Module 5: Supplementary Slides

Counting Rules

Rule 1 (Multiplicative Rule)

If an experiment comprises of n steps, with each step offering a unique set of outcomes — the first step having k_1 outcomes, the second step having k_2 possible outcomes, and so on up to the n^{th} step with k_n outcomes — the total number of potential outcomes for the entire experiment is:

$$k_1 k_2 \ldots k_n$$

Example:

You're interested in enjoying a well-rounded day out by visiting a park, having a meal at a restaurant, and watching a movie. With a selection of 3 unique parks, 4 distinct restaurants, and 6 different movie options, how many delightful combinations can be formed?

Solution:

$$2 \times 3 \times 6 = 72$$

Rule 2 (Factorial Rule)

The number of ways that n items can be arranged in order is

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

Remark:

$$0! = 1$$

Example:

You have five books to put on a bookshelf. How many different ways can these books be placed on the shelf?

Solution:

$$5! = 120$$

Rule 3 (Permutation Rule)

The arrangement of objects in a specific order is called permutation. The number of ways of selecting k objects from n different objects and arranging them in a specific order is

$$P(n,k) = \frac{n!}{(n-k)!}$$

Example:

You have 5 different books and are going to put 3 on a bookshelf. How many different ways can the books be ordered on the bookshelf?

Solution:

$$P(5,3) = \frac{5!}{(5-3)!} = 5 \times 4 \times 3 = 60$$

Rule 4 (Combination Rule)

The number of ways of selecting k objects from n objects, irrespective of order, is

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

Example:

You have 4 books and are going to select two to read. How many different combinations are there, ignoring the order in which they are selected?

Solution:

$$C(4,2) = \frac{2!}{2!\,(4-2)!} = 6$$

Rule 5 (Partitions Rule)

The formula for determining the count of arrangements possible for a set of n objects, structured into g groups or partitions, each composed of identical elements, with respective sizes denoted as $k_1, k_2, \dots, k_g$ (with $k_1 + k_2 + \cdots + k_g = n$), can be expressed as follows:

$$\binom{n}{k_1} \times \binom{n-k_1}{k_2} \times \cdots \times \binom{k_g}{k_g} = \frac{n!}{k_1!\,k_2! \dots k_g!}$$

Example:

How many distinct arrangements can we make from the characters of the word BOB?

Solution:

$$\binom{3}{2}\binom{1}{1} = \frac{3!}{2!\,1!} = 6$$

Some Terminologies Related to Bayes' Theorem with Applications in Statistics and Machine Learning

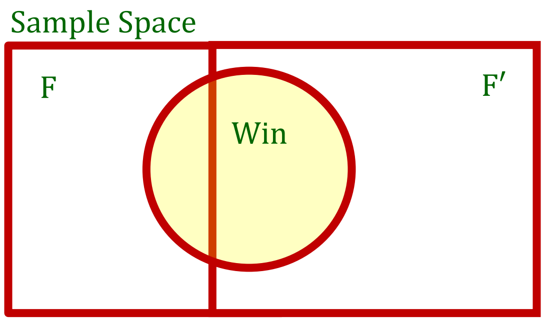
Example

A soccer team wins 60% of its games when it scores the first goal, and 10% of its games when the opposing team scores first. If the team scores the first goal about 30% of the time, what fraction of the games does it win?

Solution

$$P(\mathrm{Win}) = P(F) \times P(\mathrm{Win}|\mathrm{F}) + P(F') \times \mathrm{P}(\mathrm{Win}|F')$$

$$= 0.3 \times 0.6 + 0.7 \times 0.1 = 0.18 + 0.07 = 0.25$$



True and Fasle Positives and Negatives

- True Positive (TP): These are the cases in which the test correctly predicts the positive condition. For instance, if it's a medical test for a disease, the TP cases are those who have the disease and test positive for it.
- False Positive (FP): These are the cases where the test predicts the positive condition when it's not actually present. In the setting of a medical test, for instance, an FP case would be a person who does not have the disease but receives a positive result from the test.
- True Negatives (TN): These are the cases in which the test correctly predicts the negative condition. For instance, if it's a medical test for a disease, the TN cases are those who do not have the disease and test negative for it.
- False Negatives (FN): These are the cases in which the test incorrectly predicts the negative condition. In the context of the medical test, FN cases are those who have the disease but test negative for it.

	S	S'
T	True Positive (TP)	False Positive (FP)
T'	False Negative (FN)	True Negative (TN)

$$P(S|T) = \frac{P(S) \times P(T|S)}{P(S) \times P(T|S) + P(S') \times P(T|S')}$$

Sensitivity of a Test

Sensitivity ($P(T|S)$): It is the conditional probability that tells us how well the test identifies positive cases correctly. A high sensitivity indicates that the test correctly identifies most of the positive cases and misses fewer of them.

$$\text{Sensitivity} = P(T|S) = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Positives (FP)}}$$

Sensitivity can be important when the cost of missing a positive case (False Negative) is high. For example, in medical tests for serious diseases, a high sensitivity is crucial because failing to detect the disease in someone who has it (a false negative) can have severe consequences.

	S	S'
T	True Positive (TP)	False Positive (FP)
T'	False Negative (FN)	True Negative (TN)

$$P(S|T) = \frac{P(S) \times P(T|S)}{P(S) \times P(T|S) + P(S') \times P(T|S')}$$

Selectivity of a Test

Selectivity (Specificity) ($P(T'|S')$): It is the conditional probability that measures how well the test identifies negative cases correctly.

$$\text{Specificity} = P(T'|S') = \frac{\text{True Negatives (TN)}}{\text{False Positives (FP)} + \text{True Negatives (TN)}}$$

In situations where the cost or consequences of a false positive are high, specificity becomes critical. For example, in a diagnostic setting, falsely diagnosing someone with a severe illness can lead to unnecessary stress, invasive procedures, and unwarranted treatments.

	S	S'
T	True Positive (TP)	False Positive (FP)
T'	False Negative (FN)	True Negative (TN)

$$P(S|T) = \frac{P(S) \times P(T|S)}{P(S) \times P(T|S) + P(S') \times P(T|S')}$$

Prior Distribution

Prior (or Prevalence) ($P(S)$): It is a probability that represents the state of knowledge (or uncertainty) about a parameter before considering new evidence (data).

In epidemiology and medical research, “prevalence” is the equivalent terminology used, which refers to the proportion of a population that have a specific characteristic, condition, or disease at a specific point in time or over a specific period (before observing new evidence). For instance, the prevalence of diabetes in a population would represent the proportion of people in that population who have diabetes. (This proportion is also our prior belief in Bayesian inference).

	S	S'
T	True Positive (TP)	False Positive (FP)
T'	False Negative (FN)	True Negative (TN)

$$P(S|T) = \frac{P(S) \times P(T|S)}{P(S) \times P(T|S) + P(S') \times P(T|S')}$$

Posterior Distribution and Likelihood

Posterior ($P(S|T)$): It is the updated probability of a parameter after observing new evidence (data). It integrates prior beliefs or knowledge (expressed through the "prior" distribution $P(S)$) with current observed data (expressed through the "likelihood" function $P(T|S)$). The fundamental concept of Bayesian inference is updating beliefs in light of new data.

$$\text{Posterior} \propto \text{Prior} \times \text{Likelihood}$$

$$P(S|T) \propto P(S) \times P(T|S)$$

Likelihood ($P(T|S)$): $P(T|S)$ is the likelihood, which describes the probability of observing the data (evidence) T given/assuming a particular parameter value S is true.

	S	S'
T	True Positive (TP)	False Positive (FP)
T'	False Negative (FN)	True Negative (TN)

$$P(S|T) = \frac{P(S) \times P(T|S)}{P(S) \times P(T|S) + P(S') \times P(T|S')}$$

Accuracy

Accuracy: is a metric used to measure the performance of a classification test (e.g., medical test). Specifically, accuracy is the fraction of predictions that a classification model got right.

$$\text{Accuracy} = \frac{\text{Number of correct preditions}}{\text{Total number of predictions}} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

Equivalently, using the probability values, it can also be obtained as:

$$\text{Accuracy} = \text{P}(T \cap S) + P(T' \cap S') = \text{P(S)} \times \text{P(T|S)} + \text{P(S}') \times \text{P(T}'|\text{S}')$$

	S	S′
T	True Positive (TP)	False Positive (FP)
T′	False Negative (FN)	True Negative (TN)

$$\text{P(S|T)} = \frac{\text{P(S)} \times \text{P(T|S)}}{\text{P(S)} \times \text{P(T|S)} + \text{P(S}') \times \text{P(T|S}')}$$