Chris Piech
CS109

Practice Midterm Solutions

1.

a. The answer to this question is simply a multinomial coefficient, which can be written/computed in numerous ways (several of which are shown below).

$$\binom{12}{5,4,3} = \frac{12!}{5!\,4!\,3!} = \binom{12}{5}\binom{7}{4}\binom{3}{3} = \binom{12}{5}\binom{7}{4}$$

b. $\binom{10}{3,4,3} + \binom{10}{5,2,3} + \binom{10}{5,4,1}$

We select (remove) two candies of the same type to give to Larry and Sergey (there is only 1 way to do this for each type of candy). The remaining 10 candies are then distributed to the remaining 10 students. The three terms above correspond respectively to Snickers, Kit Kats, and M&Ms being given to Larry and Sergey.

Note that each of the multinomial coefficients could have been written in different ways (analogously to what was shown in part (a)).

c. $\binom{10}{3,4,3} + \binom{10}{5,2,3} + \binom{10}{5,4,1} + \binom{10}{4,3,3} + \binom{10}{4,4,2} + \binom{10}{5,3,2}$

We select two candies to remain in the bag and the remaining 10 candies are then distributed to the 10 students. The six terms above correspond respectively to the cases where the two candies left in the bag are: (a) 2 Snickers, (b) 2 Kit Kats, (c) 2 M&Ms, (d) 1 Snickers and 1 Kit Kat, (e) 1 Snickers and 1 M&Ms, and (f) 1 Kit Kat and 1 M&Ms.

Note that each of the multinomial coefficients could have been written in different ways (analogously to what was shown in part (a)).

2. There are multiple ways to obtain this answer; here are two:

The first (common) method is to let X = number of slices of pizza eaten immediately after last slice of cheese pizza is eaten. Note that X ~ NegBin(12, 0.5) since there are 12 slices of cheese pizza and slices of the two pizzas are equally likely to get eaten.

Now, we want to consider all cases where $12 \le X \le 21$, since at least 12 slices of pizza must be eaten in order for there to be a chance that the last cheese slice was eaten, and if no more than 21 (out of 24) slices are eaten when the last cheese slice is eaten, then at least 3 slices of pepperoni must remain. Thus, the probability we want is given by the expression:

$$\sum_{i=12}^{21} P(\mathrm{X}=i) = \sum_{i=12}^{21} \binom{i-1}{11}\left(\frac{1}{2}\right)^{12}\left(\frac{1}{2}\right)^{i-12} = \sum_{i=12}^{21} \binom{i-1}{11}\left(\frac{1}{2}\right)^{i}$$

A second method to compute the answer is to use a set of Binomial variables defined as: Y_i = number of cheese slices eaten at time when i total slices have been eaten. We have $\mathrm{Y}_i \sim \mathrm{Bin}(i, 0.5)$, since we have i trials (slices of pizza eaten), where there is a 50% chance that each slice eaten is cheese. Here, we want to compute: $\left(\frac{1}{2}\right)\sum_{i=11}^{20} P(\mathrm{Y}_i = 11)$, since we want to find the probability that when 11 slices of cheese pizza have been eaten (i.e., only one cheese slice remains), a total of 11 to 20 slices of pizza have been eaten. We then multiply by 1/2 to denote the chance that the next slice eaten is in fact the 12th (last) slice of cheese. At that time a total of 12 to 21 slices of pizza will have been eaten, with 12 of those slices having been cheese, which means there are at least 3 slices of pepperoni remaining. Solving yields:

$$\left(\frac{1}{2}\right)\sum_{i=11}^{20} P(\mathrm{Y}_i = 11) = \left(\frac{1}{2}\right)\sum_{i=11}^{20} \binom{i}{11}\left(\frac{1}{2}\right)^{11}\left(\frac{1}{2}\right)^{i-11} = \sum_{i=11}^{20} \binom{i}{11}\left(\frac{1}{2}\right)^{i+1}$$

And just to show the equivalence of this result, if we let $j = i + 1$, we can rewrite the expression immediately above in the same way we computed it in the first method:

$$= \sum_{i=11}^{20} \binom{i}{11}\left(\frac{1}{2}\right)^{i+1} = \sum_{j=12}^{21} \binom{j-1}{11}\left(\frac{1}{2}\right)^{j}$$

3.

a. Let X = the number of times the randomly chosen song is played.

Here the probability p of selecting the particular song = 1/500 and the number of independent trials (song selections) n = 200. So, we have X ~ Bin(200, 1/500). We want to compute:

$$P(\mathrm{X} > 4) = 1 - P(X \le 4) = 1 - \sum_{i=0}^{4} P(\mathrm{X}=i) = 1 - \sum_{i=0}^{4} \binom{200}{i}\left(\frac{1}{500}\right)^{i}\left(\frac{499}{500}\right)^{200-i}$$

b. Let p = probability that a randomly chosen song is played more than 4 times.

As determined in part (a): $p = 1 - \sum_{i=0}^{4} \binom{200}{i} \left(\frac{1}{500}\right)^i \left(\frac{499}{500}\right)^{200-i}$

Now, let Y = the number of songs that have been heard more than 4 times. Here, this problem set-up fits the Poisson paradigm (it is really the same as computing if 3 buckets in a hash table each have more the 4 strings hashed to them). Thus, we have: $Y \sim \text{Poi}(\lambda)$ where $\lambda = 500p$, and p is defined as above.

$$P(Y = 3) = \frac{e^{-\lambda}\lambda^3}{3!} \text{ where } \lambda = 500\left(1 - \sum_{i=0}^{4} \binom{200}{i} \left(\frac{1}{500}\right)^i \left(\frac{499}{500}\right)^{200-i}\right).$$

Note that a Normal approximation is not as appropriate as a Poisson approximation here since p is a very small value.

4.

a. Let X_i = the value rolled on die i, where $1 \le i \le 4$.

$P(X \ge k) = P(X_1 \ge k, X_2 \ge k, X_3 \ge k, X_4 \ge k,) = \left(\frac{6-k+1}{6}\right)^4$, since all four rolls must be greater than or equal to k.

b. There are two common ways to compute E[X]. The first is to directly use the definition of expectation:

$$\begin{aligned} E[X] &= \sum_{x=1}^{6} x \cdot P(X = x) = \sum_{x=1}^{6} x \cdot [P(X \ge x) - P(X \ge x+1)] \\ &= \sum_{x=1}^{6} x \cdot [\left(\frac{6-x+1}{6}\right)^4 - \left(\frac{6-x}{6}\right)^4] \end{aligned}$$

Alternatively, noting that X is non-negative, we can use the property that:

$$E[X] = \sum_{x=1}^{6} P(X \ge x) = \sum_{x=1}^{6} \left(\frac{6-x+1}{6}\right)^4 = \left(\frac{6}{6}\right)^4 + \left(\frac{5}{6}\right)^4 + \left(\frac{4}{6}\right)^4 + \left(\frac{3}{6}\right)^4 + \left(\frac{2}{6}\right)^4 + \left(\frac{1}{6}\right)^4$$

The two expressions to compute E[X] above are, indeed, equivalent.

c. E[S] = E[T – X] = E[T] – E[X]

Let X_i = the value rolled on die i, where $1 \le i \le 4$. As computed in class, we know that $E[X_i] = 3.5$ for all $1 \le i \le 4$.

$E[T] = E[X_1 + X_2 + X_3 + X_4] = E[X_1] + E[X_2] + E[X_3] + E[X_4] = 4(3.5) = 14$

So, E[S] = 14 – E[X], where E[X] is as computed in part (b).

5.

Let X = lifetime of screen in our laptop.
Let event A = manufacturer A produced the screen.
Let event B = manufacturer A produced the screen.

a. We want to compute P(A | X > 18). Using Bayes Theorem, we have:

$$P(A \mid X > 18) = \frac{P(X > 18 \mid A)P(A)}{P(X > 18)} = \frac{(1 - P(X \le 18 \mid A))(0.5)}{P(X > 18)}$$

Noting that (X | A) ~ N(20, 4), we have:

$$P(A \mid X > 18) = \frac{(0.5)(1 - P(\frac{X-20}{2} \le \frac{18-20}{2}))}{P(X > 18)} = \frac{(0.5)\Phi(1)}{P(X > 18)} = \frac{(0.5)(0.8413)}{P(X > 18)}$$

Now, we need to compute P(X > 18):

$$\begin{aligned}P(X > 18) &= P(X > 18 \mid A)P(A) + P(X > 18 \mid B)P(B) \\ &= P(X > 18 \mid A)(0.5) + P(X > 18 \mid B)(0.5) \\ &= 0.5(1 - P(\frac{X-20}{2} \le \frac{18-20}{2}) + 0.5(1 - (1 - e^{-\frac{18}{20}})) \\ &= 0.5(1 - P(Z \le -1)) + 0.5(e^{-\frac{9}{10}}) = 0.5(1 - (1 - P(Z \le 1))) + 0.5(e^{-\frac{9}{10}}) \\ &= 0.5\Phi(1) + 0.5(e^{-\frac{9}{10}}) = 0.5(0.8413) + 0.5(e^{-\frac{9}{10}})\end{aligned}$$

Substituting P(X > 18) into the expression for P(A | X > 18), yields the answer:

$$P(A \mid X > 18) = \frac{(0.5)(0.8413)}{P(X > 18)} = \frac{0.8413}{0.8413 + e^{-\frac{9}{10}}}$$

b. Here, we want to compute P(B | X > 18). Using Bayes Theorem, we have:

$$P(B \mid X > 18) = \frac{P(X > 18 \mid B)P(B)}{P(X > 18)} = \frac{(1 - P(X \le 18 \mid B))(0.5)}{P(X > 18)}$$

Noting that (X | B) ~ Exp(1/20), we have:

$$P(B \mid X > 18) = \frac{0.5(1 - (1 - e^{-\frac{18}{20}}))}{P(X > 18)} = \frac{0.5(e^{-\frac{9}{10}})}{P(X > 18)}$$

Substituting the previously computed value for P(X > 18) into the expression for P(B | X > 18), yields the final answer:

$$P(B \mid X > 18) = \frac{0.5(e^{-\frac{9}{10}})}{P(X > 18)} = \frac{e^{-\frac{9}{10}}}{0.8413 + e^{-\frac{9}{10}}}$$

6.

a. Since X and Y are independent, we have: $f_{X,Y}(x, y) = f_X(x)\, f_Y(y) = \frac{1}{T} \cdot \frac{1}{T} = \frac{1}{T^2}$

b. (Note that this problem is really a variant of the "Joy of Meetings" problem we did in class regarding the time that two people arrive to a meeting.)

Two packets collide when: $| X - Y | \leq \delta$

So we want to compute:

$$\begin{aligned}
P(| X - Y | \leq \delta) &= 1 - P(| X - Y | > \delta) \\
&= 1 - [P(X - Y > \delta) + P(Y - X > \delta)] \\
&= 1 - 2P(X - Y > \delta) \qquad \text{By symmetry} \\
&= 1 - 2P(X > Y + \delta) \\
&= 1 - 2\int_{y=0}^{T-\delta} \int_{x=y+\delta}^{T} f_{X,Y}(x, y)\, dx\, dy = 1 - 2\int_{y=0}^{T-\delta} \int_{x=y+\delta}^{T} \frac{1}{T^2}\, dx\, dy \\
&= 1 - \frac{2}{T^2}\int_{y=0}^{T-\delta} \int_{x=y+\delta}^{T} dx\, dy = 1 - \frac{2}{T^2}\int_{y=0}^{T-\delta} \left[x \Big|_{y+\delta}^{T} \right] dy = 1 - \frac{2}{T^2}\int_{y=0}^{T-\delta} [T-(y+\delta)]\, dy \\
&= 1 - \frac{2}{T^2}\left[Ty - \frac{y^2}{2} - \delta y \Big|_{0}^{T-\delta} \right] = 1 - \frac{2}{T^2}\left[T(T-\delta) - \frac{(T-\delta)^2}{2} - \delta(T-\delta)\right] \\
&= 1 - \left(1 - \frac{2\delta}{T} + \frac{\delta^2}{T^2}\right) = 1 - \left(1 - \frac{\delta}{T}\right)^2 = \frac{2T\delta - \delta^2}{T^2}
\end{aligned}$$

Note that getting the bounds for the integrals right was perhaps the most critical aspect of this problem. To understand the bounds on the integral, sometimes it helps to visualize the region being integrated over. Here we shade the region being integrated over (which is also symmetric with the unshaded upper triangle):