Exercise Sheet 13 – Random Graphs
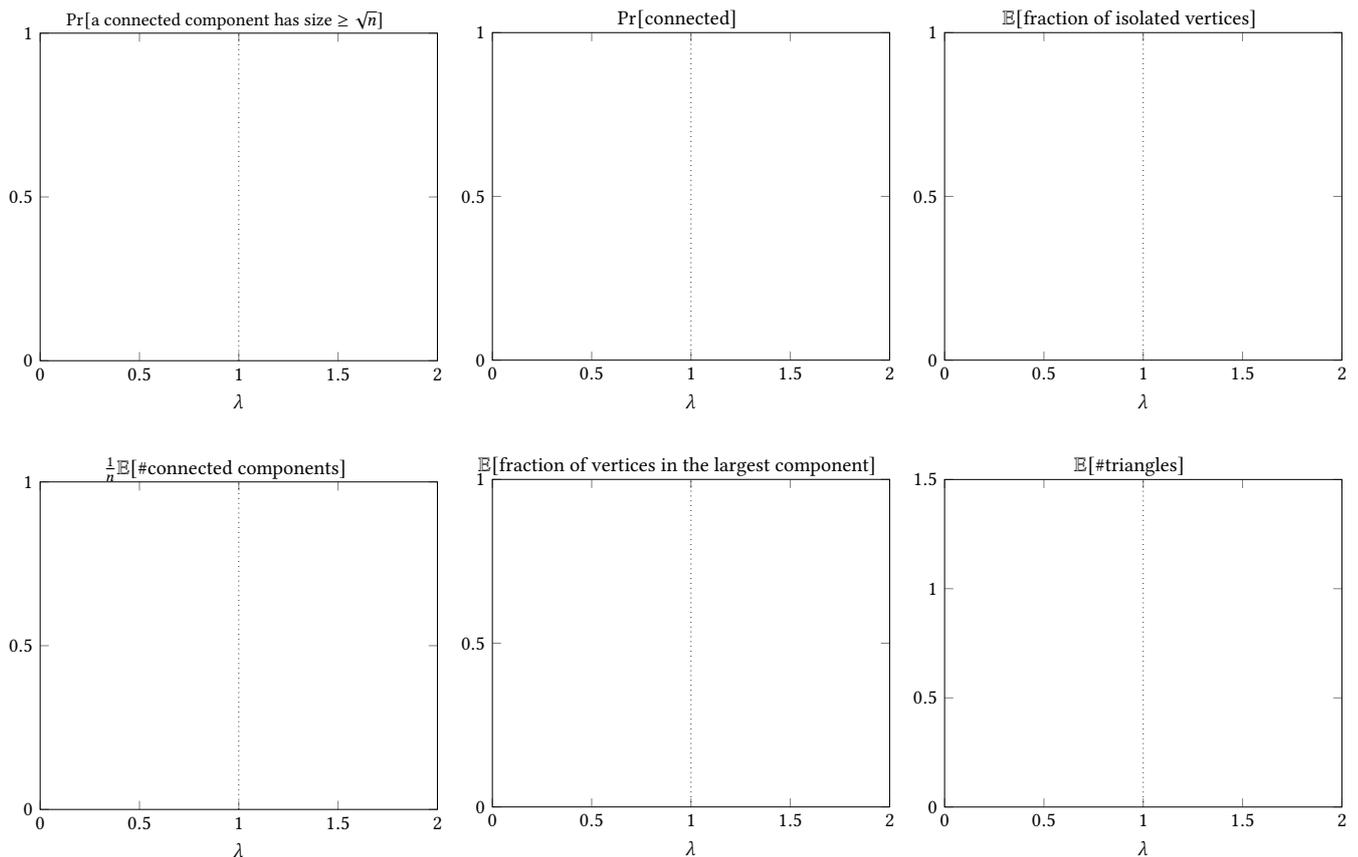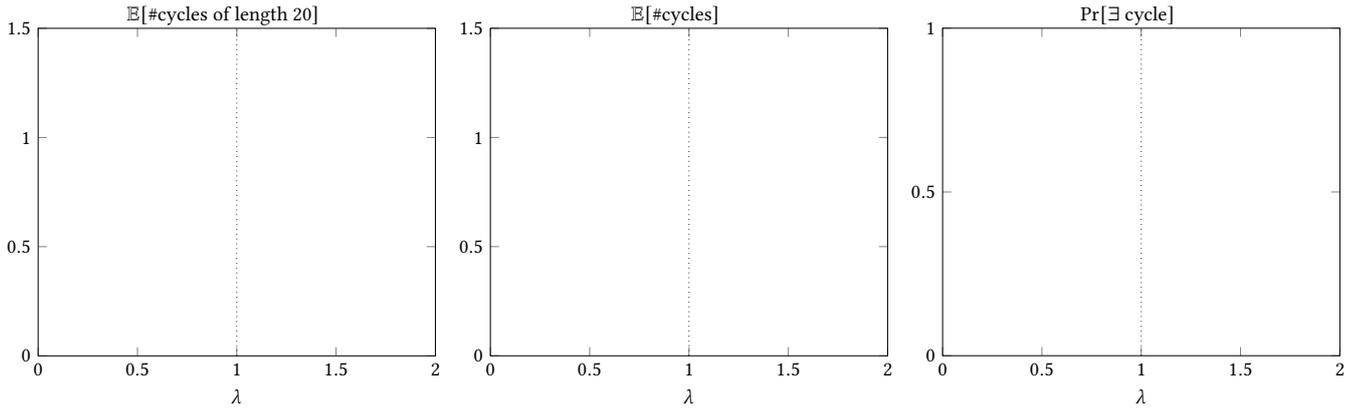
Probability and Computing

Exercise 1 – Intuition for Erdős–Rényi Graphs

We consider the Erdős–Rényi graph $G(n, \lambda n/2)$ or the Gilbert graph $G(n, \lambda/n)$ (both lead to the same result). The expected vertex degree is therefore $\lambda \pm O(1/n)$.

(i) Sketch the behavior of the following probabilities and expectations for $n \to \infty$ as a function of $\lambda \in [0, 2]$.

Hint: This is not about numerical exactness but about the qualitative behavior. Where does the curve take the value 0, 1, or ∞? Does anything special happen at $\lambda = 1$?

Pr[a connected component has size $\geq \sqrt{n}$]

Pr[connected]

$\mathbb{E}$[fraction of isolated vertices]

$\frac{1}{n}\mathbb{E}$[#connected components]

$\mathbb{E}$[fraction of vertices in the largest component]

$\mathbb{E}$[#triangles]

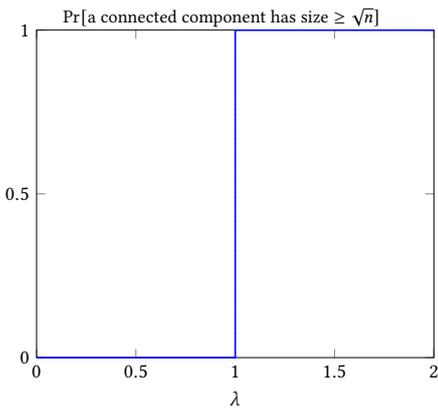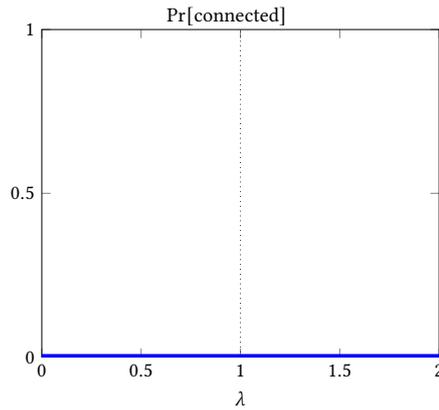E[#cycles of length 20]　　　E[#cycles]　　　Pr[∃ cycle]

(ii) Let $\lambda = \Theta(1)$. Guess: What holds with high probability for (the order of magnitude of) the minimum and maximum degree as a function of n?

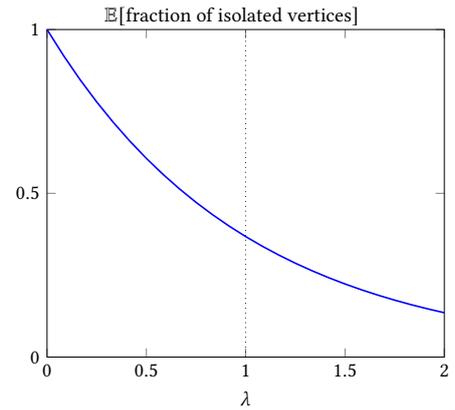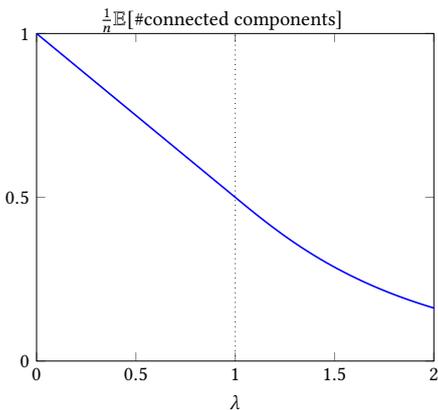$$\min_{v \in [n]} \deg(v) = \text{\dots\dots} \qquad \max_{v \in [n]} \deg(v) = \Theta\Big(\text{\dots\dots}\Big)$$

Solution 1

Pr[a connected component has size $\geq \sqrt{n}$]

follows from the "Sudden Emergence" result

Pr[connected]

see next question

E[fraction of isolated vertices]

$\Pr_{X \sim \text{Po}(\lambda)}[X = 0] = e^{-\lambda}$

$\frac{1}{n}\mathbb{E}[\text{\#connected components}]$

experimental results
for $n = 1000000$

E[fraction of vertices in the largest component]

$\frac{\lambda + W(-e^{-\lambda}\lambda)}{\lambda}$ for $\lambda \geq 1$,
see Exercise 3

E[#triangles]

$\binom{n}{3}\left(\frac{\lambda}{n}\right)^3 \to \lambda^3/6$

$\mathbb{E}[\#\text{cycles of length }20]$	$\mathbb{E}[\#\text{cycles}]$	$\Pr[\exists\text{ cycle}]$
$\frac{n(n-1)\cdot\ldots\cdot(n-19)}{2\cdot20}\cdot\left(\frac{\lambda}{n}\right)^{20}\to\lambda^{20}/40$	$\to\sum_{k=3}^{\infty}\frac{\lambda^k}{2k}=\frac{1}{4}(-\lambda^2-2\lambda-2\ln(1-\lambda))$	experimental results for $n=1000000$
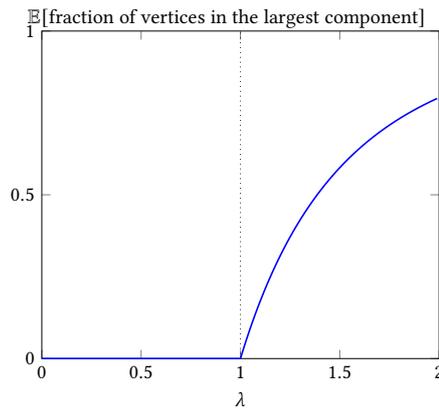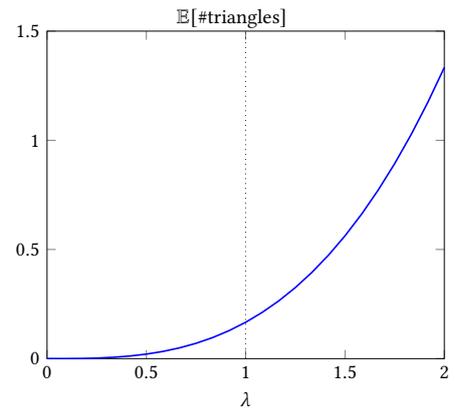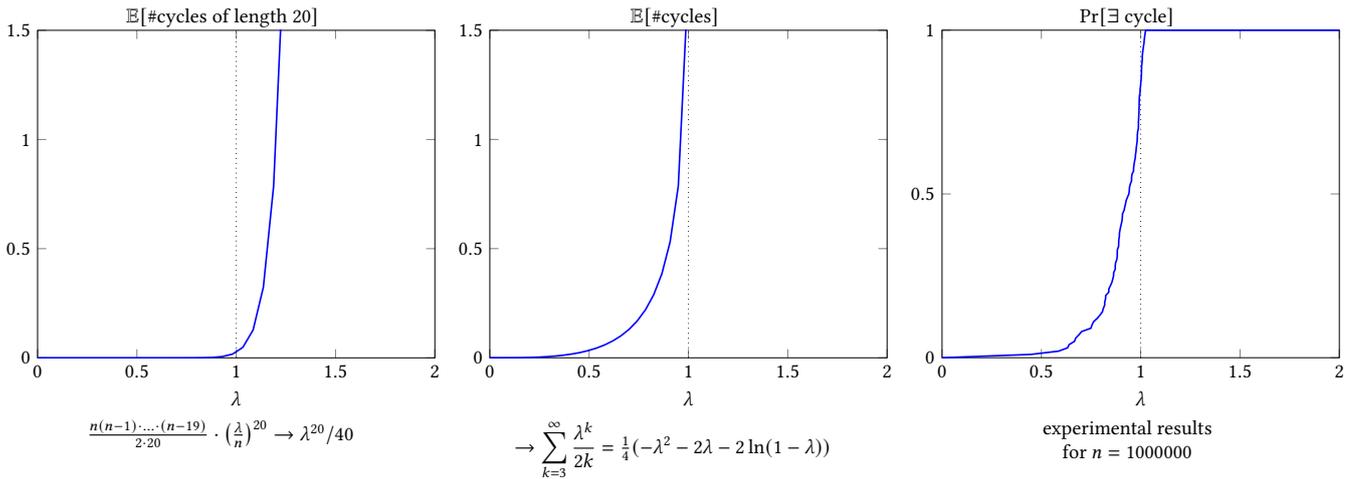
(ii) The minimum degree is 0 with high probability. In part (a) we even saw that there are in expectation $\Theta(n)$ isolated vertices. A formal proof can be somewhat tricky.

The maximum degree is $\Theta(\frac{\log n}{\log\log n})$ with high probability. This can be seen by noting that the expected number of vertices of degree d is approximately $\mathbb{E}[N_d]\approx n\cdot e^{-\lambda}\frac{\lambda^d}{d!}$. Setting $\mathbb{E}[N_d]=1$ yields $d=\Theta(\frac{\log n}{\log\log n})$. The main issue here is which d is needed so that $d!\approx n$ holds.

Exercise 2 – Vertex Degrees in Erdős–Rényi Graphs

On slide 15 of the section on balls into bins and Poissonization we showed that $\text{Bin}(n,\frac{\lambda}{n})$ *converges to* $\text{Pois}(\lambda)$ *as* $n\to\infty$ *(or, respectively, that the CDFs converge). In the following exercise you may use without proof:*

> **Lemma.** Let $t_1,t_2,\cdots\in\mathbb{N}$ and $p_1,p_2,\cdots\in(0,1)$ as well as $\lambda\in\mathbb{R}_+$. Furthermore let $X_n\sim\text{Bin}(t_n,p_n)$ for all $n\in\mathbb{N}$ and $X\sim\text{Pois}(\lambda)$. If $t_n\to\infty$ and $t_n\cdot p_n\to\lambda$ as $n\to\infty$, then $X_n\xrightarrow{\;\text{d}\;}X$ as $n\to\infty$.

Let $\lambda>0$ and $X\sim\text{Pois}(\lambda)$. We consider variants of Erdős–Rényi graphs from the lecture. We set the expected degree to approximately λ and want to show (using the above lemma) that the distribution of a single vertex converges asymptotically to $\text{Pois}(\lambda)$ as $n\to\infty$ (while λ remains constant).

(i) Let X_n be the degree of vertex 1 in $G(n,p)$ with $p=\lambda/n$. Show $X_n\xrightarrow{\;\text{d}\;}X$.

(ii) Let X_n be the degree of vertex 1 in $G^{\text{UE}}(n,m)$ with $m=\lfloor\lambda n/2\rfloor$. Show $X_n\xrightarrow{\;\text{d}\;}X$.

(iii) Let X_n be the degree of vertex 1 in $G(n,m)$ with $m=\lfloor\lambda n/2\rfloor$. Show $X_n\xrightarrow{\;\text{d}\;}X$.

Hint: The last part is by far the most difficult. It is helpful to "trap" $G(n,m)$ using a coupling between two Gilbert graphs $G^-=(n,p^-)$ and $G^+=(n,p^+)$.

3

Solution 2

(i) Since each of the $n - 1$ edges incident to vertex 1 is present with probability λ/n, we have $X_n \sim \text{Bin}(n - 1, \lambda/n)$. We apply the lemma with $t_n = n - 1$ and $p_n = \lambda/n$. Clearly $t_n \to \infty$ and $t_n p_n \to \lambda$. Thus $X_n \to X$ as desired.

(ii) Since each of the $2m$ edge endpoints is attached to vertex 1 with probability $1/n$, we have $X_n = \text{Bin}(2m, \frac{1}{n})$. We apply the lemma with $t_n = 2m = 2\lfloor \lambda n/2 \rfloor$ and $p_n = 1/n$. Clearly $t_n \to \infty$ and $t_n p_n \to \lambda$. Thus $X_n \to X$ as desired.

(iii) Let $G = G(n, m)$. We additionally consider the Gilbert graphs $G^- = G(n, p^-)$ and $G^+ = G(n, p^+)$ where $p^- := \frac{\lambda}{n} - n^{-4/3}$ and $p^+ := \frac{\lambda}{n} + n^{-4/3}$. Since p^- and p^+ differ only slightly from $p = \lambda/n$, the degree X_n^- of vertex 1 in G^- and the degree X_n^+ of vertex 1 in G^+ converge as in part (i), i.e. $X_n^- \xrightarrow{\text{d}} X$ and $X_n^+ \xrightarrow{\text{d}} X$. To "trap" X_n between X_n^- and X_n^+ we use a coupling.

Let $\mathcal{E}$ be the set of the $\binom{n}{2}$ possible edges. In a common probability space, for each $e \in \mathcal{E}$ there is an independent random variable $Z_e \sim \mathcal{U}([0, 1])$. The edge sets E^-, E, and E^+ of G^-, G, and G^+ (or formally of their "copies" $G^{-\prime}$, G', and $G^{+\prime}$ with the same distribution) are defined as:

$$E^- = \{e \in \mathcal{E} \mid Z_e \leq p^-\}$$
$$E\ = \{e \in \mathcal{E} \mid Z_e \text{ belongs to the } m \text{ smallest numbers in } (Z_e)_{e \in \mathcal{E}}\}$$
$$E^+ = \{e \in \mathcal{E} \mid Z_e \leq p^+\}$$

It is clear that this yields the correct distributions, i.e. that this is a valid coupling. Let $m^- = |E^-|$ and $m^+ = |E^+|$. We have:

$$m^- \sim \text{Bin}\left(\binom{n}{2}, p^-\right)$$
$$\mathbb{E}[m^-] = \binom{n}{2}p^- = \frac{n(n-1)}{2}\left(\frac{\lambda}{n} - n^{-4/3}\right) = \frac{\lambda n}{2} - \Theta(n^{2/3}) = m - \Theta(n^{2/3})$$
$$\Pr[m^- \geq m] \leq \Pr[|m^- - \mathbb{E}[m^-]| \geq \Theta(n^{2/3})] \overset{\text{Cheb.}}{\leq} \frac{\text{Var}(m^-)}{\Theta(n^{4/3})} = \frac{\Theta(n)}{\Theta(n^{4/3})} = \Theta(n^{-1/3}).$$

Analogously one shows that $\Pr[m^+ \leq m] = \Theta(n^{-1/3})$. We now consider the event $\text{succ} = \{m^- \leq m \leq m^+\}$. If succ occurs, then by construction $E^- \subseteq E \subseteq E^+$ and thus $X_n^- \leq X_n \leq X_n^+$. By the above calculation and a union bound we have $\Pr[\text{succ}] = 1 - \Theta(n^{-1/3})$. It follows for $i \in \mathbb{N}_0$:

$$\Pr[X_n \leq i] = \Pr[X_n \leq i \wedge \text{succ}] + \Pr[X_n \leq i \wedge \overline{\text{succ}}] \leq \Pr[X_n^- \leq i \wedge \text{succ}] + \Theta(n^{-1/3})$$
$$\leq \Pr[X_n^- \leq i] + \Theta(n^{-1/3}) \longrightarrow \Pr[X \leq i]. \quad \text{// since } X_n^- \xrightarrow{\text{d}} X$$

Similarly one obtains

$$\Pr[X_n \leq i] \geq \Pr[X_n \leq i \wedge \text{succ}] \geq \Pr[X_n^+ \leq i \wedge \text{succ}] = \Pr[X_n^+ \leq i] - \Pr[X_n^+ \leq i \wedge \overline{\text{succ}}]$$
$$\geq \Pr[X_n^+ \leq i] - \Theta(n^{-1/3}) \longrightarrow \Pr[X \leq i]. \quad \text{// since } X_n^+ \xrightarrow{\text{d}} X$$

Hence $\Pr[X_n \leq i] \to \Pr[X \leq i]$ for all $i \in \mathbb{N}_0$, and thus $X_n \xrightarrow{\text{d}} X$.

Exercise 3 – Extinction Probability in Galton–Watson Trees

Let $\mathrm{GWT}(\lambda)$ be the Galton–Watson tree with offspring distribution $\mathrm{Pois}(\lambda)$.

(i) Let n_i be the number of vertices in level i of $\mathrm{GWT}(\lambda)$. The root level is level 0. What is $\mathbb{E}[n_i]$?

(ii) For $i \in \mathbb{N}_0$ let p_i be the probability that $\mathrm{GWT}(\lambda)$ has at least one vertex at level i. Express p_{i+1} in terms of p_i.
Hint: Use Exercise 3 (iv) from Sheet 9.

(iii) Determine, for $\lambda \in \{0, 0.5, 1, 1.1, 1.5\}$ (or for arbitrary λ), approximations for the probability $s(\lambda)$ that $\mathrm{GWT}(\lambda)$ is infinite. A computer algebra system may be useful (e.g. Wolfram Alpha).

Additional consideration: What is the expected number of vertices at level i?

Solution 3

(i) Each vertex has expected λ children in the next level. It is quite intuitive that therefore $\mathbb{E}[n_i] = \lambda^i$. Formally one can use conditional expectations and induction. Written compactly:

$$\mathbb{E}[n_i] = \mathbb{E}[\mathbb{E}[n_i \mid n_{i-1}]] = \mathbb{E}[\lambda n_{i-1}] = \lambda \mathbb{E}[n_{i-1}] \overset{\text{Ind.}}{=} \lambda\lambda^{i-1} = \lambda^i.$$

(ii) Let $\mathrm{depth}(T) \in \mathbb{N}_0 \cup \{\infty\}$ be the depth of a tree T. Let $X \sim \mathrm{Pois}(\lambda)$ be the number of children of the root of $\mathrm{GWT}(\lambda)$ and let $T^{(1)}, T^{(2)}, \ldots, T^{(X)}$ be the subtrees starting at these children. Let $Y = |\{i \in \{1, \ldots, X\} \mid \mathrm{depth}(T^{(i)}) \geq i - 1\}|$ be the number of subtrees of depth at least $i-1$. Since the subtrees have the same distribution as $\mathrm{GWT}(\lambda)$, we have $\Pr[\mathrm{depth}(T^{(i)}) \geq i - 1] = p_{i-1}$ for all $i \in \{1, \ldots, X\}$. Since the subtrees are independent, we furthermore have $Y \sim \mathrm{Bin}(X, p_{i-1})$. By Exercise 3 (iv) from Sheet 9 it follows that $Y \sim \mathrm{Pois}(\lambda p_{i-1})$. Hence:

$$p_i = \Pr[\mathrm{depth}(\mathrm{GWT}(\lambda)) \geq i] = \Pr[\exists j \in \{1, \ldots, X\} : \mathrm{depth}(T^{(j)}) \geq i - 1]$$
$$= \Pr[Y > 0] = 1 - \Pr[Y = 0] = 1 - e^{-\lambda p_{i-1}}.$$

(iii) It is intuitive that $s(\lambda) = \lim_{i \to \infty} p_i$, but we nevertheless show this formally. The probability that $\mathrm{GWT}(\lambda)$ has *exactly* i levels is $q_i = p_i - p_{i+1}$. Clearly $s(\lambda) + q_0 + q_1 + \cdots = 1$. It follows:

$$s(\lambda) = 1 - \lim_{i \to \infty} \sum_{j=0}^{i-1} q_j = \lim_{i \to \infty} \left(p_0 - \sum_{j=0}^{i-1} (p_j - p_{j+1}) \right) = \lim_{i \to \infty} p_i.$$

By (ii) we have $p_0 = 1$ and $p_{i+1} = 1 - e^{-\lambda p_i}$ for all $i \geq 0$. It is convenient to consider the function $f(x) = 1 - e^{-\lambda x}$. Iterating this function starting at $x = 1$ generates the

sequence $(p_i)_{i\in\mathbb{N}}$. Since f is monotonically decreasing and our starting value satisfies $f(p_0) < p_0$, the sequence is monotonically decreasing. Hence $s(\lambda) = \lim_{i\to\infty} p_i$ is the largest fixed point of f. For $\lambda \leq 1$ we have $f(x) = 1 - e^{-\lambda x} \leq 1 - e^{-x} \leq 1 - (1 - x) = x$ with equality only for $x = 0$. For such λ we therefore have $s(\lambda) = 0$.

For $\lambda > 1$ the largest solution of $x = 1 - e^{-\lambda x}$ differs from the trivial solution $x = 0$. A computer algebra system can determine this as $s(\lambda) = \frac{\lambda + W(-\lambda e^{-\lambda})}{\lambda}$, where W denotes the so-called Lambert W-function. Numerically one obtains $s_{1.1} = 0.176134$ and $s_{1.5} = 0.582812$. A plot can be found in Exercise 1.