Midterm: Solutions

1. (15pt) Let Ω represent the collection of all the books in the Science Library at Wesleyan. We define $\mathscr{F}$ as the entire power set of Ω, and P is a probability measure such that each individual book has an equal probability of being selected. Please provide three examples of random variables defined on the probability space $(\Omega, \mathscr{F}, \mathsf{P})$.

   Solution: Let's say $\Omega = \{b_1, b_2, \ldots, b_{|\Omega|}\}$ be the collection of all $|\Omega|$ books in the Science Library. We can view each element of Ω as an outcome of randomly choosing a book from the library, so that for any i, $\mathsf{P}(\{b_i\}) = \frac{1}{|\Omega|}$.

   Recall that a random variable assigns to each outcome in Ω a *numerical* value. Here are three examples of random variables defined on this probability space:

   (i) $X :=$ number of pages in the book.
   In this case, for example, the probability that a randomly chosen book has at most 300 pages is

   $$\mathsf{P}\{X \leq 300\} = \mathsf{P}(\{b_i \in \Omega : b_i \text{ has at most 300 pages}\}),$$

   where $\{b_i \in \Omega : b_i \text{ has at most 300 pages}\}$ is a subset of Ω and therefore an element of $\mathscr{F}$.

   (ii) $$Y := \begin{cases} 0 & \text{if the book was published before 1500,} \\ 1 & \text{if the book was published between 1501-1600,} \\ 2 & \text{if the book was published between 1601-1700,} \\ 3 & \text{if the book was published between 1701-1800,} \\ 4 & \text{if the book was published between 1801-1900,} \\ 5 & \text{if the book was published between 1901-2000,} \\ 6 & \text{if the book was published after 2001.} \end{cases}$$

   (iii) $Z :=$ weight (in grams) of the book.

   X and Y are discrete random variables, whereas Z is a continuous random variable.

2. (15pt) Within an urn, there are balls numbered from 1 to m. Let's consider the random process of drawing these balls, with replacement, n times. We define a random variable ξ as the largest number observed during these n draws. Please determine the probability distribution of ξ. (Hint: first determine $\mathsf{P}(\{\xi \le k\})$.)

   Solution: The sample space is $\Omega = \underbrace{\{1,\dots,m\} \times \cdots \times \{1,\dots,m\}}_{n \text{ times}} = \{1,\dots,m\}^n$. So $|\Omega| = m^n$.

   So, for example, $(2,2,\dots,2) \in \Omega$ is the outcome that ball #2 was picked in all n drawings, and for this outcome, $\xi = 2$.

   Note that the possible values of ξ are $1,\dots,m$. Let $k \in \{1,\dots,m\}$. Let's first calculate $\mathsf{P}(\{\xi \le k\})$, the probability that the largest number observed during n drawings is at most k. This happens only if in *each* of the n drawings, the ball drawn has number at most k. So the event $\{\xi \le k\} = \underbrace{\{1,\dots,k\} \times \cdots \times \{1,\dots,k\}}_{n \text{ times}}$, which has size k^n.

   Therefore,
$$\mathsf{P}(\{\xi \le k\}) = \frac{|\{\xi \le k\}|}{|\Omega|} = \frac{k^n}{m^n}.$$

   Here's an alternative approach: For $i \in \{1,\dots,n\}$, let η_i be the number of the ball drawn on the i^{th} drawing. On any given drawing, the probability that the selected ball is numbered at most k is $\mathsf{P}(\{\eta_i \le k\}) = \frac{|\{1,\dots,k\}|}{|\{1,\dots,m\}|} = \frac{k}{n}$. Since balls are replaced after each drawing, the outcomes of the n drawings are *independent* of each other. Therefore,

$$\begin{aligned}
\mathsf{P}(\{\xi \le k\}) &= \mathsf{P}\Big(\{\eta_1 \le k\} \cap \cdots \cap \mathsf{P}\{\eta_n \le k\}\Big) \\
&= \mathsf{P}(\{\eta_1 \le k\}) \cdots \mathsf{P}(\{\eta_n \le k\}) && \text{(by independence)} \\
&= \underbrace{\frac{k}{m} \cdots \frac{k}{m}}_{n \text{ times}} = \left(\frac{k}{m}\right)^n.
\end{aligned}$$

   Now, for $k \in \{1,\dots,m\}$,

$$\begin{aligned}
\mathsf{P}(\{\xi = k\}) &= \mathsf{P}(\{\xi \le k\}) - \mathsf{P}(\{\xi \le k-1\}) \\
&= \frac{k^n}{m^n} - \frac{(k-1)^n}{m^n} \\
&= \frac{k^n - (k-1)^n}{m^n}.
\end{aligned}$$

3. (20pt) In the United States, each of the 50 states is represented by 2 federal senators. Now, suppose a committee of 20 senators is randomly formed. Let ξ be the random variable representing the number of senators from Connecticut (CT) in the committee, and let η be a random variable indicating whether Massachusetts (MA) is represented in the committee (where $\eta = 1$ if MA is represented, and $\eta = 0$ otherwise). Please determine the joint distribution of ξ and η.

Solution: Ω contains all the ways 20 senators can be picked from 100 possible federal senators, so $|\Omega| = \binom{100}{20}$. The joint distribution of ξ and η will be computed as follows:

η \ ξ	0	1	2
0	a	c	e
1	b	d	f

(a) We need 0 CT senator and 0 MA senator, so we're counting the number ways to choose 20 senators from the 96 remaining candidates, which gives us $\binom{96}{20}$:

$$\mathsf{P}_{(\xi,\eta)}(0,0) = a = \frac{\binom{96}{20}}{\binom{100}{20}}$$

(b) We need 0 CT senator and at least one MA senator, so we can either have exactly 1 MA senator or exactly 2 MA senators: $\binom{2}{1} \cdot \binom{96}{19} + \binom{2}{2} \cdot \binom{96}{18}$. So:

$$\mathsf{P}_{(\xi,\eta)}(0,1) = b = \frac{\binom{2}{1} \cdot \binom{96}{19} + \binom{2}{2} \cdot \binom{96}{18}}{\binom{100}{20}}$$

(c) We need 1 CT senator and 0 MA senator, so we count the number of ways to choose 1 from 2 CT senators and then 19 from the rest of the 96 remaining candidates, which gives us $\binom{2}{1} \cdot \binom{96}{19}$:

$$\mathsf{P}_{(\xi,\eta)}(1,0) = c = \frac{\binom{2}{1} \cdot \binom{96}{19}}{\binom{100}{20}}$$

(d) We need exactly 1 CT senator. And as in (b) we want at least one MA senator, so we can either have exactly 1 or exactly 2 of them:

$$\mathsf{P}_{(\xi,\eta)}(1,1) = d = \frac{\binom{2}{1} \cdot \binom{2}{1} \cdot \binom{96}{18} + \binom{2}{1} \cdot \binom{2}{2} \cdot \binom{96}{17}}{\binom{100}{20}}$$

(e) We need 2 CT senators and 0 MA senator, so we count the number of ways to choose 2 from 2 CT senators and then 18 from the rest of the 96 remaining candidates:

$$\mathsf{P}_{(\xi,\eta)}(2,0) = e = \frac{\binom{2}{2} \cdot \binom{96}{18}}{\binom{100}{20}}$$

(f) We need exactly 2 CT senator. And as in (b) and (d) we want at least one MA senator, so we can either have exactly 1 or exactly 2 of them:

$$\mathsf{P}_{(\xi,\eta)}(2,1) = f = \frac{\binom{2}{2} \cdot \binom{2}{1} \cdot \binom{96}{17} + \binom{2}{2} \cdot \binom{2}{2} \cdot \binom{96}{16}}{\binom{100}{20}}$$

4. (20pt) Let (ξ, η) be a random vector whose joint distribution is given by:

$\eta \backslash \xi$	-1	0	1
-1	0.1	0	0.2
0	0.1	0.2	0.1
1	0	0.2	0.1

(a) What is the probability for $\xi\eta < 0.5$, given that $\xi + \eta > -0.5$?

(b) What is the probability that the equation $x^2 + \xi x + \eta = 0$ has no real root?

Solution:

(a) We need to find

$$\mathsf{P}\{\xi\eta < 0.5 \mid \xi + \eta > -0.5\} = \frac{\mathsf{P}\Big(\{\xi\eta < 0.5\} \cap \{\xi + \eta > -0.5\}\Big)}{\mathsf{P}\{\xi + \eta > -0.5\}}.$$

- $\{\xi\eta < 0.5\} = \{(-1, 0), (-1, 1), (0, -1), (0, 0), (0, 1), (1, -1), (1, 0)\}$
- $\{\xi + \eta > -0.5\} = \{(-1, 1), (0, 0), (1, -1), (0, 1), (1, 0), (1, 1)\}$
  so $\mathsf{P}\{\xi + \eta > -0.5\} = 0 + 0.2 + 0.2 + 0.2 + 0.1 + 0.1 = 0.8$
- $\{\xi\eta < 0.5\} \cap \{\xi + \eta > -0.5\} = \{(-1, 1), (0, 0), (1, -1), (0, 1), (1, 0)\}$
  so $\mathsf{P}\Big(\{\xi\eta < 0.5\} \cap \{\xi + \eta > -0.5\}\Big) = 0 + 0.2 + 0.2 + 0.1 + 0.2 = 0.7$

Therefore,

$$\mathsf{P}\{\xi\eta < 0.5 \mid \xi + \eta > -0.5\} = \frac{\mathsf{P}\Big(\{\xi\eta < 0.5\} \cap \{\xi + \eta > -0.5\}\Big)}{\mathsf{P}\{\xi + \eta > -0.5\}} = \frac{0.7}{0.8} = 0.875.$$

(b) Recall that the quadratic equation $ax^2 + bx + c = 0$ has no real roots iff $b^2 - 4ac < 0$. So, we want to find the probability that $\xi^2 - 4\eta < 0$. Note that

$$\{\xi^2 - 4\eta < 0\} = \{(-1, 1), (0, 1), (1, 1)\}.$$

Therefore,

$$\mathsf{P}(x^2 + \xi x + \eta = 0 \text{ has no real roots}) = \mathsf{P}\{\xi^2 - 4\eta < 0\} = 0 + 0.2 + 0.1 = 0.3.$$

5. (10pt) Let ξ be a discrete random variable on a probability space $(\Omega, \mathscr{F}, \mathsf{P})$, and suppose that ξ is independent with itself. Show that there is an $x \in \mathbf{R}$ such that $\mathsf{P}\{\xi = x\} = 1$.

   Solution: Let ξ be a discrete random variable that is independent with itself. Let $S \subseteq \mathbf{R}$ be the set of values that ξ takes on.

   Recall that since ξ is independent with itself, for any $x, y \in S$,

$$\mathsf{P}\{\xi = x, \xi = y\} = \mathsf{P}\{\xi = x\} \cdot \mathsf{P}\{\xi = y\}.$$

   Then, for each $x \in S$,

$$\mathsf{P}\{\xi = x\} = \mathsf{P}\{\xi = x, \xi = x\} = \mathsf{P}\{\xi = x\} \cdot \mathsf{P}\{\xi = x\} = \Big(\mathsf{P}\{\xi = x\}\Big)^2.$$

   which implies that either $\mathsf{P}\{\xi = x\} = 0$ or $\mathsf{P}\{\xi = x\} = 1$.

   Since

$$\sum_{x \in S} \mathsf{P}\{\xi = x\} = 1,$$

   we know *all* the $P\{\xi = x\}$ in the summation can't possibly be 0, so there exists $x_0 \in S$ such that $P\{\xi = x_0\} = 1$.