Risk Assessment / Statistical Modelling

-TEAM-06

Description

This DSL is developed for risk assessment requiring detailed probabilistic and statistical analysis. This DSL will enable users to calculate expectations, probabilities (Conditional probability, Joint probability, etc..), mean time of failure (MTF), mean, median, mode, variance, and standard Deviation, and also deal with different types of distributions such as Bernoulli, Binomial, Poisson, Gaussian, etc.

Aims to improve

Here are some aims for improvement that this DSL could achieve in the domain of risk assessment:

1. **Better Decision-Making:** The DSL helps users make smarter choices by quickly providing accurate risk information and calculations.
2. **Faster Analysis:** It makes complex calculations easier and quicker, so users can complete risk assessments without wasting too much time.
3. **User-Friendly:** The DSL is designed for everyone, even those without programming skills, making it easier for more people to do detailed risk analysis.
4. **More Accurate Results:** By using standard methods for calculations, the DSL reduces mistakes, ensuring that risk assessments are reliable and consistent every time.

Planned types and operators:

Sample:

Unordered_map
Check_Continuous
String Distribution

Mean()
Mode()
Median()
Variance()
Standard deviation()
Probability()
Expectation()

Random variable:

1. Discrete Random Variables:
    a. Rolling a Die: (X) = the number shown on the die. Possible Values: 1, 2, 3, 4, 5, 6
    b. Number of Heads in Coin Tosses: (Y) = the number of heads obtained in three coin flips. Possible Values: 0, 1, 2, 3
2. Continuous Random Variables:
    a. Height of a Person: (H) = the height of a randomly selected person. Possible Values: Any value within a range (e.g., 150 cm to 200 cm)
    b. The temperature on a Given Day: (T) = the temperature on a given day. Possible Values: Any real number (e.g., -10°C to 40°C)

Event:

1. Rolling a Die:
    a. Event: The outcome is an even number. (Possible outcomes: 2, 4, 6)
    b. Event: {X > 4}. (Possible outcomes: 5, 6)
2. Height of a Person:
    a. Event: {H > 170 cm}. (Possible outcomes: Any height above 170 cm, such as 171 cm, 175.5 cm, etc.)
    b. Event: {H > 160 cm ^ H < 180 cm}. (Possible outcomes: Any height within the range, like 162.3 cm, 170 cm, 179.8 cm, etc.).

Planned operators:

Example: Consider the following sample on rolling a die

- S = { 1, 2, 3, 4, 5, 6}
- Random Variable X: Outcome of the die.
- **Event A = X > 2** and **Event B = X < 5**.

Union (+):

- The union of two events A and B includes all outcomes that are in either event A, event B, or both.
- Event C = A + B.

Complement (~):

- The complement of event A includes all outcomes in the sample space that are not in event A.
- Event E = ~ A.

Conditional (|):

- The conditional probability of event A given that event B has occurred. It quantifies the likelihood of A occurring under the condition that B is true.
- Event F = (A | B).

Joint or Intersection (^):

- The intersection of two events A and B includes all outcomes that are in both event A and event B.
- Event D = A ^ B.

Multi Staging:

- We will use static variables which are evaluated in compile time.
- Variables related to probabilities or statistical distributions (Probability, Distribution)and parameters (mean, variance) can be declared as static variables so that they are evaluated at compile time.

Possible Optimizations:

- We will try to implement efficient algorithms for calculating mean, median, standard deviation, variance and other methods.
- We will use optimized data structures for storing distributions and calculations (e.g., use arrays or hash maps for access).
- We will not calculate that may not be needed immediately.