Combinatorics

#dataScience

Permutations

Permutations represent the number of different possible ways we can *arrange* a number of elements.

$$P(n) = n \times (n-1) \times (n-2) \times \cdots \times 1$$

P = Permutations.
n = Options for who we put first.
$(n-1)$ = Options for who we put second.
1 = Options for who we put last.

Characteristics of Permutations:

- Arranging all elements within the sample space.
- No repetition.
- $P(n) = n \times (n-1) \times (n-2) \times \cdots \times 1 = n!$ (Called “n factorial”).

Example:

- If we need to arrange 5 people, we would have $P(5) = 120$ ways of doing so.

Factorials

Factorials express the *product* of all integers from 1 to n and we denote them with the “!” symbol.

$$n! = n \times (n-1) \times (n-2) \times \cdots \times 1$$

Key Values:

- $0! = 1.$
- If $n < 0$, $n!$ does not exist.

Rules for factorial multiplication. (For $n > 0$ and $n > k$)

- $$n + k! = n! \times (n+1) \times \cdots \times (n+k)$$
- $$n - k! = \frac{n!}{(n-k+1) \times \cdots \times (n-k+k)} = \frac{n!}{(n-k+1) \times \cdots \times n}$$
- $$\frac{n!}{k!} = \frac{k! \times (k+1) \times \cdots \times n}{k!} = (k+1) \times \cdots \times n$$

Examples: $n = 7$, $k = 4$

- $(7+4)! = 11! = 7! \times 8 \times 9 \times 10 \times 11$

- $(7-4)! = 3! = \frac{7!}{4\times5\times6\times7}$
- $\frac{7!}{4!} = 5\times6\times7$

Variations

Variations represent the number of different possible ways we can *pick* and *arrange* a number of elements.

$$\hat{V}(n,p) = n^p$$

$\hat{V}$ = Variations with repetition.
n = Number of different elements available.
p = Number of elements we are arranging.

$$V(n,p) = \frac{n!}{(n-p)!}$$

V = Variations without repetition.
n = Number of different elements available.
p = Number of elements we are arranging.

Intuition behind the formula. (*With Repetition*).

- We have n-many options for the first element.
- We *still have n-many options* for the second element because repetition is allowed.
- We have n-many options for each of the pmany elements.
- $n\times n\times n\ldots n = n^p$

Intuition behind the formula. (*Without Repetition*).

- We have n-many options for the first element.
- We only have *(n-1)-many options* for the second element because we cannot repeat the value for we chose to start with.
- We have less options left for each additional element.
- $n\times(n-1)\times(n-2)\ldots(n-p+1) = \frac{n!}{(n-p)!}$

Combinations

Combinations represent the number of different possible ways we can *pick* a number of elements.

$$C(n,p) = \frac{n!}{(n-p)!p!}$$

C = Combinations.
n = Total number of elements in the sample space.
p = Number of elements we need to select.

Characteristics of Combinations:

- Takes into account double-counting. (Selecting Johny, Kate and Marie is the same as selecting Marie, Kate and Johny)
- All the different permutations of a single combination are different variations.
- $C = \frac{V}{P} = \frac{n!/n-p!}{p!} = \frac{n!}{(n-p)!p!}$

- Combinations are symmetric, so $C_p^n = C_{n-p}^n$, since selecting p elements is the same as omitting $n - p$ elements.

Combinations with separate sample spaces

Combinations represent the number of different possible ways we can *pick* a number of elements.

$$C = n_1 \times n_2 \times \cdots \times n_p$$

C = Combinations.
n_1 = Size of the first sample space.
n_2 = Size of the second sample space.
n_p = Size of the last sample space.

Characteristics of Combinations with separate sample spaces:

- The option we choose for any element does not affect the number of options for the other elements.
- The order in which we pick the individual elements is arbitrary.
- We need to know the size of the sample space for each individual element. $(n_1, n_2 \ldots n_p)$

Winning the Lottery

To win the lottery, you need to satisfy two distinct independent events:

- Correctly guess the “Powerball” number. (From 1 to 26)
- Correctly guess the 5 regular numbers. (From 1 to 69)

$$C = \frac{69!}{64!5!} \times 26$$

C = Total number of Combinations.
$\frac{}{64!5!}$ = $C_{5\text{ numbers}}$
26 = $C_{\text{Powerball Numbers}}$

Intuition behind the formula:

- We consider the two distinct events as a combination of two elements with different sample sizes.
    - One event has a sample size of 26, the other has a sample size of C_5^{69}.
- Using the “favoured over all” formula, we find the probability of any single ticket winning equals $1/(\frac{69!}{64!5!} \times 26)$.