Solutions to Homework 6

Math 5010-1, Summer 2010

July 7, 2010

p. 202–207, #2. Clearly, Y has a binomial distribution with parameters $n = 3$ and $p = \frac{1}{2}$. That is,

k	P{Y = k}
0	$\binom{3}{0}(\frac{1}{2})^3 = \frac{1}{8}$
1	$\binom{3}{1}(\frac{1}{2})^3 = \frac{3}{8}$
2	$\binom{3}{2}(\frac{1}{2})^3 = \frac{3}{8}$
3	$\binom{3}{3}(\frac{1}{2})^3 = \frac{1}{8}$

Therefore,

$$E(Y^2) = \left(0^2 \times \frac{1}{8}\right) + \left(1^2 \times \frac{3}{8}\right) + \left(2^2 \times \frac{3}{8}\right) + \left(3^2 \times \frac{1}{8}\right) = \frac{24}{8} = 3.$$

Also,

$$E(Y^4) = \left(0^4 \times \frac{1}{8}\right) + \left(1^4 \times \frac{3}{8}\right) + \left(2^4 \times \frac{3}{8}\right) + \left(3^4 \times \frac{1}{8}\right) = \frac{132}{8} = \frac{33}{2}.$$

Therefore,

$$\text{Var}(Y^2) = E(Y^4) - \left\{E(Y^2)\right\}^2 = \frac{33}{2} - 3^2 = \frac{15}{2}.$$

p. 202–207, #3. Know: $EX = EY = EZ = 1$ and $\text{Var}X = \text{Var}Y = \text{Var}Z = 2$.

(a) $E(2X + 3Y) = 2E(X) + 3E(Y) = 5$.

(b) $\text{Var}(2X + 3Y) = \text{Var}(2X) + \text{Var}(3Y)$, by independence. Therefore, $\text{Var}(2X + 3Y) = 4\text{Var}X + 9\text{Var}Y = 26$.

(c) We know that if X_1 and X_2 are independent, then $E(X_1X_2) = E(X_1)E(X_2)$. This and induction together prove that if $X_1, \ldots, X_n$ are independent, then $E(X_1 \cdots X_n) = E(X_1) \cdots E(X_n)$. Therefore, in particular,

$$E(XYZ) = E(X) \cdot E(Y) \cdot E(Z) = 1.$$

(d) We write

$$\text{Var}(XYZ) = E(X^2Y^2Z^2) - \{E(XYZ)\}^2 = E(X^2) \cdot E(Y^2) \cdot E(Z^2) - 1.$$

Now, $2 = \text{Var}(X) = E(X^2) - (EX)^2 = E(X^2) - 1$. Therefore, $E(X^2) = 2 + 1 = 3$. Similarly, $E(Y^2) = E(Z^2) = 3$, and consequently,

$$\text{Var}(XYZ) = 3^3 - 1 = 26.$$

p. 202–207, #10. (a) Clearly,

$$E(X^k) = \left(1^k \times \frac{1}{n}\right) + \cdots + \left(n^k \times \frac{1}{n}\right) = \frac{1^k + \cdots + n^k}{n} = \frac{s(k\,, n)}{n}.$$

And

$$E\left[(X+1)^k\right] = \frac{2^k + \cdots + (n+1)^k}{n} = \frac{s(k\,, n+1) - 1}{n}.$$

(b) By the binomial theorem,

$$\begin{aligned}(X+1)^k &= \binom{k}{0} X^k 1^0 + \binom{k}{1} X^{k-1} 1^1 + \binom{k}{2} X^{k-2} 1^2 + \cdots + \binom{k}{k} X^0 1^k \\ &= X^k + kX^{k-1} + \binom{k}{2} X^{k-1} + \cdots + 1.\end{aligned}$$

Therefore,

$$kX^{k-1} + \binom{k}{2} X^{k-2} + \cdots + 1 = (X+1)^k - X^k.$$

Take expectations to find that

$$E\left[kX^{k-1} + \binom{k}{2} X^{k-2} + \cdots + 1\right] = \frac{s(k\,, n+1) - 1}{n} - \frac{s(k\,, n)}{n}.$$

The right-hand side is $\frac{1}{n}$ times $s(k\,, n+1) - s(k\,, n) - 1 = (n+1)^k - 1$, whence follows the desired result.

(c) $E(X) = (1 + \cdots + n)/n$, and it is easy to see that

$$1 + \cdots + n = \frac{n(n+1)}{2}.$$

For instance, note that $2(1+\cdots+n) = n(n+1)$ because we can write $2(1+\cdots+n)$ as

$$\begin{array}{ll} 1+2 & +\cdots+n \\ +n+(n-1) & +\cdots+1, \end{array}$$

but sum in columns to see that every column is $n+1$, and there are n columns.

Therefore,

$$E(X) = \frac{n+1}{2}.$$

(d) Apply (b) with $k := 3$ to find that

$$E\left[3X^2 + 3X + 1\right] = \frac{(n+1)^3 - 1}{n}.$$

Now, Recall that

$$E(X) = \frac{n+1}{2} \quad \Rightarrow \quad \underbrace{3E(X^2) + \frac{3(n+1)}{2} + 1}_{E[3X^2+3X+1]} = \frac{(n+1)^3 - 1}{n}.$$

Solve to find that

$$\begin{aligned} E(X^2) &= \frac{1}{3}\left[\frac{(n+1)^3 - 1}{n} - \frac{3(n+1)}{2} - 1\right] \\ &= \frac{1}{3}\left[\frac{n^3 + 3n^2 + 3n}{n} - \frac{3n+3}{2} - 1\right] \\ &= \frac{1}{3}\left[n^2 + 3n + 3 - \frac{3}{2}n - \frac{3}{2} - 1\right] \\ &= \frac{1}{3}\left[n^2 + \frac{3}{2}n + \frac{1}{2}\right] = \frac{1}{6}\left[2n^2 + 3n + 1\right] = \frac{(n+1)(2n+1)}{6}. \end{aligned}$$

Because the left-hand side is $\frac{s(2\,,n)}{n} = \frac{1^2+\cdots+n^2}{n}$, we have $s(2\,,n) = n(n+1)(2n+1)/6$. In other words, this gives a probabilistic proof of the following classical identity [due to the Archimedes]:

$$1^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}.$$

(e) $\text{Var}X = E(X^2) - (EX)^2$. Therefore,

$$\begin{aligned}\text{Var}X &= \frac{(n+1)(2n+1)}{6} - \left(\frac{n+1}{2}\right)^2 \\ &= \frac{n+1}{2}\left(\frac{2n+1}{3} - \frac{n+1}{2}\right) \\ &= \frac{n+1}{2}\left(\frac{(4n+2)-(3n+3)}{6}\right) \\ &= \frac{n+1}{2} \times \frac{n-1}{6} = \frac{n^2-1}{12}.\end{aligned}$$

(f) Directly check.

(g) Now we apply (b) with $k = 4$:

$$E\left[4X^3 + 6X^2 + 4X + 1\right] = \frac{(n+1)^4 - 1}{n}. \qquad (1)$$

Because $E(X) = \frac{n+1}{2}$ and $E(X^2) = \frac{(n+1)(2n+1)}{6}$, the left-hand side is

$$\begin{aligned}&4E(X^3) + 6E(X^2) + 4E(X) + 1 \\ &\qquad = 4\frac{s(3\,,n)}{n} + (n+1)(2n+1) + 2(n+1) + 1.\end{aligned}$$

Plug this into (1) to find that

$$4\frac{s(3\,,n)}{n} + \underbrace{(n+1)(2n+1) + 2(n+1) + 1}_{(2n^2+3n+1)+(2n+2)+1=2n^2+5n+4} = \underbrace{\frac{(n+1)^4-1}{n}}_{\frac{n^4+4n^3+6n^2+4n}{n}=n^3+4n^2+6n+4}.$$

Equivalently,

$$4\frac{s(3\,,n)}{n} + 2n^2 + 5n + 4 = n^3 + 4n^2 + 6n + 4,$$

which simplifies to

$$4\frac{s(3\,,n)}{n} = n^3 + 2n^2 + n = n(n^2 + 2n + 1) = n(n+1)^2.$$

Thus,

$$s(3\,,n) = \left[\frac{n(n+1)}{2}\right]^2 = [s(1\,,n)]^2.$$

This is another famous formula [due to Al Karaji]:

$$1^3 + \cdots + n^3 = \left[\frac{n(n+1)}{2}\right]^2.$$

p. 202–207, #13. Let $\mu = 100$ and $\sigma = 10$ respectively denote the mean and the SD.

(a) Select one person at random; call his or her IQ score X. Now $EX = \mu = 100$ and SD$X = \sigma = 10$. Because $130 = 100 + (3 \times 10) = \mu + (3\sigma)$,

$$P\{X > 130\} \le P\left\{|X - \mu| > 3\sigma\right\} \le \frac{\text{Var X}}{(3\sigma)^2} = \frac{1}{9}.$$

But $P\{X > 130\}$ is the total number of scores that exceed 130 divided by the population size. Therefore, the number of scores that exceed 130 is at most (1/9)th of the total population size.

(b) By symmetry,

$$P\{X > 130\} = \frac{1}{2} P\left\{|X - \mu| > 3\sigma\right\} \le \frac{1}{2}\frac{\text{Var X}}{(3\sigma)^2} = \frac{1}{18}.$$

Therefore, the number of scores that exceed 130 is at most (1/18)th of the total population size.

(c) If X is approximately normal, then we can compute [instead of estimate, using Chebyshev inequality],

$$\begin{aligned} P\{X > 130\} &= P\left\{\frac{X - 100}{10} > 3\right\} = 1 - \Phi(3) \\ &\approx 1 - 0.9987 = 0.0013. \end{aligned}$$

Therefore, the number of scores that exceed 130 is approximately 0.13 percent of the total population size.

p. 217–221, #5. Let X_i denote the number of tosses required for the ith person to get his or her first heads. We know that each X_i is geometrically distributed:

$$P\{X_i = k\} = q_i^{k-1} p_i \qquad \text{for } k = 1, 2, \dots\,,$$

where $q_i := 1 - p_i$.

(a) Mary is the second person. Therefore,

$$P\{X_2 > n\} = \sum_{k=n+1}^{\infty} q_2^{k-1} p_2 = p_2 \sum_{k=n+1}^{\infty} q_2^{k-1} = p_2 \cdot \frac{q_2^n}{1 - q_2},$$

thanks to properties of geometric series. Because $p_2 = 1 - q_2$, it follows that $P\{X_2 > n\} = q_2^n$.

(b) Let Y denote the minimum of X_1, X_2, and X_3. We are asked to find $P\{Y > n\}$. But

$$\begin{aligned} P\{Y > n\} &= P\left\{X_1 > n\ ,\ X_2 > n\ ,\ X_3 > n\right\} \\ &= P\{X_1 > n\} \cdot P\{X_2 > n\} \cdot P\{X_3 > n\}, \end{aligned}$$

by independence. Plug in the probabilities [from (a)] to find that

$$P\{Y > n\} = q_1^n q_2^n q_3^n = (q_1 q_2 q_3)^n.$$

(c) Because $P\{Y = n\} + P\{Y > n\} = P\{Y > n-1\}$, we have

$$\begin{aligned} P\{Y = n\} &= P\{Y > n-1\} - P\{Y > n\} \\ &= (q_1q_2q_3)^{n-1} - (q_1q_2q_3)^n \\ &= (q_1q_2q_3)^{n-1}\left[1 - q_1q_2q_3\right]. \end{aligned}$$

In other words, the random variable Y has a geometric distribution with parameter $1 - q_1q_2q_3$!

(d) We want $P\{X_1 > X_2\,, X_3 > X_2\}$. Once again,

$$\begin{aligned} P\{X_1 > X_2\,, X_3 > X_2\} &= \sum_{n=1}^{\infty} P\{X_2 = n\,, X_1 > n\,, X_3 > n\} \\ &= \sum_{n=1}^{\infty} P\{X_2 = n\}P\{X_1 > n\}P\{X_3 > n\} \\ &= \sum_{n=1}^{\infty} q_2^{n-1} p_2 q_1^n q_3^n. \end{aligned}$$

This expression can be simplified as follows:

$$\begin{aligned} P\{X_1 > X_2\,, X_3 > X_2\} &= p_2q_1q_3 \sum_{n=1}^{\infty} (q_2q_1q_3)^{n-1} = p_2q_1q_3 \sum_{k=0}^{\infty} (q_2q_1q_3)^k \\ &= \frac{p_2q_1q_3}{1 - q_1q_2q_3}. \end{aligned}$$

p. 217–221, #10. We will need, for this problem, two identities that were discussed in the lectures. Namely, that if $0 < p < 1$, then:

$$\sum_{k=0}^{\infty} kp^{k-1} = \frac{d}{dp}\sum_{k=0}^{\infty} p^k = \frac{d}{dp}\left(\frac{1}{1-p}\right) = \frac{1}{q^2}; \qquad (2)$$

and because $k^2 = k(k-1) + k$,

$$\begin{aligned}\sum_{k=0}^{\infty} k^2 p^{k-2} &= \sum_{k=0}^{\infty} k(k-1)p^{k-2} + \sum_{k=0}^{\infty} kp^{k-2} \\ &= \frac{d^2}{dp^2}\sum_{k=0}^{\infty} p^k + \frac{1}{p}\sum_{k=0}^{\infty} kp^{k-1} = \frac{2}{q^3} + \frac{1}{pq^2}.\end{aligned} \tag{3}$$

(a) First, note that $P\{X = 2\} = P(S_1 F_2) + P(F_1 S_2) = pq + qp$ $[= 2pq]$. Also, $P\{X = 3\} = P(S_1 S_1 F_2) + P(F_1 F_2 S_3) = p^2 q + q^2 p$. And now we keep going to find that

$$P\{X = k\} = p^{k-1}q + q^{k-1}p \qquad \text{for all } k \geq 2.$$

(b) We follow the definition of expectation:

$$\begin{aligned}E(X) &= \sum_{k=2}^{\infty} k\left(p^{k-1}q + q^{k-1}p\right) \\ &= q\sum_{k=2}^{\infty} kp^{k-1} + p\sum_{k=2}^{\infty} kq^{k-1}.\end{aligned}$$

Eq. (2) above tells us that

$$\sum_{k=0}^{\infty} kp^{k-1} = \frac{1}{q^2} \quad \Rightarrow \quad \sum_{k=2}^{\infty} kp^{k-1} = \frac{1}{q^2} - \sum_{k=0}^{1} kp^{k-1} = \frac{1}{q^2} - 1.$$

Similarly, $\sum_{k=2}^{\infty} kq^{k-1} = p^{-2} - 1$. Because $p + q = 1$, it follows that

$$E(X) = \frac{1}{q} - q + \frac{1}{p} - p = \frac{1}{q} + \frac{1}{p} - 1 = \frac{1}{pq} - 1.$$

(c) We compute

$$\begin{aligned}E(X^2) &= \sum_{k=2}^{\infty} k^2\left(p^{k-1}q + q^{k-1}p\right) \\ &= q\sum_{k=2}^{\infty} k^2 p^{k-1} + p\sum_{k=2}^{\infty} k^2 q^{k-1} \\ &= pq\sum_{k=2}^{\infty} k^2 p^{k-2} + pq\sum_{k=2}^{\infty} k^2 q^{k-2}.\end{aligned}$$

Thanks to eq. (3),

$$\sum_{k=2}^{\infty} k^2 p^{k-2} = \frac{2}{q^3} + \frac{1}{pq^2} - \sum_{k=0}^{1} k^2 p^{k-2} = \frac{2}{q^3} + \frac{1}{pq^2} - \frac{1}{p}.$$

Similarly,

$$\sum_{k=2}^{\infty} k^2 q^{k-2} = \frac{2}{p^3} + \frac{1}{qp^2} - \frac{1}{q}.$$

Therefore,

$$E(X^2) = \frac{2p}{q^2} + \frac{1}{q} - q + \frac{2q}{p^2} + \frac{1}{p} - p.$$

Now $p + q = 1$ and $p^{-1} + q^{-1} = (p+q)/pq = (pq)^{-1}$. Therefore,

$$E(X^2) = \frac{2p}{q^2} + \frac{2q}{p^2} + \frac{1}{pq} - 1 = \frac{2(p^3+q^3)}{p^2q^2} + \left(\frac{1}{pq} - 1\right).$$

This can be simplified even further: By the binomial theorem,

$$1 = (p+q)^3 = p^3 + 3p^2q + 3pq^2 + q^3.$$

Therefore,

$$p^3 + q^3 = 1 - 3p^2q - 3pq^2 = 1 - 3pq(p+q) = 1 - 3pq.$$

Consequently,

$$E(X^2) = \frac{2}{p^2q^2}(1 - 3pq) + \left(\frac{1}{pq} - 1\right).$$

And

$$\operatorname{Var}(X) = \frac{2}{p^2q^2}(1 - 3pq) + \left(\frac{1}{pq} - 1\right) - \left(\frac{1}{pq} - 1\right)^2.$$

Let $\mu := 1/(pq)$ to see that $E(X) = \mu - 1$, and

$\operatorname{Var}(X) = 2\mu^2 - 6\mu + (\mu - 1) - (\mu - 1)^2 = \mu^2 - 3\mu - 2.$

p. 217–221, #12. (a) Let $q_i := 1 - p_i$ to find that

$$\begin{aligned} P\{W_1 = W_2\} &= \sum_{k=1}^{\infty} P\{W_1 = W_2 = k\} = \sum_{k=1}^{\infty} P\{W_1 = k\}P\{W_2 = k\} \\ &= \sum_{k=1}^{\infty} q_1^{k-1} p_1 q_2^{k-1} p_2 = p_1p_2 \sum_{k=1}^{\infty} (q_1q_2)^{k-1} = p_1p_2 \sum_{j=0}^{\infty} (q_1q_2)^j \\ &= \frac{p_1p_2}{1 - q_1q_2}. \end{aligned}$$

(b) Once again,

$$P\{W_1 < W_2\} = \sum_{k=1}^{\infty} P\{W_1 = k < W_2\} = \sum_{k=1}^{\infty} P\{W_1 = k\}P\{W_2 > k\}$$
$$= \sum_{k=1}^{\infty} q_1^{k-1} p_1 q_2^k = p_1 q_2 \sum_{k=1}^{\infty} (q_1 q_2)^{k-1} = \frac{p_1 q_2}{1 - q_1 q_2}.$$

(c) $P\{W_1 > W_2\}$ is the same as $P\{W_2 > W_1\}$ but with the roles of $(p_1\,, q_1)$ and $(p_2\,, q_2)$ reversed. That is,

$$P\{W_1 > W_2\} = \frac{p_2 q_1}{1 - q_1 q_2}.$$

(d) Let $\underline{M} := \min(W_1\,, W_2)$. We saw in #5 that $\underline{M}$ is geometric with parameter $1 - q_1 q_2$. Explicitly said:

$$P\{\underline{M} > n\} = P\{W_1 > n\}P\{W_2 > n\} = q_1^n q_2^n = (q_1 q_2)^n.$$

Therefore,

$$P\{\underline{M} = n\} = P\{\underline{M} > n-1\} - P\{\underline{M} > n\} = (q_1 q_2)^{n-1} - (q_1 q_2)^n.$$

Factor to find that

$$P\{\underline{M} = n\} = (q_1 q_2)^n \left[1 - (q_1 q_2)\right].$$

This is of the form $q^{n-1}p$; therefore, $\underline{M}$ has a geometric distribution with parameter $1 - q_1 q_2$.

(e) Let $\overline{M} := \max(W_1\,, W_2)$. Then,

$$P\left\{\overline{M} < k\right\} = P\{W_1 < k\,, W_2 < k\} = P\{W_1 < k\}P\{W_2 < k\},$$

by independence. Now, $P\{W_1 < k\} = 1 - P\{W_1 > k + 1\} = 1 - q_1^{k+1}$; see #5(a). Similarly, $P\{W_2 < k\} = 1 - q_2^{k+1}$. Therefore, for all $k \geq 1$,

$$P\left\{\overline{M} < k\right\} = \left(1 - q_1^{k+1}\right) \cdot \left(1 - q_2^{k+1}\right).$$

From this we find the distribution of $\overline{M}$ as follows: $P\{\overline{M} = k\} = P\{\overline{M} < k + 1\} - P\{\overline{M} < k\}$ (why?), whence for all $k \geq 1$,

$$P\left\{\overline{M} = k\right\} = \left(1 - q_1^{k+2}\right) \cdot \left(1 - q_2^{k+2}\right) - \left(1 - q_1^{k+1}\right) \cdot \left(1 - q_2^{k+1}\right).$$

p. 233–236, #8. Suppose that pulses arrive independently in time; then the "Random Scatter Theorem" (p. 230) tells us that the total number of pulses in a given half-minute period is distributed according to the Poisson distribution with $\lambda = 5$.

p. 233–236, #10. For parts (a) and (b) it might help to recall that $E(X) = \lambda$ and $\mathrm{Var}(X) = \lambda$.

(a) $E(3X + 5) = 3E(X) + 5 = 3\lambda + 5$.

(b) $\mathrm{Var}(3X + 5) = 9\mathrm{Var}X = 9\lambda$.

(c) This part requires a direct computation:

$$
\begin{aligned}
E\left(\frac{1}{1+X}\right) &= \sum_{k=0}^{\infty}\left(\frac{1}{1+k}\right)\frac{e^{-\lambda}\lambda^k}{k!} = \sum_{k=0}^{\infty}\frac{e^{-\lambda}\lambda^k}{(k+1)!} \\
&= \frac{e^{-\lambda}}{\lambda}\sum_{j=1}^{\infty}\frac{\lambda^j}{j!} \qquad [j := k+1] \\
&= \frac{e^{-\lambda}}{\lambda}\left(\sum_{j=0}^{\infty}\frac{\lambda^j}{j!} - 1\right) = \frac{e^{-\lambda}}{\lambda}\left(e^{\lambda} - 1\right) = \frac{1 - e^{-\lambda}}{\lambda}.
\end{aligned}
$$