MTHE/STAT 351 — Fall 2016: Solutions to Homework Assignment 6

1. [Let $X$ be a random variable . . .]

Recall that a random variable is a constant with probability 1 if and only if its variance is zero. We will show that since the random variable $X$ given to us satisfies $E[X] = a$ and $E[X(X-1)] = a(a-1)$, we must have $\text{Var}(X) = 0$.

Observe that using linearity of expectation, we can write


Therefore, $E[X(X-1)] = a(a-1)$ can be re-written as $E[X^2] - a = a(a-1)$, from which we get that $E[X^2] = a^2$. Therefore, the variance of $X$ is

$$\text{Var}(X) = E[X^2] - (E[X])^2 = a^2 - (a)^2 = 0.$$ 

Hence, $X$ is a constant, with probability 1, and since $E[X] = a$, that constant must be $a$. In other words, $X = a$ with probability 1.

2. [We have two biased coins]

(a) Let $I_i$, $i = 1, 2$ denote the event that the $i$th coin is picked and let $A_7$ event that the coin we pick lands on heads in exactly 7 of the 10 flips. Since the number of heads in 10 flips of a biased coin is a binomial random variable with parameters 10 and $p$, we know that

$$P(A_7|I_1) = \binom{10}{7}(0.4)^7(0.6)^3, \quad P(A_7|I_2) = \binom{10}{7}(0.7)^7(0.3)^3.$$ 

Thus the law of total probability gives

$$P(A_7) = \sum_{i=1}^{2} P(A_7|I_i)P(I_i) = \frac{1}{2}\binom{10}{7}(0.4)^7(0.6)^3 + \frac{1}{2}\binom{10}{7}(0.7)^7(0.3)^3 \approx 0.155.$$ 

(b) Let $B$ denote the event that the first flip lands heads. Then

$$P(A_7|B) = \frac{P(BA_7)}{P(B)}.$$ 

From the law of total probability

$$P(B) = \sum_{i=1}^{2} P(B|I_i)P(I_i) = \frac{1}{2}(0.4 + 0.7) = 0.55.$$
Similarly
\[ P(BA_7) = \sum_{i=1}^{2} P(BA_7|I_i)P(I_i). \]

Now \( BA_7 \) is the event that the first flip landed heads and exactly 6 of the remaining 9 flips are heads. Thus, by independence,
\[ P(BA_7|I_i) = p_i \left( \begin{array}{c} 9 \\ 6 \end{array} \right) (p_i)^6 (1-p_i)^3 \]
where \( p_1 = 0.4 \) and \( p_2 = 0.7 \), so that
\[ P(BA_7) = \frac{1}{2} \cdot 0.4 \left( \begin{array}{c} 9 \\ 6 \end{array} \right) (0.4)^6 (0.6)^3 + \frac{1}{2} \cdot 0.7 \left( \begin{array}{c} 9 \\ 6 \end{array} \right) (0.7)^6 (0.3)^3 = 0.1082. \]

Finally, we obtain
\[ P(A_7|B) = \frac{P(BA_7)}{P(B)} = \frac{0.1082}{0.55} \approx 0.197. \]

3. [Suppose that 100 numbers . . .]

Let \( X \) be the number of 1's picked out of the 100 numbers selected (with replacement) from the set \( \{-1, 1\} \). Since the numbers are picked independently of one another, \( X \) is a binomial random variable with parameters \( n = 100 \) (total number of trials) and \( p = 1/2 \) (probability of “success” in a particular trial).

Note that \( E[X] = np = 100 \times 1/2 = 50 \), while \( \text{Var}(X) = np(1-p) = 100 \times (1/2) \times (1/2) = 25 \).

(a) Note that the sum of the 100 numbers picked is zero if and only if exactly 50 of the numbers are 1's and the remaining 50 are –1's. So, the probability that the sum of the 100 numbers is zero is nothing but \( P(X = 50) \), which is given by
\[ P(X = 50) = \left( \begin{array}{c} 100 \\ 50 \end{array} \right) (1/2)^{50} (1/2)^{50} = \left( \begin{array}{c} 100 \\ 50 \end{array} \right) / 2^{100}. \]

(b) Let \( S \) be the sum of the 100 numbers picked. Since the number of 1's picked is \( X \), the number of –1's picked is \( 100 - X \). The sum of all the 1's picked is simply \( X \), while the sum of all the –1's picked is \(- (100 - X) = X - 100 \). So, the sum of all the numbers picked is given by
\[ S = X + X - 100 = 2X - 100. \]

Therefore, using linearity of expectation,
\[ E[S] = E[2X - 100] = 2E[X] - 100 = 2 \times 50 - 100 = 0. \]

To get the standard deviation of \( S \), we compute the variance first, using the fact that \( \text{Var}(aX + b) = a^2 \text{Var}(X) \).
\[ \text{Var}(S) = \text{Var}(2X - 100) = 2^2 \text{Var}(X) = 4 \times 25 = 100. \]

Hence, the standard deviation of the sum of the 100 numbers is \( \sqrt{\text{Var}(S)} = 10 \).
4. [Suppose you continually collect coupons...]

(a) First note that the possible values of \( X \) are 2, 3, 4, \ldots. Let \( A_i, i = 1, 2, \ldots \) denote the event that the \( i \)th coupon collected is a type A coupon and let \( B_i \) be the event that the \( i \)th coupon is a type B coupon. Then for all \( n \geq 2 \) the event \( \{ X = n \} \) is the union of two mutually exclusive events:

\[
\{ X = n \} = A_1 A_2 \cdots A_{n-1} B_n \bigcup B_1 B_2 \cdots B_{n-1} A_n
\]

and so

\[
P(X = n) = P(A_1 A_2 \cdots A_{n-1} B_n) + P(B_1 B_2 \cdots B_{n-1} A_n)
\]

\[
= P(A_1) P(A_2) \cdots P(A_{n-1}) P(B_n) + P(B_1) P(B_2) \cdots P(B_{n-1}) P(A_n)
\]

\[
= \left( \frac{1}{3} \right)^{n-1} \cdot \frac{2}{3} + \left( \frac{2}{3} \right)^{n-1} \cdot \frac{1}{3}
\]

where the second equality holds since by the independence assumption on the type of coupons obtained.

Thus the pmf of \( X \) is

\[
P(X = n) = \left( \frac{1}{3} \right)^{n-1} \cdot \frac{2}{3} + \left( \frac{2}{3} \right)^{n-1} \cdot \frac{1}{3}, \ n = 2, 3, 4, \ldots
\]

(b) The expected value of \( X \) is

\[
E(X) = \sum_{n=2}^{\infty} n P(X = n)
\]

\[
= \sum_{n=2}^{\infty} n \left[ \left( \frac{1}{3} \right)^{n-1} \cdot \frac{2}{3} + \left( \frac{2}{3} \right)^{n-1} \cdot \frac{1}{3} \right]
\]

\[
= \sum_{n=2}^{\infty} n \left( \frac{1}{3} \right)^{n-1} \cdot \frac{2}{3} + \sum_{n=2}^{\infty} n \left( \frac{2}{3} \right)^{n-1} \cdot \frac{1}{3}
\]

\[
= \frac{1}{3} \sum_{k=1}^{\infty} (k+1) \left( \frac{1}{3} \right)^{k-1} \cdot \frac{2}{3} + \frac{2}{3} \sum_{k=1}^{\infty} (k+1) \left( \frac{2}{3} \right)^{k-1} \cdot \frac{1}{3}
\]

where in the last equality we used the change of summation variable \( k = n - 1 \). Now notice that \( \left( \frac{1}{3} \right)^{k-1} \cdot \frac{2}{3}, k = 1, 2, \ldots \) is the pmf of a geometric random variable with parameter \( p = \frac{2}{3} \), while \( \left( \frac{2}{3} \right)^{k-1} \cdot \frac{1}{3} \) is the pmf of a geometric random variable with parameter \( p = \frac{1}{3} \). Let these
random variables be $Y$ and $Z$, respectively. Then from the above

$$E(X) = \frac{1}{3} \sum_{k=1}^{\infty} (k+1) \left( \frac{1}{3} \right)^{k-1} \cdot \frac{2}{3} + \frac{2}{3} \sum_{k=1}^{\infty} (k+1) \left( \frac{2}{3} \right)^{k-1} \cdot \frac{1}{3}$$

$$= \frac{1}{3} E(Y + 1) + \frac{2}{3} E(Z + 1)$$

$$= \frac{1}{3} E(Y) + \frac{1}{3} + \frac{2}{3} E(Z) + \frac{2}{3}$$

$$= \frac{1}{3} \cdot \frac{7}{3} + \frac{1}{3} + \frac{2}{3} \cdot \frac{1}{3} + \frac{2}{3}$$

$$= 3.5.$$  

5. [Ghahramani, 5.3, # 14]

Let’s call the event that a senior citizen enters the mall a success. Since 15% of the population are senior citizens, the probability of success is $p = 0.15$. Assuming people enter the mall independently of each other, the number of people that enter the mall until the 10th senior citizen enters is the number of Bernoulli trials until a total of 10 successes accumulate. This number, $Y$, is a negative binomial random variable with parameters $(10, 0.15)$. Note that the random variable $X$ is given in terms of $Y$ as

$$X = Y - 10.$$  

Recall that the pmf of $Y$ is

$$P(Y = n) = \binom{n-1}{9} (0.15)^{10} (0.85)^{n-10}, \quad n = 10, 11, 12, \ldots$$

Thus we obtain

$$P(X = i) = P(Y - 10 = i) = P(Y = i + 10) = \binom{i+9}{9} (0.15)^{10} (0.85)^i, \quad i = 1, 2, 3, \ldots$$

6. [Ghahramani, 5.2, # 12]

Let $N(t)$ be the number of earthquakes in Japan during a time period of $t$ consecutive weeks. We are given that $\{N(t), t \geq 0\}$ is a Poisson process with rate $\lambda = 3$. The probability that the next earthquake occurs after two weeks is equal to the probability that there are no earthquakes within the next two weeks. This is given by

$$P(N(2) = 0) = e^{-(3)(2)} \frac{(3 \cdot 2)^0}{0!} = e^{-6} \approx 0.002479.$$
7. [Ghahramani, 5.2, # 16] This problem is similar to Example 5.16 on page 209 of the text. The solution to that problem uses more technical language than absolutely necessary. The solution below tries to be more intuitive.

Let \( N(t) \) be the number of customers arriving at the bookstore in a time period of duration \( t \) hours. We are given that \( N(t) \) is a Poisson process with rate \( \lambda = 6 \).

The event “exactly one customer arrives between 9:30am and 10:00am, and exactly 10 customers arrive between 9:30am and noon”, can be re-worded as “exactly one customer arrives between 9:30am and 10:00am, and exactly 9 customers arrive between 10:00am and noon”. So, if \( A \) is the event “exactly one customer arrives between 9:30am and 10:00am”, and \( B \) is the event “exactly 9 customers arrive between 10:00am and noon”, we are looking for the probability \( P(AB) \).

Note that the time interval from 9:30 to 10:00 am is disjoint from the time interval from 10:00am to noon, so events occurring within these two time intervals are independent. This is just the independent increments assumption that we used to derive the Poisson process.

Thus, \( A \) and \( B \) are independent events, so \( P(AB) = P(A)P(B) \). Now,

\[
P(A) = P(N(1/2) = 1) = \frac{e^{-(1/2) \lambda}[(1/2) \lambda]^1}{1!} = 3e^{-3}
\]

and

\[
P(B) = P(N(2) = 9) = \frac{e^{-(2) \lambda}[(2) \lambda]^9}{9!} = \frac{12^9 e^{-12}}{9!}.
\]

So, the required probability is

\[
P(AB) = P(A)P(B) = (3e^{-3}) \left( \frac{12^9 e^{-12}}{9!} \right) \approx 0.01305.
\]

Bonus question [Ghahramani, 4.5, # 14] We are given that two discrete random variables \( X \) and \( Y \) taking values in the set \( \{a_1, a_2, \ldots, a_n\} \) have the same \( r \)th moments for \( r = 1, 2, \ldots, n-1 \). That is,

\[
E(X^r) = E(Y^r), \quad r = 1, \ldots, n-1.
\]

To avoid the trivial case, let’s assume that \( n \geq 2 \). Let \( p_i = P(X = a_i) \) and \( q_i = P(Y = a_i) \). Then the equality of the \( r \)th moments is is equivalent to

\[
a_1^r p_1 + a_2^r p_2 + \cdots + a_n^r p_n = a_1^r q_1 + a_2^r q_2 + \cdots + a_n^r q_n, \quad r = 1, \ldots, n-1.
\]

Note that since \( \sum_{i=1}^{n} p_i = \sum_{i=1}^{n} q_i = 1 \), the equation above also holds if \( r = 0 \). Letting \( x_i = p_i - q_i \), we the have the following \( n \) equations

\[
a_1^r x_1 + a_2^r x_2 + \cdots + a_n^r x_n = 0, \quad r = 0, 1, \ldots, n-1.
\]
or, in matrix form,

$$Ax = 0$$  \hfill (*)

where we view $$x = (x_1, \ldots, x_n)$$ as a column vector and

$$A = \begin{bmatrix}
1 & 1 & \ldots & 1 \\
1 & a_2 & \ldots & a_n \\
a_1^2 & a_2^2 & \ldots & a_n^2 \\
\vdots & \vdots & \ddots & \vdots \\
a_1^{n-1} & a_2^{n-1} & \ldots & a_n^{n-1}
\end{bmatrix}$$

We now see that $$A$$ is the transpose of a Vandermonde matrix, and therefore its determinant (by the hint given) is

$$\det(A) = \prod_{1 \leq i < j \leq n} (a_j - a_i).$$

Since the $$a_i$$ are distinct numbers, the determinant is nonzero, which implies that $$A$$ is nonsingular. Therefore the only solution of (*) is $$x = 0$$, the zero vector in $$\mathbb{R}^n$$. Thus the equality of the $$r$$th moments for $$r = 1, \ldots, n-1$$ implies that $$p_i = q_i$$ for all $$i$$, i.e.,

$$P(X = a_i) = P(Y = a_i), \quad i = 1, \ldots, n$$

as claimed.