Intro to discrete random variables (part 2)

Lecture 2b (STAT 24400 F24)

Discrete distributions

Any valid probability function specifies a distribution. However in applications, some specific probability functions are common and come up in many scenarios, so they are given special names.

Some common discrete distributions:

- Bernoulli
- Binomial
- Geometric
- Negative binomial
- Poisson
- (Hypergeometric)

Bernoulli distribution

X is a Bernoulli random variable if its only possible values are 0 and 1.

Its distribution is parametrized by the probability p, the probability of "success", i.e. getting a value 1.

The **Bernoulli**(p) distribution — with parameter $p \in [0, 1]$

The PMF is

$$\mathbb{P}(X = 0) = 1 - p, \quad \mathbb{P}(X = 1) = p$$

(and $\mathbb{P}(X = x) = 0$ for any other value of x)

Examples of experiments measuring a binary quantity:

- Flip a fair coin and record Heads (success) or Tails (failure): Bernoulli($\frac{1}{2}$)
- Roll a fair die and record whether we get a 1 (success) or not (failure): Bernoulli($\frac{1}{6}$)

Binomial distribution

The **Binomial**(n, p) distribution — parameters $n \geq 1$ (integer) & $p \in [0, 1]$.

It is the distribution of the number of successes from n trials, which are all independent and each have a probability p of success.

Example: Roll a dice 3 times and count the number of 1's: Binomial($3, \frac{1}{6}$).

The PMF for the binomial distribution is

$$\mathbb{P}(X = k) = \binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k}$$

for each $k = 0, \ldots, n$.

To understand the binomial formula: There are $\binom{n}{k}$ possible sequences that have exactly k 1's and $(n - k)$ 0's; each one has $p^k(1 - p)^{n-k}$ probability.

Examples of Binomial distribution

- Canonical prototype:
  Flip a coin 10 times and count number of Heads: Binomial$(10, \frac{1}{2})$.
- Common application in survey model:
  Survey 100 people chosen at random in Chicago, and ask a yes/no question — this type of experiment is often modeled as Binomial (even though technically the sampling is usually without replacement). The approximation is justified by relatively large size of the population.
- An example that is *not* binomial:
  Draw 10 cards from a deck and count the number of red cards.
  Here the sampled 10 cards is not small compare to the total of 52 cards.
- An example that is *not* binomial:
  Roll a dice until the first time you get a 6. How many rolls did it take?

Geometric distribution

The **Geometric**(p) distribution — parameter $p \in [0, 1]$.

It is the distribution of the number of trials needed to reach the first success, when the trials are all independent and each have a probability p of success.

The PMF is

$$\mathbb{P}(X = k) = (1-p)^{k-1} \cdot p$$

for each $k = 1, 2, 3, \ldots$

To understand this formula:
How many times do you have to flip a coin to get the first Heads?
Suppose p = probability of Heads.

$$\mathbb{P}(k \text{ times}) = \mathbb{P}(\text{first } k-1 \text{ are T, } k\text{th one is H}) = (1-p)^{k-1} \cdot p$$

(Note: Sometime X-1 is also called of geometric distribution.)

Memoryless property of Geometric distribution

Important property: the geometric distribution is **memoryless**.

Example of memoryless

Flip a coin until the first Heads (success).

Given that you've flipped the coin 100 times with no successes so far, the distribution of the additional number of flips needed to reach the first success is unchanged.

Example of *not* memoryless

Draw cards from a deck (without replacement) until the first red.

Because the trials of draws are not independent from each other, the number of draws needed are not of geometric distribution,

Negative binomial distribution

The NegativeBinomial(r, p) distribution — parameters $r \geq 1$ (integer) & $p \in [0, 1]$.

It is the distribution of the number of trials needed to reach the rth success, when trials are independent and each has probability p of success.

This is a generalization of the geometric distribution
— Geometric(p) and NegativeBinomial$(1, p)$ are the same distribution.

The PMF is

$$\mathbb{P}(X = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}$$

for each $k = r, r+1, r+2, \ldots$

To understand this formula (esp. compared with the binomial formula):
For the rth success to occur on the kth trial, we need $r-1$ successes among the first $k-1$ trials, and then we need one more success on the kth trial.

Poisson distribution

The Poisson(λ) distribution — parameter $\lambda > 0$

The PMF is

$$\mathbb{P}(X = k) = \frac{\lambda^k e^{-\lambda}}{k!} \qquad \text{for each} \quad k = 0, 1, 2, \ldots$$

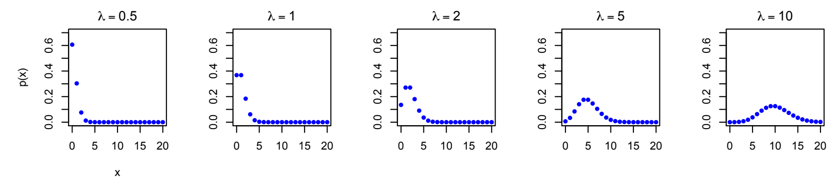
The Poisson distribution is a natural distribution for data that is a count that can be arbitrarily large, which occurs in many natural applications.

Example: photon emission from an X-ray source roughly follows this distribution.

Poisson paramter λ and PMF

Larger $\lambda \rightsquigarrow X$ is likely to be larger.
(λ corresponds to the intensity of the process.)

Examples of Poisson PMF:



Example - Binomial r. v.

- If you roll a fair die 50 times, what is the distribution of the # of 1's rolled? And, how likely is it that you will get no more than 10% 1's?

  Let X = number of 1's. Since
  - The trials of rolls are independent of each other
  - Each trial has 1/6 probability of success
  - The number of trials, 50, is determined ahead of time

  Then $\underbrace{X \sim \text{Binomial}(50, 1/6)}_{\text{i.e. the distribution of } X \text{ is the Binomial}(50,1/6) \text{ distribution}}$

$$\mathbb{P}(X \leq 5) = \sum_{k=0}^{5} \mathbb{P}(X = k) = \sum_{k=0}^{5} \binom{50}{k} (1/6)^k (5/6)^{50-k} = 0.139$$

Example (Hypergeometric distribution)

- If you draw 10 cards from a standard deck, what is the distribution of the number of Kings you draw?

  Let X = number of Kings.
  The distribution of X is *not* Binomial (since the trials are not independent).

  The PMF is
$$\mathbb{P}(X = k) = \frac{\binom{4}{k} \cdot \binom{48}{10-k}}{\binom{52}{10}},$$
  for each $k = 0, 1, 2, 3, 4$.
  This is an example of the Hypergeometric distribution.

Example - Poisson

- For a low intensity X-ray beam with Poisson distribution with $\lambda = 3$, what is the probability that at least one photon is emitted?

$$\mathbb{P}(X \geq 1) = 1 - \mathbb{P}(X = 0) = 1 - \frac{3^0 e^{-3}}{0!} = 0.95.$$

which is equivalent but simpler than

$$\mathbb{P}(X \geq 1) = \sum_{k=1}^{\infty} \mathbb{P}(X = k) = \sum_{k=1}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!}$$

Example - Geometric and Binomial

- Consider the following game:
  At each round, you roll a red die & a blue die.
  If the red die is even, you win a prize, otherwise you win nothing.
  If the blue die is a 1, then you stop playing, otherwise you continue.

  Q1. What is the distribution of the # of rounds you play?
  Q2. What is the distribution of the # of times you win?

Q1. Let $X = \#$ rounds you play.
Then $X = \#$ of rolls needed for the blue die to reach the first 1, so

$$X \sim \text{Geometric}\left(\tfrac{1}{6}\right)$$

i.e. The PMF is

$$\mathbb{P}(X = k) = \left(1 - \tfrac{1}{6}\right)^{k-1} \cdot \tfrac{1}{6} = \left(-\tfrac{5}{6}\right)^{k-1} \cdot \tfrac{1}{6} \quad \text{for each} \quad k = 1, 2, 3, \ldots$$

Example - Geometric and Binomial (cont.)

Q2. Let $X = \#$ rounds you play, and $Y = \#$ times you win.
To calculate the PMF of Y: for each $k = 0, 1, 2, \ldots,$

$$\begin{aligned}
\mathbb{P}(Y = k) = \sum_{\ell \geq 1} \mathbb{P}(Y = k \text{ and } X = \ell) \\
&= \sum_{\ell \geq 1} \mathbb{P}(X = \ell) \cdot \mathbb{P}(Y = k \mid X = \ell) \\
&= \sum_{\ell \geq k} \mathbb{P}(X = \ell) \cdot \mathbb{P}(Y = k \mid X = \ell) \\
&= \sum_{\ell \geq k} \underbrace{\left(\tfrac{5}{6}\right)^{\ell-1} \cdot \tfrac{1}{6}}_{\text{from Geom. distrib.}} \cdot \underbrace{\binom{\ell}{k} \left(\tfrac{1}{2}\right)^k \left(\tfrac{1}{2}\right)^{\ell-k}}_{\text{from Binomial distrib.}}
\end{aligned}$$

This example illustrates the usefulness of conditioning.