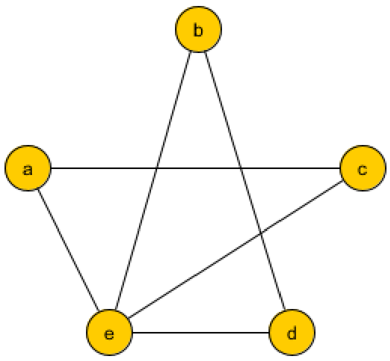
Sujet (niveau 2)
Graphes aléatoires d'Erdös-Renyi

On considère un espace probabilisé (Ω, A, P).
Pour définir un graphe aléatoire non orienté G on se donne :

- $S = \{s_1, ..., s_n\}$, un ensemble fini de $n \geq 2$ sommets ;
- pour toute paire de sommets $\{s_i, s_j\}$ avec $i < j$, $T_{i,j}$ est une variable de Bernoulli de paramètre p sur (Ω, A, P) ;

Les arêtes du graphe sont les paires $\{s_i, s_j\}$ telles que $T_{i,j} = 1$ avec $i < j$. Les variables $T_{i,j}$ sont supposées indépendantes.
Voici un exemple de graphe aléatoire avec $S = \{a, b, c, d, e\}$ et $p = 0,4$:



On peut considérer qu'un graphe aléatoire est un modèle très simplifié de réseau social à un instant donné.
1. Quel est le nombre maximal d'arêtes de G?
2. Ecrire une fonction Python `listAdj(S,p)` qui génère la liste des listes d'adjacence d'un tel graphe aléatoire ayant `S` pour liste de sommets.
Le graphe dessiné correspond à la liste des listes d'adjacence suivante, `[['c', 'e'], ['d', 'e'], ['a', 'e'], ['b', 'e'], ['a', 'b', 'c', 'd']]`, avec la liste des sommets qui est `S='abcde'`.
3. Pour tout $k \in [[1, n]]$, on considère la variable aléatoire D_k égale au degré du sommet s_k. Déterminer la loi de D_k.
4. On dit que s_k est isolé si $D_k = 0$ est réalisé. On note Z_n la variable aléatoire comptant les sommets isolés de G et X_k la variable aléatoire de Bernoulli qui vaut 1 si s_k est isolé et 0 sinon.
a) Exprimer Z_n en fonction de $(X_1, \ldots, X_n)$. En déduire que $E(Z_n) = n(1-p)^{n-1}$.
b) Montrer que : $Z_n^2 = \sum_{k=1}^{n} X_k + 2 \sum_{1 \leq i < j \leq n} X_i X_j$.
c) Justifier que pour tout $i \neq j$, $P([X_i = 1] \cap [X_j = 1]) = (1-p)^{2n-3}$. En déduire que $E\left(Z_n^2\right) = n(1-p)^{n-1} + n(n-1)(1-p)^{2n-3}$.

On suppose désormais que $p = p_n = c.\dfrac{\ln(n)}{n}$, avec $c > 0$, $c \neq 1$.

5.a) écrire une fonction Python Z qui renvoie le nombre de sommets isolés d'un graphe donné par sa liste de listes d'adjacence lst . On importera les bibliothèques numpy as np et numpy.random as rd.

b) On souhaite estimer l'influence de la valeur de c sur le nombre de sommets isolés.

En exécutant le script suivant

```
list2c=[0.3,0.5,0.7,1.3,1.5,1.7];n=1000;res=[]

for c in list2c:
    s=0
    for k in range(200):
        if Z(listAdj(range(1,n+1),c*np.log(n)/n))==0:
            s+=1
    res.append(s/200)
print(res)
```

on obtient [0.0, 0.0, 0.0, 0.91, 0.975, 0.99] après de longues minutes .

Quelle conjecture sur la valeur d'une probabilité pouvez-vous faire lorsque $c < 1$ et $c > 1$? Justifier.

6.a) Montrer que $(1 - p_n)^{n-1} \sim_{+\infty} (1 - p_n)^n$ puis que $(1 - p_n)^{n-1} \sim_{+\infty} n^{-c}$.

b) On rappelle que l'inégalité de Markov affirme que si X est une variable aléatoire positive admettant une espérance et $a > 0$,

$P([X \geq a]) \leq \dfrac{E(X)}{a}$.

Si $c > 1$, en déduire la limite de $P([Z_n \geq 1])$ puis de $P([Z_n = 0])$, lorsque $n \longrightarrow +\infty$.

c) En utilisant l'inégalité de Bienaymé-Tchebychev, montrer que $P([Z_n = 0]) \leq \dfrac{V(Z_n)}{E(Z_n)^2}$.

En déduire que si $c < 1$, $\lim\limits_{n \to +\infty} P([Z_n = 0]) = 0$.

d) Votre conjecture est-elle correcte ?