CS 70 Discrete Mathematics and Probability Theory
Summer 2020 Course Notes **Discrete Distributions Sagnik**

1 What is the Distribution?

5pm

5

8

6pm

What distribution would best model each of the following scenarios? Choose from Binomial, Poisson, Geometric. No justification needed.

(a) $\boxed{\text{Poisson}}$ Number of taxis passing the corner of Euclid Ave and Hearst Ave between 5 pm and 6 pm on a weekday.

(b) $\boxed{\text{Geometric}}$ Number of customers who purchase a lottery ticket before someone hits the jackpot.

(c) $\boxed{\text{Binomial}}$ Number of girls in a family with 6 kids.

2 Poisson Quiz

In a semester when we got to meet in person, CS70 teaching staff were trying to infer how many GSIs would need to be present at the final exam to be able to answer questions students might have on the exam. Each GSI recorded the statistics of how often students enrolled in their sections ask questions during the last midterm exams. You can assume the students ask questions independently from each other, and the number of questions students ask follows Poisson distribution.

- Sagnick's students ask N_1 number of questions per every 60 minutes on average.
- Agnibho's students ask N_2 number of questions per every 90 minutes on average.
- Shahzar's students ask N_3 number of questions per every 30 minutes on average.

Note that there are no other students outside of these three sections.

(a) What is the probability that the students ask $\boxed{\text{at least two}}$ questions throughout three hour (180 min) final exam?

$$\lambda_1 = \frac{N_1 \text{ q}}{60 \text{ mins}} \cdot \frac{3 \quad 60 \text{ mins}}{180 \text{ min}} = 3N_1 \qquad \text{q}/180 \text{ mins}$$

$$Y = \underbrace{X_1}_{\text{Pois}(\lambda_1)} + \underbrace{X_2}_{\text{Pois}(\lambda_2)} + \underbrace{X_3}_{\text{Pois}(\lambda_3)}$$

$$\lambda_2 = \frac{N_2 \text{ q}}{90 \text{ min}} \cdot \frac{2 \quad 60 \text{ mins}}{180 \text{ mins}} = 2N_2$$

$$\lambda_3 = \frac{N_3 \text{ q}}{30 \text{ mins}} \cdot 6 \; \frac{60 \text{ mins}}{180 \text{ mins}} = 6N_3$$

Y be # of questions total. $Y \sim \text{Poisson}(\lambda_1 + \lambda_2 + \lambda_3) = \text{Pois}(3N_1 + 2N_2 + 6N_3)$

$$\Pr[Y \geq 2] = 1 - \Pr[Y=1] - \Pr[Y=0] = 1 - \exp(-(3N_1 + 2N_2 + 6N_3))(1 + 3N_1 + 2N_2 + 6N_3)$$

(b) If we let Z be the number of questions asked by students from one of the three sections with the least number of questions, what is the probability that $Z \geq 2$?

$Q_1 \sim \text{Pois}(3N_1)$

$\dot{Q}_1 \quad \dot{Q}_2 \quad \dot{Q}_3$

$1 - \Pr[\min(Q_1, Q_2, Q_3) < 2]$

$Z = \min(Q_1, Q_2, Q_3)$

$$\Pr[Q_2 < 2] \neq \Pr[Q_2 \leq 2]$$
$$= \Pr[Q_2 = 0] + \Pr[Q_2 = 1]$$

$$\Pr[Z \geq 2] = \Pr[\min(Q_1, Q_2, Q_3) \geq 2]$$
$$= \Pr[Q_1 \geq 2 \cap Q_2 \geq 2 \cap Q_3 \geq 2]$$
$$= \Pr[Q_1 \geq 2]\Pr[Q_2 \geq 2]\Pr[Q_3 \geq 2]$$
$$= (1 - \Pr[Q_1 < 2])(1 - \Pr[Q_2 < 2])(1 - \Pr[Q_3 < 2])$$
$$= (1 - e^{-3N_1} - 3N_1 e^{-3N_1})(\qquad)(\qquad)$$

3 A roll of the dice

Consider a single roll of two dice, one red and one blue.

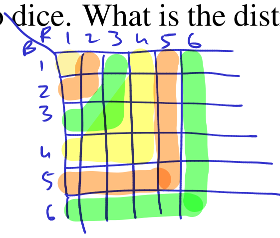
1. Let R be the value of the red die. What is the distribution of R? What is the expectation of R?

what vals?	what prob?
1	1/6
2	1/6
3	1/6
4	1/6
5	1/6
6	1/6

$$E[R] = 1 \cdot \tfrac{1}{6} + 2 \cdot \tfrac{1}{6} + \cdots + 6 \cdot \tfrac{1}{6}$$
$$= \tfrac{1}{6}(1+2+3+4+5+6)$$
$$= 3.5$$

2. Let M be the maximum of the numbers on the two dice. What is the distribution of M? What is the expectation of M?

$(1,2) \; (2,1) \; (2,2)$



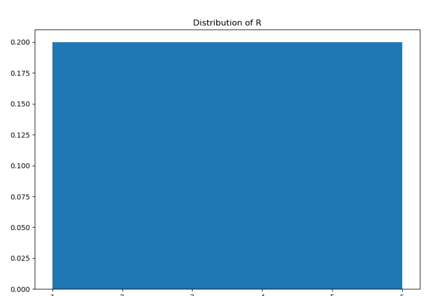
what vals?	what prob?
1	1/36
2	3/36
3	5/36
4	7/36
5	9/36
6	11/36

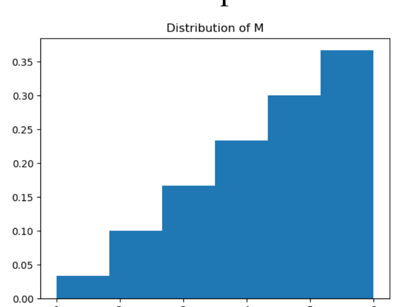
$E[M] \approx 4.47$

$i \in \mathbb{N}$

$\Pr[X = i] = \frac{1}{2^{i+1}}$

3. How do the distribution and expectation of M compare to that of R?





4 Dice Distributions

A fair die with k faces ($k \geq 2$), numbered $1, \ldots, k$, is rolled n times, with each roll being independent of all other rolls. Let X_i be a random variable for the number of times the ith face shows up. For each of the following parts, your answers should be in terms of n and k.

(a) What is the size of the sample space? (How many possible outcomes are there?)

k^n

$\underline{1}\ \underline{5}\ \underline{9}\ \underset{n}{\underline{7}}\ \underline{k}\ \underline{100}\ \underline{k-20}\ \underline{k-1}\ \underline{2}\ \underline{2}$

(b) What is the distribution of X_i, for $1 \leq i \leq k$?

$X_i \sim \text{Binom}(n, \frac{1}{k})$

MISSISSIPPI

(c) What is the joint distribution of $X_1, X_2, \ldots, X_k$?

$k = 3 \quad n = 8$

$\_\ \_\ \_\ \_\ \_\ \_\ \_\ \_$

$x_1 = 3 \quad 111$

$x_2 = 2 \quad 22$

$x_3 = 3 \quad 333$

$\Pr[X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k] = ?$

$x_i \in [0, n],\ x_i \in \mathbb{Z}$

if $x_1 + x_2 + \cdots + x_k \neq n$, $\Pr[X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k] = 0$

else $x_1 + x_2 + \cdots + x_k = n$ $\Pr[X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k] = \left(\frac{n!}{x_1! x_2! \cdots x_k!}\right) / k^n$

(d) Are X_1 and X_2 independent random variables?

No.

$\Pr[X_1 = x_1, X_2 = x_2] \stackrel{?}{=} \Pr[X_1 = x_1]\Pr[X_2 = x_2]$

$\Pr[X_1 = n, X_2 = n] = 0 \neq \Pr[X_1 = n]\Pr[X_2 = n] = \left(\frac{1}{k^n}\right)\left(\frac{1}{k^n}\right)$