STAT410 - 0101

Ash
ash@ash.lgbt

Last updated 2024-11-07.

Contents

Combinatorics

Generalize the Basic Principle of Counting

If r experiments are to be performed, and the experiment can result in n_1 many different ways and for each possible outcome (of n_1), there are n_2 different ways the second experiment can result, ..., for each possible outcome of n_{r-1} there are n_r different ways the rth experiment can result.

Then, the r experiments can end in $n_1 \times n_2 \times n_3 \times \cdots \times n_r$ many different ways.

Example

In a club, there are:

- 5 Freshman
- 10 Sophomores
- 5 Juniors
- 12 Seniors

A committee of 4 consisting of one person from each class is to be formed. How many different committees are possible?

Solution

The first experiment is choosing the freshman, the second experiment is choosing the sophomore, the third experiment is choosing the junior, and fourth experiment is choosing the senior.

$$n_1 = 5, n_2 = 10, n_3 = 5, n_4 = 12$$
$$\Rightarrow 5 \times 10 \times 5 \times 12 = 3000$$

Example

How many different 7-place plate numbers such that the first three places are letters and the remaining four are numbers?

Solution

Choose a letter, then a letter, then a letter, then a number, then a number, then a number, and finally a number:

$$26 \times 26 \times 26 \times 10 \times 10 \times 10 \times 10 = 175760000$$

Example

How many different 7-place plate numbers such that the first three places are letters and the remaining four are numbers, such that letters and numbers do not repeat?

Solution

Choose a letter, then a letter, then a letter, then a number, then a number, then a number, and finally a number, but remove one option each time:

$$26 \times 25 \times 24 \times 10 \times 9 \times 8 \times 7 = 78624000$$

Permutations

Each ordered arrangement is called a permutation.

There are $n \times (n-1) \times (n-2) \times \cdots \times 1 = n!$ many permutations for n objects.

Formational Example

Given $\{a, b, c\}$, how many different ways can you arrange it?

Solution

Give yourself 3 spaces:

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

Select any first space for a, then you have 2 spaces left over for b and 1 space left over for c.

$$3 \times 2 \times 1 = 6$$

The possibilities are: $abc, acb, bac, bca, cab, cba$

Example

In a class, there are 6 women and 4 men. They will be ranked according their scores. No two people get the same score, how many different rankings are possible?

Solution

$(4 + 6)! = 3628800$

Part B

If men are to be ranked among themselves, and women are ranked among themselves, how many possibilities are there?

Solution

For the first experiment, rank the men. There are $4!$ ways to do so.

For the second experiment, rank the women. There are $6!$ ways to do so.

Apply the basic principle of counting: there are $4! \times 6! = 17280$ possible results.

Example

A professor has

- 4 math books
- 3 chemistry
- 2 history
- 1 language

These books will be arranged such that the books on the same subject are together.

Solution

$$\underbrace{4!}_{\text{arrange the subjects}} \times \underbrace{4!}_{\text{arrange math books}} \times \underbrace{3!}_{\text{chemistry}} \times \underbrace{2!}_{\text{history}} \times \underbrace{1!}_{\text{language}} = 6912$$

Permutations with repeated elements

There are $\frac{n!}{n_{1!}n_{2!}\cdots n_{r!}}$ different permutations of n objects of which there is a group of n_1 alike, n_2 alike, ..., n_r alike.

Formational Example

How many different letter arrangements can be formed using the letters in the word “PEPPER”?

Solution

Treat them as individuals, and in that case, the answer is $6!$.

Then, you need to remove the number of repeated cases, so divide by the rearrangements of the 3 Ps, 2 Es, and 1 R.

$$\frac{6!}{3!2!1!} = 60$$

Example

In a chess competition, there are 10 competitors:

- 4 Russians
- 3 Brazilians
- 2 Englishmen
- 1 Grecian

The results only list the nationalities. In how many different results are possible:

Solution

$$\frac{10!}{4! \times 3! \times 2! \times 1!} = 12600$$

Example

You have a line of thread that you want to put 10 beads on:

- 3 red beads
- 5 green beads
- 2 yellow beads

Assuming the beads are indifferentiable, how many arrangments are possible?

Solution

$$\frac{10!}{3! \times 5! \times 2!} = 2520$$

Number of subsets

In general, there are

$$\frac{n(n-1)\cdots(n-r+1)}{r!} = \frac{n!}{(n-r)!r!} = \binom{n}{r}$$

ways to select r items from a collection of n.

Formational Example

Given $\{A, B, C, D, E\}$, how many different subsets of 3 are possible?

Solution

Choose 3 out of the 5 with order: $5 \times 4 \times 3$. But you have order, so A, B, C and A, C, B are counted as different. Therefore, we need to divide out the order:

$$\frac{5 \times 4 \times 3}{3!} = 10$$

Example

Given 20 individuals, in how many different ways can we form a group of 5?

Solution

$$\binom{20}{5} = 15504$$

Example

We have a group of 6 women and 8 men. How many different committees consisting of 2 women and 3 men are possible?

Solution

First experiment, choose the women, second experiment, choose the men:

$$\binom{6}{2} \times \binom{8}{3} = 840$$

Subexample

There are two men who refuse to work together. How many committees can we form now?

Solution

$$\binom{6}{2} \times \binom{8}{3} - \binom{6}{2} \times \binom{8-2}{3-2} = 750$$

Alternatively,

$$\binom{6}{2} \times \left(\binom{6}{2} \cdot 2 + \binom{6}{3}\right) = 750$$

Annenta problem

Given n antennas, m of which are defective, how many arrangments are there where no two defective antennas are next to each other?

Solution

$$\binom{n-m+1}{m}$$

Pascal's Identity

$$\binom{n}{r} = \binom{n-1}{r-1} + \binom{n-1}{r}$$

Proof

$$
\begin{aligned}
&\binom{n-1}{r-1} + \binom{n-1}{r} \\
&= \frac{(n-1)!}{(r-1)!(n-1-(r-1))!} + \frac{(n-1)!}{(r)!(n-1-r)!} \\
&= \frac{(n-1)!}{(r-1)!(n-1-r+1)!} + \frac{(n-1)!}{r!(n-1-r)!} \\
&= \frac{(n-1)!}{(r-1)!}\left(\frac{1}{(n-1-r+1)!} + \frac{1}{r(n-1-r)!}\right) \\
&= \frac{(n-1)!}{(r-1)!}\left(\frac{1}{(n-r)!} + \frac{1}{r(n-1-r)!}\right) \\
&= \frac{(n-1)!}{(r-1)!(n-1-r)!}\left(\frac{1}{n-r} + \frac{1}{r}\right) \\
&= \frac{(n-1)!}{(r-1)!(n-1-r)!}\left(\frac{r}{r(n-r)} + \frac{n-r}{(n-r)(r)}\right) \\
&= \frac{(n-1)!}{(r-1)!(n-1-r)!}\left(\frac{n}{r(n-r)}\right) \\
&= \frac{n(n-1)!}{r(r-1)!(n-r)(n-1-r)!} \\
&= \frac{n!}{r!(n-r)!} \\
&= \binom{n}{r}
\end{aligned}
$$

Alternative

Fix an object, and consider combinations including that object. This has $\binom{n-1}{r-1}$ combinations. Consider combinations that do not include that object. There are $\binom{n-1}{r}$ combinations for that situation.

Sum the two situations because they are mutually exclusive, leading to $\binom{n}{r} = \binom{n-1}{r-1} + \binom{n-1}{r}$.

Binomial Theorem

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

Proof

The terms will have the form $\binom{n}{k}x^k y^{n-k}$. This leads to the whole sum $\sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$.

$$(x+y)^n = (x_1 + y_1)(x_2 + y_2)\cdots(x_n + y_n)$$

The reason the terms have that form is that you have n choices, and then for a single term $x^k y^{n-k}$, you can make k decisions about where the x's come from, leading to $\binom{n}{k}$ options.

Example

How many different subsets of a set with n elements are there?

Solution

The number of subsets with:

- Subsets with 0 elements is $\binom{n}{0}$
- Subsets with 1 element is $\binom{n}{1}$
- Subsets with 2 elements is $\binom{n}{2}$
- $\vdots$
- Subsets with n elements is $\binom{n}{n}$

Therefore, the total number of subsets is $\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{n}$.

$$\begin{aligned}
&\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{n} \\
&= \binom{n}{0}(1)^0 1^n + \binom{n}{1}(1)^1 1^{n-1} + \cdots + \binom{n}{n} 1^n 1^0 \\
&= (1+1)^n \\
&= 2^n
\end{aligned}$$

Alternative

Choose for each element to include or not to include it. Therefore:

$$\underbrace{2 \times 2 \times 2 \times \cdots \times 2}_{n \text{ times}} = 2^n$$

Multinomial Coefficients

A set of n distinct objects is to be divided into r distinct groups of respective sizes $n_1, n_2, \ldots, n_r$.

How many different divisions are possible?

$$\frac{n!}{n_1! n_2! n_3! \cdots n_r!}$$

Proof

$$\begin{aligned}
&\binom{n}{n_1} \times \binom{n-n_1}{n_2} \times \binom{n-n_1-n_2}{n_3} \times \cdots \times \binom{n_r}{n_r} \\
&= \frac{n!}{(n_1)!(n-n_1)!} \times \frac{(n-n_1)!}{(n_2)!(n-n_1-n_2)!} \times \frac{(n-n_1-n_2)!}{(n_3)!(n-n_1-n_2-n_3)!} \times \cdots \times \frac{(n_r)!}{(n_r)!(n_r-n_r)!} \\
&= \frac{n!}{(n_1)!\cancel{(n-n_1)!}} \times \frac{\cancel{(n-n_1)!}}{(n_2)!\cancel{(n-n_1-n_2)!}} \times \frac{\cancel{(n-n_1-n_2)!}}{(n_3)!\cancel{(n-n_1-n_2-n_3)!}} \times \cdots \times \frac{\cancel{(n_r)!}}{(n_r)!(n_r-n_r)!} \\
&= \frac{n!}{n_1! n_2! n_3! \cdots n_r!} \\
&= \binom{n}{n_1, n_2, n_3, \ldots, n_r}
\end{aligned}$$

Example
10 players are to be divided into an A team and a B team which will play in different leagues. How many different divisions are possible?

$$\binom{10}{5, 5} = \frac{10!}{5!5!} = 252$$

Example
10 children are to be divided into two teams (each 5) to play a game. How many divisions are possible?

$$\binom{10}{5, 5} \times \frac{1}{2} = \frac{10!}{2 \times 5!5!} = 126$$

Example
In the first round of knockout tournaments involving $n = 2^m$ players, n players are divided into $\frac{n}{2}$ groups. The losers are eliminated, and the winner goes to the next round. The process continues until there is only one player left.

Solution
With $n = 2^3$, how many possibilities are there for the first round?

$$\binom{8}{2, 2, 2, 2} \cdot \frac{1}{4!} \cdot 2^4 = \frac{8!}{4!}$$

How many ways can the tournament end?

$$\frac{8!}{4!} \cdot \frac{4!}{2!} \cdot 2 = 8!$$

Stars and Bars (1.6)

$$x_1 + x_2 + \cdots + x_r = n$$

Trying to find $(x_1, x_2, \cdots, x_r)$, all positive integers.

Create n dots:

$$\bullet\bullet\bullet\cdots\bullet$$

Create $r - 1$ seperators:

$$\underbrace{\bullet}_{x_1} \wedge \underbrace{\bullet\bullet}_{x_2} \wedge \cdots \wedge \underbrace{\bullet}_{x_r}$$

Therefore, there are $\binom{n-1}{r-1}$ possible solutions to the equation since there is a 1-1 correspondence to the arrangements.

What about the number of nonnegative integer solutions?

To solve the equation:

$$y_1 + y_2 + y_3 + \cdots + y_r = n$$

To transform it into the previous problem, add one to each input to make it positive:

$$(y_1 + 1) + (y_2 + 1) + \cdots + (y_r + 1) = n + r$$

So now "n" $= n + r$, leading to the solution:

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

Example
Given n antennas, m of which are defective, how many arrangments are there where no two defective antennas are next to each other?

Solution
Put down the defective antennas in some order:

▼▼▼▼▼▼

Between every two defective antennas, you have to put down at least 1 one antenna:

△▼△▼△▼△▼△▼△▼△

The ones on the ends are optional.

This is then the same as selecting solutions to

$$\underbrace{x_1 + x_2 + \cdots + x_{m-1}}_{\text{between}} + \underbrace{(x_m + 1) + (x_{m+1} + 1)}_{\text{edges}} = n - m + 2$$

This then leads to the solution:

$$\binom{n - m + 2 - 1}{m + 1 - 1} = \binom{n - m + 1}{m}$$

Which is the same as what was previously achieved.

Experiments

The set of all possible outcomes of ann experiment is called the sample space, denoted S.

A subset E of a sample space is called an event. If the outcome belongs to E, then we say that E has occurred.

Formational Experiment: rolling d6
Roll a die.

The sample space of the output is:

$$S = \{1, 2, 3, 4, 5, 6\}$$

Assume you want even numbers:

$$E = \{2, 4, 6\}$$

E occurs if the outcome of the experiment is $\in E$, IOW if it is 2, 4 or 6.

Example: childbirth
Childbirth: the outcome is determined by the sex of the child.

Then $S = \{\text{boy}, \text{girl}, \text{intersex}\}$.

Define $E = \{\text{boy}\}$. Then E occurs if you have a boy.

Example: 7 horse race
$H = \{h_1, h_2, ..., h_7\}$. 7 horses are in a race: the outcome is determined by the finishing order of the horses.

$S = \{\text{all permutations of } \ H\}$

Example: flipping 2 coins
The experiment consists of flipping 2 coins:

$$S = \{(h, h), (h, t), (t, h), (t, t)\}$$

At least one head:

$$E_1 = \{(h, h), (h, t), (t, h)\}$$

At least one tail:

$$E_2 = \{(h, t), (t, h), (t, t)\}$$

Example: Lifetime of a transistor

$$S = \{x \mid x \geq 0\}$$

The transistor lasts less than 4 units of time:

$$E = \{x \mid 0 \leq x \leq 4\}$$

Union

Let E_1, E_2 be two events associated with a sample space S.

The union $E_1 \cup E_2$ is the event consisting of all outcomes that are in E_1, E_2 or both.

In other words, $E_1 \cup E_2$ occurs if either E_1 occurs, E_2 occurs or both occur.

Example

$$S = \{(h, t), (t, h), (h, h), (t, t)\}$$

Let E_1 be the set of at least one head, and let E_2 be the set of at least one tail.

Then, this means that $E_1 \cup E_2 = \{(h, t), (t, h), (h, h), (t, t)\} = S$

Intersection

Given E_1, E_2 their intersection $E_1 E_2$ ($E_1 \cap E_2$) consists of all outcomes belonging to both E_1 and E_2.

Example
Use the coin example again:

$$E_1E_2 = \{(h,t),(t,h)\} = \text{exactly one head and one tail}$$

Complement

Given E in S, the complement event is denoted by $E^\complement$ or E' and consists of all outcomes in S but not in S.

In other words, $E^\complement$ occurs if E does not occur, and if E occurs, $E^\complement$ does not occur.

Null event

$\varnothing$ denotes the null event and never occurs.

2.3

Mutually Exclusive Events

Events $E_1, E_2, ..., E_r$ are mutually exclusive iff:

$$\forall i, j \in \mathbb{Z} \cap [1, r], i \neq j \to E_iE_j = \varnothing$$

Probability

For each event E in S we assume that a value $P(E)$ is defined such that:

1. $0 \leq P(E) \leq 1$
2. $P(S) = 1$
3. For any set of mutually exclusive events $E_1, E_2, ..., E_r$:

$$P\left(\bigcup_{i=1}^{r} E_i\right) = \sum_{i=1}^{r} P(E_i)$$

Formational Example: Relative Frequency

Assume E is an event, and an experiment defined by the sample space S is repeated n times.

Let $n(E)$ denote the number of times E occurs.

Then the relative frequency of E is $\frac{n(E)}{n}$.

The probability of E is then $\lim_{n\to\infty} \frac{n(E)}{n}$.

But how do we know that this limit exists?

Conjugate

Given E and $E^\complement$:

$$P(E^\complement) = 1 - P(E)$$

Proof

E and $E^\complement$ are mutually exclusive.

By axiom 3, $P(E \cup E^\complement) = P(E) + P(E^\complement)$.

By axiom 1, $P(E \cup E^\complement) = P(S) = 1$.

Therefore:

$$P(E) + P(E^{\complement}) = 1$$
$$P(E^{\complement}) = 1 - P(E)$$

Sample Spaces with equally likely outcomes

Assume a sample space S consists of equally likely outcomes. Then,

$$P(E) = \frac{\#\text{ outcomes belonging to }\ E}{\#\text{ outcomes in }\ S}$$

Formational Example: Probabibility of a Die

Considering a 6-sided die such that all sides are equally likely, each side has probability $\frac{1}{6}$

Proof

Consider a 6-sided die such that all sides are equally likely.

Consider the mutally exclusive events, $E_1 = \{1\}, E_2 = \{2\}, ..., E_6 = \{6\}$.

By the axioms, $P(S) = 1$

$$S = E_1 \cup E_2 \cup \dots \cup E_6$$
$$1 = P(S) = P(E_1 \cup E_2 \cup \dots \cup E_6) = P(E_1) + P(E_2) + \dots + P(E_6)$$
$$1 = 6P(E_1)$$
$$\frac{1}{6} = P(E_1)$$

Therefore, for example, $P(\{2,4,6\}) = P(E_2) + P(E_4) + P(E_6) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6} = \frac{1}{2}$.

Example

If 3 balls are randomly drawn from a bowl containing 6 white and 5 black balls, compute the probability that 1 is white and 2 are black.

Solution

$E =$ 1 white, 2 black.

The number of events in S is $\binom{5+6}{3}$. The number of events in E is $\binom{6}{1}\binom{5}{2}$.

Therefore:

$$P(E) = \frac{\binom{6}{1}\binom{5}{2}}{\binom{5+6}{3}} = \frac{60}{165} = \frac{4}{11}$$

Alternatively, the number of events in S is $\frac{11!}{(11-3)!}$, and there are $\underbrace{6}_{\text{white}} \times \underbrace{5 \times 4}_{\text{black}} \times \underbrace{3}_{\text{move white around}}$ events in E.

Therefore:

$$P(E) = \frac{6 \times 5 \times 4 \times 3}{\frac{11!}{(11-3)!}} = \frac{360}{990} = \frac{4}{11}$$

Probability of the union of two events

$$P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1E_2)$$

Proof

$$\begin{aligned} \text{I } &= E_1 \setminus (E_1E_2) \\ \text{II} &= E_1E_2 \\ \text{III} &= E_2 \setminus (E_1E_2) \end{aligned}$$

We know $E_1 \cup E_2 = \text{I } \cup \text{II} \cup \text{III}$ and that the events are mutually exclusive.

$$\begin{aligned} P(E_1) &= P(E_1 \setminus (E_1E_2)) + P(E_1E_2) \\ P(E_1) - P(E_1E_2) &= P(E_1 \setminus (E_1E_2)) \\ P(E_1) - P(E_1E_2) &= P(I) \end{aligned}$$

$$\begin{aligned} P(E_2) &= P(E_2 \setminus (E_1E_2)) + P(E_1E_2) \\ P(E_2) - P(E_1E_2) &= P(E_2 \setminus (E_1E_2)) \\ P(E_2) - P(E_1E_2) &= P(\text{III}) \end{aligned}$$

$$\begin{aligned} P(E_1 \cup E_2) &= P(\text{I }) + P(\text{II}) + P(\text{III}) \\ &= (P(E_1) - P(E_1E_2)) + (P(E_1E_2)) + (P(E_2) - P(E_1E_2)) \\ &= P(E_1) + P(E_2) - P(E_1E_2) \end{aligned}$$

DeMorgan's Laws

$$\left(\bigcup_{i=1}^{n} E_i\right)^{\mathsf{C}} = \bigcap_{i=1}^{n} E_i^{\mathsf{C}}$$

$$\left(\bigcap_{i=1}^{n} E_i\right)^{\mathsf{C}} = \bigcup_{i=1}^{n} E_i^{\mathsf{C}}$$

Example

Jude is taking two books along on her vacation.

She likes the first book with a probability of 0.4, the second with a 0.2 probability and she likes both with a 0.1 probability.

Compute the probability she likes neither.

Solution

Probability of the union of three events

$$P(E_1 \cup E_2 \cup E_3) = P(E_1) + P(E_2) + P(E_3) - P(E_1E_2) - P(E_1E_3) - P(E_2E_3) + P(E_1E_2E_3)$$

Proof

$$
\begin{aligned}
P(E_1 \cup E_2 \cup E_3) &= P(E_1 \cup E_2) + P(E_3) - P((E_1 \cup E_2) \cap E_3) \\
&= P(E_1) + P(E_2) - P(E_1E_2) + P(E_3) - P(E_1E_3 \cup E_2E_3) \\
&= P(E_1) + P(E_2) + P(E_3) - P(E_1E_2) - (P(E_1E_3) + P(E_2E_3) - P(E_1E_3E_2E_3)) \\
&= P(E_1) + P(E_2) + P(E_3) - P(E_1E_2) - (P(E_1E_3) + P(E_2E_3) - P(E_1E_2E_3)) \\
&= P(E_1) + P(E_2) + P(E_3) - P(E_1E_2) - P(E_1E_3) - P(E_2E_3) + P(E_1E_2E_3)
\end{aligned}
$$

Inclusion-Exclusion Principle

$$
P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{i=1}^{n} P(E_i) - \sum_{i<j} P\left(E_iE_j\right) + \sum_{i<j<k} P\left(E_iE_jE_k\right) + \cdots + (-1)^{n+1} P(E_1E_2 \cdots E_n)
$$

Example: Matching Problem

Suppose that each of the N men at a party throws his hat into the center of the room. Then, the hats are mixed up, and each man randomly selects a hat. What is the probability that no man selects his own hat?

Solution

Let E_i, $1 \leq i \leq N$ denote the event where the ith person selects his own hat.

Then, the probability that no man selects his own hat is:

$$
\begin{aligned}
& P\left(E_1^{\mathsf{C}} \cap E_2^{\mathsf{C}} \cap E_3^{\mathsf{C}} \cup \cdots \cup E_N^{\mathsf{C}}\right) \\
&= P\left((E_1 \cup E_2 \cup \cdots \cup E_N)^{\mathsf{C}}\right) \\
&= 1 - P(E_1 \cup E_2 \cup \cdots \cup E_N)
\end{aligned}
$$

$$
\begin{aligned}
P\left(\bigcup_{i=1}^{N} E_i\right) = \\
& \sum_{i=1}^{N} P(E_i) \\
& - \sum_{i_1<i_2} P\left(E_{i_1}E_{i_2}\right) \\
& + \sum_{i_1<i_2<i_3} P\left(E_{i_1}E_{i_2}E_{i_3}\right) \\
& \vdots \\
& +(-1)^{k+1} \sum_{i_1<i_2<\cdots<i_k} P\left(E_{i_1}E_{i_2} \cdots E_{i_k}\right) \\
& \vdots \\
& +(-1)^{N+1} P(E_1E_2 \cdots E_N)
\end{aligned}
$$

Define probabilities:

$$P(E_i) = \frac{(N-1)!}{N!}$$

$$P\left(E_i E_j\right) = \frac{(N-2)!}{N!}$$

$$\vdots$$

$$P\left(E_{i_1} E_{i_2} \cdots E_k\right) = \frac{(N-k)!}{N!}$$

Use them:

$$\begin{aligned}
&\sum_{i=1}^{N} P(E_i) \\
&- \sum_{i_1 < i_2} P\left(E_{i_1} E_{i_2}\right) \\
&+ \sum_{i_1 < i_2 < i_3} P\left(E_{i_1} E_{i_2} E_{i_3}\right) \\
&\vdots \\
&+(-1)^{k+1} \sum_{i_1 < i_2 < \cdots < i_k} P\left(E_{i_1} E_{i_2} \cdots E_{i_k}\right) \\
&\vdots \\
&+(-1)^{N+1} P(E_1 E_2 \cdots E_N) \\
&= \binom{N}{1} \frac{(N-1)!}{N!} - \binom{N}{2} \frac{(N-2)!}{N!} + \binom{N}{3} \frac{(N-3)!}{N!} + \cdots \\
&+ (-1)^{k+1} \binom{N}{k} \frac{(N-k)!}{N!} + \cdots \\
&+ (-1)^{N+1} \binom{N}{N} \frac{(N-N)!}{N!}
\end{aligned}$$

Simplify:

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

Therefore, the probability nobody gets their hat is:

$$\begin{aligned}
&1 - \sum_{k=1}^{N} \frac{(-1)^{k+1}}{k!} \\
&= \sum_{k=0}^{N} \frac{(-1)^k}{k!}
\end{aligned}$$

Furthermore,

$$\lim_{N\to\infty}\sum_{k=0}^{N}\frac{(-1)^k}{k!}=\frac{1}{e}\approx 0.367879$$

Example: Straight

$$P(\text{straight})=\frac{10\times(4^5-4)}{\binom{52}{5}}$$

Overlapping Birthdays

Assume there are n people, with 365 possible birthdays. Compute the probability that they have different birthdays.

$$\frac{365\times 364\times\ldots\times(365-n-1)}{365^n}=\frac{\frac{365!}{(365-n)!}}{365^n}$$

Conditional Probability

Conditional Probability

$P(E|F)$ denotes the conditional probability that E occurs given F has occured.

Assume $P(F)>0$

$$P(E|F)=\frac{P(EF)}{P(F)}$$

Formational Example

Two dice are rolled. Assume that the first die shows a 3. Compute the probability that the sum of the two dice is 7.

Solution

The second dice must be a 4. Therefore, $P=\frac{1}{6}$.

Alternatively:

$$S=\{(1,1),(1,2),\ldots,(6,5),(6,6)\}$$
$$S'=\{(3,1),(3,2),(3,3),(3,4),(3,5),(3,6)\}$$

With the sums:

$$S'_{\sum}=\{4,5,6,7,8,9\}$$

Leading to a $\frac{1}{6}$ probability.

Example

Joe is 80% certain that his missing key is in one of two pockets. He is 40% sure it is in the left pocket and 40% sure it is in the right pocket.

Given he checked his left pocket, and there was no key, what is the probability it was in the right pocket?

Solution

Let L be the event the key is in the left pocket and R for the right pocket.

$$P(R|L^{C}) = \frac{P(R \cap L^{C})}{P(L^{C})} = \frac{0.4}{1-0.4} = \frac{0.4}{0.6} = \frac{4}{6} = \frac{2}{3} = 0.666...$$

Conditional Probability for Equally Likely Sets

$$P(E|F) = \frac{|E \cap F|}{|F|}$$

Example

A coin is flipped twice. All four outcomes are equally likely. What is the conditional probability that you get two heads given:

1. The first coin is heads:

$$S = \{((h,h)), (h,t), (t,h), (t,t)\}$$
$$F = \{(h,h), (h,t)\}$$
$$E = \{(h,h)\}$$

$$\frac{1}{2}$$

2. At least one coin is heads:

$$S = \{((h,h)), (h,t), (t,h), (t,t)\}$$
$$F = \{(h,h), (h,t), (t,h)\}$$
$$E = \{(h,h)\}$$

$$\frac{1}{3}$$

Multiplication of Probabilities

$$P(EF) = P(F) \cdot P(E|F)$$

This can be generalized to:

$$P(E_1E_2E_3) = P(E_3|E_2E_1)P(E_2 \mid E_1)P(E_1)$$
$$\vdots$$
$$P(E_1E_2\cdots E_n) = P(E_n \mid E_1E_2\cdots E_{n-1})\cdots P(E_2 \mid E_1)P(E_1)$$

Example

An urn contains 8 red and 4 white balls. We draw 2 balls out without replacement. Compute the probability that

1. Both are white

$$P(E_1 \mid E_2) = P(E_2 \mid E_1)P(E_1) = \left(\tfrac{3}{11}\right) \cdot \left(\tfrac{4}{12}\right) = \tfrac{1}{11}$$

Hat Matching Problem

We have computed the probability that nobody gets their own hat: $\sum_{k=0}^{N} \frac{(-1)^k}{k!}$.

Compute the probability that exactly k of the N men pick his own hat.

Fix a particular order of people: $\{1, 2, 3, ..., k, k+1, ..., N\}$.

$$P(\text{set of k people take their own hat}) = P(\text{remaining take someone else's} \mid \text{the fixed set picked their own})$$

Furthermore, we can create a recurrence:

$$F(k, N) = \frac{1}{N} F(k-1, N-1)$$
$$F(0, N) = \sum_{k=0}^{N} \frac{(-1)^k}{k!}$$

Resulting in:

$$\begin{aligned} &\frac{1}{N} \cdot \frac{1}{N-1} \cdot \frac{1}{N-2} \cdots \frac{1}{N-(k-1)} \cdot \sum_{i=0}^{N-k} \frac{(-1)^i}{i!} \\ &= \frac{(N-k)!}{N!} \sum_{i=0}^{N-k} \frac{(-1)^i}{i!} \\ &\Rightarrow \frac{1}{k!} \sum_{i=0}^{N-k} \frac{(-1)^i}{i!} \end{aligned}$$

Something

Take two sets, E and F.

$E = EF \cup EF^{\mathsf{C}}$. This is a disjoint union.

$$\begin{aligned} P(E) &= P(EF) + P(EF^{\mathsf{C}}) \\ &= P(E \mid F)P(F) + P(E \mid F^{\mathsf{C}})P(F^{\mathsf{C}}) \end{aligned}$$

Example

There are two types of people, people who are accident-prone, and those who are not accident-prone.

If someone is accident-prone, they will have a 0.4 probability of having an accident. If someone is not accident-prone, the probability of having an accident is 0.2. Assume 30% of the society is accident-prone. Compute the probability that a new customer will have an accident.

$$
\begin{aligned}
P(A_1) &= P(A_1 \cap \text{accident-prone}) \cup P(A_1 \cap \text{not accident-prone}) \\
&= P(A_1 \cap \text{accident-prone}) + P(A_1 \cap \text{not accident-prone}) \\
&= P(A_1 \mid \text{accident-prone}) + P(A_1 \mid \text{not accident-prone}) \\
&= P(A_1 \mid \text{accident-prone})P(\text{accident prone}) + P(A_1 \mid \text{not accident-prone})P(\text{not accident-prone}) \\
&= (0.4)(0.3) + (0.2)(0.7) \\
&= 0.26
\end{aligned}
$$

Given a customer had an accident, what is the probability that he/she is an accident-prone person?

$$
P(\text{accident-prone} \mid A_1) = \frac{P(A_1 \cap \text{accident prone})}{P(A_1)} = \frac{0.12}{0.26}
$$

Probability when breaking down set

$F_1, F_2, ..., F_n$ are disjoint and $\bigcup_{i=1}^{n} F_i = S$.

$$
\begin{aligned}
E &= EF_1 \cup EF_2 \cup \cdots \cup EF_n \\
P(E) &= P(EF_1) + P(EF_2) + \cdots + P(EF_n) \\
&= P(E|F_1)P(F_1) + P(E|F_2)P(F_2) + \cdots + P(E|F_n)P(F_n)
\end{aligned}
$$

Example

There are three types of flashlights in a bin, I, II and III.

The probability that a type I flashlight will give more than 100 hours of light is 0.4, for type II this is 0.4 and for type III this is 0.3.

Suppose 20% of the flashlights are in the bin are type I, 30% are type II and 50% are type III.

$$
P(E) = P(E|T_1)P(T_1) + P(E|T_2)P(T_2) + P(E|T_3)P(T_3)
$$

Independent Events

Two events E, F are independent if

$$
P(EF) = P(E|F)P(F) = P(E)P(F)
$$

$P(E|F) = P(E)$.

Example

Two coins are flipped. E is the event the first coin lands on heads. F is the event that the second coin lands on tails.

Check if E, F are independent.

$EF = \{(H, T)\}$

$E = \{(H, H), (H, T)\}$ $F = \{(H, T), (T, T)\}$

$$P(EF) \stackrel{?}{=} P(E)P(F)$$
$$\frac{1}{4} \stackrel{?}{=} \left(\frac{1}{2}\right)\left(\frac{1}{2}\right)$$
$$\frac{1}{4} = \frac{1}{4}$$

Yes

E, F are independent.

Example
You are rolling two dice. Let E be the event that the first one shows 4, and F be the event that the sum is 4.

$P(E) = \frac{1}{6}$,$P(F) = \frac{\binom{4-1}{2-1}}{6^2}$. $P(EF) = 0$

$$P(EF) \stackrel{?}{=} P(E)P(F)$$
$$\frac{1}{6} \cdot \frac{3}{6^2} \stackrel{?}{=} 0$$
$$\frac{3}{6^3} \neq 0$$

Not independent events.

Independence of n events

$$P(EF) = P(E)P(F)P(EG) = P(E)P(G)P(FG) = P(F)P(G)P(EFG) = P(E)P(F)P(G)$$

n events $E_1, E_2, ..., E_n$ are independent if for any subset $\left\{E_{i_1}, E_{i_2}, ..., E_{i_k}\right\}$:

$$P\left(E_{i_1}E_{i_2} \cdots E_{i_k}\right) = P\left(E_{i_1}\right)P\left(E_{i_2}\right) \cdots P\left(E_{i_k}\right)$$

Example
An infinite sequence of independent trials will be performed. Each trial ends with a success with a probability p. There will be a failure with probability $1 - p$.

1. At least one success occurs in the first n trails.

$$P(\text{at least one success}) = 1 - P(\text{all n trails are failures})$$

Let E_i be the event that the trial i is a failure. $P(\text{all n trails are failures}) = P\left(\bigcap_{i=1}^{n} E_i\right)$.

$$P\left(\bigcap_{i=1}^{n} E_i\right) = P(E_1)P(E_2) \cdots P(E_n)$$
$$\Rightarrow 1 - P(E_1)P(E_2) \cdots P(E_n)$$
$$= 1 - (1-p)(1-p) \cdots (1-p) = 1 - (1-p)^n$$

Example
$BB \to$ brown eyes, $Bb \to$ brown eyes, $bb \to$ blue eyes.

When you get your genes, they are independently selected from both your parents.

Smith's parents both have brown eyes, and his sister has blue eyes. He also has brown eyes. This implies that both his parents have Bb genes.

Compute the probability that he possesses the blue-eyed gene.

	B	b
B	BB	bB
b	Bb	bb

Therefore, $\frac{2}{3}$ is the probability.

If Smith's first child has brown eyes, what is the probability that his second child has brown eyes as well? He has a wife with blue eyes.

$$P(\text{second child has brown eyes}|\text{first child has brown eyes}) = \frac{P(\text{first and second child have brown eyes})}{P(\text{first child has brown eyes})}$$

$$= \frac{P(\text{1st \& 2nd brown eyes} \cap \text{Smith blue-eyed gene}) + P(\text{1st \& 2nd have brown eyes} \cap \text{Smith does not have blue-eyed gene})}{P(\text{1st has brown eyes} \cap \text{smith has blue eyed gene}) + P(\text{1st has brown eyes} \cap \text{Smith does not have blue-eyed gene})}$$

$$= \frac{\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{2}{3} + 1 \cdot 1 \cdot \frac{1}{3}}{\frac{1}{2} \cdot \frac{2}{3} + 1 \cdot \frac{1}{3}} = \frac{3}{4}$$

	B	b
b	Bb	bb
b	Bb	bb

Random Variable

Real valued functions defined on the sample space are called random variables.

Formational Example

Take the sum of numbers observed when you toss two dice together.

$(1, 1) \to 2$, so $f(x, y) = x + y$ is the random variable.

This has the sample space $\{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$.

Example

Suppose 3 coins are tossed together. The random variable y is defined as the number of heads observed in the experiment.

$(H, T, H) \to 2$
$(T, T, T) \to 0$

Find $P(y = 0) = \left(\frac{1}{2}\right) \cdot \left(\frac{1}{2}\right) \cdot \left(\frac{1}{2}\right)$.

Find $P(y = 1) = \frac{3!}{2!} \cdot \left(\frac{1}{2}\right) \cdot \left(\frac{1}{2}\right) \cdot \left(\frac{1}{2}\right)$

Find $P(y = 2) = \frac{3!}{2!} \cdot \left(\frac{1}{2}\right) \cdot \left(\frac{1}{2}\right) \cdot \left(\frac{1}{2}\right)$

Find $P(y=3)=\left(\frac{1}{2}\right)\cdot\left(\frac{1}{2}\right)\cdot\left(\frac{1}{2}\right)$

Example

Let O be the event that the older client dies in the following year. $P(O)=0.1$. Let Y be the event that the younger client dies, $P(Y)=0.05$. O, Y are independent events. For each death $100k is paid to the beneficiaries. X is the total amount of money paid by the agent to the beneficiaries of these two clients in the following year, in units of $100k.

$X \in \{0,1,2\}$.

$P(X=0)=P(O'Y')=(1-0.1)\cdot(1-0.05)=0.855$

$P(X=1)=P(O'Y)+P(Y'O)=(1-0.1)(0.05)+(0.1)(1-0.05)=0.14$

$P(X=2)=P(OY)=0.1\cdot 0.05=0.005$

Example

Experiments draw 4 random balls from a bowl containing all balls numbered from 1 to 20. Let X be the largest of the 4 numbers obtained.

$$X \in \{4,5,6,...,19,20\}$$

$$P(X=i)=\frac{\binom{i-1}{3}}{\binom{20}{4}}$$

$$P(X>10)=1-P(X\leq 10)=1-\frac{\binom{10}{4}}{\binom{20}{4}}$$

Example

A coin (with probability p of landing on heads). It is flipped continuously until a head occurs or the maximum number of flips n has occurred.

Let X be the number of flips in the experiment.

$$X \in \{1,2,...,n\}$$

$$P(X=1)=p$$
$$P(X=2)=(1-p)p$$
$$k<n, P(X=k)=(1-p)^{k-1}p$$
$$P(X=n)=(1-p)^n+(1-p)^{n-1}p=(1-p)^{n-1}(1-p+p)=(1-p)^{n-1}$$

Show the sum of the options is 1.

$$P(X<n)=\sum_{k=1}^{n-1}(1-p)^{k-1}p=1-(1-p)^{n-1}$$
$$P(X\leq n)=P(X<n)+P(X=n)=1-(1-p)^{n-1}+(1-p)^{n-1}=1$$

Alternative done in class:

$$P(X \leq n) = p\sum_{k=0}^{n-1}(1-p)^k + (1-p)^n = p\left(\frac{1-(1-p)^n}{1-(1-p)}\right) + (1-p)^n$$
$$= \not{p}\left(\frac{1-(1-p)^n}{\not{p}}\right) + (1-p)^n$$
$$= 1-(1-p)^n + (1-p)^n = 1$$

Discrete Random Variable

A random variable that can take on at most a countable number of possible values is called discrete.

For each random variable, we define a probability mass function $p(a) = P(X = a)$.

Let $X \in \{x_1, ..., x_n\}$. Then $\sum_{i=1}^{n} p(x_i) = 1$.

Example

Let X be a discrete random variable, where λ is positive and fixed.

$$p(i) = c\frac{\lambda^i}{i!}$$

$$\sum_{i=0}^{\infty} p(i) = c\sum_{i=0}^{\infty}\frac{\lambda^i}{i!} = ce^{\lambda}$$

Since the sum should equal 1, $c = e^{-\lambda}$.

$$p(0) = e^{-\lambda}$$

$$P(X \leq 2) = p(0) + p(1) + p(2) = e^{-\lambda}\left(1 + \lambda + \frac{\lambda^2}{2}\right)$$

Expected Value

If X is a discrete random variable with a probability mass function $p(x)$, the expected value of X denoted $E[X]$ is defined by:

$$E[X] = \sum xp(x)$$

$E[x]$ is a weighted average of the possible values for X.

Example

A 6 sided die is rolled. Let X be the outcome of the experiment.

$$E[X] = 1\cdot\frac{1}{6} + 2\cdot\frac{1}{6} + 3\cdot\frac{1}{6} + 4\cdot\frac{1}{6} + 5\cdot\frac{1}{6} + 6\cdot\frac{1}{6} = \frac{7}{2}$$

Example

Let A be an event associated to an experiment.

Let

$$I = \begin{cases} 1 \text{ if } \ A \text{ occurs} \\ 0 \text{ if } A \text{ does not occur} \end{cases}$$

Then:

$$E[I] = 1 \cdot p(A) + 0(1 - p(A)) = p(A)$$

Example
On bus 1 there are 36 students, on bus 2 there are 40 students, and on bus 3 there are 44 students.

Let X be the number of students in the bus of a randomly selected student.

$$E[X] = 36 \cdot \frac{36}{120} + 40 \cdot \frac{40}{120} + 44 \cdot \frac{44}{120} = \frac{604}{15} \approx 40.2667 > 40$$

Example
The expected value for a function of a random variable X, $p(X)$,

If X is discrete, that takes on one of the values $x_i, i \in \mathbb{N}$ with respective probability $p(x_i)$, then for a function g,

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

Example
Assume X is a random variable with possible values $-1, 0, 1$, with probabilities $0.2, 0.5, 0.3$.

Compute $E[X^2]$.

$$E[X^2] = (-1)^2 \cdot 0.2 + (0)^2 \cdot 0.5 + (1)^2 \cdot 0.3 = 0.5$$

Furthermore, $E[X^2] \geq (E[X])^2$ for any random variable (by triangle inequality presumably).

Corollary about the expected value of a linear function of a random variable

$$E[aX + b] = \sum (ax + b) \cdot p(x) = \sum axp(x) + \sum bp(x) = a \sum xp(x) + b \sum p(x) = aE[X] + b \cdot 1 = aE[x] + b$$
$$\therefore E[aX + b] = aE[x] + b$$

This result could also be achieved by the linearity of the sum operator.

Example
Assume X is the number of guests to attend a party. Let Y be the cost of the party. Then, $Y = 10X + 20$ if for each person the party costs 10 and there is a constant 20 spend on decoration.

If $E[X] = 5$ (you expect 5 people to go to your party), then $E[Y] = 10 \cdot 5 + 20 = 70$.

Shortcut

The shortcut to computing the probability that a sum of 5 occurs before a sum of 7.

Let E be the event that a 5 occurs before a sum of 7.

$$P(E) = \frac{4}{36} + 0 \cdot P(E) + P(E) \cdot \left(1 - \frac{10}{36}\right)$$
$$\Rightarrow P(E) = \frac{2}{5}$$

Variance

Let X be a random variable where $E[X] = \mu$. Then, the variance is defined as $E[(X - \mu)^2]$.

Alternative Form

Assume X has probability mass function $p(x)$.

$$\begin{aligned}
\mathrm{Var}(X) = E[(X-\mu)^2] &= E(X^2 - 2\mu X + \mu^2) \\
&= \sum_x (x^2 - 2\mu x + \mu^2)p(x) \\
&= \sum_x (x^2)p(x) - \sum_x 2\mu x p(x) + \sum_x \mu^2 p(x) \\
&= \sum_x (x^2)p(x) - 2\mu \sum_x x p(x) + \mu^2 \sum_x p(x) \\
&= E[X^2] - 2\mu\mu + \mu^2(1) \\
&= E[X^2] - 2\mu^2 + \mu^2 \\
&= E[X^2] - \mu^2
\end{aligned}$$

Therefore $E[X^2] \geq (E[X])^2$

Example

What is the variance of the value a die lands on?

$$\begin{aligned}
E[X] &= 3.5 \\
E[X^2] &= 1p(1) + 2^2p(2) + 3^2p(3) + 4^2p(4) + 5^2p(5) + 6^2p(6) \\
&= \frac{1}{6} + \frac{4}{6} + \frac{9}{6} + \frac{16}{6} + \frac{25}{6} + 6 = \frac{91}{6}
\end{aligned}$$

$$\mathrm{Var}(x) = E[X^2] - E[X]^2 \Rightarrow \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12} = 2.91\overline{6}$$

Example

Suppose there are m days in a year, and each person is independently born a day r with probability p_r. This then implies that $\sum_{r=1}^{m} p_r = 1$.

Let $A_{i,j}$ be the event that person i and person j are born on the same day.

1. Find $P(A_{1,3})$.

$$= p_1 p_1 + p_2 p_3 + \cdots + p_m p_m = \sum_r^m p_r^2$$

2. Find $P(A_{1,3} | A_{1,2})$.

$$= \frac{\sum_r^m p_r^3}{\sum_r^m p_r^2}$$

3. Show $P\left(A_{1,3}|A_{1,3}\right) \geq P\left(A_{1,3}\right)$.

   Define $X = p_r$ with probability p_r.

   Then, $\sum_r^m p_r^3 = E[X^2]$ and $\sum_r^m p_r^2 = E[X]$.

   So then, $\frac{E[X^2]}{E[X]} \overset{?}{\geq} E[X] \Leftrightarrow E[X^2] \geq E[X]^2$ which is proven.

Bernoulli & Binomial Random Variables

For Bernoulli (1 trial binomal):

$$X = \begin{cases} 1 & \text{with probability } p \\ 0 & \text{with probability } 1-p \end{cases}$$

For Binomial, with n independent trials with each having probability of success p:

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

Then then creates the expected value:

$$\begin{aligned} E[X] &= \sum_{i=0}^{n} i \binom{n}{i} p^i (1-p)^{n-i} \\ &= \sum_{i=0}^{n} i \frac{n!}{i!(n-i)!} p^i (1-p)^{n-i} \\ &= \sum_{i=0}^{n} i \frac{n(n-1)!}{i(i-1)!(n-i)!} p^i (1-p)^{n-i} \\ &= \sum_{i=0}^{n} i \frac{n(n-1)!}{i(i-1)!((n-1)-(i-1))!} p^i (1-p)^{n-i} \\ &= \sum_{i=0}^{n} i \frac{n(n-1)!}{i(i-1)!((n-1)-(i-1))!} p p^{i-1} (1-p)^{(n-1)-(i-1)} \\ &= np \sum_{i=0}^{n-1} \frac{(n-1)!}{j!((n-1)-j)!} p^j (1-p)^{(n-1)-j} \\ &= np \underbrace{\quad \sum_{i=0}^{n-1} \binom{n-1}{j} p^j (1-p)^{(n-1)-j}}_{\text{sum of probabilities for a binomial distribution}} \\ &= np \end{aligned}$$

Show the sum of probabilities for the binomial distribution is actually 1:

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

Variance:

$$\begin{aligned} \text{Var } X &= E[X^2] - \mu^2 \\ &= np(1-p+np) - (np)^2 \\ &= np(1-p) \end{aligned}$$

Example

Let the probability of a girl in each birth be 0.55. Assume births are independent.

If a couple has 4 children, the number of girls they will have is represented by $X \sim B(n=4, p=0.55)$.

Poisson distribution

If X is a discrete random variable with a probability mass function

$$i \in \{0, 1, 2, ...\}$$
$$p(i) = e^{-\lambda} \frac{\lambda^i}{i!}$$

Then we say X has a Poisson distribution with parameter λ.

1. The probability that exactly one event occurs in a given interval of length h is $\lambda h + o(h)$
   - $\lim_{h \to 0} \frac{o(h)}{h} = 0$
2. The probability that 2 or more events occur is $\Theta(h)$
3. The events occur independently in nonoverlapping intervals.

If a time interval has a length t, $X \sim \text{Poisson}(\lambda t)$, and X is the number of events in the interval.

Expected Value

$$E[X] = \sum_{i=1}^{\infty} i e^{-\lambda} \frac{\lambda^i}{i!} = \lambda \underbrace{\sum_{i=1}^{\infty} e^{-\lambda} \frac{\lambda^{i-1}}{(i-1)!}}_{\text{sum of all probabilities}} = \lambda$$

Variance

This is similarly easy to prove:

$$\text{Var } X = \lambda$$

Example

Earthquakes occur according to a Poisson process at the rate of 2 per week. Let X be the number of earthquakes in 1 week.

$X \sim \text{Poisson}(2)$

What is $P(X \geq 4)$?

$$P(X \geq 4) = 1 - P(X \leq 3) = 1 - \left(e^{-2}\frac{2^0}{0!} + e^{-2}\frac{2^1}{1!} + e^{-2}\frac{2^2}{2!} + e^{-2}\frac{2^3}{3!} \right)$$

Furthermore, if Y is the number of earthquakes in 3 weeks, $Y \sim \text{Poisson}(6)$.

Example

A textbook has 500 pages. For each page, the probability of having a typo is 0.01.

Compute the probability that there are 2 pages with typographical errors.

Let X be the number of pages with errors.

$$X \sim \text{Binomial}(n = 500, p = 0.01)$$

This is similar to the Poisson distribution with $\lambda = 5$, because over the entire book, you expect there to be $n \cdot p = 5$ pages with errors.

Note that the variance is 5, and for the binomial distribution, it is 4.95, which is about 5.

$$\sim (X) \sim \text{Poisson}(5)$$

$$P(X = 2) \approx P(\sim (X) = 2) = e^{-5}\frac{5^2}{2!} \approx 0.0842243$$

$$P(X = 2) \approx 0.083631$$

This method, in this instance, has an error of about 0.000593304.

Theorem on Approximating Binomial distribution with Poisson

Given a X a random variable with a binomial distribution with parameters n, p, if n is large, and p is small enough, The distribution of X can be approximated by the Poisson distribution where $\lambda = np$.

$$P(X = x) \approx P(\sim (X) = x)$$

$$\binom{n}{x} p^x (1-p)^{n-x} \approx e^{-\lambda}\frac{\lambda^x}{x!}$$

$$\binom{n}{x} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x} \approx e^{-\lambda}\frac{\lambda^x}{x!}$$

$$\frac{n!}{x!(n-x)!}\frac{\lambda^x}{n^x}e^{-\lambda} \approx e^{-\lambda}\frac{\lambda^x}{x!}$$

$$\frac{n!}{(n-x)!n^x}e^{-\lambda}\frac{\lambda^x}{x!} \approx e^{-\lambda}\frac{\lambda^x}{x!}$$

$$e^{-\lambda}\frac{\lambda^x}{x!} \approx e^{-\lambda}\frac{\lambda^x}{x!}$$

Geometric Random Variable

X is the number of trials until and including the success. Trials are independent, and there is a probability of success of p.

$$X \in \{1, 2, 3, ...\}$$

Then, $p(x) = (1-p)^{x-1}p$.

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

$$\mathrm{Var}(X) = \frac{1-p}{p^2}$$

Example

Assume the probability of getting a girl or a boy is 0.5 each, and births are independent.

In a town, each couple will have children until they get a girl. On average, how many kids will a household have?

$$E[X] = \frac{1}{p} = \frac{1}{0.5} = 2$$

Negative Binomial Random Variable

Let X be the number of trials until and including the rth success.

In other words, there needs to be r successes, and you should stop as soon as you get r successes.

Trials are independent, and the probability of success is p.

$$X \in \{r, r+1, r+2, ...\}$$

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

There are two parameters, r and p.

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

$$\mathrm{Var}(X) = \frac{1-p}{p^2}$$

Expected values sum by the linearity of the sum operator.

Hypergeometric Distribution

Overall let there be N items, with m special.

Draw a random sample of n items (without replacement is implied). Let X be the number of special items selected.

Because it is drawn without replacement, trials are not independent.

$$p(x) = \frac{\binom{N-m}{n-x}\binom{m}{x}}{\binom{N}{n}}$$

$$E[X] = \frac{nm}{N}$$

$$\text{Var}(X) = n\left(\frac{m}{N}\right)\left(1 - \frac{m}{N}\right)\left(1 - \frac{n-1}{N-1}\right)$$

Continuous Distributions

Expected Value/PDF

For a discrete value, the expected value is:

$$\sum_x xp(x)$$

This corresponds to

$$\int_{-\infty}^{\infty} xp(x)\,\mathrm{d}x$$

but $p(x)$ is a probability density function here.

For discrete values,

$$\sum_x p(x) = 1$$

Which corresponds to

$$\int_{-\infty}^{\infty} p(x) = 1$$

Example

Let X have the pdf

$$f(x) = \begin{cases} \frac{x}{2} & \text{if } 0 \le x \le 2 \\ 0 & \text{otherwise} \end{cases}$$

$$E[X] = \int_{-\infty}^{\infty} xf(x)\,\mathrm{d}x = \int_0^2 x\frac{x}{2}\,\mathrm{d}x = \int_0^2 \frac{x^2}{2}\,\mathrm{d}x = \frac{1}{6}(2^3 - 0^3) = \frac{4}{3}$$

Functions

Given X, a continuous random variable with pdf $f(x)$, for any function $H(X)$,

$$E[H(X)] = \int_{-\infty}^{\infty} H(x)f(x)\,\mathrm{d}x$$

Example

Let $Y = e^X$, where X is defined by the pdf

$$f(x) = \begin{cases} \frac{x}{2} & \text{if } 0 \le x \le 2 \\ 0 & \text{otherwise} \end{cases}$$

$$\begin{aligned}
&\int_{-\infty}^{\infty} e^x f(x)\,\mathrm{d}x \\
&= \int_0^2 e^x \frac{x}{2}\,\mathrm{d}x \\
&= \frac{1}{2}\int_0^2 xe^x\,\mathrm{d}x \\
&= \frac{1}{2}\left(xe^x\big|_0^2 - \int_0^2 e^x\,\mathrm{d}x \right) \\
&= \frac{1}{2}\Big((xe^x - e^x)\big|_0^2 \Big) \\
&= \frac{1}{2}((2e^2 - e^2) - (0e^0 - e^0)) \\
&= \frac{1}{2}(e^2 + 1) \\
&= \frac{e^2 + 1}{2}
\end{aligned}$$

Variance

$$\text{Var } X = E[(X - \mu)^2]$$

$$\begin{aligned}
E[(X-\mu)^2] &= \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,\mathrm{d}x \\
&= \int_{-\infty}^{\infty} (x^2 - 2x\mu + \mu^2) f(x)\,\mathrm{d}x \\
&= \int_{-\infty}^{\infty} x^2 f(x)\,\mathrm{d}x - 2\mu \int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x + \mu^2 \int_{-\infty}^{\infty} f(x)\,\mathrm{d}x \\
&= E[X^2] - 2\mu E[X] + \mu^2(1) \\
&= E[X^2] - 2\mu^2 + \mu^2 \\
&= E[X^2] - \mu^2
\end{aligned}$$

Example

X is defined by the pdf

$$f(x) = \begin{cases} \frac{x}{2} & \text{if } 0 \le x \le 2 \\ 0 & \text{otherwise} \end{cases}$$

What is the variance?

$E[X] = \frac{4}{3}$

$$
\begin{aligned}
E[X^2] &= \int_{-\infty}^{\infty} x^2 f(x)\,\mathrm{d}x = \int_0^2 x^2 \frac{x}{2}\,\mathrm{d}x \\
&= \frac{x^4}{4\cdot 2}\bigg|_0^2 = \frac{2^4}{2^3} - \frac{0}{2^3} = 2
\end{aligned}
$$

$$
\mathrm{Var}\ X = E[X^2] - \mu^2 = 2 - \left(\frac{4}{3}\right)^2 = \frac{18}{9} - \frac{16}{9} = \frac{2}{9}
$$

Cumulative distribution function

Given X, a continuous random variable, the cumulative distribution function, $F(u) = P(x \leq u) = \int_{-\infty}^{u} f(x)\,\mathrm{d}x$

Example

Let X be a random variable with pdf $f_X(x)$. Let $Y = 2X$, and derive the probability density function $f_Y(y)$ for y.

$$
\begin{aligned}
F_Y(y) &= P(Y \leq y) \\
&= P(2X \leq y) \\
&= P\Big(X \leq \frac{y}{2}\Big) \\
&= \int_{-\infty}^{y/2} f_X(x)\,\mathrm{d}x
\end{aligned}
$$

Therefore,

$$
f_Y(y) = f_X\Big(\frac{y}{2}\Big) \cdot \frac{1}{2}
$$

Uniform Distrubtion

$$
f(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & \text{otherwise} \end{cases}
$$

The CDF is then:

$$
F(x) = \begin{cases} 0 & x \leq a \\ \frac{x-a}{b-a} & a < x < b \\ 1 & b \leq x \end{cases}
$$

The expected value is $E[X] = \frac{1}{2}(a+b)$

$$
E[X^2] = \frac{1}{3}(a^2 + ab + b^2)
$$

The variance is

$$
\mathrm{Var}\ X = E[X^2] - E[X]^2 = \frac{(b-a)^2}{12}
$$

Example

Buses arrive at the station every 15 minutes, starting at 7 am. If a passenger arrives at a time that is uniformly distributed between 7 and 7:30 am, what is the probability that they will wait less than 5 minutes?

To arrive to wait less than 5 minutes, he needs to arrive between 7:10 and 7:15 or between 7:25 and 7:30.

$$f(x) = \begin{cases} 0 & x \leq 0 \\ \frac{1}{30} & 0 < x < 30 \\ 0 & 30 \leq x \end{cases}$$

Then

$$\int_{10}^{15} f(x)\,\mathrm{d}x + \int_{25}^{30} f(x)\,\mathrm{d}x = \frac{1}{3}$$

is the probability that they will wait less than 5 minutes.

Normal Distribution

We say X has a normal distribution with parameters μ, σ^2 is:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}$$

$$\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}\,\mathrm{d}x \qquad y = \frac{x-\mu}{\sigma}$$

$$= \int \frac{1}{\cancel{\sigma}\sqrt{2\pi}} e^{-\frac{1}{2}y^2} \cancel{\sigma}\,\mathrm{d}y$$

$$I = \int_{-\infty}^{\infty} e^{-\frac{1}{2}y^2}\,\mathrm{d}y$$

$$I = \int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2}\,\mathrm{d}x$$

$$\Rightarrow I^2 = \int_{-\infty}^{\infty} e^{-\frac{1}{2}y^2}\,\mathrm{d}y \int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2}\,\mathrm{d}x$$

$$\begin{aligned}
I^2 &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{1}{2}(x^2+y^2)}\,\mathrm{d}x\,\mathrm{d}y \\
&= \int_0^{2\pi}\int_0^{\infty} re^{-\frac{1}{2}r^2}\,\mathrm{d}r\,\mathrm{d}\theta \\
&= \int_0^{2\pi}\int_0^{\infty} e^{-u}\,\mathrm{d}r\,\mathrm{d}\theta \\
&= \int_0^{2\pi} \left(-e^{-u}\big|_0^{\infty}\right)\mathrm{d}\theta \\
&= \int_0^{2\pi} -e^{-\infty} + e^0\,\mathrm{d}\theta \\
&= \int_0^{2\pi} 1\,\mathrm{d}\theta \\
&= 2\pi
\end{aligned}$$

$$I^2 = 2\pi \Rightarrow I = \sqrt{2\pi}$$

And therefore

$$\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}\,\mathrm{d}x = 1$$

Moving Normal

If X is normal with parameters μ and σ^2, $Y = aX + b$, Y is normal with parameters $a\mu + b$ and $a^2\sigma^2$.

$$F_Y(y) = P(Y \leq y) = P(aX + b \leq y) = P\left(X \leq \frac{y-b}{a}\right) = \int_{-\infty}^{\frac{y-b}{a}} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\,\mathrm{d}x$$

$$\begin{aligned}
f_Y(y) &= \frac{1}{a\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{y-b-a\mu}{a\sigma}\right)^2} \\
&= \frac{1}{a\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{y-(a\mu+b)}{a\sigma}\right)^2}
\end{aligned}$$

Therefore, Y is normal with parameters $(a\mu + b, a^2\sigma^2)$.

Standard Normal

If $X \sim \mathrm{Normal}(\mu, \sigma^2)$

$$Z = \frac{X-\mu}{\sigma} \sim \mathrm{Normal}\left(\frac{\mu}{\sigma} - \frac{\mu}{\sigma}, \frac{1}{\sigma^2}\sigma^2\right) = \mathrm{Normal}(0, 1)$$

The normal distribution with parameters 0 and 1 is the standard normal. It is usually denoted Z.

You can also know $E[X] = \sigma E[Z] + \mu$ and $\mathrm{Var}(X) = \sigma^2\ \mathrm{Var}(Z)$.

$$E[Z] = \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}\,\mathrm{d}x = -\lim_{t\to\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}\Big|_{-t}^{t} = -\lim_{t\to\infty}\left(\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}t^2} - \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}t^2}\right) = -0 = 0$$

$$E(Z^2) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x \cdot x e^{-\frac{1}{2}x^2}\,\mathrm{d}x$$

$$\begin{aligned} f &= x \qquad & g &= -e^{-\frac{1}{2}x^2} \\ \mathrm{d}f &= 1\,\mathrm{d}x \qquad & \mathrm{d}g &= x e^{-\frac{1}{2}x^2}\,\mathrm{d}x \end{aligned}$$

$$\begin{aligned} &= \frac{1}{\sqrt{2\pi}}\left(0 - \int_{-\infty}^{\infty} -e^{\frac{1}{2}x^2}\,\mathrm{d}x\right) \\ &= \frac{1}{\sqrt{2\pi}}\sqrt{2\pi} = 1 \end{aligned}$$

And therefore, the $\mathrm{Var}(Z) = 1$.

$$F_Z(x) = \Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2}\,\mathrm{d}y$$

Note that $\Phi(x) = 1 - \Phi(-x) \Rightarrow \Phi(-x) = 1 - \Phi(x)$.

$$F_X(x) = P(X < x) = P(\sigma Z + \mu < x) = P\left(Z < \frac{x-\mu}{\sigma}\right) = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

Example

If X is normal with $\mu = 3$ and $\sigma^2 = 4$:

1. $P(2 < x < 5)$

$$\begin{aligned} P(2 < x < 5) = F_X(5) - F_X(2) = \Phi\left(\frac{5-3}{4}\right) - \Phi\left(\frac{2-3}{4}\right) = \Phi(1) - \Phi\left(-\frac{1}{2}\right) = \Phi(1) - 1 + \Phi\left(\frac{1}{2}\right) \\ = 0.841345 - 1 + 0.691462 = 0.532807 \end{aligned}$$

2. $P(|X-3| > 4)$

$$\begin{aligned} P(|X-3| > 4) = P(X - 3 > 4) + P(X - 3 < -4) = P(Z > 2) + P(Z < -2) \\ = 2 \times \Phi(-2) = 2 \times (1 - \Phi(2)) = 2 \times (1 - 0.97725) = 0.0455003 \end{aligned}$$

Central Limit Theorem

If S_n denotes the number of successes that occur when n independent trials where the probability of success for each trial is p are performed, then for any a, b:

$$\lim_{n\to\infty} P\left(a \le \frac{S_n - np}{\sqrt{np(1-p)}} \le b\right) = \Phi(b) - \Phi(a)$$

In other words, if n is large,

$$S_n \sim \mathrm{Normal}(np, np(1-p))$$

Example
With a fair coin, what is the probability of getting 48 or more heads in 100 tosses?

$$X \sim \text{Binomial}(100, 0.5)$$

$$X \sim \text{Normal}(100 \cdot 0.5, 100 \cdot 0.5(1-0.5)) = \text{Normal}(50, 25)$$

$\sigma^2 = 25 \Rightarrow \sigma = 5$

$$P(X \geq 48) = P(X > 47.5) = P\left(\frac{X-50}{5} > \frac{47.5-50}{5}\right) \approx P(Z > -0.5) = 1 - \Phi(-0.5) = \Phi(0.5) = 0.691462$$

The difference between the real solution and the approximation is -0.000112168.

If X has pdf

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

Assume events occur according to a Poisson process at a rate of λ per unit time.

Let X denote the wait time until the first event occurs.

$$P(X > t) = e^{-\lambda t} \frac{(\lambda t)^0}{0!} = e^{-\lambda t}$$

$$F_X(t) = P(X < t) = 1 - P(X \geq t) = 1 - e^{-\lambda t}$$

$$f_x(t) = \frac{\mathrm{d}}{\mathrm{d}t}(1 - e^{-\lambda t}) = \lambda e^{-\lambda t}$$

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

$$\text{Var}(X) = \frac{1}{\lambda^2}$$

Gamma Distribution

$$f(x) = \begin{cases} \frac{\lambda e^{-\lambda x}(\lambda x)^{\alpha-1}}{\Gamma(\alpha)} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

For integer n, $\Gamma(n) = (n-1)!$.

For any real number α, $\Gamma(1) = (\alpha - 1)\Gamma(\alpha - 1)$

$$E[X] = \frac{\alpha}{\lambda}$$

$$\text{Var } X = \frac{\alpha}{\lambda^2}$$

Example
k Given a Poisson Process with rate λ per unit time.

Compute the probability it takes less than t units of time for α events to occur.

Let $X \sim \text{Gamma}(\alpha, \lambda)$. The probability is then $P(X < t) = F_X(t)$.

This is because

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_{k=\alpha}^{\infty} \frac{e^{\lambda t}(\lambda t)^k}{k!}\right) = \frac{\lambda e^{-\lambda x}(\lambda x)^{\alpha-1}}{(\alpha-1)!}$$

Consequently, the gamma distribution is like the Poisson distribution but requires α events.

Jointly Distributed Random Variables

Joint Cumulative Probability Distribution

The joint cumulative probability distribution function of X and Y

$$F(a,b) = P(X \leq a, Y \leq b)$$

By the above:

$$P(a_1 < X \leq a_2, b_1 < y \leq b_2) = F(a_2,b_2) - F(a_1,b_2) - F(a_2,b_1) + F(a_1,b_1)$$

Joint Probability Mass Function

Given discrete X, Y:

$$p(x,y) = P(X = x, Y = y)$$

You can get an individual probability mass function from the joint one:

$$p_X(x) = P(X = x) = \sum_{j=1}^{\infty} P\big(X = x, Y = y_j\big) = \sum_{j=1}^{\infty} p\big(x, y_j\big)$$

$$p_Y(y) = \sum_{j=1}^{\infty} p\big(x_j, y\big)$$

Example

You have an urn with 3 red, 4 white and 5 blue balls. 3 balls are to be drawn randomly in an experiment. Let X be the number of red balls drawn and Y be the number of white balls drawn.

Compute $p(x,y)$, the joint probability mass function.

$$p(x,y) = \frac{\binom{3}{x}\binom{4}{y}\binom{5}{3-(x+y)}}{\binom{3+4+5}{3}}$$

Compute $p(1,1)$:

$$p(1,1) = \frac{3 \cdot 4 \cdot 5}{4 \cdot 5 \cdot 11} = \frac{3}{11}$$

Definition of discrete jointly distributed random variables

Given discrete X, Y, they are jointly distributed if there exists a

$$p(x, y) = P(X = x, Y = y)$$

Example

15% of families have no children, 20% have one child, 35% have 2 children, and 30% have 3 children.

Assume for each family and each birth, the probability of having a girl or a boy is equal, and births are independent.

X\Y	0	1	2	3
0	0.15	$0.20 \times \frac{1}{2}$	$0.35 \times \frac{1}{2}$	$0.30 \times \frac{1}{8}$
1	$0.20 \times \frac{1}{2}$	$0.35 \times \frac{2}{4}$	$0.30 \times \frac{3}{8}$	0
2	$0.35 \times \frac{1}{4}$	$0.30 \times \frac{3}{8}$	0	0
3	$0.30 \times \frac{1}{8}$	0	0	0

Definition of jointly distributed random variables

We say two random variables are jointly distributed random variables if there is a $f(x, y)$ such that

$$\iint_C f(x, y) \, \mathrm{d}x \, \mathrm{d}y = P((X, Y) \in C)$$

Integral is 1

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \, \mathrm{d}y \, \mathrm{d}x = 1$$

Integrals for jointly distributed random variables

$$P(X \leq a, Y \leq b) = F(a, b) = \int_{-\infty}^{b} \int_{-\infty}^{a} f(x, y) \, \mathrm{d}x \, \mathrm{d}y$$

$$\frac{\partial F(a, b)}{\partial a \partial b} = f(a, b)$$

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) \, \mathrm{d}y$$

Example

The joint pdf of X and Y is given as

$$f(x, y) = \begin{cases} 2e^{-x}e^{-2y} & 0 \leq x < \infty, 0 \leq y < \infty \\ 0 & x < 0, y < 0 \end{cases}$$

Compute $P(X > 1, Y < 1)$.

Compute $P(\frac{X}{Y} < a)$

$$P\left(\frac{X}{Y} < a\right) = P(X < aY) = \int_0^\infty \int_0^{ay} (2e^{-x}e^{-2y})\,\mathrm{d}x\,\mathrm{d}y$$
$$= \frac{a}{2+a}$$

Let $R = \frac{X}{Y}$.

Then we know $P(R < a) = F_R(a) = \frac{a}{2+a}$. This implies that $f_R(a) = \frac{2}{(2+a)^2}$.

What is the expected value of $E\left(\frac{X}{Y}\right)$?

$$\int_0^\infty a\left(\frac{2}{(a+2)^2}\right)\mathrm{d}a \to \infty$$

Joint Distribution of n random variables

Joint cumulative probability distribution function

$$X_1, X_2, X_3, X_4, ..., X_n$$

$$F(a_1, a_2, ..., a_n) = P(X_1 \le a_1, X_2 \le a_2, ..., X_n \le a_n)$$

$$\underbrace{\int\int\int\cdots\int}_{n} f(x_1, x_2, x_3, ..., x_n)\,\mathrm{d}x_n \cdots \mathrm{d}x_3\,\mathrm{d}x_2\,\mathrm{d}x_1$$

Multinomial Distribution

An experiment can end in one of the r outcomes $\{o_1, o_2, ..., o_r\}$, with respect to probabilities $\{p_1, ..., p_r\}$. The experiment is repeated n times.

$$X_1 = \#o_1$$
$$X_2 = \#o_2$$
$$\vdots$$
$$X_r = \#o_r$$

$$p(a_1, a_2, ..., a_r) = P(X_1 = a_1, ..., X_r = a_r)$$

$$p(a_1, a_2, ..., a_r) = p_1^{a_1} p_2^{a_2} \cdots p_r^{a_r} \binom{n}{a_1, a_2, a_3, ..., a_r}$$

Independent random variables

Example

A man and a woman arrive independently at a time uniformly distributed between 12-1. Compute the probability that the one arriving first has to wait more than 10 minutes.

Distribution of Sum of Independent Random Variables

Let X and Y be continuous random variables. The probability density of $X + Y$ is then

$$
\begin{aligned}
F_{X+Y} &= P(X+Y \le a) \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{a-y} f_X(x) f_Y(y) \,\mathrm{d}x \,\mathrm{d}y \\
&= \int_{-\infty}^{\infty} F_X(a-y) f_Y(y) \,\mathrm{d}y
\end{aligned}
$$

$$
f_{X+Y}(a) = \int_{-\infty}^{\infty} f_X(a-y) f_Y(y) \,\mathrm{d}y
$$

This appears to be a convolution

Example

Let X, Y be two uniform distributions in $[0, 1]$.

What is the distribution of $X + Y$?

$$
0 \le a \le 1 \to f_{X+Y}(a) = \int_0^a 1 \cdot 1 \,\mathrm{d}y = a
$$

$$
1 < a \le 2 \to f_{X+Y}(a) = \cdots = 2 - a
$$

Expected value

Given X, Y as two random variables,

$$
E[h(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, y) f(x, y) \,\mathrm{d}x \,\mathrm{d}y
$$

Example

An accident occurs at a position X, where X is uniformly distributed in $[0, L]$. An ambulance is at position Y, and Y is also uniformly distributed in $[0, L]$. X, Y are independent.

$$
\begin{aligned}
E[|X - Y|] &= \frac{1}{L^2} \int_0^L \int_0^L |x - y| \,\mathrm{d}x \,\mathrm{d}y \\
&= \frac{1}{L^2} \int_0^L \int_0^y |x - y| \,\mathrm{d}x + \int_y^L |x - y| \,\mathrm{d}x \,\mathrm{d}y \\
&= \frac{L}{3}
\end{aligned}
$$

Linearity of Expected Value

For any two random variables X, Y

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

$$
\begin{aligned}
E[X+Y] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y)f(x,y)\,\mathrm{d}x\,\mathrm{d}y \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xf(x,y)\,\mathrm{d}x\,\mathrm{d}y + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}yf(x,y)\,\mathrm{d}x\,\mathrm{d}y \\
&= E[X]+E[Y]
\end{aligned}
$$

And therefore

$$E[X_1+X_2+\cdots+X_n] = E[X_1]+E[X_2]+\cdots+E[X_n]$$

Example

Let X be a binomial random variable with parameters (n,p).

$$X = X_1+X_2+\cdots+X_n$$

Where

$$X_i = \begin{cases} 1 \text{ if the } i\text{th trial is a success} \\ 0 \text{ if the } i\text{th trial is a failure} \end{cases}$$

and $P(i\text{th trial}) = p$.

$$
\begin{aligned}
E[X] &= E[X_1]+E[X_2]+\cdots+E[X_n] \\
&= \underbrace{p+p+\cdots+p}_{n \text{ times}} \\
&= np
\end{aligned}
$$

Covariance

$$\mathrm{Cov}(X,Y) = E[(X-E[X])(Y-E[Y])]$$

Covariance Computation

$$\mathrm{Cov}(X,Y) = E[XY]-E[X]E[Y]$$

$$
\begin{aligned}
\mathrm{Var}\left(\sum_{i=1}^{n} x_i\right) &= \mathrm{Cov}\ \left(\sum_{i=1}^{n} X_i, \sum_{k=1}^{n} X_k\right) \\
&= \sum_{i=1}^{n}\sum_{k=1}^{n}\mathrm{Cov}\ (X_i,X_k) \\
&= \sum_{i=1}^{n}\mathrm{Var}(X_i) + 2\sum_{i<k}\mathrm{Cov}(X_iY_k)
\end{aligned}
$$

Example

Consider m indepdent trials, each of which can end in one of the outcomes $\Theta_1, \Theta_2, ..., \Theta_r$ with respective probabilties, $p_1, p_2, ..., p_r$.

Let N_i be the number of trials ending in outcome i.

What is $\mathrm{Cov}(N_i, N_j)$?

Define

$$I_i(\ell) = \begin{cases} 1 & \text{if trial } \ell \text{ ends in outcome } i \\ 0 & \text{otherwise} \end{cases}$$

Then,

$$N_i = \sum_{\ell=1}^{m} I_i(\ell)$$

And

$$\begin{aligned} \operatorname{Cov}(N_i, N_j) &= \operatorname{Cov}\left(\sum_{\ell=1}^{m} I_i(\ell), \sum_{\ell=1}^{m} I_j(\ell)\right) \\ &= \sum_{\ell=1} \operatorname{Cov}(I_i(\ell), I_j(\ell)) + \sum_{k\neq\ell} \operatorname{Cov}(I_i(\ell), I_j(k)) \\ &= \sum_{\ell=1} \operatorname{Cov}(I_i(\ell), I_j(\ell)) \\ &= \sum_{\ell=1} (E[I_i(\ell)I_j(\ell)] - E[I_i(\ell)]E[I_j(\ell)]) \\ &= -mp_ip_j \end{aligned}$$

Conditional probability distribution

For X, Y discrete, jointly distributed random variables with probability mass function $p(x, y)$:

$$P_{X|Y} = P(X = x|Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)} = \frac{p(x, y)}{p_Y(y)}$$

$$E[X \mid Y = y] = \sum_x x p_{X \mid Y}(x|y)$$

For X, Y continuous, $f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y}(y)$

$$E[X|Y = y] = \int_{-\infty}^{\infty} x f_{X|Y}(x|y) \, \mathrm{d}x$$

Example

Suppose that the joint probability distribution function of X and Y is

$$f(x, y) = \frac{e^{-x/y} e^{-y}}{y}$$

for $y > 0$. Compute $E[X|y]$.

$$f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)}$$

$$f_Y(y) = \int_{-\infty}^{\infty} \frac{e^{-x/y} e^{-y}}{y} \, \mathrm{d}x = e^{-y}$$

Therefore

$$f_{X|Y}(x|y) = \frac{\frac{e^{-x/y}e^{-y}}{y}}{e^{-y}} = \frac{e^{-x/y}}{y}$$

And then:

$$E[X|y] = \int_0^\infty x\frac{e^{-x/y}}{y} = y$$

Binomial

If $X \sim B(n, p)$ and $Y \sim B(m, p)$, $X + Y \sim B(m + n, p)$

Prop

$E[X] = E[E[X|y]]$

Proof

If Y is discrete:

$$E[X] = \sum_y E[X|y]p_Y(y)$$

if Y is continuous:

$$E[X] = \int_{-\infty}^\infty E[X|y]f_Y(y)\,\mathrm{d}y$$

Example

Let there be 3 doors. Going through door 1 will take 3 hours to leave, going through door 2 will take 5 hours, and going through door 3 will take 7 hours.

Let X be the number of hours it takes for the number to get out.

$$E[X] = E[X|\text{Door 1}]p(\text{door 1}) + E[X|\text{Door 2}]p(\text{door 2}) + E[X|\text{Door 3}]p(\text{door 3})$$

Then, $p(\text{door } i) = \frac{1}{3}$:

$$E[X] = 3\text{ hours} \times \frac{1}{3} + 5\text{ hours} \times \frac{1}{3} + 7\text{ hours} \times \frac{1}{3} = 5\text{ hours}$$