# Understanding the SAS FCF Risk Ranking Module

How the out-of-the-box FCF Risk Raking Module scores are calculated

## Contents



CONFIDENTIAL

## Introduction

The SAS Anti-Money Laundering solution includes an optional risk ranking module (i.e. fcf_risk_ranking) which calculates risk at the scenario focus entity level (e.g. account, party, or household). The purpose of this module is to calculate a relative risk score across the entities in order to provide a method to rank order alerts in terms of their relative likelihood of representing a productive alert.

This module is typically used for alert prioritization within the investigative review process. Alerts receive a score between 0 and 999; where higher scores represent a greater likelihood of resulting in a productive alert. It's not the absolute value of an alert's score that is important, but rather how the alert's score compares to the other alerts generated during the same replication period. In other words, it's used as a comparative relative score and NOT an absolute score value.

This module uses Bayes Theorem to calculate the conditional probabilities of the alerts representing money laundering (i.e. productive) and then uses these probabilities to calculate lift factors and ultimately a final risk score for the scenario's focal entity. The purpose of this white paper is to explain the calculations for the alert probabilities and to explain how these values are used to generate the final risk score for an entity.

In addition, this paper discusses methods to set the parameter inputs within the scenario administrator as well as listing some limitations surrounding the use of this module. Information regarding the implementation of the ranking module can be found in the Risk Classification Section of the SAS Anti-Money Laundering Installation, Configuration, and Administration Guide.

## System Inputs

The risk ranking module uses inputs from the scenario administration to calculate a conditional probability for an alert generated by each scenario. If the scenario is marked as Risk Factor, the

conditional probability will be calculated and included in the final score of the focal entity even though an alert will not be generated. Each Scenario has required inputs for a Money Laundering **Bayes Weight** and an **Execution Probability** (Figure 1). There is an optional input for a Terror Financing Bayes Weight.

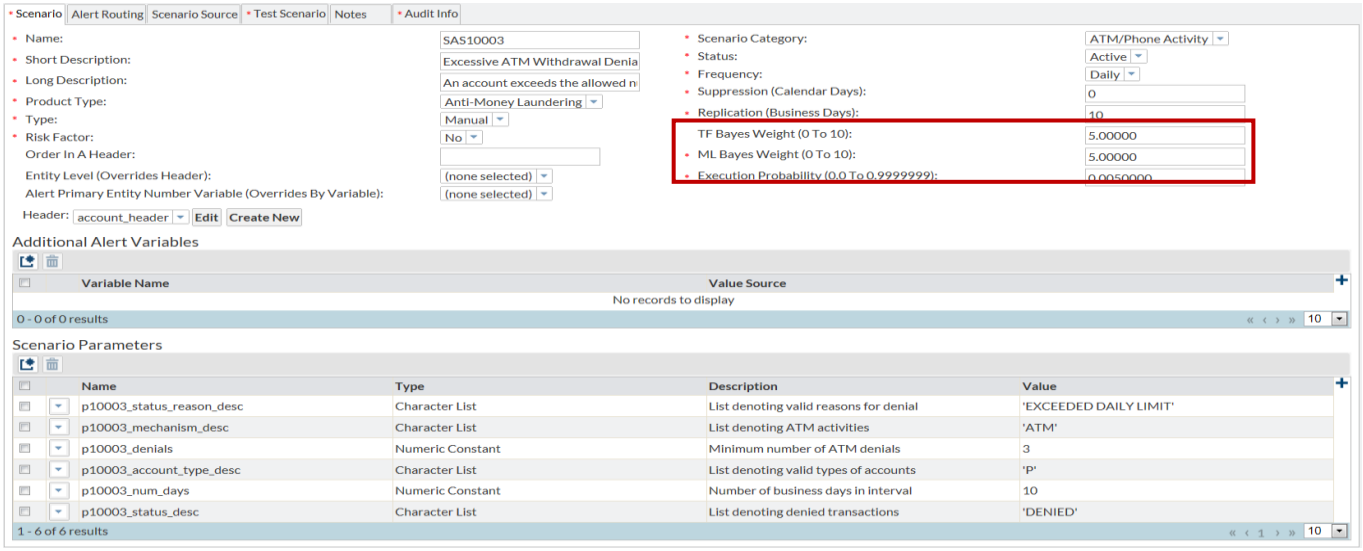


Figure 1. SAS Version 6.2 Anti-Money Laundering Scenario Administrator

CONFIDENTIAL

## Alert Probability Calculation

The alert probability calculation is based on Bayes Theorem:

$$P(ML|RF) = \frac{P(RF|ML)\,P(ML)}{P(RF)}$$

*Where: ML = Money laundering RF = Risk factor or Alert*

The risk ranking module first uses Bayes Theorem to calculate the probability that an event is money laundering (i.e. will be deemed productive) given that a particular scenario generates an alert or risk factor *(note that risk factors are treated the same as alerts within the fcf risk ranking module so the terms are used interchangeably within this document)*. It should be noted that system suppressed alerts are being included in the calculation, so it's the productivity rate calculated using both open and suppressed alerts. This probability is calculated using the following values:

- P(ML) is the probability that the event is money laundering (i.e. productive) and is hard coded with the assumption that the overall probability of money laundering within the entity population is 7%. Note that this assumption doesn't impact the rank order of the scores - only the absolute value of the individual scores.
- P(RF|ML) is the **Bayes Weight** factor from the scenario administration divided by a value of ten. This is the conditional probability that an alert or risk factor will be generated GIVEN that the behavior represents money laundering. There are separate values allowed within the scenario administration for money laundering and terror financing. Note that the Bayes Weight is divided by 10 within the module in order to transform it into a probability (i.e. Bayes Weights of 1, 2, 3, and 10 would represent conditional probabilities of 10%, 20%, 30%, and 100% respectively).

- P(RF) is the **Execution Probability** in the Scenario Definition. This is the probability that a focal entity will generate an alert during the replication period.
- P(ML|RF) is the Alert Probability calculated from the above formula. This is the conditional probability that an alert represents money laundering GIVEN that an alert or risk factor has been generated by the scenario.

### Bayes Weight

The Bayes Weight is a zero to 10 ranking (decimals are allowed) that is assigned for each scenario being registered. The weight reflects the percentage of money laundering suspects which would engage in the behavior described by the scenario. In other words, if the ENTIRE entity population were engaged in money laundering, the Bayes Weight is the percent of the entities on average that the scenario would generate an alert on. A higher ranking implies the scenario is more likely to be associated with money laundering than a lower ranking. A Bayes Weight of zero implies that the scenario provides no indication as to whether or not money laundering exists and the event will not contribute to the final risk ranking score (note that setting parameter values to zero is an easy way to exclude particular scenarios from contributing to the calculation of the risk scores).

The SAS Anti-Money Laundering Scenario Administration Guide recommends that the Bayes weights for all scenarios and risk factors should be determined at the same time to create a consistent scale across the scenarios and risk factors. More information on setting the Bayes Weights can be found within the Scenario Administration Guide.

CONFIDENTIAL

### Execution Probability Weight

The execution probability weight is the probability that a randomly chosen entity will generate an alert for the scenario or trigger the risk factor at least once during the replication period. This parameter should be consistent with regards to the alert volume generated by the scenario or risk factor. More information on setting the Execution Probabilities can be found within the Scenario Administration Guide.

## Score Calculation

The final score calculation for each entity begins with determining the Lift factor for each entity. The lift factor is essentially the probability of money laundering given a risk factor was generated minus the probability of money laundering given that a risk factor wasn't generated.

For each scenario and risk factor, the lift is calculated as follows:

$$Lift = Max\left[P(ML|RF) - \frac{\big(1 - P(RF|ML)\big)P(ML)}{1 - P(RF)}, 0\right]$$

OR

$$Lift = Max\left[P(ML|RF) - \frac{P(ML \cap RF^C)}{P(RF^C)}, 0\right]$$

OR

$$Lift = Max[P(ML|RF) - P(ML|RF^C), 0]$$

The Max Lift score is calculated as the Sum of Lift of all Scenarios:

$$Max\ Lift = \sum_{i=1}^{n} Lift(i)$$

where n = the number of scenarios (*including all segments*)

Each Entity's Lift score is calculated by summing the lift for all scenarios that alerted for the particular entity. If a particular individual scenario alerted multiple times for the entity, the alert score is only included one time. For example, if an entity had one alert on scenario 4, two alerts on scenario 10, and one alert on scenario 14, the entity lift score would be calculated as follows:

$$Entity\ Lift = Lift(4) + Lift(10) + Lift(14)$$

The final rank score for the Entity is calculated such that the final value is similar to a credit score.

$$Score = Int(999 * \left(\frac{Entity\ Lift}{Max\ Lift}\right)^{0.2})$$

CONFIDENTIAL

The ratio of the Entity Lift to Max Lift transforms the ranking to a scale from zero to one. If no alerts were triggered for a given entity, the ratio would be zero and if all alerts triggered (which is generally impossible due to segmentation), the value would be one. Raising the value to the 1/5th root does not change the order but does skew the concentration of ranks to a narrow range toward the upper end.

## Example

The example below illustrates the calculation details for a verity of situations:

| Scenario | P(ML) | Bayes Weight | P(RF\|ML) | Execution Probability [ i.e. P(RF)] | P(ML\|RF) | P(ML & NOT RF) / P(NOT RF) | Lift |
|---|---|---|---|---|---|---|---|
| **SAS10001** | 0.07 | 1 | 0.10000 | 0.00500 | 1.40000 | 0.06332 | 1.33668 |
| **SAS10002** | 0.07 | 3 | 0.30000 | 0.00500 | 4.20000 | 0.04925 | 4.15075 |
| **SAS10003** | 0.07 | 5 | 0.50000 | 0.00500 | 7.00000 | 0.03518 | 6.96482 |
| **SAS10004** | 0.07 | 7 | 0.70000 | 0.00500 | 9.80000 | 0.02111 | 9.77889 |
| **SAS10005** | 0.07 | 9 | 0.90000 | 0.00500 | 12.60000 | 0.00704 | 12.59296 |
| **SAS10006** | 0.07 | 1 | 0.10000 | 0.10000 | 0.07000 | 0.07000 | 0.00000 |
| **SAS10007** | 0.07 | 3 | 0.30000 | 0.10000 | 0.21000 | 0.05444 | 0.15556 |
| **SAS10008** | 0.07 | 5 | 0.50000 | 0.10000 | 0.35000 | 0.03889 | 0.31111 |
| **SAS10009** | 0.07 | 7 | 0.70000 | 0.10000 | 0.49000 | 0.02333 | 0.46667 |
| **SAS10010** | 0.07 | 9 | 0.90000 | 0.10000 | 0.63000 | 0.00778 | 0.62222 |
| | | | | | | **Sum** | **36.37968** |

| Scenerio | Focal Entity 1 | Focal Entity 2 | Focal Entity 3 | Focal Entity 4 | Focal Entity 5 | Focal Entity 6 | Focal Entity 7 |
|---|---|---|---|---|---|---|---|
| **SAS10001** | 1 | 1 | 0 | 1 | 0 | 0 | 0 |
| **SAS10002** | 0 | 1 | 1 | 1 | 0 | 0 | 0 |
| **SAS10003** | 0 | 0 | 0 | 1 | 0 | 1 | 0 |
| **SAS10004** | 0 | 0 | 1 | 1 | 0 | 1 | 0 |
| **SAS10005** | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| **SAS10006** | 0 | 0 | 1 | 1 | 0 | 0 | 0 |
| **SAS10007** | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| **SAS10008** | 0 | 0 | 1 | 1 | 0 | 0 | 1 |
| **SAS10009** | 0 | 0 | 0 | 1 | 1 | 1 | 0 |
| **SAS10010** | 0 | 0 | 1 | 1 | 1 | 1 | 0 |
| **Entity Lift** | **1.33668** | **5.48744** | **14.86298** | **36.37968** | **1.08889** | **17.83261** | **0.31111** |
| **Ratio** | **0.03674** | **0.15084** | **0.40855** | **1.00000** | **0.02993** | **0.49018** | **0.00855** |
| **Score** | **515** | **684** | **835** | **999** | **495** | **866** | **385** |

CONFIDENTIAL

## Limitations

The purpose of this algorithm is to assist in prioritization of investigations by providing a rank ordered score for each entity to be investigated. This algorithm is not designed predict if an entity is involved in Money Laundering. For example, if Entity A has a score of 500 and Entity B has a score of 750, the score shows that Entity B has a higher chance of being involved in Money Laundering than Entity A. The score does NOT imply that Entity B is 50% more likely to be involved in Money Laundering.

This algorithm doesn't account for the severity of the action that caused the alert. For example, if Entity A triggered a scenario by exceeding a threshold with a transaction of $15,000 and Entity B triggered the same scenario by exceeding a threshold with a transaction of $100,000, both Entity A and Entity B would receive the same score for that specific scenario.

This algorithm ONLY considers the risk factors and the particular scenarios that alerted for an entity and the estimated Bayes Weights and Execution Probabilities assigned to those scenarios when calculating the score. Other entity risk characteristics are ignored.

The Max Lift is the sum of all registered scenario/segment lift values. When segmentation is being used the maximum score that an entity can achieve may be well less than 999 due to ineligibility to generate alerts for certain registered scenarios.