# The number of small weakly connected components in random directed acyclic graphs

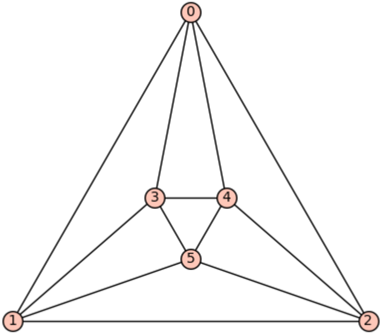
Masreshaw Temere

Stellenbosch University

7 Deccember 2022

# Introduction

A graph is a pair of $G = (V,E)$, where V is a set whose elements are called vertices, and E is a set of pairs of elements of V, whose elements are called edges.



$\mathbb{G}(n, p)$ each pairs of nodes is connected with probability $p$.

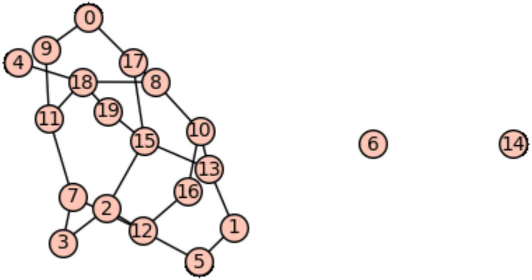
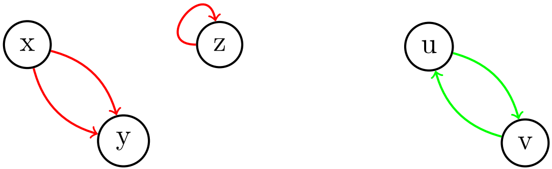



Figure : $n = 20, p = 0.09$

# Directed graphs(digraphs)

A directed graph G = (V, E) consists of a nonempty set of nodes V and a set of directed edges E. Each edge $e$ of E is specified by an ordered pair of vertices $u, v \in V$ . A directed graph is simple if it has no loops and no multiple edges.




The right is forbidden, and the left is allowed.

# Random digraph model

Model $\mathbb{D}(n, p)$: Each of the $n(n-1)$ possible edges occurs independently with probability $p \in [0, 1]$.

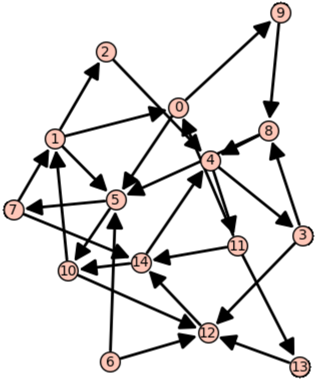
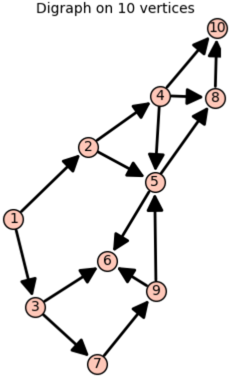
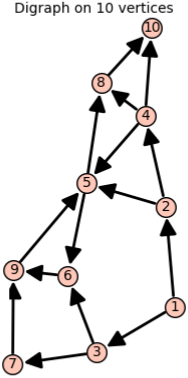



Figure : $n = 15, p = 0.09$

A directed acyclic graph(DAG) is a digraph which has no directed cycles.




(a) A DAG




(b) Not A DAG

The random directed acyclic graph $\mathbb{D}_{ac}(n,p)$ is simply $\mathbb{D}(n,p)$ conditioned to be acyclic. we restricted our selves to the spares case where $p = \frac{\lambda}{n}$, for $\lambda > 0$ is fixed.

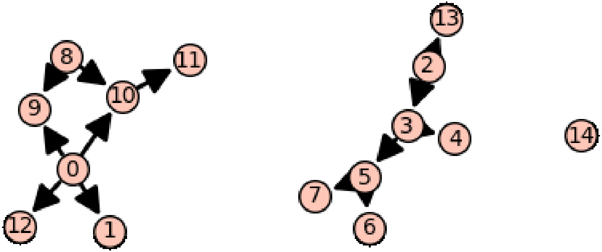



Figure : $n = 15, p = 0.09$

How many connected components of a given size are there in $\mathbb{D}_{ac}(n,p)$?

# Continued

Let $\mathcal{E}$ = Finite class of DAGs
$\delta$ = count the number of occurrences of the elements of $\mathcal{E}$ in $\mathbb{D}_{ac}(n,p)$.

For example, if $\mathcal{E} = \left\{ \textcircled{1} \right\}$, we are counting connected components of order 1. We have a theorem for such connected component.

## Theorem 1 (Naina, (2022))

*Let $\delta(D)$ denote the number of connected component in an acyclic digraph D. Define*

$$\mu^*(\lambda) = \begin{cases} e^{-2\lambda} & \text{if} \quad \lambda < 1 \\ e^{\frac{-(\lambda+1)}{\lambda}} & \text{if} \quad \lambda \geq 1 \end{cases}$$

*and*

$$\sigma^*(\lambda)^2 = \begin{cases} e^{-2\lambda}(1 + (2\lambda - 1)e^{-2\lambda}) & \text{if} \quad \lambda < 1 \\ \lambda^{-1}e^{-(\lambda+1)}(1 + e^{-(\lambda+1)}) & \text{if} \quad \lambda \geq 1 \end{cases}$$

*Then, for a fixed $\lambda > 0$, $\mathbb{E}(\delta(\mathbb{D}_{ac}(n, \lambda/n))) \sim \mu^*(\lambda)n$ as $n \to \infty$. Moreover, we have*

$$\frac{\delta(\mathbb{D}_{ac}(n, \lambda/n)) - \mu^*(\lambda)n}{\sqrt{\sigma^*(\lambda)^2 n}} \xrightarrow{d} \mathcal{N}(0,1)$$

# Generating function of DAGs

If we denote by $e(D)$ the number of edges in a digraph D, then we define

$$A_n(y) = \sum_D y^{e(D)}$$

where the sum is taken over all acyclic digraphs on $[n] = \{1, 2, 3, ..., n\}$. The corresponding exponential generating function is

$$A(x, y) = \sum_{n=0}^{\infty} A_n(y) \frac{x^n}{n!}$$

The coefficients $A_n(y)$ was studied by Robinson(1973), Kassie, Gessel, Chrstina, Xuming(2020)

To include counts on the number of connected components, for an acyclic digraph D, let $\delta(D)$ be the number of connected components from $\mathcal{E}$. Then consider the polynomial

$$A_n(y,u) = \sum_{D} y^{e(D)} u^{\delta(D)}$$

The corresponding exponential generating function is given by

$$A(x,y,u) = \sum_{n=0}^{\infty} A_n(y,u) \frac{x^n}{n!}$$

**Theorem 2**

*We have*

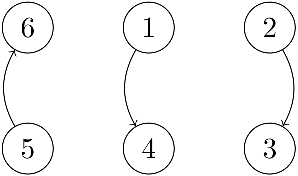
$$\sum_{n=0}^{\infty} A_n(y,u) \frac{x^n}{n!} = e^{(u-1)f(x,y)} \sum_{n=0}^{\infty} A_n(y) \frac{x^n}{n!}$$

*Where*

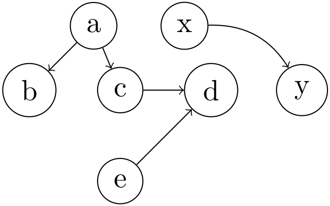
$$f(x,y) = \sum_{D \in \mathcal{E}} y^{e(D)} \frac{x^{|D|}}{|D|!}$$

# Proof

Consider the directed acyclic graphs given below




(a) set($\mathcal{E}$)




(b) DAG

- EGF for the objects on the left: $e^{uf(x,y)}$
- EGF for the objects on the right:

$$\sum_{n=0}^{\infty} A_n(y) \frac{x^n}{n!}$$

The Cartesian product of the two objects

$$
\begin{aligned}
C(x,y,u) &= \sum_{n=0}^{\infty} C_n(y,u)\frac{x^n}{n!} \\
&= set(\mathcal{E}) \times DAG \\
&= e^{uf(x,y)} \sum_{n=0}^{\infty} A_n(y)\frac{x^n}{n!}
\end{aligned}
$$

On the other hand

$$
C(x,y,u) = \sum_{n=0}^{\infty} A_n(y,1+u)\frac{x^n}{n!}
$$

The objects in the cartesian can be regarded DAGs where a subset of the components that come from $\mathcal{E}$ is distinguished..

# Continued

Which implies

$$\sum_{n=0}^{\infty} A_n(y, 1+u)\frac{x^n}{n!} = e^{uf(x,y)} \sum_{n=0}^{\infty} A_n(y)\frac{x^n}{n!}$$

Replacing $u$ by $u-1$

$$\sum_{n=0}^{\infty} A_n(y, u)\frac{x^n}{n!} = e^{(u-1)f(x,y)} \sum_{n=0}^{\infty} A_n(y)\frac{x^n}{n!}$$

# Estimate of the average of connected components of DAGs

The smallest connected component of a DAG is a connected component of order 1.
The expectation of the number of connected component of order 1 in $\mathbb{D}_{ac}(n,p)$ satisfies the asymptotic estimate $\mathbb{E}(I(\mathbb{D}_{ac}(n,\frac{\lambda}{n}))) \sim \mu^*(\lambda)n$ as $n \to \infty$ as stated in Theorem 1

**Lemma 3 (Naina , 2022)**

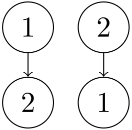
*We have*

$$\mathbb{E}\left(u^{\delta(\mathbb{D}_{ac}(n,p))}\right) = \frac{A_n\left(\frac{p}{1-p}, u\right)}{A_n\left(\frac{p}{1-p}\right)}$$

Therefore,

$$\mathbb{E}(\delta(\mathbb{D}_{ac}(n,p))) = \frac{\partial_u A_n(y,u)|_{u=1}}{A_n(y)}$$

Now, we consider a connected components of DAG of order 2.

$\mathcal{E}= \Big\{$ 

$\Big\}$

Correspondingly the exponential generating function that represents the elements of $\mathcal{E}$ is

$$f(x, y) = yx^2$$

# Continued

Using Theorem 2

$$\sum_{n=0}^{\infty} A_n(y,u)\frac{x^n}{n!} = e^{(u-1)yx^2}\sum_{n=0}^{\infty} A_n(y)\frac{x^n}{n!}$$

Differentiating with respect to $u$ and substituting $u = 1$, we get

$$\begin{aligned}\sum_{n=0}^{\infty} \partial_u A_n(y,u)\frac{x^n}{n!}|_{u=1} &= (yx^2)\sum_{n=0}^{\infty} A_n(y)\frac{x^n}{n!} \\ &= \sum_{n=0}^{\infty} n(n-1)A_{n-2}(y)y\frac{x^n}{n!}\end{aligned}$$

When we compare the coefficient of $x^n$ in the right and left we get

$$\partial_u A_n(y,u)|_{u=1} = yn(n-1)A_{n-2}(y)$$

# Continued

$\delta$ be the random variable that counts the number of occurrences of the elements of $\mathcal{E}$. Then

$$\mathbb{E}(\delta(\mathbb{D}_{ac}(n,p))) = \frac{\partial_u A_n(y,u)|_{u=1}}{A_n(y)}$$
$$= yn(n-1)\frac{A_{n-2}(y)}{A_n(y)}$$

**Lemma 4 (Naina, (2022))**

$$\frac{A_{n-j}(y)}{A_n(y)} = (1+o(1))\mu^*(\lambda)^j \times \begin{cases} e^{\frac{\lambda j^2}{n}} & \textit{if} \quad \lambda < 1 \\ e^{\frac{1}{2}\frac{(\lambda+1)j^2}{n}} & \textit{if} \quad \lambda \geq 1 \end{cases}$$

$$\mathbb{E}(\delta(\mathbb{D}_{ac}(n,p))) = (1+o(1))\mu^*(\lambda)^2 \times \begin{cases} e^{\frac{4\lambda}{n}} & \text{if} \quad \lambda < 1 \\ e^{\frac{1}{2}\frac{2(\lambda+1)}{n}} & \text{if} \quad \lambda \geq 1 \end{cases}$$
$$\sim \lambda n \mu^*(\lambda)^2 \quad \text{as} \quad n \to \infty$$

# Continued

In general, if $\delta(D)$ counts the number of components of order k in D for which $k > 0$ is fixed. We obtain the following theorem.

**Theorem 5**

*Let $\delta(D)$ denote count the number of components of acyclic digraph D with order k. Then for a fixed $\lambda > 0$,*

$$\mathbb{E}(\delta(\mathbb{D}_{ac})(n, \lambda/n)) \sim \frac{k^{k-2}2^{k-1}}{k!}\lambda^{k-1}\mu^*\lambda^k n \quad as \quad n \to \infty.$$

*Moreover, we have*

$$\frac{\delta(\mathbb{D}_{ac}(n, \lambda/n)) - \frac{k^{k-2}2^{k-1}}{k!}\lambda^{k-1}\mu^*\lambda^k n}{\sqrt{\sigma^*(\lambda)^2 n}} \xrightarrow{d} \mathcal{N}(0,1)$$

# References

Ralaivaosaona, Dimbinaina. "The Number of Sources and Isolated Vertices in Random Directed Acyclic Graphs." In 33rd International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA 2022). Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2022.

Thank you!