Karlsruhe Institute of Technology

Probability & Computing

Continuous Probability Spaces & Random Geometric Graphs

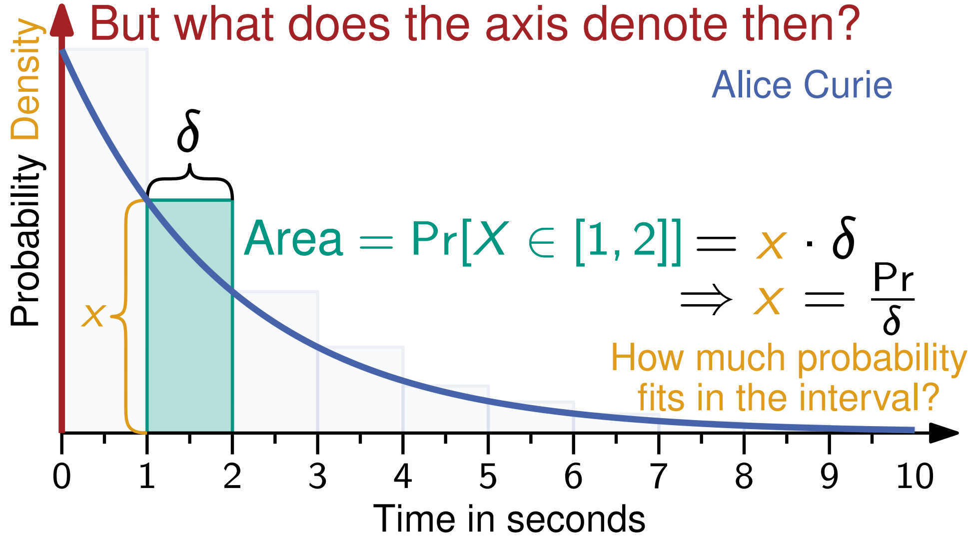
Motivation – Radioactive Decay

- Two physicists study radioactive material that emits particles every now and then
- Both compete to get the most accurate model describing the emission
- "We could do this forever!" Could they really?
- They measure with infinite precision...
  - What is $Pr[X = 2.71828182846]$?
  - What is $Pr[X = 2.7182818284\mathbf{7}]$?

  $\Big\} > 0?$ Emission could happen at any time...
  - But then the "sum" over uncountably infinite non-zero values is ∞ This is not a probability distribution!
- For continuous spaces we need to adjust how we measure probabilities

  We assign probabilities to *intervals* instead of individual values!

  The probability is the *area* of the bar, *not* the height
- As bars get thinner, areas (probabilities) decrease
- We describe distributions using probability density functions



But what does the axis denote then?

$$\text{Area} = \Pr[X \in [1,2]] = x \cdot \delta$$
$$\Rightarrow x = \frac{\Pr}{\delta}$$

How much probability fits in the interval?

youtube.com/watch?v=ZA4JkHKZM50

Working in Continuous Probability Spaces

Discrete Random Variable X

- Cumulative distribution function

  $F_X(x) = \Pr[X \leq x] = \sum_{y \leq x} f_X(y)$
- Probability mass function

  $f_X(x) = \Pr[X = x] \geq 0$ $\quad \sum_x \Pr[X=x]=1$
- Expectation

  $\mathbb{E}[X] = \sum_x x \cdot \Pr[X = x]$

Continuous Random Variable X

- Cumulative distribution function

  $F_X(x) = \Pr[X \leq x] = \int_{-\infty}^{x} f_X(y)\mathrm{d}y$
- Probability density function

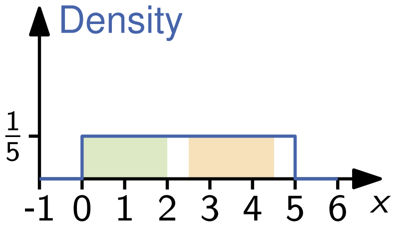
  $f_X(x) \geq 0$ $\quad \int_{-\infty}^{\infty} f_X(x)\mathrm{d}x = 1$
- Expectation

  $\mathbb{E}[X] = \int x \cdot f_X(x)\mathrm{d}x$

Example: Uniform Distribution ⟶ Over $[0, 5]$

- You build a fence that is at least 2m tall at each point
- In the hardware store they have 5m planks
- The staff member cutting your planks wears hearing protection and cuts uniformly at random
- What is the probability that you get two $\geq$ 2m boards out of one 5m plank?

$$f_X(x) = \begin{cases} \frac{1}{5}, \text{if } x \in [0, 5] \\ 0, \text{o.w.} \end{cases}$$



$$\int_{-\infty}^{\infty} f_X(x)\mathrm{d}x = \int_0^5 \frac{1}{5}\mathrm{d}x = \left[\frac{x}{5}\right]_0^5 = 1 \checkmark$$

$$\int_a^b f_X(x)\mathrm{d}x = \left[\frac{x}{5}\right]_a^b = \frac{1}{5}(b - a) \checkmark$$

for $a \leq b \in [0, 5]$

Working in Continuous Probability Spaces

Discrete Random Variable X

- Cumulative distribution function

$$F_X(x) = \Pr[X \leq x] = \sum_{y \leq x} f_X(y)$$

- Probability mass function

$$f_X(x) = \Pr[X = x] \geq 0$$

$\sum_x \Pr[X = x] = 1$

- Expectation

$$\mathbb{E}[X] = \sum_x x \cdot \Pr[X = x]$$

Continuous Random Variable X

- Cumulative distribution function

$$F_X(x) = \Pr[X \leq x] = \int_{-\infty}^{x} f_X(y)\mathrm{d}y$$

- Probability density function

$$f_X(x) \geq 0$$

$\int_{-\infty}^{\infty} f_X(x)\mathrm{d}x = 1$

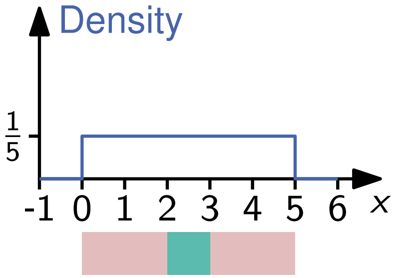
- Expectation

$$\mathbb{E}[X] = \int x \cdot f_X(x)\mathrm{d}x$$

Example: Uniform Distribution → Over $[0, 5]$

- You build a fence that is at least 2m tall at each point
- In the hardware store they have 5m planks
- The staff member cutting your planks wears hearing protection and cuts uniformly at random
- What is the probability that you get two $\geq$ 2m boards out of one 5m plank?

$$f_X(x) = \begin{cases} \frac{1}{5}, \text{ if } x \in [0, 5] \\ 0, \text{ o.w.} \end{cases}$$



$$\Pr[X \in [2, 3]] = \Pr[X \leq 3] - \Pr[X \leq 2]$$
$$= \int_0^3 \frac{1}{5}\mathrm{d}x - \int_0^2 \frac{1}{5}\mathrm{d}x$$
$$= \left[\frac{x}{5}\right]_0^3 - \left[\frac{x}{5}\right]_0^2 = \frac{3}{5} - \frac{2}{5} = \frac{1}{5}$$ ✓

Working in Continuous Probability Spaces

Discrete Random Variable X

- Cumulative distribution function
  $F_X(x) = \Pr[X \leq x] = \sum_{y \leq x} f_X(y)$
- Probability mass function
  $f_X(x) = \Pr[X = x] \geq 0$ $\sum_x \Pr[X=x]=1$
- Expectation
  $\mathbb{E}[X] = \sum_x x \cdot \Pr[X = x]$

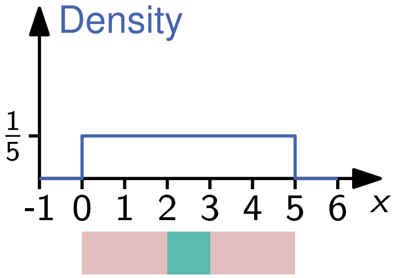
Continuous Random Variable X

- Cumulative distribution function
  $F_X(x) = \Pr[X \leq x] = \int_{-\infty}^{x} f_X(y)\mathrm{d}y$
- Probability density function
  $f_X(x) \geq 0$ $\int_{-\infty}^{\infty} f_X(x)\mathrm{d}x = 1$
- Expectation
  $\mathbb{E}[X] = \int x \cdot f_X(x)\mathrm{d}x$

Example: Uniform Distribution → Over $[0, 5]$

- You build a fence that is at least 2m tall at each point
- In the hardware store they have 5m planks
- The staff member cutting your planks wears hearing protection and cuts uniformly at random
- What is the probability that you get two $\geq$ 2m boards out of one 5m plank?

$$f_X(x) = \begin{cases} \frac{1}{5}, \text{if } x \in [0,5] \\ 0, \text{o.w.} \end{cases}$$

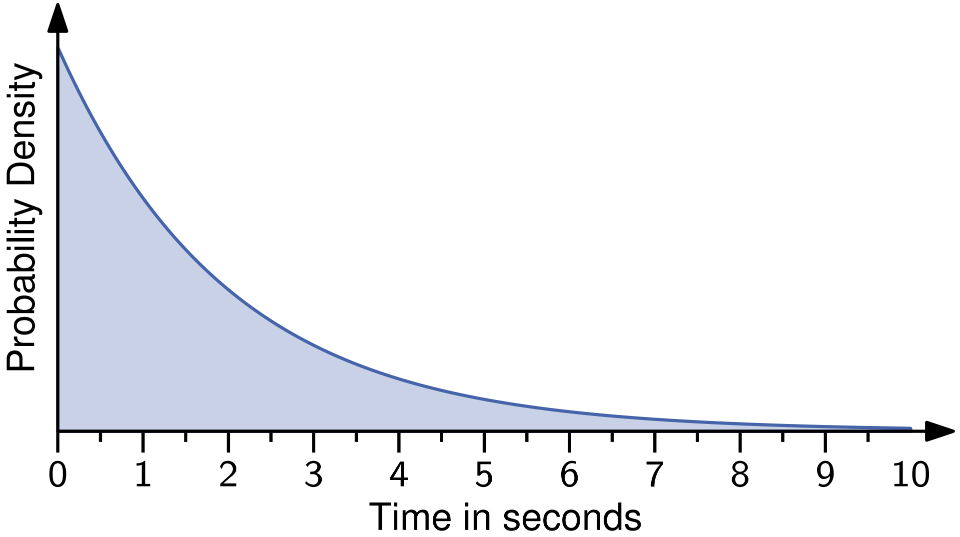


- In general: $X \sim \mathcal{U}([a, b])$
  $\Pr[X \in [c, d] \subseteq [a, b]] = \frac{d-c}{b-a}$

Example: Radioactive Decay

Exponential Distribution $X \sim \text{Exp}(\lambda)$

- “Rate” parameter $\lambda > 0$
- Continuous equivalent to geometric distribution
- “Time until first success”
- Probability density function $f_X(x) = \begin{cases} \lambda e^{-\lambda x}, \text{if } x \geq 0 \\ 0, \text{o.w.} \end{cases}$
- Cumulative distribution function
  $F_X(x) = \int_{-\infty}^{x} f_X(y)\mathrm{d}y = 1 - e^{-\lambda x}$



Integration by Parts
$\int uv'\mathrm{d}x = uv - \int u'v\mathrm{d}x$

Characterization via Moments (n-th moment: $\mathbb{E}[X^n]$)

- $\mathbb{E}[X] = \int_{-\infty}^{\infty} x \cdot f_X(x)\mathrm{d}x = \lambda \int_0^{\infty} xe^{-\lambda x}\mathrm{d}x$

$$= \lambda \left(\left[x \cdot \frac{1}{-\lambda} e^{-\lambda x} \right]_0^{\infty} - \int_0^{\infty} \frac{1}{-\lambda} e^{-\lambda x} \cdot 1\mathrm{d}x \right)$$

$$= \cancel{\lambda} \left(\frac{1}{\cancel{\lambda}} \left[xe^{-\lambda x} \right]_{\infty}^{0} + \frac{1}{\cancel{\lambda}} \int_0^{\infty} e^{-\lambda x}\mathrm{d}x \right)$$

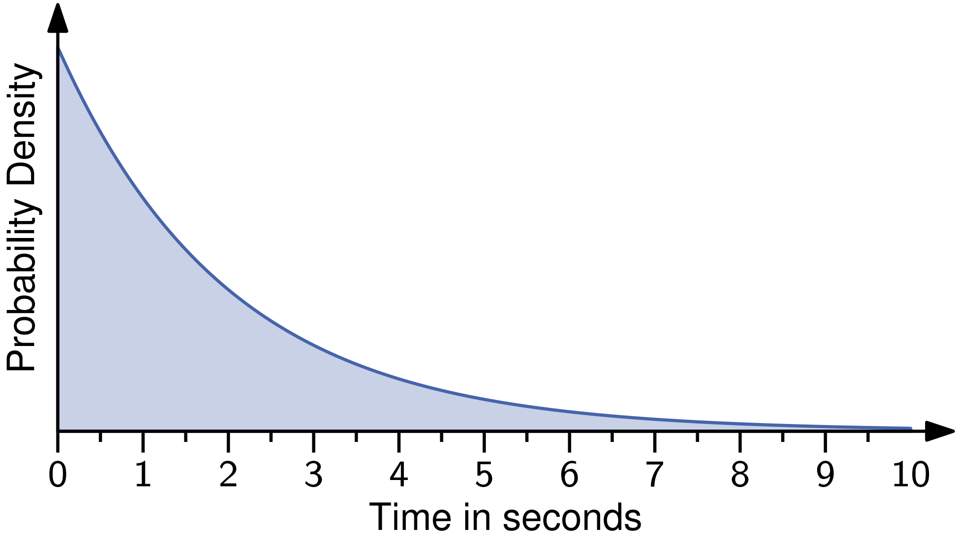
$$= 0 + 0 + \frac{1}{-\lambda} \left[e^{-\lambda x} \right]_0^{\infty} = \frac{1}{\lambda} \left[e^{-\lambda x} \right]_{\infty}^{0} = \frac{1}{\lambda}[1 - 0] = \frac{1}{\lambda}$$

Example: Radioactive Decay

Exponential Distribution $X \sim \text{Exp}(\lambda)$

- "Rate" parameter $\lambda > 0$
- Continuous equivalent to geometric distribution
- "Time until first success"
- Probability density function $f_X(x) = \begin{cases} \lambda e^{-\lambda x}, \text{if } x \geq 0 \\ 0, \text{o.w.} \end{cases}$
- Cumulative distribution function

  $F_X(x) = \int_{-\infty}^{x} f_X(y)\mathrm{d}y = 1 - e^{-\lambda x}$



Integration by Parts
$\int uv'\mathrm{d}x = uv - \int u'v\mathrm{d}x$

Characterization via Moments (n-th moment: $\mathbb{E}[X^n]$)

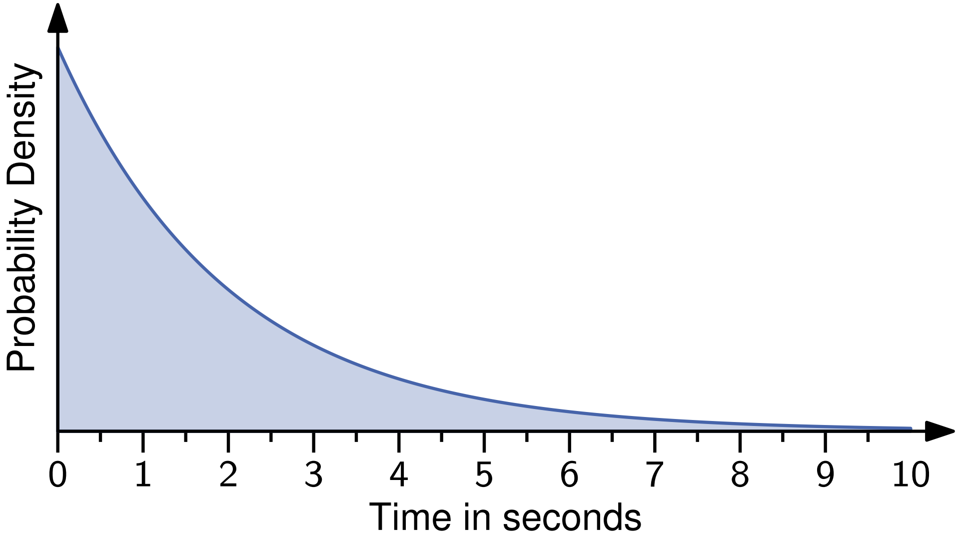
- $\mathbb{E}[X] = \int_{-\infty}^{\infty} x \cdot f_X(x)\mathrm{d}x = \lambda \int_0^\infty xe^{-\lambda x}\mathrm{d}x = \frac{1}{\lambda}$
- $\mathbb{E}[X^2] = \int_{-\infty}^{\infty} x^2 \cdot f_X(x)\mathrm{d}x = \lambda \int_0^\infty x^2 e^{-\lambda x}\mathrm{d}x$

$$= \lambda \left(\left[x^2 \frac{1}{-\lambda} e^{-\lambda x}\right]_0^\infty - \frac{2}{-\lambda} \overbrace{\int_0^\infty x \cdot e^{-\lambda x}\mathrm{d}x}^{\frac{1}{\lambda^2}} \right) = \lambda([0+0] + \frac{2}{\lambda^3}) = \frac{2}{\lambda^2}$$

Example: Radioactive Decay

Exponential Distribution $X \sim \mathrm{Exp}(\lambda)$

- “Rate” parameter $\lambda > 0$
- Continuous equivalent to geometric distribution
- “Time until first success”
- Probability density function $f_X(x) = \begin{cases} \lambda e^{-\lambda x}, \text{if } x \geq 0 \\ 0, \text{o.w.} \end{cases}$
- Cumulative distribution function
  $F_X(x) = \int_{-\infty}^{x} f_X(y)\mathrm{d}y = 1 - e^{-\lambda x}$



Integration by Parts
$\int uv'\mathrm{d}x = uv - \int u'v\mathrm{d}x$

Characterization via Moments (n-th moment: $\mathbb{E}[X^n]$)

- $\mathbb{E}[X] = \int_{-\infty}^{\infty} x \cdot f_X(x)\mathrm{d}x = \lambda \int_{0}^{\infty} xe^{-\lambda x}\mathrm{d}x = \frac{1}{\lambda}$
- $\mathbb{E}[X^2] = \int_{-\infty}^{\infty} x^2 \cdot f_X(x)\mathrm{d}x = \lambda \int_{0}^{\infty} x^2 e^{-\lambda x}\mathrm{d}x = \frac{2}{\lambda^2}$
- $\mathrm{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{2}{\lambda^2} - (\frac{1}{\lambda})^2 = \frac{1}{\lambda^2}$

Exponential Distribution: Memorylessness

Motivation

- What is the probability of having to wait longer than an additional time $s > 0$ after already having waited time $t > 0$?

$$\Pr[X > s+t \mid X > t] = \frac{\Pr[X > s+t \wedge X > t]}{\Pr[X > t]} \qquad X > s+t \Rightarrow X > t$$

$$= \frac{\Pr[X > s+t]}{\Pr[X > t]} = \frac{1 - \Pr[X \leq s+t]}{1 - \Pr[X \leq t]}$$

$$= \frac{e^{-\lambda(s+t)}}{e^{-\lambda t}} = e^{-\lambda s} = \Pr[X > s]$$

$X \sim \mathrm{Exp}(\lambda)$

$f_X(x) = \lambda e^{-\lambda x}$

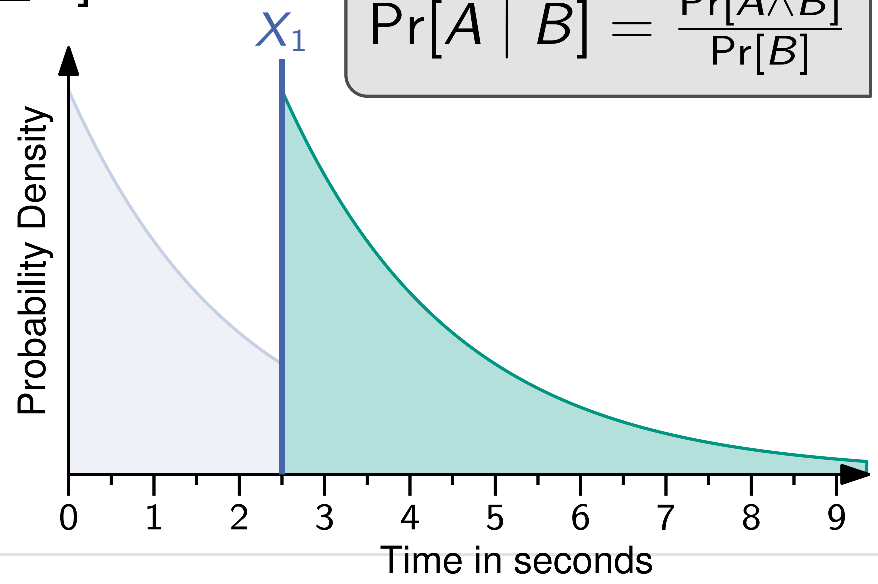
$F_X(x) = 1 - e^{-\lambda x}$

$$\Pr[A \mid B] = \frac{\Pr[A \wedge B]}{\Pr[B]}$$

- No matter how long we already waited, waiting time is distributed as if we just started

Observing Multiple Particles

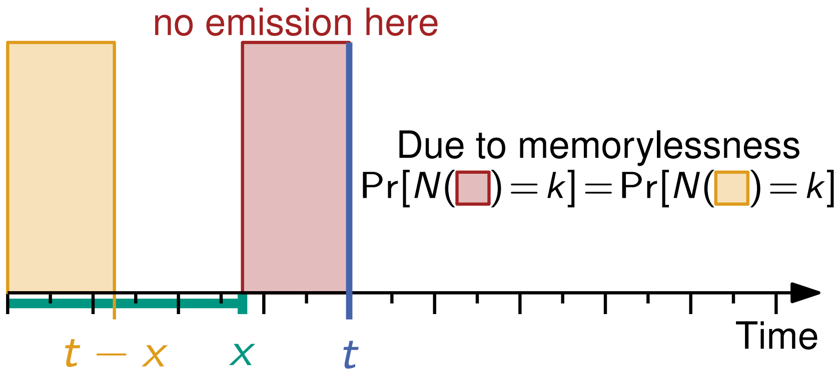
- How long do we have to wait for the second particle after having just seen the first?



Counting Decays

Motivation

- Count number of particles emitted within a given time t
- Let $X_1, X_2, X_3, \ldots \sim \mathrm{Exp}(\lambda)$ be independent waiting times
- Let $N(a, b)$ be the number of emissions in $[a, b]$
- Let $N_t = N(0, t)$ be the number of emissions until t



Specific Values

Law of Total Probability: $\Pr[A] = \int_{-\infty}^{\infty} \Pr[A \mid X = x] \cdot f_X(x)\mathrm{d}x$

$\boldsymbol{X \sim \mathrm{Exp}(\lambda)}$

$f_X(x) = \lambda e^{-\lambda x}\mathbb{1}_{x \geq 0}$

$F_X(x) = 1 - e^{-\lambda x}$

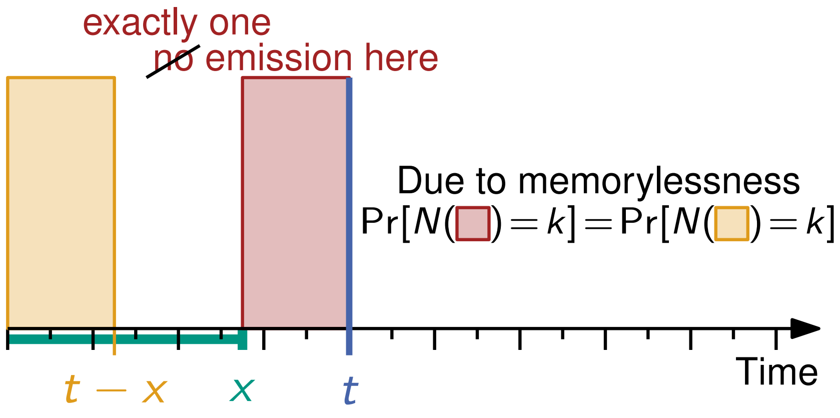
$$\Pr[N_t = 0] = e^{-\lambda t}$$

$$\begin{aligned}
\Pr[N_t = 1] &= \int_{-\infty}^{\infty} \Pr[X_1 \leq t \wedge N(x,t) = 0 \mid X_1 = x] f_{X_1}(x)\mathrm{d}x \\
&= \int_{-\infty}^{\infty} \Pr[X_1 \leq t \wedge N(x,t) = 0 \mid X_1 = x] \lambda e^{-\lambda x}\mathbb{1}_{x \geq 0}\mathrm{d}x \\
&= \int_0^t \Pr[\cancel{X_1 \leq t} \wedge N(x,t) = 0 \mid X_1 = x] \lambda e^{-\lambda x}\cancel{\mathbb{1}_{x \geq 0}}\mathrm{d}x \\
&= \int_0^t \Pr[N(x,t) = 0 \mid \cancel{X_1 = x}] \lambda e^{-\lambda x}\mathrm{d}x \\
&= \int_0^t \Pr[N(x,t) = 0] \lambda e^{-\lambda x}\mathrm{d}x = \int_0^t \overbrace{\Pr[N_{t-x} = 0]}^{e^{-\lambda(t-x)}} \lambda e^{-\lambda x}\mathrm{d}x = \int_0^t e^{-\lambda(t-x)} \cdot \lambda e^{-\lambda x}\mathrm{d}x \\
&= \lambda e^{-\lambda t} \int_0^t 1\mathrm{d}x = \lambda t e^{-\lambda t}
\end{aligned}$$

Counting Decays

Motivation

- Count number of particles emitted within a given time t
- Let $X_1, X_2, X_3, \ldots \sim \text{Exp}(\lambda)$ be independent waiting times
- Let $N(a, b)$ be the number of emissions in $[a, b]$
- Let $N_t = N(0, t)$ be the number of emissions until t



Due to memorylessness
$\Pr[N(🟥) = k] = \Pr[N(🟨) = k]$

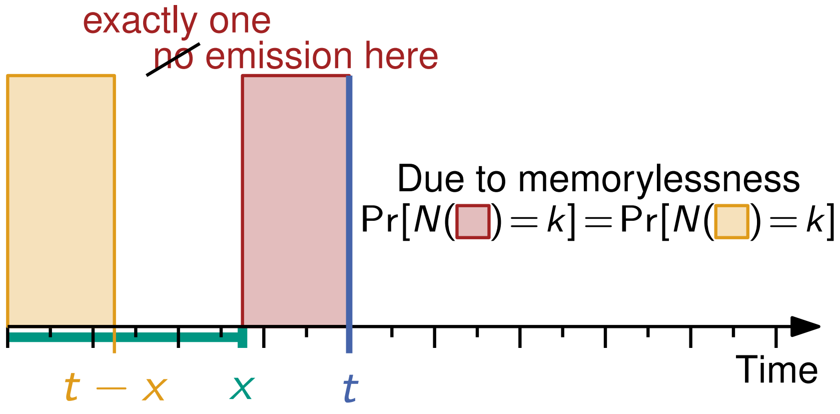
Specific Values

Law of Total Probability: $\Pr[A] = \int_{-\infty}^{\infty} \Pr[A \mid X = x] \cdot f_X(x)\mathrm{d}x$

$\boldsymbol{X \sim \text{Exp}(\lambda)}$
$f_X(x) = \lambda e^{-\lambda x}\mathbb{1}_{x \geq 0}$
$F_X(x) = 1 - e^{-\lambda x}$

$$\Pr[N_t = 0] = e^{-\lambda t} \quad \Pr[N_t = 1] = \lambda t e^{-\lambda t} \quad \Pr[N_t = 2] = \lambda^2 e^{-\lambda t} \cdot \frac{1}{2} t^2$$

Counting Decays



Due to memorylessness
$\Pr[N(🟥)=k]=\Pr[N(🟨)=k]$

Motivation

- Count number of particles emitted within a given time t
- Let $X_1, X_2, X_3, \dots \sim \mathrm{Exp}(\lambda)$ be independent waiting times
- Let $N(a, b)$ be the number of emissions in $[a, b]$
- Let $N_t = N(0, t)$ be the number of emissions until t

Specific Values

Law of Total Probability: $\Pr[A] = \int_{-\infty}^{\infty} \Pr[A \mid X = x] \cdot f_X(x)\mathrm{d}x$

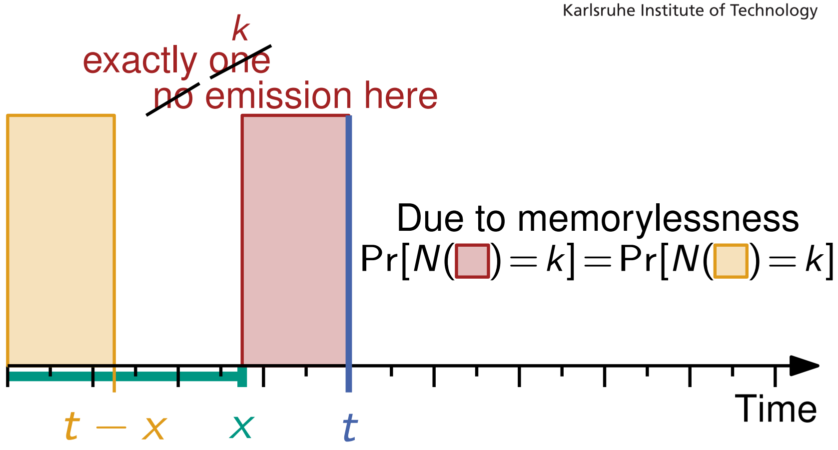
$\boldsymbol{X \sim \mathrm{Exp}(\lambda)}$
$f_X(x) = \lambda e^{-\lambda x}\mathbb{1}_{x \geq 0}$
$F_X(x) = 1 - e^{-\lambda x}$

$$\Pr[N_t = 0] = \underbrace{e^{-\lambda t}}_{\frac{(\lambda t)^0 e^{-\lambda t}}{0!}} \quad \Pr[N_t = 1] = \underbrace{\lambda t e^{-\lambda t}}_{\frac{(\lambda t)^1 e^{-\lambda t}}{1!}} \quad \Pr[N_t = 2] = \underbrace{\lambda^2 e^{-\lambda t} \cdot \frac{1}{2} t^2}_{\frac{(\lambda t)^2 e^{-\lambda t}}{2!}}$$

Counting Decays

Motivation

- Count number of particles emitted within a given time t
- Let $X_1, X_2, X_3, \ldots \sim \text{Exp}(\lambda)$ be independent waiting times
- Let $N(a, b)$ be the number of emissions in $[a, b]$
- Let $N_t = N(0, t)$ be the number of emissions until t



Due to memorylessness
Pr[N(🟥) = k] = Pr[N(🟨) = k]

Law of Total Probability: $\Pr[A] = \int_{-\infty}^{\infty} \Pr[A \mid X = x] \cdot f_X(x)\mathrm{d}x$

$X \sim \text{Exp}(\lambda)$

$f_X(x) = \lambda e^{-\lambda x}\mathbb{1}_{x \geq 0}$

$F_X(x) = 1 - e^{-\lambda x}$

Specific Values

$\Pr[N_t = 0] = e^{-\lambda t}$ $\Pr[N_t = 1] = \lambda t e^{-\lambda t}$ $\Pr[N_t = 2] = \lambda^2 e^{-\lambda t} \cdot \frac{1}{2} t^2$

General Form $\Pr[N_t = k] = \frac{(\lambda t)^k e^{-\lambda t}}{k!}$ *(proof via induction)*

↳ $N_t \sim \text{Pois}(\lambda t)$

Integration by Substitution $u = g(x)$

$$\int_a^b f(g(x))\mathrm{d}x = \int_{g(a)}^{g(b)} \frac{f(u)}{\left(\frac{\mathrm{d}g(x)}{\mathrm{d}x}\right)}\mathrm{d}u$$

$\Pr[N_t = k + 1]$

$$= \int_0^t \Pr[N_{t-x} = k] \cdot \lambda e^{-\lambda x}\mathrm{d}x$$

$$= \int_0^t \frac{(\lambda(t-x))^k e^{-\lambda(t\cancel{-x})}}{k!} \cdot \lambda \cancel{e^{-\lambda x}}\mathrm{d}x = \frac{\lambda^{(k+1)}e^{-\lambda t}}{k!}\int_0^t (t-x)^k\mathrm{d}x = \frac{\lambda^{(k+1)}e^{-\lambda t}}{k!}\int_t^0 \frac{u^k}{-1}\mathrm{d}u$$

$f(u) = u^k$

$u = g(x) = (t - x)$

$\frac{\mathrm{d}g(x)}{\mathrm{d}x} = -1$

$$= \frac{\lambda^{(k+1)}e^{-\lambda t}}{k!}\left[-\frac{1}{k+1}u^{(k+1)}\right]_t^0 = \frac{\lambda^{(k+1)}e^{-\lambda t}}{(k+1)!}\left[u^{(k+1)}\right]_0^t = \frac{(\lambda t)^{(k+1)}e^{-\lambda t}}{(k+1)!}$$ ✓

Poisson Process

Definition: A **Poisson process** with *intensity* λ is a collection of random variables $X_1, X_2, \ldots \in \mathbb{R}$ such that, if $N(a, b) = |\{i \mid X_i \in [a, b]\}|$, then

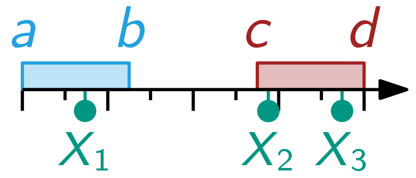
- $N(a, b) \sim \mathrm{Pois}(\lambda(b - a))$ (homogeneity)
- $a < b < c < d$: $N(a, b)$ and $N(c, d)$ are independent (independence)

- Assuming we know how many X_i are in $[\cancel{a}, b]$, *where* are they within the interval?

0 due to memorylessness

$$\Pr[N(a, b) = k] = \frac{(\lambda(b-a))^k e^{-\lambda(b-a)}}{k!}$$

- Simple case: $N(0, b) = 1$, where is X_1?



For $t \leq b$: $\Pr[X_1 \leq t \mid N(0, b) = 1] = \frac{\Pr[X_1 \leq t \wedge N(0,b)=1]}{\Pr[N(0,b)=1]}$

exactly one in $[0, b]$ *and* it is $\leq t$

$$= \frac{\Pr[N(0,t)=1 \wedge N(t,b)=0]}{\Pr[N(0,b)=1]}$$

independence of disjoint intervals

$$= \frac{\Pr[N(0,t)=1] \cdot \Pr[N(t,b)=0]}{\Pr[N(0,b)=1]}$$

for $X \sim \mathcal{U}([0, b])$

$$= \frac{(\cancel{\lambda} t) e^{\cancel{-\lambda t}} \cdot e^{\cancel{-\lambda(b}\cancel{-t})}}{(\cancel{\lambda} b) e^{\cancel{-\lambda b}}} = \frac{t}{b} = F_X(t)$$

- In general: the positions of the points are distributed uniformly in an interval

Continuous Spaces: Joint Distributions

Definition: For two random variables X, Y the **joint cumulative distribution function** is

$$F_{X,Y}(a, b) = \Pr[X \leq a \wedge Y \leq b].$$

The **joint density function** $f_{X,Y}(a, b)$ satisfies $F_{X,Y}(a, b) = \int_{-\infty}^{a} \int_{-\infty}^{b} f_{X,Y}(x, y) \mathrm{d}y \mathrm{d}x$.

Definition: The **marginal density** of X is $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) \mathrm{d}y$.

Definition: The **conditional density** of X with respect to an event A is

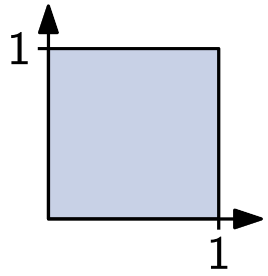
$$f_{X|A}(x) = \begin{cases} f_X(x)/\Pr[A], \text{if } x \in A, \\ 0, \text{otherwise.} \end{cases}$$

- For continuous Y, we specifically get $f_{X|Y=y}(x) = f_{X,Y}(x, y)/f_Y(y)$
- We can then write $f_{X,Y}(x, y) = f_{X|Y=y}(x) \cdot f_Y(y)$ (like the chain rule for probabilities)

Definition: Random variables X, Y are **independent** if $F_{X,Y}(x, y) = F_X(x) \cdot F_Y(y)$.

Example: $\mathcal{U}([0,1]^2)$

Uniform Distribution on the Unit Square



- We want to draw a point P uniformly at random from $[0,1]^2$
- Let X, Y be the x- and y-coordinates of P, respectively
- $f_P(x,y) = f_{X,Y}(x,y) = 1$ for $(x,y) \in [0,1]^2$ and $f_P(x,y) = 0$, otherwise

Marginal Distributions

Marginal Density
$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\mathrm{d}y$

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\mathrm{d}y = \int_0^1 1\mathrm{d}y = \left[y\right]_0^1 = 1 \qquad f_Y(y) = 1$$

- Note that $X \sim \mathcal{U}([0,1])$ and $Y \sim \mathcal{U}([0,1])$

Independence

X, Y independent if
$F_{X,Y}(x,y) = F_X(x) \cdot F_Y(y)$

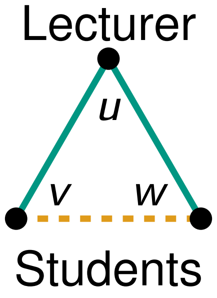
$$\begin{aligned} F_{X,Y}(a,b) = \int_{-\infty}^{a}\int_{-\infty}^{b} f_{X,Y}(x,y)\mathrm{d}y\mathrm{d}x &= \int_0^a \overbrace{\int_0^b 1\mathrm{d}y}^{\text{constant w.r.t. } x}\mathrm{d}x \\ &= \int_0^b 1\mathrm{d}y \cdot \int_0^a 1\mathrm{d}x \\ &= \int_0^b f_Y(y)\mathrm{d}y \cdot \int_0^a f_X(x)\mathrm{d}x = F_Y(b) \cdot F_X(a) \ \checkmark \end{aligned}$$

- Sample $P = (X,Y) \sim \mathcal{U}([0,1]^2)$ by independently sampling $X, Y \sim \mathcal{U}([0,1])$!

Application: Random Geometric Graphs

Motivation

- Average-case analysis: analyze models that represent the real world
- So far: Erdős-Rényi random graphs (connect two vertices independently with equal prob)
- Problem: In real networks, edges do *not* form independently
  - Two vertices are more likely to be adjacent if they have a common neighbor
    ↳ This property is called *locality* or *clustering*
  - ER-graph: $\Pr[\{v, w\} \in E \mid \{u, v\} \in E \land \{u, w\} \in E] = \Pr[\{v, w\} \in E]$ ✗



Idea

- Vertices are likelier to connect if their distance is already small

  $\Rightarrow$ Define vertex distances in advance by introducing geometry

> **Definition**: A **random geometric graph** is obtained by distributing vertices in a metric space and connecting any two with a probability that depends on their distance.

How many? Which space? Which metric? Which distribution? Which probability?

Simple & Realistic!

Application: Simple Random Geometric Graphs

Random Geometric Graph
Nodes distributed in metric space
Connection probability depends on distance

- Number: n vertices
- Space: 2-dimensional torus $\mathbb{T}^2$ (unit square with opposite sides identified)
- Metric: for $p = (p_1, p_2), q = (q_1, q_2)$: $d_i = |p_i - q_i|$
  $\hookrightarrow L_\infty$ norm: $d(p, q) = \max_{i \in \{1,2\}} \min\{d_i, 1 - d_i\}$
- Distribution: For each v independently: $P_v \sim \mathcal{U}([0, 1]^2)$
- Probability

$$\Pr[\{u, v\} \in E] = \begin{cases} 1, \text{if } d(P_u, P_v) \leq r \leftarrow \\ 0, \text{otherwise} \end{cases}$$

threshold parameter

Expected Degree of *v*

- Neighbors of v are in $\boxed{N(v)}$ (here $N(v)$ denotes the *region* in the ground space)
- $\mathbb{E}[\deg(v)] = \mathbb{E}[\sum_{u \in V \setminus \{v\}} \mathbb{1}_{\{P_u \in N(v)\}}] = \sum_{u \in V \setminus \{v\}} \Pr[d(P_u, P_v) \leq r]$
- Draw $P_u = (X, Y)$ as independent $X, Y \sim \mathcal{U}([0, 1])$

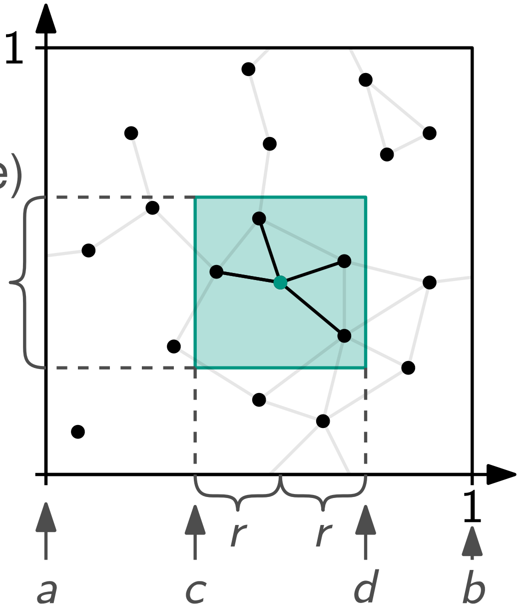
$$\Big\downarrow$$

$$= \sum_{u \in V \setminus \{v\}} \frac{2r}{1-0} \cdot \frac{2r}{1-0}$$

$$= (n-1) \cdot \underbrace{4r^2}$$

(area of the region $N(v)$)

$$X \sim \mathcal{U}([a, b]) : \Pr[X \in [c, d] \subseteq [a, b]] = \frac{d-c}{b-a}$$



and y*-coordinate of* u in here

Simple Random Geometric Graphs – Locality

Locality Realistic assumption: $r = \Theta(n^{-1/2})$ such that $\mathbb{E}[\deg(v)] = \Theta(1)$ Convention: $v = P_v$

- Two vertices v and w are likelier to connect if they have a common neighbor u

$$\Pr[\{v,w\} \in E] = \Pr[v \in N(w)] = 4r^2 = \Theta(1/n)$$

w.l.o.g assume $u = (r, r)$

$$\Pr[\{v,w\} \in E \mid \{u,v\} \in E \wedge \{u,w\} \in E]$$

$$= \Pr[w \in N(v) \mid v \in N(u) \wedge w \in N(u)] = \frac{\Pr[w \in N(v) \wedge v \in N(u) \wedge w \in N(u)]}{\Pr[v \in N(u) \wedge w \in N(u)]}$$

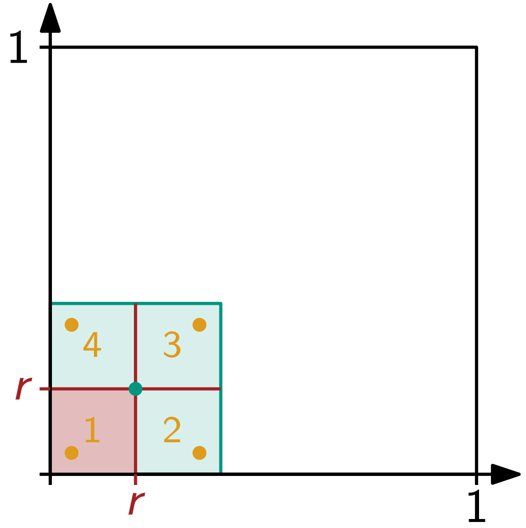
Numerator $\Pr[w \in N(v) \wedge v \in N(u) \wedge w \in N(u)]$

$$= \int_{\mathbb{R}^2} \Pr[w \in N(v) \wedge v \in N(u) \wedge w \in N(u) \mid v = (x,y)] f_{X,Y}(x,y) \mathrm{d}y \mathrm{d}x$$

$$= \int_0^{2r} \int_0^{2r} \Pr[w \in N(v) \wedge w \in [0,2r]^2 \mid v = (x,y)] \mathrm{d}y \mathrm{d}x$$

Due to symmetry the area of the intersection is the same for these 4 positions of v.

$\Rightarrow$ Integrate only one quarter and multiply by 4



Law of Total Probability
$\Pr[A] = \int_{-\infty}^{\infty} \Pr[A \mid X = x] f_X(x) \mathrm{d}x$

$(X, Y) \sim \mathcal{U}([0,1]^2)$
$f_{X,Y}(x,y) = \mathbb{1}_{\{(x,y) \in [0,1]^2\}}$

Simple Random Geometric Graphs – Locality

Locality Realistic assumption: $r = \Theta(n^{-1/2})$ such that $\mathbb{E}[\deg(v)] = \Theta(1)$ Convention: $v = P_v$

- Two vertices v and w are likelier to connect if they have a common neighbor u

$$\Pr[\{v,w\} \in E] = \Pr[v \in N(w)] = 4r^2 = \Theta(1/n)$$

w.l.o.g assume $u = (r, r)$

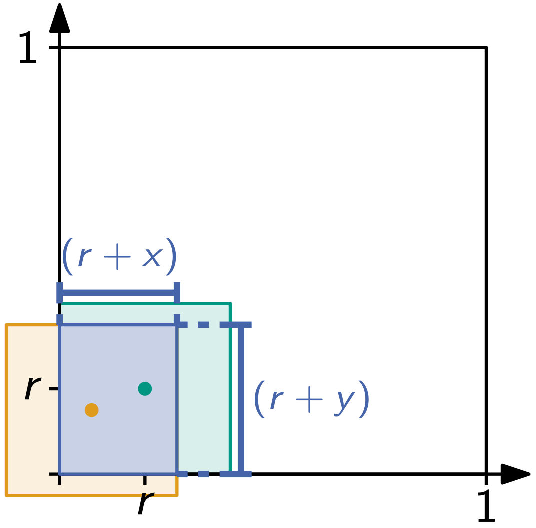
$$\Pr[\{v,w\} \in E \mid \{u,v\} \in E \wedge \{u,w\} \in E]$$

$$= \Pr[w \in N(v) \mid v \in N(u) \wedge w \in N(u)] = \frac{\Pr[w \in N(v) \wedge v \in N(u) \wedge w \in N(u)]}{\Pr[v \in N(u) \wedge w \in N(u)]}$$

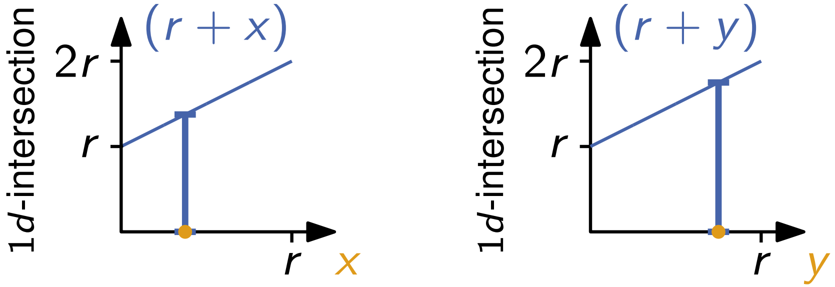
Numerator $\Pr[w \in N(v) \wedge v \in N(u) \wedge w \in N(u)]$

$$= \int_{\mathbb{R}^2} \Pr[w \in N(v) \wedge v \in N(u) \wedge w \in N(u) \mid v = (x,y)] f_{X,Y}(x,y) \mathrm{d}y \mathrm{d}x$$

$$= 4 \int_0^r \int_0^r \Pr[w \in N(v) \wedge w \in [0,2r]^2 \mid v = (x,y)] \mathrm{d}y \mathrm{d}x$$



Consider size of intersection in *one* dimension depending on position of v



$2d$-intersection is product of $1d$-intersections

$(r+x) \cdot (r+y)$

Law of Total Probability

$$\Pr[A] = \int_{-\infty}^{\infty} \Pr[A \mid X = x] f_X(x) \mathrm{d}x$$

$$(\boldsymbol{X}, \boldsymbol{Y}) \sim \boldsymbol{\mathcal{U}}([0,1]^2)$$

$$f_{X,Y}(x,y) = \mathbb{1}_{\{(x,y) \in [0,1]^2\}}$$

Simple Random Geometric Graphs – Locality

Locality

Realistic assumption: $r = \Theta(n^{-1/2})$ such that $\mathbb{E}[\deg(v)] = \Theta(1)$

Convention: $v = P_v$

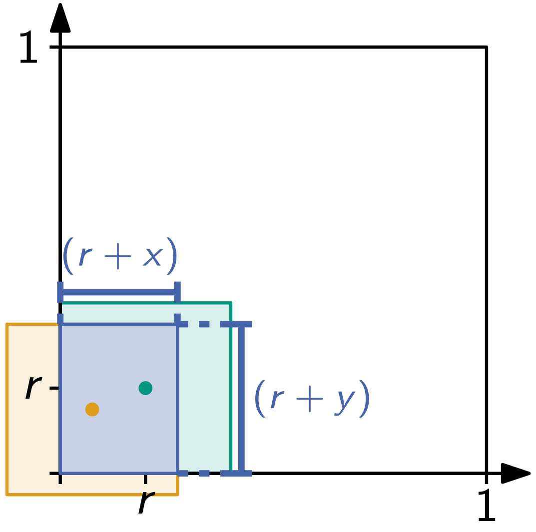
- Two vertices v and w are likelier to connect if they have a common neighbor u

$$\Pr[\{v, w\} \in E] = \Pr[v \in N(w)] = 4r^2 = \Theta(1/n)$$

w.l.o.g assume $u = (r, r)$

$$\Pr[\{v, w\} \in E \mid \{u, v\} \in E \wedge \{u, w\} \in E]$$

$$= \Pr[w \in N(v) \mid v \in N(u) \wedge w \in N(u)] = \frac{\Pr[w \in N(v) \wedge v \in N(u) \wedge w \in N(u)]}{\Pr[v \in N(u) \wedge w \in N(u)]}$$



Numerator $\Pr[w \in N(v) \wedge v \in N(u) \wedge w \in N(u)]$

$$= \int_{\mathbb{R}^2} \Pr[w \in N(v) \wedge v \in N(u) \wedge w \in N(u) \mid v = (x, y)] f_{X,Y}(x, y) \mathrm{d}y\mathrm{d}x$$

$$= 4 \int_0^r \int_0^r \Pr[w \in N(v) \wedge w \in [0, 2r]^2 \mid v = (x, y)] \mathrm{d}y\mathrm{d}x$$

$$= 4 \int_0^r \int_0^r (r + x) \cdot (r + y) \mathrm{d}y\mathrm{d}x$$

$$= 4 \int_0^r (r + x) \cdot \int_0^r (r + y) \mathrm{d}y\mathrm{d}x$$

$$= 4 \int_0^r (r + y) \mathrm{d}y \cdot \int_0^r (r + x) \mathrm{d}x$$

$$= 4 \left(\int_0^r (r + x) \mathrm{d}x \right)^2$$

$$= 4 \left(\int_0^r r \mathrm{d}x + \int_0^r x \mathrm{d}x \right)^2$$

$$= 4 \left(r \left[x \right]_0^r + \left[\tfrac{1}{2} x^2 \right]_0^r \right)^2$$

$$= 4 \left(r^2 + \tfrac{1}{2} r^2 \right)^2 = 4 \left(\tfrac{3}{2} r^2 \right)^2 = \cancel{4} \frac{9}{\cancel{4}} r^4$$

$$= 9r^4$$

Law of Total Probability

$$\Pr[A] = \int_{-\infty}^{\infty} \Pr[A \mid X = x] f_X(x) \mathrm{d}x$$

$$\boldsymbol{(X, Y) \sim \mathcal{U}([0, 1]^2)}$$

$$f_{X,Y}(x, y) = \mathbb{1}_{\{(x,y) \in [0,1]^2\}}$$

Simple Random Geometric Graphs – Locality

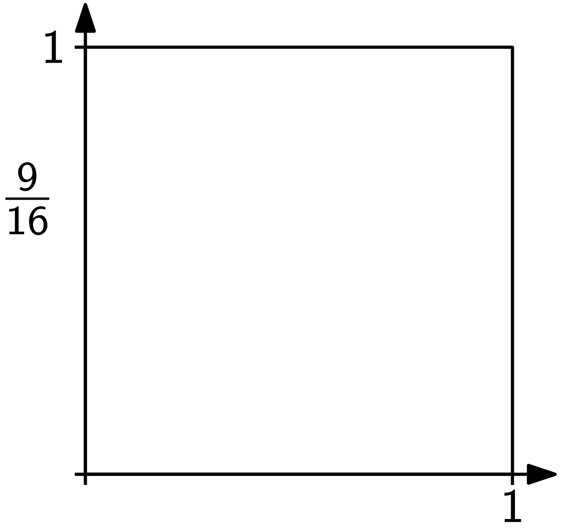
Locality Realistic assumption: $r = \Theta(n^{-1/2})$ such that $\mathbb{E}[\deg(v)] = \Theta(1)$ Convention: $v = P_v$

- Two vertices v and w are likelier to connect if they have a common neighbor u

$$\Pr[\{v, w\} \in E] = \Pr[v \in N(w)] = 4r^2 = \Theta(1/n) \; >$$

$$\Pr[\{v, w\} \in E \mid \{u, v\} \in E \wedge \{u, w\} \in E] = \Theta(1) \; ✓$$

$$= \Pr[w \in N(v) \mid v \in N(u) \wedge w \in N(u)] = \frac{\Pr[w \in N(v) \wedge v \in N(u) \wedge w \in N(u)]}{\Pr[v \in N(u) \wedge w \in N(u)]} = \frac{9}{16}$$



Numerator $\Pr[w \in N(v) \wedge v \in N(u) \wedge w \in N(u)] = 9r^4$

Denominator

$$\Pr[v \in N(u) \wedge w \in N(u)] = \Pr[v \in N(u)] \cdot \Pr[w \in N(u)]$$

positions are drawn independently

distribution identical for all vertices

$$= (\Pr[v \in N(u)])^2$$

$$= (4r^2)^2 = 16r^4$$

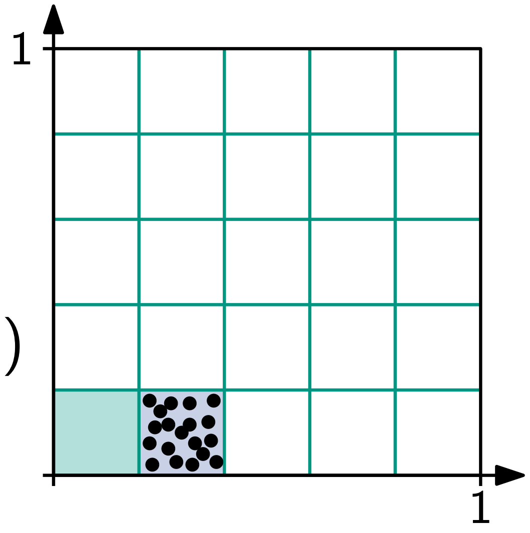
Law of Total Probability

$\Pr[A] = \int_{-\infty}^{\infty} \Pr[A \mid X = x] f_X(x) \mathrm{d}x$

$\boldsymbol{(X, Y) \sim \mathcal{U}([0, 1]^2)}$

$f_{X,Y}(x, y) = \mathbb{1}_{\{(x,y) \in [0,1]^2\}}$

Application: Simple RGGs – Fair Distribution



- Discretize the space into equally sized grid cells, such that the expected number of vertices in each cell is $\log(n)$
- Each cell C_i has width and height $\sqrt{\log(n)/n}$
- Let X_i denote the number of vertices in C_i

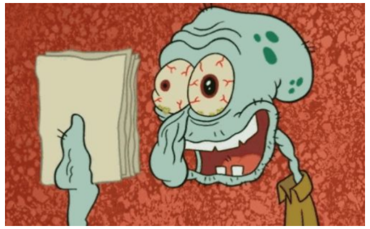
$$\mathbb{E}[X_i] = \mathbb{E}[\textstyle\sum_{v \in V} \mathbb{1}_{\{v \in C_i\}}] = n \cdot \Pr[v \in C_i] = n \frac{\sqrt{\log(n)/n}}{1-0} \frac{\sqrt{\log(n)/n}}{1-0} = \log(n)$$

- What is the probability that each cell gets *exactly* $\log(n)$ vertices?

$$\Pr[X_1 = \log(n)] = \binom{n}{\log(n)} \left(\frac{\log(n)}{n}\right)^{\log(n)} \left(1 - \frac{\log(n)}{n}\right)^{n-\log(n)}$$

- Same distribution for all X_i: $\Pr[\forall i : X_i = \log(n)] = \cancel{\prod_i \Pr[X_i = \log(n)]}$ $\neq$
- X_1 and X_2 are *not* independent $\Pr[X_1 = \log(n) \mid X_2 = n] = 0$
- Chain rule of probability:

$$\Pr[\forall i : X_i = \log(n)]$$
$$= \Pr[X_1 = \log(n)] \cdot \Pr[X_2 = \log(n) \mid X_1 = \log(n)] \cdot \Pr[X_3 = \log(n) \mid X_1 = \log(n) \wedge X_2 = \log(n)] \cdot \ldots$$


https://i.imgflip.com/1pln6k.jpg?a471949

Poissonization

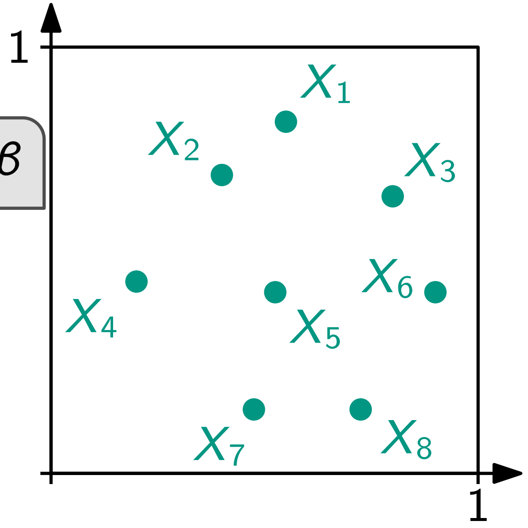
Idea

- Avoid dependencies by replacing uniform point sampling with a Poisson point process

Definition: A **Poisson *Point* process** with *intensity* λ is a collection of random variables $X_1, X_2, \ldots \in \mathbb{R}^2$ such that, if $|A|$ is the area of A and $N(A) = |\{i \mid X_i \in A\}|$, then

- $N(A) \sim \text{Pois}(\lambda|A|)$ (homogeneity)
- $A \cap B = \emptyset$: $N(A)$ and $N(B)$ are independent (independence)

(Generalizes to arbitrary dimension, $1d$ is the Poisson process seen earlier)



- Note: We do not know how many points we get!
- How do we choose λ?
  - We should at least *expect* n points in our ground space $[0,1]^2$

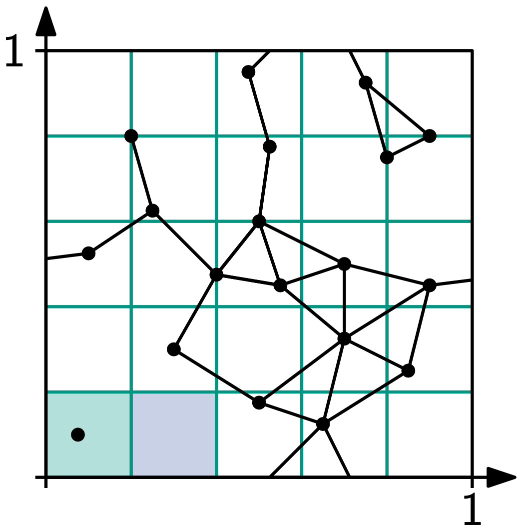
$N \sim \text{Pois}(\beta) \Rightarrow \mathbb{E}[N] = \beta$

$$n = \mathbb{E}[|\{i \mid X_i \in [0,1]^2\}|] = \mathbb{E}[N([0,1]^2)] = \lambda|[0,1]^2| = \lambda$$

- Recall: conditioned on their number, points are distributed uniformly
- Simulate PPP: sample $N \sim \text{Pois}(n)$, sample N points uniformly
- The resulting **Poissonized RGG** has n vertices in expectation

Application: Poissonized RGGs – Fair Distribution

?



- Vertices of RGG distributed using Poisson point process with intensity $\lambda = n$
- Discretize the space into equally sized grid cells, such that the expected number of vertices in each cell is $\log(n)$
- Each cell C_i has width and height $\sqrt{\log(n)/n} \Rightarrow |C_i| = \log(n)/n$
- Let X_i denote the number of vertices in $C_i \Rightarrow X_i \sim \mathrm{Pois}(\lambda|C_i|)$

$$\mathbb{E}[X_i] = \lambda|C_i| = \log(n)$$

- What is the probability that each cell gets *exactly* $\log(n)$ vertices?

$$\Pr[X_i = \log(n)] = \frac{(\lambda|C_i|)^{\log(n)} e^{-\lambda|C_i|}}{\log(n)!} = \frac{(\cancel{n}\frac{\log(n)}{\cancel{n}})^{\log(n)} e^{-\cancel{n}\frac{\log(n)}{\cancel{n}}}}{\log(n)!} = \frac{\log(n)^{\log(n)} e^{-\log(n)}}{\log(n)!}$$

$$\leq \frac{\cancel{\log(n)^{\log(n)}}\, \cancel{e^{-\log(n)}}}{e(\cancel{\frac{\log(n)}{e}})^{\log(n)}} = \frac{1}{e}$$

$N \sim \mathrm{Pois}(\lambda|A|)$

$$\mathbb{E}[N] = \lambda|A|$$

$$\Pr[N = k] = \frac{(\lambda|A|)^k e^{-\lambda|A|}}{k!}$$

$$k! \geq e(k/e)^k$$

there are $n/\log(n)$ cells

- Same distribution for all X_i: $\Pr[\forall i : X_i = \log(n)] = \prod_i \Pr[X_i = \log(n)]$

by definition, disjoint regions *are* independent $\leq e^{-n/\log(n)}$ ✓

but we cheated...

De-Poissonization

Situation

- We started with a simple RGG (n, $\mathbb{T}^2$, L_∞-norm, $P_i \sim \mathcal{U}([0,1]^2)$, $\Pr[\{u,v\} \in E] = \mathbb{1}_{\{d(u,v) \le r\}}$)
- Switched to Poissonized RGG (n is replaced by $\text{Pois}(n)$) and obtained $\Pr[\forall i\colon X_i = \log(n)] \le e^{-n/\log(n)}$
- How can we translate this result to the original model?

Recall

- Conditioned on the number of points in area A, the points are distributed uniformly in A
- So we get from the poissonized RGG to the original, by conditioning on the fact that the number of points N in $[0,1]^2$ obtained in the Poisson point process is exactly $N = n$

$$\Pr[N = n] = \frac{(\lambda|A|)^n e^{-\lambda|A|}}{n!} = \frac{n^n e^{-n}}{n!} \le \cancel{\left(\frac{n}{e}\right)^n} \cdot \frac{1}{\sqrt{2\pi n}\cancel{\left(\frac{n}{e}\right)^n}\underbrace{e^{\frac{1}{12n+1}}}_{\Theta(1)}} = \Theta(n^{-1/2})$$

$$\Pr_{\text{RGG}(n)}[\forall i\colon X_i = \log(n)] = \Pr_{\text{RGG}(\text{Pois}(n))}[\forall i\colon X_i = \log(n) \mid N = n]$$

$$= \frac{\Pr_{\text{RGG}(\text{Pois}(n))}[\forall i{:}X_i = \log(n) \wedge N = n]}{\Pr_{\text{RGG}(\text{Pois}(n))}[N = n]} \le \frac{\Pr_{\text{RGG}(\text{Pois}(n))}[\forall i{:}X_i = \log(n)]}{\Pr_{\text{RGG}(\text{Pois}(n))}[N = n]} \le \frac{e^{-n/\log(n)}}{\Theta(n^{-1/2})}$$ ✓

$\boldsymbol{N \sim \text{Pois}(\lambda|A|)}$

$$\Pr[N = k] = \frac{(\lambda|A|)^k e^{-\lambda|A|}}{k!}$$

Stirling

$$n! \ge \sqrt{2\pi n}\left(\frac{n}{e}\right)^n e^{\frac{1}{12n+1}}$$

RGG – The Bigger Picture

Random Geometric Graph
Nodes distributed in metric space
Connection probability depends on distance

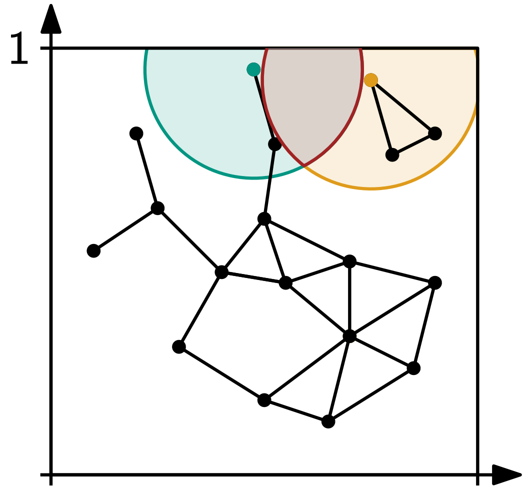
Seen so far

- Simple RGG
- n, $\mathbb{T}^2$, L_∞-norm, $P_i \sim \mathcal{U}([0,1]^2)$, $\Pr[\{u,v\} \in E] = \mathbb{1}_{\{d(u,v) \leq r\}}$
- Expected degree of a vertex is $(n-1)4r^2$
- Probability to connect given common neighbor is constant

More commonly used model

- n, $[0,1]^2$, L_2-norm, $P_i \sim \mathcal{U}([0,1]^2)$, $\Pr[\{u,v\} \in E] = \mathbb{1}_{d(u,v) \leq r}$
- Complications
  - Vertices near the boundary / corners behave differently
  - Intersections of neighborhoods are lenses or parts thereof
- Still $\mathbb{E}[\deg(v)] = \Theta(nr^2)$
- Still probability to connect given common neighbor non-vanishing

$\boxed{N(v)}$ is a disk
No wrap-around!



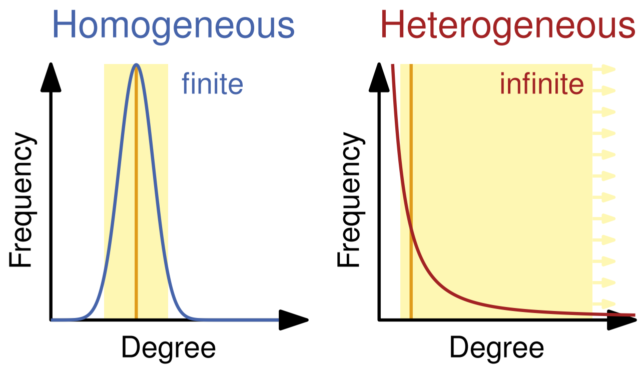
Problem: Homogeneous degree distribution does not match many real-world graphs

A Heterogeneous Distribution

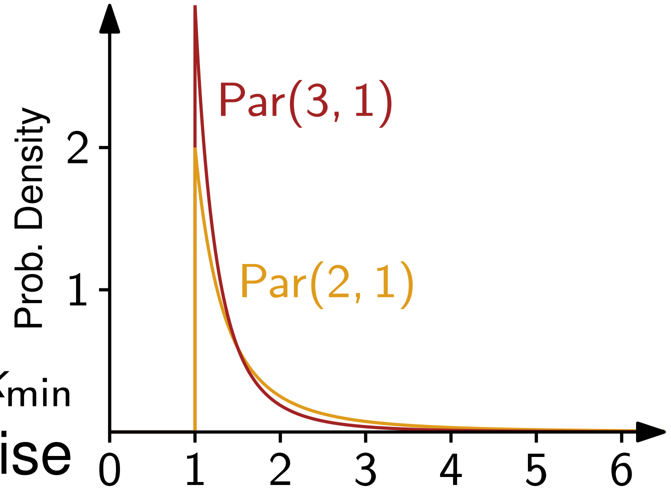
Motivation

- Distributions seen so far have finite variance
- Graphs with corresponding degree distributions are *homogeneous*

  ⇒ For constant expected degree, it is very unlikely to find a high-degree vertex
- In real-world graphs high-degree vertices are not too rare (think of celebrities in a social network)



Hard to distinguish!



Pareto Distribution

- $X \sim \mathrm{Par}(\alpha, x_{\min})$

  $x_{\min}$: minimum attainable value

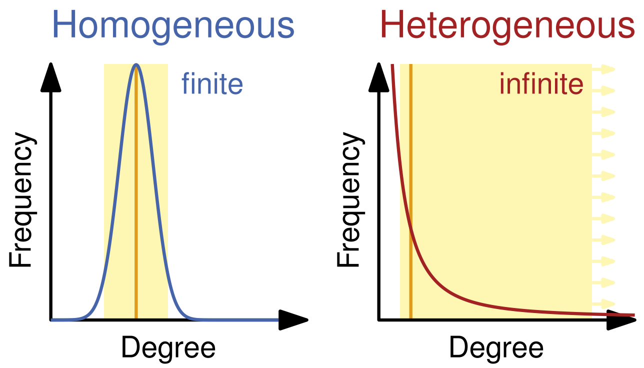
  α: shape parameter
- Probability density function: $f_X(x) = \begin{cases} \alpha x_{\min}^{\alpha} \cdot x^{-(\alpha+1)}, & \text{if } x \geq x_{\min} \\ 0, & \text{otherwise} \end{cases}$

A Heterogeneous Distribution

Motivation

- Distributions seen so far have finite variance
- Graphs with corresponding degree distributions are *homogeneous*

  $\Rightarrow$ For constant expected degree, it is very unlikely to find a high-degree vertex
- In real-world graphs high-degree vertices are not too rare (think of celebrities in a social network)



Pareto Distribution

- $X \sim \mathrm{Par}(\alpha, x_{\min})$

minimum attainable value

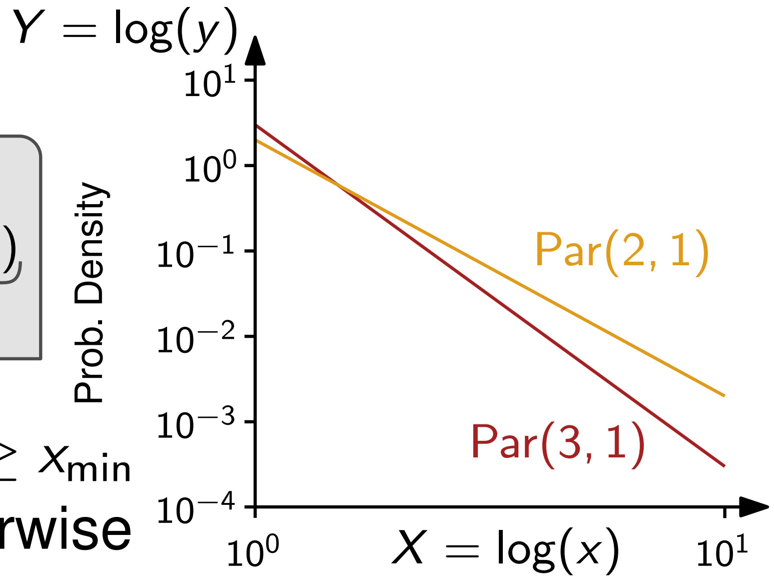
shape parameter

Log-Log-Plot $y = bx^k$

$$\underbrace{\log(y)} = \log(b) + k\,\underbrace{\log(x)}$$

$$\rightarrow Y = \log(b) + kX \leftarrow$$

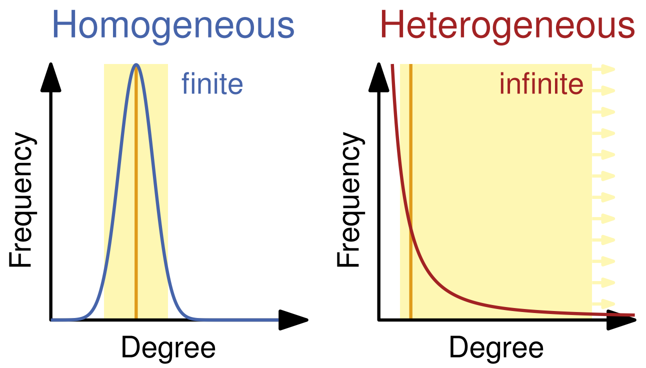
- Probability density function: $$f_X(x) = \begin{cases} \alpha x_{\min}^{\alpha} \cdot x^{-(\alpha+1)}, & \text{if } x \geq x_{\min} \\ 0, & \text{otherwise} \end{cases}$$

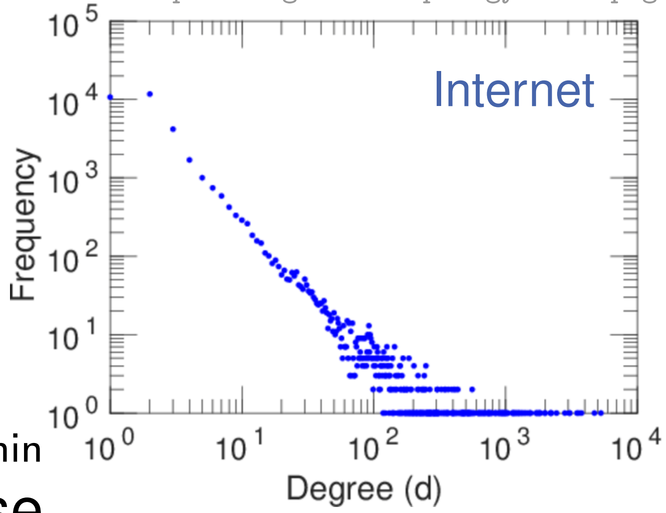


A Heterogeneous Distribution

Motivation

- Distributions seen so far have finite variance
- Graphs with corresponding degree distributions are *homogeneous*
  $\Rightarrow$ For constant expected degree, it is very unlikely to find a high-degree vertex
- In real-world graphs high-degree vertices are not too rare (think of celebrities in a social network)





konect.cc/plot/degree.a.topology.full.png

Log-Log-Plot $y = bx^k$

$$\underbrace{\log(y)}_{Y} = \log(b) + k\underbrace{\log(x)}_{X}$$

$$Y = \log(b) + kX$$

Pareto Distribution

- $X \sim \mathrm{Par}(\alpha, x_{\min})$

  $x_{\min}$: minimum attainable value

  α: shape parameter

- Probability density function: $f_X(x) = \begin{cases} \alpha x_{\min}^{\alpha} \cdot x^{-(\alpha+1)}, & \text{if } x \geq x_{\min} \\ 0, & \text{otherwise} \end{cases}$

Exercise: Determine for which values of α we have $\mathbb{E}[X] < \infty$ but $\mathrm{Var}[X] = \infty$

Conclusion

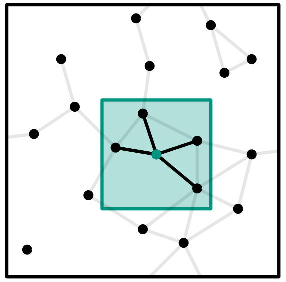
Continuous Distributions

- For our purposes they are handled like discrete versions (replacing sums with integrals)
- Seen today: Uniform distribution, exponential distribution, Pareto distribution, joint distributions

Poisson (Point) Process (not discussed in lecture)

- Yields random point set with certain properties (homogeneity & independence)
- Number of points is a random variable
- Conditioned on certain number, points are distributed uniformly
- (De-)Poissonization to circumvent stochastic dependencies

We can simulate a PPP by drawing number according to Poisson distribution and distributing as many points uniformly



Random Geometric Graphs

- Vertices distributed at random in metric space
- Edges form with probability depending on distances
- Exhibit locality (edges tend to form between vertices with common neighbors)

Outlook: More realistic extension of RGGs featuring a heterogeneous degree distribution