



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL
DEPARTAMENTO DE INGENIERÍA MATEMÁTICA

INTEGRACIÓN DE LA AVERSIÓN AL RIESGO EN JUEGOS DE STACKELBERG DE
SEGURIDAD

TESIS PARA OPTAR AL GRADO DE
MAGÍSTER EN GESTIÓN DE OPERACIONES

MEMORIA PARA OPTAR AL TÍTULO DE
INGENIERO CIVIL MATEMÁTICO

DANIEL ALFONSO CASTRO HERNÁNDEZ

PROFESOR GUÍA:
FERNANDO ORDÓÑEZ PIZARRO

MIEMBROS DE LA COMISIÓN:
JAIME ORTEGA PALMA
DENIS SAURÉ VALENZUELA

SANTIAGO DE CHILE
2019

RESUMEN DE LA MEMORIA PARA OPTAR
AL TÍTULO DE MAGÍSTER EN GESTIÓN DE OPERACIONES; INGENIERO CIVIL MATEMÁTICO
POR: DANIEL ALFONSO CASTRO HERNÁNDEZ
FECHA: 2 DE AGOSTO DE 2019
PROF. GUÍA: FERNANDO ORDÓÑEZ PIZARRO

INTEGRACIÓN DE LA AVERSIÓN AL RIESGO EN JUEGOS DE STACKELBERG DE SEGURIDAD

La presente tesis trata juegos dentro de la teoría de juegos [Jua08], los cuales son los juegos de Stackelberg. Esta variedad tiene una peculiaridad, y es que, el líder (defensor) hace la primera jugada y el seguidor (atacante) responde en base a ello. Típicamente, los modelos trabajados en este contexto suelen tomar como función objetivo la utilidad esperada. Por otro lado, los jugadores pueden tomar una distribución sobre las acciones posibles dentro del mismo (estrategia mixta). Naturalmente, a este tipo de juegos se le quiere agregar aversión al riesgo con respecto a la respuesta del atacante.

En muchas situaciones se vive un alto riesgo, por ejemplo, cuando un guardia patrulla un condominio, se defiende una página o datos de hackers, etc. Los ejemplos anteriores se pueden modelar como un juego de Stackelberg de seguridad, el que es un caso particular de juego de Stackelberg. En donde, el líder posee una cantidad de recursos, más pequeña que la cantidad de sitios, que utiliza para defender cierta cantidad de lugares. Entonces, se requiere de cierta aversión al riesgo para evitar ataques que causen mucho daño. ¿Cómo se puede agregar aversión al riesgo?, hoy en día no hay claridad sobre qué modelo es fácil de calcular y si dará un buen trade-off entre optimalidad y aversión. En la actualidad económica existen dos grandes funciones que agregan aversión al riesgo, el Value at Risk y el Conditional Value at Risk [RT00]. Un objetivo crucial era resolver la minimización de dichas funciones objetivos en un tiempo práctico y realista. También, existen otras maneras de agregar aversión al riesgo, mediante restricciones o funciones (medida entrópica [Ren16]) diferentes a las ya mencionadas.

En esta tesis se programaron diversos modelos de optimización de funciones objetivos aversivas y modelos de máxima utilidad esperada agregándoles restricciones de aversión al riesgo. Primero se intentó modelar y resolver, el VaR y CVaR, observándose que, computacionalmente, el VaR es más fácil de resolver, debido a la estructura del problema y contexto de la tesis. Además, la incertidumbre está en la respuesta del oponente y la probabilidad de obtener diferentes utilidades, y no en las posibles utilidades que están fijas. El CVaR no se logró resolver, debido a que el tiempo necesario para su resolución es demasiado alto. Estos experimentos se llevaron a cabo bajo un contexto de clara desventaja para el líder, dejándolo expuesto en muchos lugares, dándole pocos recursos.

Uno de los resultados más importantes es que fue posible incorporar exitosamente la aversión sin perder tanto tiempo, ni utilidad esperada. Por otro lado, el modelo VaR se resolvió obteniendo valiosa información en este contexto. Y por último, el Quantal Response, una función de respuesta del seguidor con respecto a la defensa del líder, capta muy bien la racionalidad del mismo, obteniéndose un buen modelo de aversión al riesgo al llevarlo a la práctica.

Una dedicatoria a mis padres, Jimena y Santiago por apoyarme siempre, y en todo.

Agradecimientos

Primeramente doy gracias a la Universidad de Chile por permitir formarme en ella y gracias a todas las personas que fueron partícipes de este proceso dentro de la misma. En particular, me gustaría dar gracias al profesor Fernando Ordóñez, mi guía en este camino, por estar siempre dispuesto a corregir mi tesis e incluso aceptar llamadas los fines de semana. Gracias a Renaud Chicoisne por haber aportado su gran conocimiento a las bases que sentaron este trabajo. Gracias a mi hermana, Pamela Castro, por apoyarme y ayudarme a lo largo de esta travesía. Gracias a mis padres y hermanos, que fueron mis mayores pilares durante esta fase. Gracias a mi pareja por siempre darme aliento y soporte en todo lo posible. Y por último, gracias a todas las amistades formadas en esta serie, sobre todo en primer y cuarto año, porque sin ellas el camino hubiera sido mucho más difícil.

Este es un momento muy especial, un cierre de ciclo, que espero me permita forjar mi largo camino con firmeza. Ojalá, algún día pueda devolver todo el apoyo y amor a aquellas personas que invirtieron su valioso tiempo en mirar mi proyecto de tesis y corregirme; a ellos, asimismo, les agradezco enormemente.

Tabla de Contenido

Índice de Tablas	vi
Índice de Ilustraciones	vii
Introducción	1
1. Marco Teórico	3
1.1. Juego de Stackelberg de Seguridad	3
1.2. Quantal Response	5
1.2.1. Modelo de Máxima Utilidad Esperada	7
1.3. Restricciones de Aversión al Riesgo	8
1.3.1. Restricción de Oportunidad	8
1.3.2. Restricción de Riesgo Distorsionado	8
1.3.3. Restricción de Primer Orden de Dominancia Estocástica	9
1.4. Funciones Objetivos de Aversión al Riesgo	10
1.4.1. Value at Risk	11
1.4.2. Conditional Value at Risk	12
1.4.3. Función de Oportunidad	13
1.4.4. Función de Riesgo Distorsionado	13
1.4.5. Función de Probabilidad	14
2. Metodología e Implementación de los Modelos de Aversión al Riesgo	15
2.1. Contexto	15
2.2. Herramientas Teóricas Utilizadas para Solucionar Problemas con Aversión al Riesgo	17
3. Algoritmos para Problemas con Aversión al Riesgo a la Estrategia del Oponente	22
3.1. Metodología CVaR	22
3.2. Metodología VaR	25
3.3. Metodología FDO	27
3.4. Metodología FRD	28
3.5. Metodología FDP	30
3.6. Metodología MUE	31
4. Resultados y Análisis	33
4.1. Resultados de Funciones de Aversión al Riesgo	34

4.2. Resultados de Restricciones de Aversión al Riesgo	47
5. Conclusiones	53
5.1. Conclusiones Funciones de Aversión al Riesgo	53
5.2. Conclusiones Restricciones de Aversión al Riesgo	55
5.3. Conclusiones Generales	56
Bibliografía	57

Índice de Tablas

1.1. Matriz de Pagos del Líder	4
1.2. Matriz de Pagos del Seguidor	4
1.3. Tabla que Representa el Contexto de Pagos del Líder	4
5.1. Tabla de Valores de Diferentes Parámetros de las Metodologías del Exp 7 . .	53
5.2. Tabla de las Diferencias en Porcentaje c/r a la Metodología MUE del Exp 7	54
5.3. Tabla de Valores de Diferentes Parámetros de las Metodologías del Exp 7 . .	55
5.4. Tabla de las Diferencias en Porcentaje c/r a la Metodología MUE del Exp 7	55

Índice de Ilustraciones

4.1. Gráfico de la Función de Distorsión	34
4.2. VaR y CVaR 0.2 del Experimento 7 sin Restricciones	35
4.3. Utilidad Esperada y Desviación Estándar del Experimento 7 sin Restricciones	36
4.4. Histograma de los 10 Peores Sitios del Experimento 7 sin Restricciones . . .	36
4.5. Función de Distribución Acumulada del Experimento 7 $VaR_{0,2}$	38
4.6. Función de Distribución Acumulada del Experimento 7 $VaR_{0,1}$	39
4.7. Función de Distribución Acumulada del Experimento 7 $VaR_{0,05}$	40
4.8. Función de Distribución Acumulada del Experimento 7 $VaR_{0,01}$	40
4.9. Función de Distribución Acumulada del Experimento 7 FRD	41
4.10. Función de Distribución Acumulada del Experimento 7 FDP50 %	42
4.11. Función de Distribución Acumulada del Experimento 7 FDP20 %	42
4.12. Función de Distribución Acumulada del Experimento 7 FDO1	43
4.13. Función de Distribución Acumulada del Experimento 7 FDO2	44
4.14. Histograma del Soporte de $VaR_{0,2}$ del Experimento 7	44
4.15. Histograma del soporte de $VaR_{0,05}$ del Experimento 7	45
4.16. Histograma del soporte de MUE del Experimento 7	46
4.17. VaR y CVaR 0.2 del Experimento 7 con Restricciones	48
4.18. Utilidad Esperada y Desviación Estándar del Experimento 7 con Restricciones	49
4.19. Histograma de los Peores Sitios del Experimento 7 con Restricciones	49
4.20. Función de Distribución Acumulada del Experimento 7 MUE RRD	50
4.21. Función de Distribución Acumulada del Experimento 7 MUE RPODE	51
4.22. Función de Distribución Acumulada del Experimento 7 MUE RO	52
5.1. VaR y CVar 0.2 del Experimento 3 sin Restricciones	58
5.2. Utilidad Esperada y Desviación Estándar del Experimento 3 sin Restricciones	59
5.3. Histograma de los Peores 10 Sitios del Experimento 3 sin Restricciones . . .	59
5.4. Función de Distribución Acumulada del Experimento 3 FRD	60
5.5. Función de Distribución Acumulada del Experimento 3 FDO1	60
5.6. Función de Distribución Acumulada del Experimento 3 FDO2	61
5.7. Función de Distribución Acumulada del Experimento 3 FDO3	61
5.8. Función de Distribución Acumulada del Experimento 3 FDP100	62
5.9. Función de Distribución Acumulada del Experimento 3 FDP50	62
5.10. Función de Distribución Acumulada del Experimento 3 FDP20	63
5.11. Función de Distribución Acumulada del Experimento 3 $VaR_{0,2}$	63
5.12. Función de Distribución Acumulada del Experimento 3 $VaR_{0,1}$	64
5.13. Función de Distribución Acumulada del Experimento 3 $VaR_{0,05}$	64

5.14. Función de Distribución Acumulada del Experimento 3 $VaR_{0,01}$	65
5.15. Histograma del Soporte de MUE del Experimento 3	65
5.16. Histograma del Soporte de $VaR_{0,2}$ del Experimento 3	66
5.17. Histograma del Soporte de $VaR_{0,05}$ del Experimento 3	66
5.18. VaR y CVar 0.2 del Experimento 3 con Restricciones	67
5.19. Utilidad Esperada y Desviación Estándar del Experimento 3 con Restricciones	67
5.20. Histograma de los Peores 10 Sitios del Experimento 3 con Restricciones . . .	68
5.21. Función de Distribución Acumulada del Experimento 3 MUE RRD	68
5.22. Función de Distribución Acumulada del Experimento 3 MUE RO	69
5.23. Función de Distribución Acumulada del Experimento 3 MUE RPODE	69
5.24. VaR y CVar 0.2 del Experimento 10 sin Restricciones	70
5.25. Utilidad Esperada y Desviación Estándar del Experimento 10 sin Restricciones	71
5.26. Histograma de los Peores 10 Sitios del Experimento 10 sin Restricciones . . .	71
5.27. Función de Distribución Acumulada del Experimento 10 FRD	72
5.28. Función de Distribución Acumulada del Experimento 10 FDO1	72
5.29. Función de Distribución Acumulada del Experimento 10 FDO2	73
5.30. Función de Distribución Acumulada del Experimento 10 FDO3	73
5.31. Función de Distribución Acumulada del Experimento 10 FDP100	74
5.32. Función de Distribución Acumulada del Experimento 10 FDP50	74
5.33. Función de Distribución Acumulada del Experimento 10 FDP20	75
5.34. Función de Distribución Acumulada del Experimento 10 $VaR_{0,2}$	75
5.35. Función de Distribución Acumulada del Experimento 10 $VaR_{0,1}$	76
5.36. Función de Distribución Acumulada del Experimento 10 $VaR_{0,05}$	76
5.37. Función de Distribución Acumulada del Experimento 10 $VaR_{0,01}$	77
5.38. Función de Distribución Acumulada del Experimento 10 $VaR_{0,05}$	77
5.39. Función de Distribución Acumulada del Experimento 10 $VaR_{0,01}$	78
5.40. Histograma del Soporte de MUE del Experimento 10	78
5.41. Histograma del Soporte de $VaR_{0,2}$ del Experimento 10	79
5.42. Histograma del Soporte de $VaR_{0,05}$ del Experimento 10	79
5.43. VaR y CVar 0.2 del Experimento 10 con Restricciones	80
5.44. VaR y CVar 0.2 del Experimento 10 con Restricciones	80
5.45. VaR y CVar 0.2 del Experimento 10 con Restricciones	81
5.46. Función de Distribución Acumulada del Experimento 10 MUE RRD	81
5.47. Función de Distribución Acumulada del Experimento 10 MUE RO	82
5.48. Función de Distribución Acumulada del Experimento 10 MUE RPODE . . .	82

Introducción

Esta tesis trabaja en el contexto de la teoría de juegos [Lue10], la que es una rama de la matemática que estudia las decisiones, en las que para que un individuo tenga éxito, debe tener en cuenta las decisiones tomadas por el resto de los agentes que intervienen en la situación. Como estudio matemático se ha utilizado en la gestión, estrategia, psicología o incluso en biología. En este caso particular se estudia los juegos de Stackelberg de seguridad [Ren16], los que son un caso particular de juego. En donde, el atacante (seguidor) espía al defensor (líder), es decir, el seguidor ataca en base a lo que defendió el líder. Este modelo considera la función objetivo valor esperado, teniendo en cuenta que, los jugadores pueden tomar distribuciones de probabilidad sobre las acciones posibles. A este tipo de modelo se le quiere agregar aversión al riesgo (miedo al peor escenario) con respecto a la respuesta del oponente. Un ejemplo de aversión al riesgo es la medida entrópica en [Ren16]. Hoy en día existen pocos trabajos que consideren aversión al riesgo con respecto a las acciones de los jugadores (existen referencias de juegos de seguridad con aversión al riesgo con respecto a cambios en las utilidades [R T00] u otros parámetros).

Existen diversas situaciones en donde se debe ser averso al riesgo, por ejemplo, en caso de defensa personal (matar o morir), inversiones en portafolios [Sta10] (el riesgo de quebrar), o bien, en un juego de Stackelberg de seguridad, en donde, el miedo a ser atacado en algún lugar importante genera la aversión al riesgo. Lamentablemente, no existe claridad de qué modelos de aversión al riesgo, en un juego de Stackelberg de seguridad, serán fáciles de calcular y permitan el trade-off eficiente entre optimalidad y aversión al riesgo. En consecuencia, nace la pregunta ¿El modelo a utilizar será la varianza, value at risk, conditional value at risk, función objetivo o restricción aversa al riesgo?*. Lo importante en esta tesis será responder a dicha pregunta, obteniéndose información sobre que modelo permitirá resolver el como integrar la aversión al riesgo con respecto a la jugada del oponente, en un juego de Stackelberg de seguridad, de forma eficiente. Conectado con lo anterior, uno de los objetivos más deseados era descubrir si era posible resolver, en un tiempo realista, el Value at Risk. También se quería averiguar si era posible lo mismo para el Conditional Value at Risk. Cabe mencionar que se intentó resolver estos dos modelos utilizando herramientas de funciones continuas convexas [R T00].

En [Ren16] se plantean modelos en que el líder de un juego de Stackelberg de Seguridad considera un objetivo de aversión al riesgo utilizando la medida de riesgo entrópica, además de los pseudo-códigos asociados a la resolución de cada modelo. En colaboración con Renaud Chicoisne se plantearon ideas para la formulación al utilizar medidas de riesgo alternativas

(Value at Risk, Conditional Value at Risk, funciones de utilidad, restricciones de aversión al riesgo) y se discutió sobre los diferentes problemas que cada medida de riesgo tenía computacionalmente. Además de la colaboración, el trabajo realizado en la tesis fue leer la literatura pertinente a juegos de seguridad y medidas de riesgo, luego implementar los códigos de los algoritmos de cada medida de aversión al riesgo, después, se obtuvieron gráficos e información interesante para luego finalizar con un estudio de dichos resultados y presentar las conclusiones (y así poder responder la pregunta *).

Para agregar la aversión al riesgo se trabajó por dos caminos diferentes, el primero fue resolver modelos de minimización de funciones que presentaban aversión al riesgo, generalmente, dándole más peso a posibles desutilidades muy grandes. La segunda opción fue agregarle restricciones de aversión al riesgo al modelo clásico de maximización de utilidad. Además, se realizó un extensivo estudio computacional que evalúa la eficiencia de cada modelo para representar aversión al riesgo y determinar la dificultad de cómputo. Lo anterior permite validar estas formulaciones computacionalmente e incluir modelos con funciones de perturbación y restricciones con aversión al riesgo, por ejemplo, la restricción de dominancia estocástica de primer orden. Las medidas de riesgo escogidas para este estudio se eligieron por dos motivos, el primero fue ver su utilidad en la práctica, o sea, qué información o resguardo útil podrían dar y, además, que tan rápido (a priori) se podría entregar la solución de dichas medidas. Lo segundo fue ver qué tan fácil era darle una interpretación a las medidas, ya que algunas tendían a ser confusas y no era claro en qué escenarios se podrían utilizar en la realidad. Existen varias medidas de riesgo que no se estudiaron en la tesis, por ejemplo, la varianza (como función objetivo) en [R T00] y la restricción de dominancia estocástica de segundo orden o tercer orden en [Lue10]

En el contexto de la tesis se dejó una cantidad de recursos disponibles, de tal forma que el líder o defensor del juego estuviese en clara desventaja con respecto al atacante. Finalmente, se escogió una cantidad suficientemente grande de sitios para generar una clara desventaja, pero no tan grande para poder resolver los modelos en un tiempo práctico.

Capítulo 1

Marco Teórico

1.1. Juego de Stackelberg de Seguridad

En seguridad ciudadana o guardia fronteriza, el *líder* o *defensor* defiende una cantidad finita de sitios u objetivos dentro de muchos posibles (debido a los recursos disponibles). Un juego de Stackelberg [Ren16] se define como un juego en donde el líder decide una estrategia mixta para maximizar la utilidad esperada, teniendo en cuenta que el *seguidor* o *atacante* lo espía. Por lo tanto, el líder elige una jugada de defensa y el seguidor sabrá cuál es la estrategia mixta elegida y en base a eso atacará. Formalmente, el juego de Stackelberg de seguridad posee los siguientes elementos:

- La existencia de un *líder* y un *atacante*.
- Cada jugador posee una matriz de pago, M_L y M_A respectivamente.
- Una cantidad n de sitios que estarán en disputa.
- Una cantidad m ($\leq n$) de sitios que puede defender el *líder*.
- El *atacante* solo puede atacar 1 sitio dentro de los n posibles.

Notemos que si el líder defiende el sitio $i \in \{1, \dots, n\}$ y el seguidor ataca el mismo, el líder es capaz de evitar el ataque y obtiene una recompensa $\bar{R}_i \geq 0$ y el atacante una pérdida $P_i \leq 0$. Al contrario, si se ataca otro sitio que no se defendió, entonces el ataque resulta exitoso y el defensor obtiene una pérdida $\bar{P}_i \leq 0$, mientras que el atacante gana $R_i \geq 0$. Por último, ambos agentes, tanto líder como seguidor, tienen la opción de optar por estrategias mixtas.

Este tipo de modelo existe en la literatura actual de teoría de juegos, por ejemplo, en [Chr09].

Ejemplo 1.1: Un ejemplo de cómo se vería una matriz de pagos y una posible defensa mixta se muestra a continuación:

$$M_L =$$

Líder	Ganancia	Pérdida
Sitio 1	7	-4
Sitio 2	13	-10
Sitio 3	20	-9

Tabla 1.1: Matriz de Pagos del Líder

$$M_S =$$

Seguidor	Ganancia	Pérdida
Sitio 1	6	-9
Sitio 2	10	-20
Sitio 3	20	-18

Tabla 1.2: Matriz de Pagos del Seguidor

Líder defiende \ Seguidor ataca	Sitio 1	Sitio 2	Sitio 3
	Sitio 1	-9 / 7	10 / -10
Sitio 2	6 / -4	-20 / 13	20 / -9
Sitio 3	6 / -4	10 / -10	-18 / 20

Tabla 1.3: Tabla que Representa el Contexto de Pagos del Líder

En este caso en particular tenemos $n = 3$ y se supondrá que $m = 1$. En cualquier juego existen diferentes tipos de estrategias, pero las dos más comunes son las mixtas y puras. Una estrategia pura es cuando un jugador escoge una acción con probabilidad 1. Por ejemplo, si el líder defiende el sitio 1 ($\rho(x) = (1, 0, 0)$) y el atacante elige atacar el sitio tres. Una estrategia mixta, en cambio, permite elegir una distribución sobre las posibles acciones, o sea, tienen probabilidades fraccionarias (y no así las estrategias puras). Por ejemplo, una estrategia mixta (y no pura) sería:

$$\Gamma(\tilde{x}) = (0.25 \quad 0.25 \quad 0.5).$$

Esto significa que cada sitio se verá defendido con frecuencias 0.25 , 0.25 y 0.5, respectivamente. Y esta frecuencia es lo único que sabrá el seguidor, por lo tanto, él podría dar una respuesta a esa defensa con un ataque (0 0.2 0.8). En [Car19] se puede ver más ejemplos o entender de forma general los juegos de Stackelberg de seguridad **Fin de Ejemplo 1.1**

Un uso particular de estos juegos se puede observar en [Bo 13], donde se obtiene una respuesta mixta óptima, la cual, para poder ser llevada a cabo diariamente, se hace un sample (tirar un dado cargado) de una distribución discreta siguiendo las probabilidades de la estrategia mixta óptima. Dado que el seguidor solo sabe la frecuencia, mas no qué sitio se defenderá algún día en particular, ni cuando se defenderá otro sitio, este podría responder con un ataque mixto.

1.2. Quantal Response

El Quantal Response se refiere a cómo el seguidor o atacante decide su estrategia mixta o frecuencia de ataque, al observar la jugada del líder. Además, se supondrá cierto nivel de racionalidad del seguidor, llámese λ . Si el líder defiende cada sitio $i \in \{1, \dots, n\}$ con frecuencia $x_i \in [0, 1]$, entonces las utilidades esperadas de un ataque en el sitio i para el líder y el atacante son, respectivamente, $\bar{U}_i(x_i) = x_i \bar{R}_i + (1 - x_i) \bar{P}_i$ y $U_i(x_i) = x_i P_i + (1 - x_i) R_i$. A partir de lo anterior, el modelo de Quantal Response utiliza la distribución de probabilidad logit para caracterizar la probabilidad con que el seguidor atacará cada sitio:

$$y_i(x) = \frac{e^{\lambda \cdot U_i(x_i)}}{\sum_{j=1}^n e^{\lambda \cdot U_j(x_j)}}$$

Observación: Al acercarse a 0 el valor de λ , el seguidor se vuelve más irracional, en donde, la probabilidad anterior se vuelve la distribución uniforme. Y si $\lambda \rightarrow \infty$ se vuelve más racional, concentrando así, la probabilidad de ataque en aquel sitio que le maximice la utilidad. La racionalidad en un juego de Stackelberg, no necesariamente de seguridad, nos dice: si un agente dentro del juego participa para maximizar su utilidad o no.

Por otro lado, al existir funciones exponenciales de por medio, computacionalmente, el atacante o seguidor será racional cuando λ sobrepase cierto umbral, que -en general- no será muy grande y dependerá de los parámetros del problema.

Con todo lo anterior ya se pueden dar las suposiciones con las cuales se trabajó a lo largo de esta tesis:

- Consideramos a $D(x)$ una variable aleatoria de desutilidad, para el **líder**, bajo el vector de decisión x en $X \subseteq \mathbb{R}^n$.
- El soporte de $D(x)$ es discreto, finito e independiente de la decisión $x \in X$, entonces:

$$Supp(D(x)) := \{(V_v)_{v \in \mathcal{V}}\}.$$

- Los valores del soporte están ordenados de forma creciente, entonces:

$$V_1 \leq V_2 \leq \dots \leq V_{|\mathcal{V}|}.$$

- El vector de decisión $x \in X$ no tiene influencia en los pagos, pero sí en la probabilidad de obtener cierto pago.

En este caso $\{V_v : v = 1, \dots, 2n\} = \{(-\bar{R}_i)_{i \in \{1, \dots, n\}}, (-\bar{P}_i)_{i \in \{1, \dots, n\}}\}$, puesto que se trabaja con desutilidades (del líder).

A continuación se presenta la probabilidad de obtener cierta desutilidad v :

$$p_v(x) := \begin{cases} x_{i(v)} \cdot y_{i(v)}(x) & \text{Si } V_v \text{ es recompensa,} \\ (1 - x_{i(v)}) \cdot y_{i(v)}(x) & \text{Si } V_v \text{ es pérdida.} \end{cases}$$

En donde, la función $i(v)$ indica el sitio correspondiente asociado al valor v , ya sea pérdida o ganancia. El sentido de estas probabilidades es que si el líder defiende un sitio (x_i) versus si el seguidor ataca dicho sitio (y_i), la probabilidad de obtener una ganancia es $x_i y_i$ y una pérdida $(1 - x_i) y_i$.

Proposición 1.1 *En un juego de Stackelberg con un oponente que sigue QR se obtiene que las probabilidades se pueden escribir como:*

$$p_v(x) := \begin{cases} x_{i(v)} \frac{\beta_{i(v)} e^{-\gamma_{i(v)} x_{i(v)}}}{\sum_{j=1}^n \beta_j e^{-\gamma_j x_j}} & \text{Si } V_v \text{ es recompensa} \\ (1 - x_{i(v)}) \frac{\beta_{i(v)} e^{-\gamma_{i(v)} x_{i(v)}}}{\sum_{j=1}^n \beta_j e^{-\gamma_j x_j}} & \text{Si } V_v \text{ es pérdida} \end{cases}$$

En donde R es la recompensa más grande del **atacante**, $\beta_i := e^{\lambda(R_i - R)}$, $\gamma_i := \lambda(R_i - P_i)$.

Demostración: Sea x una forma de defender arbitraria, λ arbitraria, y M_L y M_A matrices de pago dado y n la cantidad de sitios. Entonces:

$$\begin{aligned} y_i(x) &= \frac{e^{\lambda U_i(x_i)}}{\sum_{j=1}^n e^{\lambda U_j(x_j)}} \\ &= \frac{e^{\lambda(x_i P_i + (1-x_i) R_i)}}{\sum_{j=1}^n e^{\lambda(x_j P_j + (1-x_j) R_j)}} \quad \setminus \cdot \frac{e^{-\lambda R}}{e^{-\lambda R}} \\ &= \frac{e^{\lambda(R_i - R)} e^{-x_i \lambda(-P_i + R_i)}}{\sum_{j=1}^n e^{\lambda(R_j - R)} e^{-x_j \lambda(-P_j + R_j)}} \\ &= \frac{\beta_i e^{-x_i \gamma_i}}{\sum_{j=1}^n \beta_j e^{-x_j \gamma_j}}. \quad \square \end{aligned}$$

La razón de esta transformación es bajar el valor a las funciones exponenciales, ya que, computacionalmente, generan errores numéricos, debido a que éstas crecen rápido. Por último, notemos que $\beta_i \in (0, 1]$ y $\gamma_i \geq 0 \quad \forall i \in \{1, \dots, n\}$.

Observación: Usando el **ejemplo 1.1** y la definición anterior, se puede notar que la decisión afecta la probabilidad de obtener cierto resultado, independiente de si es pérdida o ganancia. Sea $n := 3$ y $m := 1$, y $\lambda := 0.5$, entonces $\beta_1 = e^{-7}$, $\beta_2 = e^{-5}$, $\beta_3 = 1$, $\gamma_1 = 7.5$, $\gamma_2 = 15$, $\gamma_3 = 19$. Además, el orden de los sitios en \mathcal{V} se asocia de la siguiente manera; $v=1$ o $v=5$ se asocia al sitio 3, $v=2$ o $v=6$ al sitio 2 y $v=3$ o $v=4$ al sitio 1, donde los primeros tres elementos de \mathcal{V} son ganancias y los últimos tres son pérdidas. Si defendemos como en el

ejemplo 1.1 ($\Gamma(\tilde{x})$), entonces las probabilidades de los resultados son:

$$\begin{aligned} p_1(\tilde{x}) &\approx 0,4784, \\ p_2(\tilde{x}) &\approx 0,0008, \\ p_3(\tilde{x}) &\approx 0,0315, \\ p_4(\tilde{x}) &\approx 0,0105, \\ p_5(\tilde{x}) &\approx 0,4784, \\ p_6(\tilde{x}) &\approx 0,0004. \end{aligned}$$

Notemos que es muy probable que el resultado de nuestra desutilidad sea algún elemento en \mathcal{V} , de etiquetas $v = 1$ o $v = 5$, los cuales representan el sitio 3. También existe una probabilidad baja de que obtengamos alguna ganancia en el sitio uno. Por último, es poco probable que obtengamos como ganancia el sitio dos, o pérdida del sitio dos. Notemos que el sitio tres es muy atractivo para el atacante, puesto que le reporta mucha ganancia en caso de atacarlo exitosamente. Incluso, aunque le reporte mucha pérdida al seguidor, debido a su racionalidad, buscará maximizar su utilidad y, por lo tanto, lo hará tomar esa decisión. Por otro lado, el sitio dos tiene mucha más pérdida que ganancia, lo cual lo hace menos apetecible para el seguidor. Similarmente, el sitio uno tiene mayor pérdida que ganancia para seguidor, pero no es tanto la diferencia, por lo mismo, es más deseado (teniendo en cuenta que dichos sitios se defienden mucho).

Con toda la información acerca de este tipo de juegos, se presenta a continuación el problema de maximización de utilidad esperada, al cuál se le agregará aversión al riesgo desde distintos ángulos.

1.2.1. Modelo de Máxima Utilidad Esperada

Definición 1.2 *Dado una cantidad de sitios n , la cantidad de recursos disponibles m y matrices de pago para cada jugador, la máxima utilidad esperada es:*

$$\max_{x \in X} \sum_{i=1}^n y_i(x) \bar{U}_i(x_i)$$

En donde $X := \left\{ x \in [0, 1]^n : \sum_{i=1}^n x_i \leq m \right\}$, y además, $\delta_i := \bar{R}_i - \bar{P}_i$. Y si se toma en cuenta la transformación que se hizo a $y_i(x)$, entonces el problema se puede escribir:

$$\max_{x \in X} \frac{\sum_{i=1}^n \beta_i e^{-\gamma_i x_i} (\bar{P}_i + \delta_i x_i)}{\sum_{i=1}^n \beta_i e^{-\gamma_i x_i}}$$

Observación: Este modelo es sencillo de resolver utilizando búsqueda binaria y resolución de problemas convexos, ya que la única dificultad que tiene es la función objetivo no lineal fraccionaria. Por otro lado, notemos que las restricciones son lineales. Este problema se linealiza de cierta forma (más adelante se hablará de ello) y luego se resuelven problemas con funciones objetivo convexas, y restricciones lineales.

1.3. Restricciones de Aversión al Riesgo

Como se mencionó anteriormente se le quiere agregar aversión al riesgo con respecto a la respuesta del contrincante, lo que, hasta ahora, no se había hecho numéricamente. Una forma de agregarle aversión al riesgo es mediante restricciones que obliguen a preocuparse de los peores escenarios. A continuación se presentan tres restricciones que logran influir al líder a tomar una postura más aversa al riesgo.

1.3.1. Restricción de Oportunidad

Definición 1.3 *La restricción de Oportunidad, dado un umbral $V_{2n} \geq \tilde{V} \geq V_1$ y una tolerancia $\varepsilon \in [0, 1]$, se define como sigue:*

$$\mathbb{P} \left(D(x) \geq \tilde{V} \right) \leq \varepsilon.$$

Si tomamos $\bar{v} := \arg \min_{v \in \mathcal{V}} \left\{ V_v : V_v \geq \tilde{V} \right\}$, entonces lo anterior se traduce en:

$$\sum_{v \geq \bar{v}} p_v(x) \leq \varepsilon.$$

Notemos que si el umbral lo elegimos con algún elemento V_v de \mathcal{V} , entonces se estará exigiendo que los sitios a partir del índice v se defiendan casi en su totalidad, obviamente, colocando un ε lo suficientemente pequeño.

1.3.2. Restricción de Riesgo Distorsionado

Definición 1.4 *La restricción de Riesgo Distorsionado, dado una función $d : \mathbb{R} \rightarrow \mathbb{R}$ de desutilidad, creciente y biyectiva, se define como sigue:*

$$\rho(D(x)) \leq \tilde{\rho}.$$

En donde $\rho(x) := d^{-1}(\mathbb{E}(d(x)))$ y $\tilde{\rho} \in \mathbb{R}$, que de alguna forma debe cumplir cierta restricción para que tenga validez, o sea útil la restricción.

Usando el hecho de que ' d ' es biyectiva y creciente, y además que la desutilidad es una variable aleatoria discreta, lo anterior se puede escribir como sigue:

$$\sum_{v \in \mathcal{V}} p_v(x) d(V_v) \leq d(\tilde{\rho}).$$

Se puede observar que esta restricción está relacionada con la restricción anterior, pero el contexto y cómo se presentan es bastante distinto, puesto que en la primera restricción se desea que mi umbral de probabilidad sea pequeño, mientras que en la actual se busca minimizar las probabilidades distorsionadas bajo cierto umbral.

1.3.3. Restricción de Primer Orden de Dominancia Estocástica

Definición 1.5 Se dice que una variable aleatoria de pérdida Z domina estocásticamente a otra variable de pérdida T en primer orden si:

$$F_Z(t) \geq F_T(t), \forall t \in \mathbb{R}$$

En donde F es la función de distribución acumulada.

Definición 1.6 Sea $\tilde{x} \in X$ fijo, $D(\tilde{x})$ la desutilidad asociada fija, $x \in X$ la variable independiente y $D(x)$ su desutilidad asociada. Entonces la restricción de Primer Orden de Dominancia Estocástica se define como sigue:

$$\begin{aligned} F_{D(x)}(V_{\tilde{v}}) &\geq F_{D(\tilde{x})}(V_{\tilde{v}}), \forall \tilde{v} \in \mathcal{V}, \\ \iff \sum_{v \geq \tilde{v}+1} p_v(x) &\leq 1 - F_{D(\tilde{x})}(V_{\tilde{v}}), \forall \tilde{v} \in \mathcal{V}. \end{aligned}$$

Notemos que esto quiere decir que progresivamente en los posibles resultados la desutilidad será mejor, por ejemplo, la probabilidad de que mi desutilidad sea V_v o menos, será más grande que la otra. Hay que resaltar que no significa que resultado a resultado sea mejor, si no que en conjunto lo es.

Es una manera de agregar aversión al asegurar que todos los resultados paso a paso serán más probables, lo que implica que se tenga que usar más recursos de lo habitual, antes de llegar a los peores escenarios, asegurando así una mejor defensa, descuidando aquellos sitios con pérdidas medianas o bajas.

1.4. Funciones Objetivos de Aversión al Riesgo

Los modelos que tienen en cuenta el miedo al peor escenario son varios, pero los dos que resaltan más hoy en día en muchos contextos son el VaR y el CVaR. Primero se mostrará CVaR y VaR en el mundo continuo, para luego dar a conocerlos en el mundo discreto.

El contexto en el cual se definen involucra a una función $f(x, y)$ en donde x es una elección en $[0, 1]^n$ e y es un vector aleatorio en \mathbb{R}^m . Dicha función puede ser de muchas formas, pero para darle un sentido, se dirá que es una función de pérdida, además de que cada elección cambia el soporte de la función de pérdida. También se supondrá que y posee una densidad $p(y)$ que no cambia según y , por lo tanto, la aleatoriedad vendrá dada por las elecciones en $[0, 1]^n$.

A continuación se define la probabilidad de que $f(x, y)$ no supere cierto nivel α :

$$\Phi(x, \alpha) = \int_{f(x, y) \leq \alpha} p(y) dy.$$

A partir de esto se define el VaR_α y CVaR_α :

$$\begin{aligned} \text{VaR}_\alpha(x) &= \min \{r \in \mathbb{R} : \Phi(x, r) \geq \alpha\}, \\ \text{CVaR}_\alpha(x) &= (1 - \alpha)^{-1} \int_{f(x, y) \geq \text{VaR}_\alpha(x)} f(x, y) p(y) dy. \end{aligned}$$

Lo que se puede deducir del VaR_α es que se está buscando minimizar que la pérdida no supere cierto umbral (en probabilidad), mientras que el CVaR_α calcula el promedio de la pérdida que está sobre el VaR_α y luego lo regula acorde a α . Por último, se presenta la clave para encontrar estos valores de forma efectiva:

$$F_\alpha(x, r) = r + (1 - \alpha)^{-1} \int_{y \in \mathbb{R}^m} [f(x, y) - r]_+ p(y) dy.$$

En donde $[\cdot]_+$ representa la función parte positiva.

Teorema 1.7 *Como función de r , $F_\alpha(x, r)$ es convexa y diferenciable. Luego el CVaR_α asociado a una elección x se puede calcular minimizando como sigue:*

$$\text{CVaR}_\alpha(x) = \min_{r \in \mathbb{R}} F_\alpha(x, r).$$

Además el mínimo de $F_\alpha(x, \cdot)$ se alcanza en $\text{VaR}_\alpha(x)$.

Observación: Utilizando el hecho de que al minimizar una función de dos variables, usando ambas variables, es lo mismo que minimizar primero con una y luego con la otra. Se puede mostrar que:

$$\min_{x \in X} \text{CVaR}_\alpha(x) = \min_{(x, r) \in X \times \mathbb{R}} F_\alpha(x, r).$$

Donde $X := \left\{ x \in [0, 1]^n : \sum_{i=1}^n x_i \leq 1 \right\}$. Para demostrar el teorema anterior se utilizan diferentes definiciones de funciones auxiliares, pero lo más importante que hay detrás de esto es que se utilizan el teorema de convergencia dominada y el teorema de convergencia monótona, los que requieren que el espacio de medida se mantenga constante.

Sin embargo, en esta tesis existen dos supuestos distintos a los mencionados anteriormente. Primero, las distribuciones son discretas y el soporte de cada distribución es finito e igual para todas las distribuciones. Segundo, la decisión del líder al defender afecta la distribución de desutilidad que se observa, o sea, las probabilidades de que ocurran los resultados en \mathcal{V} cambian. Así, se van obteniendo diferentes espacios de medida, pues estos van variando con cada elección del líder y, por lo tanto, no se pueden utilizar las herramientas ya conocidas de teoría de la medida [Ari18]. Con esto pasaremos a definir el VaR y CVaR en el contexto de la presente tesis.

1.4.1. Value at Risk

Definición 1.8 *El 'Value at Risk ε ' de una variable aleatoria de desutilidad $D(x)$ se define como sigue:*

$$VaR_\varepsilon(D(x)) := \inf \{ t : F_{D(x)}(t) \geq 1 - \varepsilon \}$$

En donde $F_{D(x)}(t) = \sum_{v \leq \bar{v}} p_v(x)$, con \bar{v} tal que $V_{\bar{v}} \leq t$ y $V_{\bar{v}+1} \geq t$.

Proposición 1.9 *Los únicos posibles valores que el $VaR_\varepsilon(D(x))$ puede tomar son $(V_v)_{v \in \mathcal{V}}$.*

Demostración: Primero mostremos que existe. Sea $\varepsilon > 0$ (≤ 1), si tomamos $t = V_{2n}$, entonces $F_{D(x)}(V_{2n}) = 1 \geq 1 - \varepsilon$, lo que significa que $\{t : F_{D(x)}(t) \geq 1 - \varepsilon\} \neq \emptyset$.

Luego si tomamos $t < V_1$, entonces $F_{D(x)}(t) = 0 < 1 - \varepsilon$, por lo tanto, no pertenecen al conjunto. Luego V_1 es una cota inferior de $\{t : F_{D(x)}(t) \geq 1 - \varepsilon\}$. Con esto se puede concluir que el ínfimo existe.

Para finalizar sea $t \in [V_1, V_{2n}]$, con $t \neq V_j \forall j \in \{1, \dots, 2n\}$. Se sabe que t está encajonado por dos valores del soporte, esto quiere decir que $\exists i \in \{1, \dots, 2n\}$ tal que $V_i < t < V_{i+1}$, notemos que si t es válido, teniendo en cuenta que $F_{D(x)}(t) = F_{D(x)}(V_i)$, lo será también V_i , el cual es menor que t , y recordemos que se está minimizando, entonces dicho t se descarta. Por otro lado, si t no es válido, entonces se analiza el siguiente tramo $[V_{i+1}, V_{i+2}]$. Con esto se eliminan todos los posibles valores entre cada soporte de la desutilidad, quedando únicamente $(V_v)_{v \in \mathcal{V}}$ como valores disponibles para minimizar $\{t : F_{D(x)}(t) \geq 1 - \varepsilon\}$. \square

1.4.2. Conditional Value at Risk

Definición 1.10 El 'Conditional Value at Risk ε ' de una variable aleatoria de desutilidad $D(x)$ se define como sigue:

$$CVaR_\varepsilon(D(x)) := \inf_{t \in \mathbb{R}} \{t + \varepsilon^{-1} \mathbb{E}([D(x) - t]_+)\}.$$

Proposición 1.11 El ínfimo se alcanza en el $VaR_\varepsilon(D(x))$.

Demostración: sea $\varepsilon \in [0, 1]$ y $x \in X$ arbitrarios. Primero mostremos que el valor de t debe estar en $[V_1, V_{2n}]$. Sea $f(t) := t + \varepsilon^{-1} \mathbb{E}([D(x) - t]_+)$, luego elegimos un $t < V_1 \leq 0$, entonces $f(t) = t(1 - \varepsilon^{-1}) + \varepsilon^{-1} \mathbb{E}(D(x))$, recordar que la sumatoria de todas las probabilidades suman 1. Notemos que $t(1 - \varepsilon^{-1}) \geq V_1(1 - \varepsilon^{-1})$, puesto que $(1 - \varepsilon^{-1}) \leq 0$. Por lo tanto, cualquier t más pequeño que V_1 se puede descartar. Tomemos ahora un $t > V_{2n}$, entonces $f(t) = t$, ya que la desutilidad $D(x)$ no puede ser mayor que V_{2n} , luego $f(t) = t \geq V_{2n}$, por lo tanto, cualquier t mayor que V_{2n} se puede descartar.

Supongamos que $VaR_\varepsilon(D(x)) \neq V_1, V_{2n}$, entonces $\exists i \in \{2, \dots, 2n - 1\}$ tal que $VaR_\varepsilon(D(x)) = V_i$. Con todo esto tomemos un $t > V_i$, sea \bar{v} tal que $V_{\bar{v}} \leq t \leq V_{\bar{v}+1}$. Luego:

$$\begin{aligned} f(t) &= t(1 - \varepsilon^{-1}) + \varepsilon^{-1} \sum_{v \geq \bar{v}+1} p_v(x) V_v + \varepsilon^{-1} t F_{D(x)}(V_{\bar{v}}), \\ f(V_i) &= V_i(1 - \varepsilon^{-1}) + \varepsilon^{-1} \sum_{v \geq i+1} p_v(x) V_v + \varepsilon^{-1} V_i F_{D(x)}(V_i), \\ &\implies \\ f(t) &= t(1 - \varepsilon^{-1}) + \varepsilon^{-1} \sum_{v \geq i+1} p_v(x) V_v + \varepsilon^{-1} \overbrace{\sum_{v=i+1}^{\bar{v}} p_v(x) (t - V_v)}^{(*)} + \varepsilon^{-1} t F_{D(x)}(V_i). \\ &\text{Dado que } t \geq V_{\bar{v}} \geq V_v \forall v \in \{i+1, \dots, \bar{v}\}. \text{ Aplicar en } (*), \\ &\implies \\ f(t) &\geq t(1 - \varepsilon^{-1}) + \varepsilon^{-1} \sum_{v \geq i+1} p_v(x) V_v + \varepsilon^{-1} t F_{D(x)}(V_i), \\ &= t \overbrace{(1 - \varepsilon^{-1} + \varepsilon^{-1} F_{D(x)}(V_i))}^{(**)} + \varepsilon^{-1} \sum_{v \geq i+1} p_v(x) V_v \quad (1). \end{aligned}$$

Lo que está en $(**)$ es positivo, ya que $V_i = VaR_\varepsilon(D(x))$, luego

$$\geq V_i(1 - \varepsilon^{-1} + \varepsilon^{-1} F_{D(x)}(V_i)) + \varepsilon^{-1} \sum_{v \geq i+1} p_v(x) V_v = f(V_i).$$

Notemos que podría ocurrir que $\bar{v} = i$, en ese caso no es necesario el paso en $(*)$, y se pasa directamente al paso $(**)$, ya $f(t)$ valdría lo que hay en la línea (1).

Ahora tomemos un $V_{i-1} \leq t < V_i$. Luego:

$$\begin{aligned} f(t) &= t(1 - \varepsilon^{-1}) + \varepsilon^{-1} \sum_{v \geq i} p_v(x) V_v + \varepsilon^{-1} t F_{D(x)}(V_{i-1}), \\ &\geq \varepsilon^{-1} \sum_{v \geq i} p_v(x) V_v + V_i \overbrace{(1 - \varepsilon^{-1} + \varepsilon^{-1} F_{D(x)}(V_{i-1}))}^{(***)}, \end{aligned}$$

este paso está justificado ya que lo que está en $(***)$ es negativo, gracias a $F_{D(x)}(V_{i-1}) < 1 - \varepsilon$,

$$\begin{aligned} &= \varepsilon^{-1} \sum_{v \geq i} p_v(x) V_v - \varepsilon^{-1} p_i(x) V_i + V_i(1 - \varepsilon^{-1}) + \varepsilon^{-1} V_i F_{D(x)}(V_{i-1}) + \varepsilon^{-1} p_i(x) V_i, \quad \text{se sumó 0} \\ &= \varepsilon^{-1} \sum_{v \geq i+1} p_v(x) V_v + V_i(1 - \varepsilon^{-1}) + \varepsilon^{-1} V_i F_{D(x)}(V_i) = f(V_i). \end{aligned}$$

Análogamente se puede mostrar que tomando $V_{i-2} \leq t < V_{i-1}$ se obtiene que $f(t) \geq f(V_{i-1})$. Aplicando esto recursivamente se concluye que si $t < V_i$, teniendo en cuenta que existe un \bar{v} tal que $V_{\bar{v}} \leq t \leq V_{\bar{v}+1}$, se obtiene gracias a lo anterior $f(t) \geq f(V_{\bar{v}+1}) \geq f(V_{\bar{v}+2}) \geq \dots \geq f(V_i)$.

Por último, si $VaR_\varepsilon(D(x)) = V_1$ o $VaR_\varepsilon(D(x)) = V_{2n}$, usando los mismos argumentos antes expuestos se concluye que el $VaR_\varepsilon(D(x))$ es el punto en donde $f(\cdot)$ alcanza su mínimo, terminando así la demostración. \square

1.4.3. Función de Oportunidad

Definición 1.12 *La Función de Oportunidad, dado un vector de números reales ordenados de menor a mayor $\{\xi_i\}_{i=1}^{|\mathcal{V}|}$ y un índice cualquiera $\bar{v} \in \mathcal{V}$, se define como sigue:*

$$\min_{x \in X} \sum_{v \geq \bar{v}} \xi_v p_v(x).$$

Nota: Esto quiere decir que se ponderan las probabilidades de obtener diferentes desutilidades a partir de cierto índice. Comúnmente, conviene empezar a minimizar al menos desde el índice $n + 1$, para que así las probabilidades de obtener pérdida, en general, disminuyan y, por lo tanto, se agregue así aversión al riesgo acorde a la jugada del atacante.

1.4.4. Función de Riesgo Distorsionado

Definición 1.13 *La Función de Riesgo Distorsionado, dada una función de desutilidad $d : \mathbb{R} \rightarrow \mathbb{R}$, biyectiva y creciente, se define como sigue:*

$$\min_{x \in \mathcal{X}} \sum_{v \in \mathcal{V}} p_v(x) d(V_v)$$

Nota: Más adelante se explicará el porqué esta función debe de ser biyectiva y creciente. Por otro lado, es fácil notar que, en realidad, se podría considerar como una generalización del modelo de oportunidad, pero lo importante es el cómo se presenta la aversión, puesto que ésta distorciona todas las pérdidas y ganancias, mientras que la anterior le da más peso a las pérdidas que más importan.

1.4.5. Función de Probabilidad

Definición 1.14 *La Función de Probabilidad, dado un umbral $V_{2n} \geq \tilde{V} \geq V_1$, se define como sigue:*

$$\min_{x \in \mathcal{X}} \mathbb{P} \left(D(x) \geq \tilde{V} \right)$$

Nota: Las restricciones al umbral son necesarias, puesto que si es menor a V_1 la probabilidad es siempre 1, y si es mayor que V_{2n} siempre es 0. Al igual que el modelo de oportunidad, el modelo de probabilidad se relaciona con el modelo de riesgo distorsionado, pero, nuevamente, lo importante es el cómo se presenta la aversión.

Capítulo 2

Metodología e Implementación de los Modelos de Aversión al Riesgo

A continuación se dará a conocer el paso práctico de la tesis, donde se implementaron algoritmos que calculan la máxima utilidad esperada, los modelos de aversión y, además, se implementaron las restricciones de aversión. Todo lo anterior se hizo con el objetivo de introducir aversión al riesgo a la utilidad esperada. Casos icónicos en dónde podría ser útil serían, por ejemplo, el puerto de Boston [Bo 13], donde patrullan por diferentes zonas del puerto, o bien la zona central de Santiago de Chile [Abu11], de forma que la aversión al riesgo podría interpretarse en el cómo evitar asaltos a personas mediante asesinatos o secuestro de personas.

2.1. Contexto

En general la interacción en un juego de Stackelberg de seguridad, cada jugador vela por sí mismo, o sea, tratan de maximizar su utilidad en cierto punto. En este caso si algún agente tiene poca racionalidad, entonces no jugará teniendo en cuenta la maximización de su utilidad. Para empezar, hemos supuesto que el juego es *suma cero*, lo que quiere decir que, $M_L[i][j] = -M_S[i][j] \forall i, j \in \{1, \dots, n\}$, o sea, lo que gana el líder es lo que pierde el seguidor y viceversa. La idea detrás de esto es darle la misma prioridad que el seguidor, ya que si el líder se enfoca en su prioridad, quizás estaría descuidando sitios u objetos que son muy útiles para el atacante, el cuál podría buscar un propósito mayor. Por ejemplo, supongamos que hay un ladrón que quiere entrar a robar una gran tienda. El ladrón tiene determinado los lugares claves para entrar y salir. Podría ocurrir que los lugares no tengan la misma importancia para el jefe de seguridad de la tienda, volviendo esta situación un juego asimétrico, además de poner en riesgo a la tienda, ya que no se le estaría dando la correcta prioridad a cada sitio. Entonces, si el jefe de seguridad se pusiera en los zapatos del ladrón, vería qué lugares son realmente claves o no. Ahora bien, para lograr esto, tiene que poner en sincronía sus prioridades con las del ladrón, o sea, que el juego se vuelva *suma cero*. Notemos que lo anterior es bastante difícil de lograr, pero existe una persona (Frank Abagnale Jr [Fra80]), que al ayudar

al gobierno de Estados Unidos, dió paso a la captura de muchos falsificadores de dinero, ya que él fue un criminal muy hábil que podía salir en libertad condicional si ayudaba al gobierno.

Por el motivo anterior vamos a presentar resultados computacionales de juegos de suma cero, aunque las funciones y restricciones que se consideran son de uso más general. De hecho, existen diversos documentos que optan por no suponer juego cero por otras razones [Jua08].

Los escenarios generados son leídos por programas implementados en Julia desde Excel y, para crearlos, se usó una macro, la cual genera datos de ganancias aleatorias entre $[0, 10]$ y pérdidas entre $[-10, 0]$, luego en otras dos columnas (V1 y V2) escribe los mismos datos pero multiplicados por -1, para luego sobre esas mismas columnas ordenar de menor a mayor. Posterior a eso se guarda el orden de los valores de ganancia y pérdida en V1 y V2, respectivamente (orden1 y orden2). Para finalizar se crean dos columnas adicionales para guardar el orden de V1 y V2, con respecto a las ganancias y pérdidas (ordenin1 y ordenin2). Básicamente, se calcula a mano las funciones biyectivas entre ganancia y V1, y pérdida y V2.

Un ejemplo de lo que entrega la macro anterior:

Sitios	Ganancia	Pérdida	V1	V2	orden1	orden2	ordenin1	ordenin2
1	9,54300	-2,34560	-9,54300	0,01345	1	5	1	3
2	0,23480	-9,95320	-8,06540	0,32550	5	3	5	5
3	2,34568	-0,32550	-6,43500	2,34560	4	1	4	2
4	6,43500	-5,23980	-2,34568	5,23980	3	4	3	4
5	8,06540	-0,01345	-0,23480	9,95320	2	2	2	1

En el ejemplo anterior la cantidad de sitios es $n = 5$, las pérdidas $\overline{P}_i \in [-10, 0] \forall i$ y las ganancias $\overline{R}_i \in [0, 10] \forall i$.

2.2. Herramientas Teóricas Utilizadas para Solucionar Problemas con Aversión al Riesgo

Antes de entrar de lleno a los algoritmos se presentará una variedad de resultados que nos permitirán transformar las funciones y restricciones descritas anteriormente a problemas convexos. Esto simplifica muchísimo el cálculo computacional.

Recordemos que \mathbf{n} es la cantidad de sitios a defender y \mathbf{m} es la cantidad de recursos disponibles para defender, $\mathbf{X} = \{x \in [0, 1] : \sum_{i=1}^{\mathbf{n}} x_i \leq \mathbf{m}\}$ y (**) $\tilde{X} = \{z \in \mathbb{R}^n : z_i \leq 1, z_i \geq e^{-\gamma_i} \text{ y } -\sum_{i=1}^n \frac{-\ln(z_i)}{\gamma_i} \leq m \quad \forall i \in \{1, \dots, n\}\}$, con $x_i := \frac{-\ln(z_i)}{\gamma_i}$. Con esto pasaremos a demostrar las diferentes herramientas teóricas utilizadas.

Proposición 2.1 *Dado un vector a valores reales $(\xi_v)_{v \in \mathcal{V}} \in \mathbb{R}^{2n}$ ordenados de forma creciente, un índice $\bar{v} \in \mathcal{V}$, un valor $\Xi \in \mathbb{R}$ y un $x \in X$. Consideremos la siguiente restricción:*

$$\sum_{v \geq \bar{v}} \xi_v p_v(x) \leq \Xi$$

Si el seguidor tiene un patrón QR y el vector ξ está ordenado de menor a mayor, entonces la restricción queda:

$$\sum_{i: v^P(i) \geq \bar{v}} \xi_{v^P(i)} \beta_i z_i - \sum_{i: v^R(i) \geq \bar{v}} \xi_{v^R(i)} \frac{\beta_i}{\gamma_i} z_i \ln(z_i) - \Xi \sum_{i=1}^n \beta_i z_i + \sum_{i: v^P(i) \geq \bar{v}} \xi_{v^P(i)} \frac{\beta_i}{\gamma_i} z_i \ln(z_i) \leq 0.$$

En donde $x_i := -\frac{1}{\gamma_i} \ln(z_i) \quad \forall i \in \{1, \dots, n\}$, quedando así una **restricción convexa**.

Demostración: Sean $x \in X$, un número $\Xi \in \mathbb{R}$, un vector a valores reales $(\xi_v)_{v \in \mathcal{V}} \in \mathbb{R}^n$ ordenados de forma creciente y un índice $\bar{v} \in \mathcal{V}$ arbitrarios. Definamos $v^P(i) \geq v$ como los elementos en \mathcal{V} que se asocian a pérdidas que están por encima de v , análogamente se define para $v^R(i) \geq v$, pero con ganancias. Entonces:

$$\begin{aligned} & \sum_{v \geq \bar{v}} \xi_v p_v(x) \leq \Xi, \\ \iff & \sum_{i: v^P(i) \geq \bar{v}} \xi_{v^P(i)} (1 - x_i) \frac{\beta_i e^{-\gamma_i x_i}}{\sum_{j=1}^n \beta_j e^{-\gamma_j x_j}} + \sum_{i: v^R(i) \geq \bar{v}} \xi_{v^R(i)} x_i \frac{\beta_i e^{-\gamma_i x_i}}{\sum_{j=1}^n \beta_j e^{-\gamma_j x_j}} \leq \Xi, \\ \iff & \sum_{i: v^P(i) \geq \bar{v}} \xi_{v^P(i)} (1 - x_i) \beta_i e^{-\gamma_i x_i} + \sum_{i: v^R(i) \geq \bar{v}} \xi_{v^R(i)} x_i \beta_i e^{-\gamma_i x_i} \leq \Xi \sum_{j=1}^n \beta_j e^{-\gamma_j x_j}, \end{aligned}$$

utilizando el cambio de variable $x_i := -\frac{1}{\gamma_i} \ln(z_i)$, y separando la resta en la expresión en (†) se llega a,

$$\begin{aligned} & \overbrace{\left(\sum_{i:v^P(i) \geq \bar{v}} \xi_{v^P(i)} \beta_i z_i - \sum_{i:v^R(i) \geq \bar{v}} \xi_{v^R(i)} \frac{\beta_i}{\gamma_i} z_i \ln(z_i) - \Xi \sum_{i=1}^n \beta_i z_i + \sum_{i:v^P(i) \geq \bar{v}} \xi_{v^P(i)} \frac{\beta_i}{\gamma_i} z_i \ln(z_i) \right)}^{h(z):=} \leq 0. \\ \iff & \sum_{i:v^P(i) \geq \bar{v}} \xi_{v^P(i)} \beta_i z_i - \sum_{i:v^R(i) \geq \bar{v}} \xi_{v^R(i)} \frac{\beta_i}{\gamma_i} z_i \ln(z_i) - \Xi \sum_{i=1}^n \beta_i z_i + \sum_{i:v^P(i) \geq \bar{v}} \xi_{v^P(i)} \frac{\beta_i}{\gamma_i} z_i \ln(z_i) \leq 0. \end{aligned}$$

Notemos que si un término en (*) aparece, este vuelve a aparecer en la última sumatoria, ya que $v^P(i) > v^R(i)$. Luego, consideremos un sitio i que contenga un término en (*). Su contribución a la restricción es:

$$\xi_{v^P(i)} \beta_i z_i - \Xi \beta_i z_i + \xi_{v^P(i)} \frac{\beta_i}{\gamma_i} z_i \ln(z_i) - \xi_{v^R(i)} \frac{\beta_i}{\gamma_i} z_i \ln(z_i).$$

Entonces, aquellos términos que son lineales se pueden ignorar, pues son convexos. Dentro del espacio factible, los términos no lineales de la contribución son 2 veces diferenciables, por lo tanto, aplicando la caracterización de funciones convexas se obtiene que:

$$(\xi_{v^P(i)} - \xi_{v^R(i)}) \frac{\beta_i}{\gamma_i} z_i \ln(z_i) \text{ es convexo} \iff (z_i \ln(z_i))'' \geq 0, \quad (**)$$

dado que $(\xi_{v^P(i)} - \xi_{v^R(i)}) \geq 0 \iff \xi_{v^P(i)} \geq \xi_{v^R(i)}$, pues este vector de valores reales está ordenado de forma creciente, y además se tiene que $\beta, \gamma \geq 0$. Finalmente demostremos (**):

$$(z_i \ln(z_i))' = \ln(z_i) + 1 \implies (z_i \ln(z_i))'' = \frac{1}{z_i} \geq 0 \quad \forall z \in \tilde{X}.$$

Todo lo anterior es útil, ya que la función $h(\cdot)$ es la suma de funciones independientes entre si, por lo tanto, la matriz hessiana de $h(\cdot)$ es diagonal y, como cada término de la diagonal es positivo $\forall z \in \tilde{X}$, entonces la matriz es semidefinida postvia en \tilde{X} , o sea, $h(\cdot)$ es convexa en \tilde{X} . \square

Nota: Aquellos sitios que sólo posean términos de pérdida, tendrán un aporte a la restricción de la forma $\xi_{v^P(i)} \beta_i z_i - \Xi \beta_i z_i + \xi_{v^P(i)} \frac{\beta_i}{\gamma_i} z_i \ln(z_i)$ y como estamos analizando aversión al riesgo, nunca se le dará peso negativo a la probabilidad p_v con v asociado a una pérdida. Por lo tanto, esa contribución también será convexa.

Proposición 2.2 *La restricción de Oportunidad se puede escribir como una restricción convexa utilizando la **Proposición 2.1** con $\xi_v = 1$, $\Xi = \varepsilon$ y $\bar{v} = \tilde{v}$*

Proposición 2.3 *La restricción de Riesgo Distorsionado se puede escribir como una restricción convexa utilizando la **Proposición 2.1** con $\xi_v = d(V_v)$, $\Xi = d(\tilde{\rho})$ y $\bar{v} = 1$*

Proposición 2.4 *La restricción de Primer Orden de Dominancia Estocástica se puede escribir como una restricción convexa utilizando la **Proposición 2.1** con $\xi_v = 1$, $\Xi = 1 - F_{D(\bar{x})}(V_{\bar{v}})$ y $\bar{v} = \tilde{v} + 1$ para cada $\tilde{v} \in \mathcal{V}$.*

Antes de presentar la proposición que nos dice cómo resolver los modelos del capítulo anterior se introducirá un teorema que fue demostrado en [Bo 13] (anexo), el cuál nos dice cómo linealizar y adivinar el valor, vía búsqueda binaria, de la función objetivo que es fraccionaria.

Definición 2.5 *Sean $D(\cdot)$, $N(\cdot)$ funciones continuas definidas en $\Pi \subseteq \mathbb{R}^n$ y sea $r \in \mathbb{R}$, se define CF-OPT como sigue:*

$$CF - OPT : \quad \delta_r^* := \min_{x \in \Pi} \{N(x) - rD(x)\}.$$

Lema 2.6 *Sean $N(\cdot)$, $D(\cdot)$ funciones continuas definidas en un conjunto $\Pi \subseteq \mathbb{R}^n$, cerrado y acotado. Supongamos que $D(x) > 0 \forall x \in \Pi$. Si $p^* = \min_{x \in \Pi} \frac{N(x)}{D(x)}$, $r \in \mathbb{R}$, δ_r^* definidos como en CF-OPT, entonces:*

$$r \geq p^* \iff \delta_r^* \leq 0.$$

Demostración: Sean todas las funciones y valores ya dados y arbitrarios, cumpliendo las hipótesis mencionadas.

\implies Se tendrá que $\exists \bar{x} \in \Pi \quad r \geq \frac{N(\bar{x})}{D(\bar{x})}$, pues p^* se alcanza en un punto. Entonces $N(\bar{x}) - rD(\bar{x}) \leq 0$, lo que implica que $\min_{x \in \Pi} \{N(x) - rD(x)\} \leq 0 \iff \delta_r^* \leq 0$.

\impliedby Se tendrá que $\exists \bar{x} \in \Pi \quad N(\bar{x}) - rD(\bar{x}) \leq 0$. Luego, $\exists \bar{x} \in \Pi \quad r \geq \frac{N(\bar{x})}{D(\bar{x})} \implies r \geq \min_{x \in \Pi} \frac{N(x)}{D(x)} = p^*$.

Por último, cabe mencionar que el mínimo de la función $\frac{N(x)}{D(x)}$ se alcanza porque esta última está bien definida en Π , además, es continua al ser composición de funciones continuas y, por consiguiente, alcanza el ínfimo y el supremo en un conjunto compacto (cerrado y acotado en espacios de dimensión finita). \square

Proposición 2.7 *Si el atacante sigue un patrón QR, dado un vector a valores reales $(\xi_v)_{v \in \mathcal{V}}$ tales que $\xi_1 \leq \xi_2 \leq \dots \leq \xi_{2n}$ y algún $\bar{v} \in \mathcal{V}$, se puede resolver el siguiente problema:*

$$\min_{x \in X} \sum_{v \geq \bar{v}} \xi_v p_v(x),$$

resolviendo una cantidad logarítmica de problemas de optimización convexa de la forma:

$$\min_{z \in \tilde{X}} \left\{ \sum_{i: v^P(i) \geq \bar{v}} \xi_{v^P(i)} \beta_i z_i - r \sum_{i=1}^n \beta_i z_i + \sum_{i: v^P(i) \geq \bar{v}} \xi_{v^P(i)} \frac{\beta_i}{\gamma_i} z_i \ln(z_i) - \sum_{i: v^R(i) \geq \bar{v}} \xi_{v^R(i)} \frac{\beta_i}{\gamma_i} z_i \ln(z_i) \right\}.$$

En donde $z_i = e^{-\gamma_i x_i} \forall i \in \{1, \dots, n\}$.

Demostración: Primero notemos que el **lema 2.6** nos dá un indicio de cómo solucionar el problema, puesto que nos proporciona con una herramienta de búsqueda binaria linealizando la función fraccionaria, mediante adivinanzas de lo que podría ser el óptimo de la función en cuestión. Entonces, transformemos el problema a otro equivalente como sigue:

$$\min_{x \in X} \sum_{v \geq \bar{v}} \xi_v p_v(x) = \inf_{x \in X, r \in \mathbb{R}} \left\{ r : \exists x \in \Pi, r \geq \sum_{v \geq \bar{v}} \xi_v p_v(x) \right\} := r^*$$

Esta transformación se tiene ya que si r cumple la desigualdad para algún $x \in X$, entonces $\min_{x \in X} \sum_{v \geq \bar{v}} \xi_v p_v(x) \leq r \implies \min_{x \in X} \sum_{v \geq \bar{v}} \xi_v p_v(x) \leq r^*$, pues es cota inferior de todos los r factibles.

Por otro lado, tomando como punto factible el mínimo de las probabilidades ponderadas se muestra que el ínfimo en realidad es un mínimo y además que $r^* = \min_{x \in X} \sum_{v \geq \bar{v}} \xi_v p_v(x)$. Con lo anterior se procede a solucionar el problema tomando adivinanzas sobre r y dejándolo fijo, quedando como resultado:

$$\min_{x \in X} \left\{ r : \exists x \in \Pi, r \geq \sum_{v \geq \bar{v}} \xi_v p_v(x) \right\}.$$

El cuál es un problema de factibilidad puro. Por lo tanto, usando la **proposición 2.1** y además transformando las probabilidades p_v y X con el cambio de variable $x_i := \frac{-\ln(z_i)}{\gamma_i} \forall \in \{1, \dots, n\}$, se puede escribir la restricción de forma convexa. Gracias a lo anterior, para resolver el problema de factibilidad puro se minimiza una función convexa en un conjunto convexo, en dónde el problema tiene la forma:

$$l^* := \min_{z \in \tilde{X}} \left\{ \sum_{i: v^P(i) \geq \bar{v}} \xi_{v^P(i)} \beta_i z_i - \sum_{i: v^R(i) \geq \bar{v}} \xi_{v^R(i)} \frac{\beta_i}{\gamma_i} z_i \ln(z_i) - r \sum_{i=1}^n \beta_i z_i + \sum_{i: v^P(i) \geq \bar{v}} \xi_{v^P(i)} \frac{\beta_i}{\gamma_i} z_i \ln(z_i) \right\}.$$

Si $l^* \leq 0$, entonces r es factible, se guarda la solución $z^* \in \tilde{X}$ (**) y se prosigue con la búsqueda binaria. Si nó, no se guarda ninguna información relevante y se continua con la búsqueda binaria. \square

Observación: Notemos que el problema descrito anteriormente es precisamente el modelo de oportunidad definido en el capítulo anterior.

Proposición 2.8 Sea $d : \mathbb{R} \rightarrow \mathbb{R}$ una función de desutilidad creciente y biyectiva. Entonces el problema de minimización de la medida de riesgo distorsionada es como sigue:

$$\min_{x \in X} \sum_{v \in \mathcal{V}} p_v(x) d(V_v).$$

En donde, usando la **proposición 2.7** con $\bar{v} = 1$ y $\xi_v = d(V_v)$, se puede solucionar resolviendo una cantidad logarítmica de problemas convexos.

Proposición 2.9 Sea $V_1 \leq \tilde{V} \leq V_{2n}$ un umbral. Entonces el problema de minimización de la probabilidad sobre ese umbral es como sigue:

$$\min_{x \in X} \mathbb{P} \left(D(x) \geq \tilde{V} \right).$$

Sea $\tilde{v} := \arg \min_{v \in \mathcal{V}} \{V_v : V_v \geq V_{\tilde{V}}\}$. El problema pasa a ser:

$$\min_{x \in X} \sum_{v \geq \tilde{v}} p_v(x).$$

El cuál se puede resolver, usando la **proposición 2.7** con $\bar{v} = \tilde{v}$ y $\xi_v = 1$, solucionando una cantidad logarítmica de problemas convexos.

Capítulo 3

Algoritmos para Problemas con Aversión al Riesgo a la Estrategia del Oponente

En este capítulo usaremos fuertemente las proposiciones y lemas vistos en el capítulo anterior para presentar esquemas o pseudo códigos escritos en Julia para la solución bajo el contexto descrito anteriormente.

Antes de empezar, Julia es un programa de uso abierto, el cuál posee soporte para Excel, Python, Matlab, lenguaje de optimización y una variedad de solvers, algunos abiertos al público y otros pagados. En este caso se utilizó Ipopt como solver, JuMP como lenguaje de optimización y el soporte con Excel. Todo lo relacionado con Julia se puede encontrar en la página <https://www.juliaopt.org/>.

Como dato final, se presentará el algoritmo que resuelve la minimización del CVaR, pero no se hablará acerca de sus resultados, puesto que el tiempo de resolución es demasiado lento (existe una restricción cóncava que genera problemas).

3.1. Metodología CVaR

Para poder resolver este problema se utiliza la proposición 1.11, la cual nos dice que el ínfimo de la definición 1.10 se alcanza en el $\text{VaR}_\varepsilon(D(x))$. Esto nos indica una manera de resolver, pero que veremos que es lento.

Notemos que el problema se basa en ir adivinando el valor de t en el problema:

$$\min_{x \in X, \bar{v} \in \mathcal{V}} \{V_{\bar{v}} + \varepsilon^{-1} \mathbb{E}([D(x) - V_{\bar{v}}]_+)\} \quad (0)$$

Con esto, se procede a dejar fijo el valor de \bar{v} , y se va resolviendo el siguiente problema:

$$\begin{aligned}
V_{\bar{v}} + \varepsilon^{-1} \overbrace{\min_{x \in X} \sum_{v \geq \bar{v}+1}^{u_{\bar{v}}^*} p_v(x) (V_v - V_{\bar{v}})} & \quad (00) \\
\text{s.a. } \sum_{v \geq \bar{v}} p_v(x) & \geq \varepsilon \\
\sum_{v \geq \bar{v}+1} p_v(x) & \leq \varepsilon
\end{aligned}$$

En donde la restricción es para asegurar que el $\text{VaR}_\varepsilon(D(x)) = V_{\bar{v}}$, puesto que de no colocar dicha restricción existe el problema de encontrar una solución x^* cuyo VaR_ε no concuerde con $V_{\bar{v}}$. Con esto se ve que el problema está en la primera restricción, ya que es una restricción cóncava, mientras que el resto de factores del problema son convexos. Por lo tanto, para solucionar este dilema se procede a linealizar el problema, utilizando el siguiente teorema (que simplemente se mencionará).

Teorema 3.1 *Dado una partición $l = t_0 < t_1 < \dots < t_K = u$ de $[l, u]$, se puede aproximar la función $f : [l, u] \rightarrow \mathbb{R}$ como sigue:*

$$\begin{aligned}
f(x) \approx \bar{f}(x) & := \min_{x_k, z_k} f(l) + \sum_{k=1}^K (f(t_k) - f(t_{k-1})) x_k \\
\text{s.a. } x & = l + \sum_{k=1}^K x_k (t_k - t_{k-1}) \\
x_1 & \leq 1 \\
x_K & \geq 0 \\
x_{k+1} & \leq z_k \leq x_k, \quad \forall k \in \{1, \dots, K-1\} \\
z_k & \in \{0, 1\}, \quad \forall k \in \{1, \dots, K-1\}
\end{aligned}$$

Esta forma de aproximar mediante funciones lineales por tramos se conoce como modelo incremental [Dan63]. El problema de utilizar una aproximación lineal es que agrega muchísimas restricciones y variables, transformándolo a un MIP gigantesco, si es que la cantidad de sitios crece mucho, por ejemplo, $n \geq 20$. Para aplicar esta aproximación se debe linealizar cada función en cada factor no lineal que se esté involucrado en el modelo.

Como se puede apreciar, lo anterior es básicamente resolver $|\mathcal{V}| = 2n$ problemas linealizados con muchas variables y restricciones. Finalmente, para presentar el algoritmo se define lo siguiente: la rutina 'solución(\tilde{v})' toma como argumento un índice $\tilde{v} \in \mathcal{V}$ y entrega una tupla $(x^*, u_{\tilde{v}}^*)$, en donde, respectivamente, corresponden a la solución óptima y al valor óptimo del problema (00).

Data: Ingresamos una instancia para el problema (0)

Result: Una solución óptima x^* y un valor óptimo w^* para el problema (0)

$v = 1;$

$w^* := \infty^+;$

while $v \neq 2n$ **do**

$(x, u) := \text{solución}(v);$

if $w^* > V_v + \varepsilon^{-1}u$ **then**

$x^* = x;$

$w^* = V_v + \varepsilon^{-1}u;$

end

$v = v + 1;$

end

Entregar x^* y w^* .

Algorithm 1: Conditional Value at Risk

Esto se podría mejorar eliminando objetivos en \mathcal{V} si utilizamos el hecho de que $VaR_\varepsilon(D(x)) \leq CVaR_\varepsilon(D(x))$. Si algún valor $w_i^* < V_{i+1}$, quiere decir que $w_{i+1}^* > w_i^*$, por lo tanto, V_{i+1} es descartado como adivinanza de t .

3.2. Metodología VaR

Utilizando La **proposición 1.9** el problema de minimización del Value at Risk queda como sigue:

$$\min_{x \in X, \tilde{v} \in \mathcal{V}} \left\{ V_{\tilde{v}} : \sum_{v \leq \tilde{v}} p_v(x) \geq 1 - \varepsilon \right\}. \quad (1)$$

Luego, recordemos que minimizar sobre $A \times B$ es lo mismo que minimizar sobre A primero y luego B , o viceversa. Entonces cambiando el orden de minimización, el problema queda:

$$\min_{\tilde{v} \in \mathcal{V}} \left\{ V_{\tilde{v}} + \min_{x \in X} \left\{ 0 : \sum_{v \leq \tilde{v}} p_v(x) \geq 1 - \varepsilon \right\} \right\}. \quad (2)$$

Por tanto, en el caso del VaR se harán 2 búsquedas binarias, la primera para encontrar el mínimo v en \mathcal{V} y la segunda para resolver cada subproblema de factibilidad sobre X utilizando la **proposición 2.7** y el **lema 2.6**.

Para resolver el problema de factibilidad sobre X al tener fijo un elemento en \mathcal{V} se resuelve el siguiente problema:

$$u_{\tilde{v}} := \min_{x \in X} \sum_{v \geq \tilde{v}+1} p_v(x). \quad (3)$$

La razón de esto es que si $u_{\tilde{v}} \leq \varepsilon$, entonces $1 - u_{\tilde{v}} = \sum_{v \leq \tilde{v}} p_v(x^*) \geq 1 - \varepsilon$, con x^* siendo el óptimo encontrado para resolver el problema (3). Con toda esta información se procede a explicar y mostrar el algoritmo diseñado para minimizar el VaR.

- Primero se toma una adivinanza sobre el posible valor mínimo que puede tener el VaR en \mathcal{V} . Por ejemplo, $v \in \mathcal{V}$.
- Luego se verifica si existe una forma de defender que garantice que $F_{D(x)}(V_v) \geq 1 - \varepsilon$ para algún $x \in X$ (encontrando $u_{\tilde{v}}$ usando la proposición 2.7, y lema 2.6. Y finalmente se compara con ε).
- Si este valor es menor o igual ε , entonces el subproblema de factibilidad en (2) es factible y se guarda la solución encontrada. Ahora bien, si v posee un $x \in X$ que garantice $F_{D(x)}(V_v) \geq 1 - \varepsilon$ entonces cualquier otro elemento en \mathcal{V} superior a v ya no sirve, puesto que se está minimizando el VaR. Luego se guarda el valor v como posible mínimo del VaR. Por otro lado, si no logra garantizar $F_{D(x)}(V_v) \geq 1 - \varepsilon$ para algún $x \in X$, entonces v es una cota inferior al problema y el óptimo debe de estar por encima de él.

Cómo se puede apreciar lo anterior es en sí es una búsqueda binaria sobre \mathcal{V} . Finalmente para presentar el algoritmo se define lo siguiente: la rutina 'solución(\tilde{v})' toma como argumento un índice $\tilde{v} \in \mathcal{V}$ y entrega una tupla $(x^*, u_{\tilde{v}})$, en donde respectivamente corresponden a la solución óptima y al valor óptimo del problema (3).

Comentario: Notemos que al calcular solución($|\mathcal{V}| = 2n$), esta nos entregará cualquier solución factible, ya que si exigimos que nuestra desutilidad (en probabilidad) sea más pequeño que V_{2n} , entonces la probabilidad, llámese 'p', de que eso ocurra es $p = 1 \geq 1 - \varepsilon \quad \forall \varepsilon > 0$.

Data: Ingresamos una instancia para el problema (1)

Result: Una solución óptima x^* y un valor óptimo v^* para el problema (1)

Inicialización: $(x^*, u) := \text{solución}(1)$;

if $u \leq \varepsilon$ **then**

 | entregar $(x^*, 1)$ y salir

else

 | $L := 1$;

 | $U := 2n$;

 | $(x^*, u) = \text{solución}(|\mathcal{V}|)$;

while $U > L + 1$ **do**

 | $v := \lceil (L + U)/2 \rceil$;

 | $(x, u) := \text{solución}(v)$;

 | **if** $u \leq \varepsilon$ **then**

 | $x^* = x$ y $U := v$;

 | **else**

 | $L := v$;

 | **end**

 | **end**

 | Entregar x^* y $U = v^*$.

end

Algorithm 2: Value at Risk

3.3. Metodología FDO

Notemos que el algoritmo anterior simplifica una parte vital, la cuál es encontrar la solución al problema (3). El algoritmo presentado en esta sección resuelve el problema (3) y el mínimo de la función de Oportunidad. Como se menciona en el capítulo anterior utilizaremos la **proposición 2.7** y **lema 2.6** para resolver este problema. Recordemos que el problema de minimización de la función de Oportunidad es:

$$\min_{x \in X} \sum_{v \geq \bar{v}} \xi_v p_v(x). \quad (4)$$

En donde $\{\xi_v\}_{v \in \mathcal{V}}$ están ordenados de manera creciente y \bar{v} es un índice arbitrario en \mathcal{V} . Al hacer un cambio a un problema equivalente lo anterior queda:

$$\min_{x \in X} \left\{ r : r \geq \sum_{v \geq \bar{v}} \xi_v p_v(x) \right\}. \quad (5)$$

Entonces, la idea detrás de esto es ir tomando diferentes valores de r e ir resolviendo problemas de factibilidad linealizando la función.

Primero que todo, una cota inferior asociada al problema (4) sería darle el valor de 1 a aquella probabilidad que contenga al ξ_v más pequeño, en caso de que este sea negativo, o darle valor 0 a todas las probabilidades en caso de que ningún ξ_v sea negativo. Llamaremos a esta cota inferior L , cuya definición más formal es:

$$L := \min_{e^\top p=1, p \geq 0} \sum_{i \in \{\bar{v}, \dots, 2n\}} p_i \xi_v = \min_{v \geq \bar{v}} \{\xi_v\}.$$

En donde $e_i = 1 \quad \forall i \in \{1, \dots, 2n\}$.

Con toda esta información se procede a explicar y mostrar el algoritmo diseñado para la función de Oportunidad.

- Primero se toma una adivinanza sobre el posible valor mínimo del problema (5) para tener una cota superior. Por ejemplo $U := \sum_{v \geq \bar{v}} \xi_v p_v(\bar{x})$, donde \bar{x} es un punto factible en X .
- Luego se adivina $r = (L + U)/2$. Y se procede a linealizar la función fraccionaria del problema (4) quedando de la forma $N(x) - rD(x)$. Posterior a esto, lo que queda es un problema convexo (gracias a la proposición 2.7) que se resuelve en Julia con el solver Ipopt y se obtiene una tupla (x^*, l^*) que corresponde al argumento mínimo y valor mínimo de la función linealizada respectivamente.
- Si l^* es no positivo, entonces se guarda la solución y se deja a r como cota superior (U). Si nó, se deja a r como cota inferior (L).

Observemos que r cada vez va siendo más pequeño y se va viendo si es válido como adivinanza o no. Claramente esto podría iterar indefinidamente, pero lo que se busca es una

aproximación lo suficientemente cercana al óptimo del problema (4), por lo tanto, el algoritmo itera hasta que la diferencia entre las cotas es ínfima. Definamos la rutina 'sol(r)', la cuál toma como argumento un valor $r \in \mathbb{R}$, y entrega una tupla (x^*, l^*) que representa la solución óptima y valor óptimo del problema (4), respectivamente, linealizado como indica el lema 2.6.

Data: Ingresamos una instancia para el problema (4). Y una tolerancia $\delta > 0$ pequeña
Result: Una solución óptima x^* y valor óptimo r^* para el problema (4)
Inicialización: $U := \sum_{v \geq \bar{v}} \xi_v p_v(\tilde{x})$, con \tilde{x} solución factible y $L := \min_{v \geq \bar{v}} \{\xi_v\}$;
while $U - L \geq \delta$ **do**
 $r = (U + L) / 2$;
 $(x, l) := \text{sol}(r)$;
 if $l \leq 0$ **then**
 $x^* := x, r^* := r$ y $U = r$;
 else
 $L = r$;
 end
end
Entregar x^* y r^* .

Algorithm 3: Función de Oportunidad

Nota: Este algoritmo se considera la rutina solución(\cdot) utilizada en el VaR. Lo anterior llevaría a la conclusión prematura de que resolver el VaR es más lento que resolver cualquier otra función de aversión al riesgo y, de hecho, así es, pero esto se verá al momento de observar los resultados de los experimentos. Además, más adelante se entrará en más detalle del porqué ocurre esto.

3.4. Metodología FRD

En esta sección, usando la **proposición 2.8**, se presenta un algoritmo que resuelve de manera similar el problema de minimización de la función de riesgo distorsionado. Lo que cambia es el peso que se le dá a cada probabilidad, el índice desde donde se suma y la cota inferior natural del problema. Recordemos que este problema es como sigue:

$$\min_{x \in X} \sum_{v \in \mathcal{V}} p_v(x) d(V_v) \quad (6)$$

El problema (6) posee la siguiente cota inferior:

$$L := \min_{e^\top p = 1, p \geq 0} \sum_{i \in \{1, \dots, 2n\}} d(V_i) p_i = d(V_1).$$

Ya que d es creciente, además $e_i = 1 \forall i \in \{1, \dots, 2n\}$.

Definamos la rutina 'sol(r)', la cuál toma como argumento un valor $r \in \mathbb{R}$, y entrega una tupla (x^*, l^*) que representa la solución óptima y valor óptimo del problema (6) respectivamente, linealizado como indica el lema 2.6.

Data: Ingresamos una instancia para el problema (6). Y una tolerancia $\delta > 0$ pequeña

Result: Una solución óptima x^* y valor óptimo r^* para el problema (6)

Inicialización: $U := \sum_{v \geq 1} d(V_v)p_v(\tilde{x})$, con \tilde{x} solución factible y $L := d(V_1)$;

while $U - L \geq \delta$ **do**

$r = (U + L) / 2$;

$(x, l) := \text{sol}(r)$;

if $l \leq 0$ **then**

$x^* := x, r^* := r$ y $U = r$;

else

$L = r$;

end

end

Entregar x^* y r^* .

Algorithm 4: Función de Riesgo Distorsionado

3.5. Metodología FDP

En esta sección, usando el **proposición 2.9**, se presenta un algoritmo que resuelve de manera similar el problema de minimización de la función de probabilidad. Al igual que antes cambian algunos parámetros, pero el algoritmo sigue siendo el mismo. Los parámetros que se ven alterados son los mismos que en la sección anterior. Recordemos que este problema es como sigue:

$$\min_{x \in X} \mathbb{P} \left(D(x) \geq \tilde{V} \right) \iff \min_{x \in X} \sum_{v \geq \tilde{v}} p_v(x) \quad (7)$$

En donde \tilde{v} es tal que $V_{\tilde{v}} \geq \tilde{V}$ y $V_{\tilde{v}-1} \leq \tilde{V}$.

El problema (7) tiene cota inferior trivial $L := 0$ y superior $U := 1$. Por otro lado, para este algoritmo se diseñó un programa de búsqueda binaria que encuentra en tiempo logarítmico el \tilde{v} más pequeño, que sea mayor o igual a \tilde{V} . A este algoritmo se define como 'cota(\cdot)', el cuál toma como argumento un número \tilde{V} y entrega el índice v más pequeño tal que $V_v \geq \tilde{V}$. Y por último, definamos la rutina 'sol(r, v)', la cuál toma como argumento un valor $r \in \mathbb{R}$ y un índice v en donde inicia la suma, y entrega una tupla (x^*, l^*) que representa la solución óptima y valor óptimo del problema (7) respectivamente, linealizado como indica el lema 2.6.

Data: Ingresamos una instancia para el problema (7). Y una tolerancia $\delta > 0$ pequeña

Result: Una solución óptima x^* y valor óptimo r^* para el problema (7)

Inicialización: $U := 1$ y $L := 0$;

while $U - L \geq \delta$ **do**

$r = (U + L) / 2$;

$(x, l) := \text{sol}(r)$;

if $l \leq 0$ **then**

$x^* := x, r^* := r$ y $U = r$;

else

$L = r$;

end

end

Entregar x^* y r^* .

Algorithm 5: Función de Probabilidad

3.6. Metodología MUE

Antes que nada recordemos como es el problema de maximización de la utilidad esperada:

$$\max_{x \in X} \sum_{i=1}^n y_i(x) \bar{U}_i(x_i) \quad (8)$$

Entonces, en esta sección se utilizarán 2 herramientas teóricas diferentes a las del resto, ya que se está maximizando una función, que a priori no es convexa y es fraccionaria. Estas herramientas son las siguientes:

Definición 3.2 Sean $D(\cdot)$, $N(\cdot)$ funciones continuas definidas en $\Pi \subseteq \mathbb{R}^n$ y sea $r \in \mathbb{R}$, se define $CF-OPT^*$ como sigue:

$$CF - OPT^* : \quad \delta_r^* := \min_{x \in \Pi} \{rD(x) - N(x)\}.$$

Lema 3.3 Sean $N(\cdot)$, $D(\cdot)$ funciones continuas definidas en un conjunto $\Pi \subseteq \mathbb{R}^n$, cerrado y acotado. Supongamos que $D(x) > 0 \forall x \in \Pi$. Si $p^* = \max_{x \in \Pi} \frac{N(x)}{D(x)}$, $r \in \mathbb{R}$, δ_r^* definidos como en $CF-OPT^*$, entonces:

$$r \leq p^* \iff \delta_r^* \leq 0.$$

Observemos que al linealizar la función, ésta tiene signo cambiado con respecto a la del lema 2.6, pero sigue siendo minimización. Por lo tanto, se puede seguir aplicando la **proposición 2.7**, sólo que ahora si un r es válido, es por que es cota inferior del problema, no superior.

Por otro lado, el problema (8) se puede transformar de forma equivalente a:

$$\max \left\{ r : r \leq \sum_{i=1}^n y_i(x) \bar{U}_i(x_i) \right\} \quad (9)$$

La demostración de lo anterior es análogo a lo hecho en la demostración del teorema de la proposición 2.7. adherente a la transformación, esto nos da indicios de una búsqueda binaria. La diferencia es que la adivinanza tendrá que ir subiendo, entonces, como se mencionó anteriormente, al ser válido el valor r , se le asignará su valor a la cota inferior relativa, así en la siguiente iteración la adivinanza habrá crecido.

Notemos que este problema tiene cota inferior L igual a evaluar un punto factible en la función de utilidad esperada, mientras que la cota superior es como sigue:

$$\begin{aligned} U &:= \max_{i \in \{1, \dots, n\}} \bar{R}_i \geq \sum_{i=1}^n y_i(x) x_i \left(\max_{i \in \{1, \dots, n\}} \bar{R}_i \right) \geq \sum_{i=1}^n y_i(x) x_i \left(\max_{i \in \{1, \dots, n\}} \bar{R}_i \right) + y_i(x) (1 - x_i) \bar{P}_i \\ &\geq \sum_{i=1}^n y_i(x) (\bar{R}_i x_i + (1 - x_i) \bar{P}_i) = \sum_{i=1}^n y_i(x) \bar{U}_i(x_i). \end{aligned}$$

Entonces, definamos la rutina 'sol(r)', la cual toma como argumento un valor $r \in \mathbb{R}$ y entrega una tupla (x^*, l^*) que representa la solución óptima y valor óptimo del problema (8), respectivamente, linealizado como indica el lema 2.6.

Data: Ingresamos una instancia para el problema (8). Y una tolerancia $\delta > 0$ pequeña

Result: Una solución óptima x^* y valor óptimo r^* para el problema (8)

Inicialización: $U := \max_{i \in \{1, \dots, n\}} \bar{R}_i$ y $L := \sum_{i=1}^n y_i(\tilde{x}) \bar{U}_i(\tilde{x})$, con \tilde{x} solución factible;

while $U - L \geq \delta$ **do**

$r = (U + L) / 2$;

$(x, l) := \text{sol}(r)$;

if $l \leq 0$ **then**

$x^* := x, r^* := r$ y $L = r$;

else

$U = r$;

end

end

Entregar x^* y r^* .

Algorithm 6: Máxima Utilidad Esperada

Capítulo 4

Resultados y Análisis

Se resolvieron los problemas de búsqueda binaria en las diferentes metodologías a una precisión de 10^{-4} . Además, se probaron los algoritmos con coeficiente de racionalidad del atacante $\lambda = 0.25$, $n = 400$ sitios a defender y $m = 60$ recursos disponibles. También, como se mencionó anteriormente, las ganancias (pérdidas respectivamente) de cada sitio se obtuvieron uniformemente en $(0, 10)$ (respectivamente $(-10, 0)$), utilizando el macro en Excel ya mencionado al inicio del capítulo 2. Recordar que el juego es suma cero.

El método que resuelve los subproblemas convexos es Ipopt y se programó en Julia. El equipo utilizado tiene memoria RAM de 8 GB, procesador Intel(R) Core(TM) i5-4300U CPU @ 1.90 GHz 2.49 GHz, sistema operativo de 64 bits Windows 10 Pro.

Además, cabe mencionar que en todos los algoritmos de búsqueda binaria toma como máximo una cantidad $O(\log_2(\frac{U-L}{10^{-4}}))$ de iteraciones para resolver cada problema.

Por último, la muestra de resultados se separará en dos tipos. El primero comparando la metodología MUE versus el resto de metodologías resueltas de formas diferentes. Y el segundo, mostrando la metodología MUE versus la misma metodología, pero agregando aversión al riesgo, mediante restricciones.

Cada experimento consta de cargar los datos de Excel, luego se calcula cada parámetro asociado a dichos datos. Después, se ordenan los parámetros (limpieza en el programa), se calcula la solución mediante la búsqueda binaria y finalmente, se presentan los datos en pantalla (consola) junto con tres gráficos, los cuales muestran la función de distribución acumulada, el histograma de las probabilidades $p_v(x^*)$ y el histograma de las probabilidades x^* (solución del problema).

Los gráficos que se mostrarán corresponden a los del experimento 7, de un total de 10 que se llevaron a cabo. Se observó el mismo patrón en todos y además se calculó los resultados promedios, arrojando el mismo patrón. Por esa razón, se quiso mostrar sólo un experimento, ya que sería redundante mostrar el mismo patrón varias veces.

4.1. Resultados de Funciones de Aversión al Riesgo

En esta sección se mostrará los resultados obtenidos al resolver la metodología VaR con $\varepsilon \in \{0.2, 0.1, 0.05, 0.01\}$. La metodología FDO con $\bar{v} := 320$ y el vector ξ de tres formas distintas:

- 1) $\xi_i = \begin{cases} 1 & , \text{ si } 1 \leq i \leq 639 \\ 1 + 0.1 * (i - 639) & , \text{ si no } \end{cases}$.
- 2) $\xi_i = \begin{cases} 1 & , \text{ si } 1 \leq i \leq 639 \\ 1 + (1.05)^{(i-639)} & , \text{ si no } \end{cases}$.
- 3) $\xi_i = 1 \quad \forall i \in \{1, \dots, n\}$.

O sea, sumando a partir del 40% de \mathcal{V} y modificando el peso de cada resultado a partir del índice del 80% de \mathcal{V} , ya que la aversión al riesgo vista en esta función se enfoca en el peor 20% de los sitios con mayor pérdida.

Además, se resuelve y muestra la metodología FRD utilizando una función d de desutilidad biyectiva y creciente de la siguiente forma:

$$d(x) = \begin{cases} -(-x)^{\frac{2}{3}} & , \text{ si } x < 0 \\ 2.25 \cdot x^{1.6} & , \text{ si } x \geq 0 \end{cases}$$

Esto se definió así porque se le está dando un poco más del doble de importancia a las pérdidas que a las ganancias, tanto en el valor que acompaña a x cómo a cuanto se está elevando el mismo. A continuación se presenta un gráfico que muestra la forma de esta función.

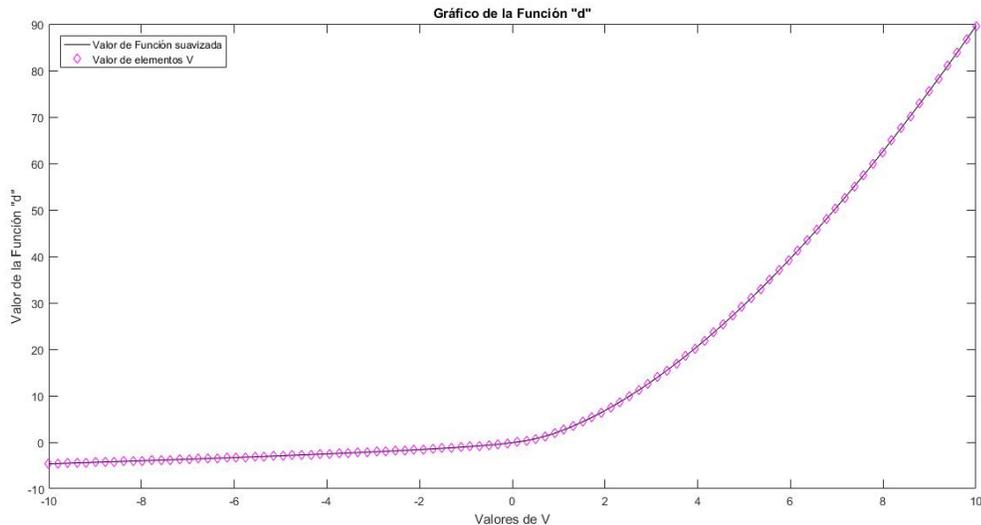


Figura 4.1: Gráfico de la Función de Distorción

Por otro lado, se resuelve y se muestra la metodología FDP utilizando tres umbrales en \mathcal{V} diferentes $v \in \{401, 600, 720\}$, que corresponden al 100 %, 50 % y 20 % de las pérdidas, respectivamente. Notemos que si utilizamos un umbral menor o igual al $(100 - \frac{m}{4})$ %, la solución que se obtendría sería trivial, defender con probabilidad 1 a aquellos sitios que posean alta pérdida dentro del $(100 - \frac{m}{4})$ % de las pérdidas más altas y el resto de sitios defenderlos con probabilidad 0.

Finalmente, se resuelve y muestra la metodología MUE para así comparar las soluciones, teniendo en cuenta la utilidad esperada, la desviación estándar, el $\text{VaR}_{0.2}$, el $\text{CVaR}_{0.2}$ y como se defienden los diez sitios con mayor pérdida (peores diez sitios).

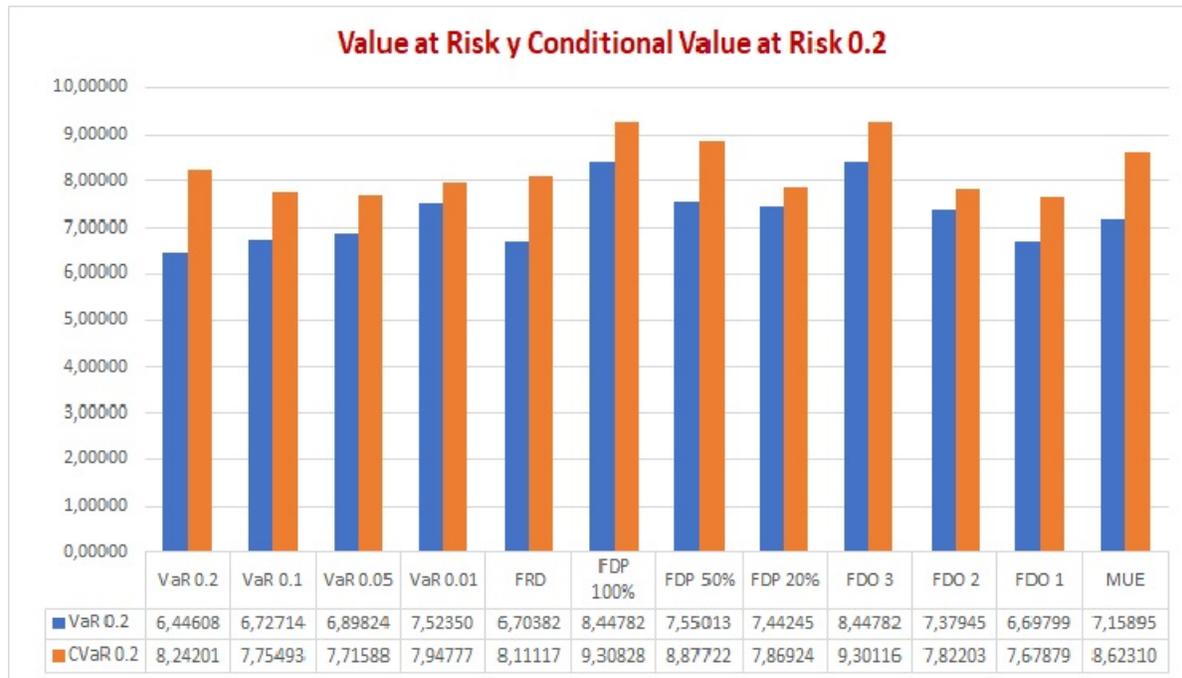


Figura 4.2: VaR y CVaR 0.2 del Experimento 7 sin Restricciones

Se puede apreciar en el gráfico 4.2 que, efectivamente, la metodología $\text{VaR}_{0.2}$ es la más pequeña en términos de $\text{VaR}_{0.2}$, pero en $\text{CVaR}_{0.2}$ no lo es. Además, hay que destacar el buen rendimiento de MUE en VaR, mas no en CVaR teniendo uno de los más grandes. Por otro lado, es fácil notar que algunas metodologías difieren mucho en cuanto a VaR y CVaR. En las primeras cuatro metodologías es evidente, ya que se exige más a medida que ε es más pequeño, de hecho, más adelante se muestra en otro gráfico que, en general, las probabilidades de resultados por sobre el VaR respectivo son 3 veces más pequeñas en $\text{VaR}_{0.01}$ que en $\text{VaR}_{0.2}$.

En las demás metodologías se puede observar que las probabilidades pasado el VaR respectivo son bajas, entonces, si se llega a estar por encima del VaR (20 % de que ocurra), en promedio, se estará muy por encima del VaR, lo cual no es deseable puesto que se tendrá más riesgo en caso de fallar.

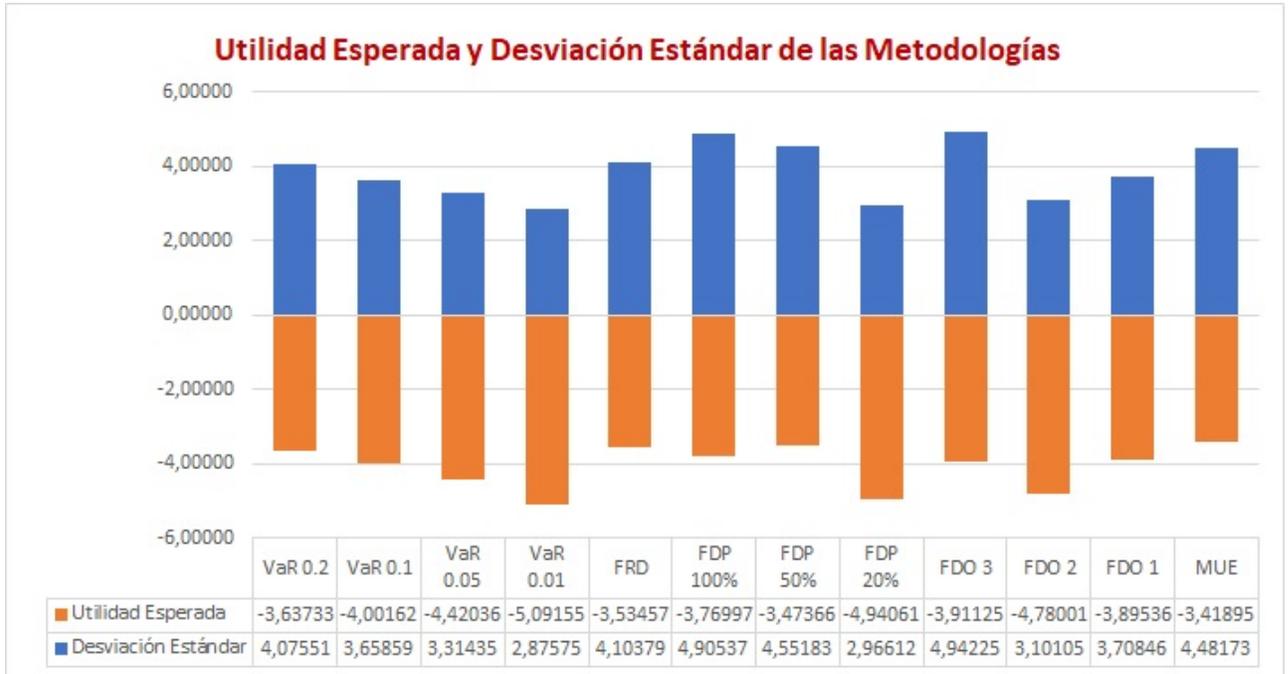


Figura 4.3: Utilidad Esperada y Desviación Estándar del Experimento 7 sin Restricciones

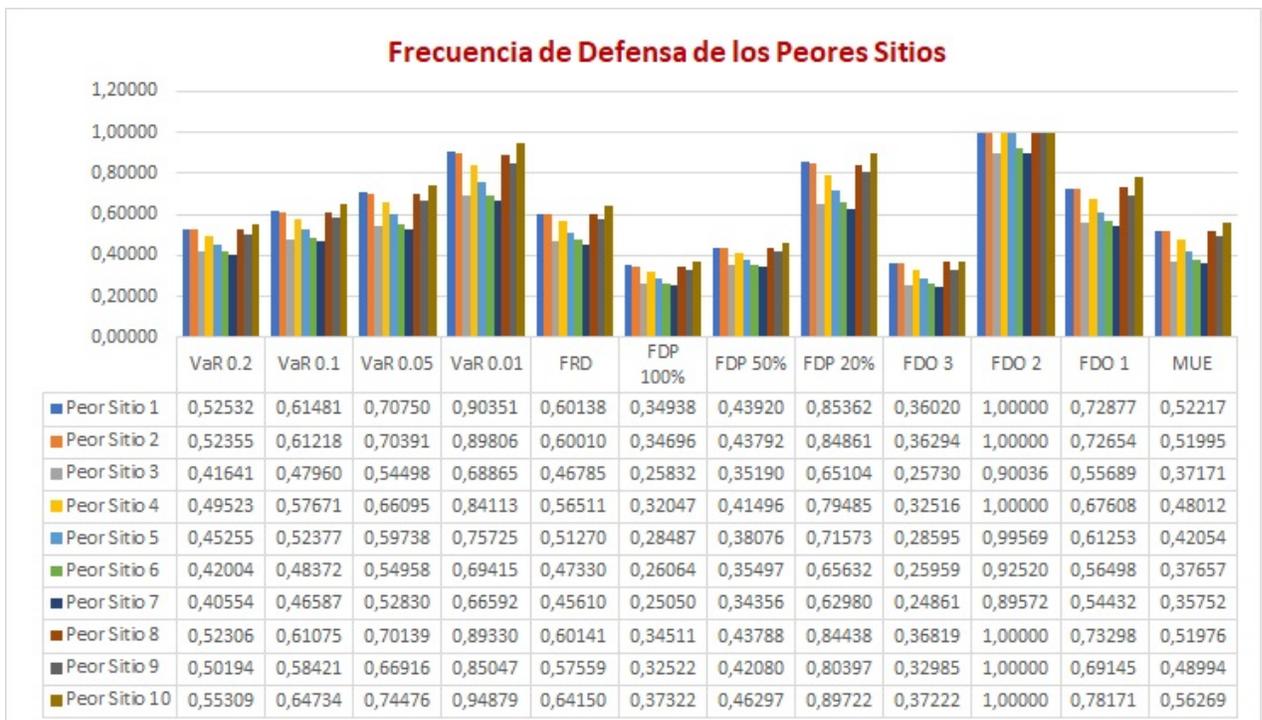


Figura 4.4: Histograma de los 10 Peores Sitios del Experimento 7 sin Restricciones

En la figura 4.3 se obtiene lo esperable, ya que la metodología MUE posee la mayor UE, pero tiene una de las desviaciones estándar más grande. Y es que mientras más se le exige a la función de aversión al riesgo, más pequeña es la utilidad esperada, ya que deja de lado muchos sitios que entregan muy buena recompensa y, al mismo tiempo, evita desviarse de la UE, puesto que se defiende de forma muy segura.

Lo que también se puede observar es que $\text{VaR}_{0,2}$, FDP 100 %, FDP 50 % y FDR se asemejan a la metodología MUE en UE, inclusive hay dos que tienen mejor desviación estándar que MUE. Esto se puede deber a que, por ejemplo, en FDR le conviene defender aquellos sitios con buena recompensa, ya que éstas agregan peso negativo a minimizar, mientras que el $\text{VaR}_{0,2}$ defiende muchos sitios antes de llegar al sitio referenciado en el índice del VaR, entre aquellos sitios existen varios con buenas recompensas, lo cual mejora la UE.

La demás metodologías mejoran mucho la desviación estándar, teniendo más seguridad del resultado, pero son poco deseables, ya que en promedio se obtendrá menos ganancia al largo plazo.

En la figura 4.4 se puede observar un patrón, esto se debe a que cada sitio otorga recompensa o peso en las funciones respectivas de igual manera, además de que depende de los valores γ , ya que al ordenar los sitios de menor a mayor con respecto a su valor γ_i , las probabilidades que se observan quedan ordenadas de mayor a menor. Esto último se debe a que los 10 últimos sitios tienen una pérdida muy similar (o ganancia para el seguidor), por lo tanto, para el seguidor le serán menos apetecibles aquellos sitios que posean mucha ganancia para el líder (mayor pérdida para el atacante), puesto que el seguidor quiere maximizar su utilidad. Además, el líder también sabe esto, entonces defenderá aquellos sitios con poca ganancia, porque son los más probables que el seguidor ataque. Lo anterior también aplica para la metodología MUE, ya que conviene gastar recursos en sitios que sean probables de que ataquen y así asegurar algún tipo de ganancia (en vez de malgastar recursos en un sitio que es menos probable que ataquen).

Notemos que la metodología MUE defiende pobremente los peores sitios, mas no tan pobremente como FDP 100 %, FDP 50 % y FDO 3, puesto que estas últimas tienen en cuenta los sitios con casi nula pérdida con el mismo peso que el resto. Además, se toma en cuenta sólo pérdidas en FDP 100 %, FDP 50 % y FDO 3, mientras que en MUE toma en cuenta las ganancias también.

Por último, notar que el sitio con pérdida más grande dentro de cada solución es el que se defiende más usualmente y se repite que el sitio que más se defiende siempre se encuentra dentro de los 10 sitios con mayor pérdida.

Ahora mostraremos diferentes funciones de distribución de cada solución y un par de histogramas de los resultados para explicar cierto patrón de comportamiento interesante.

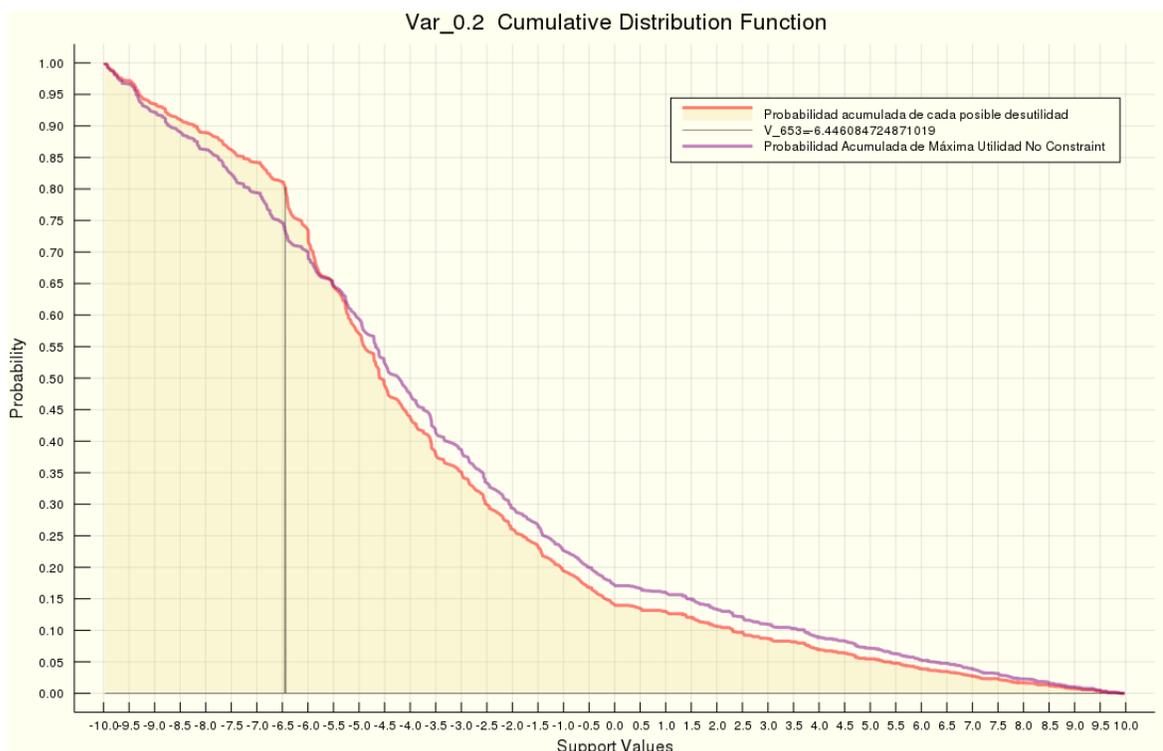


Figura 4.5: Función de Distribución Acumulada del Experimento 7 $VaR_{0,2}$

En el gráfico 4.5 se comparan la función de distribución acumulada (fda) de la metodología MUE con la metodología $VaR_{0,2}$. La línea vertical indica el valor exacto del $VaR_{0,2}$ de la metodología con el mismo nombre, cuyo valor es $V_{653} \approx 6,4461$.

Notemos que inicialmente la metodología MUE es mejor, pero a medida que se aproxima a los sitios con pérdida mayor, el $VaR_{0,2}$ lo empieza a superar justo antes de llegar al punto crítico en donde se asegure con al menos un 80% que la desutilidad será menor o igual a 6.4461. Notemos que ya habíamos visto una semejanza anteriormente entre estas dos metodologías, viendo que los peores sitios se defienden de manera similar. Por otro lado, el CVaR y la desviación estándar son mejores en la metodología $VaR_{0,2}$. Estos resultados se deben a como se afrontó cada problema, el primero busca satisfacer su utilidad al máximo sin tener en cuenta la aversión al riesgo, mientras que el segundo busca protegerse de obtener una pérdida no muy grande y, en caso de que eso ocurriera, asegura que no sea tanto más grande la pérdida, en promedio, por sobre 6.4461.

También cabe destacar que la fda de la metodología MUE es similar a una línea recta, mientras que la otra fda se ven claramente tres segmentos, el primero cuando los posibles resultados son ganancia, el segundo cuando los posibles resultados son negativos, pero no tanto. Y, el tercero, cuando se alcanza el valor 6.4461, en donde se defiende mucho más esos sitios, ya que se quiere evitar entrar en ese rango de pérdidas.

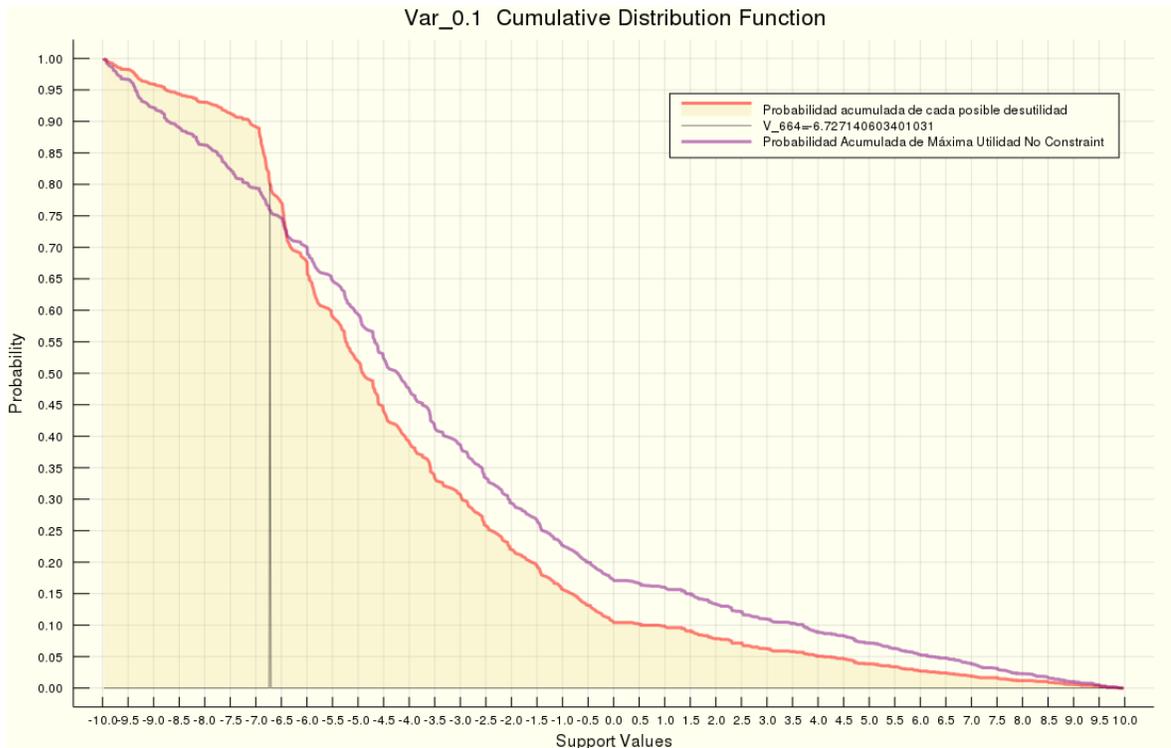


Figura 4.6: Función de Distribución Acumulada del Experimento 7 $\text{VaR}_{0,1}$

En el gráfico 4.6 se puede apreciar una diferencia mucho mayor, tanto en la fda como en la utilidad esperada, desviación estándar, etc. Esto se debe a que la metodología $\text{VaR}_{0,1}$ es 10 % más aversa al riesgo y ese extra de seguridad que se busca conlleva a tener que dejar más de lado las posibles ganancias y los sitios con pérdida mediana. Además, es fácil notar que el $\text{VaR}_{0,1}$ es mayor al $\text{VaR}_{0,2} \approx 6.7271$. Por otro lado, el $\text{CVaR}_{0,2}$ de esta metodología es menor, lo cual mejora mucho como defensa, ya que si el líder pierde, su pérdida no será tanto más grande que, incluso, 6.4461.

En los gráficos 4.7 y 4.8 se observa el mismo patrón, pero más acentuado. Lo que sí vale la pena mencionar es que el $\text{CVaR}_{0,2}$ de la metodología $\text{VaR}_{0,05}$ es uno de los más pequeños de todas las metodologías, siendo superado sólo por la metodología FDO 1. Esto se podría deber a que a partir del valor de la probabilidad del sitio representando a 701 de \mathcal{V} empieza a ser relevante ($\approx 0,488298$), pero no tanto, lo que conlleva a que estos sitios pasen a ser parte del posible resultado que pueda ser la desutilidad. Por otro lado, la metodología $\text{VaR}_{0,01}$ tiene un valor de $\text{CVaR}_{0,2}$ mayor, pero esto se explica debido a dos cosas. Lo primero es que el $\text{VaR}_{0,2}$ es más alto y lo segundo, a partir del sitio 725 el valor de p_v empieza a ser relevante ($\approx 0,8814$), pero mucho más relevante que en la metodología anterior, de hecho casi el doble. Además, esto explica el por qué la diferencia entre VaR y CVaR de la última metodología comentada es la más pequeña (a medida que empieza a tomar relevancia, ésta es fuerte).

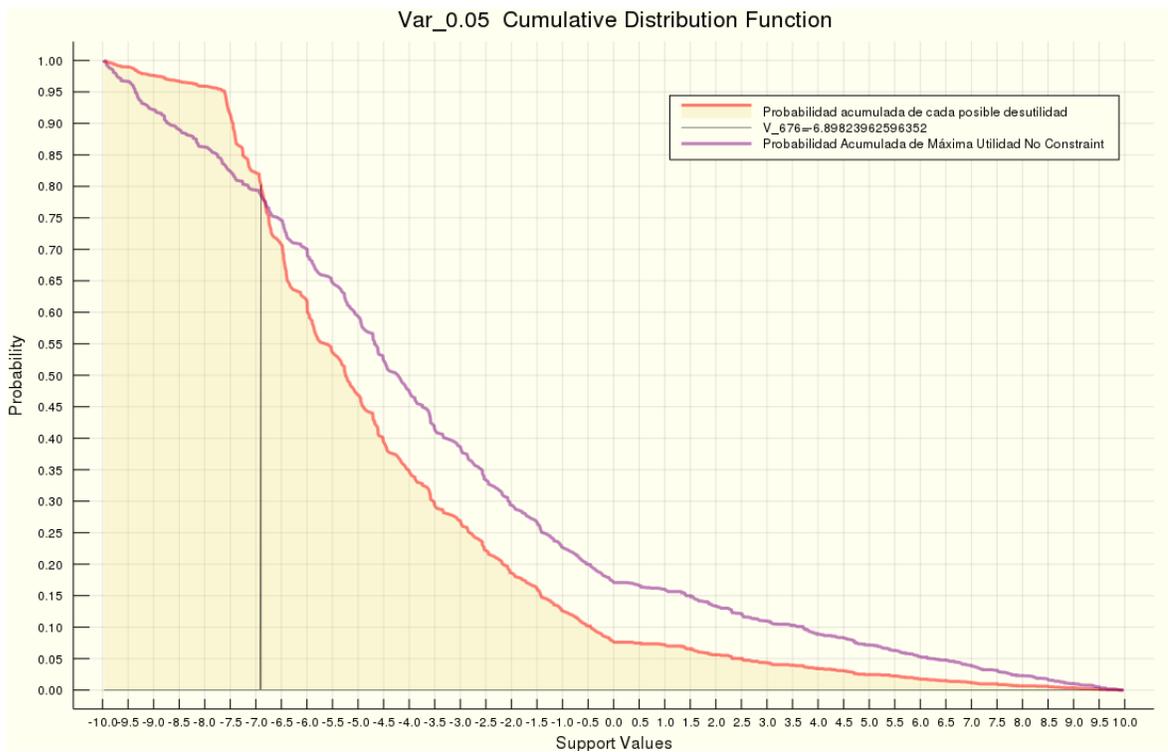


Figura 4.7: Función de Distribución Acumulada del Experimento 7 $VaR_{0,05}$

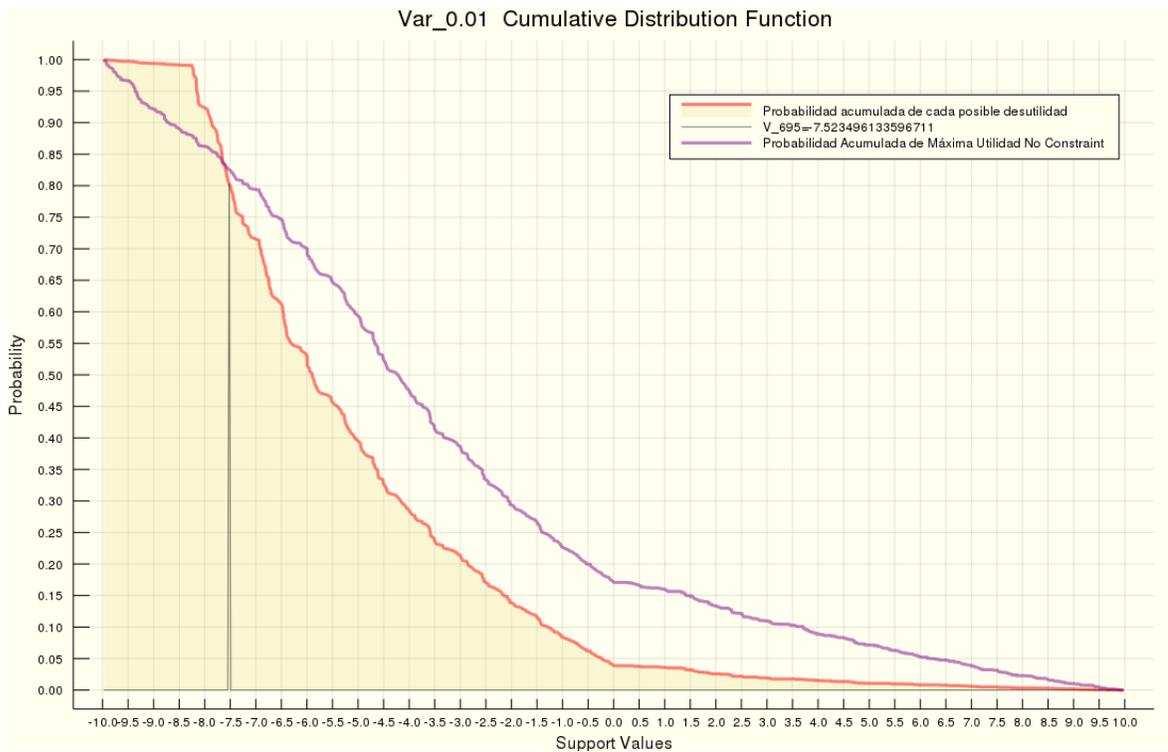


Figura 4.8: Función de Distribución Acumulada del Experimento 7 $VaR_{0,01}$

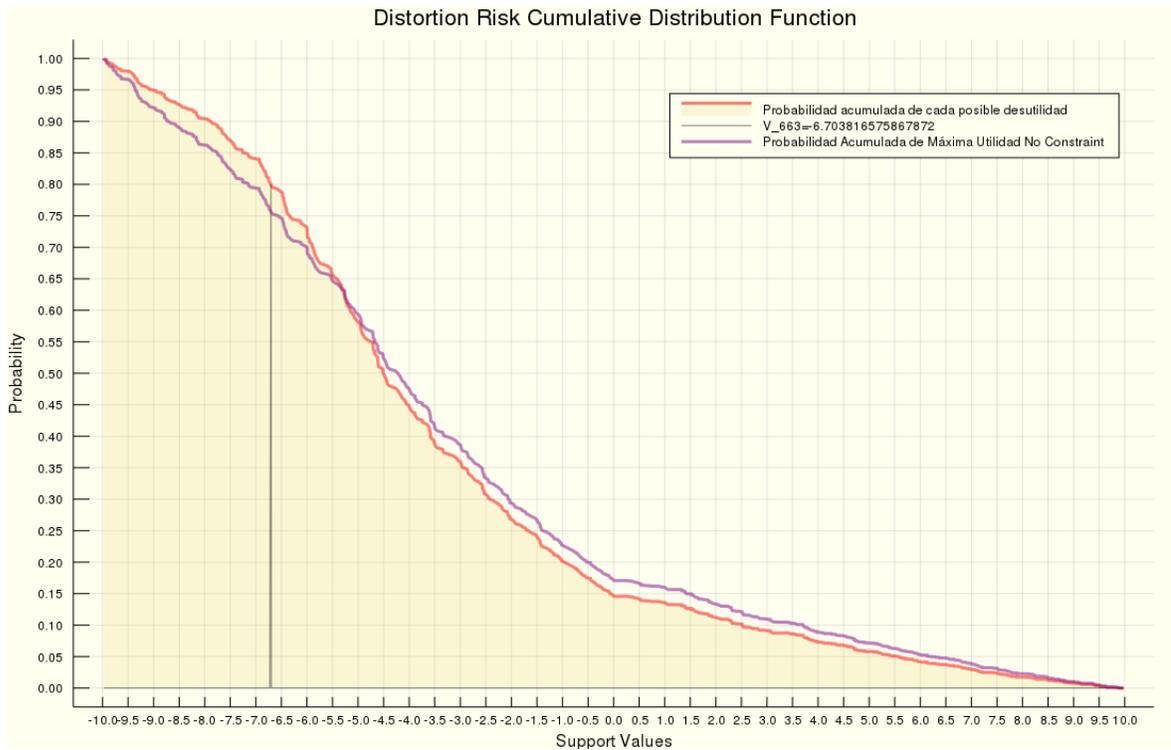


Figura 4.9: Función de Distribución Acumulada del Experimento 7 FRD

En el gráfico 4.9 se puede observar una similitud a la vista en el gráfico 4.5. A pesar de esto existen diferencias, ya que la metodología FRD tiende a buscar defender los sitios con mayor pérdida y los sitios con mayor ganancia, asemejándose más a MUE. De hecho, por esta misma razón los peores sitios se defienden más, la utilidad esperada es mejor y la desviación estándar empeora (c/r a $VaR_{0,2}$). Esta metodología podría considerarse como una de las mejores puesto que no cambia tanto la utilidad esperada en el largo plazo, mientras mejora todo lo demás.

En el gráfico 4.10 se nota mucho más claro la similitud, pero el VaR y CVaR son peores. De hecho, la defensa de los peores 10 sitios es más baja. Además, en utilidad esperada y desviación estándar la metodología es más baja que en la MUE.

El gráfico 4.11 muestra una gran diferencia en la fda, pero esto es evidente, ya que al ignorar el 80% de las pérdidas, siempre convendrá no defender los sitios asociados a los índices en $v \in \{n, n + 1, \dots, 720\}$. Esto implica que se le castigará al líder en sitios con pagos positivos o pagos negativos medios/bajos. Curiosamente, al llegar al valor $VaR_{0,2} \approx 7.4425$ de la metodología en cuestión, la forma de defender empieza a ser superior a la de la metodología MUE. Esto podría originarse dado que el valor 7.4425 está en el índice 691, cercano al índice 721. Conectado con lo anterior es muy clara la razón del por qué la metodología MUE tiene mejor valor $VaR_{0,2}$, pero peor $CVaR_{0,2}$. Es muy probable caer en las pérdidas en el rango $\{691, \dots, 720\}$ si se defiende como indica la metodología FDP20%.

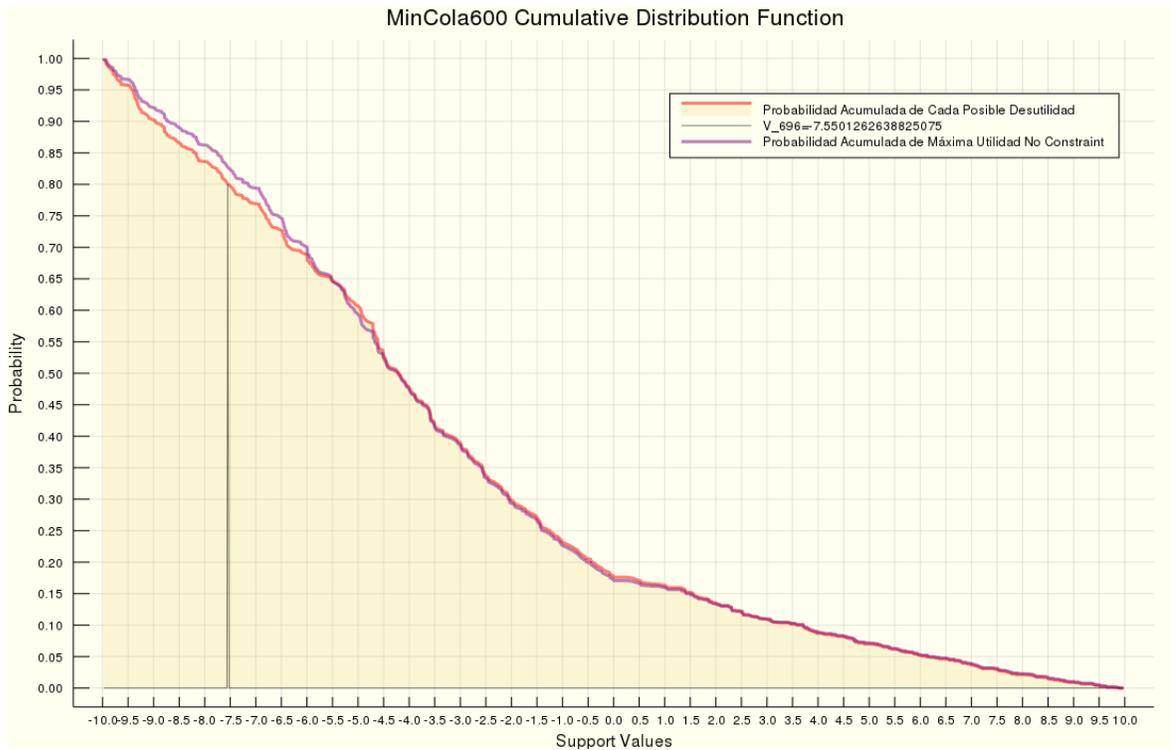


Figura 4.10: Función de Distribución Acumulada del Experimento 7 FDP50 %

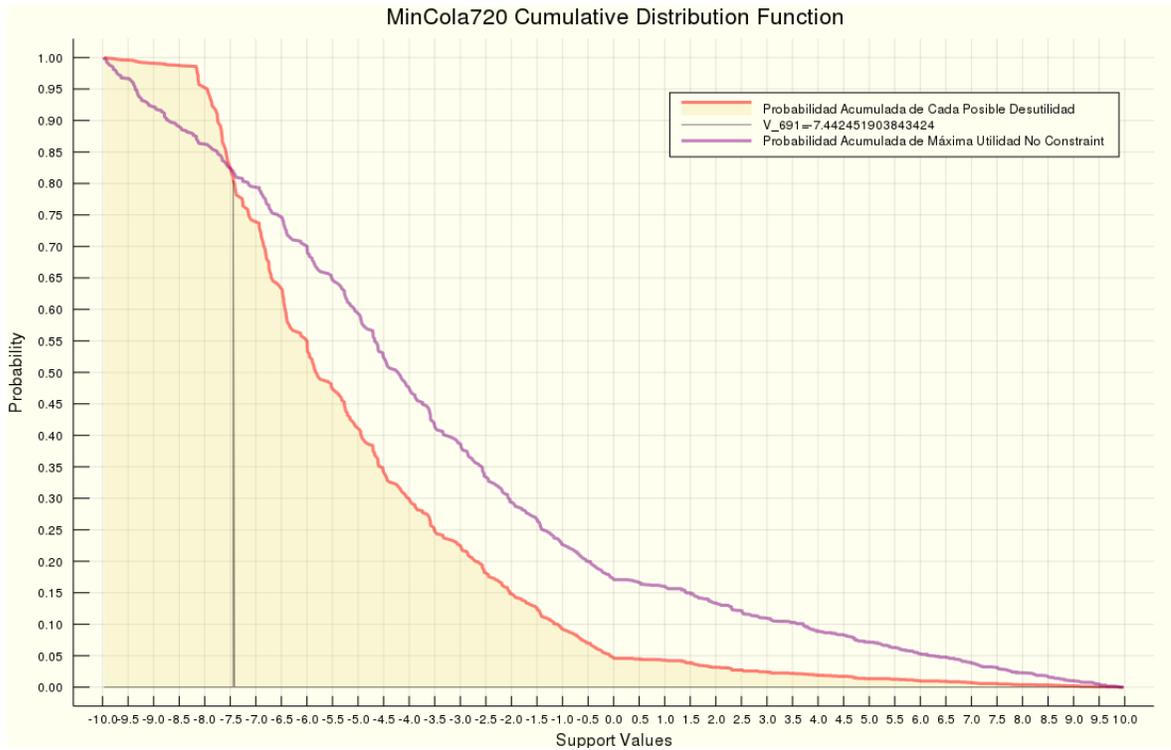


Figura 4.11: Función de Distribución Acumulada del Experimento 7 FDP20 %

En los gráficos 4.12 y 4.13 se observa similitud con la solución del gráfico 4.10, teniendo también que la mejora de la solución con aversión empieza al llegar al valor del $VaR_{0,2}$. Estas metodologías agregan más peso a medida que la pérdida es mayor. Esto explica el por qué la diferencia entre el $VaR_{0,2}$ y $CVaR_{0,2}$ es tan pequeña, aunque sólo FDO1 tiene mejor $VaR_{0,2}$. Lo último se explica puesto que el crecimiento de los pesos en FDO2 es exponencial (los 10 peores sitios se defienden casi con frecuencia ≈ 1), entonces los sitios que realmente empiezan a importar tienden a ser los que están en los índices finales de \mathcal{V} . Por esta misma razón, el $CVaR_{0,2}$ es mejor en ambas metodologías en comparación a MUE.

Finalmente, es claro que los peores sitios se defienden de mejor manera, las desviaciones estándar son mejores y ambas utilidades esperadas son peores, pero en FDO1 tan sólo por ≈ 0.4 . Y, En las metodologías FDO3 y FDP100% se ve que se empeora en todo, incluso, en las defensas de los peores sitios. Y esto es evidente, ya que se le exige demasiado al líder. En ambos casos se busca minimizar la suma de probabilidades (con el mismo peso cada probabilidad) tomando en cuenta todas las pérdidas. Sus gráficos respectivos se presentarán en el anexo.

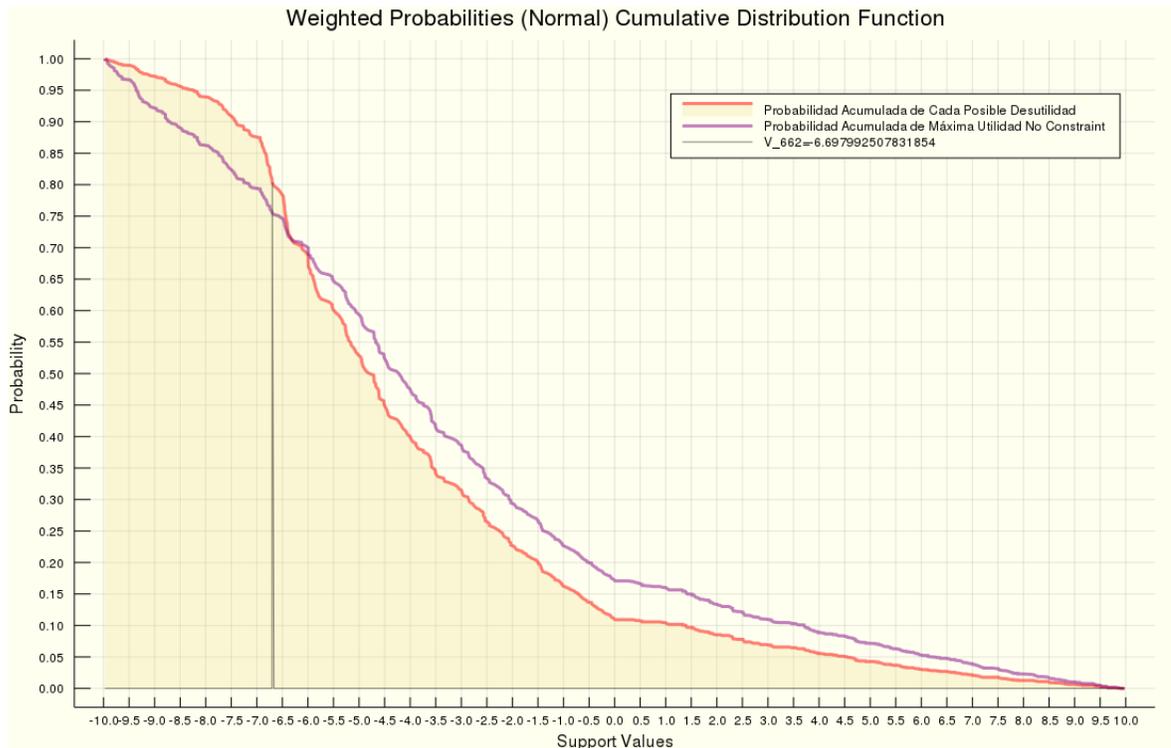


Figura 4.12: Función de Distribución Acumulada del Experimento 7 FDO1

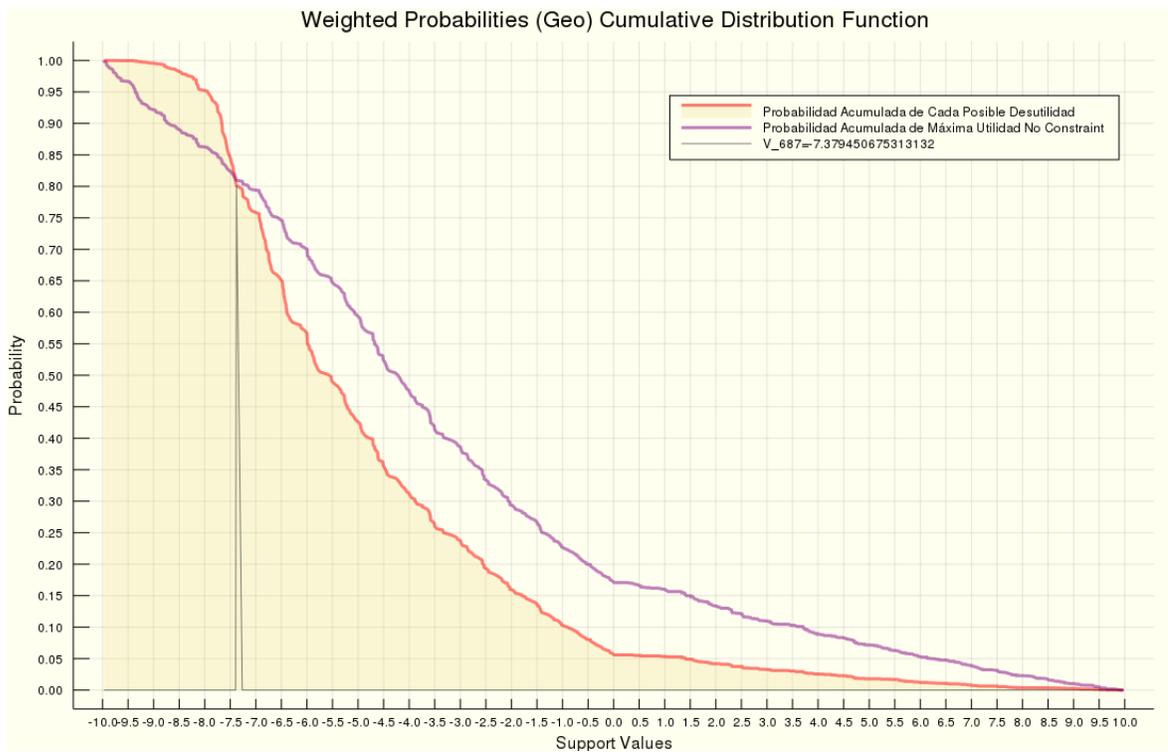


Figura 4.13: Función de Distribución Acumulada del Experimento 7 FDO2

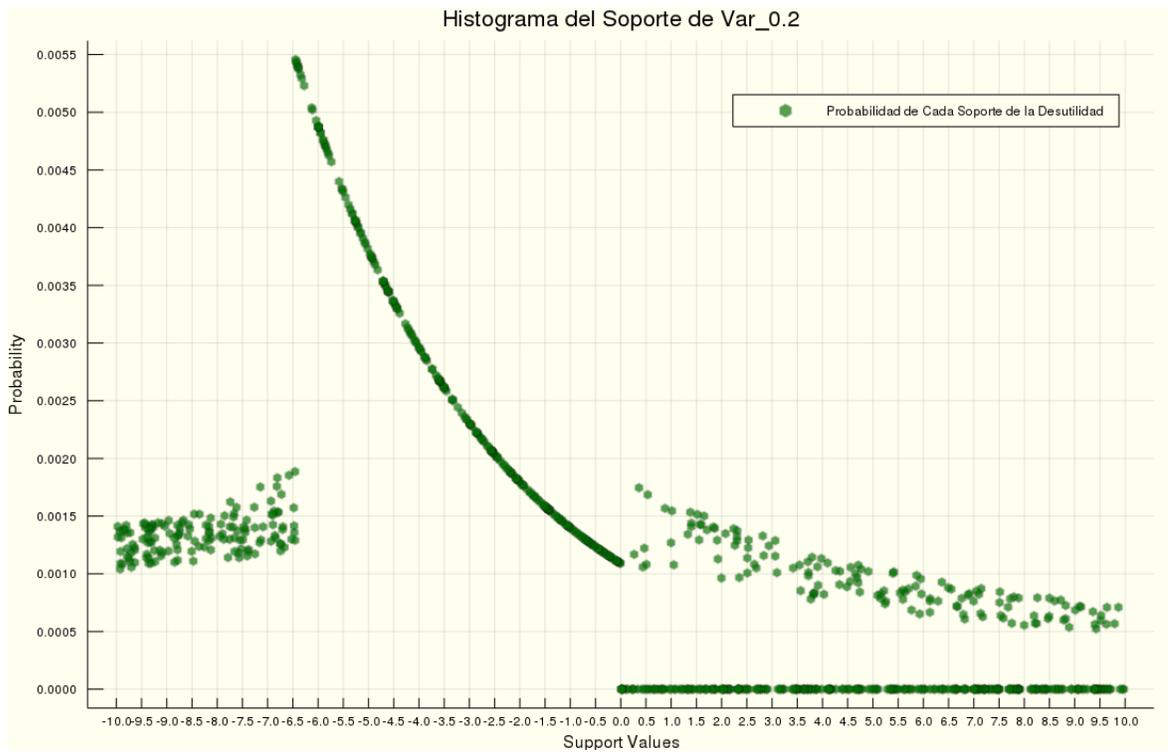


Figura 4.14: Histograma del Soporte de $VaR_{0,2}$ del Experimento 7

En el gráfico 4.14 se muestra el histograma de los posibles resultados de la desutilidad de la metodología $VaR_{0,2}$, pudiéndose apreciar un comportamiento creciente y exponencial al comenzar las pérdidas. Esto se debe a que se le exigió tener aversión al riesgo y, además, el resultado más probable concuerda con el valor del $VaR_{0,2}$, que de ahí en más decae abruptamente, lo cual es evidente, puesto que esa zona de pérdidas está asegurada en un 80 %. Por otro lado, algunas ganancias se ven posibles, pero de forma similiar a las pérdidas más grandes. Además, la solución de esta metodología va colocando en riesgo cada vez más grande a los sitios cuya pérdida va creciendo, ya que estos integran cada vez más ganancia para el seguidor.

El comportamiento exponencial de la solución se debe a la forma de las probabilidades del ataque del seguidor, las cuales son QR. Y cómo se está teniendo aversión al riesgo al ataque del oponente, no es de extrañar este patrón, que dicho sea de paso, se repite en todas las funciones que, exitosamente, entregan una mejor solución que MUE en cuanto a desviación estándar.

El gráfico 4.15 muestra el histograma del soporte de la metodología $VaR_{0,05}$, la cual podemos observar que sigue el mismo patrón, sólo que esta vez la pendiente llega hasta el valor del $VaR_{0,05}$ y luego cae. Notemos que ahora es muy poco probable obtener alguna ganancia o alguna pérdida grande con un 5 %. Esto es debido a que se está siendo mucho más averso al riesgo que antes.

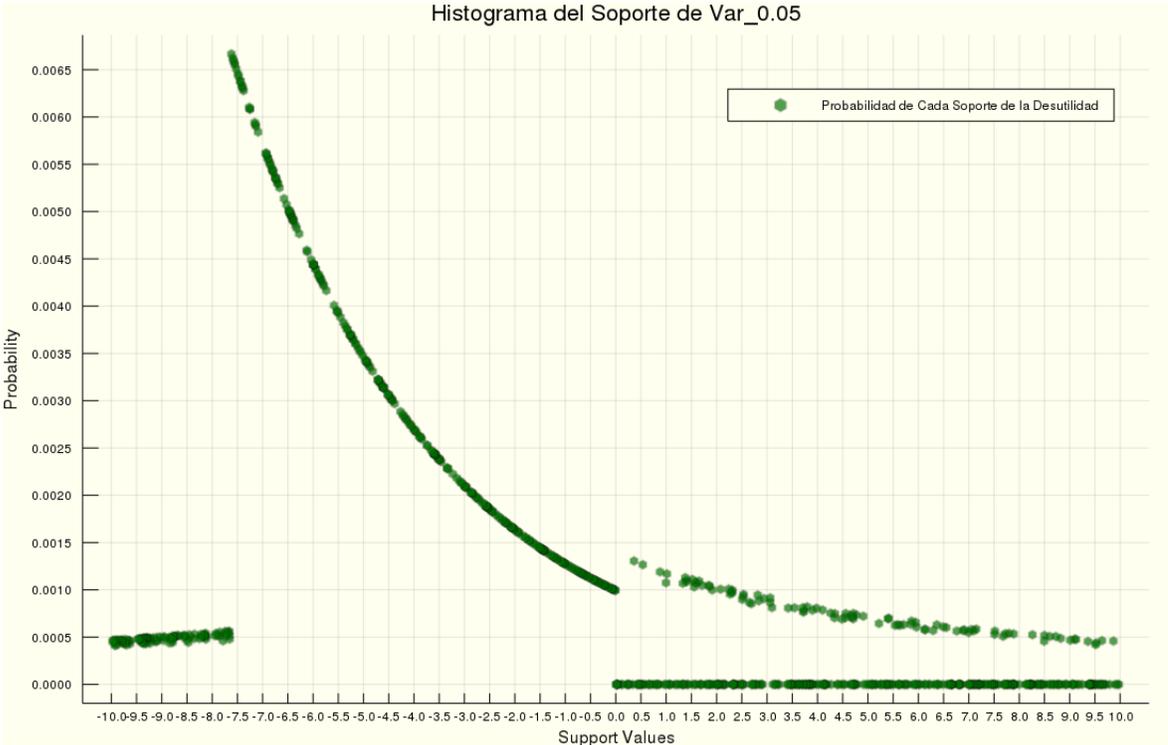


Figura 4.15: Histograma del soporte de $VaR_{0,05}$ del Experimento 7

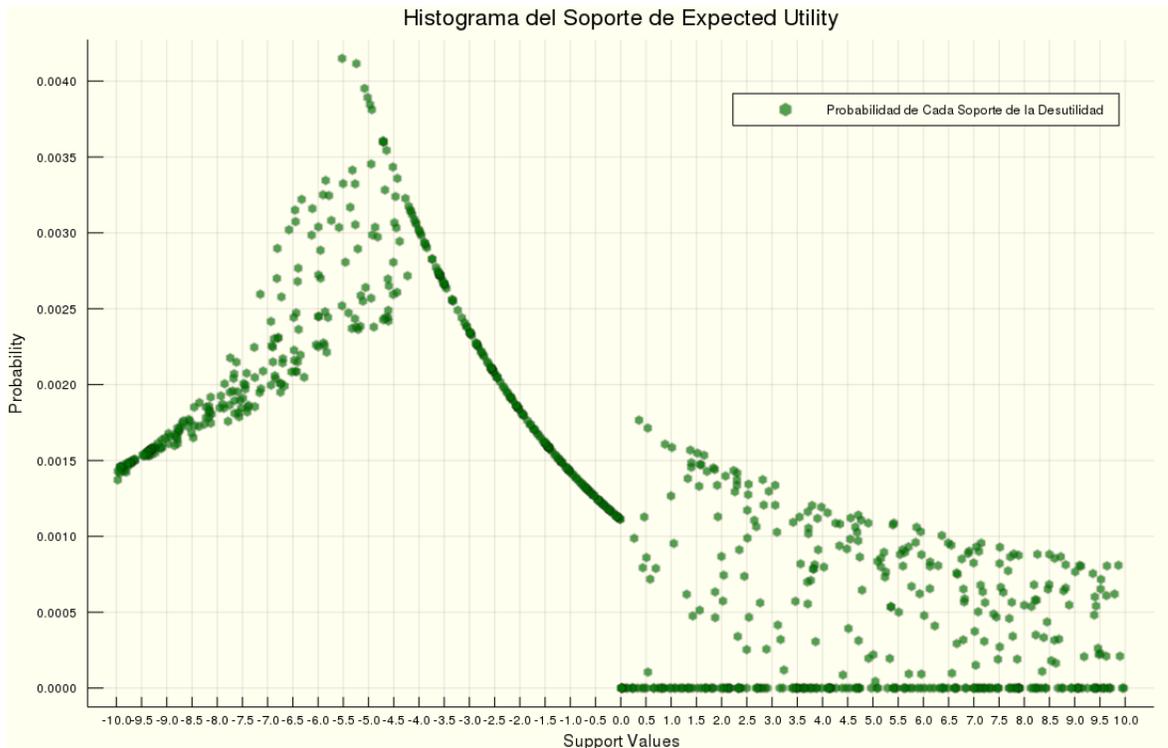


Figura 4.16: Histograma del soporte de MUE del Experimento 7

En el gráfico 4.16 se observa un patrón similar, pero tengamos en cuenta que se está maximizando la utilidad esperada. Por esta razón, los sitios con peor pérdida son claramente más probables que antes y también las ganancias. Por otro lado, es curioso notar que los sitios con pérdida mediana siguen teniendo mayor chance de ocurrir. Esto se debe a que el líder desea maximizar utilidad, entonces buscará tener buena ganancia y no tanta pérdida, dejándolo de lado los sitios con pérdida mediana. Esto es interesante, ya que se tiene cierto patrón de aversión al riesgo, sin tenerla realmente, claro está que esto podría cambiar si el sitio con pérdida pequeña tiene la mayor ganancia o viceversa, o bien si nuestro oponente es muy racional o muy irracional.

Una observación importante de todas las metodologías es que las probabilidades de los posibles resultados tienen el mismo comportamiento. Notemos que hay muchos sitios cuya ganancia jamás será posible de ganar (puntos que están en el cero). Esto es porque al tener una cantidad de recursos de 60 disponibles, mientras se defienden 400 territorios, imposibilita defender a todos los sitios de forma eficiente. Debido a lo anterior, aquellos lugares en que no se defienden ($x_i = 0$), en el contexto del atacante, él percibirá toda la ganancia en caso de atacarlos. Por esa razón, se visualiza una curva exponencial creciente a medida que la pérdida va aumentando, puesto que mientras más pérdida tenga el sitio, más ganancia tendrá el seguidor, i.e, es más probable que el líder obtenga dichas pérdidas. Al llegar a cierto punto, el líder empezará a defender los puntos que posean mucha pérdida, haciéndolos cada vez menos deseables para el seguidor, lo que conlleva a una menor probabilidad de ataque, o sea, una menor probabilidad de que dichas pérdidas se observen. Claro está que a medida que la pérdida aumenta, el líder tendrá que ir colocando una mayor cantidad de recursos a dichas localizaciones, o sea, los valores de x_i irán siendo crecientes. Conectado con este precedente, estos sitios que se defienden tendrán cierta chance de entregar su ganancia y se ven represen-

tados en los puntos sobre 0 que se observan en la parte positiva de los posibles resultados en los últimos tres gráficos. En estos, la probabilidad de que el seguidor los ataque dependerá de la pérdida que él perciba, por lo tanto, aquellos sitios que tengan menos ganancia para el líder (menor pérdida para él) son los más probables que ataque (ya que maximiza utilidad). Esto, al saberlo el líder, dispondrá de una mayor cuota de recursos a dichos sitios, obteniéndose así una curva exponencial decreciente a medida que las ganancias aumentan (y también serán decrecientes los valores de x_i). Claro está que esto último es a grandes rasgos, puesto que se pueden observar sitios con poca ganancia, la cual es casi imposible de obtener. Esto se podría deber a la forma que tiene el QR y a la pérdida asociada a dichos lugares.

4.2. Resultados de Restricciones de Aversión al Riesgo

En esta sección se procede a presentar la aversión al riesgo desde el punto de vista de restricciones, específicamente, aquellas presentadas en el capítulo 1.3. En este caso se resolvió la metodología MUE, la cual es la misma de antes. La siguiente metodología resuelta fue MUE RO, cuya restricción asociada es la restricción de Oportunidad. En este caso $\tilde{V} = \mathcal{V}(2n - 9) \implies \bar{v} = 2n - 9$, y además $\varepsilon = 10^{-3}$. Esto fuerza que la probabilidad, de que la desutilidad sea alguna de las mayores 10 pérdidas, sea más pequeña o igual a 10^{-3} .

La siguiente es la metodología MUE RPODE, la cual resuelve MUE teniendo como restricción a la restricción de Dominancia Estocástica de Primer Orden. Esta restricción es demasiado fuerte, por ejemplo, si se colocase a \tilde{x} , como sigue:

$$\tilde{x}_i = \begin{cases} 1 & \text{si } \text{orden}(i) \in \{741, \dots, 800\} \\ 0 & \text{Si no} \end{cases}$$

Lo cual significa defender los 60 sitios con mayor pérdida con frecuencia 1. Entonces, al incluir esta restricción, el único punto factible sería \tilde{x} , ya que para $\tilde{v} = 740$ se tendría que $\sum_{v \geq 741} p_v(x) \leq 0 \implies p_v(x) = 0 \forall v \geq 761 \implies x_{\text{orden}(v)} = 1 \forall v \geq 741$. Por otro lado, si la defensa utilizada a ser mejorada es la misma que la de la metodología MUE tampoco tiene sentido, ya que la solución sería la misma. En resumen, para agregar mucha aversión al riesgo se utilizó la siguiente defensa \tilde{x} :

$$\tilde{x}_i = \begin{cases} 0,96 & \text{si } \text{orden}(i) \in \{741, \dots, 800\} \\ 0,04 & \text{si } \text{orden}(i) \in \{681, \dots, 740\} \\ 0 & \text{si no} \end{cases}$$

La cual significa defender los 120 sitios con mayor pérdida, obviamente, dándole muchísimo peso a los 60 peores y no tanto a los siguientes 60.

Por último, la metodología MUE RRD, cuya restricción asociada es la restricción de Riesgo Distorsionado. La función utilizada para esta restricción es la misma que se utilizó al resolver la metodología FRD y $\tilde{\rho}$ varió en cada experimento, ya que esta restricción tuvo un comportamiento bastante sensible. Si $\tilde{\rho}$ era muy alto, esta restricción no agregaba información

útil. Si era muy bajo, esta restricción era infactible. Por lo tanto, el valor en donde realmente agrega aversión al riesgo está restringido a cierto intervalo que cambia dependiendo de los datos del problema. Entonces para lograr encontrar un valor útil se observó el valor óptimo de la metodología FRD y se encontró, probando diferentes valores de $\tilde{\rho}$, el valor útil, de tal forma que $d(\tilde{\rho})$ fuera apenas (alrededor de 0.1) más grande que el valor óptimo obtenido al resolver la metodología FRD.

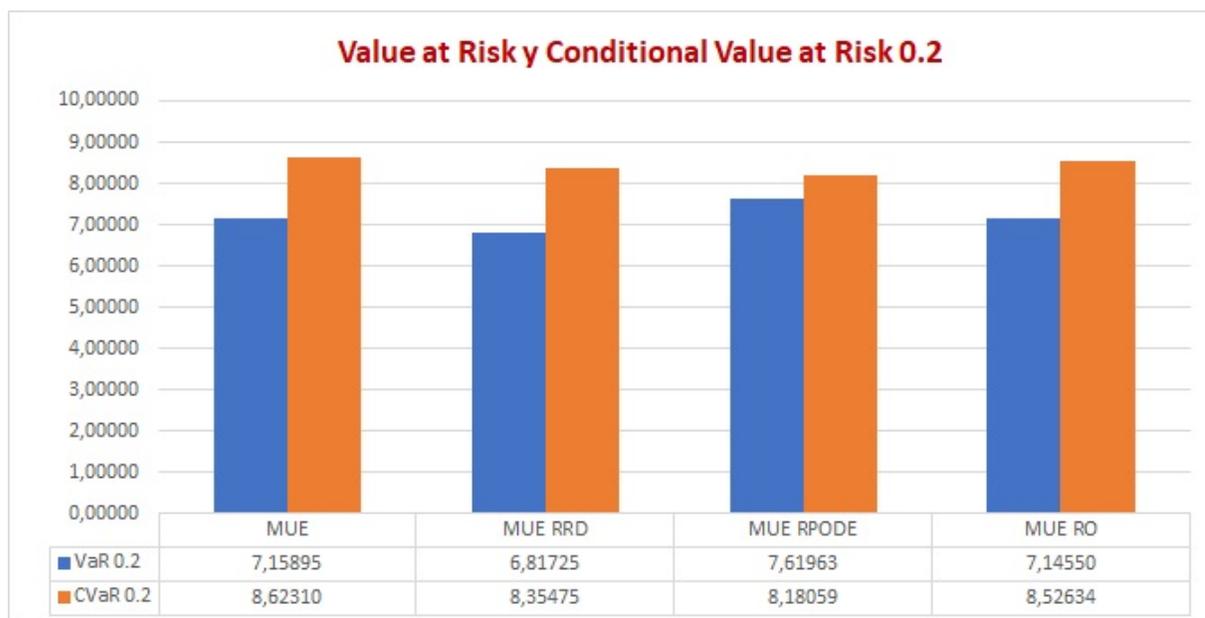


Figura 4.17: VaR y CVaR 0.2 del Experimento 7 con Restricciones

En la figura 4.17 se puede visualizar cierta semejanza con respecto al $VaR_{0,2}$ en tres metodologías y todas las metodologías con respecto al $CVaR_{0,2}$. Aquellas metodologías que mejoran el $VaR_{0,2}$, al mismo tiempo, mejoran el $CVaR_{0,2}$. Esto podría deberse a que el comportamiento de la solución es similar, pero al agregar un poco de aversión al riesgo, los sitios con mayor pérdida se tienden a defender más. Luego, al bajar el VaR, se bajó el valor del CVaR al mismo tiempo. Por otro lado, la metodología MUE RPODE aumenta el valor, ya que la solución de esta misma es muy similar a la defensa inicial que se utilizó en la restricción y dicha solución inicial es muy aversa al riesgo, es muy distinta a la forma de defender de la metodología MUE. Aún así, mejora el valor del $CVaR_{0,2}$, lo cual se explica por como se defienden los sitios un poco más adelante del valor del $VaR_{0,2}$, siendo así muy probable caer en un rango cercano a 7.6163. En general, se observa un buen comportamiento de estas metodologías porque mejoran todo, sin cambiar tanto el comportamiento de querer maximizar utilidad.

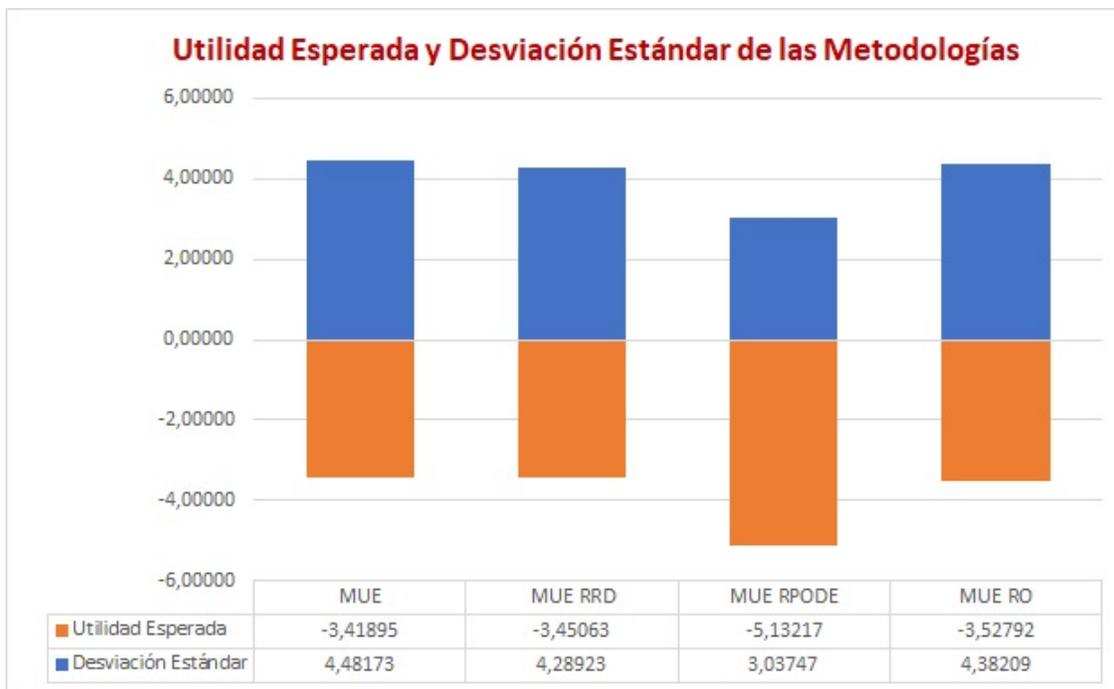


Figura 4.18: Utilidad Esperada y Desviación Estándar del Experimento 7 con Restricciones



Figura 4.19: Histograma de los Peores Sitios del Experimento 7 con Restricciones

En la figura 4.18 se puede observar el precio a pagar por haber mejorado el $VaR_{0,2}$, $CVaR_{0,2}$ y la desviación estándar. La metodología que paga mucha utilidad esperada es la MUE RPODE, mientras que las otras dos difieren muy poco en este valor, algo bastante deseable y útil.

En la figura 4.19 se observa la mejora de la defensa de estos sitios, pero la de la metodología MUE RPODE es bastante engañosa, ya que el precio a pagar es bastante alto, como se observa en los gráficos anteriores y en los siguientes a mostrar. Además, el patrón de la defensa sigue siendo similar al resto, mientras más ganancia posea el sitio, menos se defiende. Por lo tanto, al ordenar con respecto a γ_i de menor a mayor, estas probabilidades quedarán ordenadas de mayor a menor. Este gráfico es muy similar al gráfico 4.3, puesto que ocurren cosas similares con las metodologías y las explicaciones son las mismas.

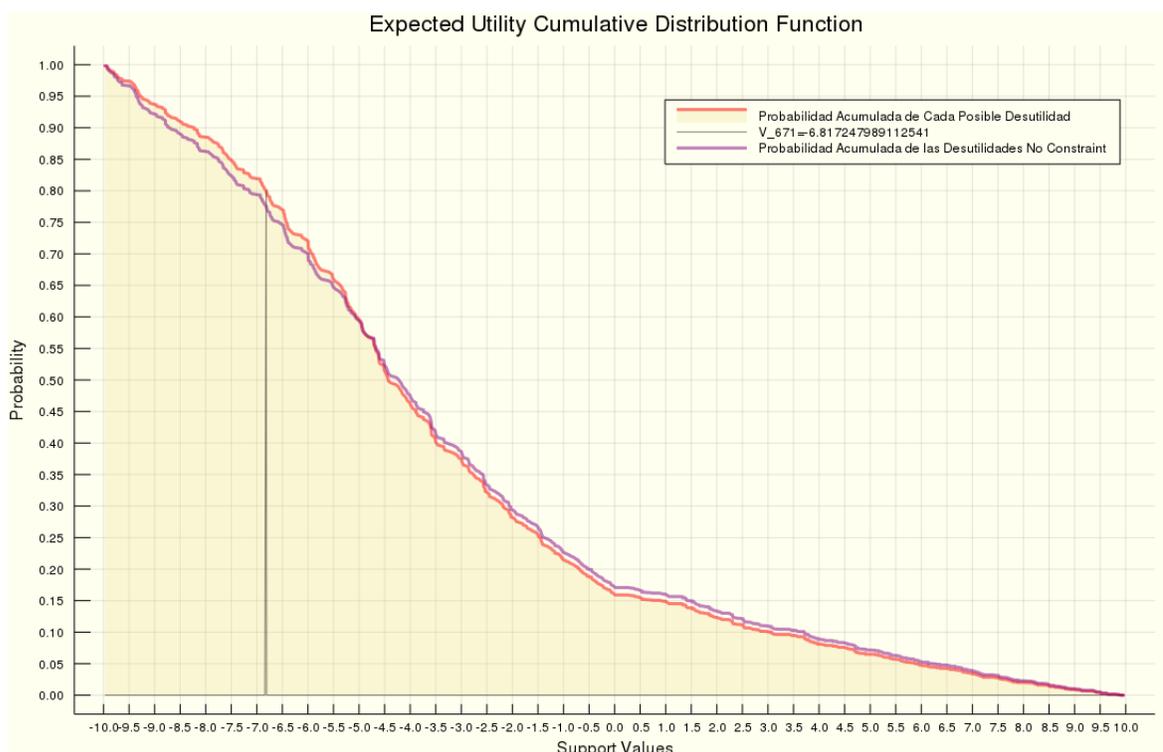


Figura 4.20: Función de Distribución Acumulada del Experimento 7 MUE RRD

El gráfico 4.20 muestra claramente lo ya antes mencionado con respecto a la similitud que tienen estas dos metodologías. Además, podemos observar que cuando las pérdidas empiezan a ser altas, la fda de la metodología MUE RRD supera a la de MUE, agregando exitosamente la aversión al riesgo que se pedía mediante la restricción de riesgo distorsionado. La razón de por qué existe esta semejanza tiene que ver con dos cosas, la primera es que al ser una restricción muy sensible la aversión al riesgo que agrega tiende a ser poca. La segunda es que esta restricción le da peso negativo a los posibles resultados de ganancia, por lo tanto, el defender sitios con grandes ganancias tiende a ser deseable, teniendo así un comportamiento similar a la de maximizar utilidad.

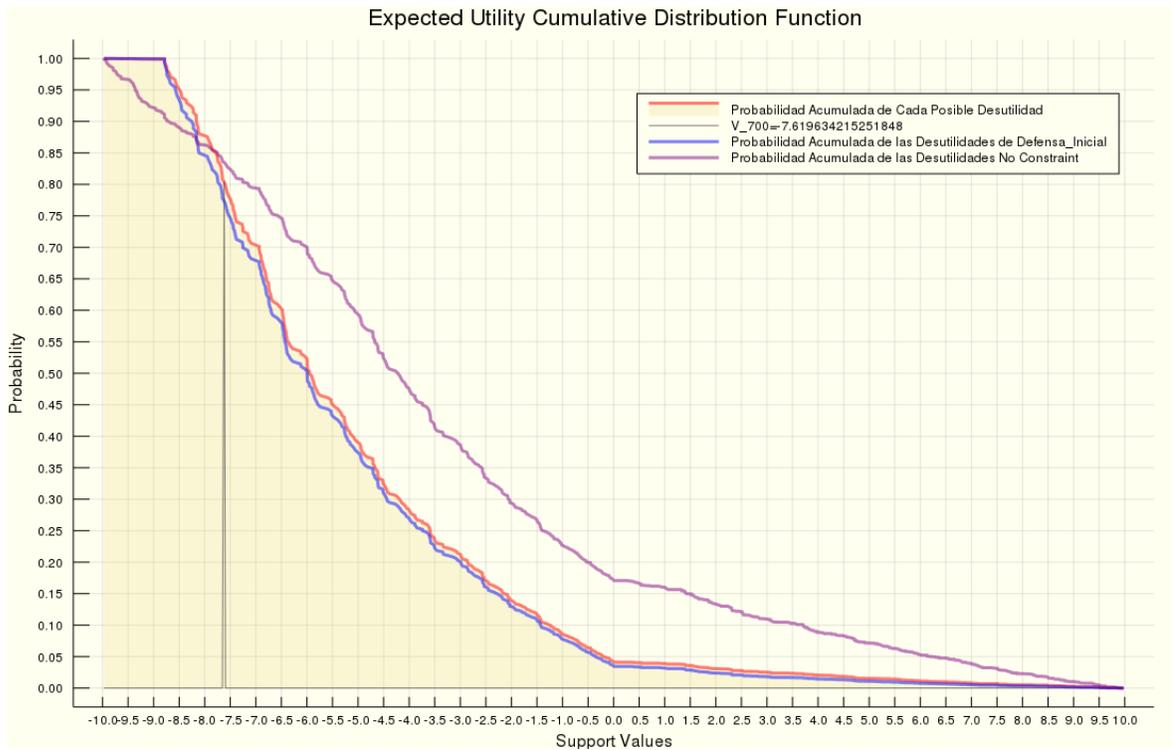


Figura 4.21: Función de Distribución Acumulada del Experimento 7 MUE RPODE

En este gráfico se observan 4 líneas, la horizontal ya se mencionó. La roja y morada son las mismas que aparecen antes, siendo las soluciones de las metodologías MUE RPODE y MUE, respectivamente. La línea azul representa la fda de la forma de defensa inicial (\tilde{x}) que se utilizó en la restricción.

Primero, es trivial que la roja está por encima que la azul en todo momento, además de lo esperable, ya que eso hace la restricción. Ahora bien, fuera de eso no aporta gran cosa, ya que como podemos observar hay una diferencia muy grande en los sitios con ganancia. Además, cuando las pérdidas son muy grandes (pasado el $VaR_{0,2}$) le empieza a ganar.

Asimismo, en el gráfico 4.22 se observa que son casi iguales, siendo distintas cuando la ganancia es poca o la pérdida es mediana. También observamos que casi al final de las pérdidas más grandes le gana, esto gracias a la restricción de Oportunidad. De todas las metodologías, la metodología MUE RO es la más similar en utilidad esperada, mejorando, al mismo tiempo $VaR_{0,2}$, $CVaR_{0,2}$, desviación estándar y la defensa de los 10 peores sitios.

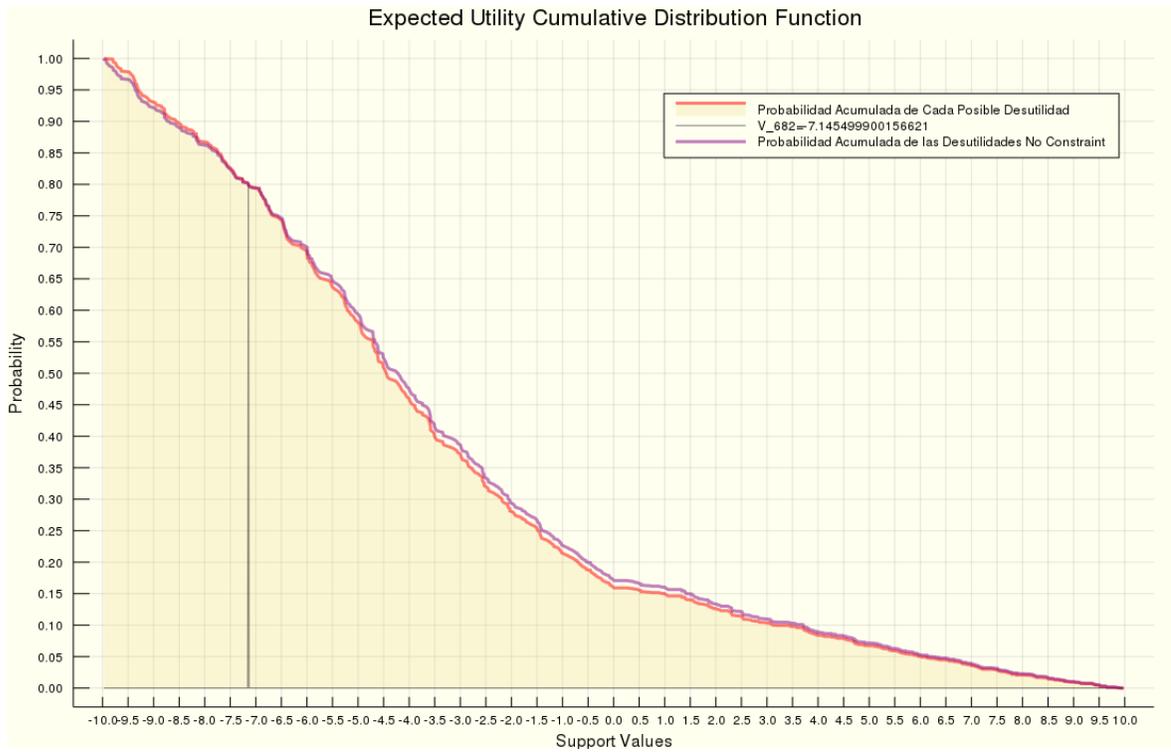


Figura 4.22: Función de Distribución Acumulada del Experimento 7 MUE RO

Capítulo 5

Conclusiones

En este capítulo mostraremos argumentos concisos y breves para determinar qué función de aversión al riesgo agrega, sin perder mucha utilidad esperada, la aversión deseada. Y también, qué restricción funciona como se desea.

Para esto presentaremos tablas mostrando los valores $VaR_{0,2}$, $CVaR_{0,2}$, Ut.Esp y Desv.Std de las diferentes metodologías, y cuánto difieren con respecto a los valores obtenidos de la metodología MUE. Además, se tendrá en cuenta lo discutido en los gráficos que muestran la función de distribución acumulada, los cuáles dan indicios de similitudes, de comportamiento en diferentes zonas de ganancia o pérdida. Por último, al mencionar las diferentes conclusiones se hablará de plazos, siendo el primero el corto plazo, que hace referencia a que se le dá más importancia a la aversión al riesgo y un poco a la maximización de utilidad. El segundo es el mediano plazo, el cual le da pesos similares a la aversión al riesgo y a la maximización de la utilidad. Finalmente, el tercero es el largo plazo, en donde se le dá mucha importancia a la maximización de la utilidad, mas no a la aversión al riesgo.

5.1. Conclusiones Funciones de Aversión al Riesgo

Primero que todo iniciemos mostrando la tabla con la información resumida, mencionando que los tiempos de resolución están en segundos:

Valor Real	$VaR_{0,2}$	$VaR_{0,1}$	$VaR_{0,05}$	$VaR_{0,01}$	FRD	FDP 100 %	FDP 50 %	FDP 20 %	FDO 3	FDO 2	FDO 1	MUE
Desviación Estándar	4.08	3.66	3.31	2.88	4.10	4.91	4.55	2.97	4.94	3.10	3.71	4.48
Utilidad Esperada	-3.64	-4.00	-4.42	-5.09	-3.53	-3.77	-3.47	-4.94	-3.91	-4.78	-3.90	-3.42
$VaR_{0,2}$	6.45	6.73	6.90	7.52	6.70	8.45	7.55	7.44	8.45	7.38	6.70	7.16
$CVaR_{0,2}$	8.24	7.75	7.72	7.95	8.11	9.31	8.88	7.87	9.30	7.82	7.68	8.62
Tiempo Resolución	38.04	39.14	39.71	39.40	5.12	4.23	3.59	3.86	6.29	7.04	6.71	5.63

Tabla 5.1: Tabla de Valores de Diferentes Parámetros de las Metodologías del Exp 7

Porcentaje c/r a MUE	VaR _{0,2}	VaR _{0,1}	VaR _{0,05}	VaR _{0,01}	FRD	FDP 100 %	FDP 50 %	FDP 20 %	FDO 3	FDO 2	FDO 1
Desviación Estándar	-9.06 %	-18.37 %	-26.05 %	-35.83 %	-8.43 %	9.45 %	1.56 %	-33.82 %	10.28 %	-30.81 %	-17.25 %
Utilidad Esperada %	-6.39 %	-17.04 %	-29.29 %	-48.92 %	-3.38 %	-10.27 %	-1.60 %	-44.51 %	-14.40 %	-39.81 %	-13.93 %
VaR _{0,2} %	-9.96 %	-6.03 %	-3.64 %	5.09 %	-6.36 %	18.00 %	5.46 %	3.96 %	18.00 %	3.08 %	-6.44 %
CVaR _{0,2} %	-4.42 %	-10.07 %	-10.52 %	-7.83 %	-5.94 %	7.95 %	2.95 %	-8.74 %	7.86 %	-9.29 %	-10.95 %
Tiempo Resolución %	576.18 %	595.73 %	605.98 %	600.31 %	-8.97 %	-24.89 %	-36.26 %	-31.41 %	11.74 %	25.09 %	19.24 %

Tabla 5.2: Tabla de las Diferencias en Porcentaje c/r a la Metodología MUE del Exp 7

Con toda la información anterior se procede a mostrar conclusiones con respecto a estas funciones.

- Las metodologías FDP 100 %, FDP 50 % y FDO 3 no sirven para agregar aversión, puesto que empeoran en todo aspecto, lo cual se puede observar en la tabla 5.4.
- En cualquier tipo de plazo, las metodologías VaR_{0,2}, VaR_{0,1}, FRD y FDO 1 agregan de manera exitosa la aversión al riesgo que se deseaba y se recomiendan usar si es que se quiere proteger algún sitio o sitios claves, perdiendo un máximo de $\approx 17\%$ en utilidad esperada, pero mejorando un máximo de $\approx 18\%$ en desviación estándar. Además, el VaR_{0,2} y CVaR_{0,2} disminuyen un mínimo de $\approx 6\%$ y $\approx 4,5\%$, respectivamente, y disminuyen un máximo de $\approx 10\%$ y $\approx 11\%$, respectivamente.
- En el largo plazo existen ciertas metodologías que terminan siendo costosas (pero no tanto) si se usan, pues son demasiado aversas al riesgo. Estas metodologías son VaR_{0,05} y FDO 2. La ventaja de estas metodologías es que en el corto y mediano plazo sirven para defender, en caso de emergencia, ciertos sitios o un sitio en particular, ya que protegen mucho mejor los peores 10 sitios. Además, la desviación estándar disminuye un poco más de un $\frac{1}{4}$ y el CVaR_{0,2} disminuye casi un 10 %. Esto último asegura de buena forma tener claro cuánto se podría perder en promedio y que, en un 80 % si todo sale mal, se sabrá cuánto más en promedio del valor VaR_{0,2} se perderá.
- Las metodologías VaR_{0,01} y FDP 20 % se descartan como buenas metodologías, puesto que el precio a pagar es demasiado alto, casi un $\frac{1}{3}$ más de pérdida en utilidad esperada. Quizás en casos muy particulares y, extremo absoluto, se podrían llegar a utilizar para proteger ciertos sitios.
- En el caso de las metodologías VaR, mientras más aversión se le agregue, más tiempo demora en resolverse (más aversión implica mayor tiempo). Por otro lado, éstas son caras en tiempo con respecto a MUE, ya que es resolución de optimización convexa vía búsquedas binarias y aproximación del óptimo.
- Conectado con el punto anterior, la metodología FRD, además de ser semejante en utilidad esperada (+8.43 %), toma menos tiempo en resolver que la metodología MUE, lo cual es deseable. Por ejemplo; si se quiere utilizar en alguna aplicación en teléfonos. Esto también es aplicable para la metodología FDP 20 %, la cual se puede usar en casos extremos absolutos.

5.2. Conclusiones Restricciones de Aversión al Riesgo

Al igual que en la sección anterior, primero se entregará la información resumida en tablas con respecto a la metodología MUE, pero agregando aversión al riesgo mediante restricciones. Asimismo, el tiempo está medido en segundos.

Valor Real	MUE	MUE RRD	MUE RPODE	MUE RO
Desviación Estándar	4.48	4.29	3.04	4.38
Utilidad Esperada	-3.42	-3.45	-5.13	-3.53
$VaR_{0,2}$	7.16	6.82	7.62	7.15
$CVaR_{0,2}$	8.62	8.35	8.18	8.53
Tiempo de Resolución	5.63	6.14	18356.72	5.12

Tabla 5.3: Tabla de Valores de Diferentes Parámetros de las Metodologías del Exp 7

Porcentaje c/r a MUE	MUE RRD	MUE RPODE	MUE RO
Desviación Estándar	-4.30 %	-32.23 %	-2.22 %
Utilidad Esperada	-0.93 %	-50.11 %	-3.19 %
$VaR_{0,2}$	-4.77 %	6.44 %	-0.19 %
$CVaR_{0,2}$	-3.11 %	-5.13 %	-1.12 %
Tiempo de Resolución	9.12 %	326217.93 %	-8.95 %

Tabla 5.4: Tabla de las Diferencias en Porcentaje c/r a la Metodología MUE del Exp 7

Al observar esta información, y tomando en cuenta lo discutido en el capítulo anterior se concluyó lo siguiente:

- Las metodologías MUE RRD y MUE RO son útiles en lo deseado, puesto que el precio a pagar es menos del 3.2 % , mejorando todos los parámetros de aversión al riesgo.
- La metodología MUE RRD mejora más los parámetros, agregando suficiente aversión pagando poca Ut.Esp. La solución de la metodología MUE RO mejora en poco porcentaje los parámetros, pero incrementa notablemente la defensa de los 10 peores sitios. Se recomiendan ambas para mejorar la aversión.
- Como era de esperarse, la metodología MUE RPODE es mala, ya que paga 50 % más de utilidad esperada y, además, paga exuberantemente en tiempo, demorándose $\approx 326217\%$ más en resolverse que la metodología MUE. Recordar que se agregaron 800 restricciones convexas. Esta metodología no es recomendable, a menos que se utilice, por ejemplo, una solución inicial similar a MUE RRD o MUE RO, las cuales cambiaron un poco en utilidad esperada y agregaron aversión (manteniéndose la lentitud en resolución, lo cual es poco práctico).

5.3. Conclusiones Generales

Con todo lo anterior se procede a mostrar lo concluido de forma global.

- Agregar aversión al riesgo no significa un gran pago en tiempo, exceptuando las metodologías VaR y la MUE RPODE.
- De las metodologías recomendadas, se pierde a lo más un 17% de utilidad esperada, mejorando en conjunto los parámetros de aversión. En el largo plazo se podría suplir de muchas formas debido a lo bajo del porcentaje.
- Sale más caro en utilidad y tiempo agregar aversión al riesgo al resolver funciones que agregar una restricción de aversión.
- Se puede mejorar los tiempos de resolución de las metodologías al usar otro solver o encontrar un método más directo.
- No se puede agregar demasiada aversión al riesgo sin perder mucha utilidad o tiempo.
- El QR representa muy bien la racionalidad del atacante y muy bien el hecho de que el líder sabe que tan racional es el seguidor.
- Un posible estudio sería lograr ver la respuesta del modelo QR al cambiar la racionalidad del atacante bajo diferentes contextos. Ver si sigue un mismo patrón, independiente de la cantidad de sitios a defender, o bien ver cómo se comporta al existir muchos sitios con alta paga para ambos contendientes.
- Se puede expandir el estudio con un modelo de respuesta distinto, que agregue más información y más realista, lo cual, al mismo tiempo, haría más complejo resolverlo numéricamente.
- Como posible trabajo a futuro se puede estudiar la misma aversión, pero incrementando la cantidad de sitios a defender. O bien, agregar distorsión a la defensa y al ataque. Por otro lado, se podría cambiar la probabilidad de ataque como una variable aleatoria discreta, más un ruido que involucre la decisión del líder. Por ejemplo, $y_i(x) = \frac{1}{n} + \ln(x_i)$. Por último, se podría trasladar la aleatoriedad al soporte de la pérdida (cambiando el soporte, dependiendo de la decisión del líder), manteniendo constante (misma variable aleatoria) la forma de ataque. Logrando así tener un modelo discreto semejante al mencionado en la sección 1.4.
- Calcular el VaR en este tipo de modelo es más sencillo, debido a que la aleatoriedad se observa en la probabilidad de que cada sitio sea atacado y no en la función de probabilidad del soporte de la desutilidad referida a la pérdida, que se mantiene constante. Lo anterior permite la transformación del problema de minimización del VaR a uno convexo, mas no al problema de minimización del CVaR, el cual no es posible de llevar a un problema convexo y altamente caro en tiempo, ya que se resolverán siempre muchos problemas lineales gigantes, demorándose incluso más que resolver el VaR.
- La racionalidad impacta fuertemente en el modelo, ya que para valores $\lambda \geq 3$, numéricamente, el solver empieza a presentar problemas por números muy grandes. Si λ es grande, entonces la solución, en los modelos de aversión, tienden a distribuir mucho los recursos (de forma inteligente) a lo largo de todos los sitios, puesto que el seguidor hará lo mismo con la distribución sobre sus acciones. Mientras que, si λ es pequeño, entonces la solución será defender con frecuencia 1 los m peores sitios.

Bibliografía

- [Dan63] Dantzig, G.B. *Linear Programming and Extensions*. Princeton: Princeton University Press, 1963.
- [Fra80] Frank Abagnale Jr., Stan Redding. *Catch Me if You Can*. 345 Hudson Street 10th Floor New York NY 10014 USA: Grosset & Dunlap, 1980.
- [R T00] R. Tyrrell Rockafellar, Stanislav Uryasev. “Optimization of Conditional Value at-Risk”. En: *American Optimal Decisions, Inc 2.3* (2000), págs. 1-21.
- [Jua08] Juan Carlos Bárcena Ruiz. *Aproximación a las Ciencias Sociales desde la Teoría de los Juegos*. Curso del Departamento de Fundamentos del Análisis Económico I, Universidad del País Vasco. 2008.
- [Chr09] Christopher Kiekintveld, Manish Jain, Jason Tsai, James Pita, Fernando Ordoñez, Milind Tambe Less. “Computing Optimal Randomized Resource Allocations for Massive Security Games”. En: *AAMAS* (2009). DOI: 10.1145/1558013.1558108.
- [Lue10] Elena Almaraz Luengo. “Cuestiones Notables de Ordenación Estocástica de Optimización Financiera”. Tesis doct. Universidad Complutense de Madrid, 2010.
- [Sta10] Stan Uryasev, Sergey Sarykalin, Gaia Serraino y Konstantin Kalinchenko. *VaR vs CVaR in Risk Management and Optimization*. Risk Management and Financial Engineering Lab, University of Florida y American Optimal Decisions. 2010.
- [Abu11] Juan F. Jara Aburto. *Modelo de Asignación de Recursos Policiacos en la Vía Pública*. Memoria Ingeniería Indutrsial, Universidad de Chile. 2011.
- [Bo 13] Bo An, Fernando Ordóñez, Milind Tambe, Eric Shieh, Rong Yang, Craig Baldwin, Joseph DiRenzo III, Kathryn Moretti, Ben Maule, Garrett Meye. “A Deployed Quantal Response-Based Patrol Planning System for the U.S. Coast Guard”. En: *Institute for Operations Research and the Management Sciences* 43.5 (2013), págs. 400-422. DOI: <http://dx.doi.org/10.1287/inte.2013.0700>.
- [Ren16] Renaud Chicoisne, Fernando Ordóñez. *Decision and Game Theory for Security*. 1201 Larimer Street, Denver, CO, 80204, USA: Springer International Publishing AG, 2016.
- [Ari18] Jaime San Martín Aristegui. *Teoría de la Medida*. Santiago: Editorial Universitaria, 2018.
- [Car19] Carlos Casorran, Bernard Fortz, Martine Labbé, Fernando Ordóñez. “A study of general and security Stackelberg game formulations”. En: *European Journal of Operational Research* 3.278 (2019). DOI: 10.1016/j.ejor.2019.05.012.

Anexo

Experimento 3: Sin Restricciones

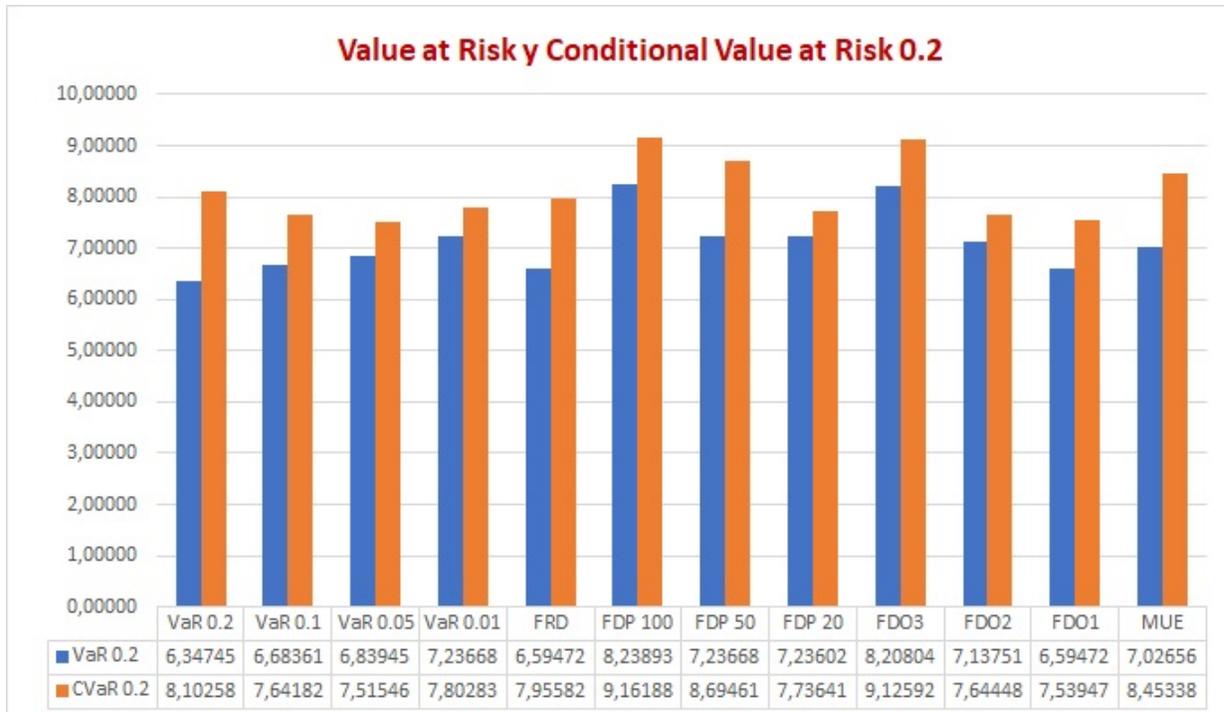


Figura 5.1: VaR y CVar 0.2 del Experimento 3 sin Restricciones

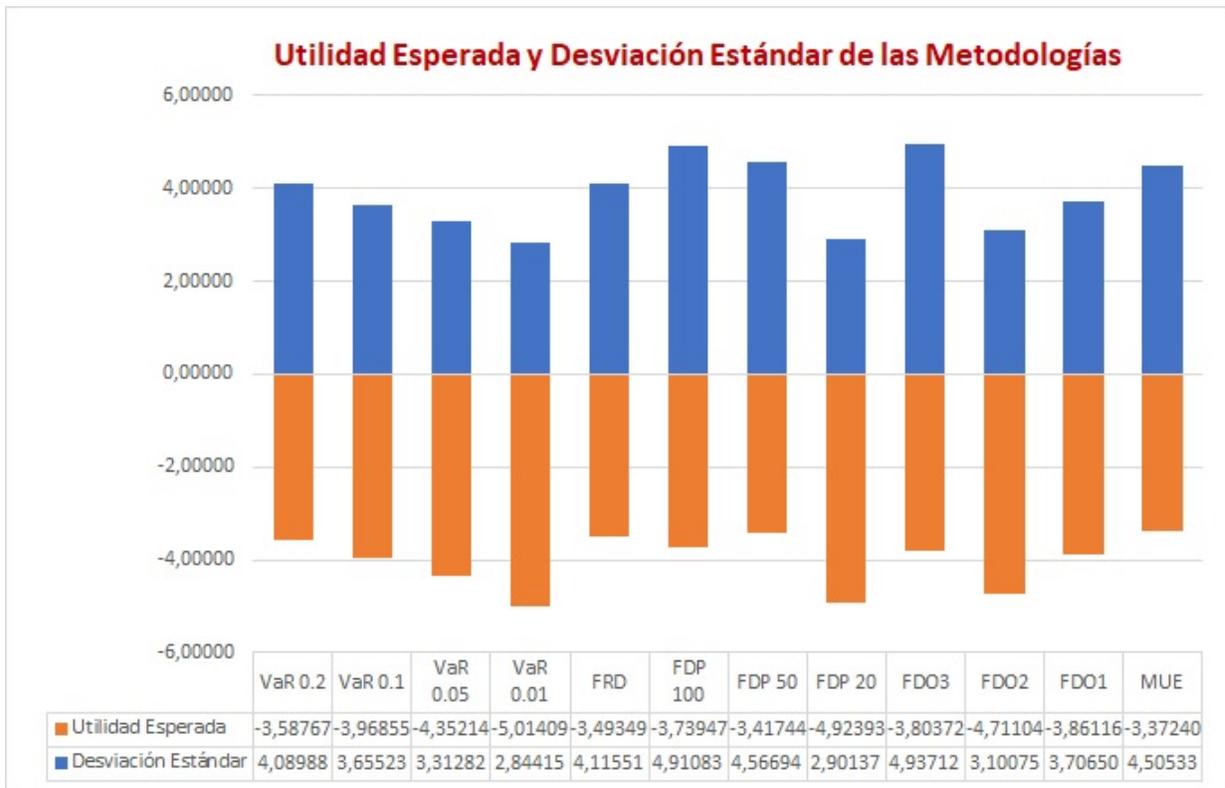


Figura 5.2: Utilidad Esperada y Desviación Estándar del Experimento 3 sin Restricciones

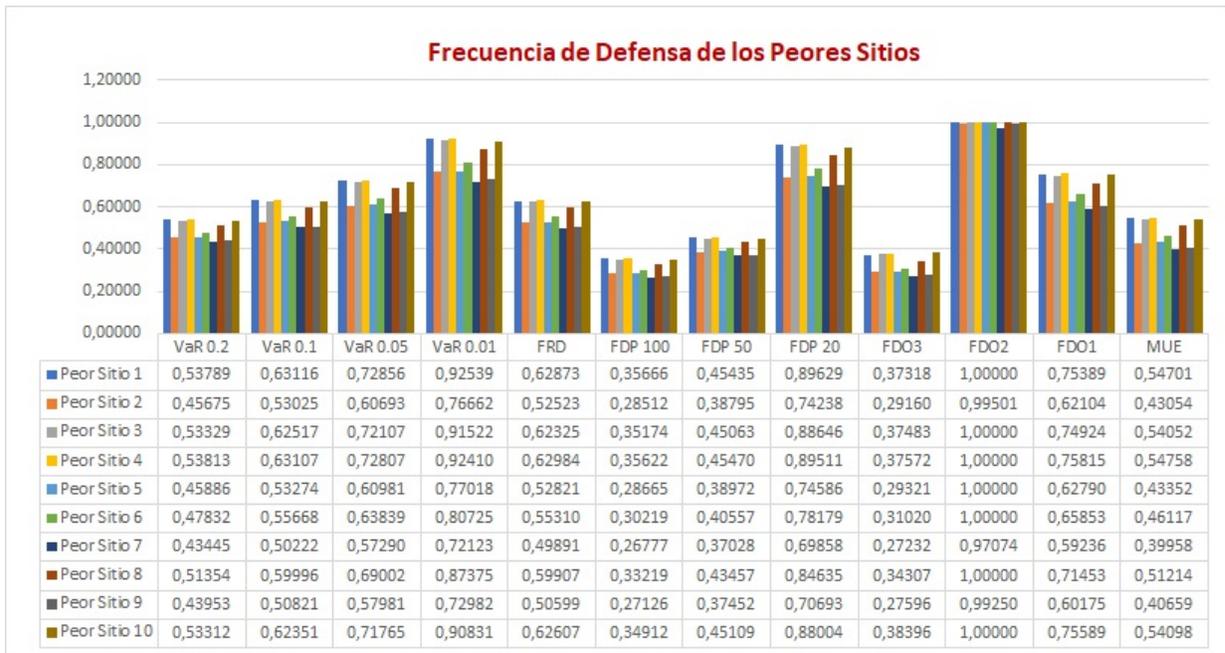


Figura 5.3: Histograma de los Peores 10 Sitios del Experimento 3 sin Restricciones

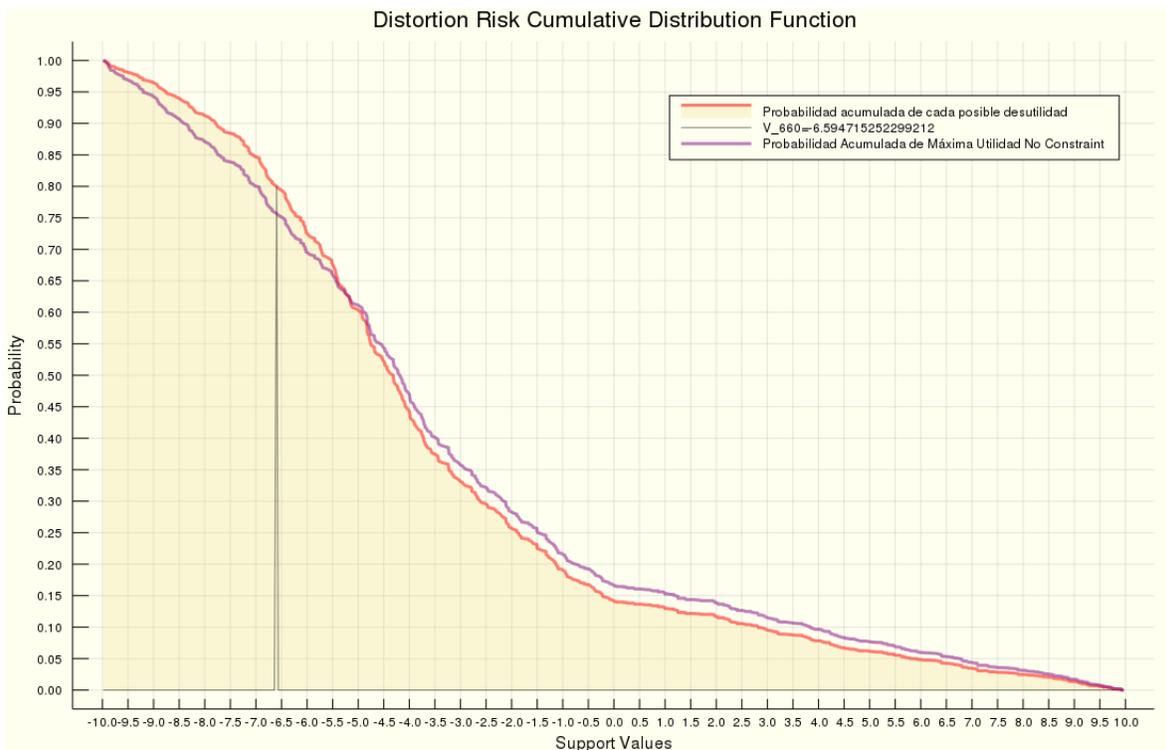


Figura 5.4: Función de Distribución Acumulada del Experimento 3 FRD

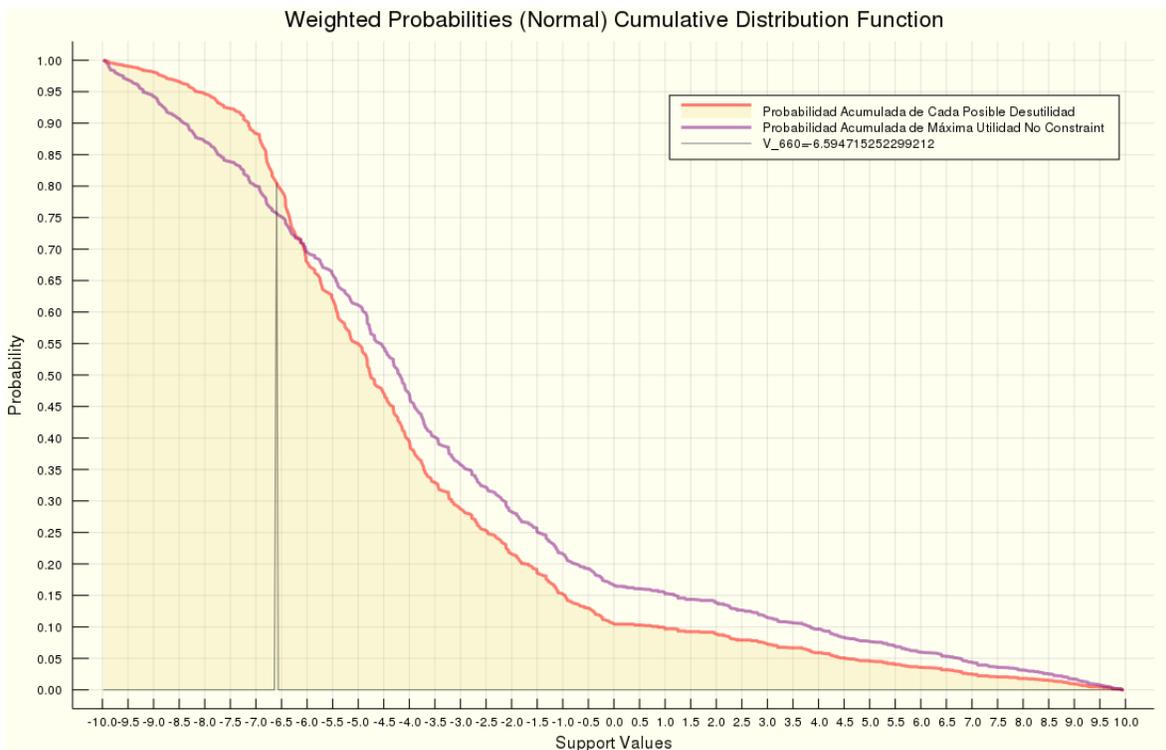


Figura 5.5: Función de Distribución Acumulada del Experimento 3 FDO1

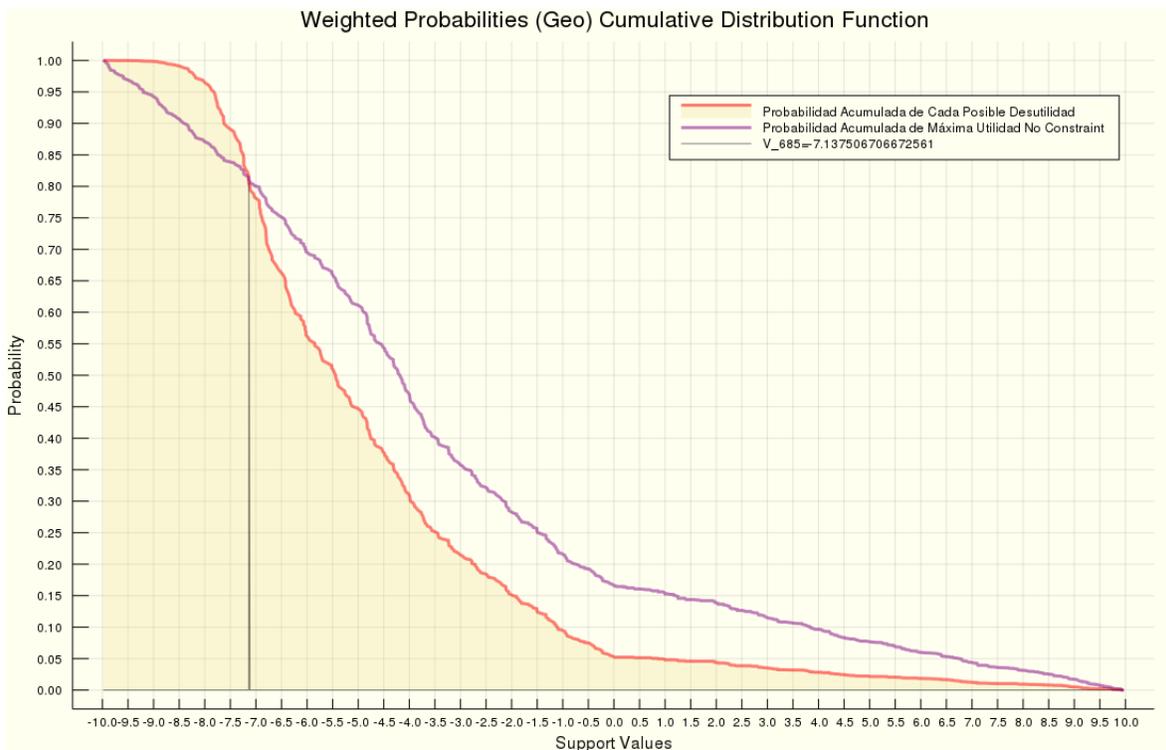


Figura 5.6: Función de Distribución Acumulada del Experimento 3 FDO2

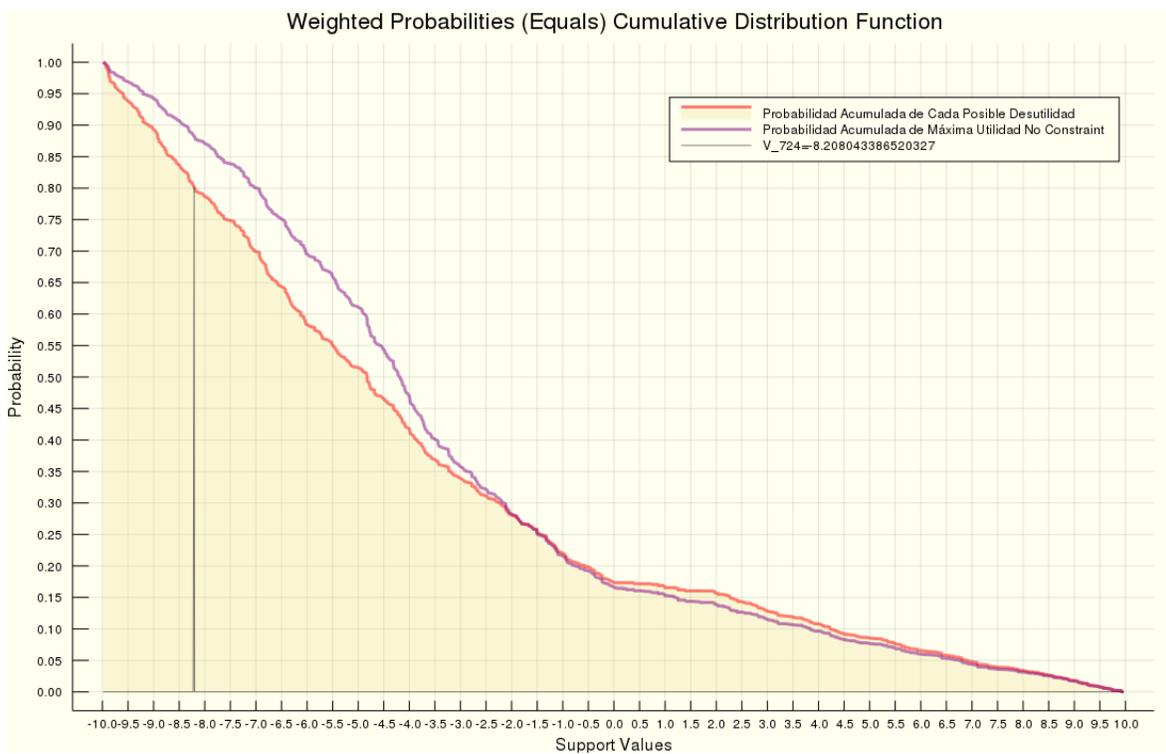


Figura 5.7: Función de Distribución Acumulada del Experimento 3 FDO3

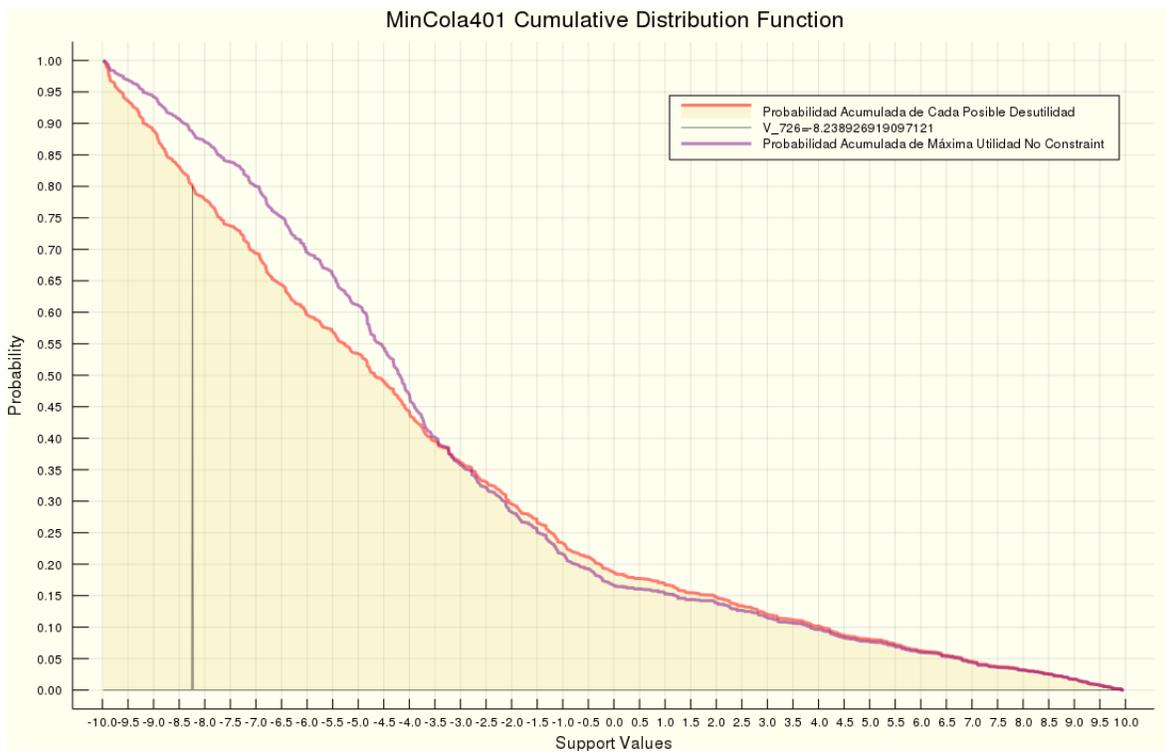


Figura 5.8: Función de Distribución Acumulada del Experimento 3 FDP100

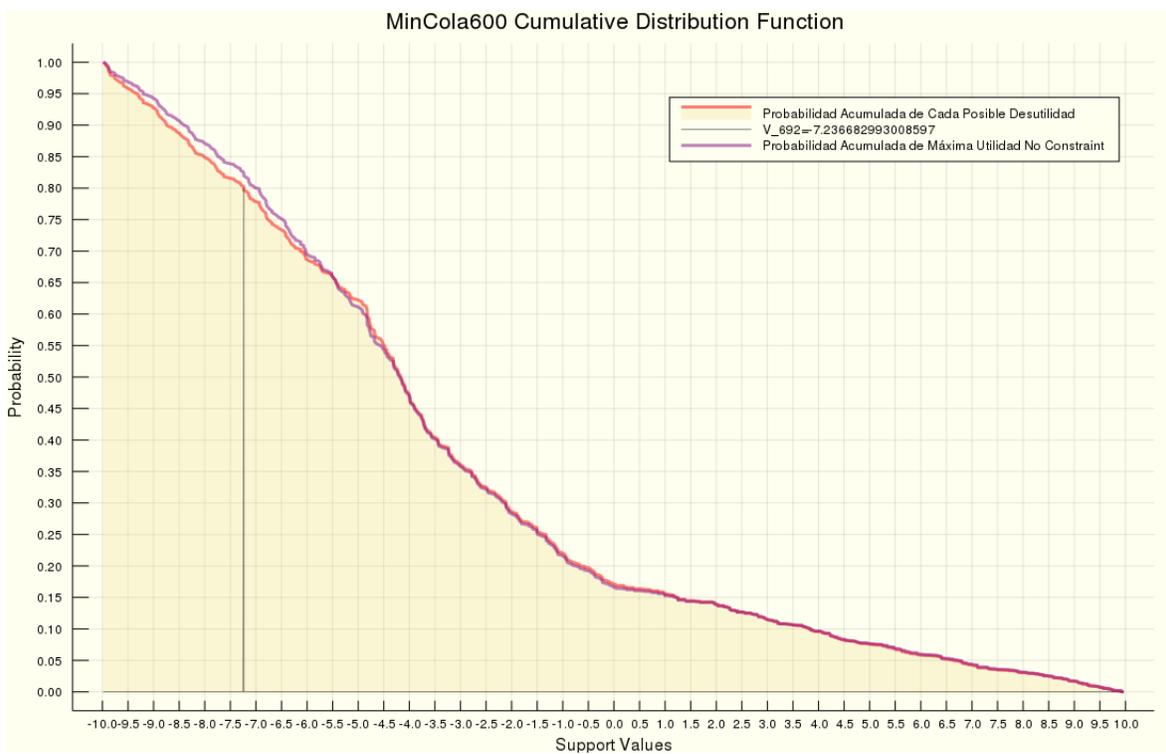


Figura 5.9: Función de Distribución Acumulada del Experimento 3 FDP50

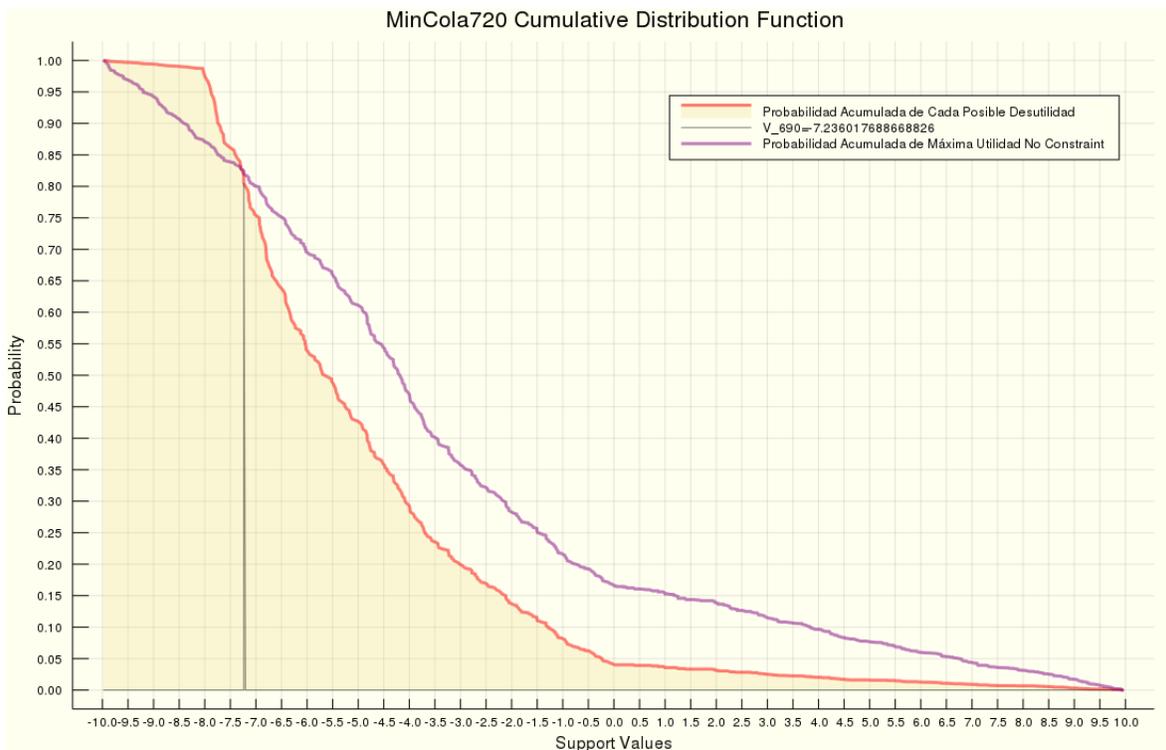


Figura 5.10: Función de Distribución Acumulada del Experimento 3 FDP20

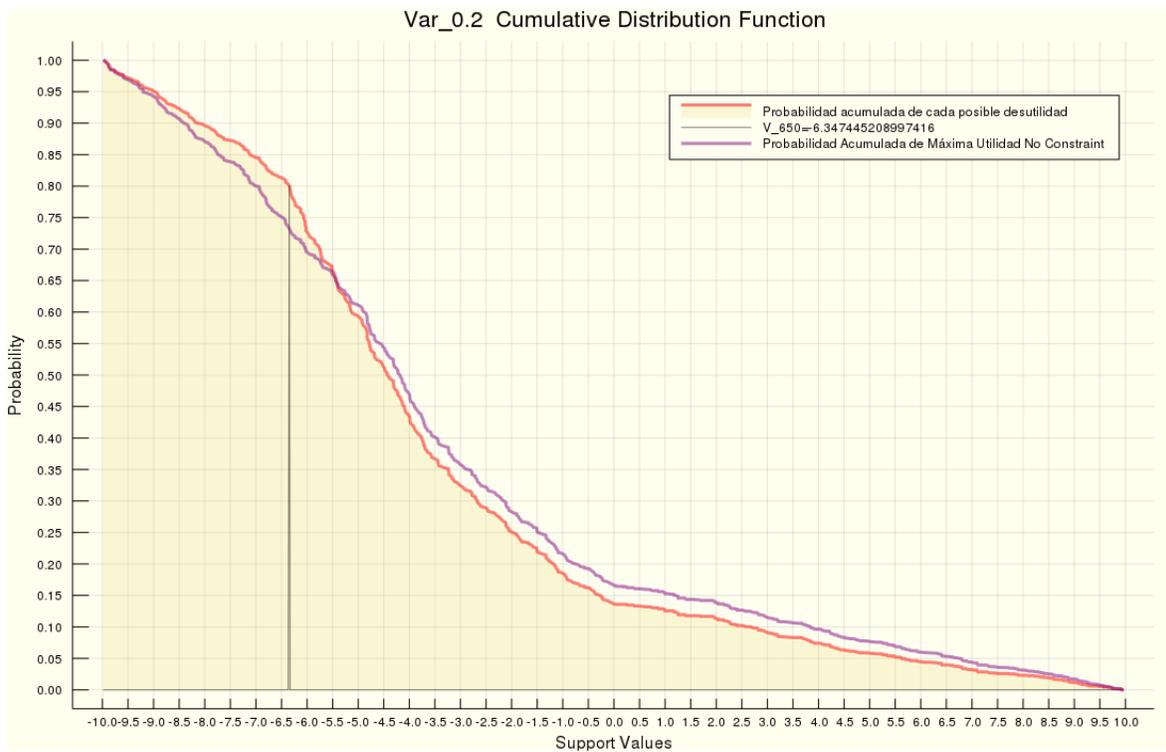


Figura 5.11: Función de Distribución Acumulada del Experimento 3 VaR_{0,2}

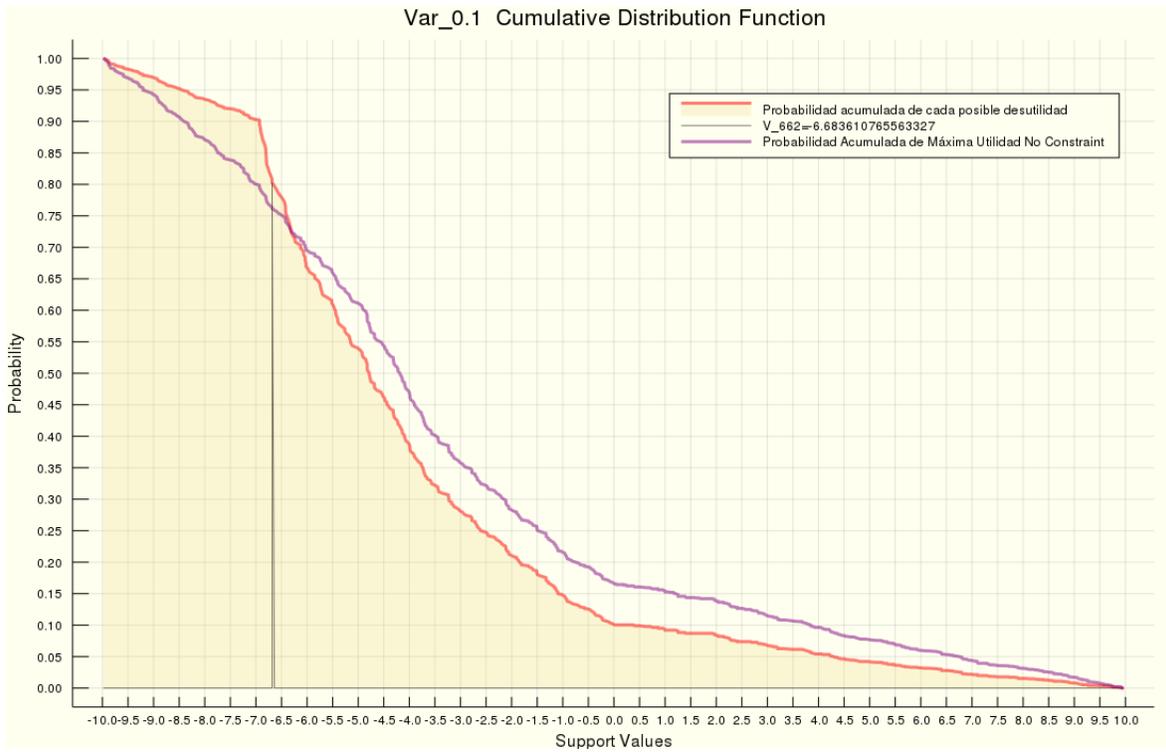


Figura 5.12: Función de Distribución Acumulada del Experimento 3 $VaR_{0,1}$

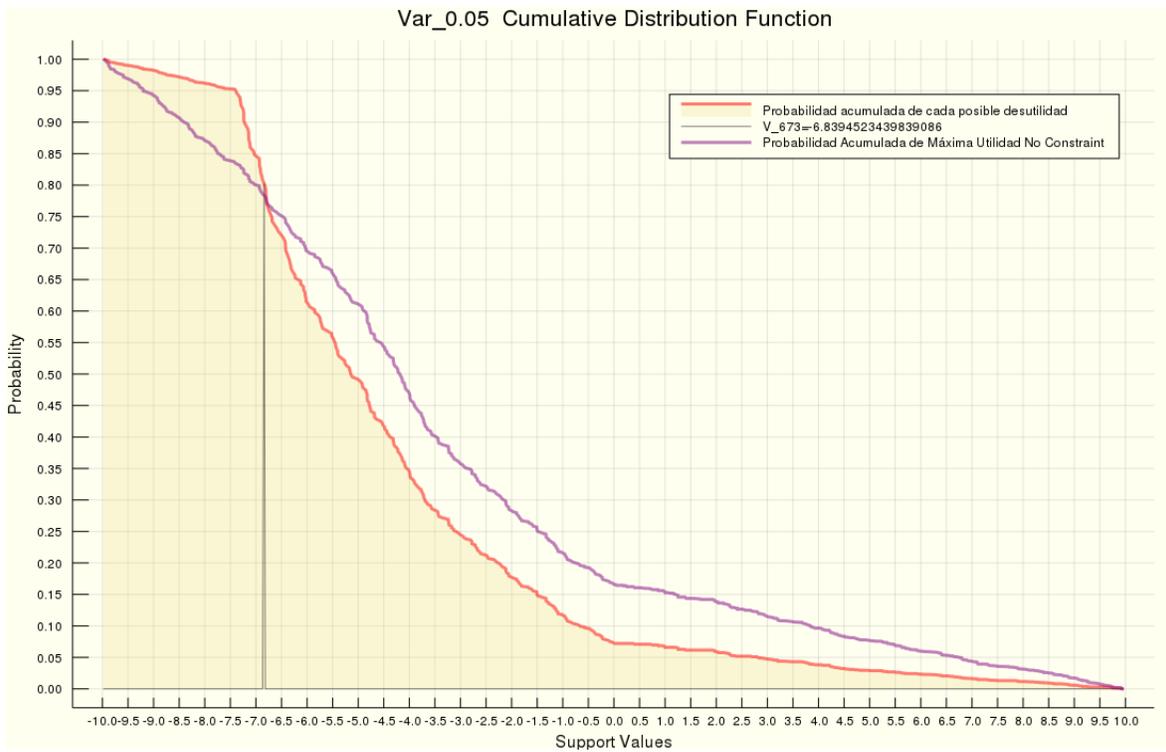


Figura 5.13: Función de Distribución Acumulada del Experimento 3 $VaR_{0,05}$

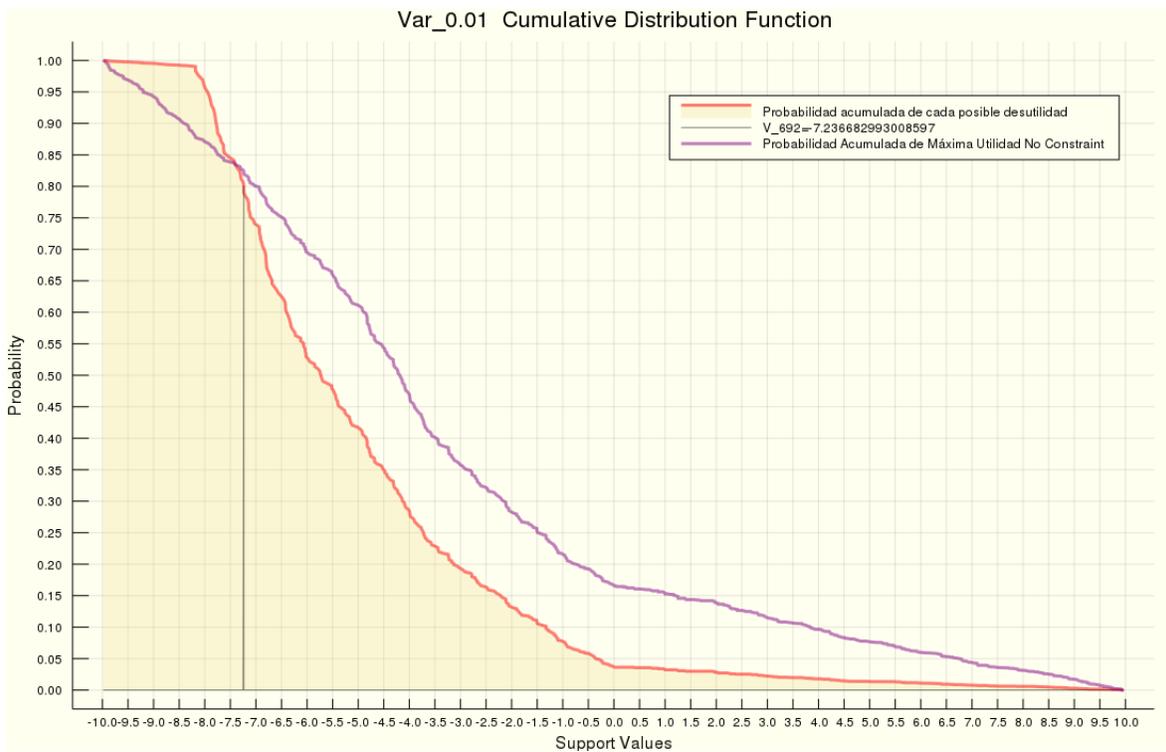


Figura 5.14: Función de Distribución Acumulada del Experimento 3 $Var_{0,01}$

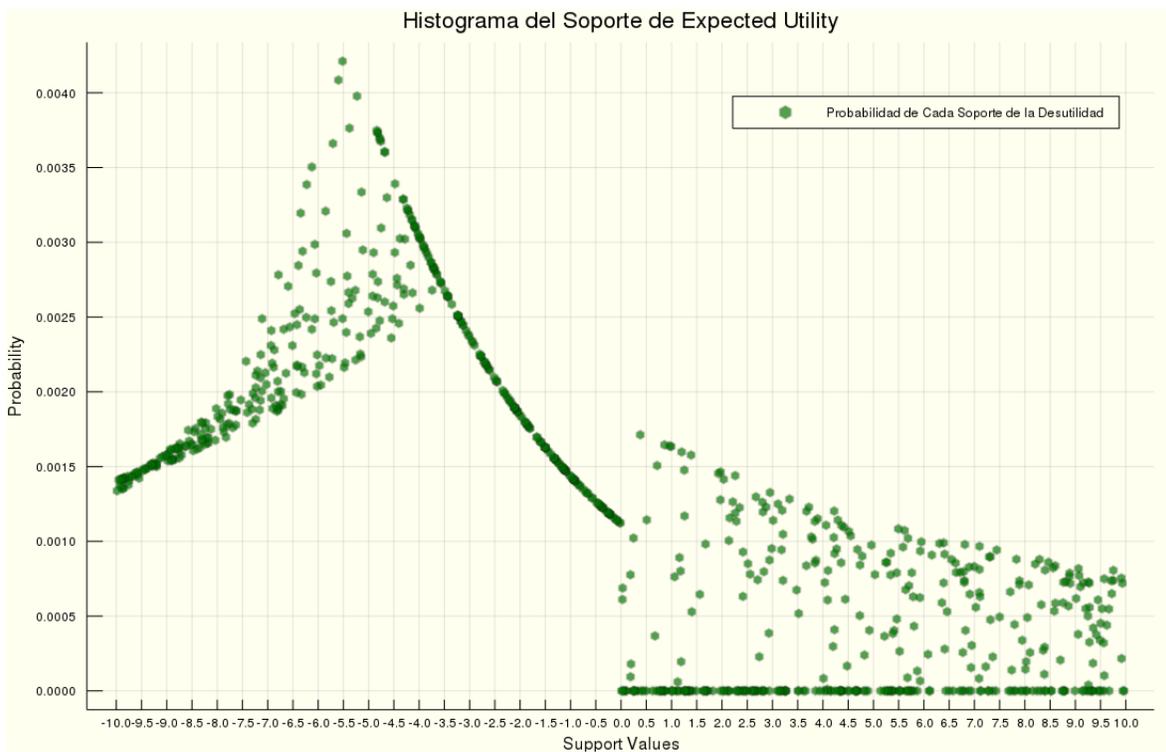


Figura 5.15: Histograma del Soporte de MUE del Experimento 3

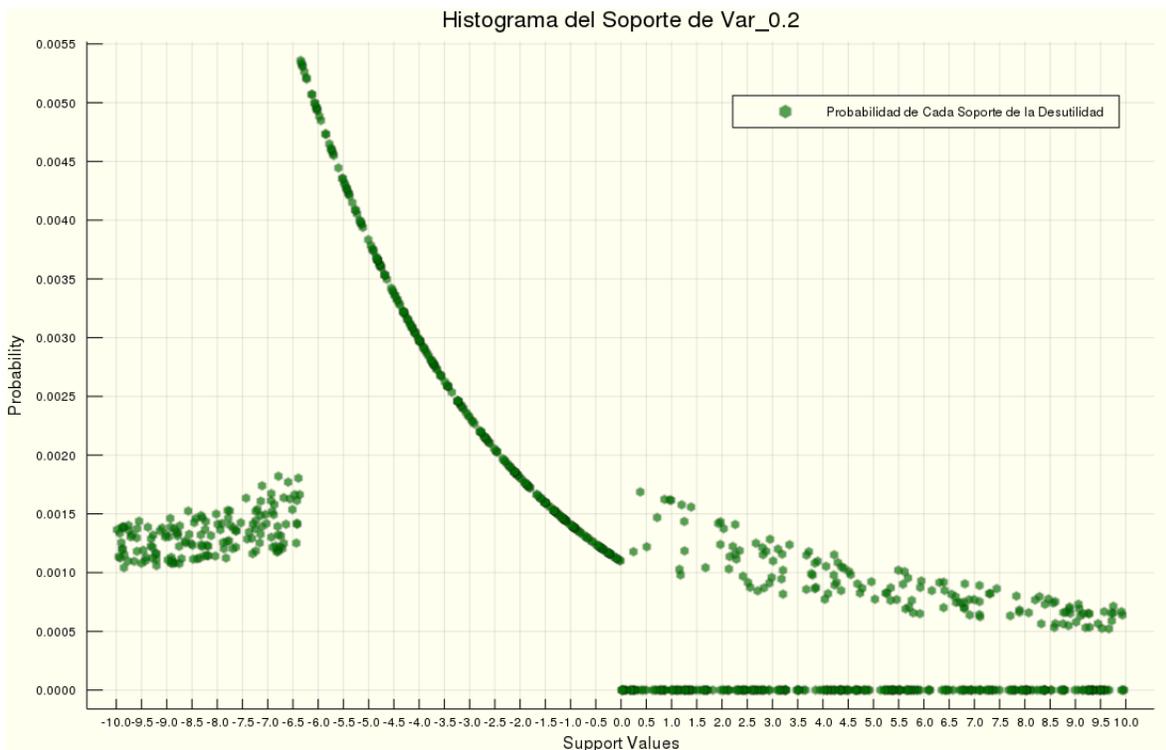


Figura 5.16: Histograma del Soporte de $VaR_{0,2}$ del Experimento 3

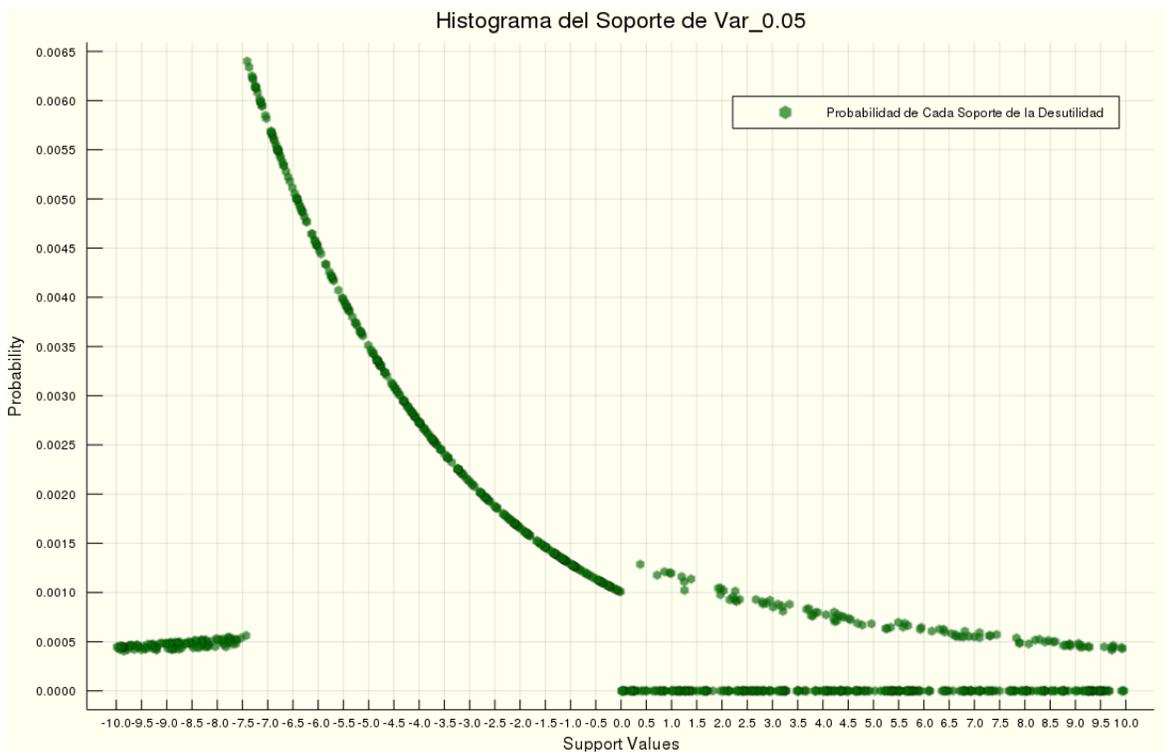


Figura 5.17: Histograma del Soporte de $VaR_{0,05}$ del Experimento 3

Experimento 3: Con Restricciones



Figura 5.18: VaR y CVar 0.2 del Experimento 3 con Restricciones

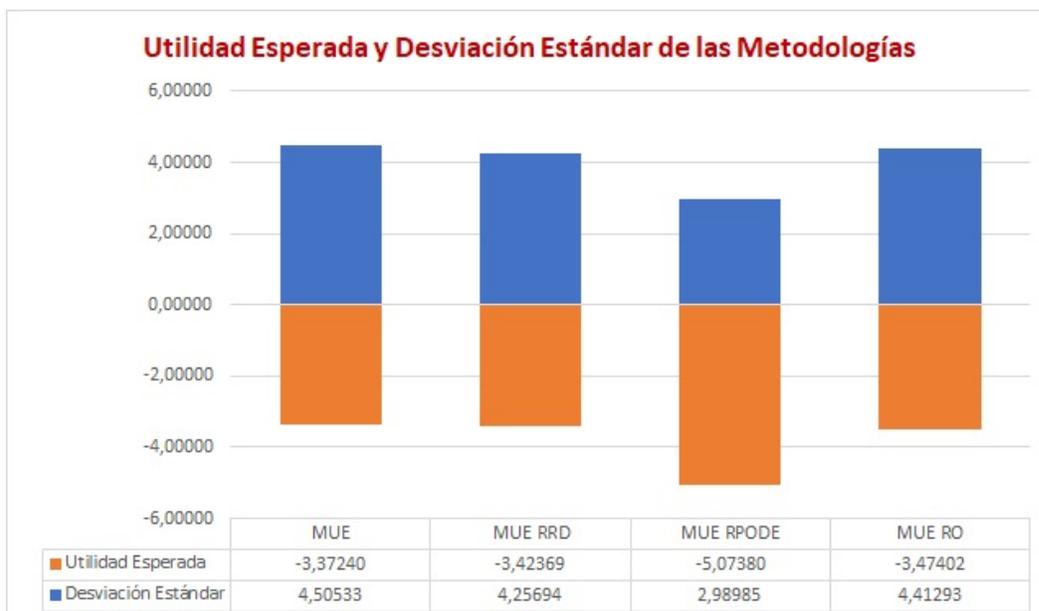


Figura 5.19: Utilidad Esperada y Desviación Estándar del Experimento 3 con Restricciones



Figura 5.20: Histograma de los Peores 10 Sitios del Experimento 3 con Restricciones

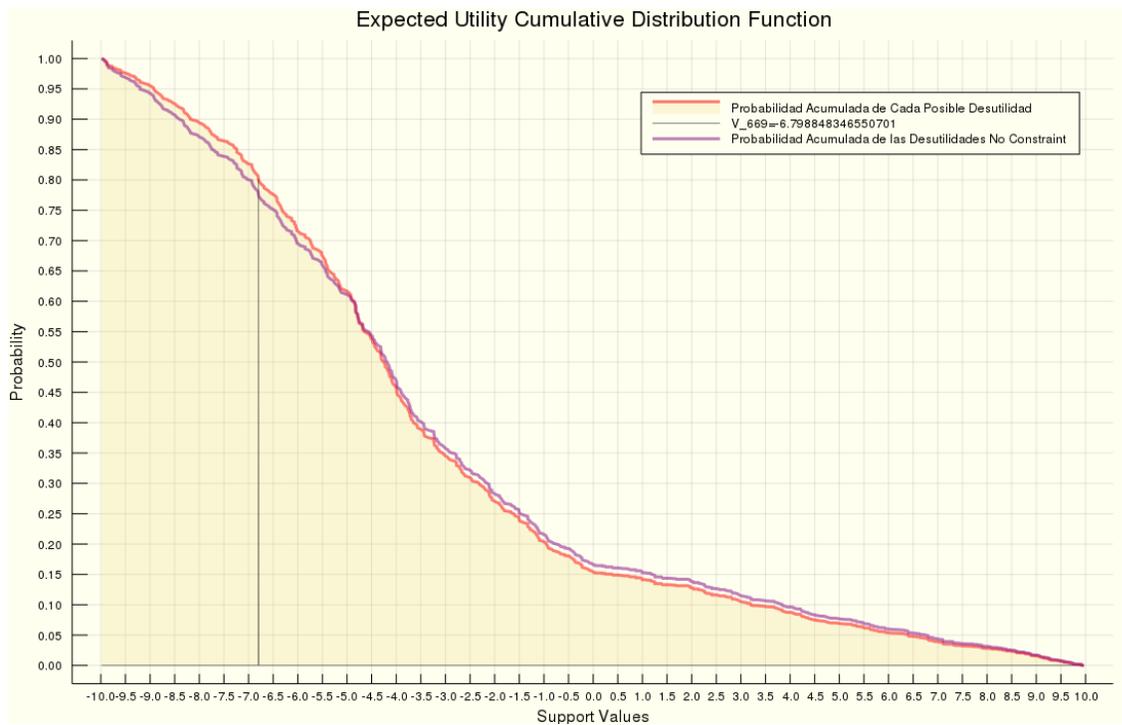


Figura 5.21: Función de Distribución Acumulada del Experimento 3 MUE RRD

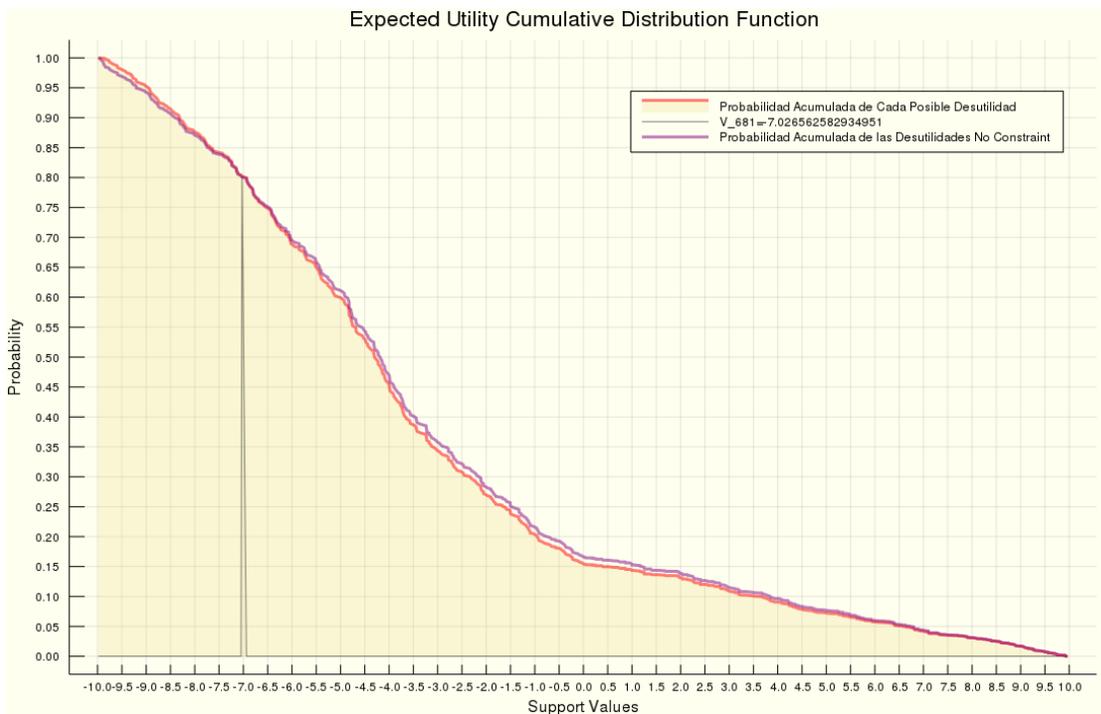


Figura 5.22: Función de Distribución Acumulada del Experimento 3 MUE RO

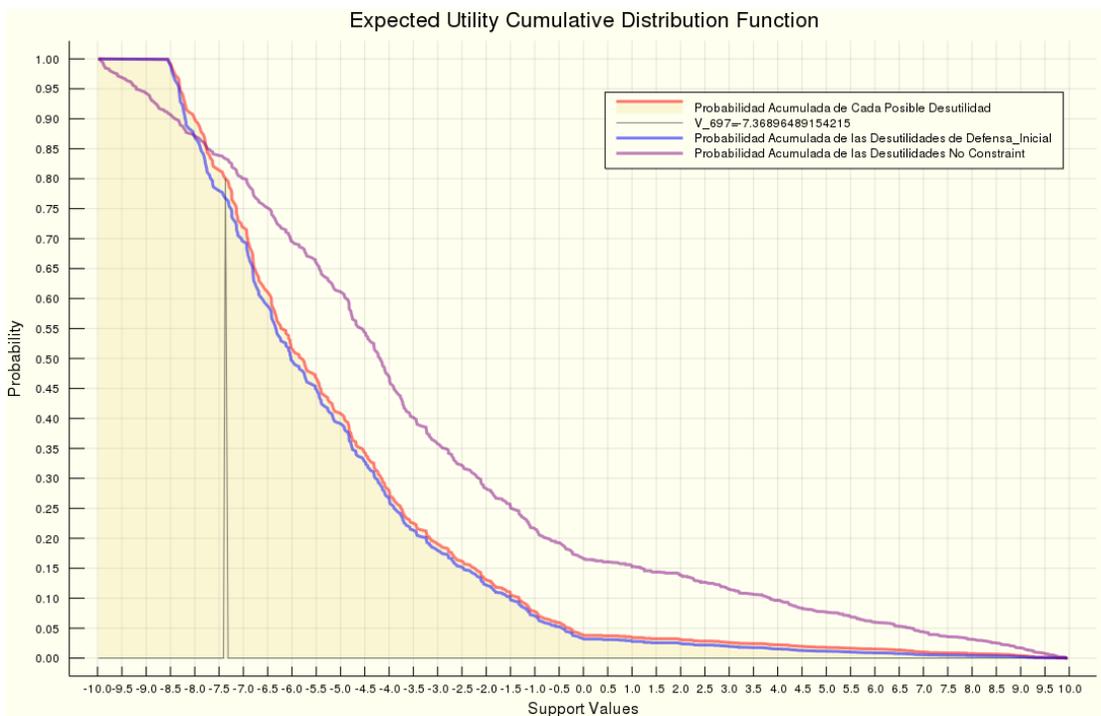


Figura 5.23: Función de Distribución Acumulada del Experimento 3 MUE RPODE

Experimento 10: Sin Restricciones

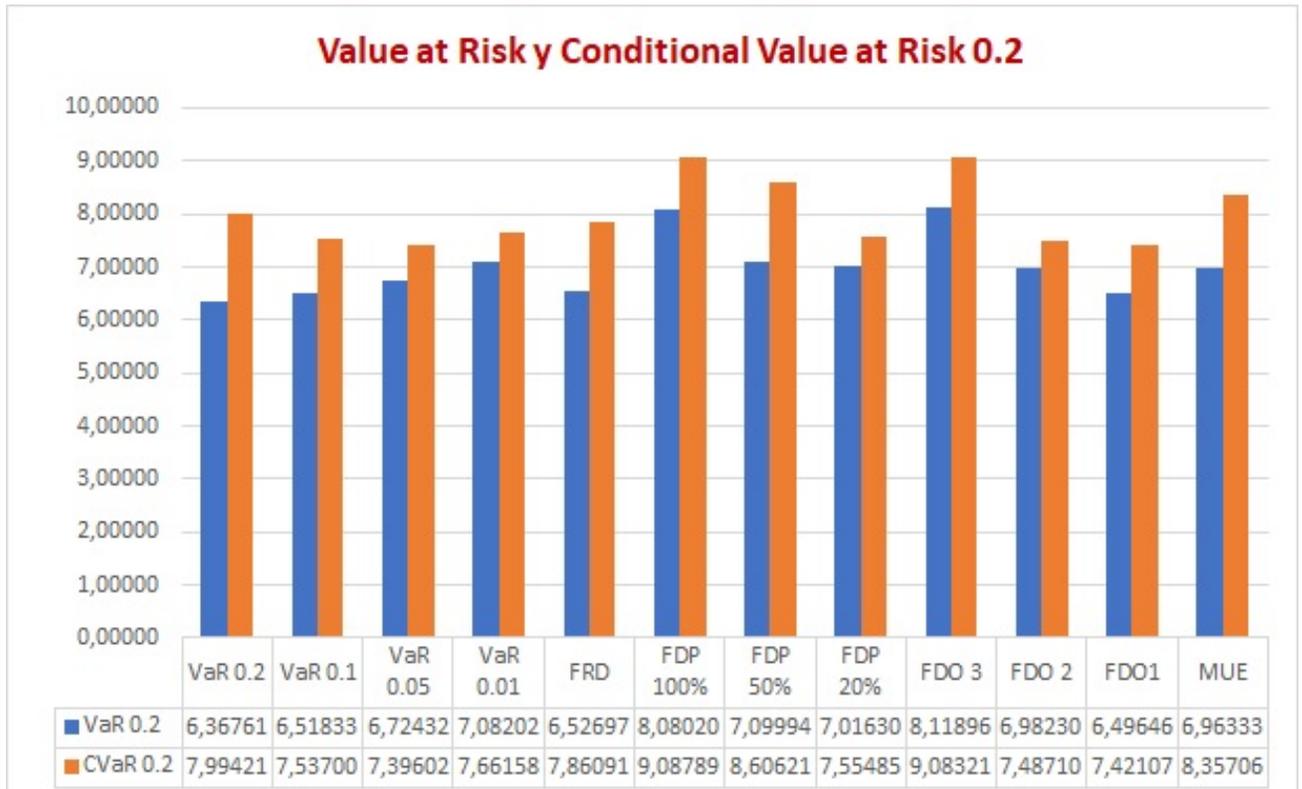


Figura 5.24: VaR y CVar 0.2 del Experimento 10 sin Restricciones

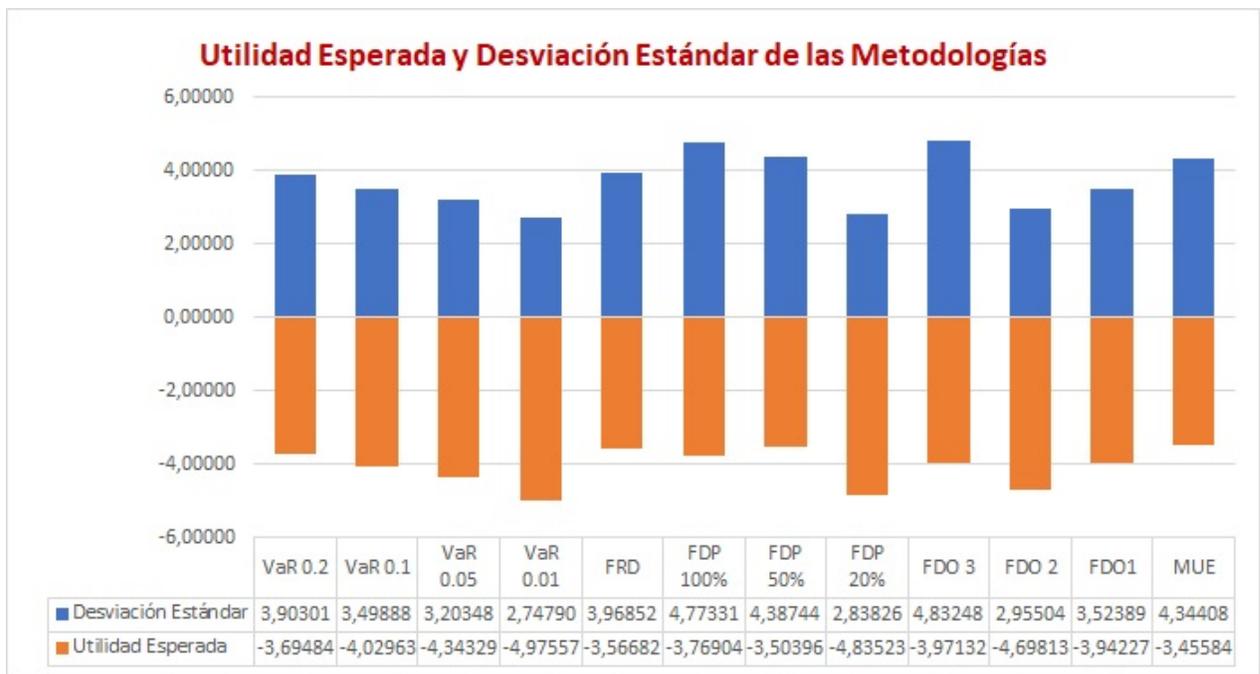


Figura 5.25: Utilidad Esperada y Desviación Estándar del Experimento 10 sin Restricciones

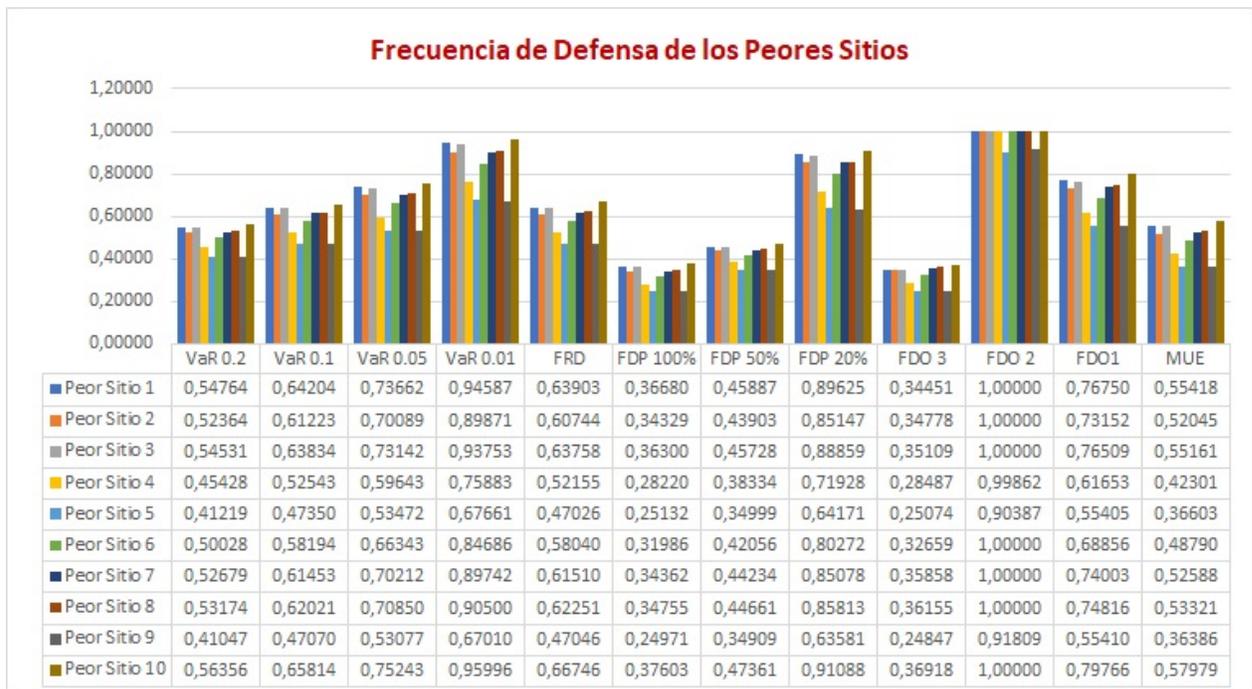


Figura 5.26: Histograma de los Peores 10 Sitios del Experimento 10 sin Restricciones

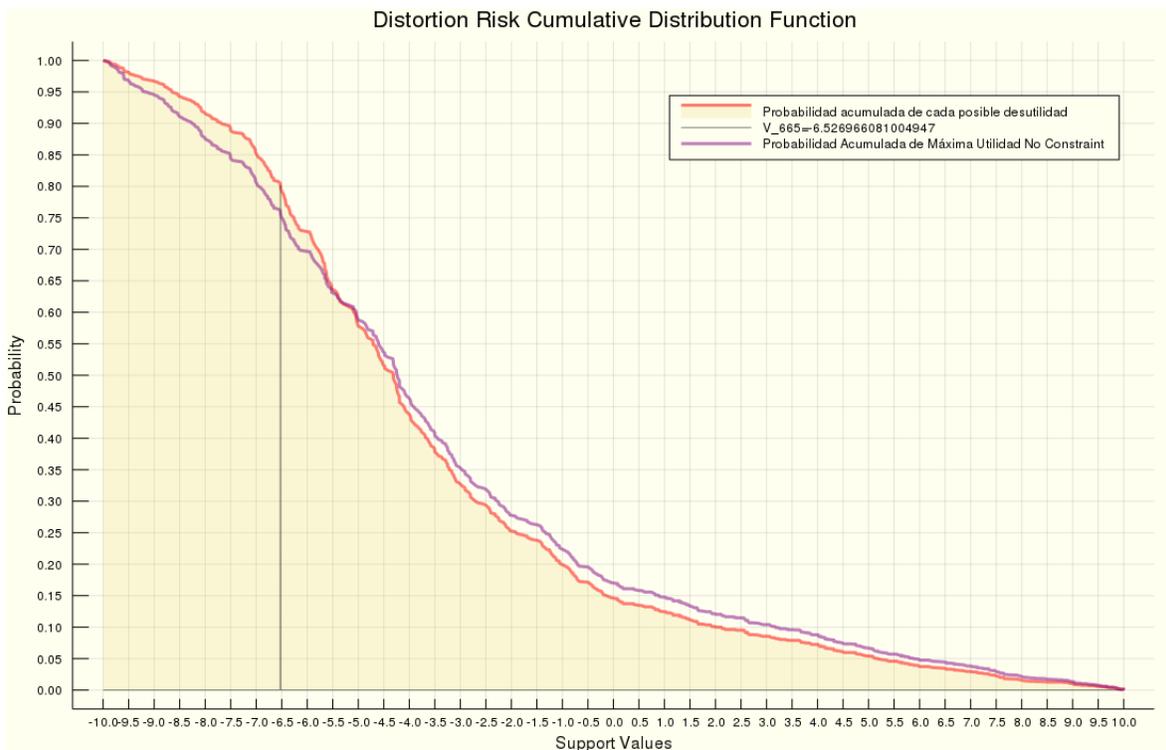


Figura 5.27: Función de Distribución Acumulada del Experimento 10 FRD

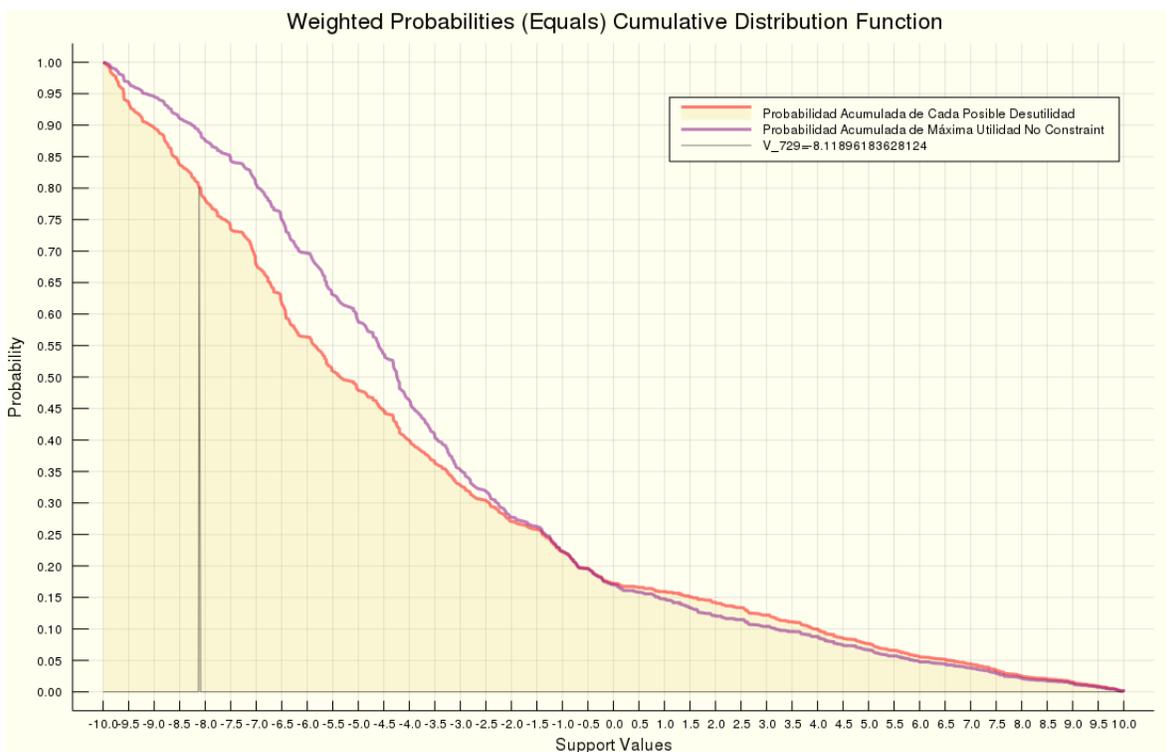


Figura 5.28: Función de Distribución Acumulada del Experimento 10 FDO1

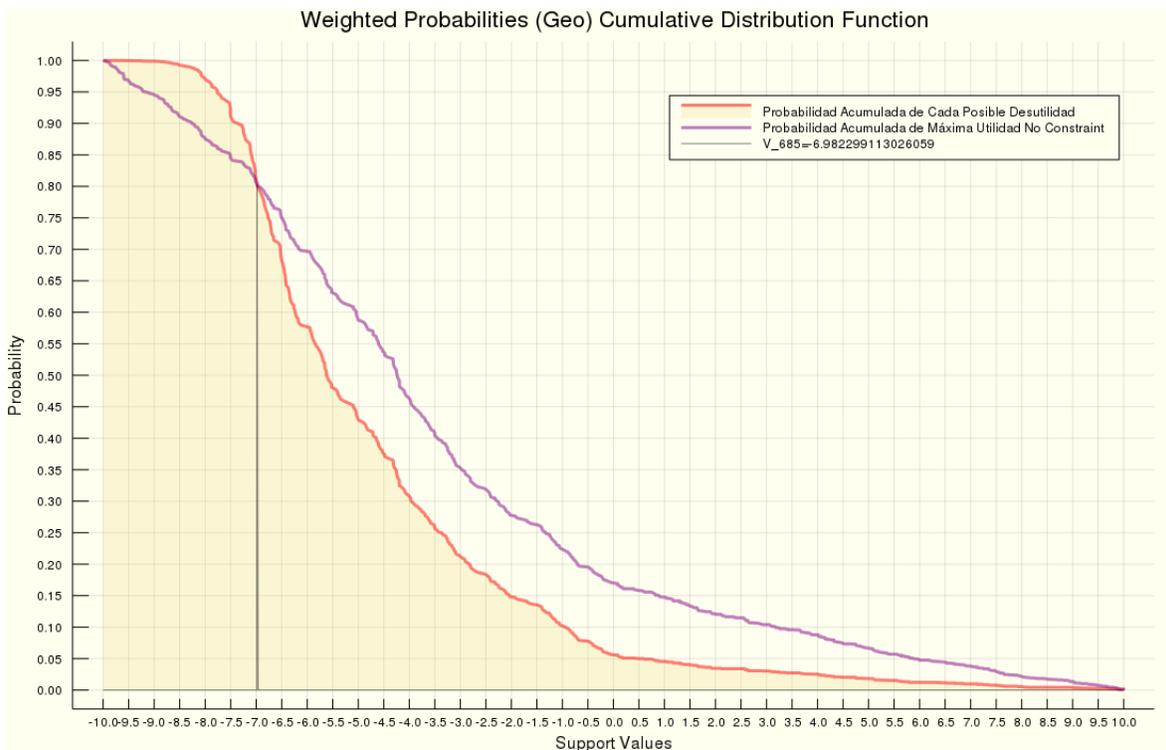


Figura 5.29: Función de Distribución Acumulada del Experimento 10 FDO2

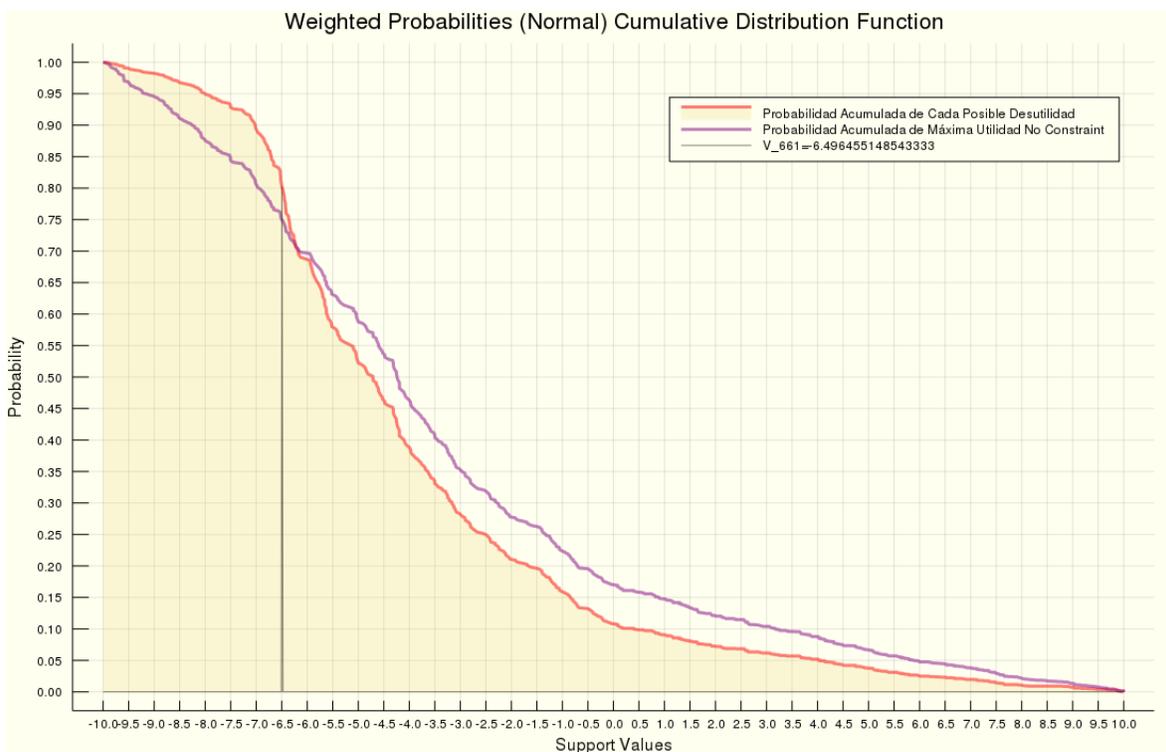


Figura 5.30: Función de Distribución Acumulada del Experimento 10 FDO3

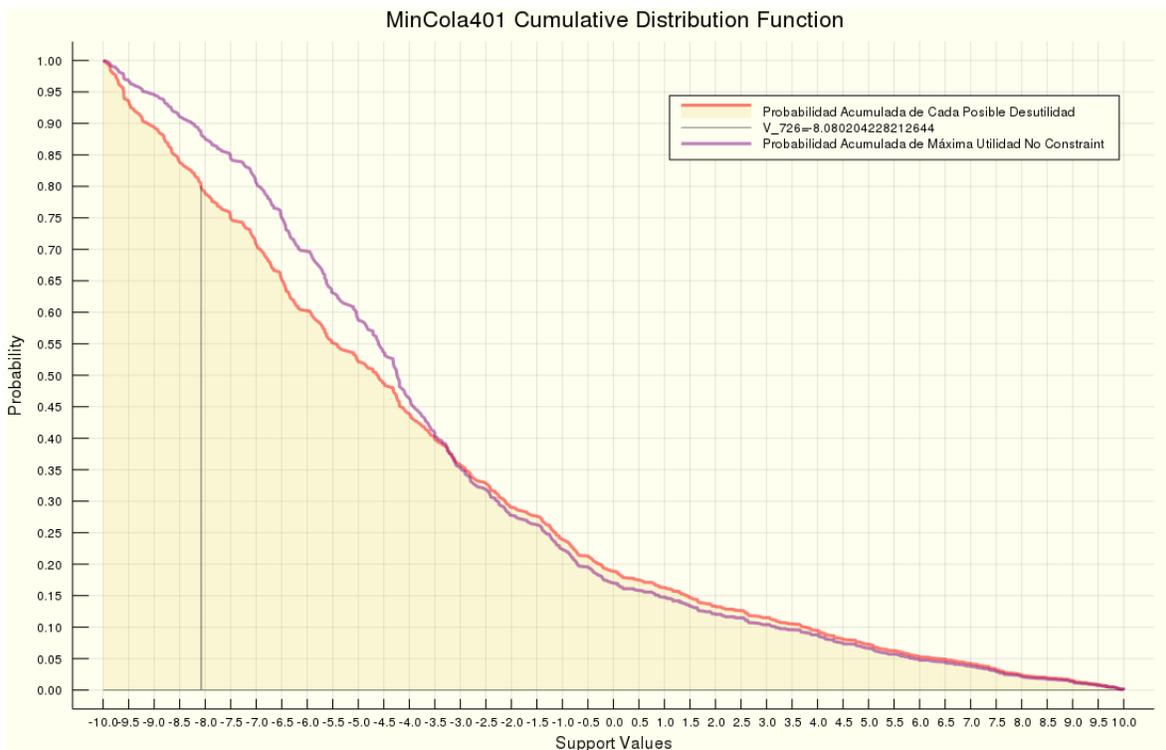


Figura 5.31: Función de Distribución Acumulada del Experimento 10 FDP100

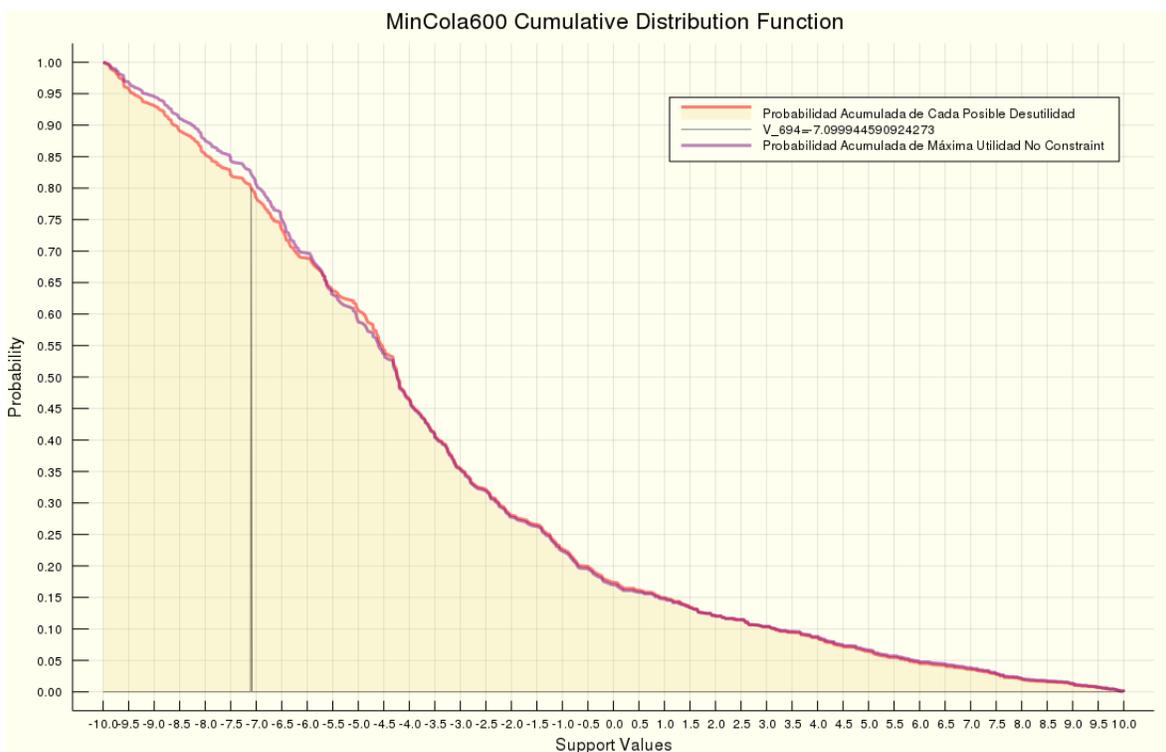


Figura 5.32: Función de Distribución Acumulada del Experimento 10 FDP50

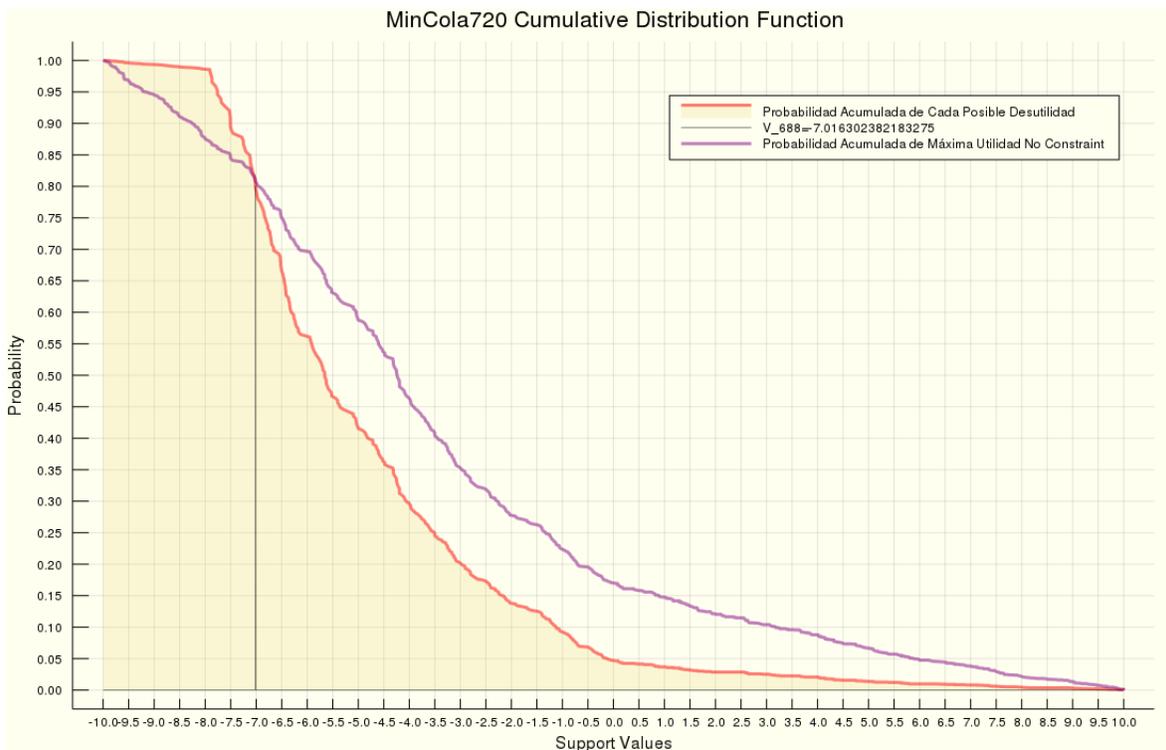


Figura 5.33: Función de Distribución Acumulada del Experimento 10 FDP20

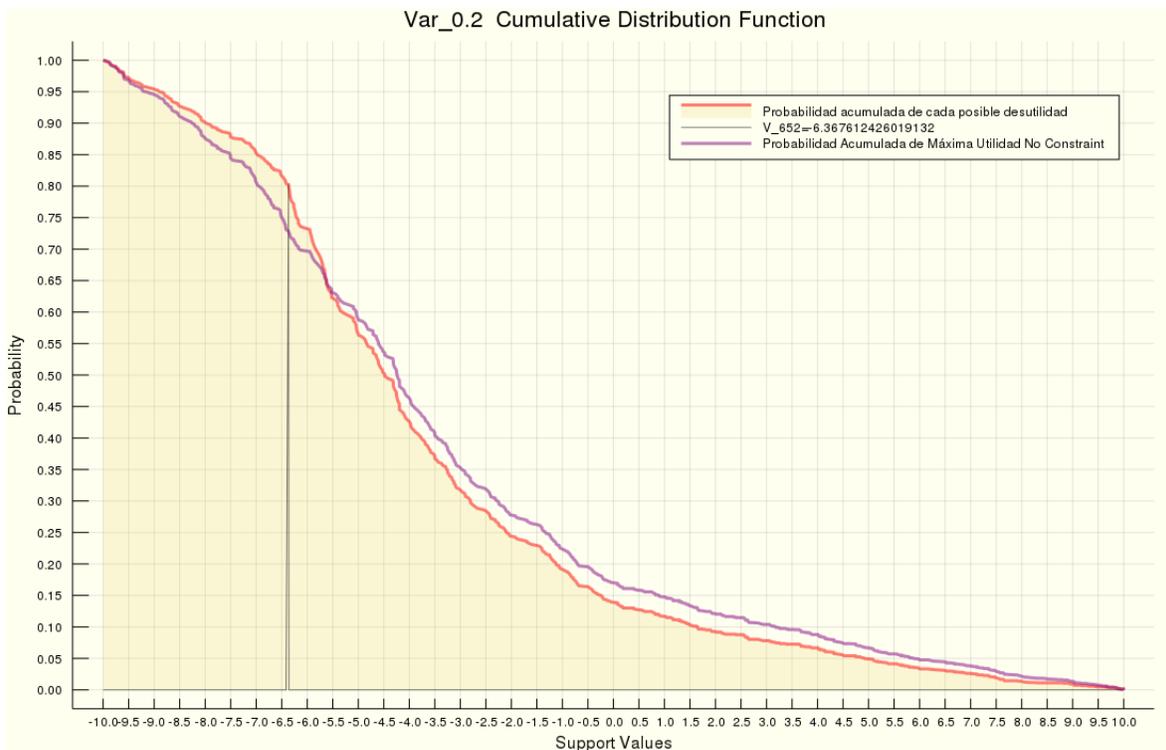


Figura 5.34: Función de Distribución Acumulada del Experimento 10 VaR_{0,2}

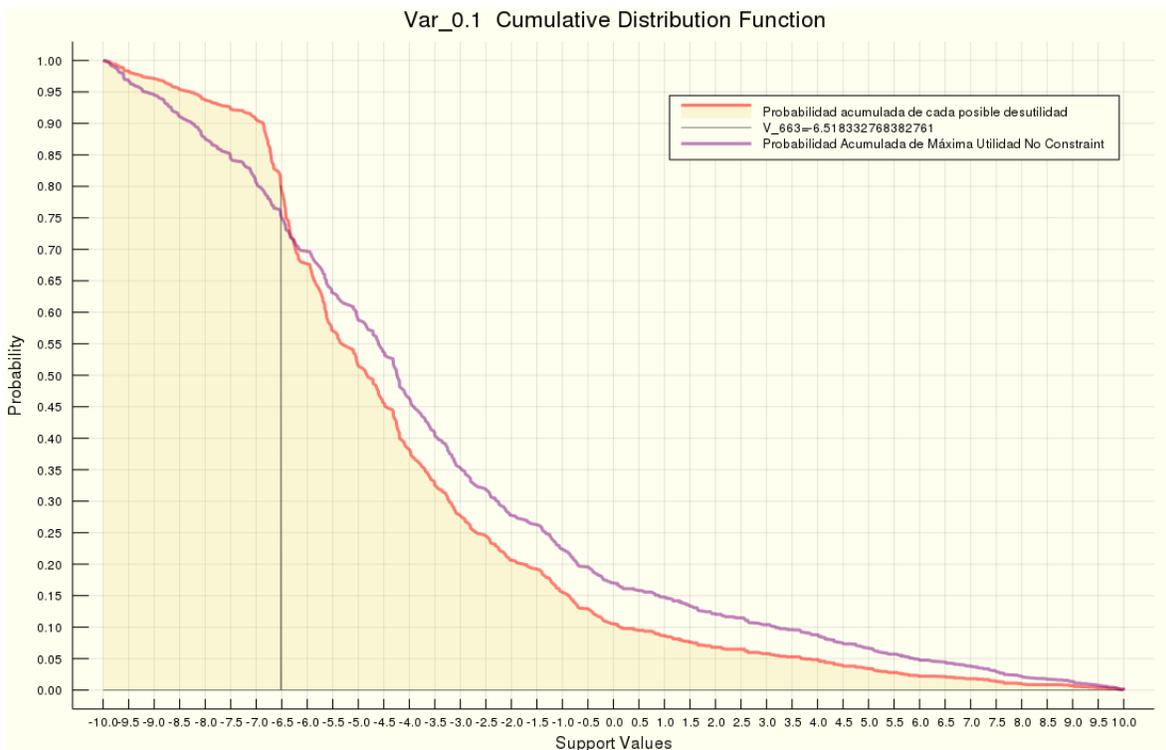


Figura 5.35: Función de Distribución Acumulada del Experimento 10 $VaR_{0,1}$

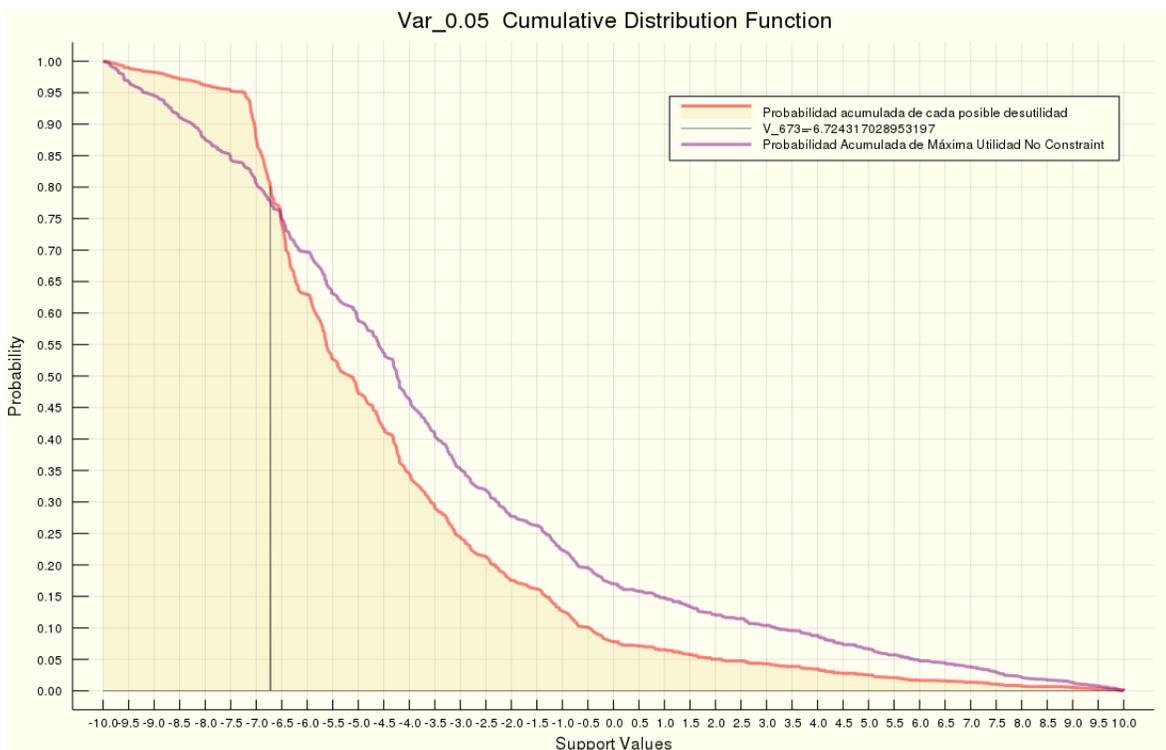


Figura 5.36: Función de Distribución Acumulada del Experimento 10 $VaR_{0,05}$

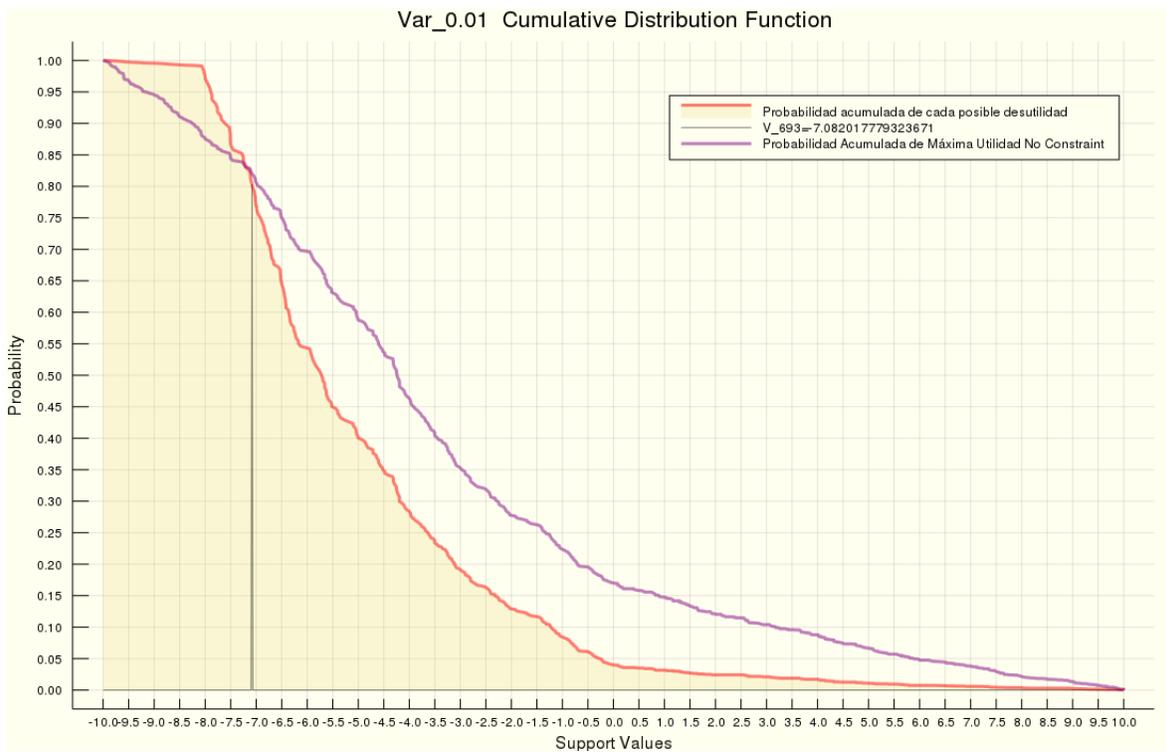


Figura 5.37: Función de Distribución Acumulada del Experimento 10 $VaR_{0,01}$

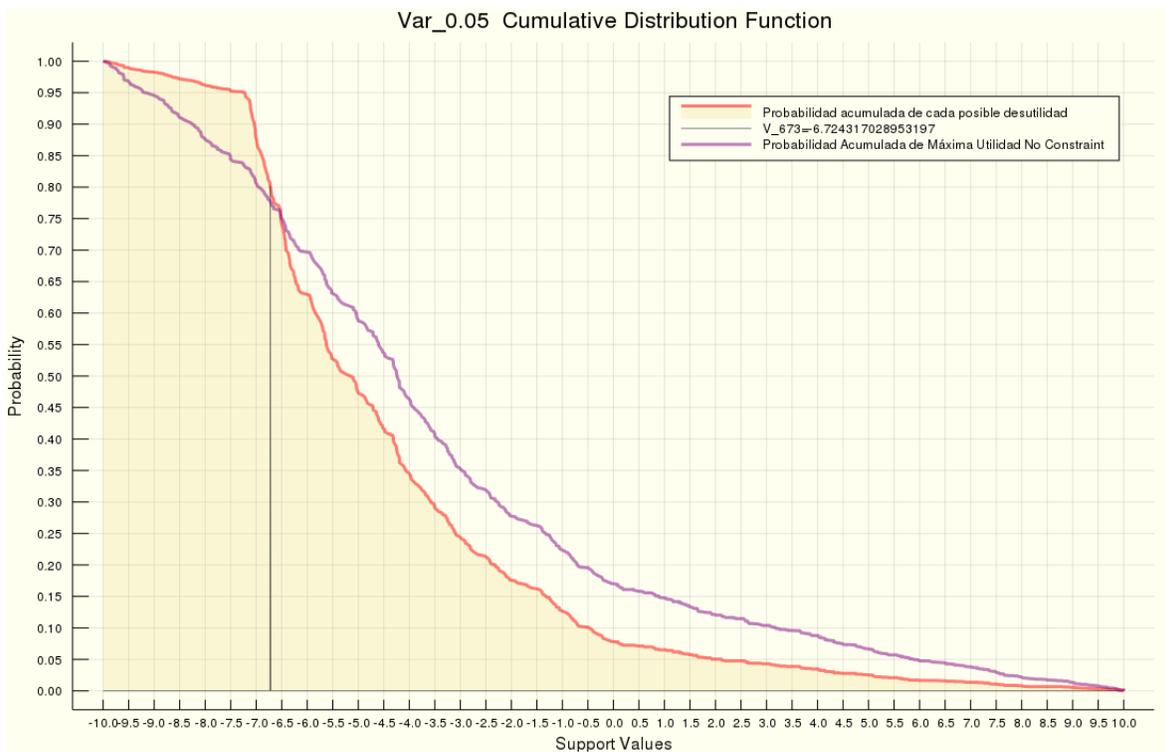


Figura 5.38: Función de Distribución Acumulada del Experimento 10 $VaR_{0,05}$

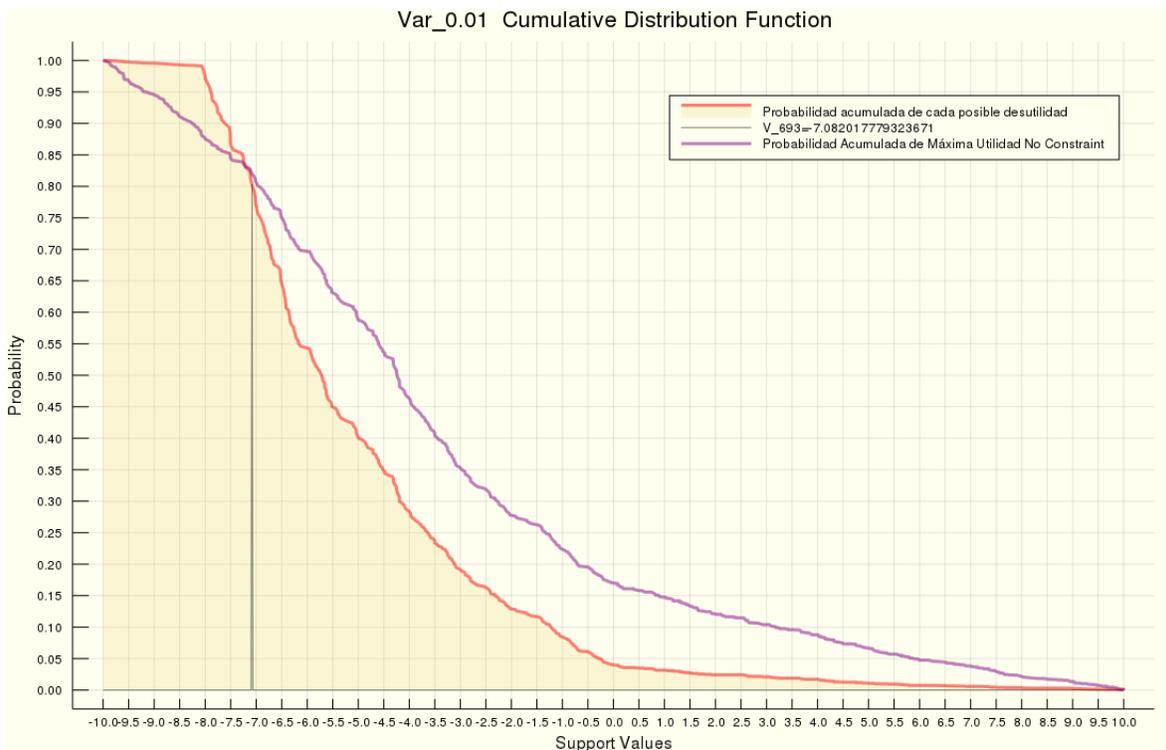


Figura 5.39: Función de Distribución Acumulada del Experimento 10 $VaR_{0,01}$

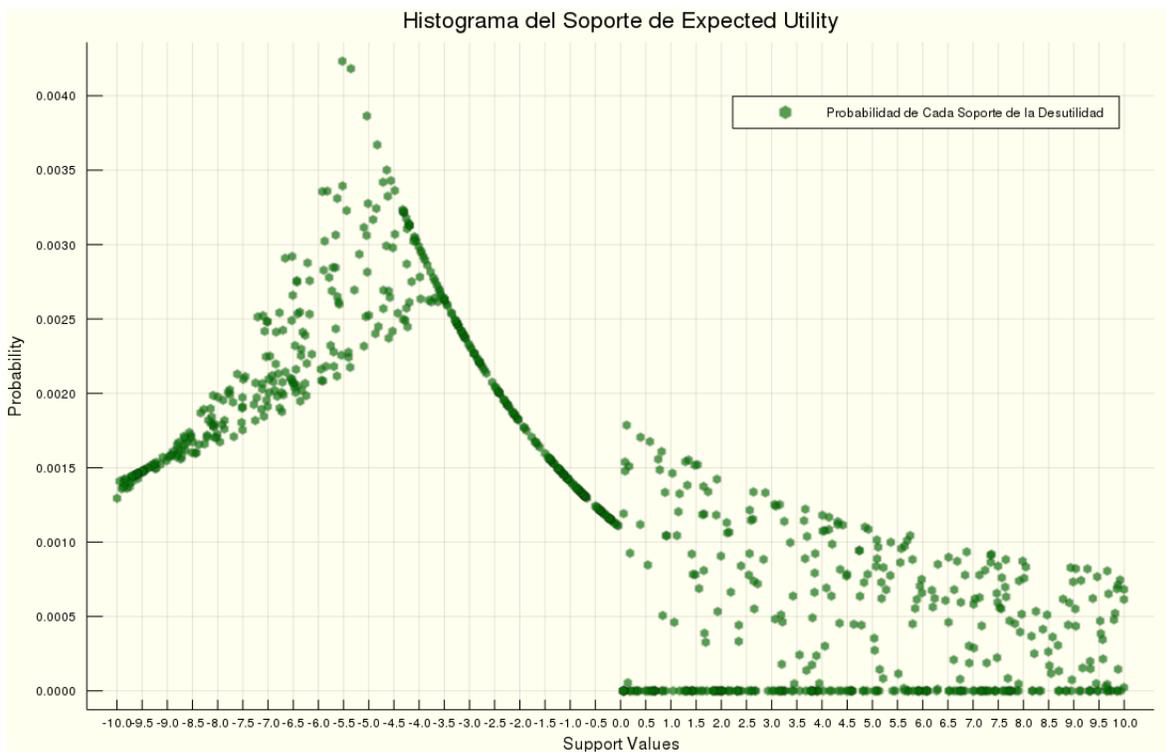


Figura 5.40: Histograma del Soporte de MUE del Experimento 10

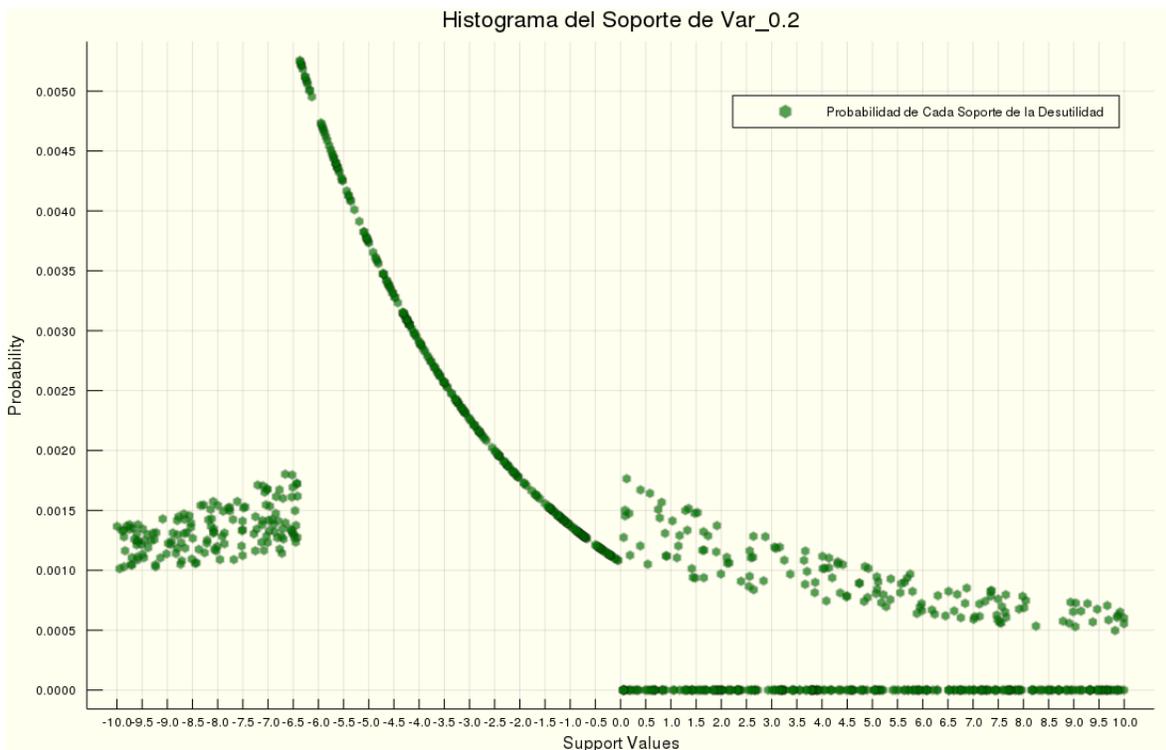


Figura 5.41: Histograma del Soporte de VaR_{0,2} del Experimento 10

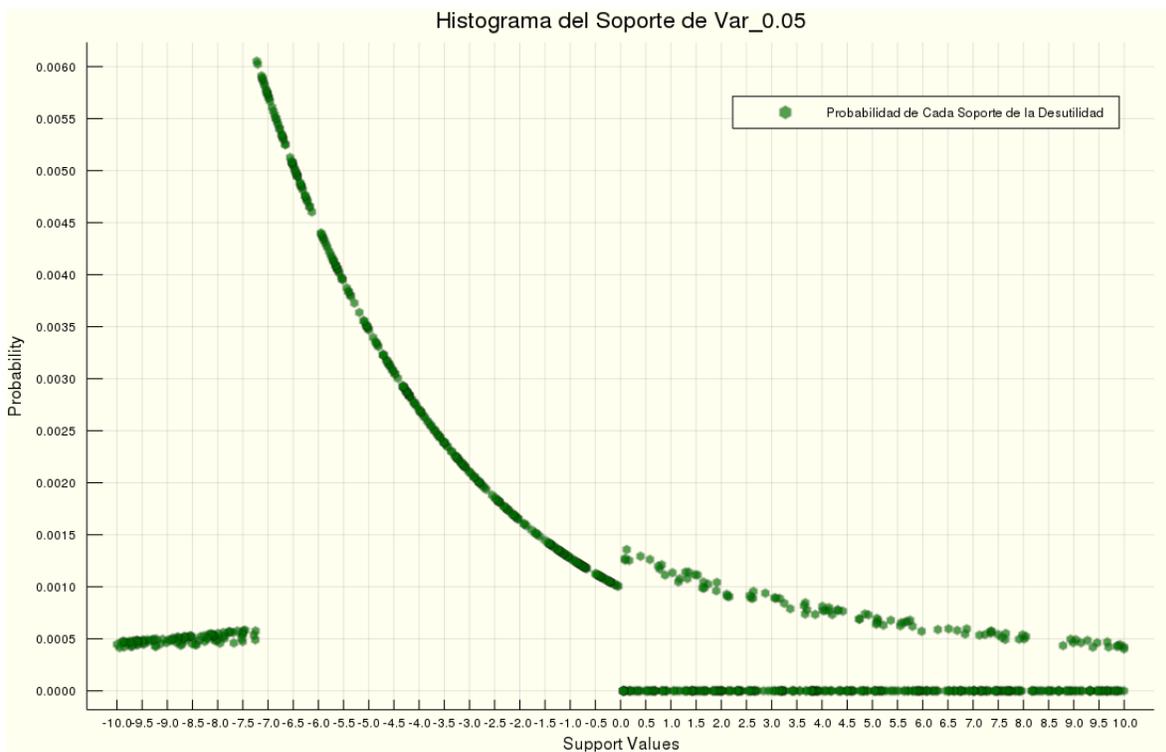


Figura 5.42: Histograma del Soporte de VaR_{0,05} del Experimento 10

Experimento 10: Con Restricciones



Figura 5.43: VaR y CVar 0.2 del Experimento 10 con Restricciones

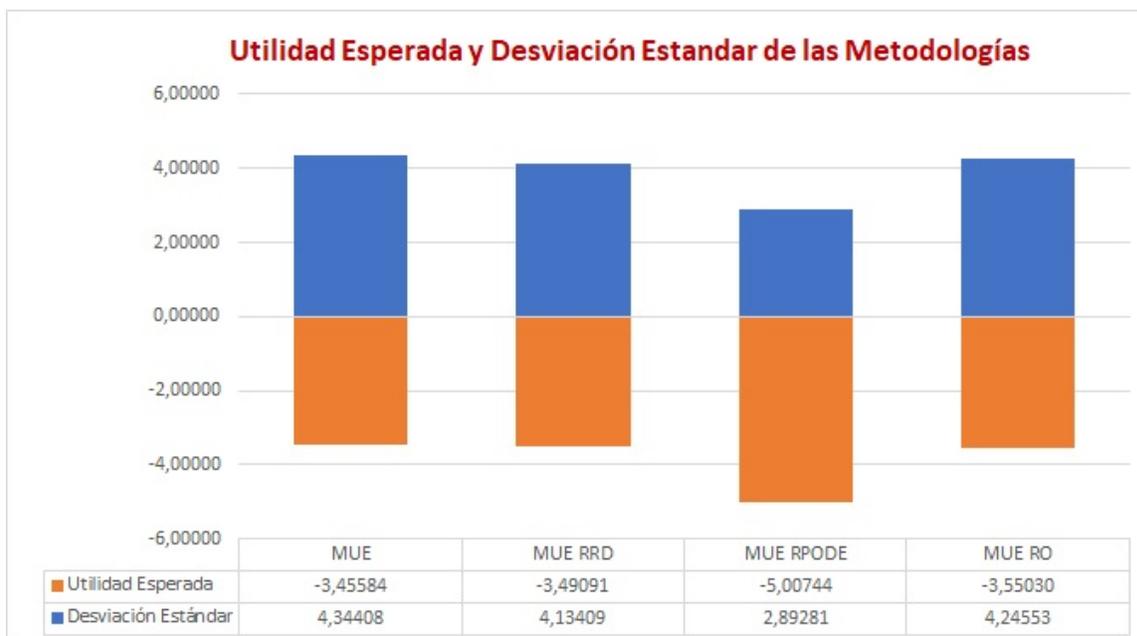


Figura 5.44: VaR y CVar 0.2 del Experimento 10 con Restricciones



Figura 5.45: VaR y CVar 0.2 del Experimento 10 con Restricciones

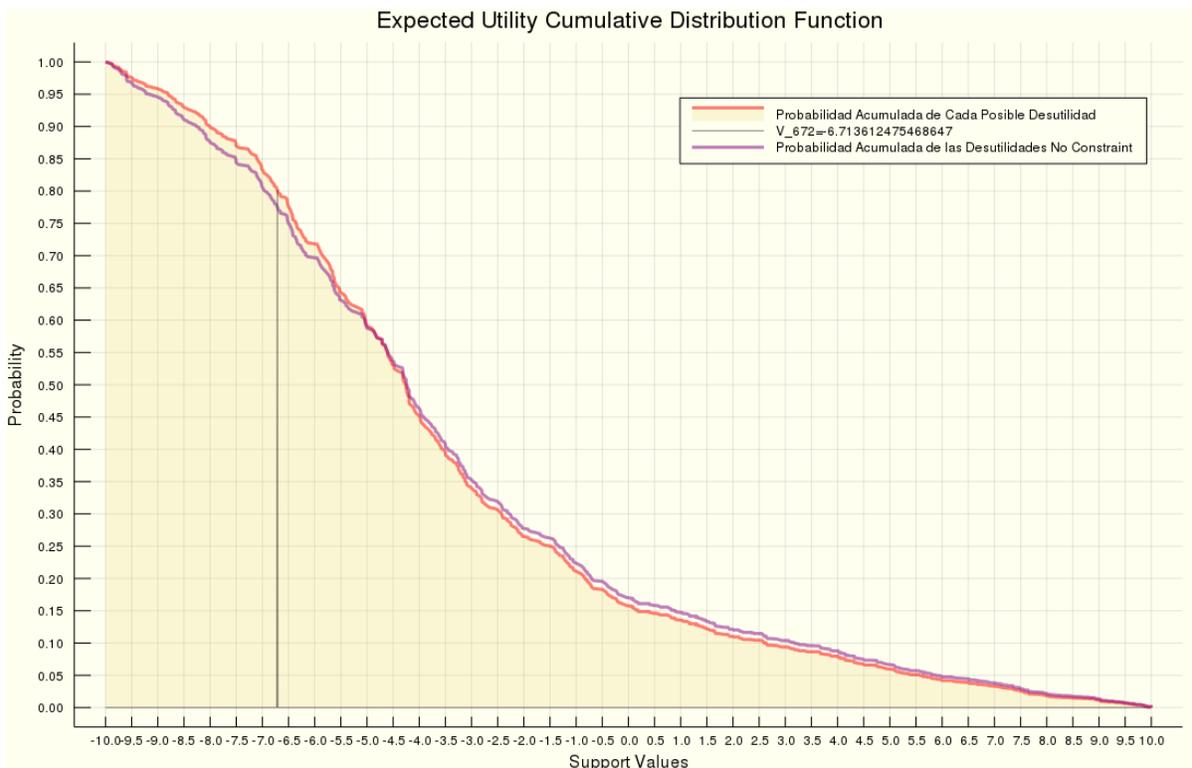


Figura 5.46: Función de Distribución Acumulada del Experimento 10 MUE RRD

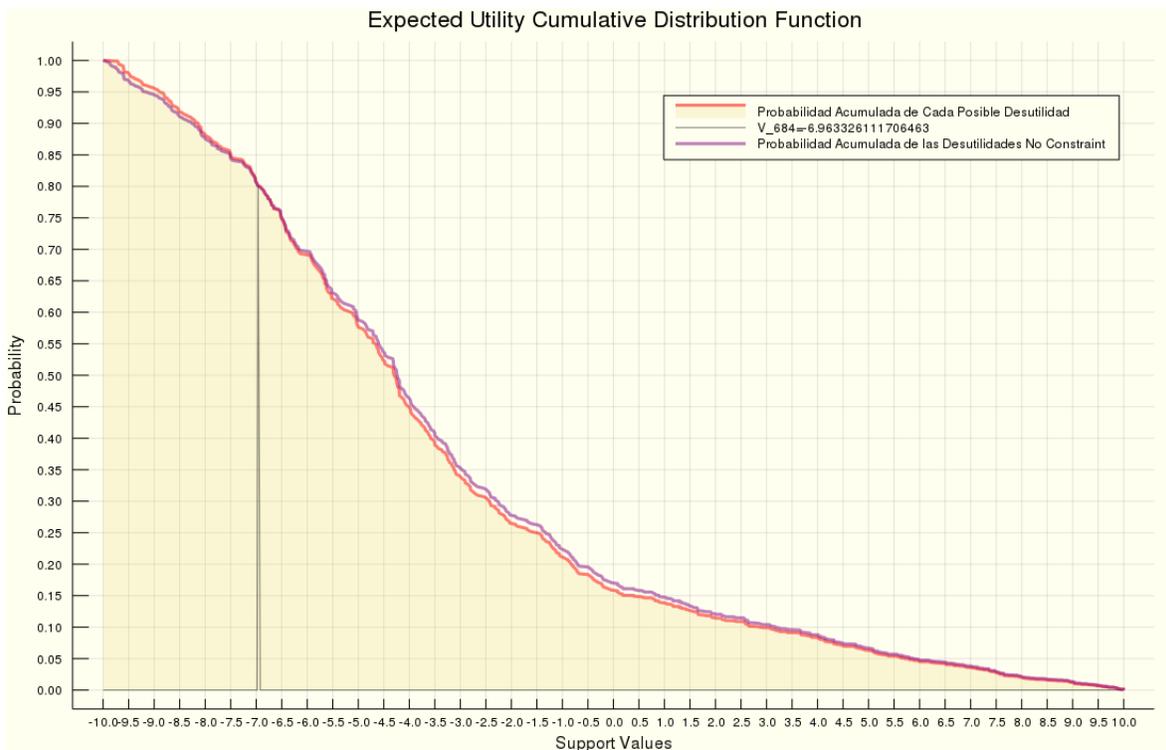


Figura 5.47: Función de Distribución Acumulada del Experimento 10 MUE RO

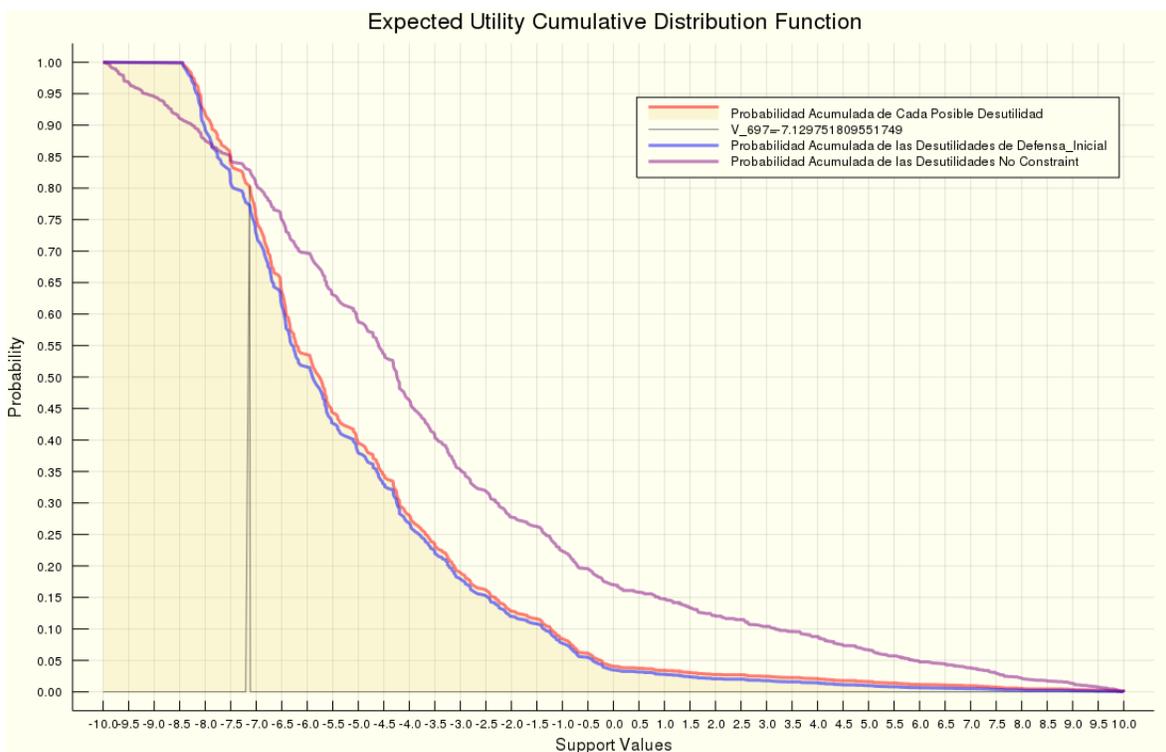


Figura 5.48: Función de Distribución Acumulada del Experimento 10 MUE RPODE