**HW4, for MATH441, STAT461, STAT561, due October 2nd**

**1**. Consider rolling a 6-sided die and a 4-sided die. Let $X$ be the value of the 6-sided die minus the value of the 4-sided die.

(a) Determine the probability mass function for $X$.

Solution

$$P(X = i) = \begin{cases} \frac{1}{24} & i = -3 \\ \frac{2}{24} & i = -2 \\ \frac{3}{24} & i = -1 \\ \frac{4}{24} & i = 0 \\ \frac{4}{24} & i = 1 \\ \frac{4}{24} & i = 2 \\ \frac{3}{24} & i = 3 \\ \frac{2}{24} & i = 4 \\ \frac{1}{24} & i = 5 \\ 0 & \text{otherwise} \end{cases}$$

(b) Determine $E[X]$.

Solution

One way is to compute $\sum_{i=-3}^{5} iP(X = i)$. Another way is to let $X_1$ be the value of the 6-sided die and $X_2$ the value of the four sided die. Then $E[X_1] = 3.5$ and $E[X_2] = 2.5$, so $E[X_1 - X_2] = 3.5 - 2.5 = 1.0$, and that is the answer.

(c) Suppose with probability $p$ I pick the 6-sided die and with probability $1 - p$ I pick the 4-sided die. Then I roll whichever die I picked up. Let $Y$ be the resulting value. Find $E[Y]$. Hint: First find the mass function for $Y$.

Solution. Let $A$ be the event that I pick the 6-sided die. Then

$$P(Y = 1) = P(Y = 1|A)P(A) + P(Y = 1|A^c)P(A^c) = (1/6)p + (1/4)(1 - p) = 1/4 - 1/(12p)$$

Also $P(Y = 1) = P(Y = 2) = P(Y = 3) = P(Y = 4)$. However,

$$P(Y = 5) = P(Y = 6) = P(Y = 6|A)P(A) + P(Y = 6|A^c)P(A^c) = p/6 + 0(1 - p) = p/6$$

Thus,

$$E[Y] = \sum_{i=1}^{6} iP(Y = i) = (1 + 2 + 3 + 4)\left[\frac{1}{4} - \frac{p}{12}\right] + \frac{5p}{6} + \frac{6p}{6} = 2.5 - \frac{5p}{6} + \frac{5p}{6} + p = 2.5 + p$$

One way to think about whether this makes sense is to think about extreme values for $p$. If $p = 0$, then you always pick the 4-sided die, and you get the value 2.5. If $p = 1$, you always pick the 6-sided die, and you get 3.5. If $p$ is in between, you get an in between answer. If $p = .5$, you get 3, which is exactly the average of the expected values for the 4- and 6-sided dice.

**2**. Let $X$ be the value of a single fair die with 4 sides.

(a) Find $\sqrt{E[X]}$.

(b) Find $E[\sqrt{X}]$.

(c) Find the standard deviation of $X$, $\sqrt{Var(X)} = \sqrt{E\{(X - E[X])^2\}}$

(d) Find $E(|X - E[X]|) = E\{\sqrt{(X - E[X])^2}\}$.

Solutions. Note that $E[X] = 2.5$. Then

$$\begin{aligned}
\sqrt{E[X]} &= \sqrt{2.5} \approx 1.58 \\
E[\sqrt{X}] &= \sum_{i=1}^{4} \sqrt{i} P(X = i) = \frac{1}{4}\left[\sqrt{1} + \sqrt{2} + \sqrt{3} + \sqrt{4}\right] \approx 1.54 \\
E[X^2] &= \frac{1}{4}\left[1^2 + 2^2 + 3^2 + 4^2\right] = 7.5 \\
Var(X) &= 7.5 - 2.5^2 = 1.25 \\
SD(X) &= \sqrt{Var(X)} = sqrt 1.25 \approx 1.118 \\
E[|X - E[X]|] &= E[|X - 2.5|] = \frac{1}{4}(|1 - 2.5| + |2 - 2.5| + |3 - 2.5| + |4 - 2.5|) = 1
\end{aligned}$$

**3**. Suppose the number of accidents per day at a certain intersection is a Poisson random variable with rate 0.1 on days that it is not raining but with rate 0.2 on days when there is rain.

(a) Find the probability that there is more than one accident on a rainy day.

(b) Find the probability that there is more than one accident on a non-rainy day.

(c) If the probability that there is rain tomorrow is 80%, find the probability that there will be more than one accident.

Solutions. Let $A$ be the event that it is a rainy day.

(a)

$$P(X > 1|A) = 1 - P(X = 0|A) - P(X = 1|A) = 1 - e^{-0.2} - e^{-0.2}(0.2) = 0.0175231$$

(b)

$$P(X > 1|A^c) = 1 - P(X = 0|A^c) - P(X = 1|A^c) = 1 - e^{-0.1} - e^{-0.1}(0.1) = 0.00467884$$

(c) Let $P(A) = .8$. The question asks for $P(X > 1)$. We use total probability.

$$P(X > 1) = P(X > 1|A)P(A) + P(X > 1|A^c)P(A^c) = (1 - e^{-0.2} - e^{-0.2}(0.2))(0.8) + (1 - e^{-0.1} - e^{-0.1}(0.1))(0.2) = 0.01495425$$

**4**. Let

$$P(X = i) = \begin{cases} e^{-3} \cdot 3^i / i! & \text{for } i \in \{0, 1, 2, \dots\} \\ 0 & \text{otherwise} \end{cases}$$

(a) Find $P(X \le 2)$

(b) Show that $\sum_{i=0}^{\infty} P(X = i) = 1$. Hint: Recall the infinite series for $e^x$ or look up and read about the Poisson distribution either in one of the textbooks or online.

(c) (Graduate students). Find $P(X \text{ is an odd number})$.

For part (c), write the probability as an infinite series, and see if you find what infinite series this corresponds to from Calc II or Wikipedia.

Solutions.

(a) $P(X \le 2) = P(X = 0) + P(X = 1) + P(X = 2) = e^{-3}\left(1 + 3 + 3^2/2\right) = 0.4231901$

(b) $\sum_{i=0}^{\infty} e^{-3} 3^i / i! = e^{-3} \sum_{i=0}^{\infty} 3^i / i! = e^{-3} e^3 = e^0 = 1$

(c)

$$P(X \text{ is an odd number}) = e^{-\lambda}\left[\frac{\lambda^1}{1!} + \frac{\lambda^3}{3!} + \cdots\right] = e^{-\lambda}\sinh(\lambda) = e^{-\lambda}\left[\frac{e^{\lambda} - e^{-\lambda}}{2}\right] = \frac{1 - e^{-2\lambda}}{2}$$

For $\lambda = 3$, this means

$$P(X \text{ is an odd number}) = \frac{1 - e^{-6}}{2} = 0.4987606$$

**5**. When I play drums, the number of times I drop a stick is a Poisson random variable with a rate of once every 20 minutes. What is the probability that I drop the stick exactly two times in a 45-minute practice session?

Solution. Since the rate is once per 20 minutes, the rate is .25 times per 5 minutes and $(0.25)(9) = 2.25$ times per 45 minutes. So the number of times I drop a stick is Poisson with $\lambda = 2.25$. Therfore, the probability of dropping a stick exactly twice is

$$e^{-2.25}(2.25)^2/2! \approx 2.27$$

**6**. Suppose I flip a fair coin $2n$ times. Find the probability that I get exactly $n$ heads when
(a) $n = 2$,
(b) $n = 4$,
(c) $n = 8$,
(d) $n = 9$,
(e) $n = 100$,

Now find the probability that, from the same fair coin, I get at least 60% of the flips are heads for the same values of $n$.

Solutions

$$\begin{aligned}
n = 2 : P(X = 1) &= \binom{2}{1}(.5)^2 = 0.5 \\
n = 4 : P(X = 2) &= \binom{4}{2}(.5)^4 = 0.375 \\
n = 8 : P(X = 4) &= \binom{8}{4}(.5)^8 \approx 0.273 \\
n = 9 : P(X = 4.5) &= 0 \\
n = 100 : P(X = 50) &= \binom{100}{50}(.5)^{100} \approx 0.080
\end{aligned}$$

To get at least 60% heads, you need 2 heads when $n = 2$, 3 heads when $n = 3$, 5 heads when $n = 8$, 6 heads when $n = 9$, and 60 heads when $n = 60$. Let $A$ be the event that there are at least 60 heads. Then

$$\begin{aligned}
n = 2 : P(A) &= \binom{2}{2}(.5)^2 = 0.25 \\
n = 4 : P(A) &= \binom{4}{3}(.5)^4 + \binom{4}{4}(.5)^4 = 0.3125 \\
n = 8 : P(A) &= \left[\binom{8}{5} + \binom{8}{6} + \binom{8}{7} + \binom{8}{8}\right](.5)^8 \approx 0.363 \\
n = 9 : P(A) &= \left[\binom{9}{6} + \binom{9}{7} + \binom{9}{8} + \binom{9}{9}\right](.5)^9 \approx 0.253 \\
n = 100 : P(A) &= 0.02844397
\end{aligned}$$

For $n = 100$, I used R where `pbinom()` is the function for the cdf.

```
  1-pbinom(59,100,.5)
[1] 0.02844397
```

**7**. Suppose you roll a (fair, six-sided) die once and record the value.

(a) Now roll the die again until something other than the first value shows up. Let $X_2$ be this waiting time (the number of new rolls until the second distinct value is observed). What is the distribution for $X_2$ (it is one of the named distribution in chapter 4 of Ross), and what is $E[X_2]$?

(b) Now that two distinct values have been rolled, keep rolling again until a third distinct value has been rolled. Let this waiting time be $E[X_3]$. What is the distribution of $X_3$ and what is $E[X_3]$?

(c) (graduate students). Define $X_i$ as the waiting time until the $i$th distinct value is observed after $i-1$ distinct values have been observed. How do you interpret the sum $\sum_{i=1}^{6} X_i$ and what is the value of $E[\sum_{i=1}^{6} X_i]$? Generalize this problem to an $m$-sided fair die.

Solutions.

(a) $X_2$ is geometric with success probability $p = 5/6$ (that is, there is a 5/6 chance that the second roll is distinct from the first). So $E[X/2] = 1/p = 6/5$.

(b) Once two distinct values have been rolled, a third occurs if any of the remaining 4 out of 6 possibilities occurs, so $X_3$ is geometric with $p = 4/6$, and $E[X_3] = 6/4$.

(c) $\sum_{i=1}^{6} X_i$ is the waiting time until all distinct values have occurred, and

$$E\left[\sum_{i=1}^{6} X_i\right] = 1 + 6/5 + 6/4 + 6/3 + 6/2 + 6/1 = 6(1/6 + 1/5 + 1/4 + 1/3 + 1/2 + 1) = 14.7$$

For $m$ dice the expected waiting time until all values occur is

$$m(1 + 1/2 + 1/3 + \cdots 1/m)$$

Note that this is $m$ times the first $m$ terms of the harmonic series. This means that waiting time approaches infinity as die approaches having infinitely many sides (which might not seem surprising), but the growth rate for the waiting is moderately slow since the growth rate for the harmonic series is logarithmic.

For this problem, $X_1$ is a "degenerate" random variable, meaning it is 1 with probability 1, since you always get a "distinct" value on your first toss.