Simulation and Modelling
Warm-up Problem

1. We can simulate the system by randomly sampling the bus inter-arrival time distribution until the next arrival time, at, say, is greater than t. Then $at - t$ forms an observation. Note that this assumes that the previous bus arrived at time 0. If we repeat this many times (here 100000) and take an average then we approximate $E(W)$. For example, for Pareto-distributed inter-arrival times with $a = 2$ and $x > 1$:

   ```
   int n = 100000;
   double sum = 0;
   double at;
   for (int i = 0; i < n; i++) {
     at = 0;
     while (at < t) {
       at = Pareto.pareto(1,2);
     }
     sum += at - t;
   }
   System.out.println("t = " + t + ", E(W) = " + sum / n);
   ```

   If we set $t = 3$, say, then the expected waiting time is $t/(a-1) = 3$ (see below). Running the program three times:

   ```
   t = 3.0, E(W) = 2.98573023400585
   t = 3.0, E(W) = 3.064193215463034
   t = 3.0, E(W) = 2.9572261602447343
   ```

2. Each estimate from the simulation is itself an average of many (here 100000) individual waiting times. By the central limit theorem, each estimate is thus (approximately) a sample from a normal distribution whose mean is $E(W)$. If we run the program n times then each estimate, E_i, $1 \leq i \leq n$ say, will be independent of the others. The *sample mean* of the estimates

$$\overline{E} = \frac{1}{n}\sum_{i=0}^{n} E_i$$

therefore has a Student's 't' distribution so the 90% confidence interval is:

$$\frac{\overline{E} \pm t_{n,0.9} S}{\sqrt{n}}$$

where $t_{n,0.9}$ comes from tables and S is the *sample standard deviation*

$$S = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (E_i - \overline{E})^2}$$

3. As we saw...

$$\begin{aligned}
P(X \leq t + x \mid X > t) &= 1 - P(X > t + x \mid X > t) \\
&= 1 - \frac{P(X > t + x \;\&\; X > t)}{P(X > t)} \\
&= 1 - \frac{P(X > t + x)}{P(X > t)} \\
&= 1 - \frac{1 - F(t + x)}{1 - F(t)} \\
&= \frac{F(t + x) - F(t)}{1 - F(t)}
\end{aligned}$$

4. (a) Plug in $F(x) = x/b$:

$$\begin{aligned}
P(X \leq t + x \mid X > t) &= \frac{F(t + x) - F(t)}{1 - F(t)} \\
&= \frac{\frac{t+x}{b} - \frac{t}{b}}{1 - \frac{t}{b}} \\
&= \frac{x}{b - t}
\end{aligned}$$

Using the definition of $E(W)$ above, noting that $0 \leq W \leq b - t$ in this case:

$$\begin{aligned}
E(W) &= \int_0^{b-t} 1 - \frac{x}{(b - t)}\, dx \\
&= \frac{b - t}{2}
\end{aligned}$$

Observe that the mean waiting time drops linearly with t.

(b) Plug in $F(x) = 1 - e^{-\lambda x}$:

$$\begin{aligned}
P(X \leq t + x \mid X > t) &= \frac{F(t + x) - F(t)}{1 - F(t)} \\
&= \frac{e^{-\lambda t} - e^{-\lambda(t+x)}}{e^{-\lambda t}} \\
&= 1 - e^{-\lambda x}
\end{aligned}$$

Amazingly, this means that W has the *same* distribution as the original bus inter-arrival times. We already know that the mean is $1/\lambda$, which is *independent* of t so

$$E(W) = \frac{1}{\lambda}$$

and we're done. This is called the *memoryless* property of the exponential distribution: the future is independent of the past.

(c) Plug in $F(x) = 1 - x^{-a}$, where $x > 1$:

$$\begin{aligned} P(X \leq t + x \mid X > t) &= \frac{F(t+x) - F(t)}{1 - F(t)} \\ &= \frac{t^{-a} - (t+x)^{-a}}{t^{-a}} \\ &= 1 - \left(\frac{t}{t+x}\right)^{a} \end{aligned}$$

Thus:

$$\begin{aligned} E(W) &= \int_0^\infty \left(\frac{t}{t+x}\right)^a dx \\ &= \left| -\frac{t^a}{(a-1)(t+x)^{a-1}} \right|_0^\infty \end{aligned}$$

Note that this is ∞ if $a \leq 1$. For $a > 1$ the mean is finite and given by

$$E(W) = \frac{t}{a-1}$$

so we assume that we will always pick an $a > 1$. Curiously, notice that the expected waiting time *increases* with t!