Worksheet 17: April 1 (Solutions)

Principles to Remember

- Remember the following terms relating to discrete probability spaces:
  - **experiment:** A procedure that yields one of a given set of possible outcomes
  - **sample space:** The set S of possible outcomes, which (in discrete probability) is finite or countably infinite
  - **event:** A subset of S
  - **probability distribution:** A function $p : S \to [0,1]$ such that $\sum_{s \in S} p(s) = 1$. We say that $p(s)$ is the **probability** of an outcome $s \in S$. The probability of an event $E \subseteq S$ is defined as $p(E) = \sum_{s \in E} p(s)$.
- **Conditional probability:** If E and F are two events, the *probability of E given F* is defined as $p(E \mid F) = \dfrac{p(E \cap F)}{p(F)}$.
- **Independence:** Two events E and F are *independent* if $p(E \cap F) = p(E)p(F)$.
- **Bayes' Theorem:** If $p(E) \neq 0$ and $p(F) \neq 0$, then

$$p(F \mid E) = \frac{p(E \mid F)p(F)}{p(E \mid F)p(F) + p(E \mid \overline{F})p(\overline{F})}.$$

Exercises

1. Let E and F be events such that $p(E) = 0.2$ and $p(F) = 0.7$. Find the largest and smallest possible values of...
   (a) $p(E \cup F)$
   Solution: Minimum 0.2 (if $E \subset F$), maximum 0.9 (if E and F are disjoint)
   (b) $p(E \cap F)$
   Solution: Minimum 0 (if E and F are disjoint), maximum 0.2 (if $E \subset F$)
   (c) $p(E \mid F)$
   Solution: Minimum 0 (if E and F are disjoint), maximum 2/7 (if $E \subset F$)
   (d) $p(F \mid E)$
   Solution: Minimum 0 (if E and F are disjoint), maximum 1 (if $E \subset F$)
2. If E and F are disjoint and independent, what can you say about their probabilities?
   Solution: At least one of them has probability 0.

3. Show that if E and F are events, then $p(E \cap F) \geq p(E) + p(F) - 1$. This is known as Bonferroni's inequality.
Solution: This is an immediate consequence from the inclusion-exclusion principle. We know that $p(E \cup F) = p(E) + p(F) + p(E \cap F)$, and $p(E \cup F) \leq 1$ by the definition of a probability distribution; simply equate and subtract.

4. Prove Bayes' Theorem.
Solution: By the definition of conditional probability, $p(F \mid E) = \dfrac{p(E \cap F)}{p(F)}$. Expand the denominator to obtain $p(F \mid E) = \dfrac{p(E \cap F)}{p(E \cap F) + p(E \cap \overline{F})}$. Also by the definition of conditional probability, $p(E \cap F) = p(E \mid F)p(F)$ and $p(E \cap \overline{F}) = p(E \mid \overline{F})p(\overline{F})$; substitute these in.

5. Suppose $p(E) = 1/3$, $p(F) = 1/2$, and $p(E \mid F) = 2/5$. Find $p(F \mid E)$.
Solution: From the definition of conditional probability,

$$p(F \mid E) = \frac{p(E \cap F)}{p(E)} = \frac{p(E \mid F)p(F)}{p(E)} = \frac{2/5 \cdot 1/2}{1/3} = \frac{3}{5}.$$

6. I have two boxes. The first box contains three red balls and two blue balls. The second box contains one red ball and five blue balls. If I reach into a box at random and pull out a red ball, what is the probability that I reached into the first box?
Solution: Let F be the event that I reached into the first box, and let R be the event that I pulled out a red ball. Then $p(F) = \frac{1}{2}$, $p(R \mid F) = \frac{3}{5}$, and $p(R \mid \overline{F}) = \frac{1}{6}$. By Bayes' Theorem,

$$p(F \mid R) = \frac{p(R \mid F)p(F)}{p(R \mid F)p(F) + p(R \mid \overline{F})p(\overline{F})} = \frac{3/5 \cdot 1/2}{3/5 \cdot 1/2 + 1/6 \cdot 1/2} = \frac{18}{23}.$$

7. Suppose that a test for opium use has a 2% false positive rate and a 5% false negative rate. That is, 2% of people who do not use opium test positive for opium, and 5% of opium users test negative for opium. Furthermore, suppose that 1% of people actually use opium.

   (a) Find the probability that someone who tests negative for opium use does not use opium.
   Solution: Let T be the event that someone tests positive for opium, and let U be the event that they use opium. Then $p(\overline{T} \mid U) = 0.05$, $p(T \mid \overline{U}) = 0.02$, and $p(U) = 0.01$. By Bayes' Theorem,

$$p(\overline{U} \mid \overline{T}) = \frac{p(\overline{T} \mid \overline{U})p(\overline{U})}{p(\overline{T} \mid \overline{U})p(\overline{U}) + p(\overline{T} \mid U)p(U)} = \frac{0.98 \cdot 0.99}{0.98 \cdot 0.99 + 0.05 \cdot 0.01} = 0.9995.$$

   (b) Find the probability that someone who tests positive for opium use actually uses opium.
   Solution: Using the same definitions as in part (a),

$$p(U \mid T) = \frac{p(T \mid U)p(U)}{p(T \mid U)p(U) + p(T \mid \overline{U})p(\overline{U})} = \frac{0.95 \cdot 0.01}{0.95 \cdot 0.01 + 0.02 \cdot 0.99} = 0.3242.$$