CPMA 521, Fall 2001
Professor Larget
Assignment #1: due September 10, 2001

0. Memorize the definitions of and understand the concepts of the following terms: *outcome*, *experiment*, *sample space*, *event*, *set union*, *set intersection*, *set union*, *set complement*, *mutually exclusive events*, *probability*, *axioms of probability*, *combined events*, *conditional probability*, *law of total probability*, *weighted average*, *independence of events*, *mutual independence of events*, *random variables*, *discrete random variables*, *probability mass function*, *(cumulative) distribution function*, *expected value*, *moments*, *variance*, *standard deviation*, *Bernoulli trial*, *Bernoulli distributions*, *geometric distribution*, *binomial distribution*, *negative binomial distribution*. You do not need to turn anything for this problem.

1. In the card game Sheepshead, there is a deck of 32 cards. There are fourteen trumps, six hearts, six spades, and six clubs. The 32 cards are distinct — the trumps can be distinguished, etcetera. A hand of six cards is dealt. The order of these five cards does not matter.
   (a) Give an example of a single outcome.
   (b) How many different outcomes are there?
   (c) What is the probability of having exactly three trumps?
   (d) What is the probability of having exactly three clubs?
   (e) What is the probability of having exactly three trumps and exactly three clubs?
   (f) Find the probability mass function for X, the number of trumps in the hand.
   (g) Find the expected value of X.
   (h) Find the standard deviation and variance of X.

Solution:

(a) The first, third, and fifth trumps, the second heart, and the third and sixth clubs is one.

(b) $\binom{32}{6} = 906,192$.

(c) $\mathsf{P}(\text{exactly three trumps}) = \frac{\binom{14}{3}\binom{18}{3}}{\binom{32}{6}} \doteq 0.3278$

(d) $\mathsf{P}(\text{exactly three clubs}) = \frac{\binom{6}{3}\binom{18}{3}}{\binom{32}{6}} \doteq 0.0574$

(e) $\mathsf{P}(\text{exactly three trumps and three clubs}) = \frac{\binom{6}{3}\binom{14}{3}}{\binom{32}{6}} \doteq 0.0080$

(f) $\mathsf{P}(X = x) = \frac{\binom{14}{x}\binom{18}{6-x}}{\binom{32}{6}}$

x	0	1	2	3	4	5	6
$\mathsf{P}(X = x)$	0.0205	0.1324	0.3073	0.3278	0.1690	0.0398	0.0033

(g) You could use the definition directly. However, a simpler approach is to realize that $X = X_1 + \cdots + X_6$ where $X_i = 1$ if the ith card is a trump and equals 0 otherwise for $i = 1, \ldots, 6$.

Then,

$$\begin{aligned}
E[X] &= E[X_1 + \cdots + X_6] \\
&= E[X_1] + E[X_2] + \cdots + E[X_6] \\
&= \mathsf{P}(\text{first is a trump}) + \mathsf{P}(\text{second is a trump}) + \cdots + \mathsf{P}(\text{sixth is a trump}) \\
&= 6 \times \mathsf{P}(\text{first card is a trump}) \\
&= \frac{21}{8} \doteq 2.625
\end{aligned}$$

(h)

$$\begin{aligned}
V[X] &= V\left[\sum X_i\right] \\
&= \sum_{i=1}^{6} V[X_i] + 2 \sum_{1 \le < j \le 6} \mathsf{Cov}(X_i, X_j) \\
&= 6 \times V[X_1] + 30 \times \mathsf{Cov}(X_1, X_2)
\end{aligned}$$

X_1 is a Bernoulli random variable with possible values only 0 and 1, so $X_1 = X_1^2$ and $E[X_1] = \mathsf{P}(X_1 = 1)$. Thus,

$$V[X_1] = E[X_1^2] - (E[X_1])^2 = E[X_1] - (E[X_1])^2 = \mathsf{P}(X_1 = 1)(1 - \mathsf{P}(X_1 = 1)) = \frac{14 \times 18}{32}$$

The covariance also has a simple expression.

$$\mathsf{Cov}(X_1, X_2) = E[X_1 X_2] - E[X_1]E[X_2]$$

The product $X_1 X_2$ is one only when both X_1 and X_2 are one. The product of two indicator random variables is an indicator of the intersection of the two events. In this case, $E[X_1 X_2] = \mathsf{P}(\text{the first two cards are trumps}) = \frac{14}{32} \times \frac{13}{31}$. So

$$\mathsf{Cov}(X_1, X_2) = \frac{14}{32} \times \frac{13}{31} - \left(\frac{14}{32}\right)^2 = -\frac{14 \cdot 18}{31 \cdot 32^2}$$

Putting it all together,

$$\begin{aligned}
V[X] &= 6 \times V[X_1] + 30 \times \mathsf{Cov}(X_1, X_2) \\
&= 6 \times \frac{14 \times 18}{32} - 30 \times \frac{14 \cdot 18}{31 \cdot 32^2} \\
&= \frac{1701}{992} \doteq 1.715.
\end{aligned}$$

The standard deviation is 1.31. This answer makes sense. The mean is between two and three trumps, and it is not too unusual to get a number one or two away from these.

2. Problem 1.6.

Solution: $X \sim \mathsf{Unif}\{1, \ldots, N\}$, $A = \{X \text{is prime}\}$, $B = \{X \geq 11\}$.

(a) When $N = 20$, $A = \{2, 3, 5, 7, 11, 13, 17, 19\}$ and $B = \{11, 12, 13, \ldots, 20\}$.

$$\mathsf{P}(A \cap B) = \mathsf{P}(\{11, 13, 17, 19\}) = \frac{4}{20} = \frac{1}{5}$$

$$\mathsf{P}(A)\mathsf{P}(B) = \frac{8}{20} \times \frac{10}{20} = \frac{1}{5}$$

Therefore, A and B are independent events.

(b) When $N = 21$, A is as before and $B = \{11, 12, 13, \ldots, 21\}$.

$$\mathsf{P}(A \cap B) = \mathsf{P}(\{11, 13, 17, 19\}) = \frac{4}{21}$$

$$\mathsf{P}(A)\mathsf{P}(B) = \frac{8}{21} \times \frac{10}{21} \neq \frac{4}{21}$$

Therefore, A and B are not independent events.

3. Problem 1.7.

Solution: Toss two dice. $A = \{\text{first is a 4}\}$, $B = \{\text{sum is 6}\}$, and $C = \{\text{sum is 7}\}$.

(a)

$$\mathsf{P}(A \cap B) = \mathsf{P}(\text{first is a 4, second is a 2}) = \frac{1}{36}$$

$$\mathsf{P}(A)\mathsf{P}(B) = \frac{1}{6} \times \frac{5}{36} \neq \frac{1}{36}$$

Therefore, A and B are not independent events.

(b)

$$\mathsf{P}(A \cap C) = \mathsf{P}(\text{first is a 4, second is a 3}) = \frac{1}{36}$$

$$\mathsf{P}(A)\mathsf{P}(C) = \frac{1}{6} \times \frac{6}{36} = \frac{1}{36}$$

Therefore, A and C are independent events.

(c) Among all pairs of die rolls that sum to seven, each possible value from 1 to 6 appears the same number of times. Thus, when the sum is seven, this information does not affect the probability that the first die is a four. Similarly, if the first die roll is a four, there is still a 1/6 chance that the next die roll will produce a sum of seven.

4. Problem 1.16. $\mathrm{P}(Y > n) = q^n$ for all $n = 1, 2, 3, \ldots$. Show that Y is geometric.

Solution:

$$\begin{aligned} \mathsf{P}(Y = n) &= \mathsf{P}(Y > n - 1) - \mathsf{P}(Y > n) \\ &= q^{n-1} - q^n = q^{n-1}(1 - q) \\ &= (1 - p)^{n-1} p \qquad \text{where} \quad p = 1 - q, \quad \text{for} \quad n = 1, 2, 3, \ldots. \end{aligned}$$

Therefore, Y has a geometric$(1 - q)$ distribution.

5. Problem 1.17. Show that the minimum of two independent geometric random variables with the same success probability ρ is also a geometric random variable.

Solution: Let X, Y be independent geometric(ρ) random variables and let $Z = \min(X, Y)$.

$$\begin{aligned}
\mathsf{P}(Z > n) &= \mathsf{P}(\min(X,Y) > n) = \mathsf{P}(X > n, Y > n) \\
&= \mathsf{P}(X > n)\mathsf{P}(Y > n) \qquad \text{(by independence)} \\
&= (1-\rho)^n(1-\rho)^n \qquad \text{by the previous problem} \\
&= ((1-\rho)^2)^n
\end{aligned}$$

Therefore, by the previous problem and the algebraic equality between $1-(1-\rho)^2$ and $\rho(2-\rho)$, Z is geometric($\rho(2-\rho)$).

6. Problem 1.34. Find the expected value of a geometric random variable by conditioning on the first coin toss.

Solution: Let $X_1, X_2, \ldots$ be a sequence of independent Bernoulli trials with success probability p. Let $X = \{\min i | X_i = 1, i = 1, 2, 3, \ldots\}$. X is a geometric(p) random variable. Condition on the first coin toss.

$$X = 1 \times 1_{\{X_1=1\}} + (1+Y) \times 1_{\{X_1=0\}}$$

where $Y = \{\min i | X_{1+i} = 1, i = 1, 2, 3, \ldots\}$. In other words, Y is the geometric random variable that counts the number of coin tosses after the first toss until the next head. When we take expectations on both sides, we will use the fact that the expected value of an indicator random variable is the probability of the event it indicates and that Y and X_1 are independent since Y does not depend on the first coin toss. Then

$$E(X) = p + E(1+Y)(1-p) = 1 + (1-p)E(X)$$

since X and Y have the same distribution. Solving this equation gives $E(X) = 1/p$.

7. Problem 1.48. Show the tail sum formula for expectations of non-negative discrete random variables is true.

Solution: It can be helpful to write out double sums in a rectangular array. Switching the order of summation corresponds to summing columns first instead of rows. Let X be a random variable taking values on the non-negative integers.

$$\begin{aligned}
\sum_{n=0}^{\infty} \mathsf{P}(X > n) &= \sum_{n=0}^{\infty}\sum_{k=1}^{\infty} \mathsf{P}(X = k) \\
&= \sum_{k=1}^{\infty}\sum_{n=0}^{k-1} \mathsf{P}(X = k) \qquad \text{(switch order of sums)} \\
&= \sum_{k=1}^{\infty} k\mathsf{P}(X = k) \qquad \text{(adding a constant } k \text{ times)} \\
&= E(X) \qquad \text{(by definition)}
\end{aligned}$$

8. Write a function gnbinom that graphs the negative binomial distribution.

Solution:

```
# Code to graph the negative binomial distribution.

gnbinom <- function(n, p, low=0, high=qnbinom(0.9999,n,p),scale=F)
{
  if(scale) {
low <- qnbinom(0.0001,n,p)
high <- qnbinom(0.9999,n,p)
  }
  values <- low:high
  probs <- dnbinom(values, n, p)
  plot(c(low,high), c(0,max(probs)), type = "n",
    xlab = "Possible Values", ylab = "Probability",
    main = paste("Negative Binomial Distribution \n", "n =",
    n, ", p =", p))
  lines(values, probs, type = "h", col = 2)
  abline(h=0,col=3)
  return(invisible())
}
```