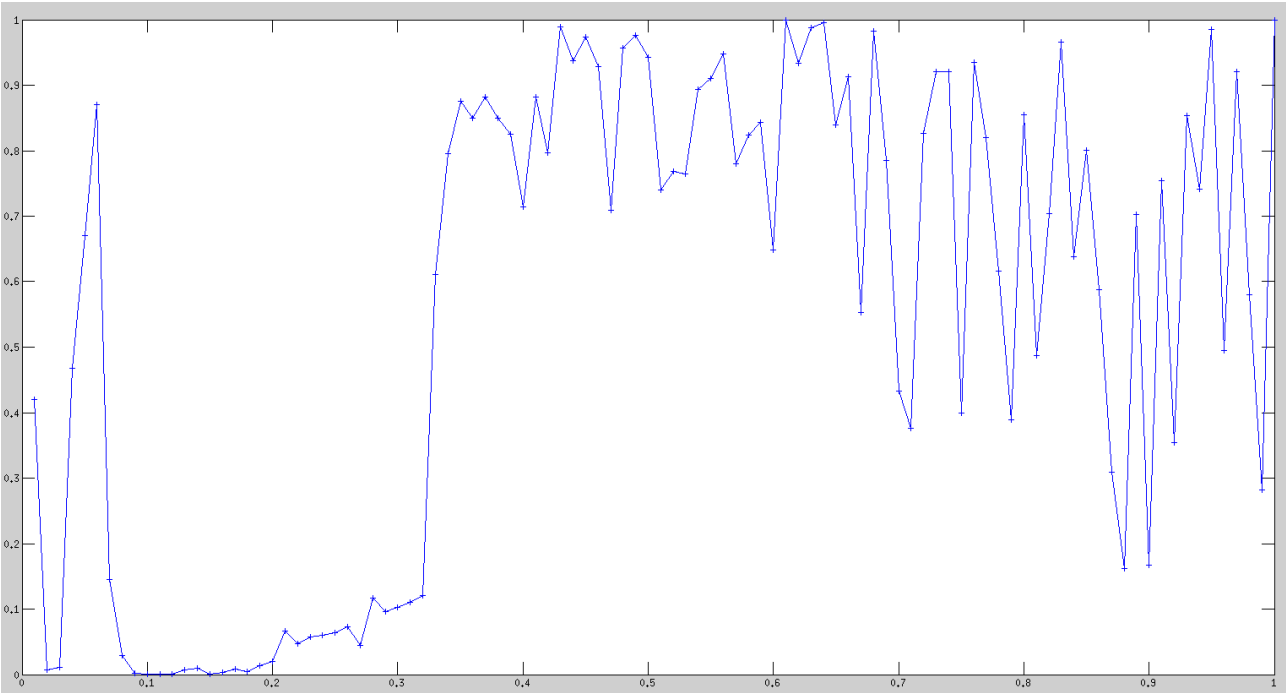


Umbral del tipo 1 (positivos):

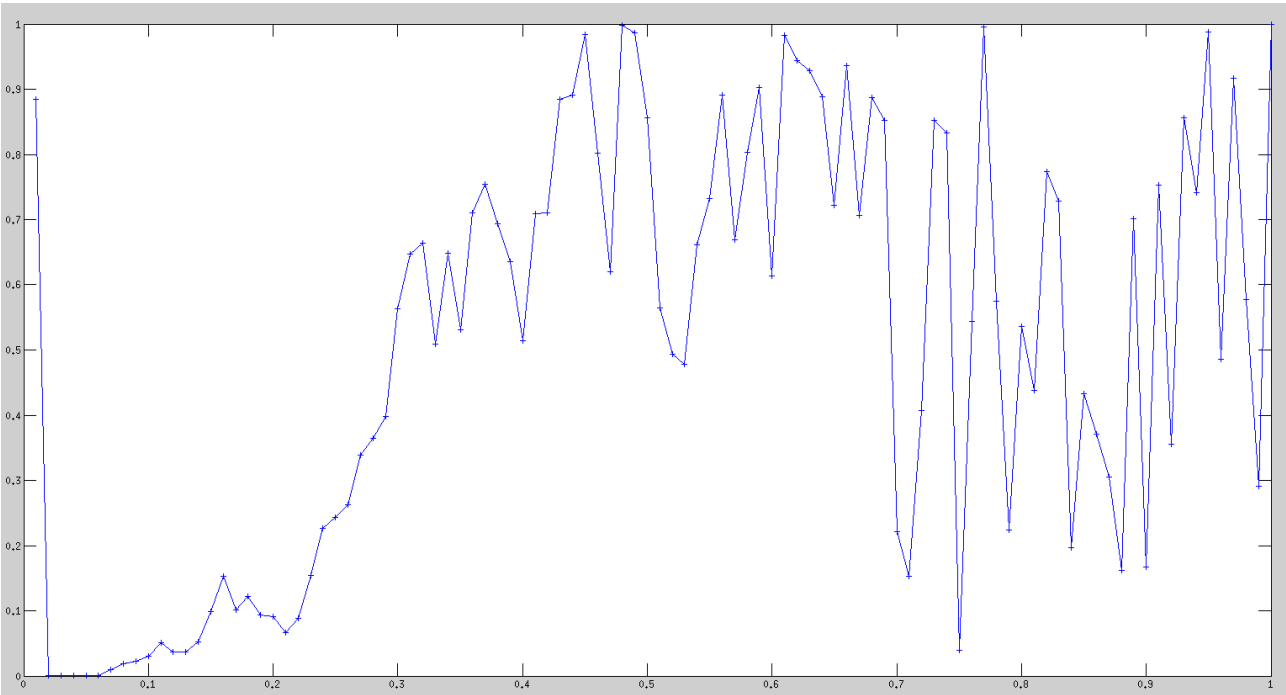
Clustering Coefficient



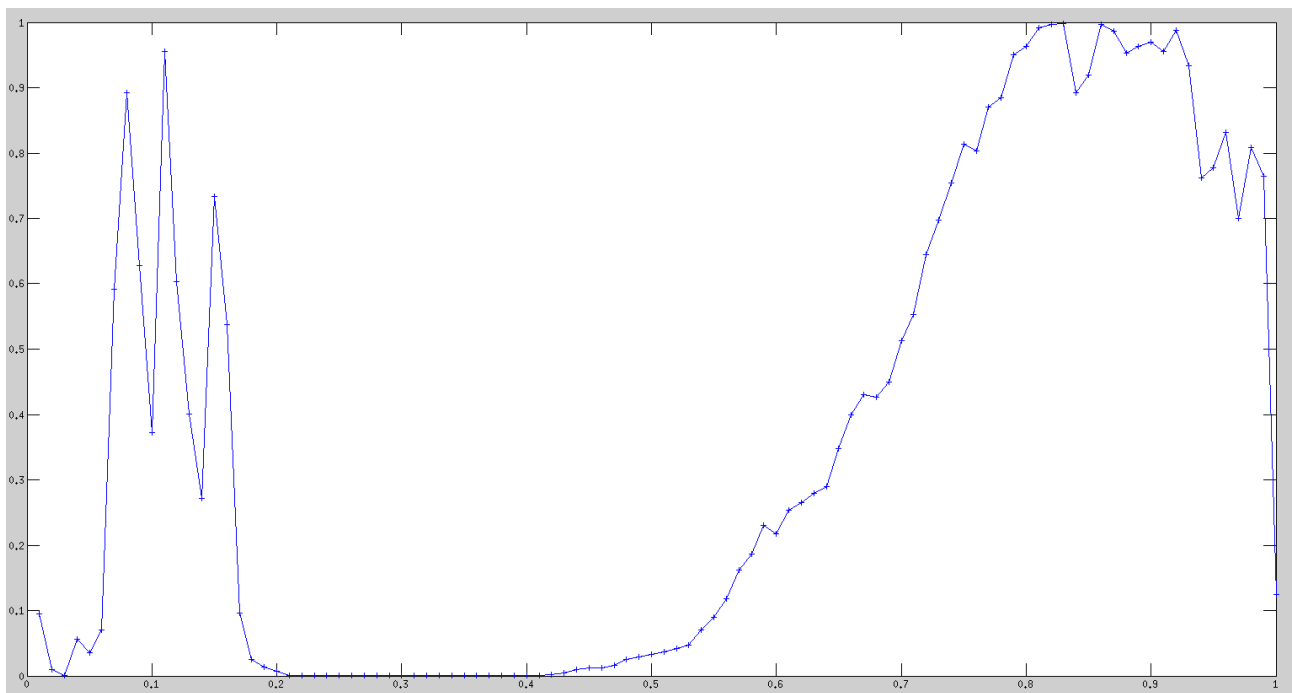
Average Mimimum Path Length



Global Efficiency

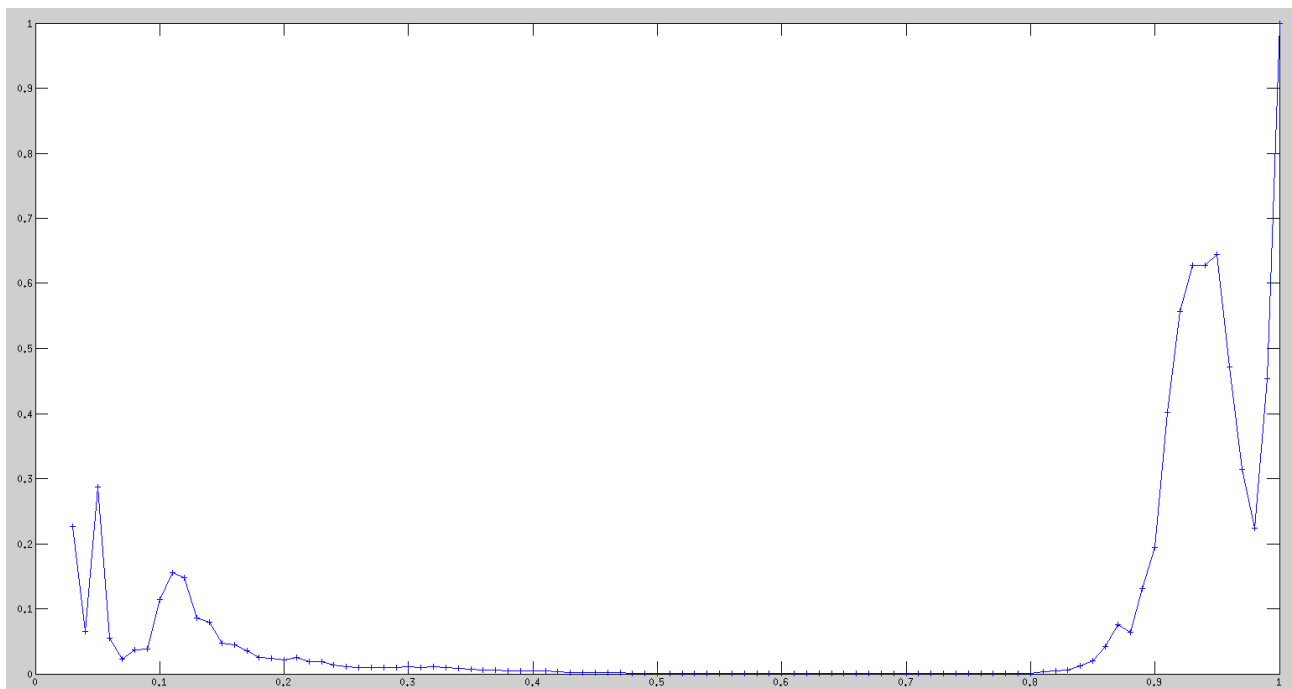


Small-World Coefficient

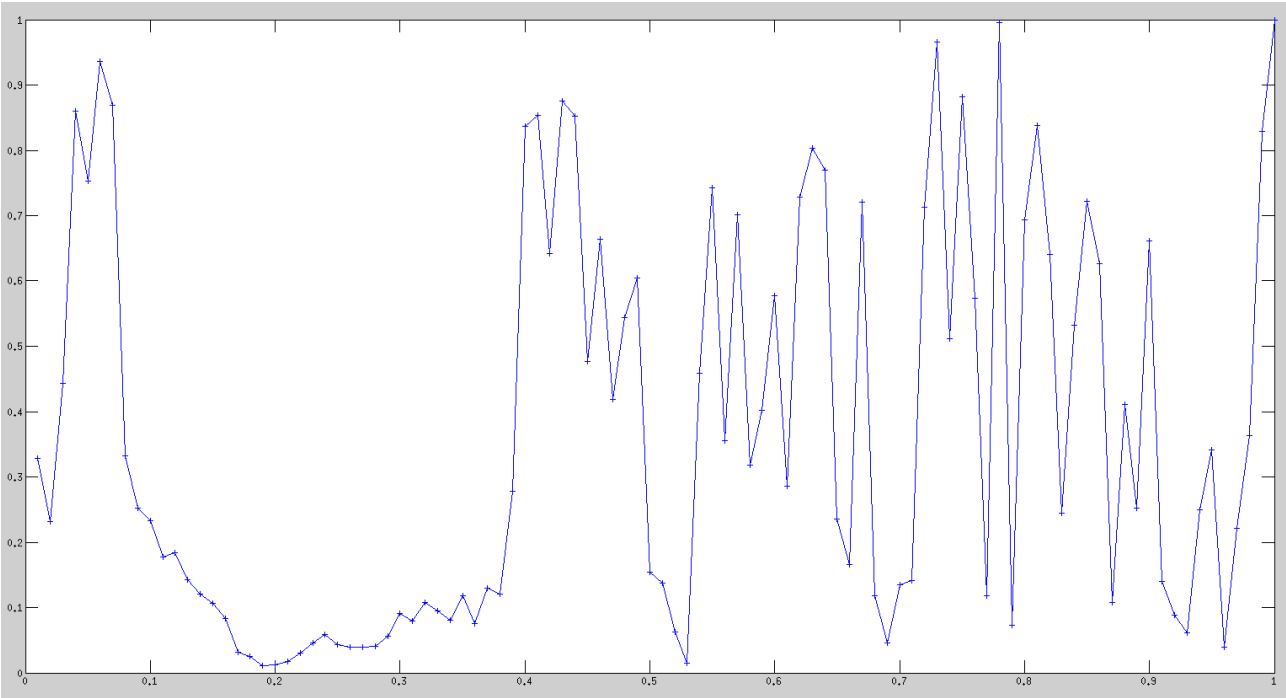


Umbral del tipo 2 (negativos):

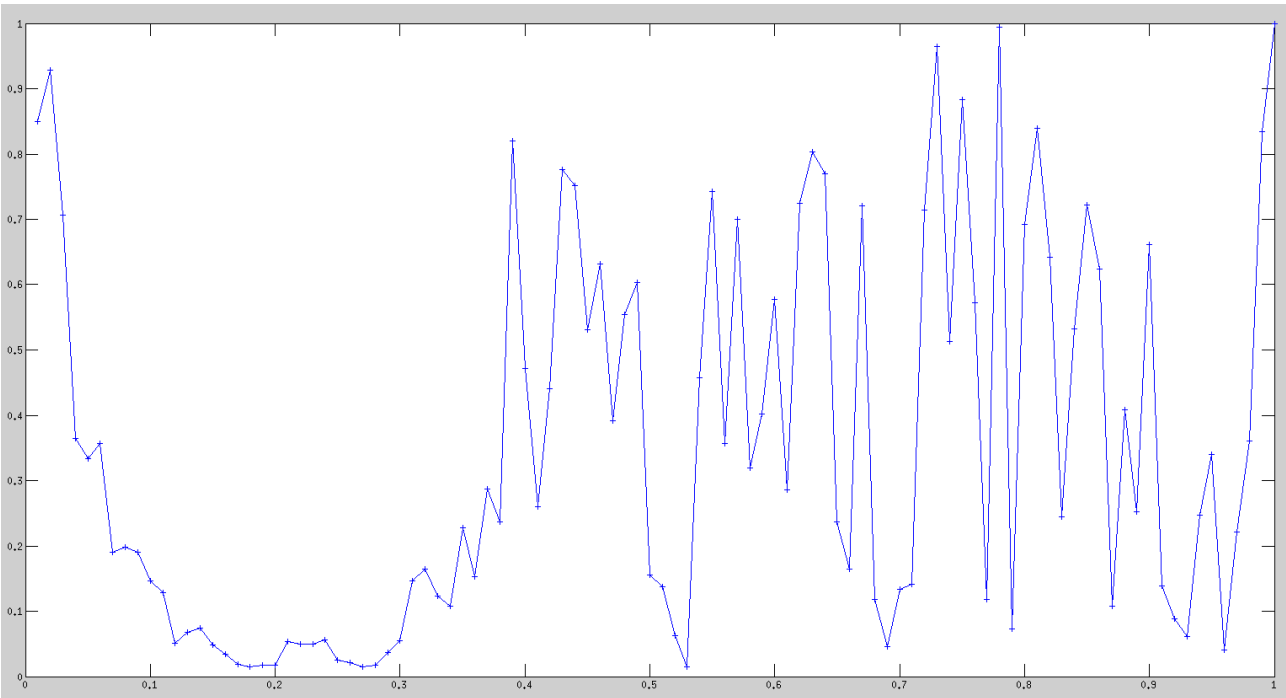
Clustering Coefficient



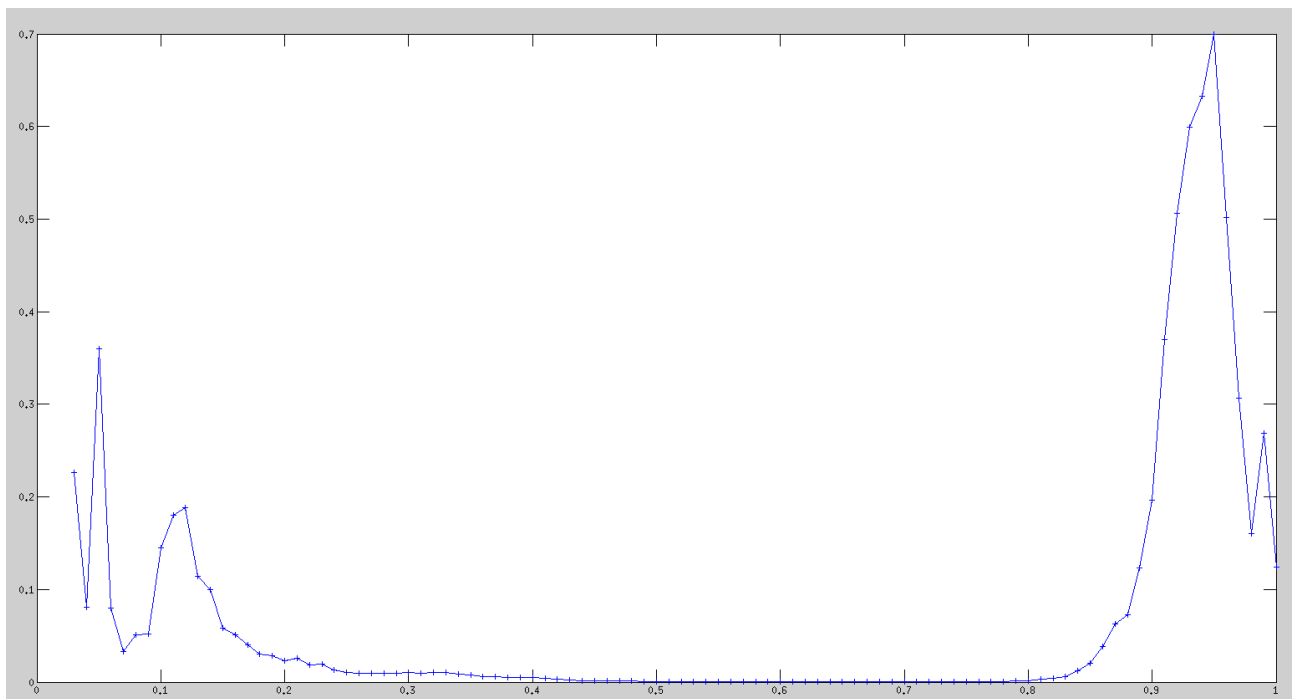
Average Mimimum Path Length



Global Efficiency

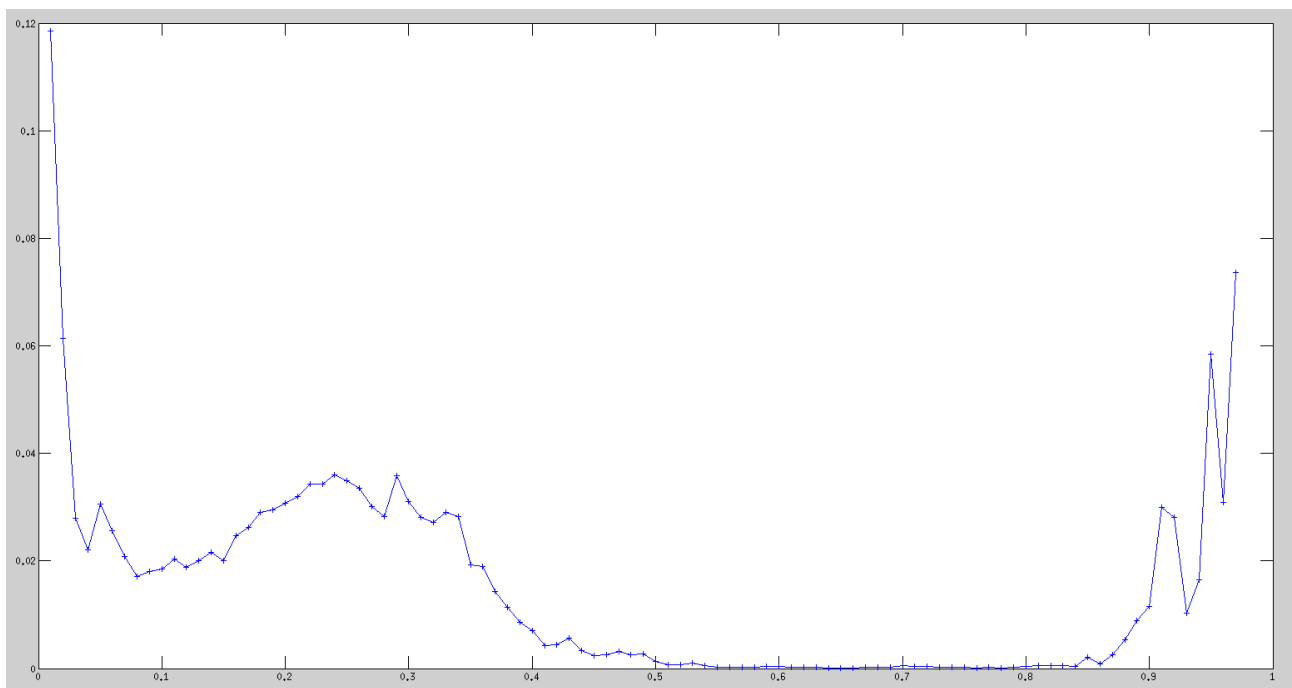


Small-World Coefficient

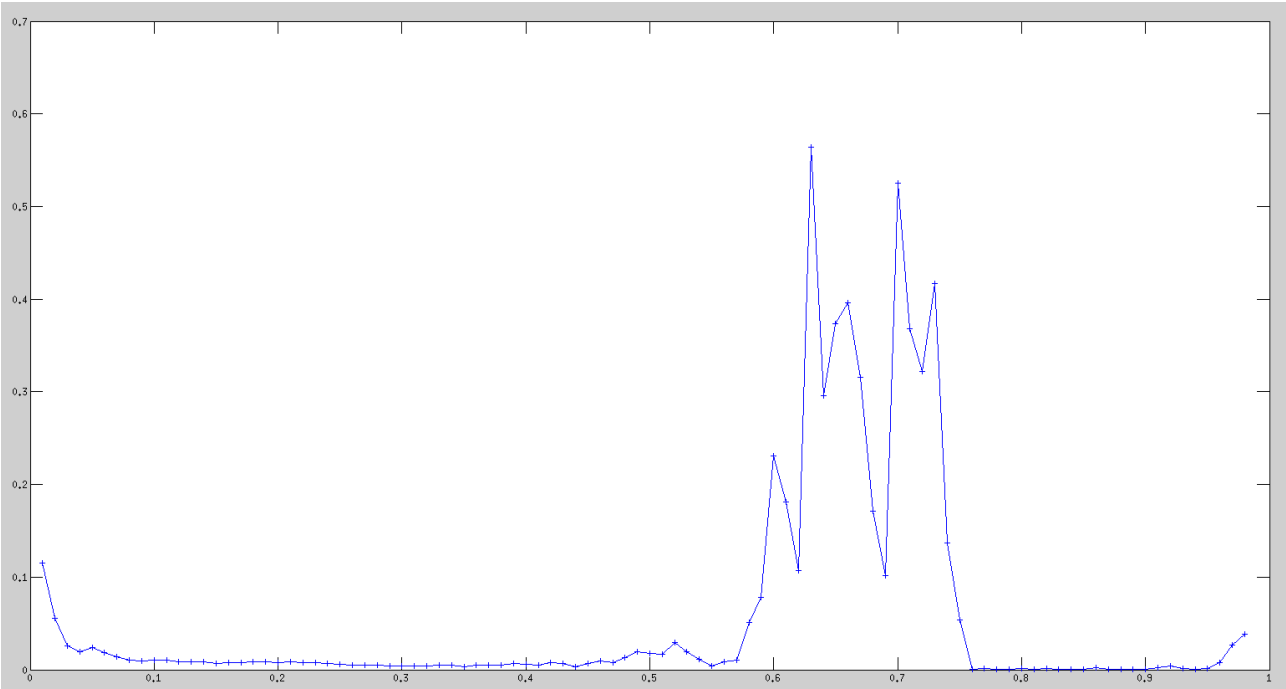


Umbral del tipo 3 (ambos - simétrico):

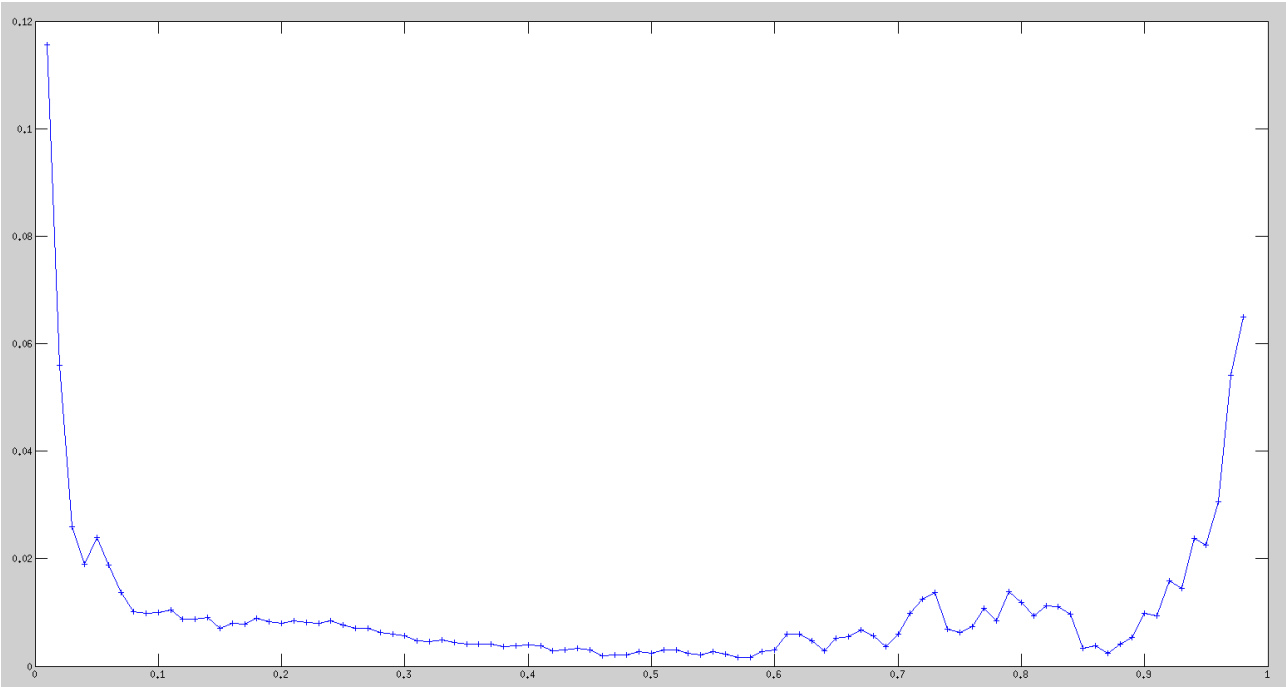
Clustering Coefficient



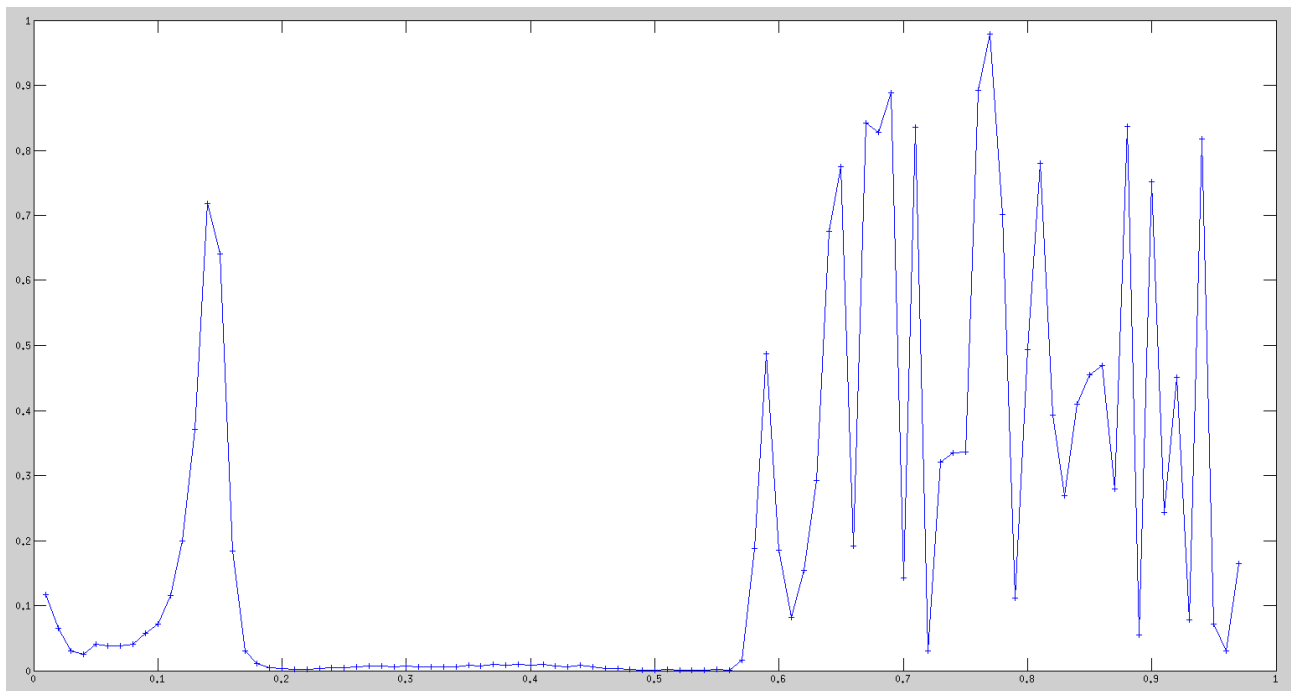
Average Mimimum Path Length



Global Efficiency



Small-World Coefficient



Comentarios, reflexiones y conclusiones

Umbral del 15% de aristas con más peso

Para los valores positivos, el P-value exacto para un umbral del 15% de aristas para el Clustering Coefficient es $3.80260332588778e-07$, mientras que para el Average Minimum Path Length es 0.000291515589991354 . Parece que ambos parámetros correlan bastante bien y sale inferior a 0.001, que era lo que esperábamos.

Para el Global Efficiency me sale 0.0987530548516353 , valor que se aleja de lo que le daba a Roser ($p < 0.001$), aunque para umbrales del 3% al 5% sí me sale inferior a 0.001. He estado comprobando el código y lo he probado con grafos pequeños y no he visto ningún error, quizá las diferencias esté en que no acabamos de elegir el umbral exactamente igual o en algún tema de precisión.

El valor P del Small World que me sale en este caso es 0.734116619174589 . Sin embargo, hay un tramo (21% al 41% de aristas) que el P-value me sale inferior a 0.001. El AMPL y el GE dan valores muy altos en este tramo, pero entre 21% y 27% el Clustering Coefficient da también valores de P inferiores a 0.001, con lo que tenemos 2 parámetros (CC y SW) que se comportan bien para umbrales 21% y 27%. ¿Podrían ser valores de umbrales a tener en cuenta?

Umbral que mantiene sólo el % de las aristas más negativas

Se observa que el Average Minimum Path Length y el Global Efficiency se comportan mal en general. En torno al 20% aparecen algunos valores inferiores al 0.05, pero en ningún punto de la gráfica hay valores inferiores a 0.001. Si comparamos con los valores de P más bajos que aparecen para el umbral del tipo 1, vemos que los de este caso son varios ordenes de magnitud superiores, lo

que me hace dudar de su importancia. Por otra parte, el Clustering Coefficient y el Small-World sí que da P's inferiores a 0.001 para umbrales entre el 48% y el 78% de aristas más negativas, aunque al ser % tan elevados ya empiezan a aparecer también las aristas positivas que están cercanas a 0, con lo que no creo que sea muy relevante.

Umbral simétrico (mirar las gráficas desde la derecha para comparar con las anteriores)

El Clustering Coefficient siempre da valores muy bajos (el valor máximo que aparece en la gráfica es 0.12). Lo mismo pasa con el Global Efficiency. Sin embargo, este último nunca presenta valores más bajos de 0.001.

El Average Minimum Path Length también tiene valores inferiores en general, pero sólo están por debajo de 0.001 para valores entre 0.8 y 0.9, aproximadamente. Para este tramo, el Clustering Coefficient también presenta algunos valores inferiores a 0.001. Este tramo se correspondería con los tramos en los que tenemos un % bajo de aristas, con lo que el comportamiento sería parecido al caso del umbral del tipo 1.

El Small-World es similar al caso del umbral del tipo 1.

Los 2 últimos puntos, unidos a lo que aparece en las gráficas del tipo 2 (aristas negativas) parecen indicar que las aristas negativas no tienen demasiada relevancia.