A Simple Formula for Calculating Probability for Third-Party Data Breach

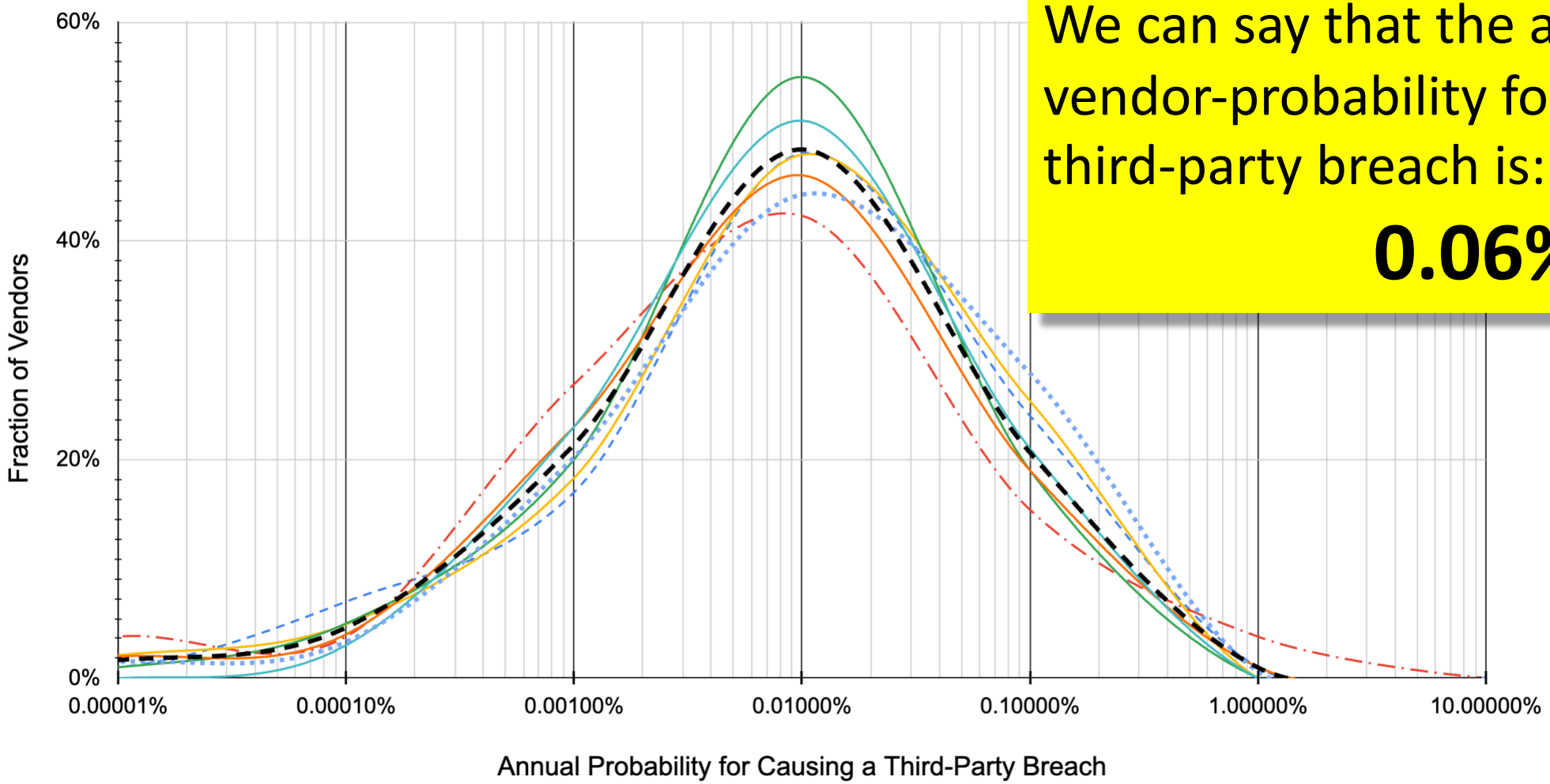
The Problem

Outsourcing is an important value for organizations, whether it is SaaS, PaaS, IaaS, low-code or no-code environments, the use of AI or simply consulting services. Because of the growing amount of outsourcing, the risk for a third-party data breach is a major concern. But organizations have no clue whether the frequency for a third-party data breach is 5-years or 5-thousand years.

The frequency for a third-party data can be calculated using *probability theory*, and an organization can use this calculation to manage this risk in a manner that both allows more outsourcing while also ensuring that a large third-party data breach will never happen.

The following slides derive a simple formula that any organization can use: N x 0.06%. We explain the basis for the constant 0.06% and we show an example calculation.

Because all companies have the same *vendor probability distribution* profile…



We can say that the average annual vendor-probability for causing a third-party breach is:

0.06%

Proof that probability for third-party data breach is N x 0.06%

Let

N Number of vendors

P_i Probability that vendor-*i* will cause a third-party data breach

$P_{Ave}=0.06\%$ Empirically found average probability that a vendor will cause a third-party breach

P_{Cum} Cumulative probability for a third-party data breach

Then

$$P_{Ave} = \sum_{i=1}^{N} P_i \Big/ N$$ Definition for average

$$P_{Cum} = \sum_{i=1}^{N} P_i$$ 3rd-axiom of probability theory

Multiply and divide by N | Isolate P_{Ave} | Substitute P_{Ave} | Substitute value for P_{Ave}

$$P_{Cum} = \sum_{i=1}^{N} P_i = N \times \sum_{i=1}^{N} P_i \Big/ N = N \times \left(\sum_{i=1}^{N} P_i \Big/ N \right) = N \times P_{Ave} = N \times 0.06\%$$ Q.E.D.

Example

A company has 400-vendors that could expose records for 1-thousand or more people and 100-vendors that could expose records for 1-million or more people.

1-thousand or more people

$P_{Cum} = 400 \times 0.06\% = 24\%$ Or once in 4-years, on average

1-Million or more people

$P_{Cum} = 100 \times 0.06\% = 6\%$ Or once in 17-years, on average