Probability & Statistics for EECS: Homework 9 # Solution

Problem 1

Show the proof of general Bayes' Rule (four cases).

	Y **discrete**	Y **continuous**
X **discrete**	$P(Y=y\|X=x)=\frac{P(X=x\|Y=y)P(Y=y)}{P(X=x)}$	$f_Y(y\|X=x)=\frac{P(X=x\|Y=y)f_Y(y)}{P(X=x)}$
X **continuous**	$P(Y=y\|X=x)=\frac{f_X(x\|Y=y)P(Y=y)}{f_X(x)}$	$f_{Y\|X}(y\|x)=\frac{f_{X\|Y}(x\|y)f_Y(y)}{f_X(x)}$

Solution

1. X discrete, Y discrete:

$$P(Y=y|X=x)P(X=x)=P(X=x,Y=y)=P(X=x|Y=y)P(Y=y)$$

Therefore,

$$P(Y=y|X=x)=\frac{P(X=x|Y=y)P(Y=y)}{P(X=x)}$$

2. X continuous, Y continuous:

$$f_{Y|X}(y|x)f_X(x)=f_{X,Y}(x,y)=f_{X|Y}(x|y)f_Y(y)$$

Therefore,

$$f_{Y|X}(y|x)=\frac{f_{X|Y}(x|y)f_Y(y)}{f_X(x)}$$

3. X discrete, Y continuous:
According to the continuous Bayes' rule, we have

$$P(Y\in(y-\varepsilon,y+\varepsilon)|X=x)=\frac{P(X=x|Y\in(y-\varepsilon,y+\varepsilon)P(Y\in(y-\varepsilon,y+\varepsilon))}{P(X=x)}.$$

By letting $\varepsilon\to 0$, we have

$$\lim_{\varepsilon\to 0}\ P(Y\in(y-\varepsilon,y+\varepsilon)|X=x)=\lim_{\varepsilon\to 0}\ f_Y(y|X=x)\cdot 2\varepsilon,$$

and

$$\lim_{\varepsilon\to 0}\frac{P(X=x|Y\in(y-\varepsilon,y+\varepsilon)P(Y\in(y-\varepsilon,y+\varepsilon))}{P(X=x)}=\lim_{\varepsilon\to 0}\frac{P(X=x|Y=y)f_Y(y)\cdot 2\varepsilon}{P(X=x)}.$$

Therefore, we can finish the proof by canceling the term 2ε in the following equation:

$$\lim_{\varepsilon\to 0}\ f_Y(y|X=x)\cdot 2\varepsilon=\lim_{\varepsilon\to 0}\frac{P(X=x|Y=y)f_Y(y)\cdot 2\varepsilon}{P(X=x)}$$
$$\Rightarrow f_Y(y|X=x)=\frac{P(X=x|Y=y)f_Y(y)}{P(X=x)}.$$

Problem 1 continued on next page...

4. X continuous, Y discrete:

$$\begin{aligned}
P(Y=y|X=x) &= \lim_{\varepsilon\to 0} P(Y=y|X\in(x-\varepsilon,x+\varepsilon)) \\
&= \lim_{\varepsilon\to 0} \frac{P(X\in(x-\varepsilon,x+\varepsilon)|Y=y)P(Y=y)}{P(X\in(x-\varepsilon,x+\varepsilon))} \\
&= \lim_{\varepsilon\to 0} \frac{2\varepsilon\cdot f_X(x|Y=y)P(Y=y)}{2\varepsilon\cdot f_X(x)} \\
&= \frac{f_X(x|Y=y)P(Y=y)}{f_X(x)}
\end{aligned}$$

Problem 2

Let X and Y be i.i.d. Geom(p), and $N = X + Y$.

(a) Find the joint PMF of X, Y, N.

(b) Find the joint PMF of X and N.

(c) Find the conditional PMF of X given $N = n$, and give a simple description in words of what the result says.

Solution

(a) Let $q = 1 - p$. Since $P(N = x + y \mid X = x, Y = y) = 1$, the joint PMF of X, Y, N is

$$P(X = x, Y = y, N = n) = P(X = x, Y = y) = pq^x pq^y = p^2 q^n$$

for x, y, n nonnegative integers with $n = x + y$.

(b) If $X = x$ and $N = n$, then $Y = n - x$. So the joint PMF of X and N is

$$P(X = x, N = n) = P(X = x, Y = n - x, N = n) = p^2 q^n$$

for x, n nonnegative integers with $x \leq n$. As a check, note that this implies

$$P(N = n) = \sum_{x=0}^{n} p^2 q^n = (n + 1) p^2 q^n$$

which agrees with the fact that $N \sim \text{NBin}(2, p)$.

(c) The conditional PMF of X given $N = n$ is

$$P(X = x \mid N = n) = \frac{P(X = x, N = n)}{P(N = n)} = \frac{p^2 q^n}{(n + 1) p^2 q^n} = \frac{1}{n + 1}$$

for $x = 0, 1, \ldots, n$ since, as noted in the solution to (b), $N \sim \text{NBin}(2, p)$. This says that, given that $N = n$, X is equally likely to be any integer between 0 and n (inclusive). To describe this in terms of the story of the Negative Binomial, imagine performing independent Bernoulli trials until the second success is obtained. Let N be the number of failures before the second success. Given that $N = n$, the $(n + 2)nd$ trial is the second success, and the result says that the first success is equally likely to be located anywhere among the first $n + 1$ trials.

Problem 3

Let $X \sim \text{Expo}(\lambda)$, and let c be a positive constant.

(a) If you remember the memoryless property, you already know that the conditional distribution of X given $X > c$ is the same as the distribution of $c + X$ (think of waiting c minutes for a "success" and then having a fresh $\text{Expo}(\lambda)$ additional waiting time). Derive this in another way, by finding the conditional CDF of X given $X > c$ and the conditional PDF of X given $X > c$.

(b) Find the conditional CDF and conditional PDF of X given $X < c$.

Solution

(a) Let F be the CDF of X. The conditional CDF of X given $X > c$ is

$$P(X \leq x \mid X > c) = \frac{P(c < X \leq x)}{P(X > c)} = \frac{F(x) - F(c)}{1 - F(c)} = \frac{e^{-\lambda c} - e^{-\lambda x}}{e^{-\lambda c}} = 1 - e^{-\lambda(x-c)},$$

for $x > c$ (and the conditional CDF is 0 for $x \leq c$). This is the CDF of $c + X$, as desired, since for $x > c$ we have

$$P(c + X \leq x) = P(X \leq x - c) = 1 - e^{-\lambda(x-c)}.$$

The conditional PDF of X given $X > c$ is the derivative of the above expression:

$$f(x \mid X > c) = \lambda e^{-\lambda(x-c)},$$

for $x > c$.
(b) For $x \geq c$, $P(X \leq x \mid X < c) = 1$. For $x < c$:

$$P(X \leq x \mid X < c) = \frac{P(X \leq x \text{ and } X < c)}{P(X < c)} = \frac{P(X \leq x)}{P(X < c)} = \frac{1 - e^{-\lambda x}}{1 - e^{-\lambda c}}.$$

The conditional PDF of X given $X < c$ is the derivative of the above expression:

$$f(x \mid X < c) = \frac{\lambda e^{-\lambda x}}{1 - e^{-\lambda c}}$$

for $x < c$ (and the conditional PDF is 0 for $x \geq c$).

Problem 4

Let U_1, U_2, U_3 be i.i.d. Unif$(0,1)$, and let $L = \min(U_1, U_2, U_3)$, $M = \max(U_1, U_2, U_3)$.

(a) Find the marginal CDF and marginal PDF of M, and the joint CDF and joint PDF of L, M.

(b) Find the conditional PDF of M given L.

Solution

(a) The event $M \leq m$ is the same as the event that all 3 of the U_j are at most m, so the CDF of M is $F_M(m) = m^3$ and the PDF is $f_M(m) = 3m^2,\ for 0 \leq m \leq 1$. The event $L \geq l, M \leq m$ is the same as the event that all 3 of the U_j are between l and m (inclusive), so

$$P(L \geq l, M \leq m) = (m-l)^3$$

for $m \geq l$ with $m, l \in [0,1]$. By the axioms of probability, we have

$$P(M \leq m) = P(L \leq l, M \leq m) + P(L > l, M \leq m)$$

So the joint CDF is

$$P(L \leq l, M \leq m) = m^3 - (m-l)^3$$

for $m \geq l$ with $m, l \in [0,1]$. The joint PDF is obtained by differentiating this with respect to l and then with respect to m (or vice versa):

$$f(l, m) = 6(m-l)$$

for $m \geq l$ with $m, l \in [0,1]$. As a check, note that getting the marginal PDF of M by finding $\int_0^m f(l,m)dl$ does recover the PDF of M (the limits of integration are from 0 to m since the min can't be more than the max).

(b) The marginal PDF of L is $f_L(l) = 3(1-l)^2$ for $0 \leq l \leq 1$ since $P(L > l) = P(U_1 > l, U_2 > l, U_3 > l) = (1-l)^3$ (alternatively, use the PDF of M together with the symmetry that $1 - U_j$ has the same distribution as U_j, or integrate out m in the joint PDF of L, M). So the conditional PDF of M given L is

$$f_{M|L}(m \mid l) = \frac{f(l,m)}{f_L(l)} = \frac{2(m-l)}{(1-l)^2},$$

for all $m, l \in [0,1]$ with $m \geq l$.

Problem 5

Let X and Y be i.i.d. Geom(p), $L = \min(X, Y)$, and $M = \max(X, Y)$.

(a) Find the joint PMF of L and M. Are they independent?

(b) Find the marginal distribution of L in two ways: using the joint PMF, and using a story.

(c) Find $E[M]$.

(d) Find the joint PMF of L and $M - L$. Are they independent?

Solution

(a) Let $q = 1 - p$, and let l and m be nonnegative integers. For $l < m$,

$$P(L = l, M = m) = P(X = l, Y = m) + P(X = m, Y = l) = 2pq^l p^m = 2p^2 q^{l+m}.$$

For $l = m$

$$P(L = l, M = m) = P(X = l, Y = l) = p^2 q^{2l}.$$

For $l > m, P(L = l, M = m) = 0$. The r.v.s L and M are not independent since we know for sure that $L \leq M$ will occur, so learning the value of L can give us information about M. We can also write the joint PMF in one expression as

$$P(L = l, M = m) = 2^{I(l<m)} p^2 q^{l+m} I(l \leq m)$$

where $I(l < m)$ is 1 if $l < m$ and 0 otherwise, and $I(l \leq m)$ is 1 if $l \leq m$ and 0 otherwise. This way of writing it makes it easier to see why the joint PMF does not factor into the product of a function of l and a function of m.
In summary,

$$P(L = l, M = m) = \begin{cases} 2p^2(1-p)^{l+m}, & \text{if } m > l \geq 0; \\ p^2(1-p)^{2l}, & \text{if } m = l \geq 0; \\ 0, & \text{otherwise.} \end{cases}$$

L and M are not independent. For example, $P(L = 1|M = 0) = 0$ while $P(L = 1) > 0$, which means $P(L = 1|M = 0) \neq P(L = 1)$.
(b) We can sum the joint PMF over all possible values of M to get the marginal distribution of L:

$$\begin{aligned} P(L = l) &= \sum_{m=l}^{\infty} P(L = l, M = m) \\ &= p^2 q^{2l} + 2p^2 q^l \sum_{m=l+1}^{\infty} q^m \\ &= p^2 q^{2l} + \left(2p^2 q^l q^{l+1}\right)/(1-q) \\ &= p^2 q^{2l} + 2pq^{2l+1} \\ &= q^{2l}\left(p^2 + 2pq\right) \end{aligned}$$

An easier way to get the same result is to use the story of the Geometric: imagining two independent sequences of independent Bern(p) trials and considering whether at time l at least one of the two trials at that time was a success, we have $L \sim$ Geom$\left(1 - q^2\right)$. This agrees with the above since the PMF of a

Geom $\left(1-q^2\right)$ r.v. is $\left(1-q^2\right)q^{2l}$ for $l=0,1,\ldots$, and $p^2+2pq+q^2=(p+q)^2=1$.

(c) We have $E[L]=q^2/\left(1-q^2\right)$ and

$$E[L]+E[M]=E(L+M)=E(X+Y)=E[X]+E[Y]=2q/p$$

so

$$E[M]=\frac{2q}{p}-\frac{q^2}{1-q^2}=\frac{(1-p)(3-p)}{p(2-p)}.$$

(d) By (a), for $l\geq 0$,

$$P(L=l,M-L=k)=P(L=l,M=k+l)=2^{I(k>0)}p^2q^{2l+k}$$

where $I(k>0)$ is 1 if $k>0$ and 0 otherwise. This factors as

$$P(L=l,M-L=k)=f(l)g(k)$$

for all nonnegative integers l,k, where

$$f(l)=\left(1-q^2\right)q^{2l},g(k)=\frac{2^{I(k>0)}p^2q^k}{1-q^2}.$$

Thus, L and $M-L$ are independent. (Since f is the PMF of L, by summing the joint PMF over the possible values of L we also have that the PMF of $M-L$ is g. The PMF g looks complicated because of the possibility of a "tie" occurring (the event $X=Y$), but conditional on a tie not occurring we have the nice result $M-L-1\mid M-L>0\sim\text{Geom}(p)$, which makes sense due to the memoryless property of the Geometric. To check this conditional distribution, use the definition of conditional probability and the fact that $p^2+q^2=1-2pq$.)