# Class Notes

STAT410

**Wenyu Chen**

Summer 2022

# 1 Axioms of Probabilities

**Definition 0.0.0.0.1**
***Experiment:*** *is repeatable task with well defined outcomes.*
***Sample Space:*** *is a xollection of all possible outcomes of the experiment.*
***Event:*** *is a subset of the sample space.*

Example: Suppose we toss a coin three times, assume coins lands on H or T.
**Experiment:** tossing the coin three times, noting the outcomes.
**Sample Space:** {HHH, HHT,HTH, THH, HTT,THT,TTH, TTT }
**Example of events:** $E_1$ = getting all heads = $\{HHH\} \subseteq S$
$E_2$ = getting exactly one H and one T = $\{\} \subseteq S$
$E_3$ = getting at least two heads = $\{HHH, HHT, HTH, THH\} \subseteq S$

We want to assign a number to each event, which is a measure of the chance or probabilities that this event happened. Our goal is to understand the process of assigning probabilities to events.

Example: **Die Roll**
**Experiment:** Roll a six sided dice two times
**Sample Space:**$\{(1,1),(1,2)...(1,6),(2,1)...(2,6)...(6,6)\}$
**Example of events:**
$E_1$ = at least one six = $\{(1,6),(2,6),(3,6),(4,6),(5,6),(6,6),(6,1),(6,2),(6,3),(6,4),(6,5)\}$
$|E_2| = 10$
$E_2$ = same numbers on both rolls = $\{(1,1),(2,2),(3,3),(4,4),(5,5),(6,6)\}$
$E_4 = \{(1,5),(3,4)\}$

Note:
**Simple Event:** is the event with exactly one outcomes. It is the cardinality of the sample space.
**Number of events:** Assume the sample space is finite of size n, the number of events is the cardinality of all possible outcomes of S = $2^n$

**Set in context**
Let an event E be describes as a subset of the sample space S.
Let E, F two events be given.
$E \cap F$ is a new event that corresponds to outcomes in both E and F
$E \cup F$ is a outcome a new event include outcomes in E or F
$E^c$ is a new event that include outcomes where E doesn't happen.

When we say an event E has happened, we means that the outcome $\omega$ of the experiment lies inside E
Example: Tossed a coin three times. In one run of the experiment, the result is HHT.
HHT $\in E_1$ = at least two heads, we said $E_1$ has happened.

**Definition 0.0.0.0.2**
*A sigma algebra $\mathscr{B}$ for a set s is a collection of subset of S that satisfies:*

1. $\emptyset \in \mathscr{B}$
2. *If $A \in \mathscr{B}$ then $A^c \in \mathscr{B}$ (closed under complement)*
3. *If $\{A_i\}_{i=1}^{\infty} \in \mathscr{B}$ then $\cap_{i=1}^{\infty} A_i \in \mathscr{B}$ (closed under)*

Sigma algebera is a collection of events to which we want to assign the probabilities.

## Axioms for a probability function

Suppose we are given the pair $(S, \mathscr{B})$ which S represents the sample space and $\mathscr{B}$ is a sigma algebra S.

**Definition 1.0.0.0.1**
*A probability function $P$ satisfies the following:*

$$P : \mathscr{B} \to \mathbb{R}$$
$$E \mapsto P(E)$$

1. $P(A) \geq 0$ *for all* $A \in \mathscr{B}$
2. $P(S) = 1$
3. *If* $A_1, A_2, A_3, ...$ *are mutually disjiont sets in* $\mathscr{B}$ *then* $P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$

**Theorem 1.1**
*Suppose the sample space $S$ is countable. To define a probability function on $(S, \mathscr{B} = P(S))$, we do the following*

1. *Find a seq* $\{p_i\}_{i=1}^{\infty}$ *such that (i)* $0 \leq p_i \leq 1$ *and (ii)* $\sum_{i=1}^{\infty} p_i = 1$
2. *Define* $P(\{s_i\}) = p_i$
3. *For any* $E \in \mathscr{B}, E = \{s_{i1}, s_{i2}, ...s_{ij}\}$

Example: Probability Functions Examples
Coin Tosseo:
**Experiment:** tossing the coin three times, noting the outcomes.
**Sample Space:** {HHH, HHT,HTH, THH, HTT,THT,TTH, TTT }
To get a probability function, we will need to work with a sigma algebra. Suppose $\mathscr{B}_1 = \mathscr{P}(S)$, note that the cardinality of $\mathscr{B}_1$ is 256.
There are infinitely many ways to choose the sequence $p_i$. One way is to choose $p_i = \frac{1}{8}$ for all i. Another way is to set $p_1 = 1$ and rest to 0.

**Tree Diagram** is a graph that describes the flow of the outcomes of each steps in an experiment.

**Definition 1.1.0.0.1**
***Conditional Probability*** *is* $P(A \mid B)$= *P(A happens given that B has already happened)*

$$= \frac{P(A \cap B)}{P(B)}$$

*And by **Multiplication Principle**,*

$$P(A \cap B) = P(A \mid B) \times P(B)$$
$$= P(B \mid A) \times P(A)$$

Assumption: 1) Probability of events when all outcomes in the sample space are equally likely. 2) Sample space is finite.

**Proposition 1.1.1**
*If every outcome in the sample space is equally likely, we can calculate the probability of $E \subseteq S$ as follow:*

$$P(E) = \frac{n(E)}{n(S)}$$

*where $n(E)$ is the number of outcomes in E.*

Example: Suppose we toss a coin that $p(H) \in (0,1)$ two times. The sample space is S = {HH, HT, TH, TT}. If the coin is fair, then $P(HH) = P(HT) = P(TH) = P(TT) = 0.25$

**Theorem 1.2**
***Fundmental Theorem of Counting:*** *Suppose a task $T$ can be performed as a sequence of subtasks: $T_1, T_2, T_3, ..., T_k$. And each $n_1, ..., n_2, n_3, ..., n_k$ is number of ways to perform $T_i$*
*Then the total number ways to perform the task $T$ is*

$$n_1 \times n_2 \times n_3 \times \cdots \times n_k$$

Typically we will have to select k objects from n distinct objects.
Example: $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$, we might be interested in knowing the total number of ways one can choose 4 digits from this 10 digits.

| | Without Replacement | With Replacement |
|---|---|---|
| Order Matters | (1,2,4,5) different from (1,5,2,4)<br>(1,1,2,5) is not possible | (1,2,4,5) different from (1,5,2,4)<br>(1,1,2,5) is possible |
| Order Does Not Matters | (1,2,3,4) is same as (4,3,2,1)<br>(1,1,2,5) not possible | (1,2,3,4) is same as (4,3,2,1)<br>(1,1,2,4) is possible |

**1. Without replacement and order matters**
Use the fundamental theorem of counting, we divide T, which is select k digits from a set of n distinct objects divide into

$$T : T_1 \to T_2 \to T_3 \to \cdots \to T_k$$

where $T_i$ is select ith object. Then, we will got

$$n \times (n-1) \times (n-2) \times (n-3) \times \cdots \times (n-k+1)$$

Then, we got

$$^nP_k = \frac{n!}{(n-k)!}$$

**2. Without replacement and order does not matter**
T = choose k objects from n distinct objects where order does not matter and without replacement.

$$T : T_1 \to T_2$$

$T_1$ is choose k objects where order matters and without replacement
$T_2$ is to get rid of all the times we have double counted.

\# ways to do $T_1 = {}^nP_k = \frac{n!}{(n-k)!}$
\# ways to do $T_2$ = number of arrangements of k objects = $k!$
Then we got

$$^nC_k = \binom{n}{k} = \frac{n!}{(n-k)!k!}$$

**3. With replacement and order does not matter**
$T$ = Choose k objects where order does not matter and with replacement.

We keep track of how many times a given object repeats in the selection and the total number of objects in the selection is equal to k.
Which is the same as choosing $n-1$ walls in a set of $(n+k-1)$, we got

$$^{n+k-1}C_k = {}^{n+k-1}C_{n-1}$$

Example: $\{1,2,3,4\}$, $k=10$, We are selecting 10 objects from $\{1,2,3,4\}$ with replacement and order does not matter.
We only care about how many times each number shows up since order does not matter and the total objects in selection are 10. To achieve that, we setup 13 spots. Such that, the extra position is how many times the given number repeats in the selection.
Such that, the

# 2 | Conditional Probability

Recall that given two events A and B, the conditional probability is

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

**Definition 0.0.0.0.1**
*WE say two events A and B are independent if and only if*

$$\begin{aligned} P(A \cap B) &= P(A) \cdot P(B) \\ P(A \mid B) &= P(A) \\ P(B \mid A) &- P(B) \end{aligned}$$

*This means that B has happened will not affect the probability of A happening. Equivalent to saying that $P(A \mid B) = P(A)$*

Note: $P(A \cap B)$ is the probability that A and B happen simultaneously.

**Theorem 0.1**

$$\begin{aligned} P(A \mid B) &= \frac{P(A \cap B)}{P(B)} \\ &= \frac{P(B \mid A) \cdot P(A)}{P(B)} \end{aligned}$$

*and*

$$\begin{aligned} P(B \mid A) &= \frac{P(B \cap A)}{P(A)} \\ &= \frac{P(A \mid B) \cdot P(B)}{P(A)} \end{aligned}$$

Exercise: Now if A and B and independent, we want to check if $A$ and $B^c$ is independent, which is checking if $P(A \cap B^c) = P(A)P(B^c)$.

Example:
**Experiment:**
Step1: Toss a coin with $P(H) = 0.4$
Step2: If we got head, we will roll a fair dice.
Else, we will roll a dice with $P(1) = P(2) = \cdots = P(5) = 0.1$ and $P(6) = 0.5$

**Question:** Given that they observed a 6, what is the probability that a fair dice was rolled?
So we need to calculate $P(H \mid 6)$ which is the probability of rolling a fair dice (toss 6) given that 6 was observed.

**Theorem 0.2**
***Law of Total Probability***
*If $A_1, A_2, \cdots, A_k$ are given, and*

Note: $\bigcup_{i=0}^{k} |\{X = i\}| = 2^n$

$$
\begin{aligned}
E(X) &= \sum_{x \in X} x \cdot p_X(x) \\
&= \sum_{k=0}^{n} k \cdot p_x(k) \\
&= \sum_{k=0}^{n} k \cdot \binom{n}{k} p^k \cdot (1-p)^{n-k} \\
&= \sum_{k=0}^{n} k \cdot \frac{n!}{(n-k)!k!} p^k \cdot (1-p)^{n-k} \\
&= \sum_{k=1}^{n} k \cdot \frac{n!}{(n-k)!k!} p^k \cdot (1-p)^{n-k} \\
&= \sum_{k=1}^{n} \frac{n!}{(n-k)!(k-1)!} p^k \cdot (1-p)^{n-k} \\
&= \sum_{k=1}^{n} n \cdot p \frac{(n-1)!}{(n-k)!(k-1)!} p^{k-1} \cdot (1-p)^{n-k} \\
&= n \cdot p \cdot \sum_{k=1}^{n} \frac{(n-1)!}{(n-k)!(k-1)!} p^{k-1} \cdot (1-p)^{n-k} \\
&= n \cdot p \cdot \sum_{r=0}^{n-1} \frac{(n-1)!}{(n-r-1)!(r)!} \cdot p^r \cdot (1-p)^{n-r-1} && \text{set } r = (k-1) \\
&= n \cdot p \cdot \sum_{r=0}^{n-1} \frac{(n-1)!}{((n-1)-r)!(r)!} \cdot p^r \cdot (1-p)^{(n-1)-r} \\
&= n \cdot p \cdot \sum_{r=0}^{n-1} \binom{n-1}{r} \cdot p^r \cdot (1-p)^{(n-1)-r} \\
&= n \cdot p && \text{note it is a pmf for binom((n-1), p), so the sum is 1}
\end{aligned}
$$

Then, let's calculate the Variance

$$
\begin{aligned}
E(X^2) &= \sum_{k=0}^{n} k^2 \cdot p_X(k) \\
&= \sum_{k=0}^{n} k^2 \cdot \binom{n}{k} p^k \cdot (1-p)^{n-k} \\
&= \sum_{k=0}^{n} k^2 \cdot \frac{n!}{(n-k)!k!} p^k \cdot (1-p)^{n-k} \\
&= \sum_{k=1}^{n} k^2 \cdot \frac{n!}{(n-k)!k!} p^k \cdot (1-p)^{n-k} \\
&= \sum_{k=1}^{n} k \cdot \frac{n!}{(n-k)!(k-1)!} p^k \cdot (1-p)^{n-k} \\
&= \sum_{k=1}^{n} k \cdot n \cdot p \frac{(n-1)!}{(n-k)!(k-1)!} p^{k-1} \cdot (1-p)^{n-k} \\
&= \sum_{k=1}^{n} k \cdot n \cdot p \frac{(n-1)!}{(n-k)!(k-1)!} p^{k-1} \cdot (1-p)^{n-k} \\
&= n \cdot p \left( \sum_{r=0}^{n-1} r \binom{n-1}{r} p^r \cdot (1-p)^{(n-1)-r} + \sum_{r=0}^{n-1} \binom{n-1}{r} p^r \cdot (1-p)^{(n-1)-r} \right) \\
&= n \cdot p((n-1)p + 1) \\
&= n(n-1)p^2 + np
\end{aligned}
$$

$$
\begin{aligned}
V(X) &= E(X^2) - E(X)^2 \\
&= n(n-1)p^2 + np - (np)^2 \\
&= np((n-1)p + 1 - np) \\
&= np(np - p + 1 - np) \\
&= np \cdot (1-p)
\end{aligned}
$$

Moment Generating Function

$$
\begin{aligned}
M_X(t) &= E(e^{tx}) \\
&= \sum_{k=0}^{n} e^{tk} \cdot p_X(k) \\
&= \sum_{k=0}^{n} e^{tk} \cdot \binom{n}{k} p^k \cdot (1-p)^{n-k} \\
&= \sum_{k=0}^{n} e^{tk} \cdot \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} \\
&= \sum_{k=0}^{n} \binom{n}{k} \cdot (e^t p)^k \cdot (1-p)^{n-k} \\
&= (pe^t + (1-p))^n
\end{aligned}
$$

**Hyper Geometric Distribution**
$N$ = population size
$M$ = number success in the population
$n$ = sample size
$p_X(k) = \frac{\binom{M}{k} \cdot \binom{N-M}{n-k}}{\binom{N}{k}}$
In this setting, $X = k$ number of successes in a sample of size n, sampled from a population of size N with M success and sampling without replacement.
$|\{X = k\}|$ = number elements in this set. Find the exactly k spots in n spots for the successes. Fill the remaining $(n-k)$ spots with Failure.

And the total number of ways to n objects from N without replacements is $\binom{N}{n}$

## The Negative Binomial Distribution

Waiting for a certain number of successes
**Experiment:** Keep tossing a coin independently, fix $r \in \mathbb{R}$
$X$ = number of tails until exactly r heads have appeared
Alternatively: Perform a trial whose outcomes are successes independently.
$X$ = number of failure until exactly r successes have appeared

We want to calculate the probability mass function of X.
Values of $\{X = x\} = \{0, 1, 2, 3, \cdots, x\}$
Example: If $r = 2$
$\{X = 0\} = \{$ All outcomes with 0 failures until 2 successes$\} = \{SSS\}$
$\{X = 0\} = \{$ All outcomes with 1 failures until 2 successes$\} = \{FSS, SFS\}$
Observe: All outcomes in the set $\{X = k\}$ is equally likely.

First, We find the probability of a single outcome in $\{X = k\}$
$\{X = k\} = \{$All outcomes with exactly k failures before the $r^{th}$ success$\}$
$\omega = FFF \cdots F_k SSS \cdots S_r$, $P(S) = p, P(F) = (1-p)$
$P(\omega) = p^k p^r$
So we have $k + r$ slots, and where $k + r$ will always be S. Then the remaining $(r-1)$ successes can happen in any of the remaining $(k + r - 1)$ slots.
So we only need to choose $(r-1)$ slots out of $(k + r - 1)$ to put the success, which is a total of

$$\binom{k+r-1}{r-1}$$

or, if we choose the failure seats, we will get

$$\binom{k+r-1}{k}$$

Therefore, we find that

$$|\{X = k\}| = \binom{k+r-1}{r-1} = \binom{k+r-1}{k} = \frac{(k+r-1)!}{(r-1)!k!}$$

So the probability mass function is

$$p_X(k) = \binom{k+r-1}{r-1} p^r \cdot (1-p)^k$$

Now, suppose $X$ NegBinom$(p, r)$, we want to calculate the **Expected Value** E(X)

$$\begin{aligned}
E(X) &= \sum_{x \in X} p_X(x) \\
&= \sum_{k=0}^{\infty} k \cdot p_X(k) \\
&= \sum_{k=0}^{\infty} k \cdot \binom{k+r-1}{k} \cdot p^r \cdot (1-p)^k \\
&= \sum_{k=0}^{\infty} k \cdot \frac{(k+r-1)!}{(r-1)!k!} \cdot p^r \cdot (1-p)^k \\
&= \sum_{k=1}^{\infty} \frac{(k+r-1)!}{(r-1)!(k-1)!} \cdot p^r \cdot (1-p)^k \\
&= \sum_{j=0}^{\infty} \frac{(j+1+r-1)!}{(r-1)!(j)!} \cdot p^r \cdot (1-p)^{j+1} && \text{Set} j = k \\
&= \sum_{j=0}^{\infty} r \cdot \frac{(j+(r+1)-1)!}{(r)!(j)!} \cdot \frac{p^{r+1}}{p} \cdot (1-p)^j \cdot (1-p) \\
&= \frac{r(1-p)}{p} \cdot \sum_{j=0}^{\infty} \frac{(j+(r+1)-1)!}{(r)!(j)!} \cdot p^{r+1} \cdot (1-p)^j \\
&= \frac{r(1-p)}{p} \text{ Note that above is the sum of all the possibilities assosicated to NegBinom}(r+1, p)
\end{aligned}$$

Similarly, we can find the **Variance** $V(X)$ to get

$$V(x) = \frac{r(1-p)}{p^2}$$

The **Moment Generating Function** is

$$M_x(t) = \left(\frac{p \cdot e^t}{1-(1-p)e^t}\right)^r$$

Exercise: Find the Moment Generating Function

## Poission Distribution

$$e^\lambda = 1 + e$$

**Definition 2.0.0.0.1**
*We say X has the **Poission Distrubution** with parameters*

$$\lambda \to rate$$

*if*

$$X = \{0, 1, 2, 3, \cdots\}$$
$$p_X(k) = e^{-\lambda}\frac{\lambda}{k!}$$

We can calculate the **Expected Value**

$$\begin{aligned} E(x) &= \sum_{k=0}^{\infty} k \cdot p_X(k) \\ &= \sum_{k=0}^{\infty} k \cdot e^{-\lambda} \cdot \frac{\lambda}{k!} \\ &= \sum_{k=0}^{\infty} e^{-\lambda} \cdot \frac{\lambda}{(k-1)!} \\ &= \lambda \sum_{k=0}^{\infty} e^{-\lambda} \cdot \frac{\lambda}{(k-1)!} \end{aligned}$$

Special Case of Binomial Distribution: $(n = 1)$ **Bernonli Ditsribution**
$x = \{0, 1\}$ $p_X(x) = \begin{cases} p & x = 1 \\ (1-p)^x = 0 \end{cases}$

## Distribution

# 3 | Continuous Variables

We say X is a continuous random variable if and only if $F_X$ is a continuous function.
Given a random variable X, there are multiple pdfs assosiated to the X.

**Definition 0.0.0.0.1**
*We say two random variables X, Y with cumulative distribution $F_X, F_Y$ respectively, are **identically distributed** if*

$$F_X(u) = F_Y(y) \forall u \in \mathbb{R}$$

## Uniform Continuous Distribution