STA 247 — Practice problem set #4 solutions

Question 1: Suppose you buy a disk drive from a manufacturer who makes them in two factories. The two factories produce equal numbers of disk drives, and one cannot tell which factory a drive was made in from looking at it. However, drives from factory A are more reliable than drives from factory B — the time (in years) from purchase to failure for a drive from factory A has the exp(1/3) distribution (with mean 3), whereas the time from purchase to failure for a drive from factory B has the exp(1/2) distribution (with mean 2).

a) Suppose your drive fails after 4 years. What is the probability that it was made in factory A?

We can use Bayes' Rule to find the probability of the event, A, of your disk being made in factory A, given that it failed after $T = 4$ years of use. For this purpose, we can treat the probability density for T as if it were a probability, since this gives the same result as treating "after 4 years" as "between 3.9999 and 4.0001 years" — the probability of the latter is close to 0.0002 times the probability density for $T = 4$, and the factor of 0.0002 will cancel in Bayes' Rule.

Let $f_{T|A}(t)$ be the probability density for the time to failure for a drive made in factory A, and let $f_{T|B}(t)$ be the probability density for the time of failure for a drive made in factory B. Since the two factories make equal numbers of drives, $P(A) = P(B) = 1/2$ (noting that $B = A^c$). Bayes' Rule gives

$$P(A|T = 4) \;=\; \frac{(1/2)\, f_{T|A}(4)}{(1/2)\, f_{T|A}(4) \;+\; (1/2)\, f_{T|B}(4)}$$

Substituting the density functions for the $\exp(1/3)$ *and* $\exp(1/2)$ *distributions, we get*

$$P(A|T = 4) \;=\; \frac{(1/2)\,(1/3)\exp(-4/3)}{(1/2)\,(1/3)\exp(-4/3) \;+\; (1/2)\,(1/2)\exp(-4/2)} \;=\; 0.5649$$

We could compute this formula in R, or instead use R's density function for exponential distributions, as follows:

```
(1/2)*dexp(4,1/3) / ((1/2)*dexp(4,1/3)+(1/2)*dexp(4,1/2))
```

b) Suppose that after 4 years of use, your drive still hasn't failed. What is the probablity that it was made in factory A?

We can again use Bayes' Rule, but now the event conditioned on is $T > 4$. We can integrate the probability density to find that

$$P(T > 4|A) \;=\; \int_4^\infty (1/3)\exp(-t/3)\, dt \;=\; \Big[-\exp(-t/3)\Big]_4^\infty \;=\; \exp(-4/3)$$

$$P(T > 4|B) \;=\; \int_4^\infty (1/2)\exp(-t/2)\, dt \;=\; \Big[-\exp(-t/2)\Big]_4^\infty \;=\; \exp(-4/2)$$

Applying Bayes' Rule, we get

$$\begin{aligned} P(A|T>4) &= \frac{(1/2)\,P(T>4|A)}{(1/2)\,P(T>4|A)+(1/2)\,P(T>4|B)} \\ &= \frac{(1/2)\,\exp(-4/3)}{(1/2)\,\exp(-4/3)+(1/2)\,\exp(-4/2)} \;=\; 0.6608 \end{aligned}$$

We can also find this answer using R's cumulative distribution function for exponential distributions, noting that $T > 4$ is the complement of $T \le 4$, whose probability is the CDF at $T = 4$:

```
(1/2)*(1-pexp(4,1/3)) / ((1/2)*(1-pexp(4,1/3))+(1/2)*(1-pexp(4,1/2)))
```

Question 2: Bits sent through a communications channel are sometimes received with the wrong value. For some channels, such errors often occur in bursts, with several errors occurring in a row. We can model such error behaviour using a Markov chain. Let E_i be the random variable having the value 1 if an error occurred in bit i, and 0 otherwise. Suppose that these errors have the Markov property, so that

$$P(E_i = e_i \,|\, E_{i-1} = e_{i-1}) \;=\; P(E_i = e_i \,|\, E_{i-1} = e_{i-1},\, E_{i-2} = e_{i-2}, \ldots, E_0 = e_0)$$

We can then specify the error behaviour of the channel by the one-step transition probabilities for this Markov chain. Suppose that these transition probabilities are the same at all times (ie, the Markov chain is homogeneous). The one-step transition probabilities will then be determined by just two numbers, $P^{(1)}(0 \to 1) = P(E_i = 1 \,|\, E_{i-1} = 0)$ and $P^{(1)}(1 \to 1) = P(E_i = 1 \,|\, E_{i-1} = 1)$.

a) Find $P(E_{i+3} = 1 \,|\, E_i = 1)$ exactly, assuming that $P^{(1)}(0 \to 1) = 0.01$ and $P^{(1)}(1 \to 1) = 0.4$.

If we start with state probabilities given by the vector $[0\ 1]$, *then the state probabilities after three transitions will be*

$$\begin{aligned} & [0\ 1] \begin{bmatrix} 0.99 & 0.01 \\ 0.6 & 0.4 \end{bmatrix} \begin{bmatrix} 0.99 & 0.01 \\ 0.6 & 0.4 \end{bmatrix} \begin{bmatrix} 0.99 & 0.01 \\ 0.6 & 0.4 \end{bmatrix} \\ &= [0.6\ 0.4] \begin{bmatrix} 0.99 & 0.01 \\ 0.6 & 0.4 \end{bmatrix} \begin{bmatrix} 0.99 & 0.01 \\ 0.6 & 0.4 \end{bmatrix} \\ &= [0.834\ 0.166] \begin{bmatrix} 0.99 & 0.01 \\ 0.6 & 0.4 \end{bmatrix} \\ &= [0.92526\ 0.07474] \end{aligned}$$

b) Find the steady-state probabilities for this Markov chain. In other words, find the limit of $P(E_i = 1)$ as i becomes very large. Express this probability as a function of $P^{(1)}(0 \to 1)$ and $P^{(1)}(1 \to 1)$ and also find its numerical value for the specific values $P^{(1)}(0 \to 1) = 0.01$ and $P^{(1)}(1 \to 1) = 0.4$.

Define e to be the equilibrium (steady-state) probability of an error — that is, the limit of $P(E_i = 1)$ as i goes to infinity. For this probability to stay the same as i increases, it must satisfy

$$e \;=\; e\,P^{(1)}(1 \to 1) \;+\; (1-e)\,P^{(1)}(0- \to 1)$$

What this says is that if the probability of error at time i has the steady-state value, e, then the probability of error at time $i+1$ must also be e. Solving this for e, we get

$$e \;=\; P^{(1)}(0 \to 1) \,/\, (P^{(1)}(0 \to 1) + 1 - P^{(1)}(1 \to 1))$$

So for $P^{(1)}(0 \to 1) = 0.01$ and $P^{(1)}(1 \to 1) = 0.4$, the steady state probability of error is $1/61$.

Question 3: Suppose that the number of cases of Bubonic Plague in Canada in a year has the Poisson(1.2) distribution. Suppose also that 3/4 of the people in Canada who get Bubonic Plague die, and that the death of one such person is independent of the death of another. Find the distribution of the number of people who die of Bubonic Plauge in Canada in a year. State and prove a theorem that generalizes this result.

A Poisson distribution for the number of Bubonic Plague cases is what we would expect if the large number, n, of people in Canada each have a small probability, p, of getting Bubonic Plague, independently. The parameter of the Poisson distribution would then be $np = 1.2$. If so, each person in Canada would have a probability of dying of Bubonic Plauge of $(3/4)p$, and the number of deaths from Bubonic Plague would have a Poisson distribution with parameter $(3/4)np = (3/4)1.2 = 0.9$.

We can generalize this to a theorem saying that if X has the Poisson(λ) distribution, and $Y|X = x$ has the binomial(x,p) distribution for any non-negative integer x, then Y has the Poisson(λp) distribution. We can prove this as follows:

$$\begin{aligned}
P(Y=y) &= \sum_{x=y}^{\infty} P(X=x)\,P(Y=y\,|\,X=x) \\
&= \sum_{x=y}^{\infty} \frac{\lambda^x}{x!} e^{-\lambda}\, \frac{x!}{y!(x-y)!} p^y (1-p)^{x-y} \\
&= \sum_{z=0}^{\infty} \lambda^{y+z} e^{-\lambda}\, \frac{1}{y!z!} p^y (1-p)^z \\
&= \frac{(\lambda p)^y}{y!}\, e^{-\lambda} \sum_{z=0}^{\infty} \frac{(\lambda(1-p))^z}{z!} \\
&= \frac{(\lambda p)^y}{y!}\, e^{-\lambda}\; e^{\lambda(1-p)} \\
&= \frac{(\lambda p)^y}{y!}\, e^{-\lambda p}
\end{aligned}$$

which is the probability mass function for the Poisson(λp) distribution.