Solutions to selected exercises from Chapter 23 of *Wasserman — All of Statistics*

(1) We have

$$\mathbb{P}(X_0 = 0, X_1 = 1, X_2 = 2) = \mathbb{P}(X_2 = 2|X_1 = 1)\mathbb{P}(X_1 = 1|X_0 = 0)\mathbb{P}(X_0 = 0)$$

and this is equal to $0 \cdot 0.2 \cdot 0.3 = 0$.

Similarly the second probability is

$$\mathbb{P}(X_2 = 1|X_1 = 1)\mathbb{P}(X_1 = 1|X_0 = 0)\mathbb{P}(X_0 = 0) = 0.1 \cdot 0.2 \cdot 0.3 = 0.006.$$

(2) If $Y_n > \max\{Y_1, \dots, Y_{n-1}\}$ then $X_n = Y_n$ and otherwise $X_n = X_{n-1} = \max\{Y_1, \dots, Y_{n-1}\}$. This shows that $\mathbb{P}(X_n = x|X_0, \dots, X_{n-1}) = \mathbb{P}(X_n = x|X_{n-1})$ and therefore $X_0, X_1, \dots$ is a Markov chain.

If $X_n = x$ then $\mathbb{P}(X_{n+1} = x) = \sum_{y \le x} \mathbb{P}(Y = y)$ and for $y > x$, $\mathbb{P}(X_{n+1} = y) = \mathbb{P}(Y = y)$. Thus the transition matrix is

$$\begin{pmatrix} 0.1 & 0.3 & 0.2 & 0.4 \\ 0 & 0.4 & 0.2 & 0.4 \\ 0 & 0 & 0.6 & 0.4 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

(3) The distribution $\pi = \frac{1}{a+b}(b, a)$ is a stationary distribution for $\mathbf{P}$. By Theorem 23.25 (related to the Perron-Frobenius theorem), $\mathbf{P}$ has limiting distribution π, i.e. $\mathbf{P}^n \to \begin{pmatrix} \pi \\ \pi \end{pmatrix}$.

(4) See the Jupyter Notebook 4.ipynb.

(5) (a) We have

$$\mathbb{E}(X_{n+1}|X_n) = \mathbb{E}\left(Y_1^{(n)}\right) + \dots + \mathbb{E}\left(Y_{X_n}^{(n)}\right) = \mu X_n.$$

By the Rule of Iterated Expectations,

$$\mathbb{E}(X_{n+1}) = \mathbb{E}(\mathbb{E}(X_{n+1}|X_n)) = \mathbb{E}(\mu X_n) = \mu\mathbb{E}(X_n) = \mu M(n).$$

The Rule of Iterated Expectations for variance states that

$$\mathbb{V}(X_{n+1}) = \mathbb{E}(\mathbb{V}(X_{n+1}|X_n)) + \mathbb{V}(\mathbb{E}(X_{n+1}|X_n)).$$

By independence of the $Y_i^{(n)}$,

$$\mathbb{V}(X_{n+1}|X_n) = \mathbb{V}(Y_1^{(n)}) + \dots + \mathbb{V}(Y_{X_n}^{(n)}) = X_n\mathbb{V}(Y) = \sigma^2 X_n.$$

The expected value of this quantity is $\sigma^2 M(n)$. By our previous computations,

$$\mathbb{V}(\mathbb{E}(X_{n+1}|X_n)) = \mu^2\mathbb{V}(X_n) = \mu^2 V(n).$$

Adding these two terms yields $V(n+1) = \sigma^2 M(n) + \mu^2 V(n)$ as claimed.

(b) It follows by induction that $M(n) = \mu^n$.
We have $V(0) = 0$ and $V(1) = \sigma^2$. Suppose for induction that $V(n) = \sigma^2\mu^{n-1}(1 + \cdots + \mu^{n-1})$. Then

$$V(n+1) = \sigma^2\mu^n + \sigma^2\mu^{n+1}(1 + \cdots + \mu^{n-1}) = \sigma^2\mu^n(1 + (\mu + \cdots + \mu^n))$$

which is the desired equation for $V(n+1)$.

(c) If $\mu \geq 1$ then $V(n) \to \infty$ as $n \to \infty$. If $\mu < 1$ then $1 + \mu + \cdots + \mu^{n-1}$ converges to a finite value c and $\mu^{n-1} \to 0$ so that $V(n) \to 0$.

(d) If there are k children of the original parent then for each child, $F(n-1)$ is the probability of all generations of their offspring eventually dying off by generation $n-1$. Hence $F(n-1)^k$ is the probability of all k lineages of children of the original parent dying off, assuming that the original parent had k children. The probability of k children is p_k. Adding up the probabilities over all possible numbers k of children yields

$$\sum_{k=0}^{\infty} p_k F(n-1)^k.$$

(e) There doesn't seem to be a nice closed form solution for $F(n)$ here.

$$F(0) = 0,\ F(1) = \frac{1}{4},\ F(2) = \frac{5^2}{2^6},\ F(3) = \frac{89^2}{2^{14}},\ F(4) = \frac{5^2 \cdot 4861^2}{2^{30}},\ \ldots$$

(6) The matrix $\mathbf{P}$ is equal to

$$\frac{1}{20}\begin{pmatrix} 8 & 10 & 2 \\ 1 & 14 & 5 \\ 1 & 10 & 9 \end{pmatrix} = \frac{1}{20}\mathbf{M}$$

Setting $(x, y, z)\mathbf{M} = 20(x, y, z)$ yields

$$\begin{aligned} -12x + y + z &= 0 \\ 10x - 6y + 10z &= 0 \\ 2x + 5y - 11z &= 0 \end{aligned}.$$

Solving this system yields $\pi = (8, 65, 31)$ up to scalar multiples.

(7) We have

$$p_{jj}(n) = \sum_{k+l+m=n} p_{ij}(m)p_{ii}(l)p_{ji}(k).$$

Choose m and k such that $p_{ij}(m) = p > 0$ and $p_{ji}(k) = q > 0$. Then for any $n \geq m + k$

$$p_{jj}(n) \geq p_{ij}(m)p_{ii}(n-m-k)p_{ji}(k) = pqp_{ii}(n-m-k).$$

Thus,

$$\sum_{n=0}^{\infty} p_{jj}(n) \geq \sum_{n=m+k}^{\infty} p_{jj}(n) \geq pq\sum_{n=0}^{\infty} p_{ii}(n) = \infty.$$

(8) The equivalence classes under $\leftrightarrow$ are $\{3,5\},\{1\},\{2\},\{4\},\{6\}$. The classes $\{3,5\}$ and $\{6\}$ are recurrent. I.e. $\mathbb{P}(X_2 = 5|X_0 = 5) = 1$, $\mathbb{P}(X_2 = 3|X_0 = 3) = 1$, and $\mathbb{P}(X_1 = 6|X_0 = 6) = 1$. We claim the other classes are transient. Note that $\{3,5\}$ and $\{6\}$ are closed sets. So if $X_0 = 1$, then with probability $2/3$, $X_i \neq 1$ for $i > 0$ and otherwise $X_1 = 1$. Thus

$$\mathbb{P}(X_i = 1 \text{ for some } i > 0|X_0 = 1) = 1/3.$$

This shows that 1 is transitive. If $X_0 = 2$ then with probability $3/4$, $X_1 = 1$ or $X_1 = 3$. In these cases $X_i \neq 2$ for any $i > 0$. Reasoning similarly to before,

$$\mathbb{P}(X_i = 2 \text{ for some } i > 0|X_0 = 2) = 1/4$$

so 2 is transitive. Finally,

$$\mathbb{P}(X_i = 4 \text{ for some } i > 0|X_0 = 4) = 0$$

so 4 is also transitive.

(9) $\pi = (1/2, 1/2)$ is clearly stationary. The resulting Markov chain does not converge since the powers of $\mathbf{P}$ alternate between $\mathbf{P}$ and the identity.

(10) The limiting distribution is

$$\begin{pmatrix} \pi \\ \pi \\ \pi \\ \pi \\ \pi \end{pmatrix} \text{ where } \pi = \frac{1-p}{1-p^5}\begin{pmatrix} 1 & p & p^2 & p^3 & p^4 \end{pmatrix}.$$

For if $\mathbf{P}^n$ converges to a matrix with all rows equal to π, then $\mathbf{P}^{n+1}$ converges to $\pi\mathbf{P}$ and therefore $\pi\mathbf{P} = \pi$. Solving the system $\pi\mathbf{P} = \pi$ yields $\pi = \begin{pmatrix} 1 & p & p^2 & p^3 & p^4 \end{pmatrix}$ up to taking a scalar multiple.

(11) We check the items in the definition of a Poisson process.

For the first item, note that $\Lambda(0) = 0$ so $Y(0) = Y(\Lambda(0)) = X(0) = 0$.

For the second item, note that Λ is monotone increasing since $\lambda(t) \geq 0$. If $0 = s_0 < s_1 < \ldots < s_n$ then we can choose $t_i \in \Lambda^{-1}(s_i)$ for each i and necessarily $0 = t_0 < t_1 < \ldots < t_n$. Thus,

$$Y(s_1) - Y(s_0), \ldots, Y(s_n) - Y(s_{n-1})$$

are equal to

$$X(t_1) - X(t_0), \ldots, X(t_n) - X(t_{n-1}).$$

Since the latter random variables are independent, so are the former.

For the last item, assuming that λ is continuous, note that

$$\int_t^{t+h} \lambda(x)dx - \lambda(t)h = o(h) \text{ since } \lim_{h\to 0} \frac{\int_t^{t+h} \lambda(x)dx}{h} = \lambda(t).$$

Given s and k, choose t and h such that $\Lambda(t) = s$ and $\Lambda(t+h) = s+k$. Then $k = \Lambda(t+h) - \Lambda(t) = \int_t^{t+h} \lambda(x)dx = \lambda(t)h + o(h)$. Thus,

$$\frac{\mathbb{P}(Y(s+k) - Y(s) = 1) - k}{k} = \frac{\mathbb{P}(X(t+h) - X(t) = 1) - \int_t^{t+h} \lambda(x)dx}{k} = \frac{o(h)}{k} = \frac{o(h)/h}{k/h}.$$

As $k \to 0$, the numerator approaches 0 and the denominator approaches $\lambda(t)$. Thus $\frac{\mathbb{P}(Y(s+k)-Y(s)=1)-k}{k} \to 0$ so that $\mathbb{P}(Y(s+k) - Y(s) = 1) = k + o(k)$. For the second part of the last item, choose s, k, t, h as before,

$$\frac{\mathbb{P}(Y(s+k) - Y(s) \geq 2)}{k} = \frac{\mathbb{P}(X(t+h) - X(t) \geq 2)}{k} = \frac{o(h)}{k}$$

and we showed this goes to zero above. Thus $\mathbb{P}(Y(s+k) - Y(s) \geq 2) = o(k)$.

(12) For $0 \leq m \leq n$, we have

$$\mathbb{P}(X(t) = m | X(t+s) = n) = \frac{\mathbb{P}(X(t) = m, X(t+s) = n)}{\mathbb{P}(X(t+s) = n)}$$

which is equal to

$$\frac{\mathbb{P}(X(t+s) - X(t) = n-m, X(t) = m)}{\mathbb{P}(X(t+s) = n)} = \frac{\mathbb{P}(X(t+s) - X(t) = n-m)\mathbb{P}(X(t) = m)}{\mathbb{P}(X(t+s) = n)}.$$

Using Theorem 23.33, this is equal to

$$\frac{e^{-m(s+t)+m(t)}e^{-m(t)}}{e^{-m(s+t)}} \frac{n!}{m!(n-m)!} \frac{((m(s+t) - m(t))^{n-m} m(t)^m}{m(s+t)^m}$$

which simplifies to

$$\binom{n}{m} \left(\frac{m(s+t) - m(t)}{m(s+t)} \right)^{n-m} \left(\frac{m(t)}{m(s+t)} \right)^m.$$

Thus, we recognize the distribution of $X(t) | (X(t+s) = n)$ as a Binomial $\left(n, \frac{m(t)}{m(s+t)}\right)$ distribution.

(13) We have $X(t) \sim \text{Poisson}(\lambda)$ where $\lambda = m(t)$. Thus,

$$\mathbb{P}(X(t) \text{ is odd}) = e^{-\lambda} \left(\frac{\lambda}{1!} + \frac{\lambda^3}{3!} + \dots \right) = e^{-\lambda} \frac{1}{2}(e^{\lambda} + e^{-\lambda}) = \frac{1}{2}(1 + e^{-2\lambda}).$$

The probability is $(1 + e^{-m(t)})/2$.