Chapter 11

11.1 If Y is a discrete random variable taking values in $\{1, 2, \cdots, n\}$ and for $i = 1, 2, \cdots n$ letX_i be a random variable with distribution function $F_i(x)$ which is independent of Y. Then $X_Y = \sum_{i=1}^n X_i \mathbb{1}(Y = i)$ has density function

$$\mathbb{P}(X_Y \le x) = \sum_{i=1}^n \mathbb{P}(X_Y \le x | Y = i)\mathbb{P}(Y = i) = \sum_{i=1}^n \mathbb{P}(X_i \le x)\mathbb{P}(Y = i) = \sum_{i=1}^n F_i(x)\mathbb{P}(Y = i).$$

Setting $P_i = \mathbb{P}(Y = i)$ brings us to the answer of the first part of the question: We simulate first Y, e.g. by simulating a uniform on $(0, 1)$, say U and setting

$$Y = \min\{y \in \mathbb{N} : \sum_{i=1}^y P_i \ge U\}$$

and then simulating from F_Y.

We can apply this to $P_1 = 1/3$, $P_2 = 2/3$, $F_1(x) = 1 - e^{-2x}$ for $x \in (0, \infty)$ and $F_2(x) = x$ for $x \in (0, 1)$.

11.5 Let X_i be a random variable with distribution function $F_i(x)$ for $i = 1, 2, \cdots, n$ and assume that the X_i's are independent. Then,

$$\prod_{i=1}^n F_i(x) = \prod_{i=1}^n \mathbb{P}(X_i \le x) = \mathbb{P}(\text{ all } X_i \text{ are at most } x) = \mathbb{P}(\max X_i \le x).$$

So, you can simulate from $\prod_{i=1}^n F_i(x)$ by simulating from the F_i's separately and take the maximum of the simulated values. Similarly,

$$1 - \prod_{i=1}^n (1 - F_i(x)) = 1 - \prod_{i=1}^n \mathbb{P}(X_i > x) = 1 - \mathbb{P}(\min X_i > x) = \mathbb{P}(\min X_i \le x).$$

So, you can simulate from $\prod_{i=1}^n F_i(x)$ by simulating from the F_i's separately and take the minimum of the simulated values.

If $F(x) = x^n$ for $x \in (0, 1)$ one can use part a) and simulate n independent random variables with distribution function $x \in (0, 1)$ (that is n independent uniforms) or one can use the inverse distribution method and simulate a single uniform U and compute $U^{1/n}$.

11.7 We are going to use a rejection algorithm. we note that

$$\frac{d}{dx} f(x) = 30(2x - 6x^2 + 4x^3) = 60x(1 - 2x)(1 - x).$$

So, $f(x)$ takes its extrima in $x = 0$, $x = 1/2$ and $x = 1$, where $f(0) = 0$, $f(1/2) = 30/16$ and $f(1) = 0$. So $f(1/2)$ is the maximum of $f(x)$ in $(0, 1)$ and we can use Section 11.2.2 with $g(y) = 1$ on $(0, 1)$ and $c = 30/16$. Then simulate two independent uniforms Y and U on $(0, 1)$ and if $U \le f(Y)/c$ then set $X = Y$ otherwise simulate new Y and U and repeat the procedure.

11.8 a) We use the rejection method for $f(x) = \frac{\lambda^n x^{n-1}}{(n-1)!} e^{-\lambda x}$ and $g(x) = (\lambda/n) e^{-(\lambda/n)x}$. We first compute c which is taken to be the maximum of $f(x)/g(x)$

$$\frac{d}{dx}\frac{f(x)}{g(x)} = \frac{d}{dx}\frac{n(\lambda x)^{n-1}}{(n-1)!} e^{-\lambda \frac{n-1}{n} x} = \left(\frac{n\lambda^{n-1} x^{n-2}}{(n-2)!} - \frac{\lambda^n x^{n-1}}{(n-2)!} \right) e^{-\lambda \frac{n-1}{n} x}$$
$$= (n - \lambda x) \frac{n\lambda^{n-1} x^{n-2}}{(n-2)!} e^{-\lambda \frac{n-1}{n} x}.$$

This derivative is positive for $x \in (0, n/\lambda)$ and negative for $x \in (n/\lambda, \infty)$ and therefore $f(x)/g(x)$ takes its maximum if $x = n/\lambda$ and the maximum is $f(n/\lambda)/g(n/\lambda) = \frac{n^n}{(n-1)!} e^{-(n-1)}$. So, we can set $c = \frac{n^n}{(n-1)!} e^{-(n-1)}$ and from the theory we know this is also the expected number of trials before we accept a proposed realisation of the random variable.

b) Stirling's formula is $n! \approx \sqrt{2\pi n} n^n e^{-n}$. Filling that in, in part a) gives

$$c \approx \frac{n^n e^{-(n-1)}}{(n-1)^{n-1} \sqrt{2\pi(n-1)} e^{-(n-1)}} = \sqrt{\frac{n-1}{2\pi}} \left(\frac{n-1}{n} \right)^{-n} = \sqrt{\frac{n-1}{2\pi}} \left(1 - \frac{1}{n} \right)^{-n}.$$

Since $\left(1 - \frac{1}{n}\right)^n \to e^{-1}$ as $n \to \infty$, we have $\frac{c}{\sqrt{n-1}} \to \frac{1}{\sqrt{2\pi}} e$ as $n \to \infty$.

c) To apply the rejection method we can generate independently an exponential Y_2 with mean 1 and a Uniform U and set $Y = nY_2/\lambda$. Note that for a constant K, K times an exponentially distributed random variable with parameter μ is an exponentially distributed random variable with parameter μ/K. Then applying the rejection method gives. if

$$U \leq \frac{n(\lambda Y)^{n-1}}{(n-1)!} e^{-\lambda \frac{n-1}{n} Y} \Big/ \frac{n^n}{(n-1)!} e^{-(n-1)},$$

that is if

$$U \leq \left(\frac{\lambda Y}{n} \right)^{n-1} e^{-(n-1)(\frac{\lambda Y}{n} - 1)} = (Y_2)^{n-1} e^{-(n-1)(Y_2 - 1)}$$

or

$$-\log U \geq (n-1)[-\log(Y_2) + Y_2 - 1],$$

then set $X = Y = nY_2/\lambda$. Otherwise repeat. Note that $-\log U$ is distributed as an exponential with mean 1 and instead of simulating U and computing $-\log U$ we could have immediately simulated Y_1 and the acceptance inequality would be:

$$Y_1 \geq (n-1)[-\log(Y_2) + Y_2 - 1].$$

d) An independent exponential can be obtained by observing that conditioned on

$$Y_1 \geq (n-1)[-\log(Y_2) + Y_2 - 1]$$

then $Y_1 - (n-1)[-\log(Y_2) + Y_2 - 1]$ is independent of $(n-1)[-\log(Y_2) + Y_2 - 1]$ and exponentially distributed with mean 1, by the memoryless property of the exponential distribution.

11.13 Let U_k be the U random variable generated in the k-th round and Y_k the Y random variable generated in the k-th round. Let A_k be the event that you accept in the k-th round. Assume that we repeat the procedure infinitely many times independently, and we only take X from the first accepted round. So we want to compute $\mathbb{P}(Y_k = i|A_k)$ and show that this is equal to $\mathbb{P}(X = i)$.

$$\mathbb{P}(Y_k = i|A_k) = \frac{\mathbb{P}(Y_k = i \cap A_k)}{\mathbb{P}(A_k)} = \frac{\mathbb{P}(Y_k = i \cap U_k < \frac{P_i}{CQ_i})}{\sum_{j=1}^n \mathbb{P}(Y_k = j \cap U_k < \frac{P_j}{CQ_j})}$$
$$= \frac{\mathbb{P}(Y_k = i)\mathbb{P}(U_k < \frac{P_i}{CQ_i})}{\sum_{j=1}^n \mathbb{P}(Y_k = j)\mathbb{P}(U_k < \frac{P_j}{CQ_j})} = \frac{Q_i \frac{P_i}{CQ_i}}{\sum_{j=1}^n Q_j \frac{P_j}{CQ_j}} = \frac{P_i}{\sum_{j=1}^n P_j} = P_i.$$

To jump from the first to the second line we have used that the U_k's are independent of the Y_k's.

11.30 In this exercise $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ and $f_t(x) = \frac{e^{tx} f(x)}{M(t)}$, where $M(t)$ can be seen as a normalizing constant.

$$e^{tx} f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2 - 2tx\sigma^2}{2\sigma^2}} = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu-t\sigma^2)^2 - t^2\sigma^4 - 2\mu\sigma^2 t}{2\sigma^2}} = e^{t^2\sigma^2/2+\mu t} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu-t\sigma^2)^2}{2\sigma^2}}.$$

Now note that the part from the fraction on is the density of a Normal distribution with mean $\mu + t\sigma^2$ and variance σ^2. The factor $e^{t^2\sigma^2/2+\mu t}$ is a constant as a function of x. Therefore, $f_t(x)$ is a constant (as function of x) times $\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu-t\sigma^2)^2}{2\sigma^2}}$, and because $f_t(x)$ should integrate to 1, the constant should be equal to 1.

11.31 We want to use variance reduction by conditioning. In order to use the method of part b we need that $\mathbb{E}[D_n|W_n] = W_n - \mu$. However

$$\mathbb{E}[D_n|W_n] = \mathbb{E}[W_n - S_n|W_n] = W_n - \mathbb{E}[S_n|W_n].$$

So, we need $\mathbb{E}[S_n|W_n] = \mu$. This is however not the case. It is even possible that we have a queue with infinitely many servers, so $\mathbb{P}(D_n = 0) = 1$, while if $W_n = S_n$ has non-zero variance, $W_n - \mu$ is with positive probability not 0.

If we want to use D_n to simulate W_n, we can use conditioning for variance reduction, since

$$\mathbb{E}[W_n|D_n] = \mathbb{E}[D_n + S_n|D_n] = D_n + \mathbb{E}[S_n|D_n] = D_n + \mu.$$

Because the service time of a customer is independent of how long he or she has been in the queue.

11.32 We use that X and Y are identically distributed and thus that $Var(X) = Var(Y)$ and $Corr(X,Y) = Cov(X,Y)/Var(X)$.

$$Var(\frac{X+Y}{2}) = \frac{1}{4}Var(X+Y) = \frac{1}{4}(VarX+VarY+2Cov(X,Y)) = \frac{Var(X)}{4}(2+2Corr(X,Y)),$$

which is in the interval $Var(X) \times [0,1]$ because the correlation takes values in $[-1,1]$.

11.33 Note that $a\mathbb{E}[X] - \mathbb{E}[X^2] = \mathbb{E}[X(a-X)] \geq 0$, because both factors in the product are in the interval $[0,a]$ and part (a) follows. Using this we obtain that

$$Var(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 \leq a\mathbb{E}[X] - (\mathbb{E}[X])^2 = \mathbb{E}[X](a - \mathbb{E}[X]).$$

The final part follows by observing that because $\mathbb{P}(0 \leq X \leq a) = 1$ and (by finding that $x(a-x)$ takes its maximum in $a/2$) that $\max_{x\in[0,a]} x(a-x) = a^2/4$.

11.23 (a) Let $m(t) = \int_0^t \lambda(s)ds$ and assume that $\lambda(s) > 0$ for $s > \infty$ $\mathbb{P}(X_1 > x) = e^{-m(x)}$, but also

$$e^{-m(x)} = \mathbb{P}(X_1 > x) = \mathbb{P}\left(\int_0^{X_1} \lambda(t)dt > \int_0^x \lambda(t)dt\right) = \mathbb{P}\left(\int_0^{X_1} \lambda(t)dt > m(x)\right),$$

which gives the desired result.

(b) We know that $\mathbb{P}(X_i - X_{i-1} > t | X_{i-1} = s) = e^{-[m(t+s)-m(s)]}$ therefore

$$\mathbb{P}\left(\int_{X_{i-1}}^{X_i} \lambda(t)dt > x | X_{i-1} = s\right) = \mathbb{P}(m(X_i) - m(X_{i-1}) > x | X_{i-1} = s) = \mathbb{P}(m(X_i) > x + m(s) | X_{i-1} = s)$$

Now the only information obtained from $X_{i-1} = s$ on the event $m(X_i) > x + m(s)$ is that $X_i > s$, thus

$$\mathbb{P}\left(\int_{X_{i-1}}^{X_i} \lambda(t)dt > x | X_{i-1} = s\right) = \mathbb{P}(X_i > m^{-1}(x + m(s)) | X_i > s) = e^{-x+m(s)}/e^{-m(s)} = e^{-x}.$$

As desired. Note that the probability above is independent of s and therefore we have independence for the different i.

11.24 Simulate two independent Poisson processes one homogeneous with rate b (e.g. by simulating i.i.d. exponentials with rate b and treat those as the interarrival times) and one inhomogeneous with rate $1/(t+a)$ and use example 11.13 to simulate the second process. Then combine the points of the two processes.

11.17 (a) Let $f(x)$ be the density associated with the distribution function F, then $\lambda(t) = f(t)/(1-F(t)$. We first compute the hazard of $X_{(1)}$ Note that this hazard is given by

$$\lim_{h\searrow 0} h^{-1}\mathbb{P}(X_{(1)} \le t+h|X_{(1)} > t) = \lim_{h\searrow 0} h^{-1}(1-\mathbb{P}(X_{(1)} > t+h|X_{(1)} > t))$$
$$= \lim_{h\searrow 0} h^{-1}\left(1 - \frac{\mathbb{P}(X_i > t+h \text{ for all } i = 1,2,\cdot,n)}{\mathbb{P}(X_i > t \text{ for all } i = 1,2,\cdot,n}\right) = \lim_{h\searrow 0} h^{-1}\left(1 - \frac{(\mathbb{P}(X_1 > t+h))^n}{(\mathbb{P}(X_1 > t))^n}\right)$$
$$= \lim_{h\searrow 0} h^{-1}\left(1 - \frac{(\mathbb{P}(X_1 > t) - f(t)h + o(h))^n}{(1-F(t))^n}\right) = \lim_{h\searrow 0} h^{-1}\left(1 - \left(1 - \frac{f(t)h + o(h)}{1-F(t)}\right)^n\right)$$
$$= n\frac{f(t)}{1-F(t)} = n\lambda(t).$$

More general assume that the j-th "arrival" is at time t_j and let $\mathcal{J}$ be the index set for which the X_i's are larger than t_j, i.e. $\mathcal{J} = \{i \in 1,2,\cdots,n; X_i > t_j\}$. Then for $t > t_j$

$$\lim_{h\searrow 0} h^{-1}\mathbb{P}(X_{(j+1)} \le t+h|X_{(j+1)} > t) = \lim_{h\searrow 0} h^{-1}(1-\mathbb{P}(X_{(1)} > t+h|X_{(1)} > t))$$
$$= \lim_{h\searrow 0} h^{-1}\left(1 - \frac{\mathbb{P}(X_i > t+h \text{ for all } i \in \mathcal{J}}{\mathbb{P}(X_i > t \text{ for all } i \in \mathcal{J}}\right) = \lim_{h\searrow 0} h^{-1}\left(1 - \frac{(\mathbb{P}(X_1 > t+h))^{n-j}}{(\mathbb{P}(X_1 > t))^{n-j}}\right)$$
$$= \cdots (\text{as above}) \cdots = (n-j)\frac{f(t)}{1-F(t)} = (n-j)\lambda(t).$$

So, conditioned on the j-th arrival being at time t_j, the $j+1$-st arrival has hazard $(n-j)\lambda(t)$ for $t > t_j$.

(b) Since F is continuous and not decreasing F^{-1} is increasing. Let $U_1, U_2, \cdots, U_n$ be independent and identically distributed uniform random variables on $(0,1)$ and for $i = 1,2,\cdots,n$ set $Y_i = F^{-1}(U_i)$. Note that Y_i is distributed as X_i and that because F^{-1} is increasing $Y_{(i)} = f^{(-1)}(U_{(i)})$. Which gives the desired result. Now the i-th order statistic of n uniforms is Beta distributed with parameters i and $n+i+1$.

(c) This is actually the order statistic property for a Poisson process with intensity 1. The numerator is distributed as the i-th point of the Poisson process, while the numerator is distributed as the $n+1$-st point. By the order statistic property, conditioned on the $n+1$-st point being at time t the positions of the first n points are distributed as n independent identically distributed uniforms on $(0,t)$ and the positions of the first n points divided by t are distributed as n independent identically distributed uniforms on $(0,1)$ as desired.

(d) Use the order statistic property as in part c. We know that $S_n = y(Y_1 + \cdots + Y_{n+1})$ and the points $S_1, \cdots S_{n-1}$ are then distributed as $n-1$ independent uniforms on $(0, S_n) = (0, y)$.

(e) Let $V_1, V_2, \cdots$ be independent identically distributed uniforms on $(0,1)$. Note that $\mathbb{P}(V_{(n)} \le x) = \prod_{j=1}^n \mathbb{P}(V_j \le x) = x^n$ for $x \in (0,1)$, while $\mathbb{P}(U_1^{1/n} \le x) = \mathbb{P}(U_1 \le x^n) = x^n$ for $x \in (0,1)$, which shows the first line of step II. The remaining part is obtained by using part (d) and noting that $U_{(j-1)}/U_{(j)}$ is distributed as the maximum of $j-1$ i.i.d. uniforms on $(0,1)$.