1 Discrete Random Variables Part II

1.1 Important Discrete Distributions (Part II)

1.1.1 Geometric Distribution: Geo(p)

The geometric distributions models the number of fails until the **first** success.

Definition 1. Suppose a Bernoulli trial has a probability of success p. The independent trials are repeated until a success has been observed. If X denotes the number of failures that we observed before the first success, then we say X follows the *geometric distribution*, and is denoted by

$$X \sim \text{Geo}(p)$$

- **PMF:**
$$f_X(x) = p(1-p)^x \quad \text{for } x = 0, 1, 2, \dots.$$
- **CDF:**
$$F_X(x) = \begin{cases} 0, & \text{if } x < 0, \\ 1-(1-p)^{\lfloor x \rfloor + 1} & \text{if } \geq 0 \end{cases}$$

Example 1. The following experiments can be modeled by a uniform distribution:

Experiment	X	Distribution
Repeated flips of a coin	# tails until the first head	Geo(0.5)
Repeated flips of a coin	# flips until the first head	Geo(0.5) + 1

Memoryless Property: Intuitively, having a long string of failures should not mean that a success is more likely when conducting independent experiments.

Theorem 1 (*Memoryless Property*)

Let $X \sim \text{Geo}(p)$ and s, t be non-negative integers. Then, the following holds

$$\mathbb{P}(X \geq s+t \mid X \geq s) = \mathbb{P}(X \geq t).$$

In fact, the geometric distribution is the only discrete distribution with this property.

1.1.2 Negative Binomial Distribution: NegBin(k, p)

The negative binomial distributions models the number of fails until a certain amount successes.

Definition 2. Suppose a Bernoulli trial has a probability of success p. The independent trials are repeated until k successes have been observed. If X denotes the number of failures that we observed before k successes, then we say X follows the *negative Binomial distribution*, and is denoted by

$$X \sim \text{NegBin}(k, p)$$

- **PMF:**
$$f_X(x) = \binom{x+k-1}{x} p^k (1-p)^x \quad \text{for } x = 0, 1, 2, \dots.$$
- **CDF:** There is no closed form in terms of elementary functions.

Example 2. The following experiments can be modeled by a negative binomial distribution:

Experiment	X	Distribution
Repeated flips of a coin	# tails until 3 heads	NegBin(3, 0.5)
Repeated flips of a coin	# flips until 3 heads	NegBin(3, 0.5) + 3

Comparison with Binomial Distribution: In the *negative binomial distribution*, you know the number of successes, but you don't know the number of trials (since # fails = # trials - # successes). In the *binomial distribution*, you know the number of trials, but you don't know the number of successes. One can also interpret the coefficient in the PMF of the negative binomial as a negative binomial coefficient (see Problem 1.11).

1.1.3 Poisson Distribution: Poi(λ)

The Poisson distribution models the number of occurrences of an event in a given period of time (or space) when the events happen one after another and the occurrence of one event does not influence another.

Definition 3. Consider counting the number of occurrences of an event that happens at random points in time (or space). Suppose the occurrence of events satisfy the following conditions

1. **Independence:** the number of occurrences in non-overlapping intervals are independent.
2. **Individuality:** for sufficiently short time periods of length Δt, the probability of 2 or more events occurring in the interval is close to zero
$$\frac{\mathbb{P}\left(2 \text{ or more events in } (t, t+\Delta_t)\right)}{\Delta_t} \to 0, \quad \Delta_t \to 0$$
3. **Homogeneity:** events occur at a uniform or homogeneous rate λ and proportional to time interval Δ_t, i.e.
$$\frac{\mathbb{P}\left(\text{one event in } (t, t+\Delta_t)\right) - \lambda\Delta_t}{\Delta_t} \to 0.$$

Suppose that for every $t \geq 0$, the random variable X_t denotes the number of events that have occurred up to time t. The family of random variables X_t is called a *Poisson process* with rate λ.

Definition 4. Let μ encode the rate of a events happening in a fixed length of time according to a *Poisson process*. If X denotes the number of events that occurs in a this time period, then we say X follows the *Poisson distribution*, and is denoted by
$$X \sim \text{Poi}(\mu).$$

- **PMF:**
$$f_X(x) = e^{-\mu}\frac{\mu^x}{x!} \qquad \text{for } x = 0, 1, 2, \dots.$$
- **CDF:** There is no closed form in terms of elementary functions.

Example 3. The following experiments can be modeled by a negative binomial distribution:

Experiment	X	Distribution
Incoming calls at a call center (at rate 2 per hour)	# of calls per hour	Poi(2)
iPhone manufacturing with failure rate 5%	# of faulty iPhones	Poi(0.05)

Poisson Approximation of the Binomial: The Poisson distribution is the limiting case of binomial distribution, where you fix $\mu = np$, and let $n \to \infty$ and $p \to 0$.

Theorem 2 (*Poisson Approximation of the Binomial*)

Given $\lambda > 0$, if $p = p_n \to 0$ in such a way such that $np_n \to \lambda$. Let $X \sim \text{Bin}(n, p_n)$ and $Y \sim \text{Poi}(\lambda)$. Then for all $k \geq 0$,
$$\lim_{n\to\infty} \mathbb{P}(X = k) = \mathbb{P}(Y = k).$$

Marginals of a Poisson Process: Let X_t be a Poisson process with rate λ. The marginals of the process satisfy
$$X_t \sim \text{Poi}(\lambda t).$$

1.2 Example Problems

1.2.1 Computing Probabilities

Problem 1.1. A website is counting visitors to their website. Suppose that visitors visit the website at random at a rate of 10 visitors per minute on average, and they visit the site independently and individually from each other. Let X denote the number of visitors after 10 minutes. What is the distribution of X?

Solution 1.1. Let t be measured in minutes. The rate is 10 visitors per minute, so X_t is a Poisson process with rate $\lambda = 10$. Therefore, number of visitors in 10 minutes is

$$X \sim X_{10} \sim \text{Poi}(10 \cdot 10) = \text{Poi}(100).$$

Problem 1.2. In the manufacturing process of commercial carpet, small faults occur at random in the carpet according to a Poisson process at an average rate of 0.95 per 20 m^2. One of the rooms of a new office block has an area of 90 m^2 and has been carpeted using the same commercial carpet described above. What is the probability that the carpet in that room contains at least 4 faults?

Solution 1.2. Let t be measured in m^2. The rate is 0.95 per 20 m^2 or equivalently a rate of $\frac{0.95}{20}$ faults per m^2, so X_t is a Poisson process with rate $\lambda = \frac{0.95}{20}$. If X denotes the faults in the room, then

$$X \sim X_{90} \sim \text{Poi}\Big(90 \cdot \frac{0.95}{20}\Big) = \text{Poi}(4.275).$$

The complement of the event of at least 4 faults is equal to 3 or less faults, so

$$\begin{aligned}\mathbb{P}(X \geq 4) &= 1 - \mathbb{P}(X \leq 3) = 1 - \mathbb{P}(X = 0) - \mathbb{P}(X = 1) - \mathbb{P}(X = 2) - \mathbb{P}(X = 3) \\ &= 1 - e^{-4.275}\Big(1 + 4.275 + \frac{4.275^2}{2} + \frac{4.275^3}{6}\Big) = 0.618.\end{aligned}$$

Problem 1.3. Website hits for a given website occur according to a Poisson process with a rate of 100 hits per minute. We say a second is a "break" if there are no hits in that second.

1. What is the probability p of a break in any given second?
2. Compute the probability of observing exactly 10 breaks in 60 consecutive non-overlapping seconds.
3. Compute the probability that one must wait for 30 seconds to get 2 breaks.

Solution 1.3. Let t be measured in seconds. The rate is 100 hits per minute, or equivalently $\frac{100}{60}$ hits per second.

1. X_t is a Poisson denotes with rate $\lambda = \frac{100}{60}$. If X is the number of hits in one sec, then

$$X \sim X_1 \sim \text{Poi}(100/60) = \text{Poi}(5/3).$$

A break means zero hits in one sec, so

$$p = \mathbb{P}(X = 0) = e^{-\frac{5}{3}} \frac{\left(\frac{5}{3}\right)^0}{0!} \approx 0.189.$$

2. Take 60 one-sec intervals. Each interval has a probability of p of having a break. Let Y be the number of one-sec intervals (from 60 one-sec intervals) with a break. Then $Y \sim Bin(60, p)$, and

$$\mathbb{P}(Y = 10) = \binom{60}{10} p^{10}(1-p)^{50} \approx 0.124$$

3. Let Z be the number of one-sec intervals one needs to wait until observing two breaks. Then, $Z \sim \text{NegBin}(2, p)$ and

$$\mathbb{P}(Z = 30) = \binom{30+2-1}{30} p^2(1-p)^{30} \approx 0.002.$$

Problem 1.4. At a super busy coffee chain, customers arrive according to a Poisson Process at a rate of $\lambda = 5$ customers per minute.

1. Find the probability that there are more than 2 customers in one minute.
2. Suppose you record the number of customers in 5 consecutive one-minute intervals. What is the probability that in at least 3 of them there were more than 2 customers?
3. Find the probability that a minute with more than 2 customers actually had 6 customers
4. Suppose you are waiting until finally, there is one minute with more than 2 customers. Denote by X the the number of minutes you need to wait. Find the PMF of X.
5. Suppose in 3 minutes, there were n customers. Find the probability that x of these came in the first two minutes.

Solution 1.4. Let t be measured in minutes, and the rate is $\lambda = 5$ customers per minute, so X_t is a Poisson process with rate $\lambda = 5$.

1. If X is the number of customers in one minute, then $X \sim Poi(5 \cdot 1)$. Thus,

$$\begin{aligned} p = \mathbb{P}(X > 2) &= 1 - \mathbb{P}(X \le 2) \\ &= 1 - \mathbb{P}(X=0) - \mathbb{P}(X=1) - \mathbb{P}(X=2) \\ &= 1 - e^{-5}\Big(1 + 5 + \frac{5^2}{2}\Big) \approx 0.875 \end{aligned}$$

2. Let Y be the number of one-minute intervals with more than two customers, then $Y \sim \text{Bin}(5, p)$ with p from earlier. Thus,

$$\begin{aligned} \mathbb{P}(Y \ge 3) &= \mathbb{P}(Y=3) + \mathbb{P}(Y=4) + \mathbb{P}(Y=5) \\ &= \binom{5}{3} p^3(1-p)^2 + \binom{5}{4} p^4(1-p) + \binom{5}{5} p^5(1-p)^0 \approx 0.984 \end{aligned}$$

3. Let X be the number of customers in one minute. Thus,

$$\mathbb{P}(X = 6 \mid X > 2) = \frac{\mathbb{P}(X = 6 \text{ and } X > 2)}{\mathbb{P}(X > 2)} = \frac{\mathbb{P}(X=6)}{\mathbb{P}(X>2)} = \frac{e^{-5}\frac{5^6}{6!}}{0.875} \approx 0.167$$

4. Let Z be the number of minutes until first minute with more than 2 customers, then $Z \sim Geo(p)$ with p from earlier. Thus,

$$f_Z(x) = \mathbb{P}(Z = x) = (1-p)^x p, \quad x = 0, 1, 2, \ldots$$

5. We want to fine

$$\mathbb{P}(x \text{ in first 2min} \mid n \text{ in 3min}) = \frac{\mathbb{P}(x \text{ in first 2min and n in 3min})}{\mathbb{P}(n \text{ in 3min})}.$$

We know that

- Denominator: The number of customers in 3 minutes follows a Poi$(5 \cdot 3)$ distribution, so

$$\mathbb{P}(n \text{ in 3 min}) = e^{-15}\frac{15^n}{n!}, \quad n = 0, 1, 2, \ldots$$

- Numerator: Since non-overlapping intervals are independent,

$$\begin{aligned}\mathbb{P}(x \text{ in first 2min and } n \text{ in 3min}) &= \mathbb{P}(x \text{ in first 2min and } n - x \text{ in last min}) \\ &= \mathbb{P}(x \text{ in first 2min}) \cdot \mathbb{P}(n - x \text{ in last min}) \\ &= e^{-10}\frac{10^x}{x!} \cdot e^{-5}\frac{5^{n-x}}{(n-x)!}\end{aligned}$$

Combining and simplifying gives

$$\mathbb{P}(x \text{ in first 2min} \mid n \text{ in 3min}) = \binom{n}{x}\left(\frac{2}{3}\right)^x\left(\frac{1}{3}\right)^{n-x}, \quad x = 0, 1, \ldots, n,$$

which is the PMF of Bin$(n, 2/3)$.

Problem 1.5. Shiny versions of Pokemon are possible to encounter and catch starting in Generation 2 (Pokemon Gold/Silver). Normal encounters with Pokemon while running in grass occur according to a Poisson process with rate 1 per minute on average. 1 in every 8192 encounters will be a Shiny Pokemon, on average.

1. Ash runs around in grass for 15 hours, what is the probability he will encounter at least one Shiny pokemon?
2. How long would Ash have to run around in grass so that he has better than 50 percent chance of encountering at least one Shiny pokemon?

Solution 1.5. Let t be measured in hours. The rate of normal encounters is 60 per hour, and the rate of shinies are is $\frac{1}{8192 \text{ encounters}}\frac{60 \text{ enounters}}{\text{hour}} = \frac{60}{8192}$ shinies per hour. Let X be number of pokemon encountered after 1 hour, Y be the number of shiny pokemon encountered after one hour. Then

$$X \sim X_1 \sim \text{Poi}(60) \qquad \text{and} \qquad Y \sim Y_1 \sim \text{Poi}\Big(\frac{60}{8192}\Big)$$

1. Let Z be the number of shiny encountered after 15 hours, then $Z \sim Y_{15} = \text{Poi}(\frac{60}{8192} \cdot 15) = \text{Poi}(0.1099)$, and
$$\mathbb{P}(Z \geq 1) = 1 - \mathbb{P}(Z = 0) = 1 - e^{-0.1099} \approx 0.104.$$
2. If Z is the number of shiny encountered after t hours, then $Z \sim \text{Poi}(\frac{60}{8192} \cdot t)$. Then
$$\begin{aligned}\mathbb{P}(Z \geq 1) = 1 - \mathbb{P}(Z = 0) \geq 0.5 &\Leftrightarrow \mathbb{P}(Z = 0) \leq 0.5 \\ &\Leftrightarrow e^{-\frac{60}{8192}\cdot t} \leq 0.5 \\ &\Leftrightarrow -\frac{60}{8192} \cdot t \leq -\log(2) \\ &\Leftrightarrow t \geq \log(2) \cdot \frac{8192}{60} \approx 94.6\end{aligned}$$
That mean Ash will have to run for at least 95 hours!

Problem 1.6. A bit error occurs for a given data transmission method independently in one out of every 1000 bits transferred. Suppose a 64 bit message is sent using the transmission system. Let p_{true} be the probability that there are exactly 2 bit errors and p_{approx} be the approximated probability that there are exactly 2 bit errors obtained through the Poisson approximation. Find p_{true} and p_{approx}.

Solution 1.6.

1. Let X be the number of errors, then $X \sim \mathrm{Bin}(64, 1/1000)$. We find

$$\mathbb{P}(X = 2) = \binom{64}{2}\left(\frac{1}{1000}\right)^2\left(\frac{999}{1000}\right)^{64-2} \approx 0.00189.$$

2. n is large and p is small, so X follows approximately a Poisson with $\mu = np = 64/1000$, so

$$\mathbb{P}(X = 2) = e^{-\frac{64}{1000}}\frac{\left(\frac{64}{1000}\right)^2}{2!} \approx 0.00192.$$

As expected by the Poisson approximation of the Binomial, both values are close.

1.2.2 Derivation of Distributions and Properties

Problem 1.7. If $p \in (0, 1)$, show that

$$f_X(x) = p(1-p)^x \quad \text{for } x = 0, 1, 2, \dots.$$

is a valid PMF.

Solution 1.7. Clearly, $f_X(x) \geq 0$ on its support. We need to check if the sum is over the support 1. Using the formula for the sum of a geometric series i.e. if $|q| < 1$, then $\sum_{k=0}^{\infty} q^k = \frac{1}{1-q}$, we see that

$$\sum_{x\geq 0} p(1-p)^x = \frac{p}{1-(1-p)} = 1$$

as required.

Problem 1.8. Find the CDF of a $\mathrm{Geo}(p)$ random variable.

Solution 1.8. For $x = 0, 1, 2, \dots$, we have by the sum of a geometric series i.e. if $|q| < 1$, then $\sum_{k=0}^{n-1} q^k = \frac{1-q^n}{1-q}$ that

$$F_X(x) = \sum_{k\leq x} p(1-p)^k = p\frac{1-(1-p)^{x+1}}{1-(1-p)} = 1-(1-p)^{x+1}.$$

Clearly, $F_X(x) = 0$ for $x < 0$ and by linearly interpolating between the discontinuities, we see that $F_X(x) = 1-(1-p)^{\lfloor x \rfloor+1}$ for $x \geq 0$.

Problem 1.9. Let $X \sim \mathrm{Geo}(p)$ and s, t be non-negative integers. Show that

$$\mathbb{P}(X \geq s+t \,|\, X \geq s) = \mathbb{P}(X \geq t).$$

Solution 1.9. Notice that for any integer r,

$$\mathbb{P}(X \geq r) = 1 - \mathbb{P}(X \leq r - 1) = 1 - F_X(r - 1) = (1 - p)^r.$$

Therefore,

$$\mathbb{P}(X \geq s + t \,|\, X \geq s) = \frac{\mathbb{P}(X \geq s + t, X \geq s)}{\mathbb{P}(X \geq s)} = \frac{\mathbb{P}(X \geq s + t)}{\mathbb{P}(X \geq s)} = \frac{(1 - p)^{s+t}}{(1 - p)^s} = \mathbb{P}(X \geq t).$$

Problem 1.10. $(*)$ Suppose that X is discrete and for all s, t be non-negative integers

$$\mathbb{P}(X \geq s + t \,|\, X \geq s) = \mathbb{P}(X \geq t).$$

Show that X must have a Geometric distribution.

Solution 1.10. By assumption, for every $n \geq 1$

$$\mathbb{P}(X \geq n + 1 \,|\, X \geq 1) = \frac{\mathbb{P}(X \geq n + 1)}{\mathbb{P}(X \geq 1)} = \mathbb{P}(X \geq n).$$

Rearranging this implies that

$$\mathbb{P}(X \geq n + 1) = \mathbb{P}(X \geq 1)\,\mathbb{P}(X \geq n) = (1 - \mathbb{P}(X = 0))\,\mathbb{P}(X \geq n) = (1 - p)\,\mathbb{P}(X \geq n)$$

where we set $p = \mathbb{P}(X = 0)$ (which is consistent with the meaning in the Geometric distribution). Since this holds for all n, we can continue inductively to see that

$$\mathbb{P}(X \geq n + 1) = (1 - p)^{n+1}.$$

Therefore,

$$\begin{aligned}\mathbb{P}(X = n) = \mathbb{P}(X \geq n) - \mathbb{P}(X \geq n + 1) &= (1 - p)^n - (1 - p)^{n+1} \\ &= (1 - (1 - p))(1 - p)^n = p(1 - p)^n.\end{aligned}$$

Problem 1.11. $(*)$ Show that

$$\binom{x + k - 1}{x} = (-1)^x \binom{-k}{x}.$$

Solution 1.11. By definition,

$$\binom{x + k - 1}{x} = \frac{(x + k - 1)(x + k - 2)\cdots(k + 1)k}{x!}.$$

We can factor out (-1) from each term in the numerator (there are a total of x of them) to conclude that

$$\begin{aligned}\frac{(x + k - 1)(x + k - 2)\cdots(k + 1)k}{x!} &= (-1)^x \frac{(-k)(-k - 1)\cdots(-k - x + 2)(-k - x + 1)}{x!} \\ &= (-1)^x \binom{-k}{x}\end{aligned}$$

if we extend the definition of the binomial coefficient to negative integers.

Problem 1.12. Show that

$$f_X(x) = e^{-\lambda}\frac{\lambda^x}{x!} \quad \text{for } x = 0, 1, 2, \dots.$$

is a valid PMF.

Solution 1.12. Clearly, $f_X(x)$ is non-negative on its support. We need to check if the sum is over the support 1. Using the exponential series $e^t = \sum_{k=0}^{\infty} \frac{t^k}{k!}$, we have

$$\sum_{x\geq 0} e^{-\lambda}\frac{\lambda^x}{x!} = e^{-\lambda}\sum_{x\geq 0}\frac{\lambda^x}{x!} = e^{-\lambda+\lambda} = 1.$$

Problem 1.13. (∗) Let $\lambda > 0$, and suppose that $p = p_n \to 0$ in such a way such that $np_n \to \lambda$. Let $X \sim \mathrm{Bin}(n, p_n)$ and $Y \sim \mathrm{Poi}(\lambda)$. Show that for $x \geq 0$,

$$\lim_{n\to\infty} \mathbb{P}(X = k) = \mathbb{P}(Y = k).$$

Solution 1.13. By assumption, $p = \frac{\lambda}{n}$. For every fixed integer $x \geq 0$,

$$\begin{aligned} f_X(x) = \binom{n}{x} p^x (1-p)^{n-x} &= \frac{n!}{x!(n-x)!}\Big(\frac{\lambda}{n}\Big)^x\Big(1-\frac{\lambda}{n}\Big)^{n-x} \\ &= \frac{\lambda^x}{x!}\Big(1-\frac{\lambda}{n}\Big)^n\Big(1-\frac{\lambda}{n}\Big)^{-x}\prod_{k=0}^{x-1}\frac{n-k}{n}. \end{aligned}$$

For each fixed x, we have

$$\lim_{n\to\infty}\Big(1-\frac{\lambda}{n}\Big)^{-x} = 1 \qquad \text{and} \qquad \lim_{n\to\infty}\prod_{k=0}^{x-1}\frac{n-k}{n} = 1.$$

Furthermore, $e^{-\lambda} = \lim_{n\to\infty}(1-\frac{\lambda}{n})^n$ by definition, so

$$\lim_{n\to\infty}\mathbb{P}(X = x) = \lim_{n\to\infty}\frac{\lambda^x}{x!}\underbrace{\Big(1-\frac{\lambda}{n}\Big)^n}_{\to e^{-\lambda}}\underbrace{\Big(1-\frac{\lambda}{n}\Big)^{-x}}_{\to 1}\underbrace{\prod_{k=0}^{x-1}\frac{n-k}{n}}_{\to 1} = e^{-\lambda}\frac{\lambda^x}{x!} = \mathbb{P}(Y = x).$$

Furthermore, $\mathbb{P}(X = x) = \mathbb{P}(Y = x) = 0$ for all other values.