10-751 Assignment 2

Due: Thursday September 26

Problem 1.

1. Show that conditional independence of A and B given C is neither implied nor implied by the independence of A and B.
2. For which events C *is* it the case that for all A and B, the events A and B are independent if and only if they are conditionally independent given C?

Solution. To show that conditional independence does not imply independence, take the 6-sided fair die and consider the events $A = 6$, $B =$ even, $C = \{4 \text{ or } 6\}$. Then $\Pr(AB \mid C) = \Pr(A \mid C) \Pr(B \mid C) = \frac{1}{2} \cdot 1$, but $\Pr(AB) = \frac{1}{6} \neq \frac{1}{6} \cdot \frac{1}{2}$. To show that independence does not imply conditional independence, take a 4-sided fair die and consider the events $A = \{1 \text{ or } 2\}$, $B = \{2 \text{ or } 3\}$, and $C = \{1 \text{ or } 3\}$. Then $\Pr(AB) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = \Pr(A) \Pr(B)$, but $\Pr(AB \mid C) = 0$ while $\Pr(A \mid C) \Pr(B \mid C) = \frac{1}{4}$. In general, the only event C for which there is equivalence is $C = S$.

Problem 2. Starting from a well-shuffled deck of 52 playing cards containing four aces, cards are turned up one-by-one off the top until the first ace appears. On average, how many cards are required to be taken to produce the first ace?

Solution. There are hard and easy ways to solve this one; here's an easy one. On average, the four aces divide up the deck into five pieces of equal length $\frac{48}{5} = 9.6$. So on average, the first ace will be the 10.6-th card.

Problem 3. A random number N of dice is thrown. Let A_i be the event that $N = i$ and assume that $\Pr(A_i) = 2^{-i}$, $i \geq 1$. Let S be the sum of the number of dots that appear on each of the N throws. Find the probability that

1. $N = 2$ given that $S = 4$;
2. $S = 4$ given that N is even;
3. $N = 2$, given that $S = 4$ and the first die showed 1;
4. the largest number shown by any die is r, where S is unknown

Solution.

1.

$$\begin{aligned}
\Pr(N = 2 \mid S = 4) &= \frac{\Pr(S = 4 \mid N = 2) \Pr(N = 2)}{\sum_i \Pr(S = 4 \mid N = i) \Pr(N = i)} \\
&= \frac{\frac{1}{12} \cdot \frac{1}{4}}{\frac{1}{6} \cdot \frac{1}{2} + \frac{1}{12} \cdot \frac{1}{4} + \frac{3}{216} \cdot \frac{1}{8} + \frac{1}{6^4} \cdot \frac{1}{16}}
\end{aligned}$$

2.

$$\begin{aligned}
\Pr(S=4 \mid N=\text{even}) &= \frac{\Pr(S=4 \mid N=2)\frac{1}{4}+\Pr(S=4 \mid N=4)\frac{1}{15}}{\Pr(N=\text{even})} \\
&= \frac{\frac{1}{12}\cdot\frac{1}{4}+\frac{1}{6^4}\cdot\frac{1}{16}}{\frac{1}{4}+\frac{1}{16}+\cdots} \\
&= \frac{4^2 3^3+1}{4^4 3^3}
\end{aligned}$$

3.

$$\begin{aligned}
\Pr(N=2 \mid S=4, \text{first}=1) &= \frac{\Pr(N=2, S=4, \text{first}=1)}{\Pr(S=4, \text{first}=1)} \\
&= \frac{\frac{1}{6}\cdot\frac{1}{6}\cdot\frac{1}{4}}{\frac{1}{6}\cdot\frac{1}{6}\cdot\frac{1}{4}+\frac{1}{6}\cdot\frac{2}{36}\cdot\frac{1}{8}+\frac{1}{6^4}\cdot\frac{1}{16}}
\end{aligned}$$

4. Let $1 \leq M \leq 6$ be the maximum number shown. Then

$$\begin{aligned}
\Pr(M \leq r) &= \sum_{j=1}^{\infty} \Pr(M \leq r \mid N=j)\frac{1}{2^j} \\
&= \sum_{j=1}^{\infty} \left(\frac{r}{6}\right)^j \frac{1}{2^j} \\
&= \frac{r}{12}\left(1-\frac{r}{12}\right)^{-} 1 \\
&= \frac{r}{12-r}
\end{aligned}$$

Thus $\Pr(M=r)=\Pr(M \leq r)-\Pr(M \leq r-1)=\frac{r}{12-r}-\frac{r-1}{12-r-1}$.

Problem 4. A bowl contains twenty cherries, exactly fifteen of which have had their pits removed. A man eats five cherries, picked at random, without noticing whether or not they contain pits. Subsequently, a cherry is picked randomly from the remaining fifteen.

1. What is the probability that this cherry contains a pit?
2. Given that this cherry contains a pit, what is the probability that the man consumed at least one pit?

Solution.

1. $\frac{5}{20}=\frac{1}{4}$.
2. Given that the last chosen cherry contains a pit, the probability that he at least one pit is

$$1-\Pr(\text{didn't eat any} \mid \text{chose from 19 cherries, 4 with pits}) \quad = \quad 1-\frac{15}{19}\cdot\frac{14}{18}\cdot\frac{13}{17}\cdot\frac{12}{16}\cdot\frac{11}{15}$$

Problem 5. Using Matlab, implement Adam Kalai's algorithm to generate factored random numbers (easily): `http://www.mit.edu/~akalai`. Using this implementation, generate a bunch of random factored numbers $1 \leq r \leq n$ (for a sufficiently large n of your choice), and make a histogram showing the empirical probability of each of the small primes (say, $p \leq 30$) dividing a random integer at least once.

Solution. Similar histogram to last time (now restricted to primes). Code omitted.