



UNIVERSIDAD DE CHILE

FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS

DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

**OPTIMIZACIÓN DE ESTRATEGIAS DE CARRERA CON COMPETENCIA
EN UN JUEGO SUMA-CERO FEEDBACK STACKELBERG**

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN GESTIÓN DE OPERACIONES
MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL

FELIPE AGUAD CASTOLDI

PROFESOR GUÍA:

CHARLES THRAVES CORTÉS-MONROY

MIEMBROS DE LA COMISIÓN:

DENIS SAURÉ VALENZUELA

FERNANDO ORDÓÑEZ PIZARRO

SANTIAGO DE CHILE

2023

OPTIMIZACIÓN DE ESTRATEGIAS DE CARRERA CON COMPETENCIA EN UN JUEGO SUMA-CERO FEEDBACK STACKELBERG

Optimizar las estrategias de paradas en pits en los deportes de motor no es trivial. La mayoría de los trabajos existentes descuidan la competencia o la consideran utilizando simulaciones o datos históricos, sin utilizar la teoría de juegos de manera completa. En este trabajo, presentamos un modelo basado en la Fórmula 1 en el que dos pilotos optimizan su estrategia de paradas en pits. Cada automóvil decide en cada vuelta si continuar en pista o hacer una parada en pits para cambiar los neumáticos a uno de los tres compuestos disponibles. Dado que las decisiones de los pilotos se afectan mutuamente debido a interacciones, tales como los adelantamientos; el problema se formula como un juego de Stackelberg de retroalimentación de suma-cero utilizando Programación Dinámica, donde en cada vuelta el piloto líder decide primero y el seguidor decide segundo. Además, los pilotos deciden simultáneamente los compuestos de neumáticos con los que comenzarán la carrera. La formulación permite la inclusión de eventos inciertos como banderas amarillas o aleatoriedad en los tiempos de vuelta.

Definimos y mostramos la existencia del equilibrio tanto del juego durante la carrera como del juego simultáneo antes de comenzar, y proporcionamos un algoritmo para encontrarlo en cada caso. Luego, resolvemos instancias numéricas del problema con cientos de millones de estados. Observamos cómo diferentes funciones objetivo de los pilotos inducen diferentes estrategias de carrera. En particular, cuando los jugadores maximizan la probabilidad de ganar en lugar de la diferencia de tiempo con respecto a su oponente, tienden a asumir acciones más riesgosas.

Nuestras instancias muestran que un piloto estratégico enfrentando a uno miope aumenta las probabilidades de ganar en más del 15 % en comparación con el caso en que ambos son estratégicos. Por último, carreras donde es más probable encontrarse con banderas amarillas tienden a aumentar las posibilidades de ganar del piloto con peor rendimiento.

*Con gratitud hacia aquellos que me guiaron y con la pasión que me impulsó;
dedico esta tesis a explorar la fascinante conexión entre el deporte y la ingeniería.*

Saludos

Agradecimientos

Quiero expresar mi más sincero agradecimiento a mi padre. Gracias a él entré a ingeniería en la Universidad de Chile; ha sido mi ídolo, inspiración y modelo a seguir desde pequeño y siempre lo será. Su incansable trabajo y sacrificio han permitido que pueda enfocarme plenamente en mis estudios. Su apoyo incondicional y su determinación para brindarme las mejores oportunidades han sido el pilar de mi educación. Le estaré eternamente agradecido por su invaluable contribución, que ha hecho posible este logro que hoy celebro en mi tesis de magíster.

Además, quiero extender mi profunda gratitud a mi amada madre, hermanos y queridos amigos. Su constante presencia, palabras de aliento y risas compartidas han sido mi fuente de fortaleza a lo largo de este camino. Gracias por creer en mí, por alentarme a seguir adelante y por recordarme constantemente que estoy en el camino correcto. Su confianza y cariño me impulsa a esforzarme cada día y me motiva a convertirme en un profesional extraordinario. Este logro es de todos nosotros.

De corazón, agradezco a mi pareja, quien ha sido un pilar imprescindible durante largos meses, brindándome apoyo y ayudándome a mantener el enfoque en mis metas. Su paciencia y aliento constante han sido fundamentales en mi camino hacia este logro.

A mis profesores, compañeros y al personal del departamento de Ingeniería Industrial, cuya buena voluntad, compromiso excepcional y excelencia en la enseñanza han hecho de la universidad un entorno gratificante para el aprendizaje y la convivencia. Su apoyo constante ha sido esencial en mi formación académica y desarrollo personal. Estoy enormemente agradecido por su dedicación y por contribuir a mi crecimiento como profesional. A cada uno de ustedes, infinitas gracias.

Tabla de Contenido

1. Introducción	1
2. Desarrollo	6
2.1. Estado del Arte	6
2.2. Metodología	8
2.3. Modelo Determinista	10
2.3.1. Función de Tiempo de Vuelta	15
2.3.2. Algoritmo	18
2.4. Modelo Estocástico	18
2.4.1. Banderas Amarillas	19
2.4.2. Tiempos de Vuelta	21
2.4.2.1. Tiempos de Vuelta bajo VSC	21
2.4.2.2. Tiempos de Vuelta bajo SC	21
2.4.2.3. Tiempos de Vuelta sin Bandera Amarilla	22
2.4.3. Programación Dinámica Estocástica	24
2.4.4. Maximizar la Probabilidad de Ganar	28
2.5. Resultados Numéricos	29
2.5.1. Jugadores Estratégicos Maximizan la Brecha de Tiempo	30
2.5.2. Maximizando las Victorias	32
2.5.3. Jugador A es Estratégico y jugador B es Miope	33
2.5.4. Variamos las Probabilidades de Bandera Amarilla	34
3. Conclusiones	36
Bibliografía	39

Anexos	42
A. Resumen de Parámetros	42
A.1. Parámetros	43
A.2. Decisiones, Estados y Variables Aleatorias	44
A.3. Decisiones, Estados y Variables Aleatorias	45
B. Mínima Brecha de Tiempo	45
C. Demostración del Teorema 2.4.2	46
D. Demostración de la Proposición 2.4.5	51
E. Escenario Alternativo con SC	51
F. Equilibrio de Nash al Maximizar la Probabilidad de Ganar	52
G. Tablas Anexas	53

Índice de Tablas

2.1.	Pago esperado de los jugadores al comienzo de la carrera, U' ; y el equilibrio de Nash mixto, π^{A^*} y π^{B^*}	30
2.2.	Distribución de victorias, tiempo promedio final de diferencia, desviación estándar, tiempo promedio final de diferencia condicionado a cada jugador ganador y estrategia de compuesto de neumáticos para iniciar la carrera (vuelta 1). . .	31
A.1.	Parámetros del problema	43
A.2.	Decisiones, estados y variables aleatorias del problema	44
A.3.	Decisiones, estados y variables aleatorias del problema estocástico	45
G.1.	Pago esperado de los jugadores al comienzo de la carrera, U'' ; y Equilibrio de Nash mixto, $\pi^{A'^*}$ y $\pi^{B'^*}$	53
G.2.	Número de estados y tiempo por vuelta para circuito de Silverstone, problema de programación dinámica.	54
G.3.	Número de estados y tiempo por vuelta para circuito de Silverstone, problema de programación dinámica.	55

Índice de Ilustraciones

2.1.	Circuito.	11
2.2.	Tiempo adicional por vuelta de los neumáticos con diferentes usos en comparación con un neumático blando nuevo.	29
2.3.	Estrategias de carrera de los jugadores (Superior), diferencia parcial de tiempo de carrera (Inferior); escenario sin bandera amarilla (Izquierda) y escenario con un VSC (Derecha).	31
2.4.	Estrategias de carrera de los jugadores (<i>Arriba</i>), diferencia de tiempo parcial de la carrera (<i>Abajo</i>), escenario en el que los jugadores quieren maximizar la brecha entre ambos pilotos (<i>Izquierda</i>), y escenario en el que ambos jugadores solo se preocupan por ganar (<i>Derecha</i>).	32
2.5.	Estrategias de carrera de los jugadores (<i>Arriba</i>), diferencia parcial de tiempo de carrera (<i>Abajo</i>), escenario donde ambos jugadores son estratégicos (<i>Izquierda</i>), y escenario donde el jugador <i>A</i> es estratégico y el jugador <i>B</i> es miope (<i>Derecha</i>).	34
2.6.	La distribución de victorias cuando la probabilidad de un SC durante la carrera va desde 0% hasta 50%. La probabilidad de tener un VSC es idéntica a la del SC.	35
E.1.	Estrategias de carrera de los jugadores (<i>Arriba</i>), diferencia parcial de tiempo de carrera (<i>Abajo</i>), escenario con SC en las vueltas 26, 27 y 28.	52
F.1.	Resultados de una carrera simulada utilizando el enfoque de maximización de la probabilidad de ganar con 50,000 escenarios.	53

Capítulo 1

Introducción

La Investigación de Operaciones (IO) y las herramientas analíticas han ampliado su alcance a varias aplicaciones en las últimas décadas. Un ejemplo de esto es su creciente presencia en múltiples deportes como el fútbol, baloncesto, béisbol y otros. Más recientemente, las carreras automovilísticas también han comenzado a captar la atención de los investigadores, en paralelo con los profesionales como los ingenieros de carreras, que están invirtiendo grandes esfuerzos en optimizar sus estrategias de carrera (ver Fortune [1]). Aunque este último implica una gran cantidad de elementos, como las vueltas en las que realizar una parada en pits¹, los compuestos de neumáticos, la gestión de combustible, la configuración del automóvil, etc.; en este artículo nos centraremos en los dos primeros pilares. En particular, enmarcaremos el problema en el contexto de la Fórmula 1; sin embargo, la metodología también se puede aplicar a otras categorías de carreras.

Las carreras de Fórmula 1 son, para muchos, el deporte más emocionante del planeta. Consisten en 44 a 78 vueltas alrededor de un circuito, tardando entre 90 a 120 minutos para completar la carrera. En promedio, cada vuelta lleva a los pilotos entre 70 y 110 segundos, dependiendo del circuito. Como regla general, cada carrera cubre alrededor de 306 kilómetros. Hay 10 equipos en competencia, cada uno con 2 pilotos participando en cada carrera. Estos 20 corredores parten desde una parrilla, con las posiciones determinadas de la sesión de clasificación realizada el día anterior, y compiten por alcanzar la mejor posición posible, con puntos distribuidos del 1 al 25 entre los 10 primeros clasificados.

El resultado de una carrera depende de varios factores además del rendimiento del automó-

¹ Los pits son el lugar al que los corredores llevan sus vehículos durante la competencia para realizar cambios de neumáticos, reparaciones mecánicas, reemplazar partes dañadas o ajustar el alerón.

vil y del piloto. Algunos de estos factores son inciertos, como los accidentes automovilísticos o las fallas del motor, mientras que otros son controlados por los ingenieros de carreras, como por ejemplo las vueltas en las que realizar paradas en pits y/o el compuesto de neumáticos con el que comenzar la carrera. Además, los eventos inciertos durante la carrera, como las banderas amarillas, hacen que la estrategia de paradas en pits sea aún más importante. (La bandera amarilla se muestra a todos los pilotos de Fórmula 1, con dos posibles escenarios: (i) Coche de Seguridad Virtual (de ahora en adelante VSC (Virtual Safety Car en inglés)), en el cual los autos deben disminuir su velocidad manteniendo la diferencia de tiempo con sus rivales, y (ii) Coche de Seguridad (de ahora en adelante SC (Safety Car en inglés)), en cuyo caso los autos se colocan en fila detrás del SC a una velocidad más baja. Las banderas amarillas generalmente se activan debido a incidentes de carrera que deben ser resueltos). Un ejemplo de esto es cómo Red Bull Racing hizo entrar a pits a Max Verstappen en el Gran Premio de Abu Dhabi en 2021 durante la bandera amarilla, cuando faltaban solo 5 vueltas, permitiéndole reemplazar un juego usado de neumáticos de compuesto duro por un nuevo juego de neumáticos de compuesto blando, lo que le permitió adelantar a Sir Lewis Hamilton (quien no entró a pits para no perder posición en la pista), ganando la carrera y el campeonato en la última vuelta. (Ver Fórmula 1 2021 [2] para más detalles). Cabe destacar que el reabastecimiento de combustible ya no se practica en la Fórmula 1 desde 2010. En el año pasado, los ingresos de la Fórmula 1 fueron de 1,800 millones de dólares, donde el 50% de esa cantidad se destinó a los equipos, con una participación igual para cada competidor (USD 42,750,000 cada uno), además de un premio monetario por rendimiento (consultar TotalSportal [3]). Esto es una motivación adicional a resolver este desafío, pues nos encontramos con cientos de miles de dólares en juego.

Los neumáticos se degradan a medida que se usan, lo que resulta en un deterioro en el rendimiento en términos de tiempos de vuelta. Para contrarrestar esto, una parada en pits brinda la oportunidad de reemplazar el conjunto actual de neumáticos desgastados por unos nuevos. En consecuencia, se espera que los tiempos de vuelta mejoren después de una parada en pits, pero el proceso en sí implica una penalización significativa de tiempo de aproximadamente 20 segundos. Esto introduce un equilibrio entre el número de paradas en pits y el nivel de degradación de los neumáticos. Otro factor a considerar es la selección de los compuestos de neumáticos para cada *stint*. Un *stint* se refiere a un conjunto consecutivo de vueltas durante las cuales se utiliza el mismo conjunto de neumáticos (por ejemplo, si

un automóvil entra en pits en las vueltas 10 y 25 en una carrera de 30 vueltas, diremos que el automóvil tiene tres *stints* de 10, 15 y 5 vueltas, respectivamente). En la Fórmula 1 hay cinco compuestos de neumáticos disponibles para condiciones de clima seco, pero solo se permiten tres para uso de todos los pilotos durante una carrera. Estos compuestos se clasifican como *soft*, *medium* y *hard* (blando, medio y duro, respectivamente). Si bien los compuestos más suaves permiten tiempos de vuelta más rápidos, también se degradan más rápidamente en comparación con los compuestos más duros. Como resultado, el rendimiento del automóvil se deteriora rápidamente cuando se utilizan neumáticos de compuesto blando. Por otro lado, los compuestos duros tienen una vida útil más larga, pero proporcionan un rendimiento ligeramente inferior, especialmente cuando los neumáticos son nuevos. Por lo tanto, el equilibrio mencionado anteriormente puede presentarse como una elección entre utilizar neumáticos más suaves para *stints* más cortos, lo que requiere paradas en pits más frecuentes, o utilizar compuestos más duros para prolongar la vida útil de los neumáticos y reducir la necesidad de paradas en pits. Para una explicación visual de estos equilibrios, consulte Chain Bear [4].

El costo en términos de pérdida de tiempo durante una parada en pits no es independiente del estado de la carrera. Eventos aleatorios como una bandera amarilla (debido a un accidente de un automóvil, por ejemplo) obligan a todos los automóviles a reducir la velocidad. En consecuencia, optar por una parada en pits bajo una bandera amarilla presenta una ventaja de tiempo en comparación con hacerlo bajo condiciones normales de carrera, ya que la velocidad en el *pit lane*² no se ve afectada por la bandera. Como resultado, ocasionalmente se observan comportamientos estratégicos, como extender el uso de los neumáticos en carreras donde la probabilidad de una parada en pits suele ser alta.

Otro factor crucial a considerar es la presencia de interacción y competencia entre los vehículos participantes. En cuanto a la interacción, existen varias formas en las que los autos se influyen mutuamente, lo que impacta en sus respectivos tiempos de vuelta. Por ejemplo, cuando un conductor adelanta a otro, ambos tiempos de vuelta se ven afectados. De manera similar, un conductor que defiende su posición contra un posible adelantamiento o que sigue de cerca a otro automóvil en frente puede tener implicaciones en los tiempos de vuelta. Esta interacción puede resultar en un aumento de los tiempos de vuelta para el seguidor durante las curvas, mientras que potencialmente tiene el efecto opuesto en las rectas.

² Pit lane (calle de pits) es la pista asfaltada que conecta el circuito con el pit stop (paradas en pits).

En consecuencia, el tiempo de vuelta de un automóvil se ve influenciado no solo por factores internos como el compuesto de neumáticos o el uso de los mismos, sino también por factores externos que involucran interacciones entre los autos. Por lo tanto, las acciones de un conductor, particularmente su estrategia de paradas en pits, deben considerar estas interacciones. Además, en términos de competencia, cada conductor optimiza su estrategia de paradas en pits y compuestos de neumáticos, teniendo en cuenta las acciones de otros conductores también.

En este trabajo presentamos un modelo de teoría de juegos en el que dos conductores optimizan sus estrategias de paradas en pits durante la carrera. Más específicamente, en cada vuelta, ambos autos deciden si continuar en pista o cambiar los neumáticos por un compuesto específico. Cada vuelta se modela como un juego de Stackelberg³, en el cual el primer conductor juega primero. Por lo tanto, el líder en cada etapa no está fijo, sino que se alterna dependiendo del conductor que esté liderando la carrera. Además de la estructura secuencial del juego, antes del inicio de la carrera, los conductores eligen en un juego simultáneo el compuesto de neumáticos con el que iniciarán la carrera.

Incluso en un escenario determinista, es decir, sin banderas amarillas ni incertidumbre en los tiempos de vuelta, la forma extendida del juego tiene un número exponencial de nodos. De hecho, consideremos una carrera con 50 vueltas y tres compuestos de neumáticos. En cada vuelta, los conductores tienen cuatro acciones posibles: cambiar los neumáticos por uno de los tres compuestos o continuar en pista. En este ejemplo, el número total de nodos terminales es igual a $4^{50} \cdot 4^{50}$, lo cual es impracticable de resolver en un tiempo razonable. Sin embargo, cada secuencia de acciones de ambos conductores hasta una vuelta en particular se puede reducir a un *estado*. Como resultado, proponemos modelar el juego como un juego de Stackelberg con retroalimentación⁴ y resolverlo mediante programación dinámica.

Las principales contribuciones de este trabajo se pueden resumir en los siguientes puntos:

- Formulamos el problema de optimización de paradas en pits con competencia para dos pilotos utilizando un juego de Stackelberg de retroalimentación de suma-cero⁵. El modelo permite incorporar escenarios estocásticos que incluyen eventos de bandera amarilla,

³ Un juego de Stackelberg es un juego secuencial en el que un jugador líder toma una decisión y luego, después de observar la acción del líder, el seguidor elige la suya.

⁴ Se llama estructura de retroalimentación cuando los jugadores conocen y toman en consideración todas las estrategias pasadas.

⁵ Se dice de suma-cero cuando la ganancia de un jugador es necesariamente la pérdida de otro jugador.

como VSC o SC, y la incertidumbre en los tiempos de vuelta de los pilotos. Además, el juego incorpora la decisión inicial de los neumáticos, que se juega de forma simultánea. El modelo también permite diferentes tipos de funciones objetivo, como maximizar la diferencia de tiempo con respecto al otro conductor o maximizar las probabilidades de ganar.

- Definimos el equilibrio de retroalimentación de Stackelberg del juego, que involucra las estrategias de paradas en pits de los jugadores, y el equilibrio de Nash de retroalimentación de Stackelberg, que agrega la estrategia inicial de los neumáticos. Mostramos la existencia de estos equilibrios y formulamos un algoritmo de inducción hacia atrás con programación dinámica que se juega secuencialmente para ambos conductores en todas las vueltas de la carrera. Mostramos que el algoritmo propuesto encuentra el equilibrio del juego.
- Podemos resolver instancias del problema con cientos de millones de estados en cada vuelta en cuestión de horas. Mostramos que la toma de decisiones estratégicas mejora el resultado de la carrera para los jugadores. De hecho, observamos casos en los que las probabilidades de ganar de un conductor aumentan en más del 15 % cuando su oponente no es estratégico. Además, mostramos cómo cambian las decisiones de los jugadores al utilizar diferentes funciones objetivo.

En resumen, el modelo propuesto aborda tres preguntas clave que surgen al prepararse para una carrera: ¿Con qué compuesto de neumáticos debería comenzar la carrera un piloto? ¿En qué vueltas debería entrar al pit lane? ¿Qué compuesto de neumáticos debería elegir para cada *stint*?

En la Sección 2.1 revisamos la literatura relacionada con este trabajo. La Sección 2.3 muestra el modelo determinístico y la Sección 2.4 muestra el modelo estocástico. Los resultados de los cálculos numéricos se presentan en la Sección 2.5. Finalmente, las conclusiones se exponen en la Sección 3. Las demostraciones se incluyen en el Apéndice.

Capítulo 2

Desarrollo

2.1. Estado del Arte

La optimización de las estrategias de carrera se ha explorado desde diferentes perspectivas. McLaren Racing Limited [5] analizó la optimización del combustible durante la carrera, lo cual incluye la cantidad inicial de combustible con la que comenzar la carrera y las vueltas y la cantidad de combustible para repostar en una parada en pits.

La optimización de la batería también ha sido objeto de estudio por parte de Liu et al. [6] y Anselma [7], donde las variables de decisión son el agotamiento y la gestión térmica de la batería.

Bekker y Lotz [8] utilizaron simulación para modelar maniobras de adelantamiento, paradas en pits y fallas de automóviles, generando el resultado de la carrera en función de los datos de entrada que describen la carrera. Al evaluar diferentes estrategias de paradas en pits (que son datos de entrada del modelo de simulación), los autores pueden encontrar la estrategia de carrera óptima. En una dirección similar, Phillips [9] desarrolló un simulador que utiliza varios datos de entrada de los pilotos, incluyendo su rendimiento en una de las prácticas libres⁶, degradación de neumáticos, variabilidad en los tiempos por vuelta y otros. También utilizando simulación, Helmeier et al. [10] amplían el trabajo anterior al considerar elementos adicionales, como utilizar tiempos de otras prácticas o la sesión de clasificación⁷ y

⁶ Cada carrera tiene tres sesiones de práctica antes de la carrera donde los pilotos pueden probar el circuito. Los datos de la práctica utilizados como entrada en este trabajo corresponden a la segunda práctica (también denominada FP2).

⁷ Las posiciones de la parrilla de salida de la carrera se determinan durante la sesión de clasificación. Consulta The Sport Rush [11] para más detalles.

verificar la brecha entre los autos en más puntos que al final de la vuelta. Helmeier et al. [12] generalizan la simulación de la carrera al considerar incertidumbre en diferentes aspectos, como los tiempos por vuelta, la ocurrencia de eventos de bandera amarilla, la duración de las paradas en pits y los accidentes y fallas. F1 Strategy Blog [13] creó un código de código abierto que simula la carrera. En este caso, se utilizan estrategias de carrera de años anteriores. Aunque las técnicas de simulación son útiles para incorporar varias variables del problema, su principal problema es la dificultad para encontrar las estrategias óptimas de paradas en pits considerando la dinámica competitiva de la carrera.

Las técnicas de aprendizaje automático también se han utilizado para predecir el rendimiento de los autos, incluyendo la estrategia de paradas en pits como una de las variables explicativas utilizadas. En este marco, Tulabandhula, Rudin [14] y Choo [15] utilizaron técnicas de aprendizaje automático entrenadas con datos de carreras anteriores para predecir los tiempos de vuelta utilizando variables como las paradas en pits, el rendimiento del auto y otras. Más recientemente, Helmeier et al. [16] combinaron redes neuronales artificiales (NN, neural networks en inglés) y simulación para encontrar la mejor estrategia de carrera para cada conductor. Las NN se entrenaron con datos de carreras de temporadas anteriores para predecir si se debe realizar una parada en pits y qué asignación de neumáticos utilizar en caso de hacerlo. Luego, se simuló la carrera, donde las acciones del conductor en cada vuelta se basaron en las predicciones de las NN. Una de las entradas utilizadas en sus NN es el número de paradas en pits restantes en la carrera, que se encontró resolviendo un programa cuadrático mixto entero asumiendo que no hay interacción entre los autos.

A diferencia de estos enfoques, nuestro trabajo difiere en que proponemos un método para calcular las estrategias óptimas de dos conductores que compiten en un entorno de teoría de juegos. Nuestro enfoque se basa en determinar las estrategias óptimas a través de un análisis de teoría de juegos, en lugar de depender de métodos de predicción calibrados en decisiones de carreras pasadas. Al considerar la naturaleza competitiva de la carrera y las interacciones estratégicas entre los conductores, nuestro modelo proporciona un marco más completo para optimizar las estrategias de paradas en pits.

Uno de los trabajos más similares a este es el realizado por Heine y Thraves [17]. Los autores presentan un enfoque de solución mediante programación dinámica para encontrar la estrategia óptima de paradas en pits en un entorno sin competencia. La principal diferencia radica en que este trabajo proporciona un modelo más general al permitir la competencia

entre dos conductores. Esto permite incorporar eventos de interacción entre los autos, al tiempo que se tiene en cuenta el aspecto de teoría de juegos en la toma de decisiones óptimas de ambos jugadores. En particular, se utiliza un juego de Stackelberg de retroalimentación de suma-cero con conjuntos finitos de acciones y horizonte de tiempo. Los juegos de Stackelberg de retroalimentación se caracterizan por tener múltiples etapas en las que en cada etapa todos los jugadores conocen el estado en el que ha evolucionado el juego y hay un jugador que actúa como líder en cada una de las etapas. Este tipo de juego fue introducido por primera vez en Simaan [18] y Cruz [19]. Las aplicaciones de los juegos de Stackelberg de retroalimentación en la gestión de operaciones se muestran en Li y Sethi [20]. Para más detalles, consulte Başar y Olsder [21].

Un aspecto interesante del problema abordado aquí es la alternancia del rol de líder en diferentes etapas (vueltas) del juego, dependiendo de quién esté liderando la carrera en la vuelta actual (a diferencia de la mayoría de los escenarios de Stackelberg dinámico en los que el mismo jugador actúa como líder en cada etapa del juego). Uno de los pocos trabajos que ha abordado situaciones con líderes alternantes es el de Nie et al. [22]. En este trabajo, los autores presentaron un marco de programación dinámica para resolver juegos de Stackelberg de retroalimentación dinámica en tiempo discreto, donde el líder en cada etapa depende de la función de pago de los jugadores en esa etapa específica. En nuestro caso, el líder de cada etapa dependerá de un componente de la variable de estado. Además, nuestra configuración del juego tiene en cuenta elementos estocásticos.

2.2. Metodología

La metodología seguida en esta tesis se dividió en varias etapas.

Abordamos este desafío desde la perspectiva de la teoría de juegos. En un principio, intentamos resolverlo utilizando la herramienta de inducción hacia atrás, pero debido a su tamaño considerable, decidimos cambiar de enfoque y utilizar programación dinámica.

Para obtener una mejor comprensión del problema, comenzamos construyendo un modelo determinista básico. Posteriormente, se fue añadiendo gradualmente complejidad al modelo, incorporando funciones de tiempos de interacción entre pilotos, desgaste no lineal de neumáticos, reglas de la FIA⁸ a considerar durante una carrera y el consumo de combustible, entre

⁸ Federación Internacional del Automóvil, FIA, es mundialmente conocida por regular las competiciones de automovilismo más importantes del mundo, como la Fórmula 1.

otros factores relevantes.

Una vez que el modelo determinista estuvo completo, se procedió a abordar el aspecto estocástico del problema, ya que se reconoció la presencia de aleatoriedad en la realidad, como lo serían las banderas amarillas, los adelantamientos en las zonas de DRS y la interacción entre pilotos vuelta a vuelta. De manera similar a la etapa anterior, se fueron agregando y refinando funciones en el modelo estocástico a medida que surgían nuevas ideas o se identificaban errores. Además, durante la fase de desarrollo del modelo estocástico, surgieron interrogantes adicionales que enriquecieron la investigación. Se exploraron escenarios en los cuales los pilotos tenían diferentes objetivos, como maximizar su probabilidad de ganar o minimizar el riesgo de perder. También se consideró la posibilidad de que uno de los jugadores fuera miope, tomando decisiones en función de minimizar su propio tiempo de carrera e ignorando rivalidad, mientras que el otro jugador adoptara una estrategia acorde las acciones de su competencia, sus propias decisiones y cómo la combinación de ambas afectan su tiempo por vuelta. Estas consideraciones añadieron una capa adicional de complejidad al modelo estocástico, permitiendo explorar diversas dinámicas estratégicas y su impacto en los resultados de las carreras simuladas. La inclusión de estas variantes en la investigación contribuyó a obtener una perspectiva más amplia y realista del problema estudiado.

El siguiente paso consistió en implementar el modelo en código. Inicialmente, se utilizó Python y se trabajó en la optimización de las funciones relacionadas con los tiempos de carrera, la interacción entre pilotos y la resolución de la programación dinámica estocástica. A medida que se avanzaba, se buscó mejorar la eficiencia del código, ya que los tiempos de ejecución eran prolongados. Se logró crear e implementar un código paralelo en C como una biblioteca que puede ser llamada desde Python, lo que permitió acelerar significativamente la velocidad de procesamiento.

Para alcanzar aún mayores niveles de rendimiento, se contó con la asistencia de la supercomputadora de NLHPC⁹. Esto facilitó la realización de miles de simulaciones, brindando una visión más completa del comportamiento de los pilotos en diferentes escenarios de carrera. Se llevó a cabo un análisis iterativo de los resultados y gráficos obtenidos, lo que permitió descubrir hallazgos interesantes y validar las simulaciones con la realidad.

En resumen, la metodología empleada en esta tesis abarcó el análisis desde la teoría de

⁹ National Laboratory for High Performance Computing (Laboratorio Nacional de Computación de Alto Rendimiento).

juegos, el desarrollo de un modelo determinista y estocástico, la implementación en código utilizando Python y C, la utilización de recursos computacionales de alto rendimiento y la realización de numerosas simulaciones para obtener resultados y gráficos relevantes.

2.3. Modelo Determinista

En esta sección presentamos un modelo de Juego Secuencial modelado con Programación Dinámica, donde en cada etapa (vuelta) los jugadores (pilotos) deciden en función del estado actual (como la asignación de compuestos de neumáticos, degradación de los neumáticos, etc.). Luego, proporcionamos detalles sobre el cálculo de los tiempos por vuelta.

Consideremos a dos pilotos, A y B , que compiten en una carrera de N vueltas. En cada vuelta n , los pilotos deciden si continuar en pista o hacer una parada para cambiar neumáticos a uno de los compuestos disponibles en el conjunto $\mathcal{T} = \{1, 2, 3\}$, donde 1=Suave, 2=Medio y 3=Duro. Suponemos que el neumático $t \in \mathcal{T}$ tiene una vida útil de u_t vueltas. Denotemos por $x_n^A, x_n^B \in \mathcal{T}_0 := \{0\} \cup \mathcal{T}$ las decisiones de los pilotos A y B en la vuelta n , respectivamente. Nuestra convención es que estas variables toman el valor cero si no hay parada en pits. Naturalmente, la secuencia de decisiones dentro de una vuelta es tal que el piloto que va delante decide primero. Es importante destacar que en este juego secuencial, el orden en el que los pilotos juegan en cada etapa, es decir, en cada vuelta, puede no ser el mismo durante toda la carrera. Por ejemplo, si el piloto A comienza la carrera en primer lugar y en algún momento es adelantado por su oponente, A pasará de ser un líder (en el sentido de la teoría de juegos) a ser un seguidor.

Sin pérdida de generalidad, asumimos que la entrada a pits está justo después de la línea de meta de la carrera¹⁰, ver Figura 2.1. Luego, una vez que el piloto que lidera la carrera observa las variables de estado (que se definirán en breve) al comienzo de una vuelta, decide su acción. Después, el piloto que está en segundo lugar decide asumiendo que conoce las variables de estado, así como la decisión de su oponente que ya ha jugado. Aquí estamos asumiendo que el piloto que está en segunda posición puede ver, no solo si el piloto de adelante hace una parada (lo cual es razonable asumir), sino también cuál será su nuevo compuesto de neumáticos. Creemos que esto es razonable por dos motivos: primero, el piloto de atrás (en particular gracias a sus ingenieros) podría ver qué tipo de neumáticos se están preparando

¹⁰ En la práctica, la entrada a pits está ligeramente antes de la línea de meta y la salida está unos cientos de metros después de esta.

para instalar en el pit stop de su oponente, y segundo, el piloto en segunda posición podría calcular racionalmente la mejor estrategia del piloto líder dada la situación actual.

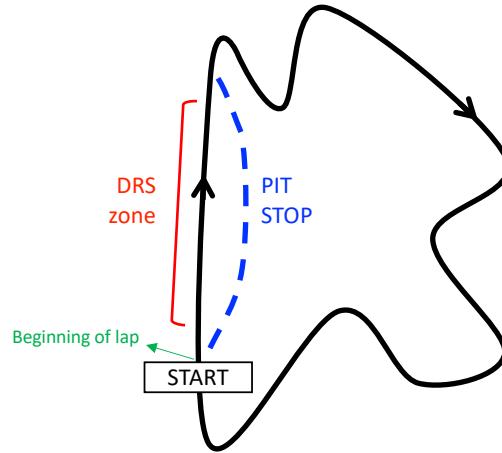


Figura 2.1: Circuito.

El estado en la vuelta n se describe mediante los siguientes elementos que dependen del estado de cada automóvil, así como de elementos que dependen de ambos automóviles. Para cada conductor $i \in \{A, B\}$ tenemos como variables de estado el compuesto de neumáticos denotado por t_n^i , el desgaste de los neumáticos denotado por w_n^i en unidades de vueltas, y una variable indicadora binaria m_n^i que toma el valor 0 si el conductor ha utilizado solo un compuesto de neumáticos, y 1 en caso contrario. Esta variable indicadora nos permite saber si el automóvil termina la carrera cumpliendo la regla que establece que se deben utilizar al menos dos compuestos de neumáticos diferentes durante la carrera. Existe una regla que establece que los conductores deben usar al menos dos compuestos de neumáticos diferentes durante la carrera, excepto si el conductor está utilizando un compuesto de neumáticos para lluvia (ver MotorSport [23] para más detalles). La última variable de estado es la diferencia de tiempo entre el conductor A y B , denotada por g_n . Nuestra convención es que un valor negativo indica que el conductor A está por delante del conductor B por $-g_n$ segundos. En cuanto a la variable de estado del desgaste de los neumáticos, w_n , su valor máximo se establece en $u_{t_n} + 1$, es decir, una vuelta adicional sobre la vida útil recomendada del neumático actual, para indicar el caso en que la vida útil del neumático ha terminado y, por lo tanto, el automóvil no puede continuar. En este caso, asumimos que la acción del conductor es el conjunto unitario $\{0\}$. Combinando todos los componentes de la variable de estado, la

variable de estado en la vuelta n se denota por:

$$s_n = (t_n^A, w_n^A, m_n^A, t_n^B, w_n^B, m_n^B, g_n). \quad (2.1)$$

La transición de las variables de estado de una etapa a la siguiente, como función de las variables de estado anteriores y las acciones de los conductores, es decir, $s_{n+1}(s_n, x_n^A, x_n^B)$, se puede expresar de la siguiente manera:

$$t_{n+1}^A = t_n^A \cdot \mathbb{1}_{\{x_n^A=0\}} + x_n^A \cdot \mathbb{1}_{\{x_n^A \neq 0\}} \quad (2.2)$$

$$w_{n+1}^A = \min\{w_n^A \cdot \mathbb{1}_{\{x_n^A=0\}} + 1, u_{t_n} + 1\} \quad (2.3)$$

$$m_{n+1}^A = \max\{m_n^A, \mathbb{1}_{\{x_n^A \neq 0, x_n^A \neq t_n^A\}}\} \quad (2.4)$$

$$g_{n+1} = g_n + \mu_n^A(s_n, x_n^A, x_n^B) - \mu_n^B(s_n, x_n^A, x_n^B), \quad (2.5)$$

donde μ_n^A y μ_n^B son las funciones de tiempo por vuelta de los conductores A y B , respectivamente (más detalles sobre esto se proporcionan en la siguiente sección). Omitimos las ecuaciones de transición para t_{n+1}^B , w_{n+1}^B , m_{n+1}^B y m_{n+1}^B , ya que estas son análogas a las de las Ecuaciones (2.2), (2.3) y (2.4). La Ecuación (2.2) establece que el compuesto de neumáticos de la siguiente vuelta será el mismo que el de la vuelta anterior en caso de no entrar a pits, de lo contrario, será el compuesto instalado en el vehículo durante la parada en pits. La Ecuación (2.3) establece que el desgaste de los neumáticos aumenta en una unidad de una vuelta a la siguiente, excepto cuando ocurre una parada en pits. En caso de una parada en pits, el desgaste de los neumáticos al comienzo de la siguiente vuelta ($n + 1$) será igual a uno. En la Ecuación (2.4), la variable indicadora de la próxima vuelta, m_{n+1} , es el máximo entre su valor anterior y uno en caso de cambiar los neumáticos por un compuesto diferente. Luego, esta variable de estado no decrece y es binaria; solo puede tomar valores $\{0, 1\}$. La actualización de la diferencia de tiempo entre los pilotos se da en la Ecuación (2.5). En particular, la brecha de tiempo en la siguiente vuelta es igual a la brecha de tiempo anterior más la diferencia de tiempo por vuelta entre ambos pilotos durante sus vueltas actuales.

El objetivo de cada jugador es terminar la carrera con la máxima diferencia de tiempo respecto a su oponente¹¹. Por lo tanto, el juego se puede modelar como un juego de suma-

¹¹ Alternativamente, podríamos resolver un modelo donde el objetivo de los jugadores sea terminar la carrera por delante de su rival, sin importar la diferencia de tiempo. Se podría demostrar fácilmente que el equilibrio de este juego también es un equilibrio del juego analizado en el artículo.

cero donde la *función de utilidad* (o más precisamente la función de utilidad para uno de los jugadores y una función de costo para el otro) es la diferencia de tiempo entre ambos pilotos al cruzar la línea de meta en la vuelta N . Nuestra convención es que esta diferencia de tiempo refleja el número de segundos en los que A está por delante de B , por lo tanto, al final de la carrera se debe cumplir $V_{N+1} = g_{N+1}$ (el índice $N + 1$ en este caso significa: al comienzo de la hipotética vuelta $N + 1$, o equivalentemente, al final de la vuelta N). El conductor A minimizará esta función de utilidad, mientras que B la maximizará. Además, la función de utilidad al final de la carrera debe tener en cuenta los casos en los que una o ambas variables m_{N+1}^A y m_{N+1}^B toman el valor cero. En este caso, el piloto correspondiente es descalificado. Entonces, la condición límite se puede resumir de la siguiente manera:

$$V_{N+1}^*(s_{N+1}) = \begin{cases} g_{N+1} & \text{si } m_{N+1}^A = 1 \text{ y } m_{N+1}^B = 1 \\ -\infty & \text{si } m_{N+1}^A = 1 \text{ y } m_{N+1}^B = 0 \\ +\infty & \text{si } m_{N+1}^A = 0 \text{ y } m_{N+1}^B = 1 \\ 0 & \text{en cualquier otro caso.} \end{cases} \quad (2.6)$$

Entonces, la ecuación de Bellman se puede expresar como:

$$V_n(s_n, x_n^A, x_n^B) = V_{n+1}^*(s_{n+1}), \quad (2.7)$$

donde, para facilitar la exposición, hemos omitido la dependencia explícita de los componentes futuros del estado (s_{n+1}) respecto al estado y las decisiones anteriores (s_n, x_n^A y x_n^B). Para $n = 2, \dots, N$

$$V_n^*(s_n) = \begin{cases} \min_{a \in \mathcal{T}_0} V_n(s_n, a, x_n^B(s_n, a)) & \text{si } g_n < 0 \\ \max_{b \in \mathcal{T}_0} V_n(s_n, x_n^A(s_n, b), b) & \text{si } g_n > 0, \end{cases} \quad (2.8)$$

donde

$$x_n^A(s_n, b) = \arg \min_{a \in \mathcal{T}_0} V_n(s_n, a, b), \quad (2.9)$$

$$x_n^B(s_n, a) = \arg \max_{b \in \mathcal{T}_0} V_n(s_n, a, b). \quad (2.10)$$

con el abuso de notación $x_n^A(s_n, \cdot)$ y $x_n^B(s_n, \cdot)$ para denotar la función de mejor respuesta

del conductor que actúa como seguidor en la vuelta n (A si $g_n > 0$, B en caso contrario) dado el estado s_n . Imponemos que no haya parada en pits en la primera vuelta (debido a la modelización de la entrada a pits) y en la última vuelta (ya que un piloto no puede terminar la carrera entrando en pits); es decir, $x_1^A = x_1^B = x_N^A = x_N^B = 0$. El estado al comienzo de la carrera es:

$$\begin{aligned} s_1 &= (t_1^A, w_1^A, m_1^A, t_1^B, w_1^B, m_1^B, g_1) \\ &= (t_1^A, 0, 0, t_1^B, 0, 0, g_1), \end{aligned} \tag{2.11}$$

donde t_1^A y t_1^B son los compuestos de neumáticos con los que comienzan la carrera, y g_1 es la cantidad de segundos en los que A aventaja a B cuando A cruza la línea de meta al comenzar la carrera¹². Asumimos que esta cantidad, g_1 , es conocida (ya que se puede estimar considerando la posición de salida en la parrilla de ambos pilotos). En cuanto a los compuestos iniciales de los neumáticos, estos son decisiones de ambos pilotos. A diferencia de las decisiones de parada en pits durante la carrera, los compuestos iniciales de los neumáticos se pueden modelar como un juego simultáneo. Esto se debe a que no hay motivo para pensar por qué un piloto elegiría su compuesto de neumáticos primero mientras que el otro lo observa y decide después.

En consecuencia, resolvemos el equilibrio del juego al comienzo de la carrera para encontrar los compuestos de neumáticos iniciales. Más precisamente, podemos encontrar el equilibrio en estrategias mixtas de este juego de suma-cero donde el espacio de acción de los jugadores es el conjunto finito \mathcal{T} . La matriz de pagos se denota por $U \in \mathbb{R}^{|\mathcal{T}| \times |\mathcal{T}|}$, de manera que $U_{t_1^A, t_1^B} := V_1^A((t_1^A, 0, 0, t_1^B, 0, 0, g_1))$ para $t_1^A, t_1^B \in \mathcal{T}$. En esta etapa, al comienzo de la carrera, el equilibrio puede resolverse utilizando el algoritmo de Lemke-Howson [24]. Sin embargo, dado que el tamaño del problema es pequeño y ambos jugadores tienen solo tres estrategias, podemos utilizar la Programación Lineal (PL). Aunque el algoritmo de Lemke-Howson es más eficiente, el tamaño del problema es lo suficientemente pequeño como para que no haya una diferencia práctica entre utilizar ese algoritmo o utilizar PL en este caso.

Denotemos por $\pi^A, \pi^B \in \Delta^{|\mathcal{T}|-1}$ las estrategias mixtas de ambos jugadores, respectivamente, donde Δ^n es la $n + 1$ dimensional simplex. Observemos que ambos jugadores resuelven el

¹²La posición de partida de los vehículos está detrás de la línea de meta. La *parrilla de salida* es el orden en el que los pilotos comenzarán la carrera.

siguiente juego minmax:

$$\begin{aligned}
& \max_{\pi^B} \min_{\pi^A} (\pi^B)^T U^T \pi^A \\
& \text{s.a.} \quad \sum_{i \in \mathcal{T}} \pi_i^B = 1 \\
& \quad \quad \pi_i^B \geq 0 \quad \forall i \in \mathcal{T} \\
& \quad \quad \sum_{i \in \mathcal{T}} \pi_i^A = 1 \\
& \quad \quad \pi_i^A \geq 0 \quad \forall i \in \mathcal{T}.
\end{aligned} \tag{2.12}$$

Entonces, podemos escribir el dual del problema de minimización del Problema de Optimización (2.12) como:

$$\begin{aligned}
& \max_{\rho} \rho \\
& \text{s.a.} \quad \rho - e_i^T U \pi^B \leq 0 \quad \forall i \in \mathcal{T},
\end{aligned} \tag{2.13}$$

donde $e_i \in \mathbb{R}^{|\mathcal{T}|}$ es un vector con un uno en la i ésima componente y cero en las demás. Combinando los Problemas de Optimización (2.12) y (2.13), obtenemos:

$$\begin{aligned}
& \max_{\rho, \pi^B} \rho \\
& \text{s.a.} \quad \rho - e_i^T U \pi^B \leq 0 \quad \forall i \in \mathcal{T} \\
& \quad \quad \sum_{i \in \mathcal{T}} \pi_i^B = 1 \\
& \quad \quad \pi_i^B \geq 0 \quad \forall i \in \mathcal{T}.
\end{aligned} \tag{2.14}$$

Una vez obtenemos la solución óptima del Problema de Optimización (2.14), π^{B*} , podemos utilizarla como estrategia del jugador B y resolver el Problema de Optimización (2.12) para A a fin de encontrar π^{A*} . De esta manera, encontramos un par de estrategias mixtas para ambos jugadores que constituye un equilibrio de Nash.

2.3.1. Función de Tiempo de Vuelta

La única parte que falta en el juego de Programación Dinámica es la definición de las funciones de tiempo por vuelta, μ_n^A y μ_n^B , que aparecen en el lado derecho de la Ecuación (2.5). A continuación, presentaremos la función de tiempo por vuelta del conductor A solamente, ya que la función para el conductor B es análoga. El tiempo por vuelta se expresa como una

función de tres elementos: (i) una *función de tiempo por vuelta fantasma*, que representa el tiempo por vuelta de un automóvil en la situación hipotética de que esté conduciendo solo (es decir, sin competencia); (ii) una *función de interacción* entre conductores, que incluye el tiempo ganado/perdido al estar cerca del automóvil oponente desde atrás o desde adelante; y (iii) una reducción de tiempo debido al uso del Sistema de Reducción de Arrastre (DRS, por sus siglas en inglés)¹³. Los conductores pueden activar el DRS cuando están a menos de un segundo del automóvil de adelante en las zonas de DRS selectivas, simplemente presionando un botón en el volante. Suponemos que esta zona de DRS está en la recta del circuito, paralela a los pits, ver Figura 2.1. Entonces, la función de tiempo por vuelta se puede escribir como:

$$\mu_n^A(s_n, x_n^A, x_n^B) = \gamma_n^A + \eta^A - t^{DRS} \cdot \mathbb{1}_{\{g_n \in [0,1]\}}, \quad (2.15)$$

donde cada término del lado derecho de la Ecuación (2.15) corresponde a (i), (ii) y (iii), respectivamente. En particular, γ_n^A denota la *función de tiempo por vuelta fantasma*, η^A es la *función de interacción* que representa el tiempo adicional por vuelta debido a la interacción entre los conductores, y t^{DRS} denota el parámetro que representa la reducción de tiempo debido al uso del DRS. Por simplicidad, omitimos la dependencia explícita de estas funciones en las variables de estado y decisión.

La *función de tiempo por vuelta fantasma* γ_n^A se puede escribir como:

$$\gamma_n^A = d_0^A + p_0^A \cdot \mathbb{1}_{\{x_n^A \neq 0\}} + \beta(t_{n+1}^A(t_n^A, x_n^A), w_n^A) - hn, \quad (2.16)$$

donde d_0^A es el tiempo base por vuelta de A asumiendo que está conduciendo con un juego nuevo de neumáticos blandos en la primera vuelta de la carrera (suponiendo que no hay interacción con otros autos). p_0^A representa el tiempo adicional relativo invertido en una parada en pits en comparación con un escenario en el que no se realiza una parada en pits. La función β captura la relación entre el compuesto de los neumáticos, su uso y el correspondiente aumento en el tiempo por vuelta en comparación con el uso de un juego nuevo de neumáticos blandos. Se ha utilizado la Ecuación (2.2) para obtener el compuesto real con el que el conductor correrá la n -ésima vuelta, ya que los neumáticos utilizados para

¹³ La tarea principal del DRS es reducir la resistencia al aire abriendo el ala trasera, lo que proporciona una velocidad adicional de 10-15 km/h en las rectas y en ciertos sectores de la carrera (también llamados zonas de DRS)

la n -ésima vuelta dependerán de los neumáticos actuales al comienzo de esta vuelta (t_n^A) y de la decisión de parada en pits (x_n^A). Finalmente, el parámetro h representa la reducción en el tiempo por vuelta por cada vuelta adicional que se completa, manteniendo todo lo demás igual. Esto tiene en cuenta el hecho de que el combustible se consume durante la carrera. Por lo tanto, el peso del automóvil disminuye a medida que se corren más vueltas y, en consecuencia, los tiempos por vuelta también. Suponemos que este efecto es lineal, aunque también podríamos utilizar formas funcionales más elaboradas para modelar este efecto.

La *función de interacción* η^A modela el tiempo adicional por vuelta para el auto A debido a la interacción entre los autos desde la salida de pits hasta el final de la vuelta. Es importante destacar que aunque puede haber una gran diferencia de tiempo entre los autos al comienzo de la vuelta n (es decir, $|g_n|$ es grande), esto puede disminuir si el auto líder hace una parada en pits mientras que el seguidor continúa en pista. Por lo tanto, los autos podrían terminar muy cerca uno del otro después de la salida de pits a pesar de estar alejados al comienzo de la vuelta. Usamos ϵ_n para denotar la diferencia de tiempo entre los autos en la vuelta n , de manera similar a la variable de estado g_n , pero en la salida de pits. Esto se puede expresar como:

$$\epsilon_n = g_n + p_0^A \cdot \mathbf{1}_{\{x_n^A \neq 0\}} - t^{DRS} \cdot \mathbf{1}_{\{g_n \in [0,1]\}} - \left(p_0^B \cdot \mathbf{1}_{\{x_n^B \neq 0\}} - t^{DRS} \cdot \mathbf{1}_{\{g_n \in [-1,0]\}} \right). \quad (2.17)$$

Entonces, la *función de interacción* se va a expresar como una función de ϵ_n . Más precisamente, se puede escribir de la siguiente manera:

$$\eta^A(\epsilon_n) = \begin{cases} \exp(-\lambda \cdot |\epsilon_n|) \cdot z^{(1)} & \text{si } \epsilon_n < 0 \\ \exp(-\lambda \cdot |\epsilon_n|) \cdot z^{(2)} & \text{en cualquier otro caso,} \end{cases} \quad (2.18)$$

donde $\lambda, z^{(1)}, z^{(2)} > 0$ son parámetros de la *función de interacción* para el auto A . El concepto detrás de incorporar diferentes parámetros, $z^{(1)}$ y $z^{(2)}$, dependiendo de la posición del oponente con respecto al auto (si está adelante o detrás) es tener en cuenta la posible variación en la pérdida de tiempo. Ser seguido de cerca por un auto puede tener un impacto diferente que tener la misma diferencia de tiempo pero estar adelante del oponente. Además, se utiliza la función exponencial para garantizar que el efecto de interacción disminuya a medida que la brecha entre los autos aumenta. La *función de interacción* para el auto B , η^B , es la misma

que la Ec. (2.18), excepto que el caso de la primera condición debe cambiarse a $\epsilon_n > 0$.

2.3.2. Algoritmo

A continuación se presenta un resumen de los pasos para resolver el juego en el Algoritmo 1. En particular, primero resolvemos el juego de Stackelberg resolviendo la Programación Dinámica para todas las vueltas de la carrera en orden inverso para ambos jugadores, y luego resolvemos el juego simultáneo antes del inicio de la carrera para obtener el equilibrio en estrategias mixtas en relación con los compuestos de neumáticos con los que comenzar la carrera. Denotamos por \mathcal{S}_n el conjunto de todos los estados en la vuelta n .

Algorithm 1 Dynamic Programming Algorithm

- 1: **Input:** Set of states $\{\mathcal{S}_n\}_{n \in \{1, \dots, N+1\}}$. Starting time gap g_1
 - 2: **Output:** Decisions $\{x_n^{A*}(s_n)\}_{s_n \in \mathcal{S}_n, n \in \{1, \dots, N\}}$, $\{x_n^{B*}(s_n)\}_{s_n \in \mathcal{S}_n, n \in \{1, \dots, N\}}$, π^{A*} and π^{B*} .
 - 3: **Foreach** $s_n \in \mathcal{S}_{N+1}$
 - 4: Compute $V_{N+1}^*(s_n)$ using Equation (2.6).
 - 5: **End For**
 - 6: **Set** $n \leftarrow N$
 - 7: **While** $n \geq 1$ **do**
 - 8: **Foreach** $s_n \in \mathcal{S}_n$
 - 9: Compute $V_n(s_n, a, b)$ from Equation (2.7) for all $a, b \in \mathcal{T}_0$.
 - 10: **If** $g_n < 0$ **Then**
 - 11: Compute $x_n^B(s_n, a)$ with Equation (2.10) for all $a \in \mathcal{T}_0$.
 - 12: $x_n^A(s_n) \leftarrow \arg \max_{a \in \mathcal{T}_0} V_n(s_n, a, x_n^B(s_n, a))$.
 - 13: $V_n^*(s_n) \leftarrow V_n(s_n, x_n^A(s_n), x_n^B(s_n, x_n^A(s_n)))$.
 - 14: **Else**
 - 15: Compute $x_n^A(s_n, b)$ with Equation (2.9) for all $b \in \mathcal{T}_0$.
 - 16: $x_n^B(s_n) \leftarrow \arg \max_{b \in \mathcal{T}_0} V_n(s_n, x_n^A(s_n, b), b)$.
 - 17: $V_n^*(s_n) \leftarrow V_n(s_n, x_n^A(s_n, x_n^B(s_n)), x_n^B(s_n))$.
 - 18: **End If**
 - 19: **End For**
 - 20: $n \leftarrow n - 1$
 - 21: **End While**
 - 22: Compute $U \in \mathbb{R}^{|\mathcal{T}| \times |\mathcal{T}|}$ such that $U_{t_1^A, t_1^B} = V_1^A((t_1^A, 0, 0, t_1^B, 0, 0, g_1))$ for $t_1^A, t_1^B \in \mathcal{T}$
 - 23: Get π^{A*} and π^{B*} by solving Optimization Problems (2.14) and (2.12).
-

2.4. Modelo Estocástico

En la sección anterior se introdujo un modelo determinista, como resultado, el resultado de la carrera está determinado de manera única por las acciones de los jugadores. En realidad,

existen varios eventos inciertos que podrían desarrollarse durante la carrera, como banderas amarillas o tiempos de vuelta. En esta sección, presentamos una formulación de programación dinámica para el juego que tiene en cuenta estos elementos. La sección se divide en tres partes. (i) Primero, introducimos la modelización de las banderas amarillas en el estado del modelo y en las variables aleatorias. (ii) Se proporcionan detalles sobre el cálculo de los tiempos de vuelta. En ambas secciones, (i) y (ii), definimos las variables aleatorias del problema. (iii) Mostramos los elementos de la programación dinámica estocástica para resolver el juego.

2.4.1. Banderas Amarillas

Los dos eventos de bandera amarilla que consideramos son el VSC y el SC. (Utilizamos la expresión *Coche de Seguridad*, es decir, sin acrónimo, para referirnos explícitamente al automóvil que se despliega en la pista para hacer cumplir la reducción de velocidad de los pilotos). En ambos casos, hay un incidente en la carrera que implica limpiar la pista de escombros. En el VSC, se obliga a los pilotos a reducir su velocidad en un 40% (ver Biswas [25]), mientras que en el SC, los pilotos deben seguir al Coche de Seguridad desplegado en la pista (en este escenario, los autos se agrupan detrás del Coche de Seguridad y se les exige mantener su posición, por lo tanto, las diferencias de tiempo y las brechas se reducen). Sea Y_n^{VSC} y Y_n^{SC} las variables aleatorias en la vuelta n que indican las vueltas restantes bajo VSC y SC, respectivamente. Para simplificar, asumimos que el incidente de bandera amarilla ocurre al final de la vuelta. Suponemos que un incidente de VSC (SC) dura l^{VSC} (l^{SC}) vueltas. Además, el DRS se habilita después de k^{VSC} (k^{SC}) vueltas después del final de un VSC (SC). Antes de definir la ley de probabilidad de estas variables aleatorias, vamos a definir las variables de estado.

Para determinar si la vuelta actual se está corriendo bajo un evento de bandera amarilla, consideramos dos componentes adicionales en las variables de estado (además de las introducidas en la DP de la Sección 2.3) que tienen en cuenta este hecho. En particular, para cada vuelta n , sea y_n^{VSC} (y_n^{SC}) el número de vueltas restantes bajo VSC (SC) sin considerar la vuelta n .

Por la definición de las variables aleatorias (Y_n^{VSC}, Y_n^{SC}) y las componentes de las variables de estado (y_n^{VSC}, y_n^{SC}), es claro que $y_{n+1}^{VSC} = Y_n^{VSC}$ y $y_{n+1}^{SC} = Y_n^{SC}$. Luego, la probabilidad de

un evento de bandera amarilla se define de la siguiente manera:

$$\begin{aligned}
\mathbb{P}\left((Y_n^{VSC}, Y_n^{SC}) = (i, 0) \mid (y_n^{VSC}, y_n^{SC}) = (i + 1, 0)\right) &= 1 \quad \forall i \in \{1, \dots, l^{VSC}\} \\
\mathbb{P}\left((Y_n^{VSC}, Y_n^{SC}) = (0, i) \mid (y_n^{VSC}, y_n^{SC}) = (0, i + 1)\right) &= 1 \quad \forall i \in \{1, \dots, l^{SC}\} \\
\mathbb{P}\left((Y_n^{VSC}, Y_n^{SC}) = (l^{VSC}, 0) \mid (y_n^{VSC}, y_n^{SC}) = a\right) &= q^{VSC} \quad \forall a \in \{(0, 0), (1, 0), (0, 1)\} \\
\mathbb{P}\left((Y_n^{VSC}, Y_n^{SC}) = (0, l^{SC}) \mid (y_n^{VSC}, y_n^{SC}) = a\right) &= q^{SC} \quad \forall a \in \{(0, 0), (1, 0), (0, 1)\},
\end{aligned}$$

y PMF 0 para los demás casos. Es decir, si quedan j vueltas con un VSC, con $j > 1$, entonces con probabilidad uno en la siguiente vuelta quedará $j - 1$ vueltas restantes bajo el mismo evento de VSC¹⁴. Sin embargo, si no hay bandera amarilla (es decir, $(y_n^{VSC} + y_n^{SC} = 0)$, o una bandera amarilla está terminando (es decir, $(y_n^{VSC} + y_n^{SC} = 1)$, entonces con probabilidad q^{VSC} (q^{SC}) ocurre un evento de VSC (SC) al final de la vuelta n . Nótese que Y_n^{VSC} y Y_n^{SC} no son independientes, ya que no puede haber VSC y SC simultáneamente. Para mayor comodidad en la notación, denotaremos por r^{VSC} (r^{SC}) la probabilidad de tener al menos un VSC (SC) durante la carrera. Así, $r^s = 1 - (1 - q^s)^{1/N}$ para $s \in \{VSC, SC\}$.

Además de los dos componentes de las variables de estado y_n^{VSC} y y_n^{SC} , se considera un tercer componente de estado para el escenario estocástico. Como se mencionó anteriormente, después de un evento de bandera amarilla, en particular bajo un SC, las brechas de tiempo entre los autos se reducen considerablemente. Como resultado, el DRS se desactiva durante la bandera amarilla y se activa después de un par de vueltas después de que la bandera amarilla finaliza (esto es para asegurar que el DRS no se active cuando los autos estén muy cerca uno del otro). Además, por razones similares, el DRS se desactiva durante las dos primeras vueltas de una carrera. Sea y_n^{DRS} un componente de variable de estado que indica las vueltas restantes en las que el DRS estará activado, tomando el valor 0 si el DRS está activado en la vuelta actual. Entonces, la variable de estado en la vuelta n , s_n , está conformada por:

$$s_n = (t_n^A, w_n^A, m_n^A, t_n^B, w_n^B, m_n^B, g_n, y_n^{VSC}, y_n^{SC}, y_n^{DRS}). \quad (2.19)$$

El impacto de las banderas amarillas en los tiempos de vuelta se explica en la siguiente sección.

¹⁴ Nótese que por "vueltas restantes bajo VSC" (y de manera similar SC), nos referimos al número de vueltas restantes bajo VSC asumiendo que no habrá más eventos de bandera amarilla en el futuro.

2.4.2. Tiempos de Vuelta

Las banderas amarillas afectan la velocidad de los vehículos y, por lo tanto, sus tiempos de vuelta. Proporcionamos expresiones de tiempo de vuelta para los siguientes tres escenarios: (i) estar bajo un VSC, (ii) estar bajo un SC y (iii) no estar bajo una bandera amarilla. Denotamos las funciones de tiempo de vuelta para estos tres escenarios como $\mu_{n,VSC}^i$, $\mu_{n,SC}^i$ y $\mu_{n,0}^i$ para cada piloto $i \in \{A, B\}$, respectivamente.

2.4.2.1. Tiempos de Vuelta bajo VSC

Cada piloto debe reducir su velocidad, no hay interacción entre ellos y no se permite adelantar en pista (aunque podría haber adelantamientos debido a las paradas en pits). Por lo tanto, los tiempos de vuelta en este caso se pueden expresar como:

$$\mu_{n,VSC}^i = d_{VSC} + p_{VSC}^i \cdot \mathbb{1}_{\{x_n^i \neq 0\}}, \quad (2.20)$$

con $i \in \{A, B\}$, donde d_{VSC} es el tiempo de vuelta de un piloto durante un VSC cuando no hay paradas en pits, y p_{VSC}^i es la pérdida de tiempo adicional del piloto i debido a hacer una parada en pits durante un VSC.

2.4.2.2. Tiempos de Vuelta bajo SC

Los autos se agrupan detrás del Coche de Seguridad. Más específicamente, un Coche de Seguridad es desplegado después de la salida de pits. Asumimos que la diferencia de tiempo entre los pilotos se reduce a 0.5 segundos. De forma más formal, consideremos la función ξ que toma el valor 1 si el piloto A termina por delante de B después de la salida de pits, y 0 en caso contrario. Esto es,

$$\xi(g_n, x_n^A, x_n^B) = \mathbb{1} \left\{ g_n + p_{SC}^A \cdot \mathbb{1}_{\{x_n^A \neq 0\}} - p_{SC}^B \cdot \mathbb{1}_{\{x_n^B \neq 0\}} < 0 \right\}, \quad (2.21)$$

donde p_{SC}^i denota el tiempo adicional perdido por el piloto $i \in \{A, B\}$ debido a una parada en pits mientras se encuentra bajo un SC. Entonces, bajo un SC, la variable de estado g_{n+1} se puede escribir de la siguiente manera:

$$g_{n+1} = 0.5 \cdot (-1)^{\xi(g_n, x_n^A, x_n^B)}. \quad (2.22)$$

La ecuación (2.22) indica que mientras el Coche de Seguridad esté en pista, uno de los pilotos, ya sea A o B , estará liderando la carrera por 0.5 segundos. Como resultado de la regla de no adelantar en pista durante una bandera amarilla, la única forma en que se produce un adelantamiento es debido a una parada en pits. Esto ocurriría si $g_n < 0$ y $\xi(g_n, x_n^A, x_n^B) = 0$ (por lo tanto, $g_{n+1} = 0.5$), o si $g_n > 0$ y $\xi(g_n, x_n^A, x_n^B) = 1$ (por lo tanto, $g_{n+1} = -0.5$). En cuanto a los tiempos de vuelta bajo un SC, el tiempo de vuelta del primer piloto siguiendo al Coche de Seguridad coincidirá con su tiempo de vuelta, denotado como d_{SC} . En cuanto al seguidor, su tiempo de vuelta también coincidirá con este tiempo, excepto en la primera vuelta del SC; en este caso, el piloto alcanzará al líder de la carrera reduciendo la diferencia de tiempo a medio segundo. Matemáticamente,

$$\mu_{n,SC}^A = \begin{cases} d_{SC} + p_{SC}^A \cdot \mathbb{1}_{\{x_n^A \neq 0\}} & \text{si } \xi(g_n, x_n^A, x_n^B) = 1 \\ d_{SC} + p_{SC}^B \cdot \mathbb{1}_{\{x_n^B \neq 0\}} - g_n + 0.5 & \text{en cualquier otro caso,} \end{cases} \quad (2.23)$$

$$\mu_{n,SC}^B = \begin{cases} d_{SC} + p_{SC}^A \cdot \mathbb{1}_{\{x_n^A \neq 0\}} + g_n + 0.5 & \text{si } \xi(g_n, x_n^A, x_n^B) = 1 \\ d_{SC} + p_{SC}^B \cdot \mathbb{1}_{\{x_n^B \neq 0\}} & \text{en cualquier otro caso.} \end{cases} \quad (2.24)$$

Podemos observar que, a partir de las ecuaciones (2.23) y (2.24), el tiempo de vuelta del piloto líder es d_{SC} más el tiempo adicional de la parada en pits, mientras que la expresión del tiempo de vuelta del seguidor (para el piloto A en el caso en que $\xi(g_n, x_n^A, x_n^B) = 1$ en la ecuación (2.23), y para el piloto B en el caso en que $\xi(g_n, x_n^A, x_n^B) = 0$ en la ecuación (2.24)) es tal que $g_{n+1} = g_n + \mu_{n,SC}^A - \mu_{n,SC}^B$.

Por ejemplo, si ocurre un SC que dura 3 vueltas al final de la vuelta $n = 15$ (es decir, $Y_{15}^{SC} = 3$), entonces g_{17} , g_{18} y g_{19} serán ± 0.5 , mientras que g_{16} tomará el valor de la diferencia de tiempo que tienen los autos cuando acaba de ocurrir el SC.

2.4.2.3. Tiempos de Vuelta sin Bandera Amarilla

Consideramos la incertidumbre en los tiempos de vuelta utilizando variables aleatorias separadas para las dos partes de la vuelta: entre el inicio de la vuelta y la salida de la parada en pits, y entre la salida de la parada en pits y el final de la vuelta. Para la primera parte, utilizamos la variable aleatoria T_n^{DRS} para denotar la reducción en el tiempo de vuelta del seguidor debido al uso del DRS durante la vuelta n . Definimos una única variable aleatoria para este efecto, en lugar de una por cada vehículo, ya que a lo sumo habrá un piloto utilizando

el DRS. Suponemos que T_n^{DRS} sigue la distribución $\mathcal{F}_{T^{DRS}}$. Para el segundo segmento de la vuelta, definimos dos variables aleatorias $Z_n^{(1)}$ y $Z_n^{(2)}$ que representan la incertidumbre en la interacción entre los pilotos. Denotemos por $\mathcal{F}_{Z^{(1)}}$ y $\mathcal{F}_{Z^{(2)}}$ la distribución de estas variables aleatorias. (Más detalles sobre estas variables aleatorias se proporcionarán más adelante). La razón de utilizar variables aleatorias separadas para los dos segmentos de la vuelta considerados es que la diferencia de tiempo entre los pilotos puede ser muy diferente entre ellos (ver la discusión antes de la Ecuación (2.17)), y por lo tanto, la dinámica entre los vehículos también es diferente. Solo mostraremos las funciones de tiempo de vuelta para el piloto A , ya que las correspondientes al piloto B son análogas. El tiempo de vuelta para el piloto A se expresa como:

$$\mu_{n,0}^A \left(s_n, x_n^A, x_n^B, T_n^{DRS}, Z_n^{(1)}, Z_n^{(2)} \right) = \gamma_n^A + \eta_n'^A(\epsilon'_n) - T_n^{DRS} \cdot \mathbb{1}_{\{g_n \in [0,1], y_n^{DRS}=0\}}, \quad (2.25)$$

donde γ^A denota la *función de tiempo de vuelta fantasma*, que es la misma que se definió para el caso determinista en la Ecuación (2.16). De manera similar a antes, ϵ'_n denota la diferencia de tiempo entre los pilotos después de la salida de la parada en pits (en el caso de que no haya bandera amarilla). Esto se calcula como:

$$\epsilon'_n = g_n + p_0^A \cdot \mathbb{1}_{\{x_n^A \neq 0\}} - T_n^{DRS} \cdot \mathbb{1}_{\{g_n \in [0,1], y^{DRS}=0\}} - \left(p_0^B \cdot \mathbb{1}_{\{x_n^B \neq 0\}} - T_n^{DRS} \cdot \mathbb{1}_{\{g_n \in [-1,0], y^{DRS}=0\}} \right). \quad (2.26)$$

Finalmente, la *función de interacción* en el caso estocástico para el piloto A , $\eta_n'^A$, se expresa como la siguiente función estocástica:

$$\eta_n'^A(\epsilon'_n) = \begin{cases} \exp(-\lambda \cdot |\epsilon'_n|) \cdot Z_n^{(1)} & \text{si } \epsilon'_n < 0 \\ \exp(-\lambda \cdot |\epsilon'_n|) \cdot Z_n^{(2)} & \text{en cualquier otro caso.} \end{cases} \quad (2.27)$$

Para el caso del conductor B , la *función de interacción* $\eta_n'^B$ se define de la misma manera que la Ecuación (2.27), excepto que la condición en el lado derecho se cambia entre los dos casos.

Juntando las Ecuaciones (2.20), (2.23), (2.24) y (2.25) de los tres casos, podemos escribir

la función de tiempo por vuelta para el conductor $i \in \{A, B\}$ como sigue:

$$\mu_n^i(s_n, x_n^A, x_n^B, T_n^{DRS}, Z_n^{(1)}, Z_n^{(2)}) = \begin{cases} \mu_{n,0}^i & \text{si } y_n^{VSC} + y_n^{SC} = 0 \\ \mu_{n,VSC}^i & \text{si } y_n^{VSC} > 0 \\ \mu_{n,SC}^i & \text{si } y_n^{SC} > 0, \end{cases} \quad (2.28)$$

donde los argumentos de las funciones en el lado derecho se han omitido por simplicidad. El modelo también se puede extender para considerar otros aspectos de las carreras, por ejemplo, podemos imponer una brecha de tiempo mínima, $\Delta > 0$, entre los conductores en una sección del circuito. Esta extensión se describe en el Apéndice B.

2.4.3. Programación Dinámica Estocástica

Recuerde las componentes adicionales de las variables de estado agregadas (ver Sección 2.4.1). Además, para mayor facilidad de notación, vamos a definir el vector aleatorio $W_n = (Y_n^{VSC}, Y_n^{SC}, T_n^{DRS}, Z_n^{(1)}, Z_n^{(2)})$. Entonces, la actualización de la variable de estado (s_{n+1}) en función del estado anterior (s_n), las decisiones (x_n^A y x_n^B), y el vector aleatorio (W_n), puede expresarse como:

$$t_{n+1}^A = t_n^A \cdot \mathbb{1}_{\{x_n^A=0\}} + x_n^A \cdot \mathbb{1}_{\{x_n^A \neq 0\}} \quad (2.29)$$

$$w_{n+1}^A = \min\{w_n^A \cdot \mathbb{1}_{\{x_n^A=0\}} + \mathbb{1}_{\{Y_n^{VSC}+Y_n^{SC}=0\}}, u_{t_n} + 1\} \quad (2.30)$$

$$m_{n+1}^A = \max\{m_n^A, \mathbb{1}_{\{x_n^A \neq 0, x_n^A \neq t_n^A\}}\} \quad (2.31)$$

$$g_{n+1} = g_n + \mu_n^A(s_n, x_n^A, x_n^B, W_n) - \mu_n^B(s_n, x_n^A, x_n^B, W_n) \quad (2.32)$$

$$y_{n+1}^{VSC} = Y_n^{VSC} \quad (2.33)$$

$$y_{n+1}^{SC} = Y_n^{SC} \quad (2.34)$$

$$y_{n+1}^{DRS} = \max\{y_n^{DRS} - 1, 0\} \cdot \mathbb{1}_{\{Y_n^{VSC}+Y_n^{SC}=0\}} + \sum_{s \in \{VSC, SC\}} (Y_n^s + k^s) \cdot \mathbb{1}_{\{Y_n^s > 0\}}, \quad (2.35)$$

y de manera similar para t_{n+1}^B , w_{n+1}^B y m_{n+1}^B . Tenga en cuenta que durante un evento de bandera amarilla, los neumáticos no se desgastan, como se muestra en la Ecuación (2.30). Esto se debe a la reducción significativa de la velocidad durante estos eventos.

El objetivo de cada conductor es maximizar la diferencia de tiempo con respecto al otro conductor al final de la carrera. Definamos el evento $H^i = \{w_{N+1}^i \leq u_{t_{i+1}}\} \wedge \{m_{N+1}^i = 1\}$

para $i \in \{A, B\}$. En consecuencia, la función de costo en la frontera puede escribirse como:

$$V_{N+1}^*(s_{N+1}) = \begin{cases} g_{N+1} & \text{si } H^A \text{ y } H^B \\ -\infty & \text{si } H^A \text{ y no } H^B \\ +\infty & \text{si no } H^A \text{ y } H^B \\ 0 & \text{en cualquier otro caso.} \end{cases} \quad (2.36)$$

Entonces, la ecuación de Bellman se puede expresar como:

$$V'_n(s_n, x_n^A, x_n^B) = \mathbb{E}_{W_n} [V_{n+1}^*(s_{n+1})]. \quad (2.37)$$

Para $n = 1, \dots, N - 1$

$$V_n^*(s_n) = \begin{cases} \min_{a \in \mathcal{T}_0} V'_n(s_n, a, x_n^B(s_n, a)) & \text{if } g_n < 0 \\ \max_{b \in \mathcal{T}_0} V'_n(s_n, x_n^A(s_n, b), b) & \text{if } g_n > 0 \end{cases}, \quad (2.38)$$

donde

$$x_n^A(s_n, b) = \arg \min_{a \in \mathcal{T}_0} V'_n(s_n, a, b), \quad (2.39)$$

$$x_n^B(s_n, a) = \arg \max_{b \in \mathcal{T}_0} V'_n(s_n, a, b). \quad (2.40)$$

Como en el caso determinista, resolvemos el juego simultáneo para la primera etapa, en la cual ambos conductores deciden el componente de neumáticos con el que iniciar la carrera. Para esto, definimos la matriz $U' \in \mathbb{R}^{|\mathcal{T}| \times |\mathcal{T}|}$ tal que $U'_{t_1^A, t_1^B} = V_1^*((t_1^A, 0, 0, t_1^B, 0, 0, g_1, 0, 0, 2))$, para $t_1^A, t_1^B \in \mathcal{T}$. Observa que estamos considerando que $y_1^{DRS} = 2$, lo cual significa que el DRS se habilitará después de dos vueltas más (en caso de no haber eventos de bandera amarilla). Luego, las estrategias mixtas para la primera etapa se encuentran resolviendo los problemas de optimización (2.12) y (2.14), pero con U' en lugar de U .

Denotamos por \mathcal{S}_n al conjunto de todos los estados en la vuelta n . Es importante tener en cuenta que, dado que el líder en cada etapa depende del estado, una estrategia de perfil de un jugador puede ser: (i) una función que depende únicamente del estado, en cuyo caso el jugador es el líder, o (ii) una función de respuesta que depende del estado y de la estrategia del otro jugador, en cuyo caso el jugador corresponde al seguidor. Por lo tanto, una estrategia

de perfil para los jugadores A y B se puede escribir como:

$$\begin{aligned} & \{x_n^{IA}(s_n)\}_{s_n \in \mathcal{S}_n^-, n \in \{1, \dots, N\}} \cup \{x_n^{IA}(s_n, b)\}_{s_n \in \mathcal{S}_n^+, b \in \mathcal{T}_0, n \in \{1, \dots, N\}} \cup \pi^{IA} \\ & \{x_n^{IB}(s_n)\}_{s_n \in \mathcal{S}_n^+, n \in \{1, \dots, N\}} \cup \{x_n^{IB}(s_n, a)\}_{s_n \in \mathcal{S}_n^-, a \in \mathcal{T}_0, n \in \{1, \dots, N\}} \cup \pi^{IB}, \end{aligned} \quad (2.41)$$

donde \mathcal{S}_n^- (\mathcal{S}_n^+) es el conjunto de estados de la vuelta n donde $g_n < 0$ (> 0). Es decir, $\mathcal{S}_n^- = \{s_n \in \mathcal{S}_n | g_n < 0\}$ y $\mathcal{S}_n^+ = \{s_n \in \mathcal{S}_n | g_n > 0\}$.

En el Algoritmo 2 se proporciona un resumen de la Programación Dinámica Estocástica.

Algorithm 2 Stochastic Dynamic Programming Algorithm

- 1: **Input:** Set of states $\{\mathcal{S}_n\}_{n \in \{1, \dots, N+1\}}$. Starting time gap g_1
 - 2: **Output:** Decisions $\{x_n^{A*}(s_n)\}_{s_n \in \mathcal{S}_n^-, n \in \{1, \dots, N\}} \cup \{x_n^{A*}(s_n, b)\}_{s_n \in \mathcal{S}_n^+, b \in \mathcal{T}_0, n \in \{1, \dots, N\}}$, $\{x_n^{B*}(s_n)\}_{s_n \in \mathcal{S}_n^+, n \in \{1, \dots, N\}} \cup \{x_n^{B*}(s_n, a)\}_{s_n \in \mathcal{S}_n^-, a \in \mathcal{T}_0, n \in \{1, \dots, N\}}$, π^{A*} and π^{B*} .
 - 3: **Foreach** $s_n \in \mathcal{S}_{N+1}$
 - 4: Compute $V_{N+1}^*(s_n)$ using Equation (2.36).
 - 5: **End For**
 - 6: **Set** $n \leftarrow N$
 - 7: **While** $n \geq 1$ **do**
 - 8: **Foreach** $s_n \in \mathcal{S}_n$
 - 9: Compute $V_n'(s_n, a, b)$ from Equation (2.37) for all $a, b \in \mathcal{T}_0$.
 - 10: **If** $g_n < 0$ **Then**
 - 11: Compute $x_n^{B*}(s_n, a)$ with Equation (2.40) for all $a \in \mathcal{T}_0$.
 - 12: $x_n^{A*}(s_n) \leftarrow \arg \max_{a \in \mathcal{T}_0} V_n'(s_n, a, x_n^{B*}(s_n, a))$.
 - 13: $V_n^*(s_n) \leftarrow V_n'(s_n, x_n^{A*}(s_n), x_n^{B*}(s_n, x_n^{A*}(s_n)))$.
 - 14: **Else**
 - 15: Compute $x_n^{A*}(s_n, b)$ with Equation (2.39) for all $b \in \mathcal{T}_0$.
 - 16: $x_n^{B*}(s_n) \leftarrow \arg \max_{b \in \mathcal{T}_0} V_n'(s_n, x_n^{A*}(s_n, b), b)$.
 - 17: $V_n^*(s_n) \leftarrow V_n'(s_n, x_n^{A*}(s_n, x_n^{B*}(s_n)), x_n^{B*}(s_n))$.
 - 18: **End If**
 - 19: **End For**
 - 20: $n \leftarrow n - 1$
 - 21: **End While**
 - 22: Compute $U' \in \mathbb{R}^{|\mathcal{T}| \times |\mathcal{T}|}$ such that $U'_{t_1^A, t_1^B} = V_1^*((t_1^A, 0, 0, t_1^B, 0, 0, g_1, 0, 0, 2))$ for $t_1^A, t_1^B \in \mathcal{T}$
 - 23: Get π^{A*} and π^{B*} by solving Optimization Problems (2.14) and (2.12) with U' instead of U .
-

Definimos un equilibrio del juego de Stackelberg de retroalimentación considerando compuestos de neumáticos iniciales fijos $t_1^A, t_1^B \in \mathcal{T}$.

Definición 2.4.1. *Un equilibrio de retroalimentación Stackelberg, dado un estado inicial s_1 , es un par de estrategias $(\{x_n^{A*}(s_n)\}_{s_n \in \mathcal{S}_n^-, n \in \{1, \dots, N\}} \cup \{x_n^{A*}(s_n, b)\}_{s_n \in \mathcal{S}_n^+, b \in \mathcal{T}_0, n \in \{1, \dots, N\}})$ y*

$\left(\{x_n^{B^*}(s_n)\}_{s_n \in \mathcal{S}_n^+, n \in \{1, \dots, N\}} \cup \{x_n^{B^*}(s_n, a)\}_{s_n \in \mathcal{S}_n^-, a \in \mathcal{T}_0, n \in \{1, \dots, N\}}\right)$ tal que para todo $n \in \{1, \dots, N\}$:

1. Si $g_n < 0$

$$V'_n(s_n, x_n^{A^*}(s_n), x_n^{B^*}(s_n, x_n^{A^*}(s_n))) \leq V'_n(s_n, x_n^{A^*}(s_n), x_n^{B^*}(s_n, x_n^{A^*}(s_n))) \quad (2.42)$$

$$V'_n(s_n, x_n^{A^*}(s_n), x_n^{B^*}(s_n, x_n^{A^*}(s_n))) \geq V'_n(s_n, x_n^{A^*}(s_n), x_n^{B^*}(s_n, x_n^{A^*}(s_n))). \quad (2.43)$$

2. Si $g_n > 0$

$$V'_n(s_n, x_n^{A^*}(s_n, x_n^{A^*}(s_n)), x_n^{B^*}(s_n)) \leq V'_n(s_n, x_n^{A^*}(s_n, x_n^{A^*}(s_n)), x_n^{B^*}(s_n)) \quad (2.44)$$

$$V'_n(s_n, x_n^{A^*}(s_n, x_n^{A^*}(s_n)), x_n^{B^*}(s_n)) \geq V'_n(s_n, x_n^{A^*}(s_n, x_n^{A^*}(s_n)), x_n^{B^*}(s_n)), \quad (2.45)$$

para todo $\left(\{x_n^{A^*}(s_n)\}_{s_n \in \mathcal{S}_n^-, n \in \{1, \dots, N\}} \cup \{x_n^{A^*}(s_n, b)\}_{s_n \in \mathcal{S}_n^+, b \in \mathcal{T}_0, n \in \{1, \dots, N\}}\right)$ y $\left(\{x_n^{B^*}(s_n)\}_{s_n \in \mathcal{S}_n^+, n \in \{1, \dots, N\}} \cup \{x_n^{B^*}(s_n, a)\}_{s_n \in \mathcal{S}_n^-, a \in \mathcal{T}_0, n \in \{1, \dots, N\}}\right)$.

A continuación, demostramos que la solución del Algoritmo 2 para cada estado de la primera vuelta, es decir, para todas las posibles combinaciones de compuestos de neumáticos de inicio, es un equilibrio de retroalimentación Stackelberg (*feedback Stackelberg equilibrium*).

Teorema 2.4.2. *Para cada par de compuestos de neumáticos de inicio $(t_1^A, t_1^B) \in \mathcal{T}^2$, la salida del Algoritmo 2 da como resultado un equilibrio de retroalimentación Stackelberg.*

DEMOSTRACIÓN. Vea Apéndice C. □

A continuación, definimos el concepto de equilibrio para todo el juego, es decir, incluyendo las decisiones simultáneas sobre los compuestos de neumáticos iniciales. Utilizamos el término *Nash-Stackelberg de retroalimentación* para indicar que la primera decisión es simultánea, mientras que todas las decisiones futuras se toman en modo de retroalimentación Stackelberg.

Definición 2.4.3. *Un equilibrio Nash-feedback Stackelberg es un par de estrategias*

$$\left(\{x_n^{A^*}(s_n)\}_{s_n \in \mathcal{S}_n^-, n \in \{1, \dots, N\}} \cup \{x_n^{A^*}(s_n, b)\}_{s_n \in \mathcal{S}_n^+, b \in \mathcal{T}_0, n \in \{1, \dots, N\}}\right) \cup \pi^{A^*} \text{ y}$$

$$\left(\{x_n^{B^*}(s_n)\}_{s_n \in \mathcal{S}_n^+, n \in \{1, \dots, N\}} \cup \{x_n^{B^*}(s_n, a)\}_{s_n \in \mathcal{S}_n^-, a \in \mathcal{T}_0, n \in \{1, \dots, N\}}\right) \cup \pi^{B^*} \text{ si:}$$

1. *Las estrategias sin la estrategia del componente de neumáticos inicial forman un equilibrio de Stackelberg con retroalimentación para todos los $(t_1^A, t_1^B) \in \mathcal{T}^2$.*

2. La estrategia del componente de neumáticos inicial de ambos jugadores, es decir, (π^{A*}, π^{B*}) , es un equilibrio mixto de Nash del juego simultáneo de suma-cero con espacio de acción \mathcal{T} para ambos jugadores y función de pago $-V_1((t_1, 0, 0, t_2, 0, 0, g_1, 0, 0, 2))$ para el jugador A (pago en el sentido de que A maximiza esta función).

Teorema 2.4.4. *La solución obtenida del Algoritmo 2 es un equilibrio de Nash-Stackelberg con retroalimentación.*

DEMOSTRACIÓN. La demostración sigue directamente de la línea 22 del Algoritmo 2, y utilizando el Teorema 2.4.2 para mostrar la primera condición de la Definición 2.4.3. \square

2.4.4. Maximizar la Probabilidad de Ganar

En esta subsección analizamos la configuración del juego en la que el objetivo de los jugadores es maximizar la probabilidad de ganar la carrera. Consideramos la misma programación dinámica estocástica presentada anteriormente, pero utilizamos la siguiente función de utilidad en la condición de borde (en lugar de la Ecuación (2.36)):

$$V_{N+1}(s_{N+1}) = \begin{cases} \mathbb{1}_{\{g_{N+1}>0\}} - \mathbb{1}_{\{g_{N+1}<0\}} & \text{si } H^A \text{ y } H^B \\ -1 & \text{si } H^A \text{ y no } H^B \\ 1 & \text{si no } H^A \text{ y } H^B \\ 0 & \text{en cualquier otro caso.} \end{cases} \quad (2.46)$$

El equilibrio en este caso se puede definir de manera similar a como se explicó anteriormente (ver Definiciones 2.4.1 y 2.4.3). Vale la pena señalar que los resultados de existencia de equilibrio también se pueden extender con modificaciones mínimas. Esto se expresa formalmente en el siguiente corolario:

Corolario 2.4.4.1. *La solución obtenida a partir del Algoritmo 2 al reemplazar la Ecuación (2.36) por la Ecuación (2.46) es un equilibrio de Nash-feedback Stackelberg para el juego donde los jugadores maximizan las probabilidades de ganar.*

DEMOSTRACIÓN. La demostración sigue el mismo argumento utilizado anteriormente para el Teorema 2.4.4. \square

Curiosamente, las probabilidades de ganar para cada conductor se pueden obtener directamente a partir del equilibrio. Para mayor simplicidad, usemos la misma notación que antes (como en la Sección 2.4.3) para el equilibrio.

Proposición 2.4.5. *La probabilidad de que el jugador A gane la carrera es igual a $(1 - (\pi^{A*})^T U' \pi^{B*})/2$.*

DEMOSTRACIÓN. Vea Apéndice D. □

2.5. Resultados Numéricos

Resolvemos una instancia del problema descrito con los siguientes parámetros: $N = 52$, $u_1 = 30$, $u_2 = 40$, $u_3 = 50$, $d_0^A = 97.22$, $d_0^B = 97.24$, $p_0^A = 20.2$, $p_0^B = 20.0$, $\lambda = 2.0$, $h = 0.02$, $l^{VSC} = 2$, $l^{SC} = 3$, $r^{VSC} = 0.3$, $r^{SC} = 0.3$, $k^{VSC} = 2$, $k^{SC} = 2$, $d_{VSC} = 170.0$, $d_{SC} = 200.0$, $p^AVSC = 10.0$, $p^BVSC = 10.0$, $p^ASC = 10.0$, $p^BSC = 10.0$, $g_0 = -0.4$, y $\Delta = 0.4$. En cuanto a las variables aleatorias $Z^{(1)}$, $Z^{(2)}$, y T^{DRS} , consideramos: $Z^{(1)}$ toma los valores 0.2, 0.4, y 0.6 con igual probabilidad; $Z^{(2)}$ toma los valores 0.5, 0.7, y 0.9 con igual probabilidad; y T^{DRS} toma los valores 0.1, 0.3, y 0.5 con las probabilidades 0.2, 0.7, y 0.1, respectivamente. La función de desgaste de los neumáticos utilizada para los compuestos blandos, medios y duros (*soft*, *medium* y *hard*, respectivamente) se describe en la Figura 2.2.

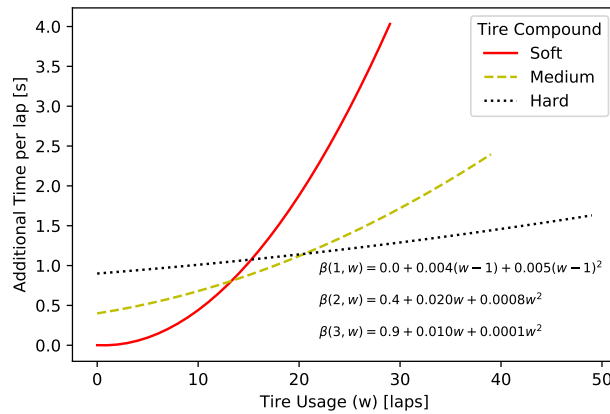


Figura 2.2: Tiempo adicional por vuelta de los neumáticos con diferentes usos en comparación con un neumático blando nuevo.

Para tener un número finito de estados, se discretiza el componente de tiempo de separación. Utilizamos una cuadrícula con una separación de 0.04 [s] entre ± 35 [s]. La instancia tiene cientos de miles de estados para vueltas mayores a 15. Se proporciona una tabla con

el número exacto de estados y el tiempo de resolución para cada vuelta en las Tablas G.2 y G.3 del Apéndice G. Las instancias se resuelven en un HPC utilizando 15 CPUs Intel de 2.10 GHz con 20 GB de RAM. El Algoritmo 2 se implementa en un código paralelo en C como una biblioteca que puede ser llamada desde Python.

2.5.1. Jugadores Estratégicos Maximizan la Brecha de Tiempo

Aplicamos el Algoritmo 2 para obtener el equilibrio de Nash-feedback Stackelberg del juego. La matriz de pagos de la decisión de los neumáticos al comienzo de la carrera (U') se muestra en la Tabla 2.1. Recuerda que estos valores representan la diferencia de tiempo esperada al final de la carrera, donde un valor negativo indica que A termina por delante de B . Las estrategias mixtas de los compuestos de neumáticos iniciales se encuentran en la columna de la derecha y en la fila inferior de la Tabla 2.1 para los jugadores A y B , respectivamente. Observamos que el jugador A elige el compuesto blando, mientras que B realiza una elección aleatoria entre los compuestos blando y medio. De hecho, a partir de la Tabla 2.1 observamos que, para ambos jugadores, la estrategia del compuesto duro está dominada por el compuesto blando.

Tabla 2.1: Pago esperado de los jugadores al comienzo de la carrera, U' ; y el equilibrio de Nash mixto, π^{A*} y π^{B*} .

		Player B			π^{A*}
		Soft	Medium	Hard	
Player A	Tire				
	Soft	-0.11	-1.32	-2.95	1.0
	Medium	-0.93	-0.08	-2.91	0.0
	Hard	1.07	0.76	-0.08	0.0
π^{B*}		0.60	0.40	0.0	

Simulamos 1,000 escenarios de carrera con los parámetros descritos al inicio de la Sección 2.5 y las decisiones de los jugadores tomadas desde el equilibrio de Nash-feedback Stackelberg. La primera fila de la Tabla 2.2 muestra diferentes estadísticas de las simulaciones. Dado que el conductor A es dos centésimas de segundo más rápido por vuelta que B , aunque B es dos décimas de segundo más rápido en los pits que A , A gana en el 68.9% de los escenarios.

Tabla 2.2: Distribución de victorias, tiempo promedio final de diferencia, desviación estándar, tiempo promedio final de diferencia condicionado a cada jugador ganador y estrategia de compuesto de neumáticos para iniciar la carrera (vuelta 1).

Strat.	Objective	A wins	Avg. g_N	Sd. g_N	Avg. g_N if		π^{A*}	π^{B*}
					A wins	B wins		
Player	Strat. Player	[%]	[s]	[s]				
Both	Time Gap	68.9	-0.63	1.77	-1.33	0.91	(1, 0, 0)	(0.60, 0.40, 0)
Both	Win Prob.	68.2	-1.88	4.59	-3.22	0.99	(1, 0, 0)	(0.61, 0.39, 0)
A	Time Gap	84.1	-1.05	1.92	-1.54	1.61	(1, 0, 0)	(0.5, 0.5, 0)
B	Time Gap	50.3	-0.03	0.59	-0.59	0.53	(0, 1, 0)	(0, 1, 0)

La Figura 2.3 muestra las estrategias de los jugadores y la diferencia de tiempo parcial (g_n) en dos escenarios: el panel izquierdo muestra una carrera sin banderas amarillas, mientras que el panel derecho muestra una carrera con un VSC en las vueltas 31 y 32. En ambos casos mostrados en la Figura 2.3, A (B) comienza con el neumático de compuesto blando (medio); por lo tanto, A planea una parada corta al principio seguida de una parada más larga, mientras que B planea la estrategia opuesta. Sin banderas amarillas, el jugador A gana la carrera por poco más de un segundo. Sin embargo, el VSC del escenario mostrado en el panel derecho de la Figura 2.3 favorece al jugador B , quien aprovecha la oportunidad de hacer una parada durante este período de tiempo perdiendo menos tiempo que al detenerse durante una vuelta regular. Como resultado, el jugador B gana en este caso por 9 segundos. Naturalmente, existen otros escenarios en los que una bandera amarilla beneficia al jugador A , consulte el Apéndice E.

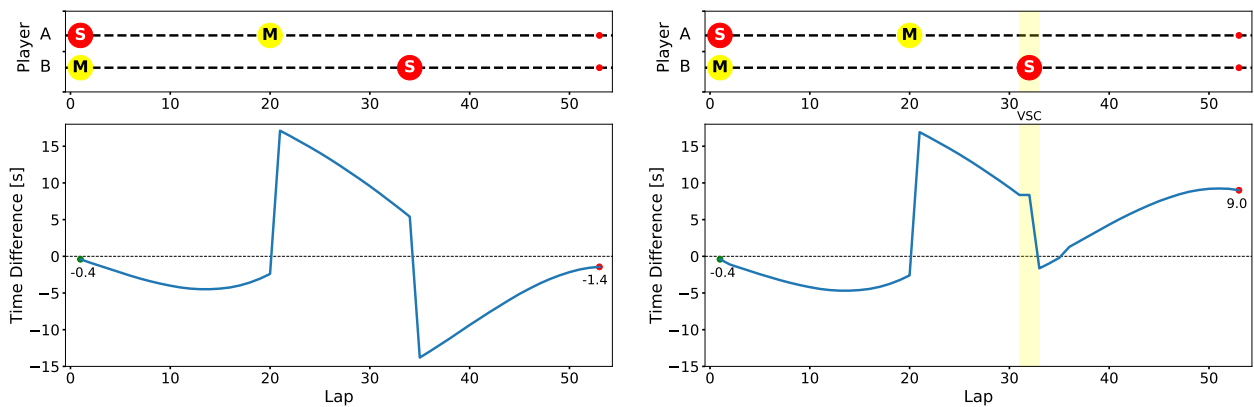


Figura 2.3: Estrategias de carrera de los jugadores (Superior), diferencia parcial de tiempo de carrera (Inferior); escenario sin bandera amarilla (Izquierda) y escenario con un VSC (Derecha).

2.5.2. Maximizando las Victorias

Analizamos la misma instancia que antes, excepto que el objetivo de ambos jugadores es maximizar la probabilidad de ganar la carrera (ver Sección 2.4.4). De la tercera columna de la Tabla 2.2, observamos que A gana el 68.8 % de las veces, de manera similar a la configuración de la Sección 2.5.1. Sin embargo, vemos que la desviación estándar de la diferencia de tiempo al final de la carrera, quinta columna de la Tabla 2.2, es significativamente mayor en este caso. La razón de esta diferencia se ilustra en el siguiente escenario de carrera. La Figura 2.4 muestra un escenario con un SC en la vuelta 17 en el que ambos pilotos cambian neumáticos. El jugador B coloca el neumático de compuesto suave, que no durará hasta el final de la carrera, requiriendo una parada adicional en el futuro. Cuando los jugadores maximizan las probabilidades de ganar (panel derecho de la Figura 2.4), el jugador B extiende el *stint* de este neumático tanto como sea posible, esperando una bandera amarilla. Como no se produce ninguna bandera amarilla después de la vuelta 20, el jugador B pierde por más de 30 segundos. En cambio, cuando los jugadores maximizan la diferencia de tiempo con respecto a su oponente, el jugador B realiza una parada en la vuelta 37, perdiendo pero solo por 2.7 segundos (ver panel izquierdo de la Figura 2.4). Entonces, si al jugador B le importa ganar (y no la diferencia de tiempo), apostará por esta oportunidad a pesar de perder por una diferencia de tiempo mayor en algunos casos.

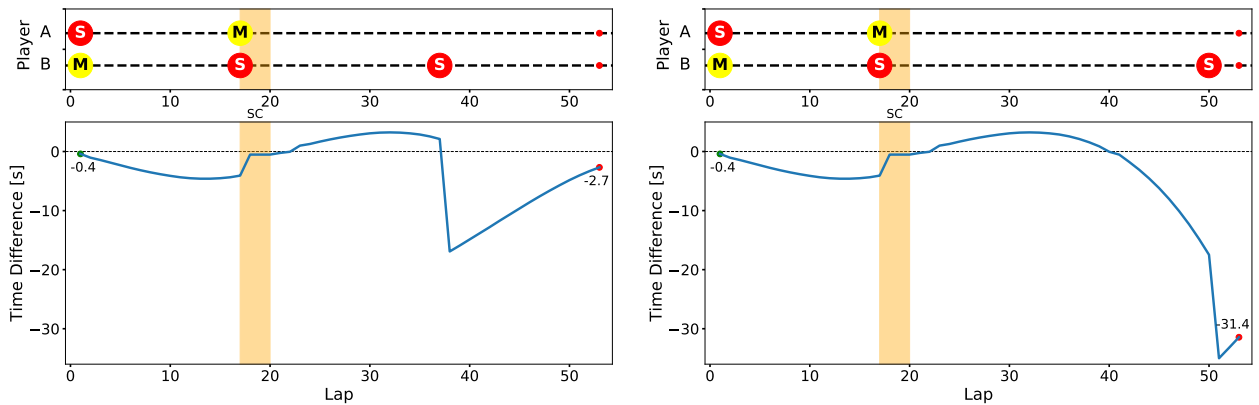


Figura 2.4: Estrategias de carrera de los jugadores (*Arriba*), diferencia de tiempo parcial de la carrera (*Abajo*), escenario en el que los jugadores quieren maximizar la brecha entre ambos pilotos (*Izquierda*), y escenario en el que ambos jugadores solo se preocupan por ganar (*Derecha*).

2.5.3. Jugador A es Estratégico y jugador B es Miope

Analizamos la misma instancia que antes, excepto que el jugador B es miope y el jugador A es consciente de esto. Mientras que el objetivo del jugador A sigue siendo maximizar la diferencia de tiempo, el objetivo del jugador B es minimizar el tiempo de carrera hasta el final de la carrera. En particular, el jugador B tiene en cuenta la incertidumbre de las banderas amarillas, pero ignora los efectos competitivos como la *función de interacción* del tiempo por vuelta (ver Ecuación (2.27)) y las acciones del jugador A . Por lo tanto, las decisiones del jugador B se obtienen resolviendo una Programación Dinámica Estocástica con la *función de tiempo de vuelta fantasma* como costo futuro (ver Ecuación (2.16)).

La Figura 2.5 muestra un escenario de carrera en el que ambos jugadores son estratégicos (panel izquierdo) y solo el jugador A es estratégico (panel derecho). Suponemos que A sabe que B no es estratégico. Observamos en este caso que la estrategia del jugador B no cambia entre las dos configuraciones. Por el contrario, cuando el jugador B es miope, el jugador A realiza una parada una vuelta antes que su oponente (vuelta 19) realizando un *undercut*¹⁵. Es decir, el jugador B lidera la carrera durante las vueltas 14 a 19. El jugador A hace una parada en la vuelta 19 para cambiar a un nuevo juego de neumáticos. Dado que los neumáticos de jugador B están desgastados, su tiempo por vuelta es mucho peor que el de jugador A en la vuelta 19 (aunque esto no se aprecia en la Figura 2.5 debido a la parada en pits de jugador A). Cuando el jugador B hace una parada en la vuelta 20, termina detrás del jugador A (es decir, se realiza el *undercut*). Entonces, el jugador A puede mantener su posición en la carrera hasta el final.

Podemos observar que la estrategia del jugador B se mantiene igual que antes, pero ahora el jugador A realiza una parada en la vuelta 19, una vuelta antes que el jugador B , para realizar un *undercut* y mantener su posición de liderazgo durante las vueltas restantes (ver panel derecho de la Figura 2.5). Recuerda que el jugador A es 0.02 segundos más rápido que el jugador B por vuelta, por lo que realizar este *undercut* le da la ventaja de esperar una victoria. Como se muestra en la tercera fila de la Tabla 2.2, el jugador A aumenta sus posibilidades de ganar en un 15.2% más, ahora que solo él es estratégico. Incluso si su

¹⁵ El *undercut* es una maniobra en la que el auto que está siguiendo de cerca a otro hace una parada algunas vueltas antes que su oponente. El piloto que sigue es capaz de lograr tiempos de vuelta más rápidos que su oponente (ya que este último está utilizando neumáticos más desgastados). Luego, cuando el líder hace una parada para cambiar neumáticos, termina detrás del auto que hizo la parada temprana.

estrategia inicial sigue siendo la misma (ver la octava columna de la Tabla 2.2), los cambios que realiza durante la carrera son más efectivos contra un jugador B miope.

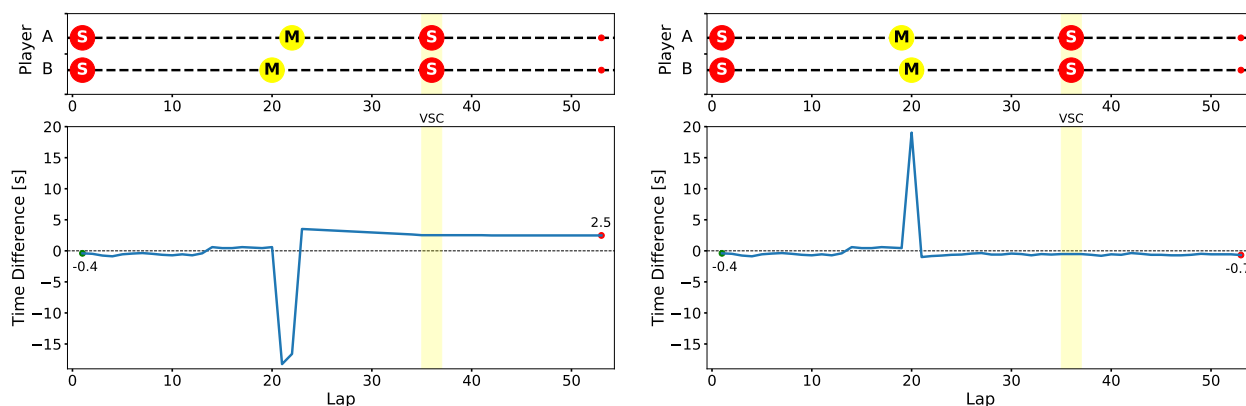


Figura 2.5: Estrategias de carrera de los jugadores (*Arriba*), diferencia parcial de tiempo de carrera (*Abajo*), escenario donde ambos jugadores son estratégicos (*Izquierda*), y escenario donde el jugador A es estratégico y el jugador B es miope (*Derecha*).

2.5.4. Variamos las Probabilidades de Bandera Amarilla

Analizamos el impacto de las banderas amarillas en las probabilidades de victoria de los pilotos. Consideramos la misma instancia descrita en la Sección 2.5.1, excepto que variamos la probabilidad de eventos de bandera amarilla. La probabilidad de tener un SC o VSC durante la carrera se considera igual, tomando valores del 0% al 50%, aumentando en un 10%. Ejecutamos 1,000 escenarios simulados para cada caso.

La Figura 2.6 muestra la probabilidad de victoria de estas simulaciones para los diferentes casos de bandera amarilla. Podemos observar que cuanto más probable es tener banderas amarillas, mayor es el beneficio para el piloto con menores probabilidades de ganar, en este caso el jugador B . El razonamiento detrás de esto es doble. En primer lugar, cuando ocurre un evento de bandera amarilla, permite al piloto rezagado acercarse a su rival, dándole una oportunidad de comenzar de nuevo y potencialmente remontar en la carrera. Esto crea una dinámica emocionante donde el resto de la carrera se convierte en una nueva oportunidad para ambos pilotos. En segundo lugar, existe la circunstancia afortunada en la que un piloto encuentra una bandera amarilla exactamente cuando planeaba hacer una parada en pits. Esta sincronización afortunada les permite ahorrar aproximadamente 10 segundos durante el cambio de neumáticos, lo que les proporciona una ventaja significativa en términos de tiempo

y posición en pista.

En contraste, cuando ambos competidores son conscientes de que la probabilidad de un evento de bandera amarilla es baja, el piloto más rápido, denominado como A en este caso, obtiene una ventaja al comenzar la carrera en la primera posición. El piloto A lidera tanto la carrera en sí como el proceso de toma de decisiones. Aunque podría parecer razonable que el jugador B intente una estrategia de *undercut*, considerando la velocidad superior del jugador A , se espera que la estrategia de *undercut* no se mantenga hasta el final y que el jugador A eventualmente recupere la ventaja. Solo en el 21.7% de los escenarios, el jugador B logrará ganar la carrera.

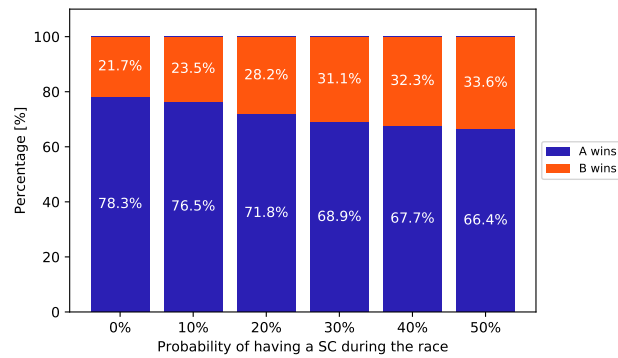


Figura 2.6: La distribución de victorias cuando la probabilidad de un SC durante la carrera va desde 0% hasta 50%. La probabilidad de tener un VSC es idéntica a la del SC.

Capítulo 3

Conclusiones

Hemos abordado con éxito tres preguntas importantes que surgen al prepararse para una carrera: ¿Con qué juego de neumáticos debería comenzar la carrera? ¿En qué vueltas debería realizar las paradas en pits? ¿Qué juego de neumáticos debería usar para el próximo tramo de la carrera?

Mostramos cómo la optimización de estrategias de carrera con competencia entre dos pilotos puede ser efectivamente modelada como un juego de Stackelberg de retroalimentación utilizando Programación Dinámica Estocástica. Esto permite tener en cuenta eventos de interacción entre pilotos, así como eventos estocásticos como escenarios de bandera amarilla, tiempos de vuelta, etc. Una vez que definimos un equilibrio de Stackelberg de retroalimentación, mostramos su existencia y proporcionamos un algoritmo para encontrarlo. Además de las decisiones de parada en pits durante la carrera (secuenciales), el modelo también considera la decisión inicial de los compuestos de neumáticos (simultánea).

Implementamos el algoritmo en un código paralelo en C y lo resolvemos para encontrar el equilibrio estratégico de la estrategia de carrera. La implementación es capaz de resolver instancias con cientos de millones de estados por vuelta en cuestión de horas. A partir de las instancias numéricas resueltas, observamos cómo diferentes escenarios de carrera afectan las decisiones de los jugadores y el resultado de la carrera. Observamos cómo las diferentes funciones objetivo de los pilotos inducen diferentes estrategias de carrera. De hecho, cuando los jugadores maximizan la probabilidad de ganar en lugar de la diferencia de tiempo con su oponente, tienden a asumir más riesgos en sus acciones. Nuestras observaciones de los cálculos numéricos revelaron que hay una ventaja significativa al ser estratégico cuando se

enfrenta a un oponente que ignora la competencia. Nuestras instancias muestran que un piloto estratégico que se enfrenta a uno miope aumenta las probabilidades de ganar en más del 15 % en comparación con el caso en que ambos son estratégicos. Finalmente, las banderas amarillas tienden a aumentar las posibilidades de ganar del piloto con menor rendimiento.

Se estima que la FIA otorga un premio monetario de USD 2,100 (consultar MotorSport [26]) a cada piloto por cada punto obtenido. Si observamos la temporada anterior, Verstappen emergió como el ganador con un total de 454 puntos. Aquellos que terminaron a mitad tabla acumularon 49 y 37 puntos (Bottas y Ricciardo, respectivamente). Esto significa que el piloto que obtuvo la mayor cantidad de puntos recibió aproximadamente USD 1 millón de la FIA, mientras que aquellos en el medio de la tabla recibieron entre 80,000 a 100,000 dólares. Logramos observar a un piloto que, a pesar de ser más lento en la pista en comparación con su rival, logra estratégicamente tener una probabilidad de ganar casi del 50 % en valor esperado. En consecuencia, si un equipo aplica esta estrategia para mejorar el rendimiento de sus pilotos "lentos", se esperaría que logre avanzar una posición cada 2 carreras, lo cual podría presenciar un aumento en el dinero de premio otorgado por la FIA, que oscilaría entre USD 24,150 y USD 169,050 por piloto por temporada¹⁶, además de un aumento en el premio monetario por rendimiento si el equipo termina en una posición más alta, que va desde 4 a 13 millones de dólares estadounidenses por equipo por temporada.

El modelo también puede adaptarse a otras aplicaciones más allá de la Fórmula 1 (o carreras), como en diversas industrias y procesos de toma de decisiones que requieren establecer la programación de un proceso mientras se enfrentan a la colaboración entre diferentes actores. Por ejemplo, problemas de competencia de participación de mercado, como un programa de televisión que optimiza la asignación de anuncios para maximizar la audiencia; o en campañas políticas donde una parada en pits podría ser una nueva posición de imagen que aumenta las posibilidades del candidato de conectarse con múltiples audiencias objetivo. En la industria minera, por ejemplo, el modelo se puede utilizar para optimizar los programas de mantenimiento de camiones y otros equipos. Al considerar factores como los patrones de uso, los costos de mantenimiento y el posible tiempo de inactividad, el modelo puede ayudar a determinar el momento óptimo para las actividades de mantenimiento, minimizando las interrupciones y maximizando la eficiencia operativa. En el sector de la salud, el modelo puede

¹⁶ Esto se debe a que cada mejora de posición proporciona entre 1 a 7 puntos adicionales, considerando el total de 23 carreras en un año y una probabilidad de ganar de 50 % contra su rival más cercano.

ayudar en la asignación y programación de recursos. Los hospitales pueden usar el modelo para optimizar la asignación de médicos y personal a diferentes departamentos o priorizar los casos de la sala de emergencias en función de varios factores, como la gravedad del paciente, los recursos disponibles y los datos históricos. Esto puede conducir a una utilización más eficiente de los recursos y una mejor atención al paciente. Incluso en el ámbito de la política, el modelo se puede aprovechar para mejorar las estrategias de campaña. Mediante el análisis de factores como la demografía de los votantes, datos históricos y calendarios de eventos, los políticos pueden identificar los días y lugares más beneficiosos para realizar mítines o eventos de campaña. Sin embargo, es importante adaptar y afinar el modelo y sus parámetros para que se ajusten al contexto específico de cada industria y aplicación. Los principios y metodologías subyacentes pueden servir como base, pero la adaptación es necesaria para abordar los desafíos y variables únicos de diferentes dominios. En general, la versatilidad y el potencial del modelo para optimizar los procesos de toma de decisiones lo convierten en una herramienta valiosa que se puede aplicar en varias industrias, lo que abre oportunidades para mejorar la eficiencia, ahorrar costos y obtener ventajas estratégicas.

Vale la pena destacar que el poder computacional desempeñó un papel fundamental en el desarrollo de esta investigación. Una computadora regular, digamos una con 12 GB de RAM y 8 núcleos, tomaría días para resolver las tareas de programación dinámica y simulación involucradas. En contraste, la supercomputadora NLHPC es capaz de resolver instancias de tamaño real (ver Tablas G.2 y G.3) en tan solo 3 a 6 horas. Una de las direcciones futuras, sin lugar a dudas, implica aprovechar computadoras aún más potentes con una mayor capacidad y memoria para manejar escenarios más complejos. Además, se harán esfuerzos para mejorar aún más el código con el fin de aumentar su eficiencia. Existen varias direcciones de investigación en las que se puede ampliar el modelo presentado. Una de ellas es considerar la competencia con más de dos pilotos. Esto claramente no es trivial, ya que el tamaño de la instancia aumentaría considerablemente. Asimismo, podría incorporar variables adicionales, como las condiciones climáticas, agregar juegos de neumáticos para condiciones de lluvia, considerar el desgaste previo de los neumáticos (no necesariamente colocan un set nuevo en una parada en pits) o también considerar que los parámetros cambian durante la carrera (por ejemplo, el rendimiento de un neumático de compuesto duro puede superar las expectativas durante la carrera). Estas mejoras tienen como objetivo acercar cada vez más el modelo a un programa del mundo real que pueda utilizarse en futuras temporadas de Fórmula 1.

Bibliografía

- [1] Fortune, “How formula 1 teams use big data to win.” <https://fortune.com/2015/11/12/big-data-formula-1-championship-race/>, 2015. (accessed December, 2022).
- [2] Formula 1, “Brilliant verstappen claims maiden title after victory in abu dhabi.” <https://www.formula1.com/en/latest/article.brilliant-verstappen-claims-maiden-title-after-victory-in-abu-dhabi-season.22Arui4l6BKK2PUsJgKqpZ.html>, 2021. accessed December, 2022.
- [3] TotalSportal, “Formula 1 prize money 2022 distribution table.” <https://www.totalsportal.com/f1/formula-one-prize-money/>, 2022. (accessed May, 2023).
- [4] Chain Bear, “Basics of f1 race strategy.” https://youtu.be/wqf-dJyU_WA, 2017. accessed April, 2021.
- [5] McLaren Racing Limited, “Formula one race strategy.” <https://www.stem.org.uk/resources/elibrary/resource/25407/formula-one-race-strategy#&gid=undefined&pid=1>, 2019. accessed December, 2021.
- [6] Liu, X., Fotouhi, A., y Auger, D. J., “Formula-e race strategy development using distributed policy gradient reinforcement learning,” Knowledge-Based Systems, vol. 216, p. 106781, 2021.
- [7] Anselma, P. G., “Optimal adaptive race strategy for a formula-e car,” Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering, vol. 236, no. 9, pp. 2185–2199, 2022.
- [8] Bekker, J. y Lotz, W., “Planning formula one race strategies using discrete-event simulation,” Journal of the Operational Research Society, vol. 60, no. 7, pp. 952–961, 2009.

- [9] Phillips, A., “Building a race simulator.” <https://f1metrics.wordpress.com/2014/10/03/building-a-race-simulator/>, 2014. accessed December, 2022.
- [10] Heilmeier, A., Graf, M., y Lienkamp, M., “A race simulation for strategy decisions in circuit motorsports,” en 2018 21st International Conference on Intelligent Transportation Systems (ITSC), pp. 2986–2993, IEEE, 2018.
- [11] The Sport Rush, “How does qualifying work in f1? qualifying system explained.” <https://thesportsrush.com/f1-news-how-does-qualifying-work-in-f1-explained/>, 2020. accessed December, 2022.
- [12] Heilmeier, A., Graf, M., Betz, J., y Lienkamp, M., “Application of monte carlo methods to consider probabilistic effects in a race simulation for circuit motorsport,” Applied Sciences, vol. 10, no. 12, p. 4229, 2020.
- [13] F1 Strategy Blog, “Race simulator: Downloadable r program code.” <https://f1strategyblog.wordpress.com/2019/05/10/race-simulator-downloadable-r-program-code/>, 2019. accessed December, 2022.
- [14] Tulabandhula, T. y Rudin, C., “Tire changes, fresh air, and yellow flags: challenges in predictive analytics for professional racing,” Big data, vol. 2, no. 2, pp. 97–112, 2014.
- [15] Choo, C. L. W., Real-time decision making in motorsports: analytics for improving professional car race strategy. PhD thesis, Massachusetts Institute of Technology, 2015.
- [16] Heilmeier, A., Thomaser, A., Graf, M., y Betz, J., “Virtual strategy engineer: Using artificial neural networks for making race strategy decisions in circuit motorsport,” Applied Sciences, vol. 10, no. 21, p. 7805, 2020.
- [17] Heine, O. F. C. y Thraves, C., “On the optimization of pit stop strategies via dynamic programming,” Central European Journal of Operations Research, pp. 1–30, 2022.
- [18] Simaan, M. y Cruz, J. B., “Additional aspects of the stackelberg strategy in nonzero-sum games,” Journal of Optimization Theory and Applications, vol. 11, no. 6, pp. 613–626, 1973.
- [19] Simaan, M. y Cruz, J. B., “On the stackelberg strategy in nonzero-sum games,” Journal of Optimization Theory and Applications, vol. 11, no. 5, pp. 533–555, 1973.
- [20] Li, T. y Sethi, S. P., “A review of dynamic stackelberg game models,” Discrete &

Continuous Dynamical Systems-B, vol. 22, no. 1, p. 125, 2017.

- [21] Başar, T. y Olsder, G. J., Dynamic noncooperative game theory. SIAM, 1998.
- [22] Nie, P.-y., Lai, M.-y., y Zhu, S.-j., “Dynamic feedback stackelberg games with alternating leaders,” Nonlinear Analysis: Real World Applications, vol. 9, no. 2, pp. 536–546, 2008.
- [23] MotorSport, “Insider’s guide: What are f1’s tyre rules?.” <https://us.motorsport.com/f1/news/insiders-guide-what-are-f1s-tyre-rules/6885459/>, 2022. (accessed March, 2023).
- [24] Von Stengel, B., “Computing equilibria for two-person games,” Handbook of game theory with economic applications, vol. 3, pp. 1723–1759, 2002.
- [25] Biswas, S., “What is the difference between a safety car and a virtual safety car in f1?” <https://www.essentiallysports.com/f1-news-what-is-the-difference-between-a-safety-car-and-a-virtual-safety-car-in-f1/>, 2022. (accessed March, 2023).
- [26] MotorSport, “Fia nets \$27m from 2023 f1 entry fees.” <https://www.motorsport.com/f1/news/fia-nets-27m-from-2023-f1-entry-fees/10433814/>, 2023. (accessed May, 2023).

Anexos

Anexo A. Resumen de Parámetros

A continuación se muestra un resumen de la notación de todos los parámetros en la Tabla A.1; y de las decisiones, estados y variables aleatorias en las Tablas A.2 y A.3.

A.1. Parámetros

Tabla A.1: Parámetros del problema

Notación	Descripción	Valor/rango
\mathcal{T}	Conjunto de neumáticos	$\{1, 2, 3\}$
N	Vueltas de la carrera	\mathbb{Z}_+
u_t	Vida útil del neumático $t \in \mathcal{T}$	\mathbb{Z}_+
d_0^i	Tiempo de vuelta para el piloto $i \in \{A, B\}$ al usar neumáticos blandos nuevos al comienzo de la carrera (sin paradas en pits, interacciones y DRS)	$[0, \infty)$
p_0^i	Tiempo adicional de vuelta para el piloto $i \in \{A, B\}$ debido a una parada en pits	$[0, \infty)$
λ	Factor que penaliza el tiempo de vuelta de los pilotos debido a estar demasiado cerca entre sí	$[0, \infty)$
h	Reducción de tiempo de vuelta entre dos vueltas consecutivas atribuido a tener un auto más ligero en la segunda vuelta (ya que tiene menos combustible que en la vuelta anterior)	$[0, \infty)$
l^{VSC}	Vueltas que dura un VSC	\mathbb{Z}_+
l^{SC}	Vueltas que dura un SC	\mathbb{Z}_+
q^{VSC}	Probabilidad de tener un VSC durante una vuelta	$[0, \infty)$
q^{SC}	Probabilidad de tener un SC durante una vuelta	$[0, \infty)$
r^{VSC}	Probabilidad de tener al menos un VSC durante la carrera	$[0, \infty)$
r^{SC}	Probabilidad de tener al menos un SC durante la carrera	$[0, \infty)$
k^{VSC}	Número de vueltas que se deben correr después del final de un VSC para habilitar el DRS	\mathbb{Z}_+
k^{SC}	Número de vueltas que se deben correr después del final de un SC para habilitar el DRS	\mathbb{Z}_+
d_{VSC}	Tiempo de vuelta durante un VSC (sin parada en pits)	$[0, \infty)$
d_{SC}	Tiempo de vuelta durante un SC (sin parada en pits)	$[0, \infty)$
p_{VSC}^i	Tiempo adicional de vuelta para el piloto $i \in \{A, B\}$ debido a una parada en pits durante un VSC	$[0, \infty)$
p_{SC}^i	Tiempo adicional de vuelta para el piloto $i \in \{A, B\}$ debido a una parada en pits durante un SC	$[0, \infty)$
t^{DRS}	Reducción de tiempo debido al uso del DRS en la vuelta n (solo en DP determinista)	\mathbb{R}
$z_n^{(1)}$	Factor de reducción de tiempo para el auto de adelante debido a las interacciones con el auto de atrás en la vuelta n (solo en DP determinista)	$[0, \infty)$
$z_n^{(2)}$	Factor de reducción de tiempo para el auto de atrás debido a las interacciones con el auto líder en la vuelta n (solo en DP determinista)	$[0, \infty)$
g_0	Diferencia de tiempo al comienzo de la carrera. Negativo (positivo) si A (B) comienza primero.	\mathbb{R}
Δ	Diferencia mínima de tiempo entre los pilotos en la salida de la parada en pits	$[0, \infty)$

A.2. Decisiones, Estados y Variables Aleatorias

Tabla A.2: Decisiones, estados y variables aleatorias del problema

Notación	Descripción	Valor/rango	Tipo	Problema
x_n^A	Neumático cambiado por el piloto A en la parada en pits en la vuelta n si $x_n^A \in \mathcal{T}$; sin parada en pits durante la vuelta n si $x_n^A = 0$	$\{0\} \cup \mathcal{T}$	Decisión	DP,SDP
x_n^B	Neumático cambiado por el piloto B en la parada en pits en la vuelta n si $x_n^B \in \mathcal{T}$; sin parada en pits durante la vuelta n si $x_n^B = 0$	$\{0\} \cup \mathcal{T}$	Decisión	DP,SDP
t_n^A	Neumático del piloto A al comienzo de la vuelta n	\mathcal{T}	Estado	DP,SDP
t_n^B	Neumático del piloto B al comienzo de la vuelta n	\mathcal{T}	Estado	DP,SDP
w_n^A	Uso (en vueltas) de los neumáticos del piloto A al comienzo de la vuelta n	$\{0, 1, \dots, u_t + 1\}$	Estado	DP,SDP
w_n^B	Uso (en vueltas) de los neumáticos del piloto B al comienzo de la vuelta n	$\{0, 1, \dots, u_t + 1\}$	Estado	DP,SDP
m_n^A	Indicador igual a 1 si A ha utilizado al menos dos compuestos de neumáticos diferentes al comienzo de la vuelta n , 0 en caso contrario	$\{0, 1\}$	Estado	DP,SDP
m_n^B	Indicador igual a 1 si B ha utilizado al menos dos compuestos de neumáticos diferentes al comienzo de la vuelta n , 0 en caso contrario	$\{0, 1\}$	Estado	DP,SDP
g_n	Diferencia de tiempo entre los autos al comienzo de la vuelta n . $g_n < 0$ si A es el primero	\mathbb{R}	Estado	DP,SDP

A.3. Decisiones, Estados y Variables Aleatorias

Tabla A.3: Decisiones, estados y variables aleatorias del problema estocástico

Notación	Descripción	Valor/rango	Tipo	Problema
y_n^{VSC}	Vueltas restantes bajo VSC desde el comienzo de la vuelta n	$\{0, 1, \dots, l^{VSC}\}$	Estado	SDP
y_n^{SC}	Vueltas restantes bajo SC desde el comienzo de la vuelta n	$\{0, 1, \dots, l^{SC}\}$	Estado	SDP
y_n^{DRS}	Vueltas restantes para habilitar el DRS desde el comienzo de la vuelta n	\mathbb{Z}_+	Estado	SDP
π^A	Vector con las probabilidades de que A elija cada tipo de neumático al comienzo de la carrera	$[0, 1]^{ T }$	Decisión	SDP
π^B	Vector con las probabilidades de que B elija cada tipo de neumático al comienzo de la carrera	$[0, 1]^{ T }$	Decisión	SDP
Y_n^{VSC}	Vueltas restantes bajo VSC en la vuelta n	$\{0, \dots, l^{VSC}\}$	V.A.	SDP
Y_n^{SC}	Vueltas restantes bajo SC en la vuelta n	$\{0, \dots, l^{SC}\}$	V.A.	SDP
T_n^{DRS}	Reducción de tiempo debido al uso del DRS en la vuelta n	\mathbb{R}	V.A.	SDP
$Z_n^{(1)}$	Factor de reducción de tiempo al auto de adelante debido a las interacciones con el auto de atrás en la vuelta n	$[0, \infty]$	V.A.	SDP
$Z_n^{(2)}$	Factor de reducción de tiempo al auto de atrás debido a las interacciones con el auto líder en la vuelta n	$[0, \infty]$	V.A.	SDP

Anexo B. Mínima Brecha de Tiempo

En esta sección, proporcionamos los detalles de los cambios necesarios en el modelo presentado en la Sección 2.4 para garantizar una diferencia de tiempo mínima entre los autos en un punto del circuito. La razón de esto es que generalmente los circuitos tienen secciones donde los autos están prácticamente obligados a seguir una *línea de carrera*, lo que implica mantener una diferencia de tiempo entre los autos. En este caso, consideramos que los autos deben tener una diferencia de tiempo mínima de $\Delta > 0$ después de salir de la parada en pits.

Los dos cambios necesarios son los siguientes: (i) En primer lugar, la *función de interacción* debe cambiarse a:

$$\eta_n^{\prime\prime A}(\epsilon'_n) = \begin{cases} \exp(-\lambda \cdot |\min\{-\Delta, \epsilon'_n\}|) \cdot Z_n^{(1)} & \text{si } \epsilon'_n < 0 \\ \exp(-\lambda \cdot |\max\{\Delta, \epsilon'_n\}|) \cdot Z_n^{(2)} & \text{en cualquier otro caso.} \end{cases} \quad (\text{B.1})$$

Para el caso del piloto B , la *función de interacción* $\eta_n^{\prime\prime B}$ se define de la misma manera que la Ecuación (B.1), excepto que la condición en el lado derecho de la ecuación cambia entre los dos casos.

(ii) El segundo cambio se refiere a la función de tiempo por vuelta, la cual se modifica de la siguiente manera:

$$\begin{aligned} \mu_{n,0}^{\prime A}(s_n, x_n^A, x_n^B, T_n^{DRS}, Z_n^{(1)}, Z_n^{(2)}) &= \gamma_n^A + \eta_n^{\prime\prime A}(\epsilon'_n) - T_n^{DRS} \cdot \mathbb{1}_{\{g_n \in [0,1], y_n^{DRS}=0\}} \\ &\quad + \max\{\Delta - \epsilon, 0\} \cdot \mathbb{1}_{\{\epsilon > 0\}}, \\ \mu_{n,0}^{\prime B}(s_n, x_n^A, x_n^B, T_n^{DRS}, Z_n^{(1)}, Z_n^{(2)}) &= \gamma_n^B + \eta_n^{\prime\prime B}(-\epsilon'_n) - T_n^{DRS} \cdot \mathbb{1}_{\{g_n \in [0,1], y_n^{DRS}=0\}} \\ &\quad + \max\{\Delta + \epsilon, 0\} \cdot \mathbb{1}_{\{\epsilon \leq 0\}}. \end{aligned}$$

Entonces, la función de tiempo por vuelta debería actualizarse de la siguiente manera:

$$\mu_n^{ii}(s_n, x_n^A, x_n^B, T_n^{DRS}, Z_n^{(1)}, Z_n^{(2)}) = \begin{cases} \mu_{n,0}^i & \text{if } y_n^{VSC} + y_n^{SC} = 0 \\ \mu_{n,VSC}^i & \text{if } y_n^{VSC} > 0 \\ \mu_{n,SC}^i & \text{if } y_n^{SC} > 0, \end{cases} \quad (\text{B.2})$$

para $i \in \{A, B\}$, y la Ecuación (2.32) debería actualizarse para que utilice la Ecuación (B.2) en lugar de la Ecuación (2.28).

Anexo C. Demostración del Teorema 2.4.2

Demostramos el resultado utilizando inducción en n desde la última vuelta hasta la primera. Antes de hacer esto, generalizamos la Definición 2.4.1 para el caso en que el juego se encuentra en una etapa intermedia (vuelta) $j \in \{1, \dots, N\}$ para cualquier estado arbitrario $s_j \in \mathcal{S}_j$. Definimos $\mathcal{S}_n(s_j)$ como el conjunto de estados alcanzables desde el estado s_j en la

etapa n . Alcanzables en el sentido de que existe una probabilidad no nula de llegar a este estado en la vuelta n , considerando todas las posibles combinaciones de acciones de ambos jugadores entre las vueltas j y $n - 1$, dado que estamos en el estado s_j en la vuelta $j \leq n$. Claramente, $\mathcal{S}_j(s_j) = \{s_j\}$. De manera similar a antes, también definimos los conjuntos $\mathcal{S}_n^-(s_j) = \{s_n \in \mathcal{S}_n(s_j) | g_n < 0\}$ y $\mathcal{S}_n^+(s_j) = \{s_n \in \mathcal{S}_n(s_j) | g_n > 0\}$ para todo $n \geq j$.

Definición C.1. *Un equilibrio de Stackelberg de retroalimentación desde un estado s_j en una etapa dada $j \in \{1, \dots, N\}$ es un par de estrategias:*

$\left(\{x_n^{A*}(s_n)\}_{s_n \in \mathcal{S}_n^-(s_j), n \in \{j, \dots, N\}} \cup \{x_n^{A*}(s_n, b)\}_{s_n \in \mathcal{S}_n^+(s_j), b \in \mathcal{T}_0, n \in \{j, \dots, N\}} \right)$ y
 $\left(\{x_n^{B*}(s_n)\}_{s_n \in \mathcal{S}_n^+(s_j), n \in \{1, \dots, N\}} \cup \{x_n^{B*}(s_n, a)\}_{s_n \in \mathcal{S}_n^-(s_j), a \in \mathcal{T}_0, n \in \{j, \dots, N\}} \right)$
tal que para todo $n \in \{j, \dots, N\}$:

1. Si $g_n < 0$

$$V'_n(s_n, x_n^{A*}(s_n), x_n^{B*}(s_n, x_n^{A*}(s_n))) \leq V'_n(s_n, x_n^A(s_n), x_n^{B*}(s_n, x_n^A(s_n))) \quad (\text{C.1})$$

$$V'_n(s_n, x_n^{A*}(s_n), x_n^{B*}(s_n, x_n^{A*}(s_n))) \geq V'_n(s_n, x_n^{A*}(s_n), x_n^B(s_n, x_n^{A*}(s_n))). \quad (\text{C.2})$$

2. Si $g_n > 0$

$$V'_n(s_n, x_n^{A*}(s_n, x_n^{A*}(s_n)), x_n^{B*}(s_n)) \leq V'_n(s_n, x_n^A(s_n, x_n^{B*}(s_n)), x_n^{B*}(s_n)) \quad (\text{C.3})$$

$$V'_n(s_n, x_n^{A*}(s_n, x_n^{B*}(s_n)), x_n^{B*}(s_n)) \geq V'_n(s_n, x_n^{A*}(s_n, x_n^B(s_n)), x_n^B(s_n)), \quad (\text{C.4})$$

para todo $\left(\{x_n^A(s_n)\}_{s_n \in \mathcal{S}_n^-(s_j), n \in \{j, \dots, N\}} \cup \{x_n^A(s_n, b)\}_{s_n \in \mathcal{S}_n^+(s_j), b \in \mathcal{T}_0, n \in \{j, \dots, N\}} \right)$ y
 $\left(\{x_n^B(s_n)\}_{s_n \in \mathcal{S}_n^+(s_j), n \in \{j, \dots, N\}} \cup \{x_n^B(s_n, a)\}_{s_n \in \mathcal{S}_n^-(s_j), a \in \mathcal{T}_0, n \in \{j, \dots, N\}} \right)$.

Es importante tener en cuenta que en la Definición C.1 estamos considerando el juego a partir de la etapa j con una condición inicial particular. Por lo tanto, esto también es un juego de Stackelberg de retroalimentación en el que las estrategias de los jugadores están involucradas desde la etapa j hasta la etapa N . Claramente, la Definición C.1 coincide con la Definición 2.4.1 en los casos en los que s_j (de la Definición C.1) se refiere al mismo estado inicial s_1 (de la Definición 2.4.1).

Denotaremos la salida del Algoritmo 2 como:

$$\begin{aligned} & \{x_n^{\diamond A}(s_n)\}_{s_n \in \mathcal{S}_n^-, n \in \{1, \dots, N\}} \cup \{x_n^{\diamond A}(s_n, b)\}_{s_n \in \mathcal{S}_n^+, b \in \mathcal{T}_0, n \in \{1, \dots, N\}} \\ & \{x_n^{\diamond B}(s_n)\}_{s_n \in \mathcal{S}_n^+, n \in \{1, \dots, N\}} \cup \{x_n^{\diamond B}(s_n, a)\}_{s_n \in \mathcal{S}_n^-, a \in \mathcal{T}_0, n \in \{1, \dots, N\}}, \end{aligned} \quad (\text{C.5})$$

donde no estamos considerando la estrategia mixta de la primera etapa.

Mostraremos primero el caso base, es decir, la última vuelta $n = N$. Además, consideremos un estado fijo $s_N \in \mathcal{S}_N$. Como comentario adicional, note que dado que solo queda una etapa en este caso, este es un juego de Stackelberg clásico, aunque el líder en este caso depende del componente g_N del estado s_N (fijo). Trataremos el caso en el que el componente $g_N < 0$, es decir, A va primero en la última vuelta. El otro caso se omite ya que su análisis es análogo. El conductor B no tiene incentivo para cambiar su estrategia, de hecho, para cualquier par de estrategias $(x_N^{\diamond A}(s_N), \{x_N^{\prime B}(s_N, a)\}_{a \in \mathcal{T}_0})$ se cumple que

$$\begin{aligned} V'_N(s_N, x_N^{\diamond A}(s_N), x_N^{\diamond B}(s_N, x_N^{\diamond A}(s_N))) &= \max_{b \in \mathcal{T}_0} V'_N(s_N, x_N^{\diamond A}(s_N), b) \\ &\geq V'_N(s_N, x_N^{\diamond A}(s_N), x_N^{\prime B}(s_N, x_N^{\diamond A}(s_N))), \end{aligned}$$

donde la igualdad se debe a la Ecuación (2.40) utilizada en la línea 11 del Algoritmo 2, y la desigualdad se debe a que estamos considerando una acción particular para B , $x_N^{\prime B}(s_N, x_N^{\diamond A}(s_N))$, en este caso. De manera similar, el jugador A no tiene incentivo para cambiar su estrategia. Consideremos cualquier par de estrategias $(x_N^{\prime A}(s_N), \{x_N^{\diamond B}(s_N, a)\}_{a \in \mathcal{T}_0})$. Entonces,

$$\begin{aligned} V'_N(s_N, x_N^{\prime A}(s_N), x_N^{\diamond B}(s_N, x_N^{\prime A}(s_N))) &= \min_{a \in \mathcal{T}_0} V'_N(s_N, a, x_N^{\diamond B}(s_N, a)) \\ &\leq V'_N(s_N, x_N^{\prime A}(s_N), x_N^{\diamond B}(s_N, x_N^{\prime A}(s_N))), \end{aligned}$$

donde la igualdad se debe a la línea 12 del Algoritmo 2, y la desigualdad se debe a que estamos considerando una acción particular para A , $x_N^{\prime A}(s_N)$.

Consideremos la hipótesis de inducción como verdadera para todas las etapas $n \geq j + 1$, para un valor fijo de $j \in \{1, \dots, N - 1\}$. (Para facilitar la notación, utilizaremos j como índice de la inducción.) En otras palabras, para todos los estados $s_n \in \mathcal{S}_n$ con $n \geq j + 1$, la salida del Algoritmo 2 (ver ecuación (C.5)), satisface la Definición C.1. Consideremos que estamos en la etapa j en un estado fijo $s_j \in \mathcal{S}_j$. Al igual que en el caso base, solo consideramos el caso en el que el componente $g_j < 0$.

Primero demostraremos que B no tiene incentivos para cambiar su estrategia. En efecto, consideremos cualquier par de estrategias

$$\begin{aligned} & \{x_n^{\diamond A}(s_n)\}_{s_n \in \mathcal{S}_n^-(s_j), n \in \{j, \dots, N\}} \cup \{x_n^{\diamond A}(s_n, b)\}_{s_n \in \mathcal{S}_n^+(s_j), b \in \mathcal{T}_0, n \in \{j, \dots, N\}} \\ & \{x_n'^B(s_n)\}_{s_n \in \mathcal{S}_n^+(s_j), n \in \{j, \dots, N\}} \cup \{x_n'^B(s_n, a)\}_{s_n \in \mathcal{S}_n^-(s_j), a \in \mathcal{T}_0, n \in \{j, \dots, N\}}. \end{aligned} \quad (\text{C.6})$$

Consideremos dos casos con respecto a la estrategia a la que B se desvía.

1. Si la estrategia de B de la Ecuación (C.6) coincide con la obtenida del Algoritmo 2 en la etapa j , es decir, $x'^B(s_j, a) = x^{\diamond B}(s_j, a)$ para todo a (nota que estamos denotando explícitamente a B como el seguidor en la etapa j ya que $g_j < 0$), entonces claramente se cumple que:

$$V_j'(s_j, x_j^{\diamond A}(s_j), x_j^{\diamond B}(s_j, x_j^{\diamond A}(s_j))) = V_j'(s_j, x_j^{\diamond A}(s_j), x_j'^B(s_j, x_j^{\diamond A}(s_j))),$$

Y por lo tanto se cumple la Ecuación (C.2). Luego, para todos los demás estados $s_n \in \mathcal{S}_n(s_j)$ para $n > j$, las Ecuaciones (C.2) y (C.4) se cumplen debido a la hipótesis de inducción. Entonces, el jugador B no tiene incentivos para desviarse en este caso.

2. Si ahora consideramos el caso en el que la estrategia de B dada por la Ecuación (C.6) no coincide con la obtenida del Algoritmo 2 en la etapa j , es decir, $x'^B(s_j, a) \neq x^{\diamond B}(s_j, a)$ para algunos a , entonces:

$$\begin{aligned} V_j'(s_j, x_j^{\diamond A}(s_j), x_j^{\diamond B}(s_j, x_j^{\diamond A}(s_j))) &= \max_{b \in \mathcal{T}_0} V_j'(s_j, x_j^{\diamond A}(s_j), b) \\ &\geq V_j'(s_j, x_j^{\diamond A}(s_j), x_j'^B(s_j, x_j^{\diamond A}(s_j))), \end{aligned}$$

donde la igualdad se debe a la Ecuación (2.40) utilizada en la línea 11 del Algoritmo 2, y la desigualdad se debe a que estamos considerando una acción particular para B , $x_j'^B(s_j, x_j^{\diamond A}(s_j))$ en este caso. Es decir, B no tiene incentivos para desviarse en la etapa j (en el estado s_j).

Ahora demostraremos que A no tiene incentivos para desviarse de la estrategia de salida

del Algoritmo 2. Consideremos cualquier par de estrategias

$$\begin{aligned} & \{x_n'^A(s_n)\}_{s_n \in \mathcal{S}_n^-(s_j), n \in \{j, \dots, N\}} \cup \{x_n'^A(s_n, b)\}_{s_n \in \mathcal{S}_n^+(s_j), b \in \mathcal{T}_0, n \in \{j, \dots, N\}} \\ & \{x_n^{\diamond B}(s_n)\}_{s_n \in \mathcal{S}_n^+(s_j), n \in \{j, \dots, N\}} \cup \{x_n^{\diamond B}(s_n, a)\}_{s_n \in \mathcal{S}_n^-(s_j), a \in \mathcal{T}_0, n \in \{j, \dots, N\}}. \end{aligned} \quad (\text{C.7})$$

Consideremos los dos siguientes casos (enumerados como (iii) y (iv)) con respecto a la estrategia a la que A se desvía:

3. Si la estrategia de A dada por la Ecuación (C.7) coincide con la obtenida del Algoritmo 2 en la etapa j , es decir, $x^A(s_j) = x^{\diamond A}(s_j)$ (nota que estamos denotando explícitamente a A como el líder en la etapa j ya que $g_j < 0$)

$$V_j'(s_j, x_j^{\diamond A}(s_j), x_j^{\diamond B}(s_j, x_j^{\diamond A}(s_j))) = V_j'(s_j, x_j^{\diamond A}(s_j), x_j'^B(s_j, x_j^{\diamond A}(s_j))),$$

y por lo tanto se cumple la Ecuación (C.1). Luego, para todos los demás estados $s_n \in \mathcal{S}_n(s_j)$ para $n > j$, las Ecuaciones (C.1) y (C.3) se cumplen debido a la hipótesis de inducción. Entonces, el jugador A no tiene incentivos para desviarse en este caso.

4. Si la estrategia de A en la Ecuación (C.7) no coincide con la obtenida del Algoritmo 2 en la etapa j , es decir, $x^A(s_j) \neq x^{\diamond A}(s_j)$, entonces:

$$\begin{aligned} V_j'(s_j, x_j^{\diamond A}(s_j), x_j^{\diamond B}(s_j, x_j^{\diamond A}(s_j))) &= \min_{a \in \mathcal{T}_0} V_j'(s_j, a, x_j^{\diamond B}(s_j, a)) \\ &\leq V_j'(s_j, x_j'^A(s_j), x_j^{\diamond B}(s_j, x_j'^A(s_j))), \end{aligned}$$

donde la igualdad se debe a la línea 12 del Algoritmo 2, y la desigualdad se debe a que estamos considerando una acción particular para A , $x_j'^A(s_j)$. Entonces, A no tiene incentivos para desviarse en la etapa j (en el estado s_j).

Dado que ningún jugador tiene incentivos para desviarse, la solución obtenida del Algoritmo 2 es un Equilibrio de Stackelberg con retroalimentación desde el estado s_j , lo que concluye la prueba.

Anexo D. Demostración de la Proposición 2.4.5

Mostraremos primero que, para cada etapa $n \in \{1, \dots, N + 1\}$, la ganancia en cada estado s_n , $V_n^{''*}(s_n)$, es tal que la probabilidad de ganar para el jugador A en ese estado (en equilibrio) está dada por $1/2 + V_n^{''*}(s_n)/2$. Utilizando inducción en orden inverso, notemos que esto se cumple para la etapa $N + 1$, es decir, la condición de borde, donde el jugador que gana es conocido y está determinado por la Ecuación (2.46). Para el paso de inducción, supongamos que para la etapa n se cumple que $1/2 + V_n^{''*}(s_n)/2$ es la probabilidad de ganar para el jugador A en el estado s_n . Consideremos cualquier estado arbitrario s_{n-1} en la etapa $n - 1$. Observamos que, según la Ecuación (2.37), se cumple que:

$$\begin{aligned} \frac{1 - V_{n-1}^{''*}(s_{n-1})}{2} &= \frac{1 - \mathbb{E}_{W_{n-1}}[V_n^{''*}(s_n)]}{2} \\ &= \mathbb{E}_{W_{n-1}} \left[\frac{1 - V_n^{''*}(s_n)}{2} \right]. \end{aligned} \quad (\text{D.1})$$

Utilizando la hipótesis de inducción, el argumento dentro de la esperanza en la Ecuación (D.1) corresponde a la probabilidad de ganar del jugador A . Entonces, la esperanza de esto también tendrá la misma interpretación. Concluyendo la inducción. Luego, aplicando las estrategias de equilibrio mixtas de la primera vuelta a la matriz de ganancias de suma-cero U'' , obtenemos que la probabilidad de ganar del jugador A está dada por:

$$(\pi^A)^T \left(\frac{\mathbf{1}_{|\mathcal{T}| \times |\mathcal{T}|} - \mathbf{U}''}{2} \right) \pi^B = \frac{\mathbf{1}_{|\mathcal{T}| \times |\mathcal{T}|} - (\pi^A)^T \mathbf{U}'' \pi^B}{2},$$

concluyendo la demostración.

Anexo E. Escenario Alternativo con SC

Mostramos un escenario de las simulaciones mencionadas en la Sección 2.5.1 donde una bandera amarilla beneficia al jugador A . La Figura E.1 muestra una carrera con un SC en las vueltas 26 a 28. Observamos que al momento del SC, el mejor neumático a elegir (hasta el final de la carrera) es el compuesto Blando (Soft). El jugador A actúa de acuerdo a esto, mientras que el jugador B no lo hace. Durante el SC, el jugador B cambia al compuesto Blando ya que debe elegir un compuesto de neumático diferente al Medio (debido a la regla

de que un auto debe usar al menos dos compuestos de neumáticos diferentes). Aunque el jugador B es más rápido que el jugador A entre las vueltas 30 y 40, esto cambia después de la vuelta 40, ya que el compuesto Blando comienza a degradarse. Como resultado, el jugador A gana la carrera por más de 4 segundos.

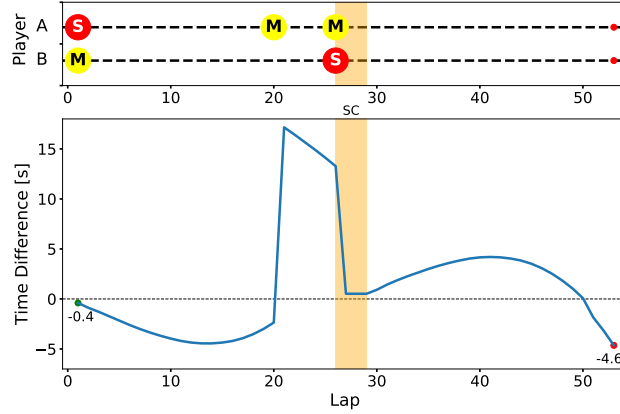


Figura E.1: Estrategias de carrera de los jugadores (Arriba), diferencia parcial de tiempo de carrera (Abajo), escenario con SC en las vueltas 26, 27 y 28.

Anexo F. Equilibrio de Nash al Maximizar la Probabilidad de Ganar

La Tabla G.1 muestra el objetivo esperado de suma-cero en la primera vuelta para todas las combinaciones de compuestos iniciales de neumáticos. Es importante destacar que, debido a la construcción de la función objetivo, en la última vuelta (que es igual a -1 si el jugador A gana y 1 si el jugador B gana), uno más los valores de pago dentro de la Tabla G.1 dividido por dos es igual a la probabilidad de ganar para el jugador B . La probabilidad de que el jugador A gane, como se deriva en la Sección 2.4.4, se expresa como $(1 - (\pi^{A*})^T U \pi^{B*})/2$. Esta fórmula cuantifica la probabilidad de que el jugador A logre la victoria en el juego. Las probabilidades de equilibrio para la primera vuelta se proporcionan en la última fila y columna de la Tabla G.1. Utilizando estos valores, calculamos que el jugador A tiene una probabilidad de ganar del 67.16%. Es importante destacar que este resultado se alinea estrechamente con el resultado simulado (68.2%) que se muestra en la segunda fila de la Tabla 2.2 en la Sección 2.5. Para aprovechar el poder computacional del supercomputador NLHPC, se le pidió al modelo que simulara una carrera con 50,000 escenarios. Los resultados

mostraron que el jugador A ganó el 66.95 % del tiempo, lo cual está aún más cerca del valor esperado basado en las probabilidades de equilibrio de la primera vuelta. Para obtener más información, consulte el histograma en la Figura F.1.

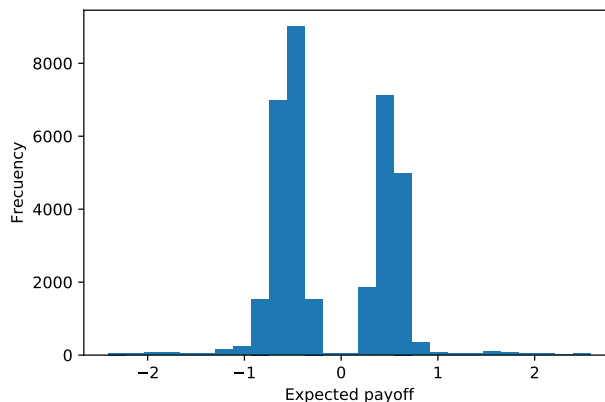


Figura F.1: Resultados de una carrera simulada utilizando el enfoque de maximización de la probabilidad de ganar con 50,000 escenarios.

Anexo G. Tablas Anexas

Tabla G.1: Pago esperado de los jugadores al comienzo de la carrera, U'' ; y Equilibrio de Nash mixto, $\pi^{A'*}$ y $\pi^{B'*}$

		Player B			$\pi^{A'*}$
Tire		Soft	Medium	Hard	
Player A	Soft	-0.14	-0.66	-0.64	1.0
	Medium	-0.48	-0.13	-0.65	0.0
	Hard	0.47	0.47	-0.13	0.0
$\pi^{B'*}$		0.61	0.39	0.0	

Tabla G.2: Número de estados y tiempo por vuelta para circuito de Silverstone, problema de programación dinámica.

Vuelta	Número de estados	Tiempo [s]
1	9	0
2	2,017,152	2
3	4,538,592	4
4	8,068,608	8
5	12,607,200	12
6	18,154,368	17
7	24,710,112	23
8	32,274,432	30
9	40,847,328	38
10	50,428,800	47
11	61,018,848	57
12	72,617,472	68
13	85,224,672	80
14	98,840,448	93
15	113,464,800	107
16	129,097,728	122
17	145,739,232	137
18	163,389,312	154
19	182,047,968	171
20	201,715,200	189
21	222,391,008	209
22	244,075,392	229
23	266,768,352	250
24	290,469,888	273
25	315,180,000	296
26	340,898,688	321
27	367,625,952	346
28	395,361,792	372
29	424,106,208	400
30	453,859,200	428

Tabla G.3: Número de estados y tiempo por vuelta para circuito de Silverstone, problema de programación dinámica.

Vuelta	Número de estados	Tiempo [s]
31	484,620,768	458
32	516,390,912	479
33	538,131,328	499
34	560,320,000	519
35	582,956,928	541
36	606,042,112	562
37	629,575,552	584
38	653,557,248	607
39	677,987,200	631
40	702,865,408	653
41	728,191,872	676
42	753,966,592	692
43	767,022,048	704
44	780,189,568	716
45	793,469,152	729
46	806,860,800	778
47	820,364,512	850
48	833,980,288	865
49	847,708,128	952
50	861,548,032	1016
51	875,500,000	1042
52	889,564,032	1046
TOTAL	20,478,351,241	5h 34' 23"