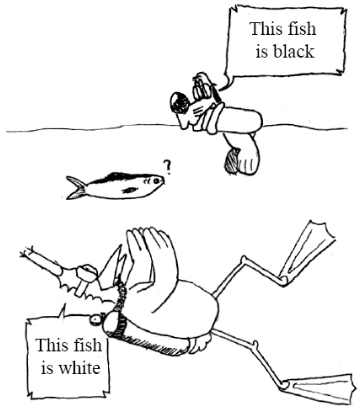
PERFORMANCE EVALUATION
EXERCISES

PALM CALCULUS

With Solutions Jean-Yves Le Boudec, Spring 2021



1. SovRail claims that only 5% of trains arrivals are late. BorduKonsum claims that 30% of train users suffer from late train arrivals. Say which is true (Hint: Use the heuristic seen in class; you may indroduce the variables $D_n = 1$ if the nth arrival is late and $D_n = 0$ if it is on time):

   (a) ☐ At least one of them uses alternative facts.
   (b) ☐ Number of passengers in a late train $\approx 1.15\times$ number of passengers in a train
   (c) ☐ Number of passengers in a late train $\approx 2.45\times$ number of passengers in a train
   (d) ☒ Number of passengers in a late train $\approx 6\times$ number of passengers in a train

   Solution. Let N be the number of train arrival events over some observation period. Let $D_n = 1$ if the nth arrival is late and $D_n = 0$ if it is on time. SovRail estimates

$$\bar{D} = \frac{\sum_{n=1}^{N} D_n}{N}$$

   Let P_n be the number of passengers leaving the train at the nth arrival event. BorduKonsum estimates

$$D^* = \frac{\sum_{n=1}^{N} P_n D_n}{\sum_{n=1}^{N} P_n}$$

   The number of passengers per train arrival event is

$$\bar{P} = \frac{\sum_{n=1}^{N} P_n}{N}$$

   and the number of passengers per late train arrival event is

$$\bar{P}_{\text{late}} = \frac{\sum_{n=1}^{N} P_n D_n}{\sum_{n=1}^{N} D_n}$$

   We see that

$$\begin{aligned}
\frac{D^*}{\bar{D}} &= \frac{\sum_{n=1}^{N} P_n D_n}{\sum_{n=1}^{N} P_n} \frac{N}{\sum_{n=1}^{N} D_n} \\
\frac{\bar{P}_{\text{late}}}{\bar{P}} &= \frac{\sum_{n=1}^{N} P_n D_n}{\sum_{n=1}^{N} D_n} \frac{N}{\sum_{n=1}^{N} P_n} = \frac{D^*}{\bar{D}}
\end{aligned}$$

   Since $\frac{D^*}{\bar{D}} = 6$ we conclude that $\frac{\bar{P}_{\text{late}}}{\bar{P}} = 6$ as well.

2. Instagram uses a clustering algorithm to classify the vacation preferences of their N users and obtains M clusters. (We have $N > M \gg 1$). We have obtained the distribution of the number of users per cluster and found that it follows an exponential distribution with rate λ. What is the PDF of the size of the cluster seen by an arbitrary user ? Compare the mean size of a cluster seen by a user and the mean size of an arbitrary cluster.

   Recall that the exponential distribution with rate λ has mean $\frac{1}{\lambda}$, variance $\frac{1}{\lambda^2}$ and PDF $f_C(x) = \lambda e^{-\lambda x}\mathbf{1}_{\{x\geq 0\}}$.

   Solution. This is the same setting as seen in the lecture with flows and packets, where flows are replaced by clusters and packets are replaced by users. The probability that a user belongs to a cluster of size x is

$$f_U(x) = \eta x f_C(x)$$

   where η is a normalizing constant. Thus

$$f_U(x) = \eta\lambda x e^{-\lambda x}\mathbf{1}_{\{x\geq 0\}}$$

   We find the value of η by the condition $\int_1^\infty f_U(x)dx = 1$:

$$\eta^{-1} = \int_0^{+\infty} \lambda x e^{-\lambda x} dx = \mathbb{E}\left(X\right) = \frac{1}{\lambda}$$

   where X is a random variable with distribution exponential with rate λ. Thus $\eta = \lambda$ and the required PDF is

$$f_U(x) = \lambda^2 x e^{-\lambda x}\mathbf{1}_{\{x\geq 0\}}$$

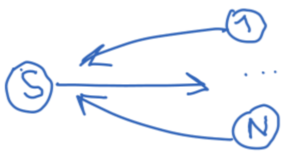
   The average cluster size $\bar{C}_U$ seen by a user is the mean of this distribution:

$$\begin{aligned}\bar{C}_U &= \int_0^{+\infty} x f_U(x)dx = \int_0^{+\infty} \lambda^2 x^2 e^{-\lambda x}dx \\ &= \lambda\mathbb{E}\left(X^2\right) = \lambda\left(\mathrm{var}X + \mathbb{E}\left(X\right)^2\right) = \lambda\frac{2}{\lambda^2} = \frac{2}{\lambda}\end{aligned}$$

   The mean size, $\bar{C}$, of an arbitrary cluster is the mean of the exponential distribution: $\bar{C} = \frac{1}{\lambda}$. We see that $\bar{C}_U = 2\bar{C}$, i.e. in average users see a cluster twice as large as Instagram does.

3. A sensor sends a broadcast poll and waits for all clients to ACK.

   The round trip times $S \to i \to S$ are iid Exp(1) and N is large. How many polls per time unit are sent ?



   (a) ☒ $\lambda \approx \frac{1}{\log N}$
   (b) ☐ $\lambda \approx \frac{1}{\sqrt{N}}$
   (c) ☐ $\lambda \approx \frac{1}{N}$
   (d) ☐ $\lambda \approx \frac{1}{N^2}$

   Solution. The answer is the intensity of the point process of polls which, by the intensity formula is

$$\lambda = \frac{1}{\mathbb{E}\left(\text{poll time}\right)}$$

The poll time is

$$T = \max\{X_1, ...X_N\}$$

where X_i is the round-trip time. Thus(see qqplots):

$$\mathbb{E}(T) = \mathbb{E}\left(X_{(N)}\right) \approx F^{-1}\left(\frac{1}{N+1}\right)$$

where F is the CDF of X_i, namely $F(x) = 1 - e^{-x}$. Thus $F^{-1}(p) = -\log p$ and

$$\mathbb{E}(T) \approx \log(N+1)$$

and thus $\lambda \approx \frac{1}{\log(N+1)} \approx \frac{1}{\log(N)}$.

4. Same question if the round trip times are iid standard Pareto with index $p = 2$.

   (a) ☐ $\lambda \approx \frac{1}{\log N}$
   (b) ☒ $\lambda \approx \frac{1}{\sqrt{N}}$
   (c) ☐ $\lambda \approx \frac{1}{N}$
   (d) ☐ $\lambda \approx \frac{1}{N^2}$

   Solution. The only difference is that now $F(x) = \frac{1}{x^2}$, thus $F^{-1}(p) = -\sqrt{p}$ and

$$\lambda \approx \frac{1}{\sqrt{N+1}} \approx \frac{1}{\sqrt{N}}$$

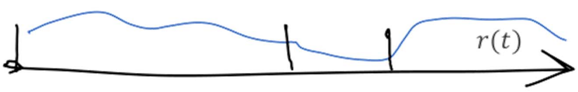
5. For the random waypoint model, the distribution of the next waypoint is uniform...

   (a) ☒ when sampled at an arbitrary waypoint
   (b) ☐ when sampled at an arbitrary point in time
   (c) ☐ both
   (d) ☐ none

   Solution. (a) is true by construction; (b)is not true, we sample more often trips that are long; long trips are more likely to have ends at the edge of the area. The next waypoint, seen at an arbitrary point in time, is more often at the edge than at the center.

6. A system downloads data over a channel with a fluctuating instantaneous rate $r(t)$. The data is sent in rounds, of average duration $\bar{T}$. The average amount of data transferred in one round is $\bar{B}$.

   We sample the channel using a Poisson process of rate λ and find a sampled data transfer rate $\bar{r}$.



   Give a formula for the value of $\bar{r}$ as a function of $\bar{T}$, $\bar{B}$ and λ. (Hint: use the large-time heuristic or the inversion formula.)

   Solution. First observe that, by PASTA, $\bar{r}$ is the time average expectation of the transfer rate.

   With Large time heuristic: We do a simulation of duration T_{tot} during which there are N_{tot} rounds. We estimate $\bar{r}$ by

$$\bar{r} \;=\; \frac{1}{T_{\text{tot}}}\int_0^{T_{\text{tot}}} r(s)ds = \frac{1}{T_{\text{tot}}}\sum_{i=1} B_i$$

where B_i is the amount of data sent in the ith round. Let also T_i be the duration of the ith round. Then

$$\begin{aligned}\bar{T} &= \frac{1}{N_{\text{tot}}}\sum_{i=1} T_i = \frac{T_{\text{tot}}}{N_{\text{tot}}} \\ \bar{B} &= \frac{1}{N_{\text{tot}}}\sum_{i=1} B_i\end{aligned}$$

hence

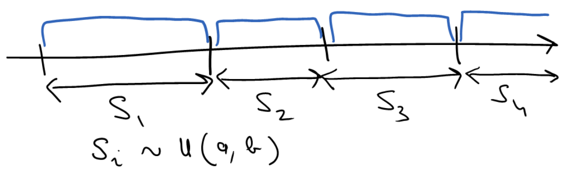
$$\bar{r} = \frac{N_{\text{tot}}\bar{B}}{T_{\text{tot}}\bar{B}} = \frac{\bar{B}}{\bar{T}}$$

With Inversion Formula: Let T_i be the beginnings of rounds and μ the intensity of this point process; by the intensity formula $\mu = \frac{1}{\bar{T}}$. By the inversion formula:

$$\bar{r} = \mathbb{E}\left(r(t)\right) = \mu\mathbb{E}^0\left(\int_{T_0}^{T_1} r(s)ds\right) = \mu\bar{B} = \frac{\bar{B}}{\bar{T}}$$

7. Sensing devices send one message at the end of the sensing interval.

The durations of sensing intervals are iid uniform between a and b. We want to write a perfect simulation of a sensor. From which distribution should we sample the residual cycle time ?



Solution. By Theorem 7.3 the residual cycle time has PDF f_X given by

$$f_X(x) = \lambda\int_x^{+\infty} f_T(t)dt$$

where f_T is the PDF of the cycle duration, i.e.

$$f_T(t) = \frac{1}{b-a}\mathbf{1}_{\{a\leq t\leq b\}}$$

and λ is the intensity of the point process, i.e.

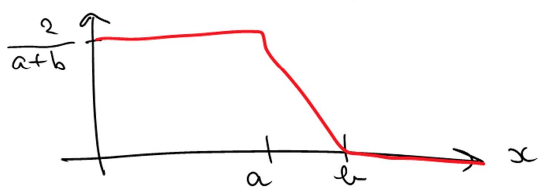
$$\lambda = \frac{1}{\text{mean cycle duration}} = \frac{2}{a+b}$$

If $x > b$ then the residual cycle time cannot be equal to x and $f_X(x) = 0$. For $0 \leq x \leq a$:

$$f_X(x) = \lambda\int_x^b f_T(t)dt = \lambda\int_a^b f_T(t)dt = \lambda$$

and for $a \leq x \leq b$:

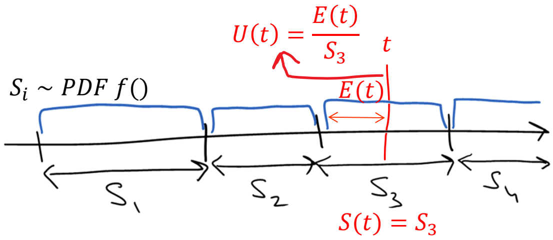
$$f_X(x) = \lambda\int_x^{+\infty} f_T(t)dt = \lambda\int_x^b f_T(t)dt = \lambda\frac{b-x}{b-a}$$



The PDF is piecewise linear, as illustrated next.

8. (Continuation) We continue with the previous example but now assume that the sensing intervals are iid with common pdf f (not necessarily uniform).

We call $S(t)$ and $E(t)$ the durations of the current interval and of the time elapsed since the beginning of the current interval, when observed at time t. Also let $U(t) = \frac{E(t)}{U(t)} \in [0; 1]$ be the fraction of the elapsed interval observed at time t.



(a) Under which condition does this simulation have a stationary regime ?

(b) When this condition holds, what is the distribution of $U(t)$?

(c) Are $U(t)$ and $S(t)$ independent, i.e. is the fraction of the interval we are in independent of the length of the interval ?

Solution.

(a) This is a modulated process, the condition is that the expected interval duration is finite, i.e.

$$\int_0^{+\infty} tf(t)dt < +\infty$$

(b) By the inversion formula, for any test function φ we have (T_i are the epochs when an interval starts):

$$\mathbb{E}\left(\varphi(U(t))\right) = \lambda \mathbb{E}^0 \left(\int_{T_0}^{T_1} \varphi(U(t))dt \right) \tag{1}$$

Now for $t \in [T_0; T_1]$, $U(t) = \frac{t}{T_1 - T_0}$; thus

$$\int_{T_0}^{T_1} \varphi(U(t))dt \quad = \quad \int_{T_0}^{T_1} \varphi\left(\frac{t}{T_1 - T_0}\right) dt = \int_0^1 \varphi\left(x\right)(T_1 - T_0)dx = (T_1 - T_0) \int_0^1 \varphi\left(x\right) dx$$

where we performed the change of variable $x = \frac{t}{T_1 - T_0}$ in the integral. Note that the last integral is non random (does not depend on T_0, T_1) thus by (1)

$$\mathbb{E}\left(\varphi(U(t))\right) = \lambda \mathbb{E}^0 \left(T_1 - T_0\right) \int_0^1 \varphi\left(x\right) dx$$

Note that (intensity formula) $\lambda^{-1} = \mathbb{E}^0 \left(T_1 - T_0\right)$ thus

$$\mathbb{E}\left(\varphi(U(t))\right) = \int_0^1 \varphi\left(x\right) dx$$

which expresses that $U(t)$ is uniform over $[0; 1]$.

(c) We prove independence by showing that, for any two test function φ, ψ we have

$$\mathbb{E}\left(\varphi(U(t))\psi(S(t))\right) = \mathbb{E}\left(\varphi(U(t))\right)\mathbb{E}\left(\psi(S(t))\right)$$

The rest is similar to the previous question; by the inversion formula:

$$\begin{aligned}
\mathbb{E}\left(\varphi(U(t))\psi(S(t))\right) &= \lambda\mathbb{E}^0\left(\int_{T_0}^{T_1}\varphi(U(t))\psi(S(t))dt\right) \\
&= \lambda\mathbb{E}^0\left(\int_{T_0}^{T_1}\varphi\left(\frac{t}{T_1-T_0})\right)\psi\left(T_1-T_0\right)dt\right) \\
&= \lambda\mathbb{E}^0\left(\psi\left(T_1-T_0\right)\int_{T_0}^{T_1}\varphi\left(\frac{t}{T_1-T_0})\right)dt\right) \\
&= \lambda\mathbb{E}^0\left(\psi\left(T_1-T_0\right)\int_0^1\varphi\left(x\right)\left(T_1-T_0\right)dx\right) \\
&= \lambda\mathbb{E}^0\left((T_1-T_0)\psi\left(T_1-T_0\right)\int_0^1\varphi\left(x\right)dx\right) \\
&= \left[\int_0^1\varphi\left(x\right)dx\right]\lambda\mathbb{E}^0\left((T_1-T_0)\psi\left(T_1-T_0\right)\right)
\end{aligned}$$

If we set $\psi=1$ in the above we obtain $\mathbb{E}\left(\varphi(U(t)\right)=\int_0^1\varphi\left(x\right)dx$ as we know already from the previous question. Similarly, by setting $\varphi=1$ we obtain

$$\begin{aligned}
\mathbb{E}\left(\psi(S(t)\right) &= \lambda\mathbb{E}^0\left((T_1-T_0)\psi\left(T_1-T_0\right)\right) \\
&= \lambda\int_0^{+\infty}s\psi(s)f(s)ds=\int_0^{+\infty}\psi(s)\left[\lambda sf(s)\right]ds
\end{aligned}$$

This shows that the distribution of $S(t)$ has PDF $\lambda sf(s)$ as we know already from Theorem 7.3. Also it follows that

$$\mathbb{E}\left(\varphi(U(t))\psi(S(t))\right)=\mathbb{E}\left(\varphi(U(t))\right)\mathbb{E}\left(\psi(S(t))\right)$$

which shows that $S(t)$ and $U(t)$ are independent.

9. (Continuation) An algorithm for the perfect simulation of a device is returning the duration of the current interval $S(t)$ and the residual time $R(t)$ until end of current interval. Which implementation is correct ?

A.
1. Draw S from distribution with PDF $\lambda sf(s)$
2. Draw $V\sim$ Unif(0, 1)
3. $R=VS$

B.
1. Draw S from distribution with PDF $f(s)$
2. Draw $V\sim$ Unif(0, 1)
3. $R=(1-V)S$

(a) ☒ A
(b) ☐ B
(c) ☐ Both
(d) ☐ None

Solution. S should be drawn from the pdf of the interval duration seen by at an arbitrary point in time, which is not equal to $f(s)$ but to $\lambda sf(s)$ hence B is wrong.

To draw R and S we can draw U and S, whom we know from the previous question are independent, and note that $R=S(1-U)$. Therefore a correct algorithm is

1. Draw S from distribution with PDF $\lambda sf(s)$
2. Draw $V\sim$ Unif(0, 1)

3. $R = (1 - V)S$

Note that since $1 - V \sim V \sim$ Unif$(0, 1)$ we can replace line 4 by

3. $R = VS$

which is Algorithm A.

10. Consider the random waypoint model, where the speed chosen at a waypoint is sampled from the pdf $f()$. Can we choose $f()$ such that

    (a) the model has a stationary regime

    (b) the distribution of speed sampled at an arbitrary point in time is uniform between 0 and $v_{\max}$?

Solution. By Theorem 7.9 we need that the mean trip duration is finite. If this holds, the distribution of speed sampled at an arbitrary point in time has PDF

$$f_{V(t)} = \frac{K}{v} f(v)$$

for some constant K. We want $f_{V(t)} = \frac{1}{v_{\max}} \mathbf{1}_{\{0 \leq v \leq v_{\max}\}}$, thus

$$f(v) = \frac{v}{K v_{\max}} \mathbf{1}_{\{0 \leq v \leq v_{\max}\}}$$

We can choose K such that the integral of f is 1 and this becomes a well defined PDF. The answer is yes and the PDF is given by

$$f(v) = \frac{2v}{v_{\max}^2} \mathbf{1}_{\{0 \leq v \leq v_{\max}\}}$$

So any speed value between 0 and $v_{\max}$ is possible but small speeds are chosen with smaller probability. With this choice the RWP has a stationary regime.