Solutions to Exercises 7.3.2 through 7.3.5, and 7.3.7 through 7.3.11.

7.3.2. (a)

$$\begin{aligned}
\mathrm{E}_{\theta,\phi}(d_N(T_N)) &= \mathrm{E}_{\theta,\phi}\{\mathrm{E}_\phi(d_N(T_N)|N,T_N)\} \\
&= \mathrm{P}(N=2,T_2=1)d_2(1)+\mathrm{P}(N=2,T_2=2)d_2(2)+\mathrm{P}(N=3,T_3=0)d_3(0) \\
&\quad +\mathrm{P}(N=3,T_3=1)d_3(1)+\mathrm{P}(N=3,T_3=2)d_3(2) \\
&= \theta(1-\theta)d_2(1)+\theta^2 d_2(2)+(1-\theta)^3 d_3(0)+2\theta(1-\theta)d_3(1)+\theta^2(1-\theta)d_3(2) \\
&= d_3(0)+\theta[d_2(1)-3d_3(0)+2d_3(1)] \\
&\quad +\theta^2[-d_2(1)+d_2(2)+3d_3(0)-4d_3(1)+d_3(2)]+\theta^3[-d_3(0)+2d_3(1)-d_3(2)].
\end{aligned}$$

(b) Equating coefficients in $\mathrm{E}_{\theta,\phi}(d_N(T_N))=\theta$, we find

$$\begin{aligned}
d_3(0) &= 0 \\
d_2(1)+2d_3(1) &= 1 \\
d_2(2)+d_3(2) &= d_2(1)+4d_3(1) \\
2d_3(1) &= d_3(2)
\end{aligned}$$

If we let $z=d_3(2)$, the class of unbiased nonrandomized estimates is: $d_2(1)=1-z$, $d_2(2)=1$, $d_3(0)=0$, $d_3(1)=z/2$ and $d_3(2)=z$ for $o\le z\le 1$.

(c) The expected loss as a function of z and θ is

$$R(\theta,z)=\mathrm{E}_{\theta,\phi}(\theta-d_N(T_N))^2=\theta(1-\theta)\{z^2(\theta+3)/2-2z+1\}.$$

(d) If $z>2/3$, then $R(\theta,z)-R(\theta,2/3)=\theta(1-\theta)\{(z-2/3)[(z+2/3)(\theta+3)/2-2]\}\ge 0$ for all θ. This shows that the unbiased estimate with $d_3(2)=z$ is improved by the estimate with $d_3(2)=2/3$.

If $z<1/2$, then $R(\theta,z)-R(\theta,1/2)=\theta(1-\theta)\{(1/2-z)[2-(1/2-z)(\theta+3)/2]\}\ge 0$ for all θ. This shows that the unbiased estimate with $d_3(2)=z$ is improved by the estimate with $d_3(2)=1/2$.

7.3.3. X_1 is $\mathcal{B}(1,1/2)$, and X_2 given $X_1=0$ is $\mathcal{B}(1,\theta)$ while X_2 given $X_1=1$ is $\mathcal{B}(1,1/2)$.

$$\begin{aligned}
R(\theta,(\phi,\delta)) &= \mathrm{E}_{\theta,\phi}(\mathrm{E}_\phi(L(\theta,d_N(X_1,\ldots,X_N))+c(\theta,X_1,\ldots,X_N)|N)) \\
&= \mathrm{P}_\phi(N=1)[(\theta-\frac{1}{2})^2+c]+\mathrm{P}_\phi(N=2)[\theta(1-\theta)^2+(1-\theta)\theta^2+2c] \\
&= \frac{1}{2}(\theta-\frac{1}{2})^2+\frac{c}{2}+\frac{1}{2}\theta(1-\theta)+c=\frac{1}{8}+\frac{3}{2}c \\
R(\theta,(\phi^0,\delta^0)) &= \mathrm{P}_{\phi^0}(X_1=1)(\theta-\frac{1}{2})^2+P_{\phi^0,\theta}(X_1=0,X_2=1)(1-\theta)^2+\mathrm{P}_{\phi^0,\theta}(X_1=0,X_2=0)\theta^2+2c \\
&= \frac{1}{2}(\frac{1}{2}-\theta)^2+\frac{1}{2}\theta(1-\theta)^2+\frac{1}{2}(1-\theta)\theta^2+2c=\frac{1}{8}+2c.
\end{aligned}$$

7.3.4. The Bayes risk for a rule that takes no observations is

$$.5L(0,a)+.5L(1,a)=.5(a^2+(1-a)^2)=a^2-a+.5$$

with minimum value 1/4, taken on at $a=.5$. For a rule that observes (X_1,X_2), the posterior distribution of θ given $X_1=1,X_2=1$ or $X_1=1,X_2=0$ is the same as the prior and so the Bayes rule for these points is $d(1,1)=d(1,0)=.5$ also. The posterior of θ for $X_1=0,X_2=1$ is degenerate at 1, so the Bayes rule is $d(0,1)=1$; similarly, the posterior of θ given $X_1=0,X_2=0$ is degenerate at zero giving $d(0,0)=0$ as the Bayes rule. Thus the rule (ϕ^0,δ^0) is Bayes with respect to this prior if its Bayes risk, $1/8+2c$, is not greater than that of taking no observations, 1/4. This reduces to the condition, $c\le 1/16$.

7.3.5. (a) First we find $\varphi_n^0(T_n)=\mathrm{P}_\varphi(N=n|N\ge n,T_n)$. Since φ always takes at least two observations, we have $\varphi_0^0=0$ and $\varphi_1^0=0$. For $n=2$, $\varphi_2^0(t)=\mathrm{P}_\varphi(N=2|T_2=t)$. If $T_2=0$, then $X_2=0$

so that $\varphi_2^0(0) = 0$. If $T_2 = 1$ then $X_1 = 0, X_2 = 1$ and $X_1 = 1, X_2 = 0$ are equally likely, so that $\varphi_2^0(1) = \mathrm{P}_\varphi(N = 2|T_2 = 1) = \mathrm{P}_\varphi(X_2 = 1|T_2 = 1) = 1/2$. If $N \geq 3$, then we know that $X_2 = 0$ so that $\varphi_3^0(t) = \mathrm{P}_\varphi(N = 3|X_2 = 0, T_3 = t)$. We find similarly, $\varphi_3^0(0) = 0$, $\varphi_3^0(1) = 1/2$ and $\varphi_3^0(2) = 1$. (If $N \geq 3$, then T_3 cannot be equal to 3.) Finally, $\varphi_4(t) \equiv 1$.

The terminal decision rule is given by $\delta_j^0(t) = \mathrm{E}_\varphi(\delta_j(X_1, \ldots, X_j)|N = j, T_j = t)$, the mixture of the distributions δ_j using the mixing distribution of $X_1, \ldots, X_j$ given $N = j$ and $T_j = t$. We never stop before stage 2, so δ_0^0 and δ_1^0 are undefined. If $N = 2$, then $X_2 = 1$ so that $\delta_2^0(t) = 1$ for all t. Similarly if $N = 3$, then $X_3 = 1$ so that $\delta_3^0(t) = 1$ for all t. If $N = 4$, then $X_2 = 0$ and $X_3 = 0$ so that $T_4 = X_1 + X_4$. We then compute $\delta_4^0(0) = 0$ w.p. 1, $\delta_4^0(1) = 0$ w.p. 1/2, and $= 1$ w.p. 1/2, and $\delta_4^0(2) = 1$ w.p. 1.

(b) To find the nonrandomized rule that improves on δ^0, we replace each δ_j^0 by its expectation. Thus, $d_2^0(t) = 1$ for all t, $d_3^0(t) = 1$ for all t, and $d_4^0(t) = 0$ for $t = 0$, $= 1/2$ for $t = 1$ and $= 1$ for $t = 1$.

7.3.7. (a) Since T_n is a sufficient statistic, and $T_n \in \mathcal{P}(n\theta)$, we have for nonnegative integers $x_1, \ldots, x_n$ such that $x_1 + \cdots + x_n = t$,

$$\begin{aligned} P(X_1 = x_1, \ldots, X_n = x_n|T_n = t) &= \frac{f_{X_1,\ldots,X_n}(x_1, \ldots, x_n|\theta)}{f_{T_n}(t|\theta)} \\ &= \frac{\exp\{-n\theta\}\theta^{x_1+\cdots+x_n}/(x_1!\cdots x_n!)}{\exp\{-n\theta\}(n\theta)^t/t!} \\ &= \frac{t!}{x_1!\cdots x_n!}(\frac{1}{n})^t. \end{aligned}$$

This is the multinomial distribution with n cells with equal probabilities $1/n$, and sample size t.

(b) By Theorem 3, $\phi(\mathbf{d}^0)$ is as good as $(\phi, \mathbf{d})$, where for $t_n \geq 2$,

$$\begin{aligned} d_n^0(t_n) &= \mathrm{E}_\phi(d_n(X_1, \ldots, X_n)|N = n, T_n = t_n) \\ &= \mathrm{E}(X_1|X_1 + \cdots + X_{n-1} < 2, T_n = t_n) \\ &= \mathrm{P}(X_1 = 1|X_1 + \cdots + X_{n-1} < 2, T_n = t_n) \\ &= \frac{\mathrm{P}(X_1 = 1, X_1 + \cdots + X_{n-1} < 2|T_n = t_n)}{\mathrm{P}(X_1 + \cdots + X_{n-1} < 2|T_n = t_n)} \\ &= \frac{\mathrm{P}(X_1 = 1, X_2 = 0, \ldots, X_{n-1} = 0|T_n = t_n)}{\mathrm{P}(X_1 + \cdots + X_{n-1} = 0|T_n = t_n) + \mathrm{P}(X_1 + \cdots + X_{n-1} = 1|T_n = t_n)} \\ &= \frac{(t_n!/(T_n - 1)!)(1/n)^{t_n}}{(1/n)^{t_n} + (n-1)(t_n!/(t_n - 1)!)(1/n)^{t_n}} = \frac{t_n}{1 + (n-1)t_n}. \end{aligned}$$

7.3.8. (a) Since T_n is a sufficient statistic, and $T_n \in \mathcal{NB}(n, \theta)$, we have for nonnegative integers $x_1, \ldots, x_n$ such that $x_1 + \cdots + x_n = t$,

$$\begin{aligned} P(X_1 = x_1, \ldots, X_n = x_n|T_n = t) &= \frac{f_{X_1,\ldots,X_n}(x_1, \ldots, x_n|\theta)}{f_{T_n}(t|\theta)} \\ &= \frac{(1-\theta)^n\theta^{x_1+\cdots+x_n}}{\binom{n+t-1}{t}(1-\theta)^n\theta^t} \\ &= \frac{1}{\binom{n+t-1}{t}}. \end{aligned}$$

(Note the misprint in the text.)

(b) For $n > 1$ and $t \geq n - 1$,

$$\mathrm{E}_\phi(X_1|N = n, T_n = t) = \cdots = \mathrm{E}_\phi(X_{n-1}|N = n, T_n = t) \quad \text{and} \quad \mathrm{E}_\phi(X_n|N = n, T_n = t) = 0$$

so that

$$t = \mathrm{E}_\phi(\sum_1^n X_i|N = n, T_n = t) = (n-1)\mathrm{E}_\phi(X_1|N = n, T_n = t)$$

which gives $d_n^0(t) = \mathrm{E}_\phi(X_1|N = n, T_n = t) = t/(n-1)$. For $n = 1$, we must have $X_1 = 0$, so $d_0^0(t) = \mathrm{E}_\phi(X_1|N = 1, T_1 = 0) = 0$.

(c) Automatically, $\phi_0^0 = \phi_0$. For $n = 1$,

$$\phi_1^0(t) = \mathrm{P}_\phi(N = 1|N \geq 1, T_1 = t) = \phi_1(t)$$

while for $n > 1$ and $t \geq n-1$,

$$\pi_n^0(t) = \frac{\mathrm{P}_\phi(N = n|T_n = t)}{\mathrm{P}_\phi(N \geq n|T_n = t)} = \frac{\mathrm{P}(X_1 > 0, \ldots, X_{n-1} > 0, X_n = 0|T_n = t)}{\mathrm{P}(X_1 > 0, \ldots, X_{n-1} > 0|T_n = t)}.$$

To compute the numerator probability, note that the number of points in the set $\{(x_1, \ldots, x_{n-1}) : x_i > 0, \sum_1^{n-1} x_i = t\}$ is the same as the number of points in $\{(x_1, \ldots, x_{n-1}) : x_i \geq 0, \sum_1^{n-1} x_i = t - (n-1)\}$, which is $\binom{(n-1)+(t-(n-1))-1}{t-(n-1)} = \binom{t-1}{t-n+1}$. Hence

$$\mathrm{P}(X_1 > 0, \ldots, X_{n-1} > 0, X_n = 0|T_n = t) = \frac{\binom{t-1}{t-n+1}}{\binom{n+t-1}{t}}.$$

Similarly,

$$\mathrm{P}(X_1 > 0, \ldots, X_{n-1} > 0|T_n = t) = \frac{\binom{t}{t-n+1}}{\binom{n+t-1}{t}}.$$

Hence,

$$\phi_n^0(t) = \frac{\binom{t-1}{t-n+1}}{\binom{t}{t-n+1}} = \frac{n-1}{t}.$$

7.3.9. We must show $\mathrm{P}(X_1 = x_1, \ldots, x_j = x_j|T_j = t, T_{j+1} = t + x)$ does not depend on x for $x = 0$ or 1.

$$\begin{aligned} \mathrm{P}(X_1 = x_1, \ldots, x_j = x_j|T_j = t, T_{j+1} = t+x) &= \frac{\mathrm{P}_\theta(X_1 = x_1, \ldots, X_j = x_j, T_j = t, X_{j+1} = x)}{\mathrm{P}_\theta(T_n = t, X_{j+1} = x)} \\ &= \frac{\mathrm{P}_\theta(X_{j+1} = x|X_1 = x_1, \ldots, X_j = x_j, T_j = t)\mathrm{P}_\theta(X_1 = x_1, \ldots, X_j = x_j, T_j = t)}{\mathrm{P}_\theta(X_{j+1} = x|T_j = t)\mathrm{P}_\theta(T_j = t)} \end{aligned}$$

The first terms in numerator and denominator are equal (when $x = 1$ both are equal to $(\theta - t)/(M - j)$). These cancel, showing T_n is transitive.

7.3.10. For any set A,

$$\begin{aligned} \mathrm{P}_{\theta,\phi}((X_1, \ldots, X_n) \in A|N = n, T_n = t) &= \frac{\mathrm{P}_{\theta,\phi}((X_1, \ldots, X_n) \in A, N = n|T_n = t)}{\mathrm{P}_{\theta,\phi}(N = n|T_n = t)} \\ &= \frac{\mathrm{E}(I_A(X_1, \ldots, X_n)\psi_n(X_1, \ldots, X_n)|T_n = n)}{\mathrm{E}(\psi_n(X_1, \ldots, X_n)|T_n = t)}. \end{aligned}$$

Note that this depends on the distribution of $X_1, \ldots, X_n$ only through the conditional distribution of $(X_1, \ldots, X_n)$ given $T_n = t$. But for $n \leq M$, $X_1, \ldots, X_n$ in Exercise 9, and $X_1, \ldots, X_n$ from independent Bernoulli trials have the same conditional distribution given $T_n = t$, hence they have the same conditional distribution given $N = n$ and $T_n = t$.

7.3.11. Let $X_1, X_2, \ldots$ be independent Bernoulli trials with $\mathrm{P}_\theta(X_1 = 1) = 1/2$, and $\mathrm{P}_\theta(X_i = 1) = \theta$ for all $i > 1$. Then $T_1 \equiv 0$, $T_2 = (X_1, X_2)$,..., $T_n = (X_1, \sum_2^n X_i)$,dots, forms a sufficient sequence for θ. But $\mathrm{E}(X_1|T_1) = 1/2$ and $\mathrm{E}(X_1|T_1, T_2 = (x_1, x_2)) = x_1$. Since these quantities differ, the sequence is not transitive.