Discrete random variables

Recall: Random Variable

- **Definition 1:** *A random variable is a function from a sample space* $S(or, \Omega)$ *to the real numbers. Conventionally, random variables are denoted with capital letters, e.g.,* $X : S \to R \coloneqq (-\infty, +\infty)$
- **If the outcome of the random experiment is ω, then the value of the random variable is $X(\boldsymbol{\omega}) \in \mathbb{R}$**
- **Examples**
  - **Toss a coin 10 times and let X be the number of Heads**
  - **Choose a random person in a class and let X be the height of the person, in inches.**

Recall: Discrete Random Variable

- **Random variables that can assume a countable number (finite or infinite) of values are called discrete.**
- **Examples**

Experiment	Random Variable	Possible Values
Make 100 Sales Calls	# Sales	0, 1, 2, ..., 100
Inspect 70 Radios	# Defective	0, 1, 2, ..., 70
Answer 33 Questions	# Correct	0, 1, 2, ..., 33

Recall: Continuous Random Variable

- **A continuous random variable is a random variable with infinitely many possible values (in an interval of real numbers).**

Experiment	Random Variable	Possible Values
Weigh 100 People	Weight	45.1, 78, ...
Measure time taken	Hours	900, 875.9, ...
Amount spent on food	$ amount	54.12, 42, ...
Measure Time Between Arrivals	Inter-Arrival Time	0, 1.3, 2.78, ...

Recall: Probability Distributions

- **Definition: Probability Distribution**
  - *A probability distribution of a random variable X is a description of the probabilities associated with the possible values of X*
- **Example**
  - **Let X = # of heads observed when a coin is flipped twice**

Number of Heads	0	1	2
Probability	1/4	2/4	1/4

Recall: Probability Mass Function (PMF)

- For a discrete random variable X with possible values $x_1, x_2, x_3, \ldots, x_n$, a probability mass function $p(x_i)$ is a function such

    1. $p(x_i) \geq 0$ for all values of x_i
    2. $\sum p(x_i) = 1$
    3. $p(x_i) = P(X = x_i)$

- A **pmf** is a function that gives the probability that a discrete random variable is exactly equal to some value.

Example (Probability Mass Function (PMF)):
Tossing a die

x	*p(x)*
1	*p(x=1)*=1/6
2	*p(x=2)*=1/6
3	*p(x=3)*=1/6
4	*p(x=4)*=1/6
5	*p(x=5)*=1/6
6	*p(x=6)*=1/6
	1.0

Recall: Cumulative Distribution Function (CDF)

- **The cumulative distribution function of a discrete random variable X, denoted as F(x), is**

$$F(x) = P(X \le x) = \sum_{x_i \le x} p(x_i)$$

- **Properties**
  1. $F(x) = P(X \le x) = \sum_{x_i \le x} p(x_i)$
  2. $0 \le F(x) \le 1, F(-\infty) = 0, F(+\infty) = 1$
  3. **If** $x \le y$, **then** $F(x) \le F(y)$

Example (CDF): Tossing a die

1
2
3
4
5
6

Recall: Summary Measures

- Expected Value (Mean of probability distribution)
  - Assume that X is a discrete random variable with possible values $x_i, i = 1, 2, \ldots$. Then, the expected value, also called expectation, average, or mean, of X is

$$\mu = E[X] = \sum_i x_i p(x_i)$$

  - For any function, $g : \mathbb{R} \to \mathbb{R}$

$$E[g(X)] = \sum_i g(x_i) P(X = x_i)$$

- Variance: Weighted average of squared deviation about mean

$$Var(X) = \sigma = E\left[(X - \mu)^2\right]$$
$$= \sum_i (x_i - \mu)^2 P(X = x_i)$$

- Standard Deviation
  - $\sigma = \sqrt{\sigma^2}$

Example

- An urn contains 20 balls numbered $1, \ldots, 20$. Select 5 balls at random, without replacement. Let X be the largest number among selected balls. Determine its p. m. f. and the probability that at least one of the selected numbers is 15 or more.
- Solution
  - Possible outcome: $\binom{20}{5}$
  - $P(X = i) = \dfrac{\binom{i-1}{4}}{\binom{20}{5}}$
  - $P\ (X \geq 15) = \sum_{i=15}^{20} P(X = i)$

Example

- An urn contains 11 balls, 3 white, 3 red, and 5 blue balls. Take out 3 balls at random, without replacement. You win \$1 for each red ball you select and lose a \$1 for each white ball you select. Determine the **p. m. f.** of X, the amount you win.
- Solution
  - Possible outcome: $\binom{11}{3} = 165$
  - X can have values -3, $-2, -1, 0, 1, 2$, and 3.
  - $P(X = 0)$
    - This can occur with one ball of each color or with 3 blue balls

$$P(X = 0) = \frac{\binom{3}{1}\binom{3}{1}\binom{5}{1} + \binom{5}{3}}{\binom{11}{3}}$$

  - Similarly, you can calculate for $P(X = -3)$, $P(X = -2)$, $P(X = -1)$, $P(X = 1)$, $P(X = 2)$, and $P(X = 3)$

Expected Value

- **Theorem. (Linearity)** If X and Y are discrete random variables and $a \in \mathbb{R}$ then
  a. $E[X+Y] = E[X] + E[Y]$
  b. $E[aX] = aE[X]$
  c. $E[aX + bY] = aE[X] + bE[Y]$

- **Proof.**

$$E[X+Y] = \sum_{\omega \in S} \big(X(\omega) + Y(\omega)\big)p(\omega)$$

$$= \sum_{\omega \in S} X(\omega)p(\omega) + Y(\omega)p(\omega)$$

$$= \sum_{\omega \in S} X(\omega)p(\omega) + \sum_{\omega \in S} Y(\omega)p(\omega)$$

$$= E[X] + E[Y]$$

- **Proof.**

$$E[aX] = \sum_{\omega \in S} \big(aX(\omega)\big)p(\omega)$$

$$= a\sum_{\omega \in S} X(\omega)p(\omega)$$

$$= aE[X]$$

Variance

- $Var(X) = \sigma^2 = E\left[(X-\mu)^2\right] = E\left[X^2 - 2.X.\mu + \mu^2\right]$
  $= E\left[X^2\right] - 2\mu E[X] + (E[X])^2$
  $= E\left[X^2\right] - 2E[X]E[X] + (E[X])^2$
  $= E\left[X^2\right] - (E[X])^2$
- **For constants a, b we have that $Var(aX + b) \ = \ a^2 Var(X)$**
- **Proof.**

$$
\begin{aligned}
Var(aX + b) &= E\left[(aX + b \ - E[aX + b])^2\right] \\
&= E\left[(aX + b \ - aE[X] - E[b])^2\right] \\
&= E\left[(aX + b \ - aE[X] - b)^2\right] \\
&= E\left[(aX - a\mu)^2\right] \\
&= a^2 E\left[(X-\mu)^2\right] = a^2 Var(X)
\end{aligned}
$$

Uniform discrete random variable

- We say X is uniform, and write this as $X \sim uniform(n)$, if $X \in \{x_1, x_2, \dots, x_n\}$
  $$p(x_i) = P(X = x_i) = \frac{1}{n}, for\ i = 1, 2, \dots, n$$
- $E[X] = \frac{x_1 + x_2 + \cdots + x_n}{n}$
- $Var(X) = \frac{x_1^2 + x_2^2 + \cdots + x_n^2}{n} - \left(\frac{x_1 + x_2 + \cdots + x_n}{n}\right)^2$
- Example
  - Let X be the number shown on a rolled fair die. Then X follow uniform distribution

Bernoulli random variable

- This is also called an indicator random variable.
- Assume that A is an event with probability p. Then, I_A, the indicator of A, is given by

$$I_A = \begin{cases} 1 \text{ if } A \text{ happens} \\ 0 \text{ otherwise.} \end{cases}$$

- $E[X] = p$
- $Var[I_A] = E[I_A^2] - (E[I_A])^2$

$$= p - p^2$$
$$= p(1-p)$$

Binomial random variable

- Characteristics of a Binomial Experiment
  - The experiment consists of n independent identical trials.
  - There are only two possible outcomes on each trial. We will denote one outcome by S (for success) and the other by F (for failure).
  - The probability of S remains the same from trial to trial. This probability is denoted by p, and the probability of F is denoted by q. Note that q = 1 – p.
- The binomial random variable X is the number of S’s in n trials.

$$P(X=i)=\binom{n}{i}p^{i}(1-p)^{n-i},$$

Binomial example

- **Take the example of 5 coin tosses. What's the probability that you flip exactly 3 heads in 5 coin tosses?**

Outcome	Probability
THHHT	$(1/2)^3 \times (1/2)^2$
HHHTT	$(1/2)^3 \times (1/2)^2$
TTHHH	$(1/2)^3 \times (1/2)^2$
HTTHH	$(1/2)^3 \times (1/2)^2$
HHTTH	$(1/2)^3 \times (1/2)^2$
THTHH	$(1/2)^3 \times (1/2)^2$
HTHTH	$(1/2)^3 \times (1/2)^2$
HHTHT	$(1/2)^3 \times (1/2)^2$
THHTH	$(1/2)^3 \times (1/2)^2$
HTHHT	$(1/2)^3 \times (1/2)^2$

10 arrangements $\times (1/2)^3 \times (1/2)^2$

$\binom{5}{3}$ ways to arrange 3 heads in 5 trials

$5!/3!2! = 10$

The probability of each unique outcome (note: they are all equal)

Binomial example

- $$\text{P(3 heads and 2 tails)} = \binom{5}{2} \times p(H)^3 \times p(T)^2$$
  $$= 10 \times \left(\frac{1}{2}\right)^5 = 0.3125$$

Binomial example

You're a telemarketer selling service contracts for Jio. You've sold 20 in your last 100 calls ($p = .20$). If you call 12 people tonight, what's the probability of

a. No sales?
b. Exactly 2 sales?
c. At most 2 sales?
d. At least 2 sales?

Binomial example

$n = 12$, $p = .20$

a. $p(0) = .0687$

b. $p(2) = .2835$

c. $p(\text{at most } 2) = p(0) + p(1) + p(2)$
$= .0687 + .2062 + .2835$
$= .5584$

d. $p(\text{at least } 2) = p(2) + p(3) \ldots + p(12)$
$= 1 - [p(0) + p(1)]$
$= 1 - .0687 - .2062$
$= .7251$

Properties

- Probability mass function: $P(X = i) = \binom{n}{i} p^i (1-p)^{n-i}$, $i = 0, 1, \dots, n$
- $\sum_{i=0}^{n} P(X = i) = 1$
  - Binomial Theorem: $\sum_{i=0}^{n} \binom{n}{i} p^i q^{n-i} = (p+q)^n$
  - $\sum_{i=0}^{n} \binom{n}{i} p^i (1-p)^{n-i} = \left(p + (1-p)\right)^n = (1)^n = 1$
- $E[X] = np$ [proof next slide]
- $Var(X) = np(1-p)$

Binomial random variable

- $$(p+q)^n = \sum_{\boldsymbol{i}=0}^{n} \binom{n}{\boldsymbol{i}} (p)^{\boldsymbol{i}} (q)^{n-\boldsymbol{i}}$$
- **By differentiating both sides with respect to p, we get:**

$$n(p+q)^{n-1} = \sum_{\boldsymbol{i}=0}^{n} i \binom{n}{\boldsymbol{i}} (p)^{\boldsymbol{i}-1} q^{n-\boldsymbol{i}}$$

$$n(p+q)^{n-1} = \frac{\mathbf{1}}{\boldsymbol{p}} \sum_{\boldsymbol{i}=0}^{n} i \binom{n}{\boldsymbol{i}} (p)^{\boldsymbol{i}} q^{n-\boldsymbol{i}}$$

Binomial random variable

- $$n(p+q)^{n-1} = \frac{\mathbf{1}}{\boldsymbol{p}}\sum_{\boldsymbol{i}=0}^{n} i\binom{n}{\boldsymbol{i}}(p)^{\boldsymbol{i}}q^{n-\boldsymbol{i}}$$

- **Substitute** $q = 1 - p$

$$np = \sum_{\boldsymbol{i}=0}^{n} i\binom{n}{\boldsymbol{i}}(p)^{\boldsymbol{i}}q^{n-\boldsymbol{i}}$$

Poisson Distribution

- Number of events that occur in an interval
  - events per unit
    - Time, Length, Area, Space
- Examples
  - Number of customers arriving in 20 minutes
  - Number of strikes per year in the India.
  - Number of defects per lot (group) of DVD’s

Poisson Distribution

1. Consists of counting number of times an event occurs during a given unit of time or in a given area or volume (any unit of measurement).
2. The probability that an event occurs in a given unit of time, area, or volume is the same for all units.
3. The number of events that occur in one unit of time, area, or volume is independent of the number that occur in any other mutually exclusive unit.
4. The mean number of events in each unit is denoted by $\lambda > 0$

Poisson Probability Distribution Function

- $$p(x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

- $p(x)$ = Probability of x given λ
- λ = Mean (expected) number of events in unit
- e = 2.71828 . . . (base of natural logarithm)
- x = Number of events per unit

Poisson Distribution

- Customers arrive at a rate of 72 per hour. What is the probability of 4 customers arriving in 3 minutes?
- Solution
  - 72 Per Hr. = 1.2 Per Min. = 3.6 Per 3 Min. Interval

$$p(x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

$$p(4) = \frac{(3.6)^4 e^{-3.6}}{4!} = .1912$$

Poisson Distribution

- **Bob receives texts on the average of two every 3 minutes. Assume Poisson.**
- **Question: What is the probability of five or more texts arriving in a 9–minute period**

$$
\begin{aligned}
\mathbb{P}(X \geq 5) &= 1 - \mathbb{P}(X \leq 4) \\
&= 1 - \sum_{n=0}^{4} \frac{e^{-6} 6^n}{n!} \\
&= 1 - .285 = .715.
\end{aligned}
$$

Properties of Poisson Distribution

- $\sum_{i=0}^{\infty} p(x_i) = 1$
  - $\sum_{i=0}^{\infty} e^{-\lambda} \frac{\lambda^i}{i!} = 1$
  - $e^{-\lambda} \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = e^{-\lambda} e^{\lambda} = 1$

- $E[X] = \sum_{i=0}^{\infty} i \, e^{-\lambda} \frac{\lambda^i}{i!}$
  - $e^{-\lambda} \lambda \sum_{i=1}^{\infty} \frac{\lambda^{i-1}}{(i-1)!} = e^{-\lambda} \lambda e^{\lambda} = \lambda$

- $Var(X) = \lambda (\lambda + 1) - \lambda^2 = \lambda$

Recall: Uniform discrete random variable

- We say X is uniform, and write this as $X \sim uniform(n)$, if $X \in \{x_1, x_2, \ldots, x_n\}$

$$p(x_i) = P(X = x_i) = \frac{1}{n}, for\ i = 1, 2, \ldots, n$$

- $E[X] = \frac{x_1 + x_2 + \cdots + x_n}{n}$
- $Var(X) = \frac{x_1^2 + x_2^2 + \cdots + x_n^2}{n} - \left(\frac{x_1 + x_2 + \cdots + x_n}{n}\right)^2$
- Example
  - Let X be the number shown on a rolled fair die. Then X follow uniform distribution

Recall: Binomial random variable

- Characteristics of a Binomial Experiment
  - The experiment consists of n independent identical trials.
  - There are only two possible outcomes on each trial. We will denote one outcome by S (for success) and the other by F (for failure).
  - The probability of S remains the same from trial to trial. This probability is denoted by p, and the probability of F is denoted by q. Note that q = 1 – p.
- The binomial random variable X is the number of S's in n trials.

$$P(X = i) = \binom{n}{i} p^i (1 - p)^{n-i},$$

Recall: Bernoulli random variable

- **Bernouilli trial: If there is only 1 trial with probability of success *p* and probability of failure *1-p*, this is called a Bernouilli distribution. (special case of the binomial with *n*=1)**
- **Probability of success:** $$P(X=1)=\binom{1}{1}p^1(1-p)^{1-1}=p$$
- **Probability of failure:** $$P(X=0)=\binom{1}{0}p^0(1-p)^{1-0}=1-p$$
- $E[X] = p$
- $Var[X] = p(1-p)$

Recall: Poisson Distribution

1. Consists of counting number of times an event occurs during a given unit of time or in a given area or volume (any unit of measurement).
2. The probability that an event occurs in a given unit of time, area, or volume is the same for all units.
3. The number of events that occur in one unit of time, area, or volume is independent of the number that occur in any other mutually exclusive unit.
4. The mean number of events in each unit is denoted by $\lambda > 0$

$$p(x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

- $p(x)$ = Probability of x given λ
- λ = Mean (expected) number of events in unit
- e = 2.71828 . . . (base of natural logarithm)
- x = Number of events per unit

Practice Example

- **As voters exit the polls, you ask a representative random sample of 6 voters if they voted for *xyz*. If the true percentage of voters who vote for *xyz* is 55.1%, what is the probability that, *in your sample,* exactly 2 voted for *xyz* and 4 did not?**

$$P(2\ yes\ votes\ exactly) = \binom{6}{2} \times (.551)^2 \times (.449)^4 = 18.5\%$$

Practice Example

- **Four fair coins are flipped. If the outcomes are assumed independent, what is the probability that two heads and two tails are obtained?**

$$P(2\ heads) = \binom{4}{2} \times \left(\frac{1}{2}\right)^2 \times \left(\frac{1}{2}\right)^2 = \frac{3}{8}$$

Geometric Distribution

- Suppose that independent trials, each having probability p of being a success, are performed until a success occurs.
  - If we let X be the number of trials required until the first success, then X is said to be a geometric random variable with parameter p.
  - $p(n) = P(X = n) = (1 - p)^{n-1}p, n = 1, 2, \ldots$
- For $p(n)$ to be pmf
  - $\sum_{n=1}^{\infty} p(n) = 1$
    - $p\sum_{n=1}^{\infty}(1 - p)^{n-1} = 1$
    - Let, $\sum_{n=1}^{\infty}(1 - p)^{n-1} = 1 + (1 - p)^1 + (1 - p)^2 + \cdots = f(x)$
    - $1 + (1 - p)(1 + (1 - p)^1 + (1 - p)^2 + \cdots) = f(x)$
    - $1 + (1 - p)f(x) = f(x) \Rightarrow f(x) = \frac{1}{p}$
    - $pf(x) = 1$

$$E[X] = \frac{1}{p}$$

Negative Binomial

- **Binomial:**
  1. Fixed number of n trials.
  2. Each trial is independent.
  3. Only two outcomes are possible (Success and Failure).
  4. Probability of success (p) for each trial is constant.
  5. A random variable Y= the number of successes.
- **Negative Binomial**
  1. The number of trials, n is not fixed.
  2. Each trial is independent.
  3. Only two outcomes are possible (Success and Failure).
  4. Probability of success (p) for each trial is constant.
  5. A random variable Y= the number of trials needed to make r successes.

Negative Binomial

- **Negative Binomial**
  1. **The number of trials, n is not fixed.**
  2. **Each trial is independent.**
  3. **Only two outcomes are possible (Success and Failure).**
  4. **Probability of success (p) for each trial is constant.**
  5. **A random variable Y= the number of trials needed to make r successes.**
- **PMF**

$$p(X=x)=\binom{x-1}{r-1}p^{r}(1-p)^{x-r}$$

- $E[X]=\frac{r}{p}, Var[X]=\frac{r(1-p)}{p^{2}}$

Hypergeometric Distribution

- Assumptions
  - Population consists of N objects (finite)
  - Each classified as a “S” or “F” and K success
  - t individuals is selected without replacement
- Random variable X, the number of successes in the sample
  - X=number of successes in a sample of size t drawn from a population consisting of K successes and N-K failures
- PMF
  - $P(X = q) = \dfrac{\binom{K}{q}\binom{N-K}{t-q}}{\binom{N}{t}}$

Hypergeometric Distribution

- Assumptions
  - Population consists of N objects (finite)
  - Each classified as a “S” or “F” and K success
  - t individuals is selected without replacement
- Random variable X, the number of successes in the sample
  - X=number of successes in a sample of size n drawn from a population consisting of K successes and N-K failures
- $E[X] = \frac{tK}{N}, Var[X] = \frac{tK}{N}\left(\frac{N-K}{N}\right)\left(\frac{N-t}{N-1}\right)$

Practice Example

- Suppose that an airplane engine will fail, when in flight, with probability $1 - p$ independently from engine to engine; suppose that the airplane will make a successful flight if at least 50 percent of its engines remain operative. For what values of p is a four-engine plane preferable to a two-engine plane?
- **Solution:**
- Probability that a four-engine plane makes a successful flight is $\binom{4}{2} p^2(1-p)^2 + \binom{4}{3} p^3(1-p)^1 + \binom{4}{4} p^4(1-p)^0$
- Probability that a two-engine plane makes a successful flight is $\binom{2}{1} p^1(1-p)^1 + \binom{2}{2} p^2(1-p)^0 = 2p(1-p) + p^2$
- $\binom{4}{2} p^2(1-p)^2 + \binom{4}{3} p^3(1-p)^1 + \binom{4}{4} p^4(1-p)^0 \geq 2p(1-p) + p^2$

Reference

- Lecture notes on Probability Theory by Phanuel Mariano
- **Schaum's Outline of Probability and Statistics, 4th Edition**