Unidad 5: Métodos cuantitativos

BLOQUE II – El análisis del ciberriesgo: enfoques cualitativos y cuantitativos

Grado en Ingeniería de la Ciberseguridad, curso 2022-2023

CONTENIDOS

1. Limitaciones de los métodos cualitativos y métodos cuantitativos.
2. Simulación de Montecarlo.
3. Curva de pérdidas esperadas.
4. FAIR y Value at Risk.

1. Limitaciones de los métodos cualitativos y métodos cuantitativos

- Las metodologías cualitativas se han utilizado durante muchos años y se siguen utilizando en la actualidad.
- Pero cada vez más se están demostrando algunas de sus limitaciones.
- Sirven para reducir incertidumbre, pero en la actualidad se intenta recurrir a métodos cuantitativos para superar estas limitaciones.

1. Limitaciones de los métodos cualitativos y métodos cuantitativos

- Las matrices de riesgo y los mapas de calor permiten priorizar pero no tomar otro tipo de decisiones.
  - Sé que un riesgo es mayor que otro, pero a partir de una matriz de este tipo es complicado tomar decisiones acerca de estrategias de gestión, inversión recomendada en contramedidas, etc.
- Las metodologías cualitativas no suelen cuantificar los impactos traduciéndolos a pérdida económica. Y si lo hacen usan una escala rígida pero no distinguen entre tipos de pérdidas.
  - Y no tiene nada que ver un impacto por pérdida de reputación, con uno por una multa o sanción, o por pérdida de productividad.
  - Un mismo riesgo puede provocar diferentes tipos de pérdidas y se debería tener en cuenta.

1. Limitaciones de los métodos cualitativos y métodos cuantitativos

- Los entregables de las metodologías cualitativas son sencillos de interpretar a primera vista (mapa de riesgos) pero mucho más complicados a la hora de realizar una comunicación en profundidad a la alta dirección.
  - ¿Cuánto nos va a costar este incidente? ¿Cuánto compensa invertir en intentar evitarlo? ¿Cómo agregamos todos estos riesgos si tenemos un año "malo"?

1. Limitaciones de los métodos cualitativos y métodos cuantitativos

- Las metodologías cualitativas se basan en gran medida en taxonomías y catálogos de activos y amenazas que suelen ser rígidos e inamovibles.
  - Y se suelen quedar obsoletos muy rápido.
- Lo mismo ocurre con los criterios para asignar probabilidades e impactos, no suele haber ningún tipo de flexibilidad para enriquecer o actualizar los escenarios que se evalúan o los criterios con los que se estiman probabilidades e impactos.
- Todo esto impide ajustar la metodología a cada organización específica y a sus necesidades en cada momento.

1. Limitaciones de los métodos cualitativos y métodos cuantitativos

Auto-evaluación de madurez	Entrevistas individuales	Grupos de discusión	Paneles de expertos externos	Talleres de calibración
Observación participante	**Recogida de información de fuentes internas y externas**	**Modelado formal**	**Simulación**	**Experimentación**

1. Limitaciones de los métodos cualitativos y métodos cuantitativos

Herramientas que suelen ser útiles:

Cuestionarios	Listas de comprobación	Matrices, tablas, mapas	Método Delphi	Catálogos, listados, glosarios
Árboles y diagramas de amenaza/ataque	Diagramas UML	Esquemas XML	**Documentación sobre incidentes pasados (internos y sectoriales)**	**Informes sobre vulnerabilidades (resultados de escaneos, pentesting, bug bounty, etc.)**
	Productos de inteligencia (internos, compartidos o comerciales)	**Técnicas estadísticas**	**ML, DL e IA**	

1. Limitaciones de los métodos cualitativos y métodos cuantitativos

- Cuidado, porque una crítica habitual a los métodos cualitativos es que son subjetivos, no repetibles.
  - Dependen en gran medida de las metodologías empleadas y de las personas que las aplican.
    - De la información de la que disponen, de sus experiencias pasadas, de sus sesgos.
- Estas limitaciones no se superan por el hecho de pasar a usar métodos cuantitativos.

1. Limitaciones de los métodos cualitativos y métodos cuantitativos

CUALITATIVOS	CUANTITATIVOS
Trabajo con catálogos de activos y de amenazas	Trabajo con escenarios personalizados
Estimación de la probabilidad con escalas (de 1 a 5, o alta/media/baja)	Estimación de la probabilidad basada en datos históricos, experimentos, métodos Bayesianos, etc.
Estimación del impacto, casi siempre técnico, con escalas (de 1 a 5, o alta/media/baja)	Estimación del impacto, casi siempre económico, con un intervalo de confianza del 90%
Entregable basado en matriz o mapa de riegos	Entregable basado en curva de pérdidas esperadas
Gestión del riesgo basada en decisiones que tienen en cuenta la ubicación de cada riesgo en esta matriz o mapa	Gestión del riesgo basada en la comparación de esta curva con la de tolerancia al riesgo y teniendo en cuenta el ROI de las inversiones

2. Simulación de Montecarlo

- La simulación puede ser una herramienta muy potente cuando queremos utilizar métodos cuantitativos.
- Nos vamos a centrar en la simulación basada en el método de Montecarlo, muy utilizada en procesos de gestión del riesgo por las ventajas que ha demostrado tener.
  - De hecho es el método que se utiliza en la metodología FAIR (de las más extendidas en la actualidad) y es una técnica muy utilizada cuando se intentan superar las matrices y mapas tradicionales.

2. Simulación de Montecarlo

- Este método estadístico se emplea cuando es muy difícil calcular el valor de una magnitud y es preferible obtener este valor de manera aproximada.
  - Porque no tenemos una expresión matemática para hacerlo.
  - O porque si la hay, resulta muy costosa de evaluar.
- Si lo expresamos de otra forma, es un método que se puede utilizar para estimar la solución de un problema complejo que depende de parámetros variables.
- Los creadores del método fueron Stanislaw Ulam y John von Neumann, que le dieron a su método el nombre de Montecarlo (la capital de Mónaco) porque allí está uno de los casinos más conocidos del mundo, y las ruletas son generadores de números aleatorios (si no están trucadas) muy sencillos de entender.

2. Simulación de Montecarlo

- Para realizar cualquier tipo de simulación, el primer paso es contar con los datos de entrada.
- En nuestro caso, supongamos que son rangos de pérdidas (impacto mínimo y máximo) y la probabilidad expresada como un porcentaje.
  - Que hemos obtenido tras la fase de análisis.
- Con el método de Montecarlo se hacen numerosos experimentos o pruebas, cada uno representa un año en la vida de la organización.
  - Suele ser suficiente realizar unos miles de experimentos.

2. Simulación de Montecarlo

- Cuando se simula un determinado año:
  - La probabilidad nos dice si ese año ocurre un incidente relacionado con ese escenario de riesgo o no.
  - Si ocurre, para saber el impacto que supone:
    - Se estima un único valor del rango que se tiene (el medio, o el más probable).
    - O se genera un número aleatorio que el 90% de las veces (si estamos trabajando con ese intervalo de confianza en nuestras estimaciones, por ejemplo) esté dentro del rango de pérdidas estimado.

2. Simulación de Montecarlo

- Por ejemplo, para un escenario que tiene que ver con la infección de los portátiles del equipo de desarrollo con un ransomware, hemos estimado una probabilidad de un 53% y un rango de pérdidas de 100.000 y 350.000 euros.
  - Si simulamos 1.000 años, 530 de ellos tendremos infecciones, 470 de ellos no las tendremos.
  - En los 530 que la tenemos, el 90% de las veces las pérdidas estarán entre los 100.000 y los 350.000 euros cada vez, usaremos un número aleatorio para simular la cifra exacta.
  - El 10% restante, estaremos fuera de este intervalo estimado.

2. Simulación de Montecarlo

- Este número aleatorio podría generarse con una distribución uniforme en la que todos los valores de pérdidas tengan exactamente la misma probabilidad de observarse.
- Pero no suele ser lo habitual, en procesos de evaluación del ciberriesgo suelen utilizarse otras funciones de distribución de probabilidad que permitan modelar algo mejor el comportamiento de las pérdidas asociadas a los incidentes de seguridad.

2. Simulación de Montecarlo

- Recordemos que una variable aleatoria es aquella cuyo valor es el resultado de un evento aleatorio y que la función de distribución de probabilidad de una variable aleatoria asigna a cada posible evento la probabilidad de que dicho evento ocurra.
- Probablemente recordéis funciones de distribución como la triangular, la normal, la lognormal, la exponencial, la Beta, la Poisson, la Weibull, la geométrica o la binomial.
- La pregunta clave es ¿de todas las distribuciones de probabilidad que podemos utilizar, cuál es la que mejor describiría el comportamiento de un escenario de ciberriesgo?

2. Simulación de Montecarlo

- En primer lugar debemos tener en cuenta el tipo de fenómeno o problema que queremos simular
- Una de las características principales de los incidentes de ciberseguridad es que puede ser que se produzcan en pocas ocasiones pero que, si se producen, el impacto (pérdidas) que generen sea elevado.
- Teniendo en cuenta esta primera característica, deberíamos utilizar una distribución que pudiera causar pérdidas elevadas en eventos con probabilidad pequeña.

2. Simulación de Montecarlo

- En segundo lugar hay que tener en cuenta que si el incidente de ciberseguridad se produce, genera un impacto (pérdida) en la organización expresada en euros, no puede tener un valor negativo.
- Por tanto, deberíamos utilizar una distribución que sólo pueda devolver valores positivos

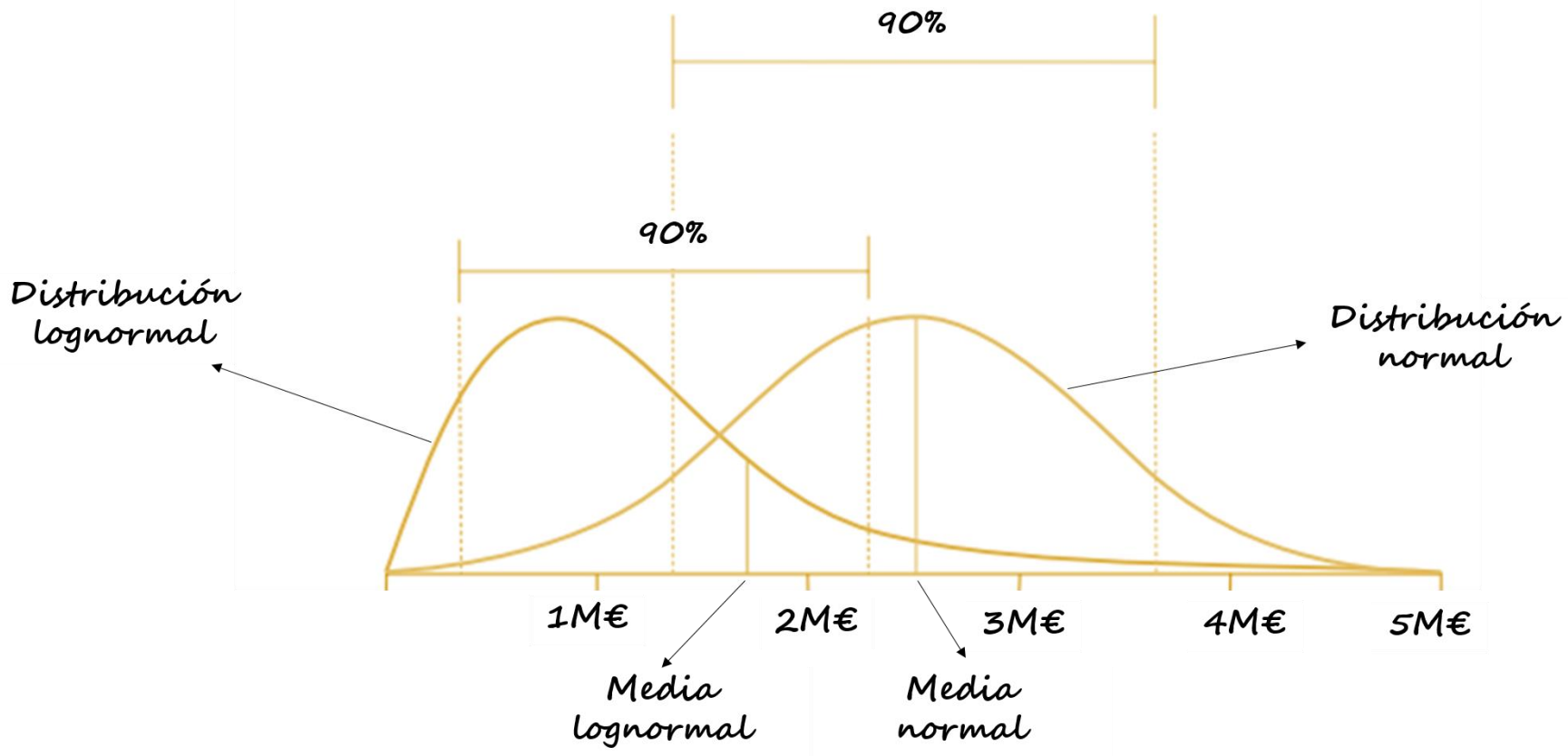
2. Simulación de Montecarlo

- Esto nos deja, habitualmente, con las distribuciones triangular y lognormal.
  - Si se trabaja con intervalos de confianza, como hemos hecho hasta ahora, lo más coherente sería utilizar la lognormal.
  - Si en lugar de trabajar con este concepto, tenemos acotadas las pérdidas mínimas y las máximas, la triangular se ajustaría mejor a nuestras necesidades para la simulación.

2. Simulación de Montecarlo

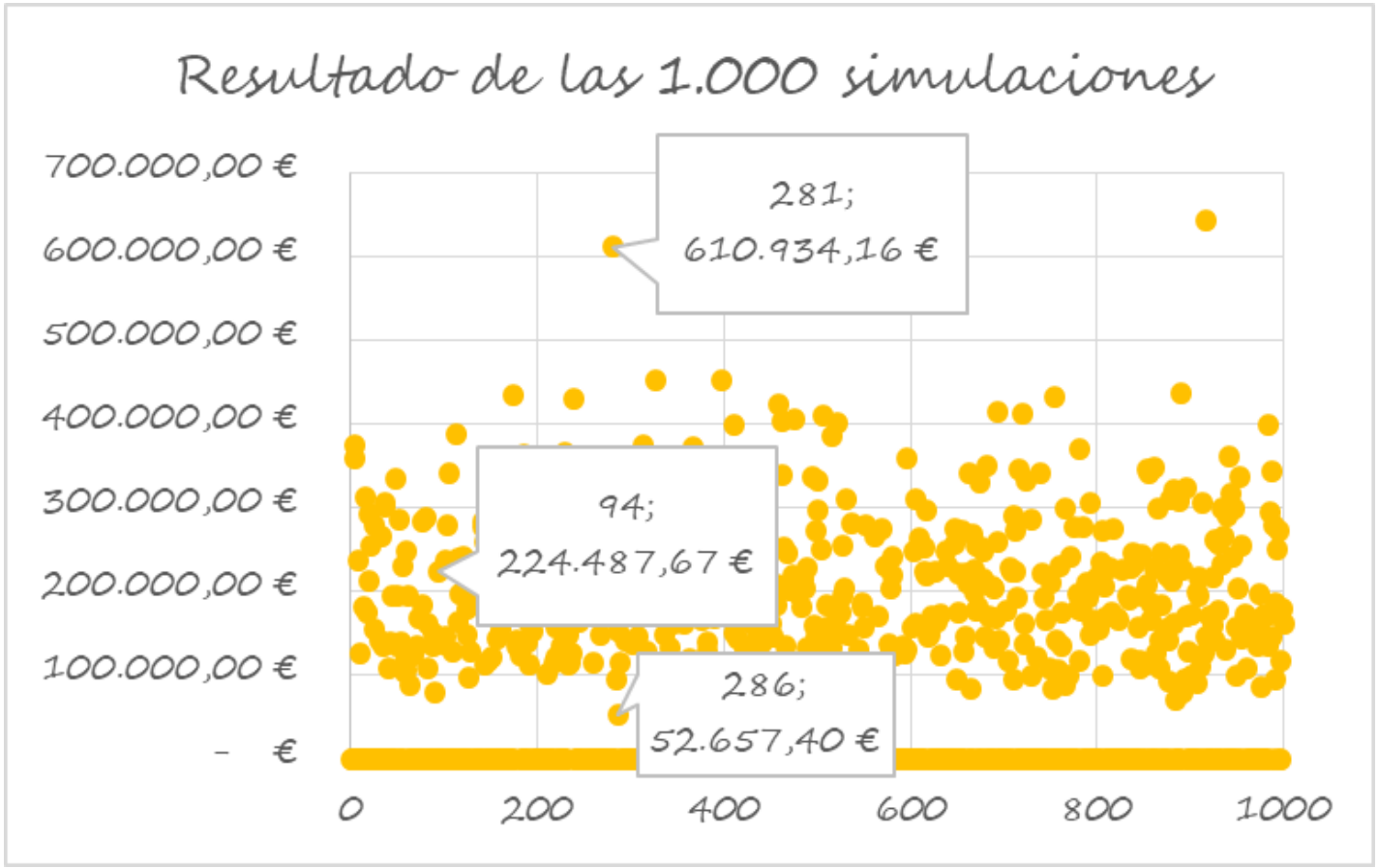
- La distribución lognormal o de Tinaut es una distribución de probabilidad de una variable aleatoria continua cuyo logaritmo se distribuye normalmente.
  - Según una distribución normal, la conocida como campana de Gauss.
- Una variable aleatoria puede modelarse con una distribución lognormal si se obtiene como el producto de muchos pequeños factores independientes.
- Este es otro motivo por el que esta distribución suele ser adecuada, ya que, lo que se modela, que son las pérdidas, suelen ser el producto de muchos pequeños factores independientes.
- La distribución lognormal, como la normal, es una distribución de dos parámetros, la media y la desviación estándar.

2. Simulación de Montecarlo



3. Curva de pérdidas esperadas

- Es muy habitual utilizar la simulación de Montecarlo u otra técnica similar para obtener resultados de este tipo una vez que se han estimado las probabilidades asociadas a los escenarios de riesgo identificados y sus rangos de impactos:

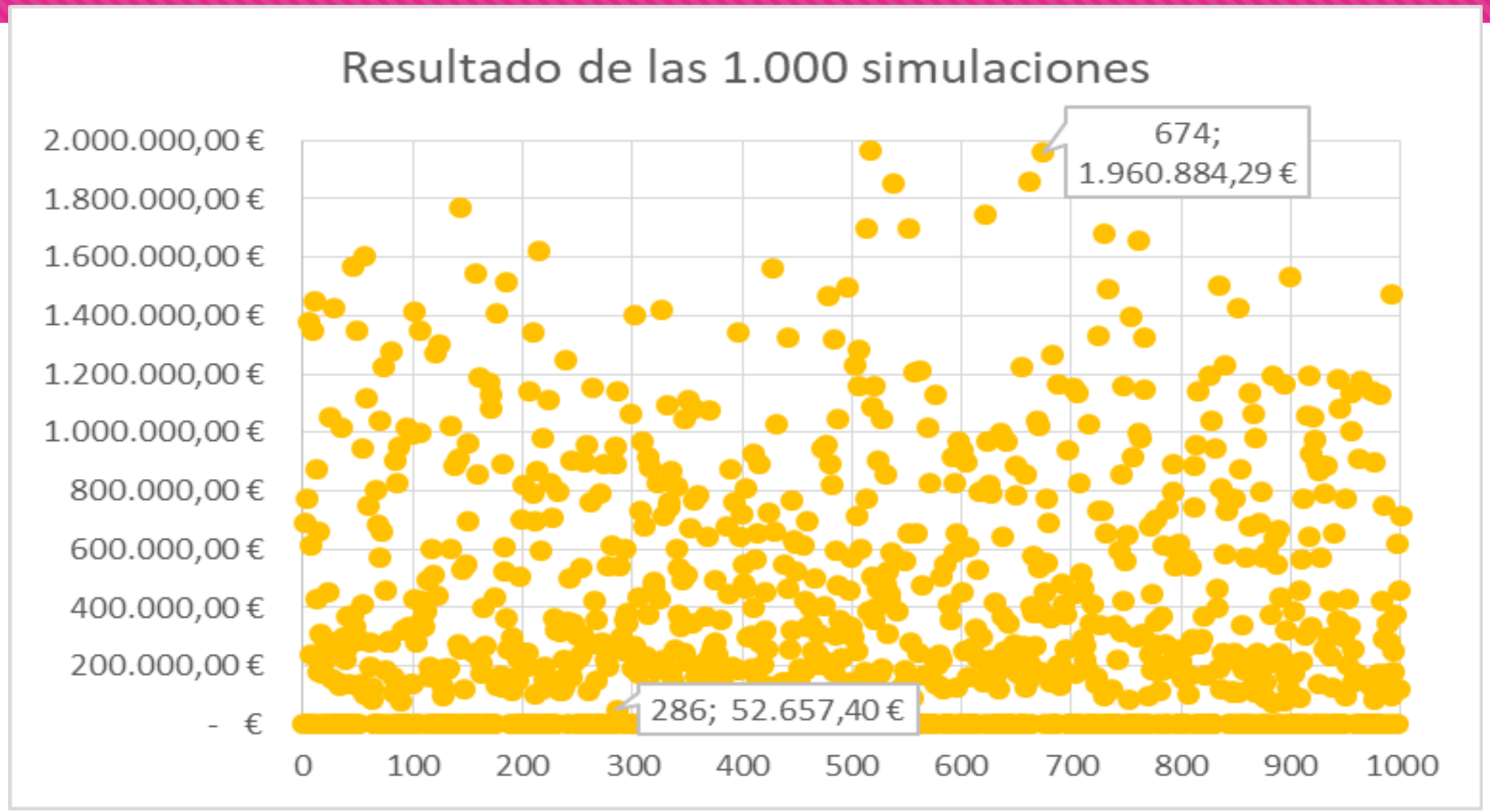


3. Curva de pérdidas esperadas

- Sólo estamos teniendo en cuenta un escenario de ciberriesgo, el del ejemplo que hemos puesto antes de la infección por ransomware.
- Supongamos que estamos evaluando los cuatro escenarios que más nos preocupan:

E	Límite inferior	Límite superior	P
E1	350.000 €	970.000 €	20%
E2	200.000 €	600.000 €	15%
E3	100.000 €	350.000 €	53%
E4	600.000 €	1.200.000 €	10%

3. Curva de pérdidas esperadas



3. Curva de pérdidas esperadas

- Utilizando la tabla donde todos estos datos de las 1.000 simulaciones se encuentran desglosados, podemos calcular la pérdida media agregada asociada a los cuatro escenarios de ciberriesgo analizados.
  - Para ello, tan sólo hay que realizar el promedio de todas las simulaciones mostradas en el gráfico de dispersión.
- En nuestro ejemplo, lo que estamos diciendo es que estos cuatro escenarios causarán anualmente de media un pérdida de 386.826€.

3. Curva de pérdidas esperadas

- Otra herramienta muy útil es la curva de exceso de pérdida o curva de pérdidas esperadas, ya que permite conocer la probabilidad de perder en un año una determinada cantidad "q" o que la cantidad perdida sea mayor que "q".
- En nuestro ejemplo, la curva de exceso de pérdida, nos permite saber que por la ocurrencia e impacto de los cuatro escenarios de ciberriesgo, existe un 36,8% de posibilidades de perder 376.584,15€ o más en un año.
  - ¿Cómo sabemos esto?

3. Curva de pérdidas esperadas

- En nuestro ejemplo, tomamos como referencia 50.000 y 2.000.000 de euros para la pérdida mínima y la pérdida máxima en un año.
- A partir de las simulaciones realizadas y de sus resultados, que tenemos detallados en una tabla de pérdidas, calculamos la probabilidad (como frecuencia, el número total de iteraciones es 1.000) de tener pérdidas iguales o superiores a "c".
  - Por ejemplo, existen 368 iteraciones cuya pérdida es mayor o igual que 376.584,15€, así que la probabilidad es un 36,8%.

3. Curva de pérdidas esperadas

- Podríamos repetir este ejercicio de simulación teniendo en cuenta los controles y contramedidas que ya tenemos desplegados en la organización.
  - Probablemente, por poco nivel de madurez que tengamos, algo estaremos haciendo para gestionar el ciberriesgo. Supongamos que para reducir la probabilidad de algunos escenarios.

E	Límite inferior	Límite superior	P	TCO_ Control	Hipótesis efectividad	P_mit
E1	350.000 €	970.000 €	20%	20.000 €	40%	12%
E2	200.000 €	600.000 €	15%	-	-	-
E3	100.000 €	350.000 €	53%	45.000 €	35%	34,5%
E4	600.000 €	1.200.000 €	10%	60.000 €	45%	5,5%

3. Curva de pérdidas esperadas

- La pérdida media agregada tras mitigación, teniendo en cuenta la efectividad de los controles desplegados, es de 257.128€.
- Realizando una sencilla operación (386.826€-257.128€) llegamos a la conclusión de que la inversión por importe de 125.000€ en mitigación (el TCO anual de los tres controles que tenemos) reduce un 33,53% la pérdida media inherente (antes de mitigación).
  - En concreto se reduce en 129.697€.

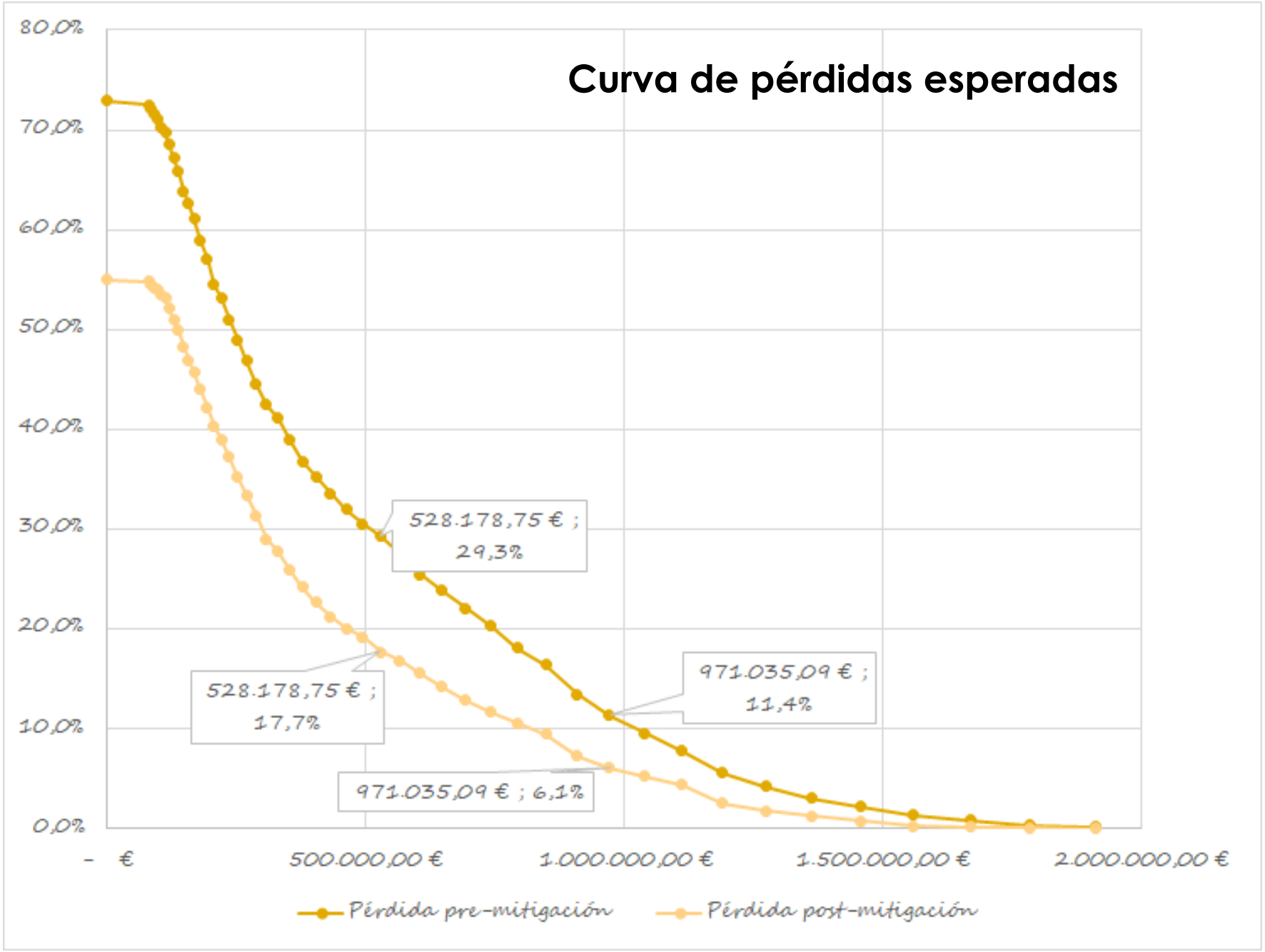

Curva de pérdidas esperadas
80,0%
70,0%
60,0%
50,0%
40,0%
30,0%
20,0%
10,0%
0,0%
528.178,75 € ; 29,3%
528.178,75 € ; 17,7%
971.035,09 € ; 11,4%
971.035,09 € ; 6,1%
- €
500.000,00 €
1.000.000,00 €
1.500.000,00 €
2.000.000,00 €
Pérdida pre-mitigación
Pérdida post-mitigación

Tabla de percentiles

Percentil	Pérdida	Probabilidad de tener esa pérdida o más	Frecuencia con la que se observa esa pérdida o más
P10	- €		
P20	- €		
P30	108.728,59 €	70%	
P40	172.268,00 €	60%	
P50	245.838,77 €	(mediana) 50%	1 vez en 2 años
P75	611.989,51 €	25%	1 vez en 4 años
P90	1.023.927,10 €	10%	1 vez en 10 años
P95	1.211.428,71 €	5%	1 vez en 20 años
P98	1.475.971,84 €	2%	1 vez en 50 años
P99	1.622.404,27 €	1%	1 vez en 100 años
P99,5	1.748.433,55 €	0,50%	1 vez en 200 años
P99,9	1.960.889,59 €	0,10%	1 vez en 1.000 años

4. FAIR y Value at Risk

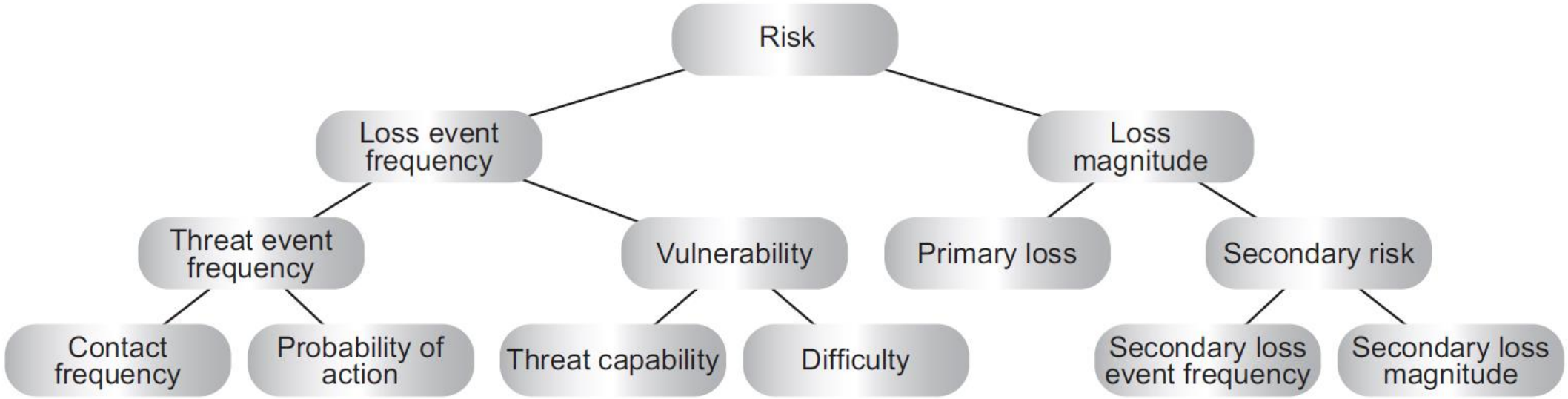
- FAIR (Factor Analysis of Information Risk) es un método cuantitativo que intenta tener en cuenta todos los factores que intervienen en el ciberriesgo para poder cuantificarlo.
  - Se centra mucho en el análisis, poco en la valoración.
  - Que se basa en la simulación de Montecarlo.
- Surgió en el año 2006.
- Está estandarizada por el Open Group.
- Es el método cuantitativo más extendido y el que está dando lugar a otros similares.

4. FAIR y Value at Risk

"Measuring and managing information risk: A FAIR approach" Jack Freund and Jack Jones, 2015

4. FAIR y Value at Risk

Factores del riesgo según FAIR



"Measuring and managing information risk: A FAIR approach" Jack Freund and Jack Jones, 2015

4. FAIR y Value at Risk

- La probabilidad de que haya pérdidas depende de cuatro factores:

Threat event frequency

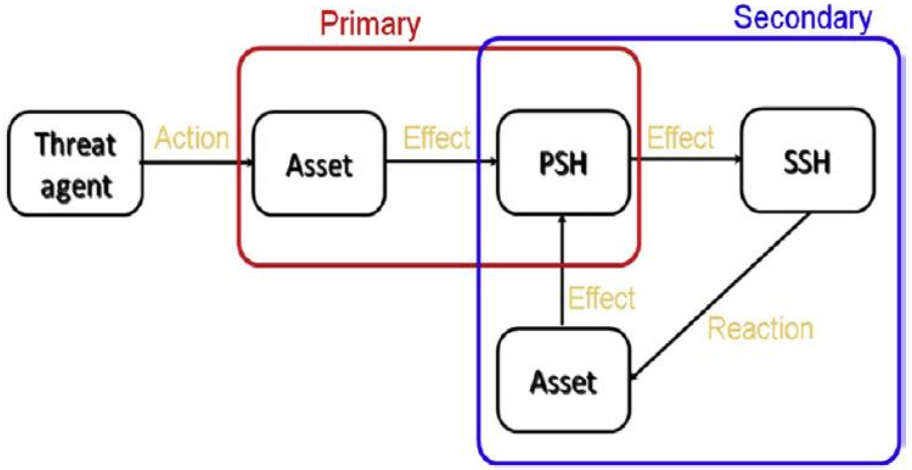
- Frecuencia, en un intervalo de tiempo, con la que la un agente de amenaza entra en contacto con los activos.
- Probabilidad de que un agente de amenaza intente llevar a cabo un ataque con un activo con el que entra en contacto.

Vulnerability

- Capacidades del agente de amenaza.
- Dificultad del ataque y fortaleza de los controles.

4. FAIR y Value at Risk

- El impacto, que se estima con pérdidas económicas, se divide en dos:
  - Impacto o pérdidas primarias (Primary StakeHolders o PSH).
  - Impacto o pérdidas secundarias (Secondary StakeHolders o SSH).



"Measuring and managing information risk: A FAIR approach" Jack Freund and Jack Jones, 2015

4. FAIR y Value at Risk

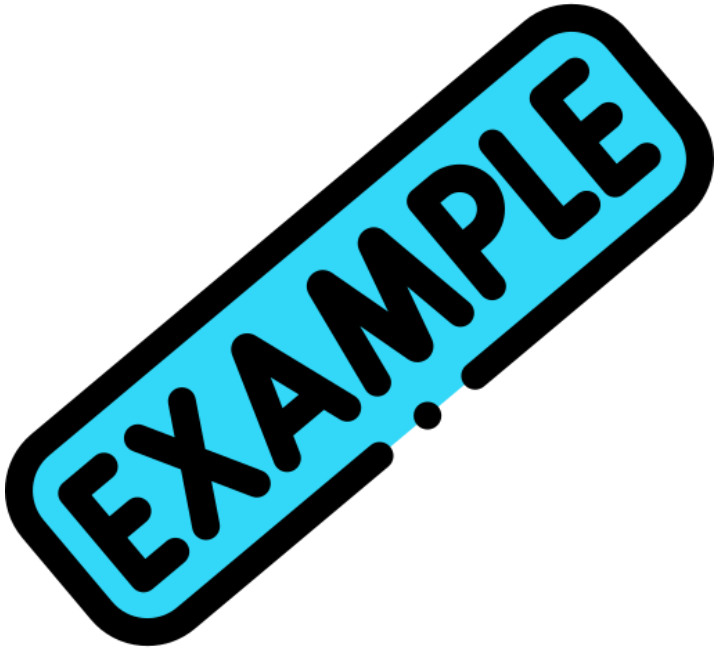
Tipos de pérdida según FAIR

Pérdida de productividad	Gastos de respuesta	Sustitución de componentes
Multas y sanciones	Pérdida de ventaja competitiva	Daño a la reputación

4. FAIR y Value at Risk

- FAIR usa la técnica PERT para estimar probabilidades e impactos:
  - Para cada factor del árbol de riesgo en cada escenario evaluado, pide tres valores a los expertos que participan en el proceso:
    - El mínimo (min).
    - El máximo (max).
    - El más probable (most likely o ML).
- Y la estimación se calcula como (min + 4·ML + max)/ 6 .

4. FAIR y Value at Risk



4. FAIR y Value at Risk

- Este enfoque de FAIR y del resto de métodos cuantitativos que han surgido de él se denomina a menudo Value at Risk.
- Este concepto proviene de las empresas de servicios financieros, se utiliza para cuantificar con métodos estadísticos el riesgo financiero de una empresa o portafolio de inversiones en un periodo de tiempo concreto.
  - Que depende de la probabilidad de que haya pérdidas, la cantidad de pérdidas y el periodo de tiempo.
- Ahora se ha trasladado al ciberriesgo.

PRÁCTICA 1

- Vais a realizar vuestro primer proceso de evaluación del ciberriesgo con métodos cuantitativos. Seguiréis trabajando con la empresa que hemos fundado para la asignatura.
  - Para ello utilizaréis FAIR.
  - Y podréis comparar con el proceso que seguisteis para CORAS.

Para leer e investigar…

1. "A System to Calculate Cyber-Value-at-Risk" Arnau Erola, Ioannis Agrafiotis, Jason R.C. Nurse, Louise Axon, Michael Goldsmith, Sadie Creese. Computers&Security (2022).

Referencias

- Fotografías
  - https://unsplash.com
- Iconos
  - https://www.flaticon.es/
- Figuras:
  - "Dirección de seguridad y gestión del ciberriesgo" Fernando Sevillano y Marta Beltrán. Colección Ciberseguridad, editorial RaMa. 2021.