CSE 107
Homework Assignment 2
Solutions

Section 1.4: Total probability and Bayes rule

1. A new test has been developed to determine whether a given student is overstressed. This test is 95% accurate if the student is not overstressed, but only 85% accurate if the student is in fact overstressed. It is known that 99.5% of all students are overstressed. Given that a particular student tests negative for stress, what is the probability that the test results are correct, and that this student is not overstressed?

   Solution:
   Define the events $T = \{\text{the test is positive}\}$ and $S = \{\text{the student is overstressed}\}$. We are given the probabilities $P(T^c|S^c) = .95$, $P(T|S) = .85$ and $P(S) = .995$, from which we deduce that $P(T|S^c) = .05$, $P(T^c|S) = .15$ and $P(S^c) = .005$. We must determine $P(S^c|T^c)$. Using Bayes rule and the total probability theorem, we have

$$P(S^c|T^C) = \frac{P(T^c|S^C)P(S^C)}{P(T^c|S^C)P(S^C) + P(T^c|S)P(S)}$$

$$= \frac{(.95)(.005)}{(.95)(.005) + (.15)(.995)}$$

$$= \mathbf{.0308}.$$

■

Section 1.5: Independence

2. Three persons each independently roll a fair n-sided die once. Let A_{ij} be the event that person i and person j roll the same face. Show that the events A_{12}, A_{13} and A_{23} are pairwise independent but are not independent.

   Proof:
   The sample space Ω for this experiment can be visualized as a $n \times n \times n$ cube, containing n^3 distinct outcomes. Since the die is fair, and the rolls are independent, each outcome is equally, likely with probability $1/n^3$. Each slice of this cube that is parallel to a face represents a fixed outcome for a particular person, and the diagonal in that slice represents the other two persons rolling the same face. We illustrate one such slice below in the case $n = 3$. Here person 1 rolls 2, and the green cells represent persons 2 and 3 rolling the same face.

211	212	213
221	222	223
231	232	233

   Since there n faces parallel to such a slice, we have $P(A_{12}) = P(A_{13}) = P(A_{23}) = n^2/n^3 = 1/n$.

   Observe that if persons 1 and 2 roll the same face, and persons 1 and 3 roll the same face, then all three have rolled the same face, and likewise for the other combinations of persons. Thus

$$A_{12} \cap A_{13} = A_{12} \cap A_{23} = A_{13} \cap A_{23} = A_{12} \cap A_{13} \cap A_{23}$$

This event $\{111, 222, 333, \ldots, nnn\}$ is the main diagonal in the cube itself, and has probability equal to $n/n^3 = 1/n^2$. We can now compute

$$P(A_{12} \cap A_{13}) = \frac{1}{n^2} = \frac{1}{n} \cdot \frac{1}{n} = P(A_{12}) \cdot P(A_{13}),$$

showing that events A_{12} and A_{13} are independent. Similarly events A_{12} and A_{23} are independent, as are events A_{13} and A_{23}. Therefore the set $\{A_{12}, A_{13}, A_{23}\}$ is ***pairwise independent***. However,

$$P(A_{12} \cap A_{13} \cap A_{23}) = \frac{1}{n^2} \neq \frac{1}{n^3} = \frac{1}{n} \cdot \frac{1}{n} \cdot \frac{1}{n} = P(A_{12}) \cdot P(A_{13}) \cdot P(A_{23})$$

so that $\{A_{12}, A_{13}, A_{23}\}$ is ***not independent***. ■

Section 2.2: Probability mass functions

3. The annual premium of a special kind of insurance starts at \$1000 and is reduced by 10% after each year where no claim has been filed. The probability that a claim is filed in a given year is 0.05, independently of preceding years. What is the PMF of the total premium paid up to and including the year when the first claim is filed?

Solution:
The following table shows what happens if a claim is first filed in year k

Year	1	2	3	…	$k-1$	k
Claim?	no	no	no	…	no	yes
Probability	.95	.95	.95	…	.95	.05
Premium	$A(.9)^0$	$A(.9)^1$	$A(.9)^2$	…	$A(.9)^{k-2}$	$A(.9)^{k-1}$

where $A = \$1000$ is the initial year premium. We see that

$$P(\text{first claim in year } k) = (.95)^{k-1} \cdot (.05),$$

and, the total premium paid X, is

$$X = \sum_{i=0}^{k-1} A(.9)^i = A \cdot \frac{1-(.9)^k}{1-(.9)} = 10{,}000 \cdot (1-(.9)^k).$$

Therefore

$$\boldsymbol{p_X(x) = \begin{cases} (.95)^{k-1}(.05) & \textbf{if } x = 10{,}000 \cdot \left(1-(.9)^k\right) \textbf{ for } k = 1,2,3,\ldots \\ 0 & \textbf{otherwise} \end{cases}}$$

■

Section 2.3: Functions of a random variable

4. Let X be a discrete random variable that is uniformly distributed over the set of integers in the range $[a, b]$, where a and b are integers with $a < 0 < b$. Find the PMF of the random variables $\max(0, X)$ and $\min(0, X)$.

Solution:
We are given that X is discrete uniform on $[a, b]$, with $a < 0$ and $b > 0$. Thus

$$p_X(k) = \begin{cases} \dfrac{1}{b-a+1} & \text{if } a \le k \le b \\ 0 & \text{otherwise} \end{cases}$$

If $Y = g(X) = \max(0, X)$, then

$$g(k) = \max(0, k) = \begin{cases} k & \text{if } 1 \le k \le b \\ 0 & \text{if } a \le k \le 0 \end{cases}$$

Thus all the mass that was on the integers $a \le k \le 0$ is mapped to 0, and the other mass stays where it was. Hence

$$\boldsymbol{p_Y(k) = \begin{cases} \dfrac{1}{b-a+1} & \textbf{if } 1 \le k \le b \\ \dfrac{1-a}{b-a+1} & \textbf{if } k = 0 \\ 0 & \textbf{otherwise} \end{cases}}$$

If $Y = g(X) = \min(0, X)$, then

$$g(k) = \min(0, k) = \begin{cases} 0 & \text{if } 0 \le k \le b \\ k & \text{if } a \le k \le -1 \end{cases}$$

In this case, all mass on $0 \le k \le b$ is mapped to 0, and the other mass is stationary. Therefore

$$\boldsymbol{p_Y(k) = \begin{cases} \dfrac{b+1}{b-a+1} & \textbf{if } k = 0 \\ \dfrac{1}{b-a+1} & \textbf{if } a \le k \le -1 \\ 0 & \textbf{otherwise} \end{cases}}$$

■

5. Let X be a discrete random variable, and let $Y = |X|$.

(a) Assume that the PMF of X is

$$p_X(x) = \begin{cases} cx^2 & \text{if } x = -3,-2,-1,0,1,2,3 \\ 0 & \text{otherwise} \end{cases}$$

where c is a suitable constant. Determine the value of c.

Solution:
We have $1 = \sum_{k=-3}^{3} p_X(k) = c(9+4+1+1+4+9) = 28c$, hence $\boldsymbol{c = 1/28}$.

(b) For the PMF of X given in part (a) calculate the PMF of Y.

Solution:

$$p_Y(y) = \sum_{\{x\,:\,y=|x|\}} p_X(x) = \begin{cases} \frac{1}{28}((-3)^2+3^2) & \text{if } y = 3 \\ \frac{1}{28}((-2)^2+2^2) & \text{if } y = 2 \\ \frac{1}{28}((-1)^2+1^2) & \text{if } y = 1 \\ \frac{1}{28}0^2 & \text{if } y = 0 \\ 0 & \text{otherwise} \end{cases} = \begin{cases} \boldsymbol{\frac{9}{14}} & \textbf{if } \boldsymbol{y = 3} \\ \boldsymbol{\frac{2}{7}} & \textbf{if } \boldsymbol{y = 2} \\ \boldsymbol{\frac{1}{14}} & \textbf{if } \boldsymbol{y = 1} \\ \boldsymbol{0} & \textbf{otherwise} \end{cases}$$

(c) Give a general formula for the PMF of Y in terms of the PMF of X.

Solution:

$$p_Y(y) = \sum_{\{x\,:\,y=|x|\}} p_X(x) = \begin{cases} \boldsymbol{p_X(-y) + p_X(y)} & \textbf{if } \boldsymbol{y > 0} \\ \boldsymbol{p_X(0)} & \textbf{if } \boldsymbol{y = 0} \\ \boldsymbol{0} & \textbf{otherwise} \end{cases}$$

Section 2.4: Expectation, mean and variance

6. Let X be a random variable that takes integer values and is symmetric, that is,

$$P(X = k) = P(X = -k)$$

for all integers k. What is the expected value of $Y = cos(X\pi)$ and $Y = sin(X\pi)$?

Solution:
We are given that $p_X(k) = p_X(-k)$, so by the expected value rule,

$$E[Y] = E[g(X)]$$

$$= \sum_{k \in \mathbb{Z}} g(k)p_X(k)$$

$$= g(0)p_X(0) + \sum_{k=1}^{\infty} g(k)p_X(k) + \sum_{k=1}^{\infty} g(-k)p_X(-k)$$

$$= g(0)p_X(0) + \sum_{k=1}^{\infty} g(k)p_X(k) + \sum_{k=1}^{\infty} g(-k)p_X(k)$$

$$= g(0)p_X(0) + \sum_{k=1}^{\infty} [g(k) + g(-k)]p_X(k).$$

If $g(X) = \cos(X\pi)$, then $g(k) = \cos(k\pi) = (-1)^k$, so that

$$E[Y] = (-1)^0 p_X(0) + \sum_{k=1}^{\infty} [(-1)^k + (-1)^{-k}]p_X(k)$$

$$= \boldsymbol{p_X(0) + 2\sum_{k=1}^{\infty} (-1)^k p_X(k).}$$

If $g(X) = \sin(X\pi)$, then $g(k) = \sin(k\pi) = 0$, and hence

$$E[Y] = \mathbf{0}.$$

■

7. A particular binary data transmission and reception device is prone to some error when receiving data. Suppose that each bit is read correctly with probability p. Find the smallest value of p such that when 10,000 bits are received, the expected number of errors is at most 10.

Solution:
Let X be the number of errors. Then X is a binomial random variable with $P(\text{Success}) = 1 - p$ (since a "success" is an error in this case), and $n = 10{,}000$. We know that $E[X] = n(1-p)$. We are to determine the smallest p such that $E[X] \leq 10$. Thus

$$n(1-p) \leq 10,$$

$$1 - p \leq \frac{10}{n},$$

$$1 - \frac{10}{n} \leq p,$$

and hence

$$p \geq \frac{n-10}{n} = \frac{9990}{100000} = .999$$

The smallest such p is $\boldsymbol{p = .999}$. ■

8. Let N be a nonnegative integer-valued random variable. Show that

$$E[N] = \sum_{i=1}^{\infty} P(N \geq i)$$

Proof:
Starting with the right hand side, we have

$$\sum_{i=1}^{\infty} P(N \geq i) = \sum_{i=1}^{\infty} \sum_{k=i}^{\infty} P(N = k)$$

$$= \sum_{i=1}^{\infty} \sum_{k=i}^{\infty} P(N = k)$$

$$= \sum_{i=1}^{\infty} \sum_{k=i}^{\infty} p_N(k)$$

$$= \sum_{k=1}^{\infty} \sum_{i=1}^{k} p_N(k)$$

$$= \sum_{k=1}^{\infty} k \cdot p_N(k)$$

$$= \sum_{k=0}^{\infty} k \cdot p_N(k) = E[N]$$

To understand the switch in the order of summation, observe that lines 3 and 4 sum over all indices (i, k) in the following infinite triangle, but in different ways.

$$\begin{array}{ccccc}
11 & 12 & 13 & 13 & \ldots \\
 & 22 & 23 & 24 & \ldots \\
 & & 33 & 32 & \ldots \\
 & & & 44 & \ldots \\
 & & & & \ddots
\end{array}$$

Line 3 sums over the rows (for each i, index k goes from i to ∞), while line 4 sums over the columns (for each k, index i goes from 1 to k).