

# Questioni di Popolarità: Dai Grafi Aleatori alle Leggi di Potenza

## 1 I Limiti dei Modelli Classici di Grafi Aleatori

### 1.1 Il Modello di Erdős-Renyi

Iniziamo con un modello fondamentale per i grafi aleatori: il **modello di Erdős-Renyi**, indicato come  $G_{n,p}$ . In questo modello, un grafo viene costruito su  $n$  nodi e un arco viene inserito tra ogni coppia di nodi distinti in modo indipendente e con una probabilità fissa  $p$ .

Una caratteristica chiave dei grafi generati da questo modello è che il numero di nodi con un grado specifico  $k$  decresce molto rapidamente all'aumentare di  $k$ . Questo decadimento è **esponenziale** in  $k$ . Per capire perché ciò accade, ci rivolgiamo a una pietra miliare della teoria della probabilità.

### 1.2 Il Teorema del Limite Centrale e la Distribuzione dei Gradi

Il rapido decadimento nella distribuzione dei gradi di un grafo di Erdős-Renyi è una conseguenza diretta del **Teorema del Limite Centrale (TLC)**. Un enunciato intuitivo del teorema è il seguente:

*La somma di un gran numero di variabili aleatorie indipendenti e identicamente distribuite sarà approssimativamente distribuita normalmente (cioè, una distribuzione Gaussiana) attorno al suo valore atteso.*

Una caratteristica fondamentale della distribuzione Gaussiana è che le sue "code" decadono in modo estremamente rapido man mano che ci si allontana dalla media.

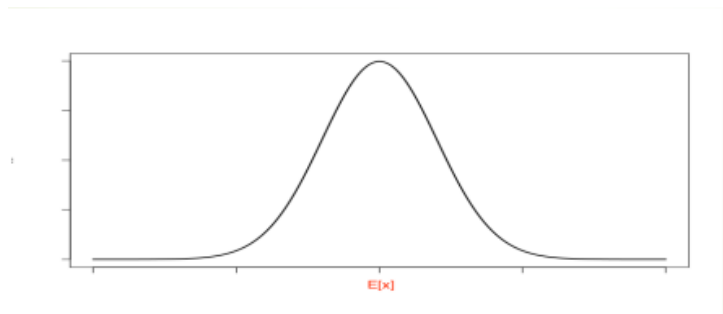


Figure 1: Una distribuzione Gaussiana (o Normale), che decresce molto velocemente allontanandosi dal suo valore atteso  $E[x]$ .

Come si collega questo al grado di un nodo in  $G_{n,p}$ ? Il grado  $\delta_i$  di un nodo specifico  $i$  è il conteggio degli archi ad esso collegati. Possiamo esprimerlo come una somma di

variabili aleatorie:

$$\delta_i = \sum_{j \in [n]-1} e_{ij}$$

dove  $e_{ij}$  è una variabile aleatoria di Bernoulli che vale 1 se l'arco  $(i, j)$  esiste (con probabilità  $p$ ) e 0 altrimenti (con probabilità  $1 - p$ ).

Poiché ogni arco nel modello  $G_{n,p}$  è formato in modo indipendente, il grado  $\delta_i$  è una somma di  $n - 1$  variabili aleatorie indipendenti. Per  $n$  grande, si applica il TLC e la distribuzione dei gradi dei nodi seguirà da vicino una curva Gaussiana. Il punto cruciale è che la probabilità che un nodo abbia un grado molto alto (lontano dalla media) è eccezionalmente piccola a causa del rapido decadimento esponenziale delle code gaussiane. La struttura del grafo già esistente non ha alcuna influenza sulla formazione dell'arco successivo.

## 2 Le Leggi di Potenza nelle Reti del Mondo Reale

### 2.1 Osservare la Realtà: il World Wide Web

Quando analizziamo reti del mondo reale, come il grafo del World Wide Web (dove i nodi sono le pagine e gli archi diretti sono i collegamenti ipertestuali), osserviamo un quadro nettamente diverso. Uno studio di Broder et al. (2000) ha analizzato la distribuzione dei gradi entranti (in-degree) delle pagine web e ha scoperto che non seguiva un decadimento esponenziale.

Hanno invece osservato che la frazione di pagine con un grado entrante  $k$  è proporzionale a  $1/k^c$  per una costante  $c$  invece che a  $1/k^k$ . Questo tipo di funzione, che decresce come l'inverso di un polinomio, è nota come "legge di potenza (power law)". La conseguenza è che il numero di nodi con un grado entrante molto elevato (pagine popolari o "hub") è enormemente maggiore di quanto prevedrebbe il modello di Erdős-Renyi.

### 2.2 Come Riconoscere una Legge di Potenza

Distinguere tra un decadimento esponenziale (come  $e^{-k}$ ) e uno polinomiale (come  $k^{-c}$ ) su un grafico lineare standard può essere difficile, poiché entrambi mostrano una tendenza al ribasso. Il metodo standard per identificare una legge di potenza è utilizzare un **grafico log-log**.

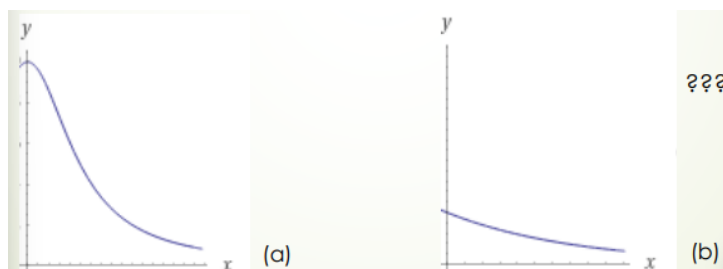
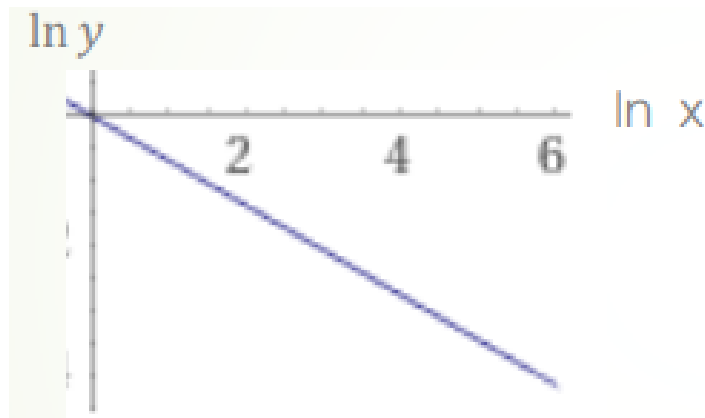


Figure 2: come fai ad accorgerti che quella funzione decresce come l'inverso di un polinomio invece che come l'inverso di una esponenziale?!

Facile: è sufficiente considerare il grafico  $\log - \log$   
 - un grafico i cui assi rappresentano  $\ln(x)$  e  $\ln(y)$

- e invece di rappresentare  $y = f(k)$ , rappresentiamo  $\ln(y) = \ln f(k)$
- dove con  $f(k)$  indichiamo la funzione che esprime il numero di nodi di grado  $k$



- In questo modo, se  $f(k) = \frac{1}{k^c}$ , il grafo sarà, grosso modo una retta:
- $\ln(y) \approx -c \ln(k)$ .
  - infatti: (a) è  $y = \frac{1}{(x^2+1)}$ .

### 3 Il Meccanismo: "The Rich Get Richer"

Il motivo per cui le reti reali si discostano dal modello di Erdős-Renyi è che l'assunzione di indipendenza viene violata. I nuovi archi non si formano con probabilità uniforme. Al contrario, è in gioco un meccanismo noto come **\*\*collegamento preferenziale (preferential attachment)\*\***, o il fenomeno del **\*\*"rich get richer"\*\*\*** (il ricco diventa sempre più ricco).

Consideriamo il processo di creazione di una nuova pagina web. Quando si decide dove aggiungere un collegamento ipertestuale, un creatore è più propenso a collegarsi a una pagina che è già nota, autorevole e popolare. Questa popolarità è spesso segnalata da un alto posizionamento nei motori di ricerca, che a sua volta è determinato dal fatto che la pagina ha molti link in entrata già esistenti. Di conseguenza, le pagine che hanno già un alto grado entrante hanno maggiori probabilità di acquisire nuovi link in entrata, aumentando ulteriormente il loro grado.

La scelta di dove aggiungere un nuovo arco dipende esplicitamente dalla struttura esistente del grafo. Questo stesso fenomeno si verifica nelle reti social (i nuovi utenti tendono a seguire persone già famose) e nelle citazioni accademiche (gli articoli molto citati hanno maggiori probabilità di essere letti e citati di nuovo).

### 4 Il Modello di Barabasi-Albert: Un Modello Generativo per le Leggi di Potenza

Per catturare la dinamica del "rich get richer", abbiamo bisogno di un modello generativo in cui la formazione degli archi dipenda dal grafo esistente. Il **\*\*modello di Barabasi-Albert (BA)\*\*** realizza questo attraverso un semplice processo basato sulla crescita e sul collegamento preferenziale.

## 4.1 Descrizione del Modello

Il grafo viene costruito in modo iterativo. Partiamo con un piccolo insieme iniziale di nodi e archi. Poi, ad ogni passo temporale, aggiungiamo un nuovo nodo e lo colleghiamo ai nodi esistenti. Il modello BA proposto nelle slide utilizza un meccanismo di "copia" che è equivalente al collegamento preferenziale.

Il processo al passo  $i > 2$ :

1. Un nuovo nodo  $i$  viene aggiunto al grafo.
2. Il nodo  $i$  formerà esattamente un arco diretto uscente.
3. Per scegliere la destinazione di questo arco, prima selezioniamo un nodo esistente  $a < i$  uniformemente a caso.
4. Quindi, viene presa una decisione basata su un parametro  $p \in [0, 1]$ :
  - Con probabilità  $p$ , viene scelto uniformemente a caso un nodo  $a < i$ . Il nuovo arco è  $(i, a)$ .
  - Con probabilità  $1 - p$ , viene scelto uniformemente a caso un nodo  $a < i$  e, detto  $(a, b)$  l'arco uscente da  $a$ , allora il nuovo arco è  $(i, b)$ .

Intuitivamente, con il passo di copia (probabilità  $1 - p$ ), un nodo  $b$  che ha un alto grado entrante è la destinazione di molti link. Pertanto, è più probabile che il nodo scelto casualmente  $a$  punti a  $b$ , rendendo  $b$  una destinazione più probabile per il nuovo link da  $i$ . Questo è precisamente l'effetto "rich get richer".

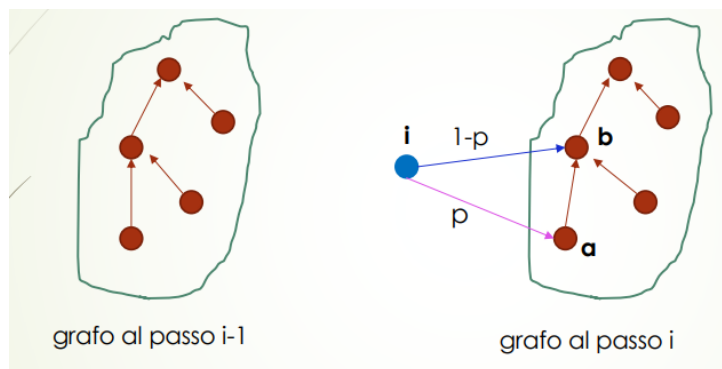


Figure 3: alla creazione del nodo  $i$ :

- viene scelto uniformemente a caso il nodo  $a$
- con probabilità  $p$  viene creato l'arco  $(i, a)$
- con probabilità  $1 - p$  viene creato l'arco  $(i, b)$

Un modo equivalente e più diretto per descrivere il processo è: Con probabilità  $p$ , scegli un nodo di destinazione uniformemente a caso. Con probabilità  $1 - p$ , scegli un nodo di destinazione con una probabilità proporzionale al suo grado entrante esistente.

## 5 Formalizzazione del Modello

### 5.1 La Probabilità di un Arco Entrante

Formalizzazione Abbiamo costruito un modello che descrive il fenomeno **Rich Get Richer**. Resta da verificare se i grafi generati esibiscono una **Power Law**.

Per farlo, introduciamo la variabile aleatoria  $d_{ij}$  per  $i > j$  e  $i \geq 2$ :

$$d_{ij} = \begin{cases} 1 & \text{se } (i, j) \in E \text{ (l'arco esiste)} \\ 0 & \text{altrimenti} \end{cases}$$

La probabilità che si crei l'arco  $(i, j)$ ,  $P(d_{ij} = 1)$ , è la somma di due eventi:

1. La probabilità che venga scelto **direttamente** il nodo  $j$ .
2. La probabilità che venga scelto un nodo  $h$  che già punta a  $j$  (e quindi si **"copi"** il suo arco).

Formula Completa La probabilità è data da:

$$P(d_{ij} = 1) = \frac{p}{i-1} + \frac{1-p}{i-1} \sum_{h:j < h < i} d_{hj}$$

**Interpretazione della Sommatoria:** La sommatoria  $\sum_{h:j < h < i} d_{hj}$  è esattamente il **grado entrante**  $k^{\text{in}}(j, i-1)$  del nodo  $j$  al passo  $i-1$ .

### 5.2 Osservazioni sul Modello

#### 5.2.1 Proprietà Fondamentali

1. **Osservazione 1:** Ha senso calcolare  $P((i, j) \in E)$  solo per  $i > 1$ , poiché il nodo 1 non ha archi uscenti per costruzione.
2. **Osservazione 2:** Ogni nodo  $i > 1$  ha **esattamente un arco uscente**.

Dimostrazione dell'Osservazione 2 (per Induzione) L'obiettivo è dimostrare che  $\sum_{j < i} P((i, j) \in E) = 1$ .

1. **Caso Base ( $i = 2$ ):** Per costruzione, l'arco  $(2, 1)$  è creato  $\implies P((2, 1) \in E) = 1$ .
2. **Ipotesi Induttiva:** Assumiamo che  $\sum_{j < a} d_{aj} = 1$  per ogni  $a < i$ .
3. **Passo Induttivo:** Si dimostra per  $i$ :

$$\begin{aligned} \sum_{j < i} P(d_{ij} = 1) &= \sum_{j < i} \left[ \frac{p}{i-1} + \frac{1-p}{i-1} \sum_{h:j < h < i} d_{hj} \right] \\ &= \sum_{j < i} \frac{p}{i-1} + \sum_{j < i} \frac{1-p}{i-1} \sum_{h:j < h < i} d_{hj} \\ &= (i-1) \frac{p}{i-1} + \frac{1-p}{i-1} \sum_{h < i} \sum_{j < h} d_{hj} \\ &= p + \frac{1-p}{i-1} \sum_{h < i} 1 \quad (\text{Per ipotesi induttiva, } \sum_{j < h} d_{hj} = 1) \\ &= p + \frac{1-p}{i-1} (i-1) = p + (1-p) = 1 \end{aligned}$$

## 5.3 Terza Interpretazione del Modello

### 5.3.1 Il Meccanismo Rich Get Richer

Osservazione 3: Interpretazione della Regola La regola per connettere il nodo  $i$  al nodo preesistente  $j$  può essere interpretata come una scelta tra due meccanismi:

1. Con probabilità  $p$ : Scegliere  $j$  **uniformemente** a caso.
2. Con probabilità  $1 - p$ : Scegliere  $j$  con probabilità **proporzionale al suo grado entrante**  $k^{in}(j, i - 1)$ .

### 5.3.2 Il Fenomeno Rich Get Richer

La formula di probabilità:

$$P(d_{ij} = 1) = \underbrace{\frac{p}{i-1}}_{\text{Scelta Casuale}} + \underbrace{\frac{1-p}{i-1} \cdot k^{in}(j, i-1)}_{\text{Scelta Proporzionale al Grado}}$$

- Con probabilità  $p$ , la scelta è **ignara** della struttura.
- Con probabilità  $1 - p$ , la scelta è **guidata dalla popolarità** (grado entrante).

Un nodo  $j$  con molti link in entrata ("**nodo ricco**") ha una **maggiore probabilità** di essere scelto come destinazione, diventando così ancora più "ricco". Questo è il meccanismo di **collegamento preferenziale** alla base della formazione delle **leggi di potenza** nelle reti reali.

## 6 Analisi Matematica del Modello di Barabasi-Albert

Dimostriamo ora che il modello BA genera grafi con una distribuzione dei gradi secondo una legge di potenza. La dimostrazione procede in quattro passaggi.

### 6.1 Passo 1: Legge Probabilistica per la Variazione del Grado

Sia  $D_j(t)$  la variabile aleatoria per il grado entrante del nodo  $j$  al passo temporale  $t$  ( $t \geq j$ ). La condizione iniziale è  $D_j(j) = 0$ . Al passo  $t + 1$ , il grado di  $j$  aumenta di 1 se e solo se il nuovo nodo  $t + 1$  si collega ad esso. La probabilità di questo evento,  $P(d_{t+1,j} = 1)$ , è:

$$P(\text{link a } j) = P(\text{scelta diretta di } j) + P(\text{copia di un link a } j)$$

Analizziamo questa probabilità al tempo  $t$ . Ci sono  $t - 1$  nodi esistenti tra cui scegliere.

$$P(d_{t,j} = 1) = \underbrace{p \cdot \frac{1}{t-1}}_{\text{Scegli } j \text{ direttamente}} + (1-p) \cdot \underbrace{\sum_{h < t, (h,j) \in E} \frac{1}{t-1}}_{\text{Scegli } h \text{ che punta a } j, \text{ poi copia}}$$

La somma  $\sum_{h < t, (h,j) \in E} 1$  è semplicemente il grado entrante del nodo  $j$  al tempo  $t - 1$ , che denotiamo con  $D_j(t - 1)$ . Quindi, la probabilità che il grado di  $j$  aumenti al passo  $t$  è:

$$P(D_j(t) - D_j(t - 1) = 1) = \frac{p}{t-1} + \frac{1-p}{t-1} D_j(t - 1)$$

Sostituendo le variabili temporali per farle corrispondere alle slide, la probabilità che il grado aumenti al passo  $t + 1$  è:

$$P(D_j(t+1) - D_j(t) = 1) = \frac{p}{t} + \frac{1-p}{t}D_j(t)$$

## 6.2 Passo 2: Approssimazione Continua e Deterministica

L'equazione per  $D_j(t)$  è probabilistica e discreta. Per renderla trattabile, approssimiamo la variabile aleatoria  $D_j(t)$  con una funzione continua e deterministica  $x_j(t)$ . La variazione attesa del grado per passo temporale è approssimata dalla derivata di  $x_j(t)$ :

$$\frac{dx_j(t)}{dt} = \frac{p}{t} + \frac{1-p}{t}x_j(t)$$

Questa è un'equazione differenziale lineare del primo ordine con la condizione iniziale  $x_j(j) = 0$ .

## 6.3 Passo 3: Risoluzione dell'Equazione Differenziale

Possiamo risolvere questa equazione separando le variabili.

$$\begin{aligned}\frac{dx_j}{dt} &= \frac{1}{t}[p + (1-p)x_j(t)] \\ \frac{dx_j}{p + (1-p)x_j} &= \frac{dt}{t}\end{aligned}$$

Ora, integriamo entrambi i lati dallo stato iniziale  $(j, x_j(j) = 0)$  a uno stato generico  $(t, x_j(t))$ :

$$\int_0^{x_j(t)} \frac{dx}{p + (1-p)x} = \int_j^t \frac{d\tau}{\tau}$$

Per risolvere l'integrale a sinistra, poniamo  $u = p + (1-p)x$ , quindi  $du = (1-p)dx$ .

$$\begin{aligned}\frac{1}{1-p} \int_p^{p+(1-p)x_j(t)} \frac{du}{u} &= [\ln(\tau)]_j^t \\ \frac{1}{1-p} [\ln(u)]_p^{p+(1-p)x_j(t)} &= \ln(t) - \ln(j) \\ \ln(p + (1-p)x_j(t)) - \ln(p) &= (1-p) \ln\left(\frac{t}{j}\right) \\ \ln\left(\frac{p + (1-p)x_j(t)}{p}\right) &= \ln\left(\left(\frac{t}{j}\right)^{1-p}\right)\end{aligned}$$

Esponenziando entrambi i lati si ottiene:

$$\frac{p + (1-p)x_j(t)}{p} = \left(\frac{t}{j}\right)^{1-p}$$

Infine, risolvendo per  $x_j(t)$ :

$$\begin{aligned}p + (1-p)x_j(t) &= p \left(\frac{t}{j}\right)^{1-p} \\ x_j(t) &= \frac{p}{1-p} \left[ \left(\frac{t}{j}\right)^{1-p} - 1 \right]\end{aligned}$$

## 6.4 Passo 4: Derivazione della Legge di Potenza

Usiamo ora la soluzione  $x_j(t)$  per trovare la frazione di nodi che hanno grado  $k$  a un tempo grande  $t$ . Per prima cosa, troviamo l'insieme di nodi  $A_t(k) = \{j \leq t : x_j(t) \geq k\}$ . Un nodo  $j$  appartiene a questo insieme se:

$$\frac{p}{1-p} \left[ \left( \frac{t}{j} \right)^{1-p} - 1 \right] \geq k$$

Riorganizzando la disuguaglianza per risolvere per  $j$ :

$$\begin{aligned} \left( \frac{t}{j} \right)^{1-p} &\geq \frac{k(1-p)}{p} + 1 \\ \frac{t}{j} &\geq \left( \frac{1-p}{p}k + 1 \right)^{\frac{1}{1-p}} \\ j &\leq t \left( \frac{1-p}{p}k + 1 \right)^{-\frac{1}{1-p}} \end{aligned}$$

Il numero di nodi con grado almeno  $k$ , denotato con  $|A_t(k)|$ , è il più grande intero  $j$  che soddisfa questa condizione. Possiamo approssimare questo conteggio come:

$$|A_t(k)| \approx t \left( \frac{1-p}{p}k + 1 \right)^{-\frac{1}{1-p}}$$

Il numero di nodi con grado \*esattamente\*  $k$  è  $N_t(k) = |A_t(k)| - |A_t(k+1)|$ . La frazione di tali nodi è  $\frac{N_t(k)}{t}$ . Per un'approssimazione continua, questa differenza può essere rappresentata dal valore assoluto della derivata:

$$\frac{N_t(k)}{t} \approx -\frac{1}{t} \frac{d|A_t(k)|}{dk}$$

Calcoliamo la derivata:

$$\begin{aligned} \frac{d|A_t(k)|}{dk} &= t \cdot \left( -\frac{1}{1-p} \right) \left( \frac{1-p}{p}k + 1 \right)^{-\frac{1}{1-p}-1} \cdot \left( \frac{1-p}{p} \right) \\ \frac{d|A_t(k)|}{dk} &= -\frac{t}{p} \left( \frac{1-p}{p}k + 1 \right)^{-\frac{2-p}{1-p}} \end{aligned}$$

Pertanto, la frazione di nodi con grado  $k$  è:

$$\frac{N_t(k)}{t} \approx \frac{1}{p} \left( \frac{1-p}{p}k + 1 \right)^{-\frac{2-p}{1-p}}$$

Per valori grandi di  $k$ , questa funzione è proporzionale a  $k^{-c}$ , dove l'esponente  $c$  è:

$$c = \frac{2-p}{1-p} = \frac{1+(1-p)}{1-p} = 1 + \frac{1}{1-p}$$

Questa è una \*\*legge di potenza\*\*. Successive analisi rigorose [Bollobas, Riordan - 2005] hanno confermato che questo risultato dell'approssimazione continua descrive accuratamente il comportamento del modello aleatorio discreto.



## 7 Implicazioni e Fenomeni Correlati

### 7.1 Imprevedibilità della Popolarità

La dinamica del "rich get richer" implica che le gare di popolarità sono estremamente sensibili alle condizioni iniziali. Un esperimento di Salganik, Dodds e Watts (2006) ha illustrato magnificamente questo concetto. Hanno creato un sito web per il download di musica con 48 canzoni sconosciute, ma hanno suddiviso i visitatori in 8 "mondi" indipendenti. In ogni mondo, i visitatori potevano vedere quante volte ogni canzone era stata scaricata dagli altri nel loro mondo.

Anche se tutti i mondi partivano in modo identico, canzoni diverse sono diventate dei successi in mondi diversi. Le scelte casuali iniziali di pochi utenti sono state amplificate dal feedback sociale, portando a risultati divergenti e imprevedibili. Il successo di una canzone non dipendeva solo dalla sua qualità intrinseca, ma anche dalle dinamiche dipendenti dal percorso del suo mondo.

### 7.2 La Lunga Coda

La distribuzione della popolarità può avere importanti effetti sulle attività commerciali, in particolar modo sui "magazzini virtuali" (società di vendita on-line che dispongono di magazzini molto vasti).

Il punto è: uno di detti venditori (diciamo, una libreria on-line) guadagna di più vendendo tante copie di pochi libri molto popolari oppure vendendo poche copie di tantissimi libri poco popolari?

Invece che interessarci a quanti sono i nodi di grado  $k$ , qui ci interessa ora sapere se, complessivamente, il **volume di affari** derivante dai nodi poco popolari (quelli che hanno grado basso, che sono tanti) equivale al volume di affari dei nodi molto popolari (quelli che hanno grado elevato, che sono pochi).

#### 7.2.1 Costruzione del Grafico e Trasformazione

Per gestire questa questione, si ordinano i nodi per popolarità non crescente, creando una "hit parade" dei nodi.

- Lungo l'asse delle **ascisse** si elencano i nodi, ordinati per popolarità non crescente (il primo nodo è il più popolare, seguito dal secondo, e così via).
- Lungo l'asse delle **ordinate** si elencano le popolarità dei nodi (il grado del nodo).

I nodi poco popolari costituiscono la **coda** del grafico. Il volume di affari che ci interessa è, grosso modo, l'area della regione compresa fra la coda del grafico e l'asse  $x$ .

#### 7.2.2 Trasformazione

Invece di elencare tutti i nodi uno per uno sull'asse delle ascisse, si elenca con un solo punto tutti i nodi che hanno la stessa popolarità. L'ascissa di tale punto è il **numero di nodi** che hanno quella popolarità (quantità di nodi).

Pensandoci bene, questo grafico è, grosso modo, ottenuto dal grafico che esprime il numero  $n$  di nodi che hanno grado  $k$  scambiando gli assi coordinati. Se  $d(k)$  è la funzione che esprime il numero di nodi di grado  $k$ , allora il grafico che ci occorre rappresenta, grosso modo, la funzione  $k = d^{-1}(n)$ , ammesso che  $d$  sia una funzione invertibile.

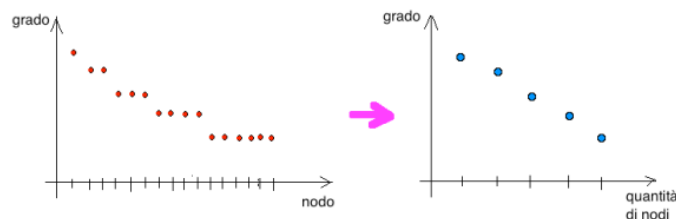


Figure 4: naturalmente, per compiere questa operazione dobbiamo assumere che a gradi di popolarità diversi corrispondano numeri di nodi diversi che hanno quelle popolarità ' ossia, non deve accadere che, ad esempio, 5 nodi abbiano popolarità  $p1$  e 5 nodi abbiano popolarità  $p2 \neq p1$

### 7.2.3 Individuazione della Power Law e Lunga Coda

È possibile dimostrare che:

- Se, come appare dai dati sperimentali, la funzione  $f$  che esprime la frazione del numero di nodi che hanno grado  $k$  è una **Power Law** ( $f(k) \approx k^{-c}$  per qualche costante  $c > 0$ ).
- Allora anche  $d$  esprime una Power Law ( $d(k) \approx k^{-a}$  per qualche costante  $a > 0$ ).

Se poniamo  $n = d(k) \approx k^{-a}$ , possiamo invertire la relazione per ottenere  $k$  in funzione di  $n$ :

$$k \approx n^{-1/a}$$

Questa funzione **decresce molto lentamente**, il che significa che il grado  $k$  non scende rapidamente anche quando si considerano nodi con un rango di popolarità  $n$  elevato.

Questo lento decadimento è ciò che definisce la **lunga coda**. Conseguentemente, il volume di affari derivante dai nodi poco popolari (l'area della regione compresa fra la coda del grafico e l'asse delle ascisse) è tutt'altro che trascurabile.

## 7.3 Il Ruolo della Ricerca Web

I motori di ricerca web hanno una relazione complessa e duplice con il fenomeno del "rich get richer".

- **Amplificazione:** Gli algoritmi di ranking spesso utilizzano il numero di link (popolarità) come segnale primario di qualità. Mostrando in primo piano le pagine popolari, indirizzano più traffico verso di esse, il che a sua volta porta a più link, amplificando il successo delle pagine già popolari.
- **Mitigazione:** La grande diversità delle query degli utenti forma una "coda lunga della ricerca". Query molto specifiche, "di nicchia", possono condurre gli utenti a pagine meno popolari che sono altamente pertinenti alla loro specifica esigenza. Questo fornisce un percorso per scoprire contenuti al di fuori dei successi principali, mitigando così il dominio dei nodi più popolari.

Strumenti moderni come i sistemi di raccomandazione sono esplicitamente progettati per sfruttare la coda lunga, utilizzando i dati degli utenti per indirizzarli verso prodotti di nicchia che potrebbero gradire, modellando e navigando così i cicli di feedback della popolarità.