

UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA

DEPARTAMENTO DE INDUSTRIAS

ANÁLISIS ECONOMÉTRICO DE LA DEMANDA Y MODELADO DE CAPACIDAD:

DETERMINACIÓN DE RECURSOS ÓPTIMOS PARA LA ATENCIÓN DE

EMERGENCIAS EN EMPRESA DE GAS NATURAL

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL

AUTOR

DIEGO IGNACIO CABRERA SEGURA

PROFESOR GUÍA

BERNARDO PINCHEIRA SARMIENTO

VALPARAÍSO, 13 DE OCTUBRE, 2025



CONSTANCIA DE VALIDACIÓN Y CONFIDENCIALIDAD DE MONOGRAFÍA A REPOSITORIO ACADÉMICO

1.- IDENTIFICACIÓN DEL TRABAJO ACADÉMICO

Tipo de monografía (marcar una opción): Memoria o trabajo de título Tesis de Postgrado

Título del trabajo: Análisis Econométrico de la demanda y modelado de capacidad: Determinación de recursos óptimos para la atención de emergencia en empresa de gas natural

Nombre del candidato(a): Diego Ignacio Cabrera Segura

Carrera / Grado: Ingeniería Civil Industrial

Campus: Casa Central **Departamento:** Industrias

2.- VALIDACIÓN DEL PROFESOR GUÍA/DIRECTOR DE TESIS

Yo, Bernardo Pincheira Sarmiento, en mi calidad de profesor(a) guía/director(a) del trabajo académico mencionado anteriormente **DEJO CONSTANCIA** que:

- He revisado esta versión del documento y corresponde a la versión final aprobada del trabajo.
- El trabajo cumple con los requisitos académicos y de formato establecidos por la institución.

3.- EVALUACIÓN DE CONFIDENCIALIDAD POR PROPIEDAD INDUSTRIAL (marcar una opción)

El trabajo **NO contiene** información que amerite confidencialidad y puede ser publicado de inmediato en repositorio con acceso abierto.

El trabajo **CONTIENE** información con potenciales implicancias de propiedad industrial o intelectual y requiere un periodo de confidencialidad (**embargo**) por (**marcar una opción**):

6 meses 12 meses 2 años 3 años 5 años 10 años

Fundamentación de la necesidad de confidencialidad (obligatorio si se solicita embargo):

4.- FIRMAS

Profesor(a) guía o director(a) de memoria o tesis:

Fecha: 05/11/2025

Firma: 

Estudiante o Candidato(a):

Fecha: 05/11/2025

Firma: 

Este formulario debe ser insertado como página 2 de la memoria o tesis, completado y firmado por estudiante y profesor(a) antes de la entrega en portal PRISMA de Biblioteca USM.

Contenidos

1. Resumen Ejecutivo.....	6
2. Introducción	9
3. Problema de Investigación.....	10
4. Objetivos.....	13
4.1 Objetivo General.....	13
4.2 Objetivos Específicos	13
5. Marco Teórico	14
5.1 El Gas Natural.....	14
5.1.1 Características	14
5.1.2 Breve Historia del Gas Natural	15
5.2 Escenario Mundial.....	16
5.2.1 Principales Países Productores	16
5.2.2 Reservas y Extracción.....	17
5.2.3 Distribución y Consumo	17
5.3 Escenario en Chile.....	18
5.3.1 Identificación Empresas de Distribución y su Rol en la Atención de Emergencias.....	19
5.4 Recursos Involucrados en la Atención de Emergencias	21
5.5 Definición, Técnicas y Herramientas para el Análisis de Datos	22
5.6 Definición y Descripción de Modelos de Capacidad en la Gestión de Operaciones.....	25
5.7 Criterios de Desempeño y Medidas de Evaluación en la Gestión de Operaciones.....	28
6. Alcance	30
6.1 Alcance Descriptivo.....	30
6.2 Alcance Explicativo.....	31
6.3 Síntesis del Alcance	31
7. Metodología.....	32
7.1 Introducción	32
7.2 Tipo y Enfoque de Investigación	32
7.3 Diseño Metodológico	33
7.3.1 Organización y Limpieza de Datos	33

7.3.2	<i>Análisis Exploratorio de Datos</i>	37
7.3.3	<i>Análisis Econométrico</i>	46
7.3.4	<i>Estimación Probabilística de la Demanda</i>	61
7.3.5	<i>Simulación de Escenarios</i>	67
7.3.6	<i>Análisis de Costo-Beneficio.</i>	71
8.	Resultados	71
8.1	Introducción	71
8.2	Análisis Exploratorio de Datos	72
8.3	Análisis Econométrico	93
8.3.1	<i>Modelos de Clasificación Binaria</i>	93
8.4	Estimación Distribución de Probabilidad	107
8.5	Simulación de Demanda	109
8.5.1	<i>Resultados de la simulación</i>	110
8.5.2	<i>Conexión Entre la Simulación y la Planificación</i>	111
8.6	Simulación y Cálculo de capacidad	113
8.6.1	<i>Demanda Proyectada por Grupos de Comunas</i>	113
8.6.2	<i>Simulación de Capacidad Operativa</i>	114
8.6.3	<i>Capacidad Operativa Para Grupo de Comunas con Baja Demanda</i>	120
8.7	Análisis Costo-Beneficio	120
8.7.1	<i>Evaluación de Aumento de Móviles</i>	122
9.	Conclusiones y Recomendaciones	126
9.1	Conclusiones	126
9.2	Recomendaciones	128
10.	Discusión	130
11.	Limitaciones	131
12.	Referencias	132
13.	Apéndices	137

Índice de Figuras

Figura 1	<i>Componentes energéticos del Gas Natural de ser procesado</i>	15
Figura 2	<i>Países con mayor producción de Gas Natural</i>	17
Figura 3	<i>Distribución de llamadas por hora del día</i>	73
Figura 4	<i>Número de llamadas por comuna</i>	74
Figura 5	<i>Distribución por tipo de emergencia</i>	75
Figura 6	<i>Tiempos de respuesta por tipo de emergencia</i>	76
Figura 7	<i>Tiempo de respuesta promedio</i>	77
Figura 8	<i>Cumplimiento global</i>	78
Figura 9	<i>Distribución de los excesos en minutos</i>	80
Figura 10	<i>Porcentaje de llamadas no cumplidas según tipo de emergencia</i>	81
Figura 11	<i>Porcentaje de no cumplimiento de tiempos de respuesta por comuna</i>	83
Figura 12	<i>Llamadas que si cumplen con los tiempos de respuesta según tipo de emergencia</i>	84
Figura 13	<i>Nivel de cumplimiento por grupo de comunas</i>	87

Índice de Tablas

Tabla 1	<i>Estructura de datos</i>	72
Tabla 2	<i>Distribución de subconjuntos de llamadas incumplidas según exceso de tiempo de respuesta</i>	82
Tabla 3	<i>Distribución de llamadas por focalización</i>	85
Tabla 4	<i>Estadísticas descriptivas de tiempos de respuesta</i>	85
Tabla 5	<i>Distribución de llamadas y nivel de cumplimiento por grupo de comunas</i>	87
Tabla 6	<i>Distribución de llamadas por simultaneidad según ventanas de tiempo</i>	89
Tabla 7	<i>Probabilidad de simultaneidad de llamadas por comuna y tiempos</i>	90
Tabla 8	<i>Resumen de cumplimientos de tiempos de respuesta por comuna y focalización</i>	91
Tabla 9	<i>Comparativa de métricas para modelos Logit y Probit</i>	95
Tabla 10	<i>Comparativa de métricas para modelos Logit y Probit</i>	100
Tabla 11	<i>Comparativa de métricas para modelos Logit y Probit con balanceo de datos</i>	105
Tabla 12	<i>Capacidad simulada por nivel de servicio</i>	111
Tabla 13	<i>Capacidad simulada por focalización, cantidad de móviles y tiempos de respuesta</i>	114
Tabla 14	<i>Cantidad de móviles y porcentaje de cumplimiento para escenarios simulados</i>	117
Tabla 15	<i>Variación de tiempos de respuesta en escenarios simulados</i>	117
Tabla 16	<i>Cantidad de móviles y porcentaje de cumplimientos con tiempos de respuesta con 45 minutos para focalización Exterior y Medidor</i>	119
Tabla 17	<i>Costos anuales estimados por incorporación de un móvil para atención de emergencias</i>	121
Tabla 18	<i>Análisis costo-beneficio al aumentar móviles con estándar base de tiempos de respuesta</i>	122
Tabla 19	<i>Análisis costo-beneficio con variación a 45 minutos para focalización en Exterior y Medidor</i>	123

1. Resumen Ejecutivo

La atención oportuna de emergencias en la distribución de gas natural constituye un aspecto esencial tanto para la seguridad de los usuarios como para el cumplimiento de estándares en la organización. En este contexto, la dotación de móviles de emergencia se configura como un recurso crítico cuya planificación debe equilibrar eficiencia operativa, costos y niveles de servicio.

El presente estudio aborda esta problemática mediante el desarrollo de un marco metodológico que combina análisis exploratorio de datos, modelos econométricos, teoría de colas, simulación estocástica y evaluación costo–beneficio. A partir de registros históricos de llamadas de emergencia de una empresa distribuidora de gas en Chile, se construyó una base de datos que permitió caracterizar la demanda, identificar patrones de tipos de emergencias y focalización, así como estimar la probabilidad de ocurrencia y simultaneidad de eventos.

En la primera etapa, el análisis exploratorio de datos permitió describir el comportamiento histórico de las emergencias, con apoyo de herramientas como RStudio, Python y Excel, evidenciando diferencias significativas entre tipos de emergencia, distinta focalización de llamadas (interior, exterior/medidor y sin suministro), carga horaria, patrones de distribución por comunas y patrones de cumplimiento de tiempos de respuesta y simultaneidad de eventos. Posteriormente, se desarrollaron modelos econométricos Logit y Probit para estimar la probabilidad de ocurrencia de emergencias y de cumplimiento de los estándares establecidos, generando un marco cuantitativo robusto.

Adicionalmente, se ajustaron distribuciones de probabilidad a la cantidad de llamadas diarias mediante técnicas de máxima verosimilitud, seleccionando aquellas con mejor ajuste (ej. Binomial Negativa, Lognormal, Gamma) y validándolas con criterios de información (AIC).

Estos resultados fueron la base para la simulación del tipo Monte Carlo, que permitió generar escenarios de demanda futura y evaluar su impacto en el cumplimiento de tiempos de respuesta.

La teoría de colas se integró en la modelación de la capacidad operativa, permitiendo calcular tiempos de espera y de servicio bajo distintos niveles de dotación de móviles. La simulación de capacidad mostró que, para emergencias de interior y sin suministro, el cumplimiento de los estándares de tiempos establecidos se logra con dotaciones de cuatro y tres móviles, respectivamente. Sin embargo, en exterior/medidor se identificó el principal cuello de botella: incluso con seis móviles el promedio supera ligeramente el estándar de 30 minutos. No obstante, el análisis de reducción de brecha de tiempos demostró mejoras sustantivas: con cuatro móviles la brecha se reduce en un 67%, con cinco en un 80% y con seis en un 83%. Asimismo, bajo un estándar de sensibilidad de 45 minutos, el cumplimiento global alcanza el 100% con cinco móviles.

Desde el punto de vista económico, el análisis costo–beneficio evidenció que el salto de tres a cuatro móviles constituye la alternativa más costo–efectiva, con un costo cercano a \$1,9 millones anuales por punto porcentual de mejora en cumplimiento. El quinto móvil no incrementa el cumplimiento general en porcentaje bajo la norma de 30 minutos para los llamados con focalización en exterior/medidor, pero sí reduce los tiempos de espera en más de un 80%. El sexto móvil, en cambio, aporta un beneficio marginal frente a un costo elevado.

Finalmente, se incorporó un análisis de riesgo asociado a la flexibilización de los estándares (40 y 45 minutos), destacando la necesidad de ponderar el ahorro de recursos frente a los riesgos regulatorios, reputacionales y de seguridad.

En conclusión, el estudio aporta un marco replicable para la toma de decisiones en la gestión de emergencias, integrando herramientas estadísticas, econométricas, de gestión de operaciones en caso de estimación de distribuciones de probabilidad y teoría de colas, simulación y análisis económico. Los resultados permiten recomendar la adecuación de pasar de tres a cuatro móviles como dotación base, la consideración de un quinto móvil en escenarios de alta demanda o con estándares más flexibles, y la reserva de quizás un sexto móvil solo para contingencias extraordinarias. De este modo, se ofrece a la organización y el proceso un soporte técnico para equilibrar costos, seguridad y cumplimiento de tiempos de respuesta en la gestión de emergencias de gas natural.

2. Introducción

La atención de emergencias constituye un componente crítico en la operación de las empresas de distribución de gas natural, dado que la seguridad de los usuarios y la continuidad del servicio dependen en gran medida de la rapidez y eficacia de la respuesta. En este contexto, la planificación adecuada de los recursos móviles se transforma en un factor decisivo para reducir riesgos, mejorar la satisfacción del cliente y cumplir con los estándares establecidos. Pese a la relevancia del tema, la gestión de emergencias en el sector suele carecer de modelos analíticos integrales que permitan planificar la demanda y asignar recursos. La evidencia demuestra que la mala gestión puede generar consecuencias negativas como incrementar los tiempos de respuesta y el riesgo de incidentes graves.

Esta investigación aborda dicha problemática mediante un enfoque cuantitativo que integra análisis exploratorio de datos, modelos econométricos y simulación de escenarios. El análisis econométrico se emplea para identificar los factores que determinan la ocurrencia y el cumplimiento de los tiempos de respuesta en emergencias, mientras que la simulación permite evaluar configuraciones alternativas de capacidad y sus efectos sobre el desempeño operativo. Se utilizan registros históricos de llamadas de emergencia de una empresa distribuidora de gas natural. A partir de esta base de datos se desarrollan modelos, estimaciones y herramientas de simulación que permiten proponer una metodología para determinar la cantidad óptima de móviles en distintos contextos de demanda. Esta información busca aportar tanto a la práctica empresarial como al ámbito académico, con un marco de análisis replicable en otras organizaciones de servicios críticos. Los resultados buscan entregar una propuesta concreta para mejorar la gestión de recursos móviles en emergencias de gas natural, conciliando la eficiencia operacional y la económica.

3. Problema de Investigación

En el contexto de la gestión de operaciones de una empresa de distribución de gas natural, la eficiencia en la atención de llamadas de emergencia es crucial para garantizar la seguridad operativa, la satisfacción del cliente y el cumplimiento de los estándares de servicio. Sin embargo, en la actualidad, se observa una ausencia de metodologías estructuradas y basadas en datos para la planificación y asignación de los recursos móviles destinados a la atención de estas emergencias. Esta carencia de análisis cuantitativo y predictivo se traduce en una gestión poco fiable, lo que puede generar ineficiencias operativas significativas.

Por un lado, una asignación subóptima podría resultar en tiempos de respuesta prolongados ante situaciones críticas, comprometiendo la seguridad y la calidad del servicio. Tal como menciona, Muñiz (2019), para una central de atención de emergencias, es importante contar con información geográfica actualizada para identificar fácilmente la procedencia de las llamadas de emergencia y así optimizar la distribución y despacho de recursos.

Además, la falta de una atención oportuna puede generar más llamadas por parte de los usuarios, lo cual incrementa los tiempos de respuesta (Lara, 2022).

Por otro lado, una sobreestimación de la demanda podría llevar a un uso ineficiente de los recursos disponibles, incrementando los costos operativos sin una mejora proporcional en la capacidad de respuesta. Esto resalta la necesidad de considerar otros aspectos y así “evaluar alternativas desde una perspectiva sistémica, es decir, incorporar planificación de recursos e infraestructura energética” (Prieto y Simonin, 2014, p.1).

A lo anterior se suma un caso particularmente grave que evidencia las consecuencias humanas y materiales que pueden surgir ante una gestión deficiente de emergencias relacionadas con gas. En abril de 2018, una explosión ocurrida en el Sanatorio Alemán de Concepción, provocada por una fuga de gas durante labores de revisión, dejó un saldo de tres personas fallecidas y alrededor de cincuenta heridos. La emergencia se desencadenó justo cuando los pacientes estaban siendo reingresados al edificio, tras una evacuación preventiva, lo que aumentó el impacto del siniestro. La onda expansiva no solo destruyó parte de la infraestructura clínica, sino que también causó daños en edificios cercanos en un radio de hasta tres cuadras. Este trágico episodio expone los riesgos asociados a una detección tardía o una respuesta inadecuada frente a emergencias de este tipo, y refuerza la importancia de contar con sistemas de asignación de recursos móviles que aseguren tiempos de reacción óptimos. Casos como este justifican la necesidad de implementar modelos cuantitativos de planificación que permitan minimizar los tiempos de respuesta y prevenir catástrofes de gran escala en el sector de distribución de gas natural (Swissinfo.ch, 2018).

La variabilidad inherente a la demanda de atención de emergencias, influenciada por factores como la hora del día, el día de la semana, las condiciones climáticas e incluso eventos específicos ocurridos, acentúa la necesidad de un enfoque dinámico y adaptable con un respaldo de análisis econométrico en la gestión de los recursos móviles. La falta de un modelo de capacidad que considere un análisis de datos con sus fluctuaciones y permita determinar la demanda impide una planificación proactiva y una asignación de recursos eficiente. Tal como se evidencia en la industria del gas y petróleo, "los eventos críticos originados en las últimas décadas han dejado graves e irreversibles daños en las personas, ambiente y en la continuidad de los negocios" (Espinosa, 2018, p.7), lo cual subraya la importancia de una gestión de emergencias eficiente.

En este sentido, surgen distintas interrogantes sobre cómo incorporar herramientas analíticas y modelos cuantitativos, específicamente la econometría y la gestión de operaciones a través de modelos de capacidad, para optimizar la asignación de recursos móviles en la atención de llamadas de emergencia en una empresa de distribución de gas natural. Entre ellas destacan las siguientes: ¿Qué técnicas econométricas son apropiadas para analizar los patrones históricos de demanda y determinar las necesidades futuras? ¿Cómo diseñar un modelo de capacidad que, a partir del análisis de datos históricos y técnicas econométricas, permita estimar la cantidad óptima de móviles requeridos en distintos periodos de tiempo, garantizando niveles adecuados de cumplimiento en los tiempos de respuesta y optimizando el uso de los recursos disponibles? Asimismo, ¿qué beneficios concretos, en términos de eficiencia operativa y económica, pueden obtenerse al implementar un modelo de estas características y cómo contribuye a la mejora del servicio de atención de emergencias?

En síntesis, este estudio busca ofrecer métricas econométricas clave y un modelo de capacidad que, mediante la integración de análisis estadístico y simulación de escenarios, permita mejorar la asignación de recursos móviles en la atención de emergencias, contribuyendo así a la eficiencia del proceso.

4. Objetivos

4.1 Objetivo General

Mejorar la eficiencia operativa en la atención de emergencias de una empresa de distribución de gas natural, mediante la aplicación de herramientas estadísticas, econométricas y de gestión de operaciones que permitan estimar la cantidad óptima de recursos móviles y facilitar la interpretación de resultados.

4.2 Objetivos Específicos

Analizar de forma descriptiva y exploratoria los datos históricos de llamadas de emergencia en la empresa de distribución de gas natural, con el fin de identificar patrones, tendencias y características que sirvan de base para el análisis econométrico.

Realizar un análisis econométrico de los datos de atención de emergencias, que incluya la selección y aplicación de técnicas para el tratamiento de datos, el modelado de la demanda y la predicción de su comportamiento futuro.

Estimar la cantidad óptima de recursos móviles requeridos para la atención de emergencias, integrando los resultados del análisis econométrico y las estimaciones probabilísticas de la demanda, considerando las restricciones operativas y los niveles de servicio deseados.

Evaluar el desempeño del sistema de atención de emergencias mediante la simulación de diferentes escenarios de demanda y disponibilidad de recursos móviles.

5. Marco Teórico

5.1 El Gas Natural

El gas natural es una mezcla de gases ligeros, principalmente compuesta por metano (CH_4), que se forma de manera natural en el subsuelo por la descomposición de materia orgánica durante millones de años. Se encuentra generalmente en yacimientos profundos, ya sea en estado libre o asociado a otros hidrocarburos líquidos como el petróleo (Aprende con Energía, s.f.).

5.1.1 Características

Entre sus características más relevantes, destaca por ser un combustible fósil de combustión más limpia, ya que al quemarse produce menos emisiones contaminantes que otros hidrocarburos. Por ejemplo, genera un 30% menos dióxido de carbono (CO_2) que el petróleo y un 45% menos que el carbón, lo que lo posiciona como un combustible de transición hacia sistemas energéticos más sostenibles (International Energy Agency, 2023)

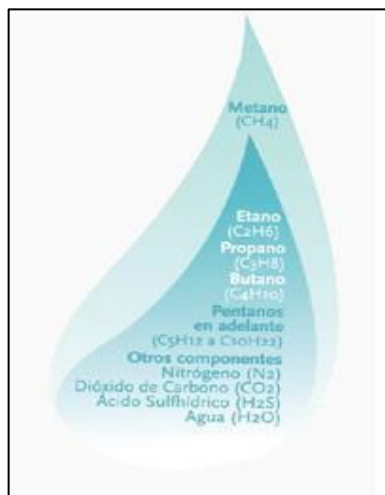
El gas natural se destaca por presentar una combinación de distintos compuestos como se observa en la Figura 1. No tiene olor ni color en su estado puro, por lo que, para su uso doméstico o industrial, se le añade una sustancia llamada mercaptano que le otorga un olor característico para facilitar su detección en caso de fugas. Además, es más liviano que el aire lo que significa que su densidad relativa es menor a 1, esto que permite que se disipe rápidamente en caso de una filtración (Gasco Educa, s.f.).

A nivel de infraestructura, su distribución requiere redes de cañerías subterráneas que permiten el suministro continuo a los usuarios, diferenciándose así del gas licuado de petróleo (GLP), que se transporta en cilindros. Esta característica le otorga ventajas en términos de comodidad, seguridad y eficiencia para su uso residencial, comercial e industrial.

Gracias a su versatilidad, el gas natural se utiliza ampliamente en sistemas de calefacción, cocinas, generación eléctrica, procesos industriales e incluso en el transporte mediante sistemas de gas natural vehicular. Su alto poder calorífico, bajo impacto ambiental y suministro estable lo han consolidado como una de las principales fuentes de energía a nivel global. ([AGN] Chile, 2022).

Figura 1

Componentes energéticos del Gas Natural de ser procesado



Nota. La figura representa los componentes del Gas natural. Tomado de Gasco Educa, s.f.

5.1.2 Breve Historia del Gas Natural

El gas natural es una fuente energética fósil compuesta principalmente por metano, que se ha utilizado desde la antigüedad. En China, hacia el año 900 a.C., ya se empleaban tubos de bambú para canalizarlo, y se reporta que en el 211 a.C. se perforó el primer pozo de gas de unos 150 metros de profundidad. En Europa, el gas fue descubierto en Inglaterra en 1659 pero por su complejidad para transportar su uso fue retrasado en 200 años, y en América, su primera explotación data de 1821 en Fredonia, Nueva York, donde se usó para iluminar calles mediante rudimentarias redes de distribución.

Durante el siglo XX, el gas natural evolucionó desde ser una fuente secundaria hasta convertirse en un pilar de la matriz energética global, esto gracias al desarrollo industrial. Su expansión fue impulsada por la construcción de redes de transporte mediante gasoductos y por su creciente competitividad frente a otros combustibles fósiles. Tras la Segunda Guerra Mundial, los avances en almacenamiento y transporte permitieron su uso masivo, particularmente en generación eléctrica, calefacción e industria (Yergin, 2009).

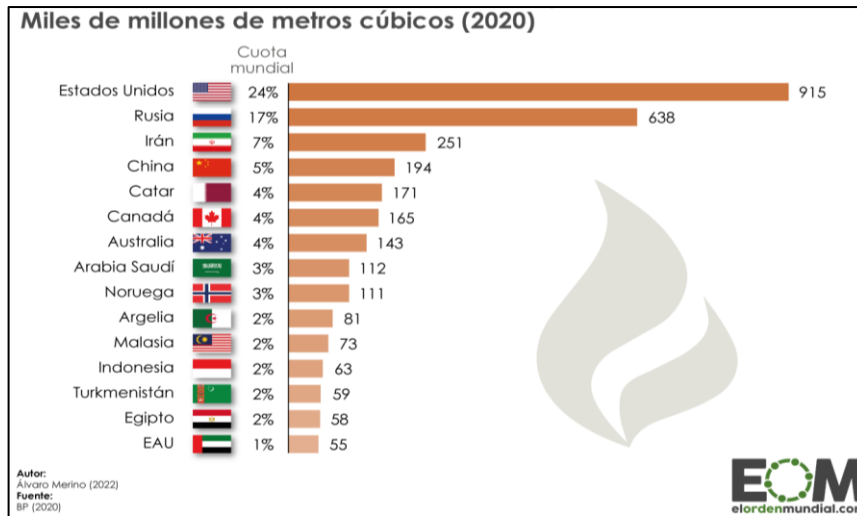
En la actualidad, el gas natural representa aproximadamente el 23% del consumo de energía primaria a nivel mundial, posicionándose como la tercera fuente energética más utilizada después del petróleo y el carbón.

5.2 Escenario Mundial

El gas natural por sus características como alta eficiencia, menor nivel de emisiones contaminantes y versatilidad en usos industriales, residenciales y de generación eléctrica lo han posicionado como un componente clave en la matriz energética global y un recurso estratégico en el contexto de la transición energética.

5.2.1 Principales Países Productores

Actualmente, la producción de gas natural se encuentra fuertemente concentrada en un grupo reducido de países ver Figura 2. Estados Unidos encabeza la lista como el mayor productor global, en gran parte debido al desarrollo del gas no convencional o shale gas. Le sigue Rusia, que posee extensas reservas y una sólida infraestructura de exportación. En tercer lugar, se encuentra Irán, con importantes recursos naturales, aunque su potencial se ve limitado por restricciones internacionales (Aprende con Energía, s.f.; El Orden Mundial, 2022).

Figura 2*Países con mayor producción de Gas Natural*

Nota. La figura representa la producción de gas natural en el mundo identificado por países productores. Tomado de El Orden Mundial, 2022.

5.2.2 Reservas y Extracción

Las reservas de gas natural están distribuidas en todos los continentes, aunque las más significativas se concentran en Rusia, Irán, Catar y Turkmenistán. En América, Estados Unidos y Venezuela cuentan con importantes volúmenes, mientras que en África destaca Nigeria. En el Cercano Oriente, países como Arabia Saudita y Emiratos Árabes Unidos también poseen reservas relevantes. No obstante, la existencia de reservas no garantiza su explotación inmediata, ya que factores técnicos, económicos y geopolíticos pueden limitar su aprovechamiento efectivo (Energía y Sociedad, s.f.).

5.2.3 Distribución y Consumo

El gas natural se utiliza en una amplia gama de sectores: residencial, industrial, comercial, transporte y generación eléctrica.

Su consumo ha aumentado considerablemente en regiones como Asia, donde ha comenzado a sustituir al carbón en muchas aplicaciones. En Europa, sigue siendo fundamental para la calefacción y la producción de energía, mientras que Estados Unidos, Rusia, China e Irán figuran entre los principales consumidores globales.

Una de las claves para su expansión ha sido el desarrollo del gas natural licuado (GNL), que permite transportar gas a largas distancias mediante buques, posibilitando el acceso al recurso en países sin gasoductos. Esta tecnología ha sido determinante para mercados como el chileno, donde no existen conexiones físicas con los grandes centros productores (Aprende con Energía, s.f.). El gas natural forma parte de una industria global estratégica, con un mercado dinámico que depende de factores geopolíticos, tecnológicos y ambientales. Este desequilibrio ha sido acentuado por eventos climáticos extremos y una limitada expansión del gas natural licuado (GNL). Se espera una mejora hacia fines de la década, con el ingreso de nueva capacidad exportadora liderada por Estados Unidos y Catar (International Energy Agency, 2023).

5.3 Escenario en Chile

En Chile, el gas natural es una fuente energética complementaria dentro de la matriz nacional, con un uso creciente en sectores como el residencial, industrial y transporte. A diferencia de otros países productores, Chile no posee yacimientos significativos de gas natural, por lo que depende totalmente de la importación para abastecer su demanda interna.

El ingreso del gas natural al país se realiza principalmente por vía marítima, en forma de gas natural licuado (GNL), a través de terminales como Quintero (Región de Valparaíso) y Mejillones (Región de Antofagasta). Allí el GNL es regasificado y distribuido por redes hacia las zonas urbanas e industriales (Gasco Educa, s.f.).

La distribución del gas natural en Chile se concentra en áreas urbanas de la zona centro, norte y sur del país, donde operan empresas concesionarias como Gasco, Lipigas, Metrogas, GasValpo y GasSur. El suministro se realiza mediante redes de cañerías subterráneas, lo que permite un suministro continuo y seguro para los usuarios. Las principales aplicaciones del gas natural en el país incluyen su uso para calefacción, agua caliente sanitaria, procesos industriales térmicos y en generación eléctrica.

El gas natural ha sido promovido como una fuente de energía de transición, dado que genera menores emisiones de CO₂ en comparación con el carbón y el petróleo. No obstante, el desarrollo del gas natural en Chile enfrenta desafíos estructurales, la competencia con otras fuentes como el gas licuado de petróleo (GLP) y las energías renovables. A pesar de ello, sigue siendo una fuente clave para asegurar el suministro energético en zonas densamente pobladas y para apoyar la transición hacia un sistema energético más limpio y resiliente ([CNE], 2023; Gobierno de Chile, 2020).

5.3.1 Identificación Empresas de Distribución y su Rol en la Atención de Emergencias

Una empresa de distribución de gas natural es la responsable de transportar y entregar el gas desde los puntos de regasificación o procesamiento hasta los consumidores finales, garantizando un suministro continuo, seguro y eficiente. Además de sus funciones técnicas y comerciales, estas empresas tienen un rol esencial en la prevención y atención de emergencias, debido a los riesgos inherentes al transporte y uso del gas natural. Las características claves de una empresa distribuidora de gas natural incluyen:

- **Red de infraestructura crítica:** Operan redes subterráneas de cañerías, estaciones de regulación de presión, válvulas, medidores y puntos de inyección de odorante.

Esta red debe estar permanentemente monitoreada, ya que cualquier falla, rotura o fuga puede generar eventos de alta peligrosidad como explosiones, incendios o intoxicaciones (Gasco Educa, s.f.).

- **Sistema de atención de emergencias:** Estas empresas deben contar con equipos móviles de respuesta disponibles las 24 horas, todos los días del año, para atender rápidamente reportes de emergencias de la ciudadanía o de sistemas automatizados. La velocidad de respuesta es crítica para mitigar riesgos a la seguridad de las personas y bienes, y forma parte de los indicadores de desempeño exigidos por la regulación ([SEC], 2014).
- **Capacidad operativa y logística:** La gestión eficiente de emergencias depende directamente de contar con una dotación óptima de móviles, personal técnico capacitado y cobertura territorial. Este aspecto es clave, ya que la falta de planificación de recursos puede traducirse en tiempos de respuesta prolongados, poniendo en riesgo la seguridad y aumentando los costos operativos (CNE, 2023).
- **Relación directa con el usuario y protocolos de seguridad:** Además de la atención técnica, las distribuidoras deben mantener canales de comunicación eficientes para la recepción de llamados, entrega de información preventiva y activación de planes de contingencia. También están obligadas a realizar inspecciones periódicas, capacitaciones y campañas sobre el uso seguro del gas en instalaciones domiciliarias.
- **Regulación estricta en gestión de emergencias:** La normativa chilena supervisada por la Superintendencia de Electricidad y Combustibles (SEC) exige a las distribuidoras protocolos formales de atención de emergencias, tiempos de respuesta definidos, trazabilidad de cada evento y reportes obligatorios ante incidentes que afecten a personas o infraestructura (SEC, 2014).

En conjunto, una empresa de distribución de gas natural no solo debe asegurar el suministro, sino también actuar como un operador crítico en la gestión de riesgos, especialmente en zonas urbanas de alta densidad. Este rol refuerza la necesidad de contar con herramientas para mejorar la eficiencia de la atención ante eventos críticos.

5.4 Recursos Involucrados en la Atención de Emergencias

La atención de emergencias en una empresa de distribución de gas natural implica la movilización coordinada de diversos recursos humanos, materiales y tecnológicos, cuyo objetivo es garantizar una respuesta oportuna, segura y eficiente ante situaciones críticas, como fugas, explosiones, fallas en la red o distintos reportes informados por los clientes.

Desde una perspectiva organizacional, los recursos humanos constituyen un eje fundamental del sistema de respuesta. Este grupo incluye al personal encargado de recepcionar las llamadas de emergencia, registrar los incidentes, despachar los móviles y ejecutar la atención directa en terreno. En distintos ámbitos se indica que los sistemas modernos de gestión de emergencias requieren no solo de operadores bien entrenados, sino también de estructuras organizativas que aseguren la coordinación entre funciones administrativas, técnicas y operativas (Haddow et al., 2016). A estos se suman los recursos materiales y tecnológicos, que hacen posible la operación del sistema. Desde el punto de vista normativo, la legislación chilena establece exigencias claras respecto a la disponibilidad y calidad de estos recursos. El Decreto Supremo Número 280 (2010), que aprueba el reglamento del servicio de distribución de gas de red, establece que las empresas concesionarias deben contar con los medios técnicos y humanos necesarios para atender emergencias, interrupciones y fallas operativas en forma continua y oportuna (Ministerio de Energía, 2010, arts. 19, 20 y 43).

Por su parte, el Decreto Supremo Número 66 (2007) exige que las instalaciones de consumo de gas cuenten con medidas de seguridad adecuadas y que las empresas involucradas en su operación establezcan protocolos y procedimientos para responder a situaciones de emergencia, tanto en instalaciones domiciliarias como industriales (Ministerio de Economía, 2007).

En conjunto, la atención de emergencias requiere una gestión integrada de múltiples tipos de recursos cuya disponibilidad, coordinación y capacidad operativa son determinantes en la eficacia de la respuesta. La falta de dimensionamiento adecuado puede derivar en tiempos de atención prolongados, riesgos a la seguridad pública y uso ineficiente de los recursos. En esta línea, el presente estudio se orienta a alcanzar dicho propósito, a través de un enfoque que integra análisis econométrico y modelación operativa.

5.5 Definición, Técnicas y Herramientas para el Análisis de Datos

El análisis de datos es un proceso sistemático que permite examinar, organizar, transformar e interpretar conjuntos de datos con el fin de extraer información útil, identificar patrones y respaldar la toma de decisiones. Dependiendo del objetivo del estudio, este análisis puede adoptar un enfoque descriptivo, exploratorio o predictivo, utilizando diversas herramientas estadísticas. En el contexto de las ciencias aplicadas, el análisis de datos constituye una etapa esencial para comprender fenómenos complejos a partir de información real.

Los datos disponibles para este caso tienen por definición el poder caracterizar la demanda de atención de emergencias en una empresa de distribución de gas natural, para luego sustentar un modelo de capacidad que permita mejorar la asignación de recursos.

Dado que la base de datos contiene observaciones con estructura temporal (fechas, horas) y espacial (por comuna), y que cada registro puede estar asociado a características específicas del evento (tipo de emergencia, clasificación, tipo de instalación, entre otros), se considera adecuado el enfoque de modelos econométricos.

Cabe señalar que, por razones de confidencialidad y resguardo de información sensible, el nombre de la empresa de gas natural que proporcionó la base de datos no será revelado en este documento. No obstante, los datos utilizados son reales, actualizados y representan fielmente las operaciones de una empresa del sector en Chile, permitiendo realizar un análisis riguroso y contextualizado.

En una primera etapa exploratoria, se observarán de forma más general técnicas de análisis descriptivo para examinar la frecuencia, distribución y comportamiento de las variables clave. Esto incluirá tablas de frecuencia por comuna y tipo de evento, distribución horaria de las llamadas, análisis de días críticos y detección de valores atípicos.

El análisis de datos constituye una etapa esencial en estudios aplicados a la gestión de emergencias, pues permite comprender el comportamiento histórico de la demanda y generar información que apoyen la planificación operativa.

Entre estas herramientas destacan los modelos de datos de panel, que integran información de múltiples unidades de observación a lo largo del tiempo, combinando variación transversal y temporal (Gujarati, 2010). Este enfoque resulta especialmente útil cuando las unidades presentan características particulares que pueden influir en los resultados, como ocurre con comunas o zonas operativas en un servicio de distribución de gas natural.

Dentro de los modelos de panel, se distinguen los efectos fijos y los efectos aleatorios. Los primeros permiten controlar por heterogeneidad no observada específica de cada unidad, bajo el supuesto de que dichas características están correlacionadas con las variables explicativas. Los segundos, en cambio, asumen que estas diferencias son aleatorias y no correlacionadas con los regresores, lo que puede ofrecer estimaciones más eficientes si se cumplen sus supuestos. La elección entre ambos enfoques suele resolverse mediante la prueba de Hausman, que contrasta la posible correlación entre los efectos no observados y las variables explicativas (Wooldridge, 2009).

La literatura también reconoce los modelos dinámicos de panel, que incorporan rezagos de la variable dependiente y permiten capturar efectos de persistencia o acumulación en el tiempo. Para la estimación de este tipo de modelos se emplea frecuentemente el Método Generalizado de Momentos (GMM), técnica que corrige problemas de endogeneidad y heterocedasticidad y que ha demostrado ser especialmente útil en contextos donde la demanda presenta fluctuaciones relevantes (Wooldridge, 2009).

Por otro lado, cuando el interés se centra en analizar la probabilidad de ocurrencia de determinados eventos, los modelos Probit y Logit constituyen herramientas econométricas ampliamente utilizadas. Estos modelos permiten vincular distintas variables explicativas como el tipo de emergencia, la localización geográfica, la hora del día con la probabilidad de que ocurra un evento crítico, lo que los convierte en técnicas pertinentes para modelar fenómenos dicotómicos en entornos operativos con incertidumbre.

El análisis de datos aplicado a la gestión de emergencias se apoya además en el uso de herramientas computacionales especializadas.

En este sentido, lenguajes como RStudio ofrecen robustez y una amplia gama de paquetes para el tratamiento estadístico y econométrico, mientras que programas como Excel permiten la exploración y organización inicial de los datos, y plataformas como Python resultan útiles para la validación de resultados y la visualización de escenarios.

En conjunto, estas técnicas constituyen la base teórica para la caracterización de la demanda de emergencias y la posterior construcción de modelos de apoyo a la toma de decisiones, los cuales buscan dimensionar adecuadamente los recursos disponibles y garantizar niveles de servicio más eficientes.

5.6 Definición y Descripción de Modelos de Capacidad en la Gestión de Operaciones

La planificación de capacidad es una función clave dentro de la gestión de operaciones, orientada a determinar la cantidad de recursos necesarios para atender una demanda esperada de manera oportuna, eficiente y con un nivel de servicio adecuado. En términos generales, la capacidad puede definirse como “la tasa máxima de producción o atención que un sistema puede alcanzar en un periodo determinado” (Krajewski y Ritzman, s.f.). En el contexto de los sistemas de respuesta a emergencias como el caso de la atención de llamadas por eventos asociados a gas natural, el problema de capacidad no solo implica contar con un número adecuado de móviles o personal técnico, sino también considerar variables como el tiempo de respuesta, la distribución geográfica de la demanda, y el comportamiento incierto del sistema.

Entre las herramientas más relevantes que constituyen específicamente modelos de capacidad se encuentran los modelos de cobertura, diseñados para optimizar la ubicación y cantidad de recursos que permitan maximizar la cobertura de la demanda dentro de un tiempo de respuesta determinado.

Estos modelos, utilizados comúnmente en la planificación de servicios de ambulancias, bomberos o policía, permiten decidir cuántos recursos disponer y dónde ubicarlos estratégicamente para mejorar la capacidad de atención. El Maximal Covering Location Problem (MCLP) es uno de los más conocidos en este ámbito, al permitir encontrar la mejor combinación entre cantidad de recursos disponibles y cobertura de la demanda (Toregas et al., 1971).

Otra técnica aplicable corresponde a los modelos de planificación de capacidad bajo incertidumbre, que buscan dimensionar recursos frente a escenarios variables de demanda. En este tipo de enfoques se integran elementos como la priorización de zonas críticas, restricciones presupuestarias y niveles de servicio deseado, permitiendo definir estrategias de respuesta más flexibles. Este tipo de planteamientos ha sido abordado en el contexto de múltiples ítems, tecnologías mixtas y procesos con recursos compartidos (Escalona y Ramírez, 2012).

Un elemento fundamental en los modelos para determinar capacidad es la estimación de la demanda, ya que constituye el insumo principal para dimensionar los recursos disponibles. La predicción mediante la identificación y ajuste de distribuciones de probabilidad a datos históricos permite anticipar escenarios futuros de carga y evaluar el desempeño del sistema bajo distintos supuestos de variabilidad (Ross, 2014). Este enfoque resulta especialmente útil en contextos de emergencias, donde la ocurrencia de eventos presenta fluctuaciones significativas y no puede ser tratada de manera determinista.

Asimismo, otra aproximación vinculada a los modelos de capacidad es la teoría de colas, la cual estudia sistemas donde múltiples unidades de demanda arriban a un conjunto limitado de servidores. Este marco teórico permite analizar tiempos de espera, niveles de utilización de los recursos y probabilidades de retraso, aspectos directamente relacionados con la gestión de emergencias (Gross & Harris, 1998; Kleinrock, 1975).

En este tipo de contextos, la dotación de móviles puede entenderse como servidores que atienden solicitudes bajo tasas de llegada variables, lo que refuerza la importancia de considerar enfoques estocásticos en la planificación de capacidad.

Por otro lado, modelos como el Machine Requirements Planning (MRP) originalmente desarrollados para entornos industriales podrían adaptarse conceptualmente al análisis de recursos móviles si se estructura el problema como una secuencia de tareas que deben atenderse bajo ciertas restricciones de capacidad y ventanas temporales. Aunque su uso no es frecuente en contextos de emergencias, su lógica de planificación puede ser útil como referencia en escenarios más estructurados o programables.

En conclusión, la literatura en gestión de operaciones ofrece diversas herramientas y modelos que pueden ser considerados para abordar el problema de dimensionamiento de capacidad en sistemas de atención de emergencias. Entre ellas destacan los modelos de cobertura, la planificación bajo escenarios inciertos, la estimación probabilística de la demanda, la teoría de colas y, eventualmente, enfoques adaptativos de planificación de recursos. Cada uno presenta ventajas y limitaciones según las características del sistema y el tipo de decisiones que se desean mejorar, lo que justifica el desarrollo de varios enfoques que integren distintos elementos para responder a contextos específicos como el de la distribución de gas natural.

5.7 Criterios de Desempeño y Medidas de Evaluación en la Gestión de Operaciones

Evaluar los modelos o métodos utilizados en la toma de decisiones operativas es un paso fundamental en el proceso de diseño e implementación de soluciones. En el caso de los modelos para dimensionar capacidad, cobertura u otro, la evaluación permite establecer criterios de desempeño que indiquen qué tan efectivo es un modelo para cumplir con los objetivos operacionales planteados, tales como minimizar el tiempo de respuesta, mejorar en el uso de recursos, garantizar niveles de servicio o maximizar la cobertura de demanda. Sin una adecuada evaluación, la utilidad del modelo queda reducida a su formulación teórica, sin evidencias sobre su aplicabilidad o eficiencia en escenarios reales o simulados (Krajewski & Ritzman, s.f.).

La evaluación de herramientas o modelos busca responder a preguntas como: ¿Qué nivel de demanda logra cubrir?, ¿Qué tan sensibles son los resultados ante cambios en la demanda o condiciones operativas?, ¿Se cumplen los estándares definidos por la operación o por normativas? En ese sentido, las medidas de evaluación permiten comparar alternativas, identificar puntos críticos y tomar decisiones informadas sobre la viabilidad y desempeño de la solución propuesta.

Entre las medidas más comunes para evaluar modelos de capacidad, cobertura y otros dentro de una categoría más específica de modelos de capacidad, destacan las siguientes:

- **Porcentaje de cobertura o nivel de servicio:** mide qué proporción de las llamadas o incidentes esperados pueden ser atendidos por los recursos disponibles dentro de un umbral de tiempo (por ejemplo, atender el 90% de las llamadas en menos de 15 minutos). Es una métrica general en modelos para empresas de emergencias y como base de los problemas como el Maximal Covering Location Problem (MCLP), desarrollado por Toregas et al. (1971).

- **Tasa de utilización de recursos:** permite evaluar si los recursos están siendo aprovechados. En modelos de capacidad, una utilización muy alta puede indicar riesgo de saturación, mientras que una muy baja puede implicar ineficiencia (Krajewski & Ritzman, s.f.).
- **Tiempo promedio de respuesta o atención:** se refiere al tiempo que transcurre desde que se registra un evento hasta que un recurso lo atiende. Esta medida es clave en servicios de emergencia y puede ser evaluada mediante simulaciones o datos históricos.
- **Costo de operación o relación costo-beneficio:** resulta relevante en modelos de capacidad bajo restricciones presupuestarias, donde se busca minimizar costos operativos manteniendo niveles de servicio. En un enfoque aplicado, esta medida puede simplificarse en términos de la conveniencia económica de aumentar o no la cantidad de recursos disponibles.
- **Robustez y sensibilidad:** evalúa cómo cambia el desempeño del modelo ante variaciones en la demanda, cambios en la localización de eventos, reducción o aumento de recursos. Este tipo de evaluación permite identificar la flexibilidad o fragilidad del modelo ante escenarios futuros.

Según el enfoque metodológico utilizado sea analítico, econométrico, de simulación u otro, las métricas pueden calcularse o estimarse mediante herramientas computacionales.

Esta evaluación se orienta principalmente hacia indicadores de tiempo de respuesta, nivel de servicio y análisis de sensibilidad, complementados con un análisis costo–beneficio de la incorporación de móviles adicionales, lo que asegura coherencia con los objetivos definidos para el estudio.

6. Alcance

La presente investigación se enmarca principalmente en los alcances descriptivo y explicativo, de acuerdo con la naturaleza del problema planteado y los objetivos definidos. Estos enfoques permiten, por un lado, caracterizar en detalle la demanda de emergencias y, por otro, explicar cómo la dotación de recursos móviles incide directamente en la eficiencia operativa y en el cumplimiento de los tiempos de respuesta.

6.1 Alcance Descriptivo

El proceso presenta un componente descriptivo, en la medida que se analizan y caracterizan los datos históricos de llamadas de emergencia en la empresa de distribución de gas natural. Este nivel permite identificar patrones, tendencias y características relevantes de la demanda, tales como:

- La distribución de emergencias por comuna, tipo de evento y otros.
- La focalización de los incidentes o no.
- La variación de la demanda.
- Los tiempos de respuesta registrados en la atención.

El alcance descriptivo entrega una visión integral del comportamiento de la demanda en el pasado, constituyendo la base sobre la cual se fundamenta el análisis econométrico y las proyecciones probabilísticas utilizadas en etapas posteriores.

6.2 Alcance Explicativo

La investigación también se sitúa en un alcance explicativo, ya que no se limita únicamente a describir patrones históricos, sino que busca determinar por qué y en qué medida la cantidad de recursos móviles disponibles influye en el nivel de cumplimiento de los tiempos de respuesta y en la eficiencia operativa.

Para lograrlo, se integran herramientas de análisis econométrico, distribuciones de probabilidad y simulación de escenarios, además de incorporando herramientas desde la gestión de operaciones. De esta manera, se explica cómo distintas configuraciones de recursos impactan en la atención de emergencias.

6.3 Síntesis del Alcance

En síntesis, el estudio combina un nivel descriptivo, orientado a la caracterización detallada de la demanda histórica, con un nivel explicativo, centrado en la identificación de factores y condiciones que determinan la eficiencia operativa del sistema de respuesta a emergencias. Esta doble perspectiva asegura una comprensión tanto del que ocurre en la operación como del por qué ocurre, en coherencia con el carácter aplicado de la investigación.

7. Metodología

7.1 Introducción

En este capítulo se presenta el enfoque metodológico que dio forma al desarrollo de esta investigación, la cual surge desde una necesidad concreta: mejorar la asignación de recursos móviles ante emergencias en una empresa de distribución de gas natural para para lograr una mejora operativa. Este apartado detalla el tipo de investigación, el diseño metodológico general, las técnicas utilizadas, la fuente de los datos y los recursos computacionales que permitieron estructurar el análisis. La metodología fue pensada para ofrecer una solución concreta, realista y útil desde una perspectiva cuantitativa, aplicada a un contexto crítico como lo es la atención de emergencias que pueden involucrar un alto riesgo.

7.2 Tipo y Enfoque de Investigación

Esta investigación se desarrolla bajo un enfoque cuantitativo, de tipo aplicado y con un diseño no experimental. Se basa en el análisis de datos reales y estructurados, extraídos del registro histórico de emergencias gestionadas por una empresa de distribución de gas natural. El estudio busca, por un lado, comprender qué factores inciden en la ocurrencia de ciertos tipos de emergencias y, por otro, proponer un enfoque que mejore la capacidad de respuesta en situaciones críticas a través de los tiempos de repuesta. El carácter aplicado se basa en su propósito concreto: entregar una propuesta útil para la toma de decisiones operativas. Para ello, se utilizarán modelos econométricos como Probit, Logit, que permiten capturar relaciones entre variables clave en un contexto donde los recursos, el tiempo de respuesta y la seguridad están directamente relacionados.

Además, se integran herramientas de la gestión de operaciones, tales como la identificación

de distribuciones de probabilidad, la estimación de la demanda futura y la simulación de escenarios de capacidad, con el fin de evaluar los niveles de cumplimiento alcanzables en distintos contextos de dotación de móviles.

7.3 Diseño Metodológico

El diseño metodológico de esta investigación se estructuró considerando tanto los objetivos planteados como la realidad operativa del servicio de atención de emergencias en una empresa de distribución de gas natural. El proceso se desarrolló en seis etapas principales, que abarcan desde la organización inicial de los datos hasta la simulación del modelo, incorporando además el análisis costo–beneficio asociado a la estimación de la cantidad de móviles requeridos bajo distintos escenarios.

7.3.1 Organización y Limpieza de Datos

La base de datos utilizada en esta investigación proviene de los registros históricos de emergencias atendidas por la empresa de distribución de gas natural, correspondientes al período comprendido entre los años 2022 y 2024. En su versión original, la planilla contenía aproximadamente 8.000 registros, estructurados en múltiples columnas, algunas de las cuales no resultaban relevantes para los objetivos del estudio. Entre las variables seleccionadas para el análisis se destacan las siguientes:

- **Dirección:** ubicación geográfica donde se origina el llamado de emergencia.
- **Fecha:** día que se generó el llamado de emergencia.

- **Comuna:** identifica la localización geográfica de la llamada. El conjunto de datos incluye registros de comunas como Valparaíso, Viña del Mar, Quilpué, Villa Alemana, Concón, Casablanca, Quillota, La Cruz, La Calera, Llay Llay, Los Andes, Coquimbo, La Serena, Talca, Rancagua, La Higuera, Catemu.
- **Tipo de emergencia:** clasifica la naturaleza del evento reportado. Los registros contemplan categorías como fuga de gas, olor a gas, sin suministro, Codec (Corte por defecto crítico), Inflamación, Detección de trabajos por empresas externas sobre redes de gas, Falla de instrumentos o equipos y otro tipo de emergencias.
- **Focalización:** describe el lugar específico donde se origina la emergencia respecto a las instalaciones. Se distinguen tres categorías principales: interior, sin suministro, exterior y medidor.
- **Clasificación:** se refiere a que si efectivamente se trató de una emergencia real. Se contemplan dos categorías, incidente y falsa alarma.
- **Hora de recepción:** momento en que la emergencia es registrada por el sistema de atención.
- **Hora de check in:** hora en que el móvil llega al lugar de la emergencia y atiende la emergencia.
- **Tiempo de respuesta:** diferencia de tiempo entre la hora de recepción y la hora de llegada del móvil al sitio de la emergencia.

Durante la limpieza inicial, se detectó que la planilla presentaba celdas vacías, inconsistencias de formato y registros duplicados. Por este motivo, se aplicaron los siguientes procedimientos:

- Eliminación de filas sin información mínima (ejemplo: registros sin hora de recepción o sin tipo de emergencia).

- Corrección de inconsistencias horarias (casos donde la hora de llegada era anterior a la hora de recepción).
- Depuración de duplicados.
- Estandarización de formatos de fecha y hora.

Tras la aplicación de estos procesos de depuración y estandarización, el número de registros disponibles se redujo respecto a la base original. Finalmente, se obtuvo un conjunto de 7.936 registros de llamadas de emergencia dentro del período de tiempo considerado, los cuales cumplieron con los criterios de completitud y consistencia requeridos para el análisis.

Asimismo, se generaron nuevas variables que facilitaron los análisis posteriores:

- **Tiempo de respuesta:** para estandarizar la variable tiempo de respuesta, se construyó en Excel una fórmula que permitió transformar automáticamente los registros, independiente del formato en que estuvieran almacenados. La fórmula aplicada fue:

```
=SI.ERROR(VALOR(E2),TIEMPO(IZQUIERDA(E2,ENCONTRAR(":",E2)-1),DERECHA(E2,LARGO(E2)-ENCONTRAR(":",E2)),0)).
```

Donde E2 corresponde a la celda con el valor original del tiempo de respuesta. En caso de que el dato se encontrara como número, se recuperaba directamente mediante la función VALOR; en caso contrario, se utilizaban funciones de texto (IZQUIERDA, DERECHA, ENCONTRAR, LARGO) para separar horas y minutos, y posteriormente la función TIEMPO para generar el valor en formato horario estándar. De esta forma, todos los tiempos de respuesta quedaron transformados a un formato numérico coherente y comparable para los cálculos posteriores.

- **Emergencia ocurrida:** creado para distinguir si en efecto ocurrió realmente una emergencia. Se creó como variable binaria (1 = incidente, 0 = falsa alarma). El código aplicado en Excel fue:

=SI (N2="Incidente", 1, 0).

- **Cumple tiempos de respuesta:** variable determinada según los tiempos máximos establecidos para cada tipo de emergencia y focalización: 30 minutos para emergencias en exterior o medidor, 60 minutos para emergencias en interior y 180 minutos para emergencias sin suministro y otras. Con ello se creó una variable binaria (1 = cumple, 0 = no cumple). El código utilizado en Excel fue:

=SI(O(M2="Medidor",M2="Exterior"),SI(F2<=30/1440,1,0),SI(M2="Interior",SI(F2<=60/1440,1,0),SI(M2="Sin Suministro", SI(F2<=180/1440,1,0),0))).

Donde la columna M2 corresponde a la focalización de la emergencia y F2 al tiempo de respuesta en formato minutos/días.

En síntesis, el proceso de organización y limpieza de datos permitió transformar una base dispersa e incompleta en un conjunto estructurado de 7.936 registros válidos, listos para el análisis. La estandarización de tiempos, la creación de variables y la eliminación de inconsistencias aseguraron que la información empleada en las siguientes etapas fuera coherente, confiable y representativa del comportamiento real de las emergencias registradas. Este paso, aunque de carácter preliminar, resulta fundamental para garantizar la validez de los análisis estadísticos, econométricos y de simulación para desarrollar en la investigación.

7.3.2 *Análisis Exploratorio de Datos*

Una vez obtenida la base de datos, el siguiente paso metodológico consiste en llevar a cabo un análisis exploratorio de datos. Esta etapa tiene como propósito describir y comprender el comportamiento histórico de las llamadas de emergencia, identificando patrones generales, posibles anomalías y relaciones entre variables. Esta etapa priorizó la estandarización de variables, la consistencia de categorías y la construcción de métricas comparables (tiempo de respuesta, cumplimiento y exceso de tiempo), acompañadas de visualizaciones que facilitan la validación y la trazabilidad.

Para esta etapa se empleará el software RStudio, seleccionado por su capacidad de manejar grandes volúmenes de datos, la disponibilidad de librerías específicas para análisis estadístico como: “dplyr”, “tidyr”, lubridate y la generación de gráficos de alta calidad mediante ggplot2. En comparación con Excel, RStudio permite automatizar procesos, reducir errores de manipulación manual y replicar los análisis de forma eficiente. El procedimiento del análisis exploratorio se desarrolló en los siguientes pasos:

7.3.2.1 Preparación y Consistencia de Categorías. Antes de definir y estandarizadas las variables base y métricas centrales, fue necesario realizar un trabajo de revisión y corrección que asegurara la consistencia de las categorías incluidas en la base de datos.

En el caso de la variable Comuna, se detectaron diferencias en la forma de escritura entre registros, principalmente por el uso de mayúsculas, tildes o abreviaturas (por ejemplo, “VINA DEL MAR” en lugar de “Viña del Mar”). Estas inconsistencias fueron corregidas para asegurar uniformidad, de manera que cada comuna se representara con un único valor válido y coherente.

Respecto a la variable tipo de emergencia, se identificaron etiquetas equivalentes o con variaciones menores de escritura. Para evitar que estas diferencias distorsionaran el conteo de emergencias, se consolidaron bajo categorías comunes como: fuga de gas, olor a gas, sin suministro, corte por defecto crítico (Codec), inflamación, detección de trabajos sobre redes, falla de instrumentos o equipos y otras emergencias. Asimismo, se creó la categoría “No especificado”, con el fin de mantener visibles los casos donde la información estaba ausente o registrada con valores irreconocibles, pero sin excluirlos del análisis general.

En la variable Focalización, se normalizó la existencia de cuatro categorías: interior, exterior, medidor y sin suministro. Esto permitió que el análisis de los tiempos de respuesta y del cumplimiento de los estándares pudiera ser segmentado de manera uniforme y sin distorsiones.

Finalmente, en la variable Clasificación, se verificó la correcta aplicación del indicador binario emergencia ocurrida, distinguiendo entre incidentes reales (1) y falsas alarmas (0). Esta validación fue clave para poder diferenciar entre llamadas que efectivamente representaron una situación de riesgo y aquellas que, aunque registradas, no implicaron una intervención operativa.

7.3.2.2 Análisis Exploratorio de Datos. En primer lugar, fue necesario preparar las variables que posteriormente serían utilizadas, a su vez se caracterizan algunas métricas necesarias para el análisis:

- i. **Normalización temporal:** se transformaron las variables de hora de recepción y hora de llegada del móvil al formato estándar de fechas y horas en RStudio, lo que permitió calcular diferencias temporales de forma consistente. Además, se generó una nueva variable denominada “hora del día”, que asigna a cada registro un valor entre 0 y 24 según la hora en la que se recibió el llamado.

-
- ii. **Cálculo del tiempo de respuesta:** a partir de las variables ya estandarizadas en Excel, se verificó en RStudio la diferencia en minutos entre la hora de llegada y atención y la hora de recepción, generando así la variable de tiempo de respuesta.
 - iii. **Distribución de llamadas por horario:** el análisis de la distribución total de llamadas según la hora del día permite identificar los horarios de mayor concentración de solicitudes. Esta información es clave para reconocer las franjas horarias críticas en las que se intensifica la demanda y, por lo tanto, resulta necesario reforzar la capacidad de respuesta.
 - iv. **Distribución de llamadas por comuna:** esta métrica desagrega la demanda histórica de llamadas a nivel comunal, lo que facilita apreciar la carga que enfrenta cada una. Con ello, es posible determinar de manera más precisa el peso operacional que asume cada comuna y orientar la planificación de recursos hacia aquellas que concentran mayor actividad.
 - v. **Distribución de llamadas por tipo de emergencia:** la determinación de esta métrica permite clasificar y cuantificar las llamadas según el tipo de emergencia reportada. De esta forma, se identifican cuáles son las categorías de incidentes que se presentan con mayor frecuencia dentro del total, lo que aporta una visión clara sobre los eventos más recurrentes y prioritarios para la gestión operativa.
 - vi. **Tiempo de respuesta por tipo de emergencia:** el uso de esta variable de estudio permite clasificar y cuantificar las llamadas según el tipo de emergencia reportada. Así se identifican cuáles son los incidentes más frecuentes dentro del total, aportando una visión clara sobre los eventos recurrentes y prioritarios para la gestión operativa.
 - vii. **Promedio de tiempos de respuesta:** este indicador permite obtener una visión global del desempeño en la atención de emergencias, a través del cálculo del tiempo promedio de respuesta considerando el total de llamadas registradas.

Complementariamente, se incorporan otras medidas estadísticas descriptivas como mínimo, máximo, mediana y desviación estándar con el fin de caracterizar de manera más completa la distribución de los tiempos.

- viii. Definición de cumplimiento:** se utilizó la variable binaria “cumple tiempos de respuesta”, que asigna un valor de 1 si el móvil llegó dentro del tiempo máximo permitido para cada tipo de emergencia, y 0 en caso contrario. Los estándares considerados fueron: 30 minutos para emergencias ocurridas en exterior o medidor, 60 minutos para emergencias en interior y 180 minutos para emergencias por sin suministro y otras.
- ix. Cálculo del exceso de tiempo en minutos sobre el estándar:** para aquellos registros en que el tiempo de respuesta fue mayor al estándar correspondiente, se calculó un valor denominado “exceso”, que corresponde a la cantidad de minutos adicionales que se tardó en atender la emergencia más allá del límite establecido. Este exceso se clasificó en rangos de 0 a 10 minutos, de 10 a 20 minutos, de 20 a 30 minutos, de 30 a 60 minutos y más de 60 minutos.
- x. Cumplimiento global:** a partir de estas variables se calcularon indicadores globales de desempeño, tales como el número total de llamadas registradas en la base de datos, la cantidad de emergencias que cumplieron con los tiempos de respuesta, la cantidad de emergencias que no cumplieron, y el porcentaje de cumplimiento y no cumplimiento respecto del total de emergencias analizadas.
- xi. Desglose por tipo de emergencia:** posteriormente, se realizó un desglose de los casos de no cumplimiento según el tipo de emergencia (fuga de gas, olor a gas, sin suministro, entre otros). Esto permitió identificar qué categorías presentaban mayores dificultades para ser atendidas en el tiempo establecido.

Para cada tipo de emergencia se calcularon tanto los conteos absolutos de no cumplimiento como los porcentajes que representan sobre el total de llamadas. Además, se elaboraron tablas

con los rangos de exceso de tiempo (0–10, 10–20, etc.) dentro de cada tipo de emergencia, lo que aportó información sobre la gravedad relativa del incumplimiento en cada categoría.

- xii. Desglose por comuna:** se calcularon tasas de no cumplimiento por comuna, considerando para cada localidad el total de llamadas recibidas, la cantidad de ellas que no cumplieron con los tiempos establecidos y el porcentaje que estas representan. Con este análisis fue posible observar las diferencias territoriales en el desempeño de la atención de emergencias, lo que aporta una visión más granular de la problemática.
- xiii. Análisis de los casos que sí cumplen:** se realizó un análisis complementario orientado a las emergencias que sí cumplieron con los tiempos de respuesta, desagregado también por tipo de emergencia. Esto permitió conocer qué categorías presentan mayores porcentajes de cumplimiento y cuáles, aun cumpliendo, representan un desafío operativo menor o mayor.
- xiv. Proporción de llamadas por focalización:** este indicador permite identificar la distribución de las llamadas según los distintos tipos de focalización: interior, medidor, exterior y sin suministro.
- xv. Estadísticas de tiempos de respuesta:** este indicador se utiliza para calcular medidas de tendencia central y de dispersión asociadas a los tiempos de respuesta, tales como media, mediana, mínimo, máximo y desviación estándar.

En resumen, para este apartado, la preparación de variables y la definición de métricas claras es un paso indispensable para dar solidez al análisis. Contar con datos estandarizados, tiempos de respuesta bien calculados y criterios objetivos de cumplimiento permitió reducir la complejidad de la base original y transformarla en datos comprensibles y utilizables.

La claridad alcanzada en este punto no solo garantiza la consistencia de los resultados, sino que además abre el camino para profundizar en los análisis econométricos y de simulación. Sin

esta base estructurada, cualquier interpretación posterior correría el riesgo de ser ambigua o poco confiable.

7.3.2.3 Agrupación y Análisis por Grupo de Comunas. Con el fin de facilitar un análisis operativo más realista y preparar la base para una futura distribución de recursos de emergencia, se estableció un procedimiento de agrupación de comunas en grupos del 1 al 5 y dejando un grupo como otros que son los que tiene muy poca influencia en el general. Ya que poseen muy pocas llamadas. Estos grupos se definieron siguiendo tres criterios principales:

- **Cercanía geográfica:** se consideró la ubicación territorial y la conectividad vial entre comunas, con el objetivo de generar grupos que puedan ser cubiertos por un mismo móvil.
- **Volumen histórico de llamadas:** se agruparon comunas de acuerdo con su carga de emergencias, de manera de equilibrar en cada grupo tanto la densidad de incidentes como la posibilidad de respuesta con los recursos disponibles.

La definición final respecto a la distribución consideró cinco grupos principales y una categoría de “otros” para comunas con menor volumen de llamadas, distribuidos de la siguiente forma:

- **Grupo 1:** Viña del Mar, Valparaíso, Quilpué, Villa Alemana, Concón, Casablanca, Curauma.
- **Grupo 2:** Quillota, Calera, La Cruz.
- **Grupo 3:** Llay Llay, Los Andes.
- **Grupo 4:** La Serena, Coquimbo.
- **Grupo 5:** Talca.
- **Otros:** Catemu, La Higuera, Rancagua.

Para cada grupo de comunas se calcularon métricas clave que sirven como indicadores para la planificación operativa:

- Número total de llamadas históricas.
- Tasa de llamadas diarias promedio.
- Porcentaje de cumplimiento de tiempos de respuesta.
- Distribución por tipo de focalización (interior, exterior, medidor, sin suministro).
- Porcentaje de cumplimiento por grupo y focalización.

El cálculo se realizó en RStudio, empleando principalmente funciones de manipulación como “`dplyr::group_by()`” y “`summarise()`”, y herramientas de visualización (`ggplot2`) para generar gráficos comparativos por grupo de comunas.

Este procedimiento constituye la base metodológica para evaluar escenarios de distribución de móviles de emergencia, ya que entrega una visión agregada del desempeño y del peso relativo de cada zona en la demanda global.

7.3.2.4 Análisis de Simultaneidad de Eventos. Además de las métricas tradicionales, se desarrolló un análisis específico orientado a medir la probabilidad de ocurrencia simultánea de llamadas de emergencia dentro de distintos intervalos de tiempo. Este enfoque resulta especialmente relevante, ya que refleja la presión operativa que enfrenta el proceso cuando dos o más emergencias deben ser atendidas en un periodo reducido, lo cual puede comprometer el cumplimiento de los tiempos de respuesta establecidos.

i. Definición del criterio de simultaneidad

Se definió como “evento simultáneo” todo aquel que ocurre dentro de un intervalo de entre 10 y 120 minutos respecto de otro evento. Para capturar con mayor detalle este fenómeno, se consideraron ventanas móviles en incrementos de 10 minutos. Ósea de 0 a 10 min, de 10 a 20 min y así sucesivamente hasta llegar a los 120 minutos.

ii. Procedimiento en programa RStudio

El cálculo se implementó con programa RStudio, donde se realizaron una serie de pasos necesarios para su procesamiento, son los siguientes:

- Se ordenaron cronológicamente todas las emergencias según la hora de recepción.
- Se calculó la diferencia en minutos entre cada evento y el siguiente (difftime).
- Para cada ventana de tiempo definida (10, 20, hasta 120 minutos), se identificó cuántas llamadas eran clasificadas como simultáneas y se estimó la probabilidad de simultaneidad (probabilidad simultaneidad = número simultáneos / número eventos).
- Este procedimiento se aplicó tanto a nivel global con todas las emergencias en conjunto como a nivel desagregado por comuna, de manera de observar diferencias territoriales en la ocurrencia de simultaneidad.

iii. Métricas generadas

Dentro del análisis es necesario encontrar métricas que permitan complementar las antes calculadas a partir de los datos, para esto se calculan las siguientes:

- **Número total de eventos y número de simultáneos:** En cada intervalo temporal se obtiene el total de eventos y número de llamadas simultáneas.

- **Probabilidad de simultaneidad global:** Muestra cómo esta probabilidad crece a medida que la ventana de análisis se amplía (ejemplo: 0,18 para 10 minutos; 0,37 para 30 minutos; 0,55 para 60 minutos; hasta 0,75 para 120 minutos).
- **Probabilidad de simultaneidad por comuna:** Permite identificar zonas donde la superposición de emergencias es más recurrente (por ejemplo, Viña del Mar, Valparaíso o Quilpué) y otras donde prácticamente no se registran casos simultáneos.

iv. Relevancia metodológica

En consecuencia, los resultados obtenidos a partir de este procedimiento son indicadores indispensables para la etapa de simulación de modelos y escenarios, en la cual se evalúa la dotación mínima de móviles requerida para mantener niveles adecuados de cumplimiento bajo distintas condiciones de demanda.

7.3.2.5 Tratamiento de Valores Atípicos y Faltantes. Como parte del aseguramiento de calidad de los datos, se realizó la identificación y tratamiento de valores atípicos y datos faltantes, con el fin de evitar sesgos en los análisis posteriores. En el caso de los valores atípicos, se aplicó el criterio del rango intercuartílico (IQR), detectando tiempos de respuesta demasiado bajos o altos que podían alterar las medidas de tendencia central.

Estos registros no fueron eliminados de manera automática, sino que se mantuvieron en la base de datos, ya que reflejan situaciones reales en la operación de emergencias, como un móvil ubicado cerca del lugar del llamado o, por el contrario, demoras inusuales por condiciones externas del traslado u otros. Respecto a los datos faltantes, se identificaron dos situaciones:

- Registros incompletos en variables críticas (fecha/hora de recepción, hora de llegada, tipo de emergencia), los cuales fueron descartados por no permitir el cálculo de métricas clave.

- Valores ausentes en variables no críticas (por ejemplo, comuna o focalización), que fueron etiquetados bajo la categoría “No especificado”, de modo que permanecieran en el análisis sin distorsionar la consistencia general.

7.3.2.6 Productos del Análisis Exploratorios. El análisis exploratorio genera un conjunto de tablas y gráficos que sirven para identificar de manera clara los valores obtenidos en cada métrica y facilitar la interpretación de resultados. Permite transformar una base de registros operativos en un conjunto de métricas estructuradas y confiables.

Este proceso no solo se orientó a describir la información disponible, sino que estableció las bases necesarias para los análisis posteriores: por un lado, la modelación econométrica, que requiere datos claros y consistentes para explorar relaciones causales; y por otro lado, la simulación de escenarios, que demanda métricas bien definidas y realistas para dimensionar la dotación óptima de recursos.

7.3.3 Análisis Econométrico

Una vez completada la etapa exploratoria, se procedió a la aplicación de modelos econométricos Logit y Probit con el objetivo de explicar y predecir los factores que inciden en el cumplimiento de los tiempos de respuesta y en la ocurrencia de emergencias críticas. Esta fase metodológica resulta fundamental, ya que entrega una base cuantitativa sólida para la posterior estimación de recursos móviles y la simulación de escenarios de capacidad.

i. Modelo Logit

El modelo Logit asume que la probabilidad de que ocurra el evento $Y_i = 1$ (por ejemplo, que se cumpla el tiempo de respuesta) está dada por la función logística:

$$P(Y_i = 1 | X_i) = \frac{e^{\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki}}}{1 + e^{\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki}}}$$

donde:

- $P(Y_i = 1 | X_i)$ es la probabilidad de que el evento ocurra, dada la información X_i
- β_0 es el intercepto del modelo
- β_j son los coeficientes asociados a las variables explicativas X_j
- e es la base del logaritmo natural.

De forma equivalente, el modelo puede expresarse en términos del Logit (logaritmo de las probabilidades):

$$\ln\left(\frac{P(Y_i = 1)}{1 - P(Y_i = 1)}\right) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki}$$

Esta transformación permite que la variable dependiente se exprese en una escala continua, facilitando la interpretación de los efectos.

ii. Modelo Probit

En el modelo Probit, la probabilidad de ocurrencia se modela mediante la función de distribución acumulada de una variable normal:

$$P(Y_i = 1 | X_i) = \Phi(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki})$$

donde:

- $\Phi(\cdot)$ representa la función de distribución acumulada normal estándar.
- Los coeficientes β_j tienen interpretación análoga a la del modelo Logit, aunque en este caso la función de enlace es la normal acumulada.

El modelo Probit se utiliza cuando se asume que existe una variable no observada Y_i^* , que determina la ocurrencia del evento:

$$Y_i^* = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + \varepsilon_i$$

Donde $\varepsilon_i \sim N(0,1)$.

Así, el evento se observa si:

$$Y_i = \begin{cases} 1, & \text{si } Y_i^* > 0 \\ 0, & \text{si } Y_i^* \leq 0 \end{cases}$$

7.3.3.1 Modelos de Clasificación Binaria. La primera aproximación econométrica se orientó a evaluar la capacidad de modelos Logit y Probit para distinguir entre emergencias reales y no reales o cual es el grado de influencia de las variables dentro del registro histórico de llamadas.

Esta etapa fue considerada fundamental, ya que permite filtrar la demanda efectiva a partir de las llamadas recibidas, diferenciando aquellas que corresponden a incidentes que requieren atención en terreno de aquellas que constituyen falsas alarmas u otros reportes sin implicancia operativa.

7.3.3.1.1 Modelos para Explicar la Ocurrencia de Emergencias Reales. La primera etapa del análisis econométrico se orientó a identificar los factores que influyen en que una llamada corresponda a una emergencia real dentro del registro histórico. Esta aproximación resulta esencial, ya que permite entender cómo determinadas características de la llamada, del evento o del contexto territorial afectan la probabilidad de que el caso deba ser efectivamente atendido.

i. Definición de la variable dependiente

Se construyó una variable binaria denominada emergencia ocurrida, donde se asignó el valor 1 cuando el registro correspondía a un incidente real y 0 cuando se trataba de una falsa alarma u otro evento, aunque el llamado al ser una incertidumbre respecto si es cierto o no, igual se debe confirmar asistencia al lugar.

ii. Selección de variables explicativas

Se consideraron como predictores aquellas variables con mayor potencial para influir en la ocurrencia de emergencias reales:

- Tipo de emergencia reportada (fuga, olor a gas, sin suministro, entre otros).

-
- Focalización del evento (interior, exterior, medidor, sin suministro).
 - Comuna donde se origina la llamada, para capturar diferencias territoriales y operativas.
 - Hora de recepción del llamado, como aproximación a patrones temporales en la ocurrencia de emergencias.

Estas variables se estandarizaron y revisaron previamente en el análisis exploratorio y en el análisis de datos, lo que aseguró consistencia y comparabilidad en la estimación.

iii. Estimación de los modelos

Con la variable dependiente y los predictores definidos, se procedió a la estimación de modelos Logit y Probit.

- Ambos modelos permiten modelar la probabilidad de ocurrencia de un evento binario, pero con funciones de enlace distintas: logística en el caso del Logit y normal acumulada en el Probit.
- Se optó por estimar ambos enfoques para comparar sus desempeños y seleccionar el más adecuado según criterios de ajuste y utilidad práctica.
- La estimación se realizó en el software RStudio, mediante métodos de máxima verosimilitud, que permiten obtener los coeficientes asociados a cada variable y sus niveles de significancia.

Ambos enfoques fueron estimados de manera complementaria con el propósito de contrastar sus desempeños y determinar cuál ofrecía un mejor ajuste y utilidad para el contexto operativo.

iv. Evaluación y comparación de modelos

Con el fin de determinar cuál de los modelos Logit o Probit ofrecía un mejor desempeño, se establecieron criterios de comparación a partir de los resultados de ajuste y la significancia de los parámetros estimados. Entre los principales criterios considerados se encuentran:

- **Deviance (null y residual):** permite observar la reducción en la variabilidad explicada por el modelo en comparación al caso nulo. Una menor deviance residual refleja un mejor ajuste.
- **Criterio de información de Akaike (AIC):** se utilizó como medida de equilibrio entre ajuste y complejidad, donde valores más bajos indican un modelo con mejor balance.
- **Significancia de las variables:** se evaluó el nivel de significancia estadística (p-valor) de cada predictor, con el fin de identificar cuáles factores presentaban un efecto relevante en la probabilidad de ocurrencia de emergencias reales.

En particular, los criterios de información AIC (Akaike Information Criterion) y BIC (Bayesian Information Criterion) se utilizan ampliamente en estadística para comparar modelos alternativos. Ambos se calculan a partir de qué tan bien el modelo logra ajustarse a los datos observados, pero incluyen un ajuste que considera la cantidad de parámetros, con el fin de evitar que se favorezcan modelos demasiado complejos.

- Un valor más bajo de AIC o BIC indica un mejor equilibrio entre buen ajuste y simplicidad del modelo.
- Mientras que el AIC pone mayor énfasis en la capacidad de predicción, el BIC aplica un ajuste más fuerte por el número de parámetros, lo que normalmente lleva a preferir modelos más simples (Burnham & Anderson, 2004).

Estos criterios permitieron comparar directamente el desempeño de ambos modelos bajo la misma base de datos. Si bien Logit y Probit suelen entregar resultados muy similares en contextos aplicados, esta evaluación fue necesaria para respaldar metodológicamente la selección del modelo a utilizar en las etapas siguientes del estudio.

v. Rol dentro de la investigación

Esta primera aproximación entrega una visión explicativa sobre los determinantes de las emergencias reales. Conocer qué factores aumentan o reducen la probabilidad de que una llamada corresponda a un incidente efectivo constituye un insumo clave para la planificación posterior.

7.3.3.1.2 Modelos para Evaluar la Capacidad Predictiva. En esta etapa, el interés metodológico se centró en evaluar la habilidad de los modelos Logit y Probit para clasificar correctamente las observaciones. Es decir, además de explicar los factores asociados a la ocurrencia de emergencias, se buscó determinar en qué medida estas herramientas podían anticipar si un nuevo registro corresponde a un incidente real o no.

i. Selección de variables

Para este procesamiento de modelos se mantuvo la misma definición de la variable dependiente (emergencia ocurrida) y el mismo conjunto de variables explicativas descritas en la sección anterior, de modo de garantizar comparabilidad en los resultados.

ii. Evaluación y comparación de modelos

Para este fin, se generaron matrices de clasificación a partir de los valores estimados por los modelos. A partir de dichas matrices, se calcularon indicadores de desempeño.

- **Exactitud global (Accuracy):** proporción de predicciones correctas sobre el total de casos.

- **Sensibilidad (Recall):** capacidad de identificar correctamente las emergencias reales.
- **Especificidad:** capacidad de reconocer adecuadamente los casos que no corresponden a emergencias.
- **Precisión (Precision):** proporción de emergencias predichas que efectivamente resultaron ser casos reales.

iii. **Justificación del enfoque**

La evaluación predictiva es fundamental porque la utilidad de los modelos no se limita a explicar relaciones estadísticas, sino también a apoyar la toma de decisiones operativas en tiempo real. En este contexto, un modelo con buena capacidad de predicción permite filtrar la demanda y estimar la proporción de llamadas que requieren efectivamente un recurso móvil.

iv. **Rol dentro de la investigación**

Este segundo bloque metodológico refuerza el valor práctico de los modelos econométricos. Al validar su capacidad predictiva, se genera una herramienta que puede ser aplicada sobre escenarios de demanda futura para estimar con mayor realismo cuántas llamadas se transformarán en emergencias efectivas, independiente se pueda dimensionar para la capacidad operativa. Esto constituye un indicador directo para la etapa siguiente de modelación de capacidad, donde se buscará determinar la dotación óptima de móviles bajo distintos escenarios de operación.

7.3.3.2.1 Modelos para Explicar el Cumplimiento de Tiempos de Respuesta. Al igual que en la etapa anterior, en esta sección se emplean modelos econométricos de tipo Logit y Probit, estimados mediante máxima verosimilitud, con el fin de analizar un nuevo fenómeno de interés: el cumplimiento de los tiempos de respuesta. Dado que ambos modelos comparten una estructura

metodológica común en función de enlace logística o normal acumulada, selección de variables explicativas y criterios de ajuste como AIC y deviance, se mantuvo la misma base de datos y conjunto de predictores, modificando únicamente la definición de la variable dependiente.

Esta integración metodológica permite garantizar la comparabilidad entre los modelos de ocurrencia de emergencias reales y los de cumplimiento de tiempos de respuesta, evaluando cómo las mismas características operativas y territoriales influyen tanto en la necesidad de atención como en la eficiencia del proceso de respuesta.

i. Definición de la variable dependiente

La variable dependiente utilizada en esta etapa se definió como cumple tiempos de respuesta (cumple_tiempo en RStudio), la cual toma el valor 1 cuando la atención de la emergencia se realizó dentro del tiempo de respuesta exigido, y 0 en caso contrario.

7.3.3.2 Modelos para Evaluar la Capacidad Predictiva. Esta etapa busca principalmente medir y optimizar la capacidad de los modelos Logit y Probit para predecir si una emergencia cumplirá o no con el tiempo de respuesta establecido según cada tipo de emergencia. A diferencia del bloque explicativo anterior, aquí el foco es estrictamente predictivo-operacional.

i. Definición de variables y consistencia con etapas previas

Se mantuvo la misma variable dependiente binaria (cumple_tiempo: 1 = cumple; 0 = no cumple) y el mismo conjunto de predictores (tipo de emergencia, focalización, comuna y hora de

recepción), con el fin de garantizar comparabilidad metodológica y aislar el efecto de las técnicas de balanceo y del umbral de decisión sobre el desempeño predictivo.

ii. Diagnóstico inicial: desbalance de clases

El análisis de la variable objetivo mostró un desbalance (mayoría de casos “no cumple”). En un escenario así, un umbral estándar de 0,5 tiende a favorecer la clase mayoritaria, logrando altas tasas de acierto en “no cumple” pero perdiendo sensibilidad para detectar los casos que sí cumplen. Este sesgo hacia la clase 0 fue documentado y motivó la aplicación de estrategias de reequilibrio y ajuste de umbral.

En modelos de regresión binaria como Logit y Probit, la estimación entrega probabilidades de ocurrencia para cada observación. Para convertir esas probabilidades en una decisión binaria (cumple / no cumple), es necesario fijar un umbral de clasificación (τ).

Por convención, se utiliza como referencia inicial un umbral de 0,5, es decir, se predice que una emergencia cumplirá el tiempo de respuesta si su probabilidad estimada es mayor o igual a 50%, y no lo cumplirá en caso contrario. Este valor es el más común porque divide en partes iguales el rango de probabilidad y es adecuado cuando las clases están balanceadas y el costo de los errores es similar. Por lo tanto, es necesario especificar claramente el proceso utilizado y porque es necesario la reestructuración de la habilidad del modelo para predecir el cumplimiento de los tiempos de respuesta.

iii. Estimación base de los modelos

El proceso comenzó con el ajuste de modelos Logit y Probit utilizando el conjunto de datos completo. La estimación se realizó mediante el método de máxima verosimilitud en RStudio, lo que permite verificar y obtener para cada observación una probabilidad estimada de cumplimiento de los tiempos de respuesta. Es importante destacar que en esta etapa no se aplicó ningún tipo de corrección por desbalance ni modificación en los parámetros de decisión; el objetivo fue construir una línea base sobre la cual evaluar los resultados posteriores y entender cómo se comporta el modelo en su forma más simple.

iv. Evaluación inicial con umbral estándar (0,5)

Una vez obtenidas las probabilidades, fue necesario convertirlas en una predicción binaria: “cumple” o “no cumple”. Por convención, se utilizó el umbral de 0,5, lo que significa que se predijo “cumple” cuando la probabilidad estimada era mayor o igual al 50%, y “no cumple” en caso contrario.

Con este criterio, se construyó la matriz de confusión y se calcularon las métricas de desempeño más utilizadas en problemas de clasificación: Exactitud, Sensibilidad (Recall) para la clase “cumple”, Especificidad para la clase “no cumple”, Precisión (PPV) y el F1-Score, que combina la Precisión y la Sensibilidad.

Los resultados de esta primera validación confirmaron lo esperado respecto a los datos observados y en un escenario con desbalance de clases: el modelo alcanzaba buena especificidad y precisión, pero presentaba una baja sensibilidad, es decir, no lograba identificar con suficiente éxito los casos en que efectivamente se cumplían los tiempos de respuesta.

v. Ajuste del umbral de decisión (τ)

Dado que el diagnóstico inicial evidenció un fuerte desbalance entre clases con una proporción significativamente mayor de casos “no cumple” (0) respecto a “cumple” (1), se consideró necesario complementar el ajuste de los modelos mediante la modificación del umbral de decisión (τ).

Con el fin de mejorar la capacidad del modelo para detectar los cumplimientos, se implementó una exploración sistemática de distintos valores de umbral, todos ellos inferiores a 0,5. El rango evaluado fue aproximadamente $\tau \in [0,15; 0,40]$. La lógica de este procedimiento es que, al bajar el umbral, el modelo clasifica como “cumple” un mayor número de observaciones, lo que teóricamente podría incrementar la sensibilidad, aunque podría reducir la especificidad y la precisión. Para determinar el umbral más adecuado se utilizaron herramientas métricas:

- **Curva ROC (Receiver Operating Characteristic):** representa la relación entre la tasa de verdaderos positivos (Sensibilidad) y la tasa de falsos positivos (1 - Especificidad) para todos los posibles valores de umbral. Esta curva permite observar el desempeño global del modelo al variar τ . Cuanto mayor es el AUC, mejor es la capacidad del modelo para distinguir entre las dos clases.
- **Curva PR (Precision–Recall):** muestra la relación entre la precisión y la sensibilidad para diferentes valores de umbral. Esta curva es especialmente útil cuando existe un desbalance, como en este caso, porque refleja directamente la relación entre detectar más casos positivos (Recall) y mantener un nivel adecuado de confianza en esas predicciones (Precisión).

La selección final del umbral se basó en la observación de estas curvas y, particularmente, en el F1-Score de la clase positiva (“cumple”), bajo el criterio operativo de que es preferible no

dejar de identificar emergencias que cumplen los tiempos de respuesta, aun cuando ello implique aceptar un mayor número de falsas alarmas.

vi. **Balanceo mediante submuestreo de datos**

Aun cuando la modificación del umbral puede mejorar la sensibilidad (Recall), pero puede persistir el problema del desbalance de clases en el conjunto de entrenamiento, lo que puede sesgar el modelo hacia la clase mayoritaria (“no cumple”). Para mitigar este sesgo, se aplica un procedimiento de submuestreo controlado. En concreto, se redujo el número de observaciones de la clase mayoritaria manteniendo aproximadamente un tercio de ellas, en relación con el total de observaciones de la clase minoritaria. De esta forma, el modelo se procesó en un contexto más equilibrado.

vii. **Reprocesamiento y selección conjunta**

En esta fase se procedió a comparar distintos enfoques metodológicos con el objetivo de optimizar el desempeño del modelo en la predicción del cumplimiento de tiempos de respuesta. Se definieron tres configuraciones principales:

- **Solo ajuste de umbral:** se utilizó el conjunto de datos original, aplicando únicamente la exploración de umbrales inferiores a 0,5. Este enfoque buscó mejorar la Sensibilidad incrementando la detección de casos de “cumple”, aunque mantuvo el desbalance de clases en el entrenamiento.
- **Solo balanceo de clases:** se entrenaron nuevamente los modelos Logit y Probit sobre un conjunto balanceado mediante submuestreo controlado, manteniendo el umbral estándar

de 0,5 como criterio de clasificación. Este escenario permitió evaluar en qué medida el equilibrio entre clases por sí solo era suficiente para mejorar la capacidad predictiva.

- **Balanceo más ajuste de umbral:** finalmente, se combinaron ambas estrategias. Primero se entrenaron los modelos sobre el conjunto balanceado, y luego se exploró la rejilla de valores de umbral ($\tau \in [0,15; 0,40]$) para ajustar el criterio de decisión.

Este enfoque permitió atacar simultáneamente el sesgo de entrenamiento derivado del desbalance y la rigidez del umbral estándar, logrando un modelo más flexible y adaptado al objetivo operativo.

En cada uno de estos escenarios se evaluaron los mismos indicadores de desempeño: sensibilidad de la clase “cumple”, F1-Score como medida de equilibrio entre sensibilidad y precisión, y un nivel mínimo de especificidad aceptable para evitar un exceso de falsas alarmas. Tanto Logit como Probit fueron estimados y comparados bajo estas tres configuraciones.

viii. Criterios de comparación y decisión

La selección del modelo más adecuado se sustentó en un conjunto de criterios metodológicos y operativos:

- **Métricas de clasificación:** se utilizaron como referencia la exactitud, el Recall de la clase “cumple”, la especificidad, la precisión y el F1-Score, con especial énfasis en el equilibrio entre estas medidas.
- **Curvas ROC y PR:** sirvieron para identificar el umbral operativo óptimo, evaluando la relación entre verdaderos positivos y falsos positivos, así como el compromiso entre precisión y sensibilidad.

- **Consideración operativa de errores:** se priorizó la reducción de falsos negativos, dado que en la gestión de emergencias resulta más crítico no reconocer un cumplimiento que aceptar una falsa alarma. Esta decisión metodológica responde a la lógica de proteger los niveles de servicio exigidos.
- **Robustez entre modelos:** se incluyó la estimación tanto en modelo Logit como en Probit para contrastar resultados y verificar la consistencia metodológica de los hallazgos.

ix. **Justificación del enfoque**

El diseño metodológico aplicado respondió a una lógica de doble corrección:

- Solo ajustar el umbral podría solo incrementar la sensibilidad y mejorar la detección de casos positivos teóricamente, pero no corregir el sesgo de generado por el fuerte desbalance de clases en los datos observados.
- Balancear las clases y ajustar el umbral en conjunto podría ofrecer una solución más robusta teóricamente, ya que el modelo aprende en un entorno más equilibrado y, al mismo tiempo, podría aplicar un criterio de decisión adaptado al objetivo operativo: maximizar la identificación de emergencias que cumplen con los tiempos de respuesta.

Este enfoque integral podría permitir obtener un modelo con mayor sensibilidad y un F1-Score más sólido, sacrificando solo una fracción controlada de especificidad teóricamente.

x. **Rol dentro de la investigación**

Esta etapa metodológica resulta crucial porque no solo podría entregar un modelo con buen desempeño estadístico. En particular, el clasificador obtenido permite:

- Detectar con alta probabilidad las emergencias que efectivamente cumplirán con los tiempos de respuesta, reduciendo al mínimo los falsos negativos.
- Mantener un equilibrio razonable entre sensibilidad y precisión, garantizando que las predicciones sean operativamente confiables.
- Proyectar el cumplimiento esperado en función de variables como comuna, franja horaria o tipo de emergencia, lo que enriquece el análisis.

De esta manera, los resultados de esta etapa constituyen un insumo directo para el modelo de capacidad, ya que permiten evaluar distintos escenarios de dotación de móviles bajo metas de nivel de servicio concretas.

7.3.4 *Estimación Probabilística de la Demanda*

La cuarta etapa metodológica corresponde a la estimación probabilística de la demanda de emergencias, cuyo propósito es modelar la ocurrencia de los llamados en términos estadísticos para proyectar escenarios futuros de atención. Esta fase resulta fundamental, ya que la demanda constituye la principal fuente de presión sobre el sistema.

A diferencia de un enfoque determinista basado únicamente en promedios históricos, el análisis probabilístico permite capturar la variabilidad e incertidumbre inherentes al fenómeno de las emergencias. De esta forma, es posible no solo estimar la cantidad esperada de llamados en un periodo dado, sino también considerar la probabilidad de ocurrencia de situaciones extremas, tales como días con una carga de trabajo inusualmente alta.

Para ello se analizaron las distribuciones de probabilidad que mejor se ajustan a los datos

observados, considerando tanto el tiempo entre la ocurrencia de llamadas como el número de llamadas recibidas en intervalos definidos. La identificación y validación de estas distribuciones se realizó mediante técnicas de ajuste en software estadístico RStudio y Python, contrastando múltiples alternativas y seleccionando aquellas que ofrecieran el mejor desempeño en términos de bondad de ajuste y criterios de información.

i. Selección de distribuciones de probabilidad

Para identificar y después modelar adecuadamente la demanda de emergencias, se consideró necesario evaluar diferentes distribuciones de probabilidad, tanto para las variables de conteo (número de llamadas por intervalo de tiempo) como para las variables continuas positivas (tiempos entre llamadas). La elección de estas distribuciones se sustentó en sus características teóricas, en la literatura sobre procesos de llegadas y en aplicaciones similares en el ámbito de la gestión de operaciones y de servicios de emergencia.

En primer lugar, se evaluó la distribución Poisson, que constituye el modelo base en la teoría de procesos de llegadas. Esta distribución asume que los eventos ocurren de manera independiente y con una tasa constante, y que la media es igual a la varianza. Debido a su sencillez y amplia utilización en sistemas de colas y planificación de servicios, se considera el punto de partida natural para modelar llamadas de emergencia (Cameron & Trivedi, 2013). Sin embargo, diversos estudios han señalado que en datos de demanda real suele presentarse sobre dispersión (es decir, una varianza mayor que la media), lo que limita la capacidad del modelo Poisson para representar adecuadamente la variabilidad observada (Hilbe, 2011).

Frente a este problema, se incorporó la distribución Binomial Negativa, la cual introduce un parámetro adicional de dispersión que permite ajustar escenarios donde la varianza excede a la media. Este modelo se ha utilizado ampliamente en la literatura aplicada para el análisis de conteos

en salud y emergencias, ya que ofrece mayor flexibilidad que el Poisson y es capaz de manejar la heterogeneidad no observada en los datos (Lawless, 1987; Cameron & Trivedi, 2013).

En paralelo, para el análisis de los tiempos entre llamadas, se consideraron distribuciones continuas positivas. La primera de ellas fue la distribución Exponencial, que supone que los intervalos entre eventos son independientes y siguen un patrón con tasa constante. Este modelo es el más simple dentro de los procesos de Poisson, aunque suele resultar demasiado restrictivo en contextos reales donde los tiempos presentan mayor variabilidad (Ross, 2014).

Para superar estas limitaciones, se analizaron dos distribuciones más flexibles: la distribución Gamma y la distribución Lognormal. La Gamma se caracteriza por tener dos parámetros (forma y escala), lo que le permite modelar con precisión distribuciones sesgadas hacia la derecha y adaptarse a diferentes grados de variabilidad. Por estas razones, ha sido empleada en estudios de fiabilidad, riesgo y tiempos de espera en sistemas de servicio (Lawless, 1987). Por su parte, la Lognormal surge naturalmente cuando el tiempo de ocurrencia de un evento es el resultado de la multiplicación de múltiples factores aleatorios positivos. Su capacidad para capturar colas largas y valores extremos la convierte en una candidata adecuada en escenarios donde existen intervalos de espera considerablemente mayores que el promedio (Limpert, Stahel & Abbt, 2001).

En síntesis, la selección de distribuciones candidatas incluyó tanto modelos clásicos como extensiones que permiten manejar escenarios más complejos.

De esta manera, se buscó garantizar que la representación probabilística de la demanda reflejara adecuadamente no solo los valores promedio, sino también la dispersión y los casos extremos propios de la ocurrencia de emergencias. Una vez definidas las distribuciones candidatas,

se procedió a realizar un proceso sistemático de ajuste y comparación, con el fin de identificar aquella que mejor representara la dinámica de la demanda de emergencias.

ii. Preparación de los datos

La base de datos de llamadas fue previamente organizada en intervalos de observación adecuados para el análisis:

- Conteos de llamadas por unidad de tiempo, por día y por hora para el estudio de distribuciones discretas.
- Intervalos de tiempo entre llamadas consecutivas, medidos en minutos, para el análisis de distribuciones continuas.

iii. Estimación de parámetros

Cada una de las distribuciones candidatas fue ajustada utilizando las funciones disponibles en librerías especializadas de RStudio y Python, como “fitdistrplus” y “scipy.stats”.

Estas herramientas permiten obtener automáticamente los parámetros de cada distribución a partir de los datos, empleando procedimientos de estimación internos en la mayoría de los casos basados en máxima verosimilitud sin necesidad de programar el algoritmo manualmente. El ajuste y selección en ambos programas se realizó de manera independiente, lo que permitió contrastar y validar la consistencia de los resultados obtenidos. Este enfoque aseguró que las conclusiones no dependieran de un único software.

iv. Evaluación del ajuste

Una vez estimados los parámetros, se evaluó el desempeño de cada distribución mediante

dos estrategias complementarias:

- **Criterios de información (AIC y BIC):** se calcularon para todas las distribuciones ajustadas, con el fin de comparar su calidad relativa. Estos criterios funcionan en base a la complejidad del modelo y permiten identificar la alternativa que ofrece un mejor equilibrio entre ajuste y simplicidad.

El uso conjunto de los criterios AIC y BIC, proporcionó una base objetiva y práctica para la comparación de las distribuciones. Esta combinación permitió evitar la dependencia de un único indicador y garantizó que la selección final no solo se sustentara en métricas estadísticas, sino también en la capacidad del modelo para representar fielmente el comportamiento empírico de la demanda.

v. **Comparación y selección**

Con los indicadores calculados y las comparaciones realizadas, se procedió a contrastar el desempeño de las distintas distribuciones candidatas. Este proceso no se limitó a identificar la que obtuviera los valores más bajos de AIC o BIC, sino que también consideró la capacidad del modelo para reproducir de forma realista la forma y dispersión observada en los datos. De este modo, la selección final respondió a un criterio integral: por un lado, un ajuste estadístico adecuado, y por otro, una coherencia práctica respecto al fenómeno de las llamadas de emergencia.

vi. **Definición de escenarios probabilísticos**

Una vez identificada la distribución que mejor representaba la demanda de emergencias, se utilizaron sus parámetros estimados. Este procedimiento buscó no limitarse a un único valor

esperado, sino capturar la variabilidad propia del fenómeno y anticipar situaciones de distinta intensidad operativa. El proceso se desarrolló en algunos pasos principales:

- **Cálculo de parámetros actuales:** a partir de los parámetros de la distribución seleccionada, se determinó la media teórica. Este escenario representa la cantidad promedio de llamadas que podrían esperarse en un periodo típico.
- **Obtención de percentiles de la distribución:** para representar la dispersión de los datos e identificar el escenario actual, se calcularon distintas probabilidades de la distribución seleccionada.
- **Escenarios de baja demanda:** identificados a partir de escenarios de baja demanda y sus probabilidades de ocurrencia.
- **Escenarios de alta demanda:** asociados a mayor cantidad de llamadas, representan situaciones críticas en las que el volumen de llamadas supera la media y exige una mayor capacidad de respuesta.

En conjunto, este procedimiento entregó un marco probabilístico sólido para describir la demanda bajo diferentes condiciones. Los escenarios resultantes no solo reflejan lo que ocurre en promedio, sino también los extremos que deben considerarse en la planificación de capacidad.

vii. Rol dentro de la investigación

La estimación de la demanda actual cumple un papel central en la investigación, ya que provee datos necesarios para trasladar el análisis desde un plano meramente descriptivo hacia una

perspectiva más profunda. Al evaluar la ocurrencia de llamadas mediante distribuciones de probabilidad, no solo se obtiene un valor esperado de demanda, sino también una representación explícita de su variabilidad e incertidumbre.

Esto resulta fundamental en el contexto operativo de una empresa de distribución de gas natural, donde la planificación de recursos móviles no puede basarse únicamente en promedios históricos, sino que debe considerar también los escenarios de mayor exigencia. Gracias a este enfoque probabilístico, podría ser posible anticipar situaciones de alta demanda y evaluar si la dotación de móviles resulta suficiente para mantener los niveles de servicio comprometidos.

7.3.5 Simulación de Escenarios

Esta fase metodológica consistió en la simulación de escenarios de demanda y capacidad, cuyo propósito fue evaluar cómo varía el desempeño del sistema de atención de emergencias bajo diferentes niveles de llamadas simuladas y configuraciones de móviles disponibles. Esta etapa se centra en la modelación estocástica, entendida como el estudio de sistemas donde los fenómenos se describen a través de variables aleatorias y distribuciones de probabilidad (Escalona, 2024). En términos prácticos, la situación corresponde a un problema de sistemas de espera o colas, donde las llamadas de emergencia representan los arribos y los móviles disponibles constituyen los servidores. La literatura en gestión de operaciones reconoce que este tipo de problemas no puede abordarse de manera determinista, ya que la demanda es incierta y sujeta a variabilidad (Gross & Harris, 1998; Kleinrock, 1975).

En consecuencia, la simulación estocástica ofrece un enfoque más flexible que la aplicación directa de fórmulas, ya que permite integrar las distribuciones estimadas en el apartado 5.3.4 y reflejar la realidad operativa con mayor fidelidad (Banks et al., 2010).

i. Procedimiento metodológico

La simulación de escenarios se diseñó para analizar el desempeño del sistema de atención de emergencias bajo distintos niveles de demanda futura y configuraciones de móviles disponibles.

Para abordar la incertidumbre en la demanda, se utilizó la simulación Monte Carlo, técnica ampliamente empleada en estadística aplicada y gestión de operaciones. La lógica de Monte Carlo consiste en generar múltiples réplicas de un mismo proceso, produciendo valores aleatorios a partir de una distribución de probabilidad previamente ajustada en este caso seleccionada en el punto anterior 5.3.4. Con ello es posible observar el rango de resultados posibles y calcular escenarios representativos en función de percentiles (Kroese et al., 2014). En este caso, se implementó en el software RStudio, utilizando funciones de librerías como “fitdistrplus” y “stats” para la generación de valores aleatorios. A través de ciclos de repetición se simularon “días futuros” de llamadas, obteniendo escenarios de baja, media y alta demanda que luego fueron usados como demanda para la evaluación de capacidad.

El sistema de atención se conceptualizó bajo la lógica de los modelos de colas. En la formulación clásica, un modelo M/M/c describe un sistema con llegadas Poisson (tasa λ), tiempos de servicio exponenciales (tasa μ) y “c” servidores en paralelo. Este marco permite calcular métricas como:

- **Probabilidad de espera (Erlang-C):**

$$P_{\text{espera}} = \frac{\frac{(\lambda/\mu)^c}{c! (1 - \rho/c)}}{\sum_{k=0}^{c-1} \frac{(\lambda/\mu)^k}{k!} + \frac{(\lambda/\mu)^c}{c! (1 - \rho/c)}}$$

donde:

$$\rho = \frac{\lambda}{c\mu}$$

Mide la utilización promedio del sistema (proporción de tiempo en que los servidores están ocupados).

- **Tiempo promedio en cola:**

$$W_q = \frac{P_{\text{espera}}}{c\mu - \lambda}$$

- **Tiempo total de respuesta:**

$$W = W_q + \frac{1}{\mu}$$

Estas expresiones reflejan cómo los tiempos de espera dependen de la relación entre la carga de la demanda (λ) y la capacidad instalada ($c\mu$), y constituyen el fundamento teórico para el análisis de sistemas de servicio (Gross & Harris, 1998; Kleinrock, 1975).

El análisis probabilístico de la demanda en el punto 5.3.4 mostró que los datos no se ajustaban a un proceso Poisson puro, ya que presentaban sobre dispersión, lo cual justificó la elección de una distribución Binomial Negativa para modelar los conteos de llamadas. En consecuencia, aunque se tomó como referencia conceptual la teoría de colas, la implementación práctica se realizó mediante simulación estocástica de Monte Carlo en RStudio, generando llegadas conforme a la distribución de probabilidad seleccionada posterior al análisis de datos y desde la estimación de la demanda. Este procedimiento permitió mantener la coherencia con la teoría de colas, pero adaptándola a la realidad del sistema de emergencias.

La construcción de los escenarios se apoyó en los resultados obtenidos en etapas anteriores de la investigación. Del análisis exploratorio en el punto 5.3.2, se incorporó la caracterización de comunas, franjas horarias de llamadas y tiempos de cumplimiento, focalización de emergencias, grupo de comunas, lo que permitió desagregar las llamadas simuladas de acuerdo con patrones

territoriales y tiempos reales establecidos.

De los modelos econométricos en el punto 5.3.3, se consideraron las probabilidades de que un llamado fuera una emergencia real considerando los tipos de emergencia y comunas, además de que cumpliera los tiempos de respuesta, utilizadas como factores de corrección para enriquecer la simulación.

De la estimación probabilística de la demanda en el punto 5.3.4 se obtuvieron los parámetros de la Binomial Negativa, base para la generación de los escenarios simulados de baja, media y alta demanda. Finalmente, se incorporó la simultaneidad de llamadas, reconociendo que los incumplimientos suelen producirse cuando coinciden varios eventos en paralelo; en la simulación se implementó una regla que permite que una fracción de estos choques de llamadas sea resuelta al aumentar la dotación de móviles.

Una vez generadas las llamadas diarias simuladas e integrados estos insumos, la demanda se desagregó geográficamente según la proporción histórica de emergencias en cada grupo de comunas. Se mantuvo la diferenciación por tipo de focalización (interior, exterior/medidor o sin suministro), dado que cada categoría posee distintos tiempos de respuesta establecidos. Posteriormente, la atención de llamadas se modeló como un sistema con capacidad limitada, en el cual los móviles atienden de manera secuencial los requerimientos entrantes. Se incorporaron los dos componentes principales del tiempo total: el tiempo fijo de traslado, independiente del número de móviles, y el tiempo de espera en cola, que sí depende de la dotación.

El tiempo de traslado se define como una variable fija dentro del modelo, dado que depende directamente de la distancia entre comunas, previamente agrupadas en el análisis exploratorio de datos. Para efectos de la modelación, se establece como referencia el tiempo máximo observado por aplicaciones móviles, correspondiente al traslado entre las comunas más alejadas. De esta

manera, se considera el escenario más desfavorable, asegurando que las estimaciones del modelo sean conservadoras y reflejen una condición límite de operación.

Finalmente, para cada escenario y nivel de capacidad se registraron indicadores, incluyendo el tiempo promedio de espera, número de móviles y el porcentaje de llamadas que cumplen especialmente determinado para los grupos de comunas. Estas métricas sirvieron para comparar configuraciones de capacidad, evaluando las relaciones entre número de recursos asignados y nivel de servicio respecto al cumplimiento establecido. Además, se introducen variaciones en aquellas variables que desempeñan un papel relevante dentro del sistema de atención de emergencias, con el propósito de analizar cómo estas modificaciones impactan en las medidas de evaluación y en los niveles de cumplimiento establecidos para el sistema.

7.3.6 *Análisis de Costo-Beneficio.*

A partir de los resultados de cumplimiento obtenidos en la simulación, se incorpora un análisis costo–beneficio orientado a evaluar en qué medida resulta conveniente aumentar o no la dotación de móviles, además de considerar otras variaciones operacionales que pueden aportar para mejorar el cumplimiento en la atención de emergencias desde una doble perspectiva: económica y operacional. Este análisis busca determinar si el incremento en los costos asociados a disponer de móviles adicionales o no se justifican frente a la mejora lograda en los niveles de cumplimiento de los tiempos de respuesta y en los porcentajes globales de atención dentro del sistema.

8. Resultados

8.1 Introducción

En este capítulo se presentan los principales resultados obtenidos a partir de la aplicación de la metodología descrita en el capítulo anterior. La exposición de resultados se organiza de

manera progresiva, comenzando con el análisis exploratorio de la base de datos, lo que permite comprender patrones iniciales en la ocurrencia de llamadas de emergencia y en los tiempos de respuesta observados. Posteriormente, se incluyen los resultados derivados de los modelos econométricos aplicados, que aportan evidencia sobre los factores que inciden en la ocurrencia y distribución de emergencias. A continuación, se presentan los hallazgos de la simulación de escenarios, en los que se evaluaron distintos niveles de dotación de móviles y su efecto sobre los porcentajes de cumplimiento de los estándares de respuesta definidos. Finalmente, se integra un análisis costo–beneficio que complementa la visión operativa, al contrastar la mejora obtenida en los niveles de servicio con el impacto económico asociado al aumento de móviles. De esta forma, el capítulo no solo entrega cifras y métricas de desempeño, sino que también interpreta su alcance en relación con los objetivos de investigación, particularmente en lo que respecta a la eficiencia operativa y la toma de decisiones estratégicas en la gestión de emergencias de gas natural.

8.2 Análisis Exploratorio de Datos

El análisis exploratorio de datos permitió identificar patrones en las llamadas de emergencia y en los tiempos de respuesta según focalización y comuna. La base consolidada se estructuró en variables clave, resumidas en la Tabla 1, que sirvieron como insumo para el análisis posterior.

Tabla 1

Estructura de datos

Variable	Descripción	Tipo de dato	Ejemplo
Fecha	Día en que se registró la emergencia	Fecha	2022-06-06
Hora de Recep.	Hora de recepción de la llamada	Hora (HH:MM:SS)	15:50:00

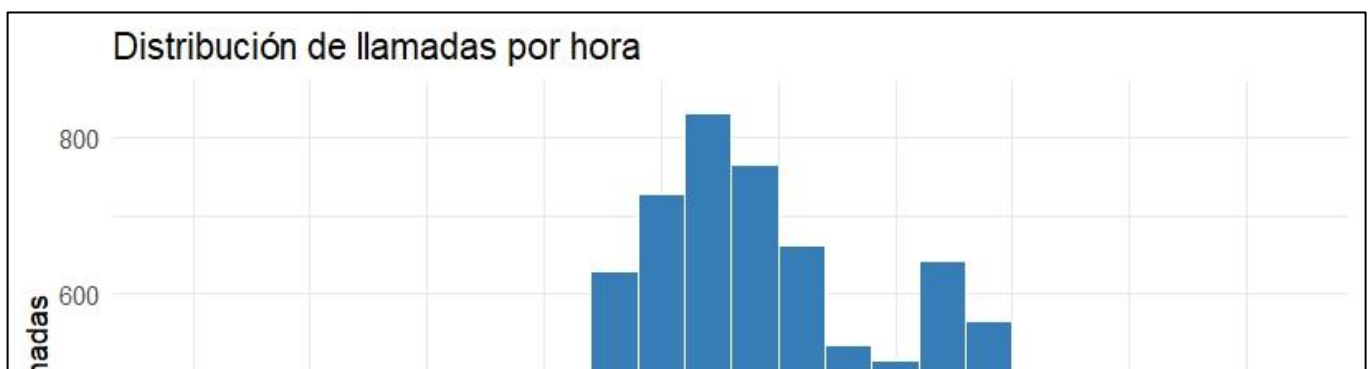
N° Gestión	Código único de identificación de la emergencia	Numérica	2803524
Hora de Check in	Hora en que el móvil reporta inicio de atención	Hora (HH:MM:SS)	15:55:00
Tiempo de respuesta	Diferencia entre la recepción y el inicio de atención	Tiempo	00:05:00
Dirección / Comuna	Ubicación exacta y comuna del evento	Texto	Viña del Mar
Tipo Emergencia	Clasificación inicial de la emergencia (ej. fuga, olor a gas, otras)	Catagórica	Fuga
Focalización	Lugar asociado al evento (interior, exterior, medidor, sin suministro)	Catagórica	Medidor
Clasificación	Tipo de resultado del evento (incidente, falsa alarma, etc.)	Catagórica	Incidente
Emergencia ocurrida	Indicador de ocurrencia real (1 = sí, 0 = no)	Binaria	1
Cumple tiempos de respuesta	Indicador de cumplimiento según estándar definido (1 = sí, 0 = no)	Binaria	1

Nota. La tabla resume la estructura de la base de datos utilizada en el análisis.

En la Figura 3 se observa un patrón claramente diurno y unimodal, con una concentración entre 09:00 horas y 13:00 horas, donde se registra el mayor volumen de llamadas. La actividad es marginal durante la madrugada y decrece gradualmente a partir de la tarde, lo que sugiere picos operativos concentrados en la jornada laboral y respalda la necesidad de dimensionar recursos con foco en ese tramo horario.

Figura 3

Distribución de llamadas por hora del día



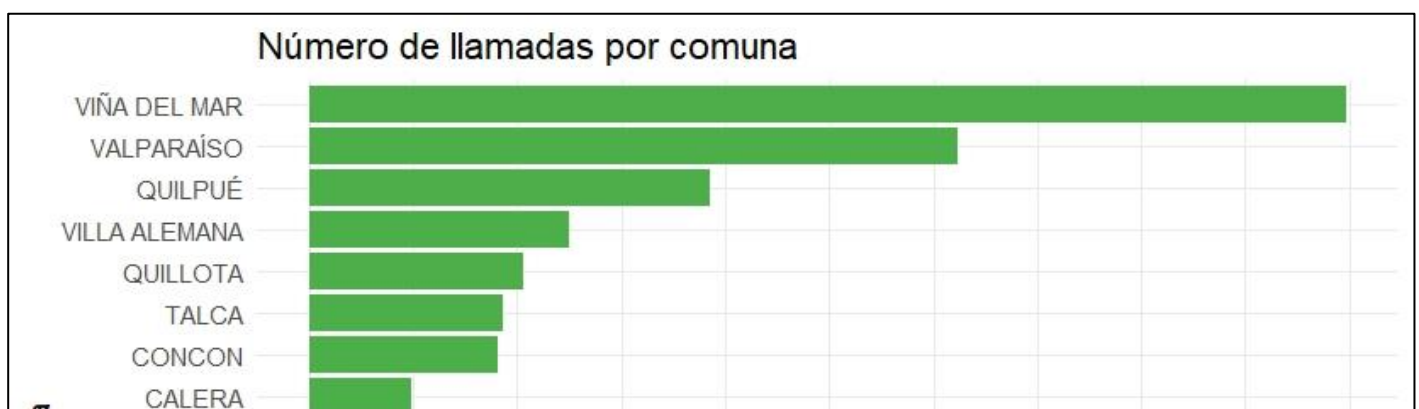
Nota. La figura muestra el conteo de llamadas por hora del día en el período analizado.

Elaboración propia a partir de base de datos.

El volumen está altamente concentrado en unas pocas comunas: Viña del Mar y Valparaíso encabezan los registros, seguidas por Quilpué, Villa Alemana y Quillota. Este patrón territorial indica extremos de demanda que son relevantes para el diseño de turnos como se parecía en la Figura 4, la asignación de móviles y potenciales puntos de ubicación antes de la asignación.

Figura 4

Número de llamadas por comuna



Nota. La figura presenta el total de llamadas por comuna, se evidencia concentración en centros urbanos de mayor densidad. Elaboración propia a partir de base de datos.

En la Figura 5 predominan las categorías Fuga y Olor a gas, seguidas por Sin suministro y CODEC. La presencia de casos NA (no clasificados) sugiere la conveniencia de depurar o reclasificar esos registros para fortalecer los análisis por tipo. La jerarquía observada orienta prioridades operativas y focos de capacitación técnica.

Figura 5

Distribución por tipo de emergencia



Nota. La figura muestra la distribución de llamadas por tipo de emergencia. Elaboración propia a partir de base de datos.

Para los tiempos de respuesta según tipo de emergencia, se aprecian diferencias sistemáticas entre categorías. “Sin suministro” exhibe mayores tiempos medianos y mayor dispersión, lo que sugiere complejidad operativa (diagnóstico y resolución más extensos). “Inflamación” presenta tiempos bajos y consistentes, coherentes con su alta prioridad. “Fuga” y “Olor a gas” muestran medianas elevadas con variabilidad moderada como se aprecia en la Figura 6. Los puntos atípicos indican eventos puntuales con demoras prolongadas que conviene revisar, pueden ser por logística, distancia, coordinación entre otras.

Figura 6

Tiempos de respuesta por tipo de emergencia

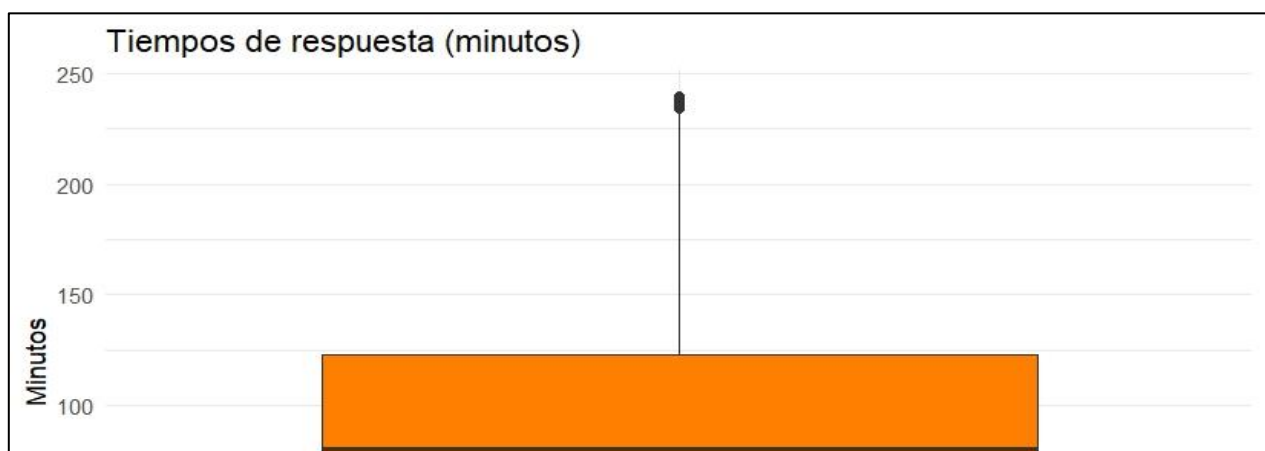


Nota. Los diagramas de Boxplots resumen la mediana, el rango intercuartílico y valores atípicos de los tiempos de respuesta por tipo. Elaboración propia a partir de base de datos.

La distribución global según la Figura 7 presenta mediana en torno a 80 minutos y un rango intercuartílico amplio, aproximadamente del orden de 50 a 120 minutos, con outliers cercanos a las 4 horas. Este comportamiento evidencia variabilidad significativa y refuerza la pertinencia de evaluar escenarios de dotación de móviles y umbrales de cumplimiento en el capítulo.

Figura 7

Tiempo de respuesta promedio



Nota. El diagrama de Boxplots global sintetiza la distribución de los tiempos de respuesta en todo el período. Elaboración propia a partir de base de datos.

El gráfico de la Figura 8 muestra la proporción de llamadas que cumplen y no cumplen con el estándar de tiempo de respuesta. Se observa que solo un 34% de los casos cumplen, mientras que el 66% restante no lo hace, lo que refleja un nivel de servicio insuficiente respecto al estándar definido. Esta evidencia confirma la necesidad de explorar alternativas de planificación de capacidad que permitan aumentar el cumplimiento y reducir los tiempos de espera.

Figura 8

Cumplimiento global



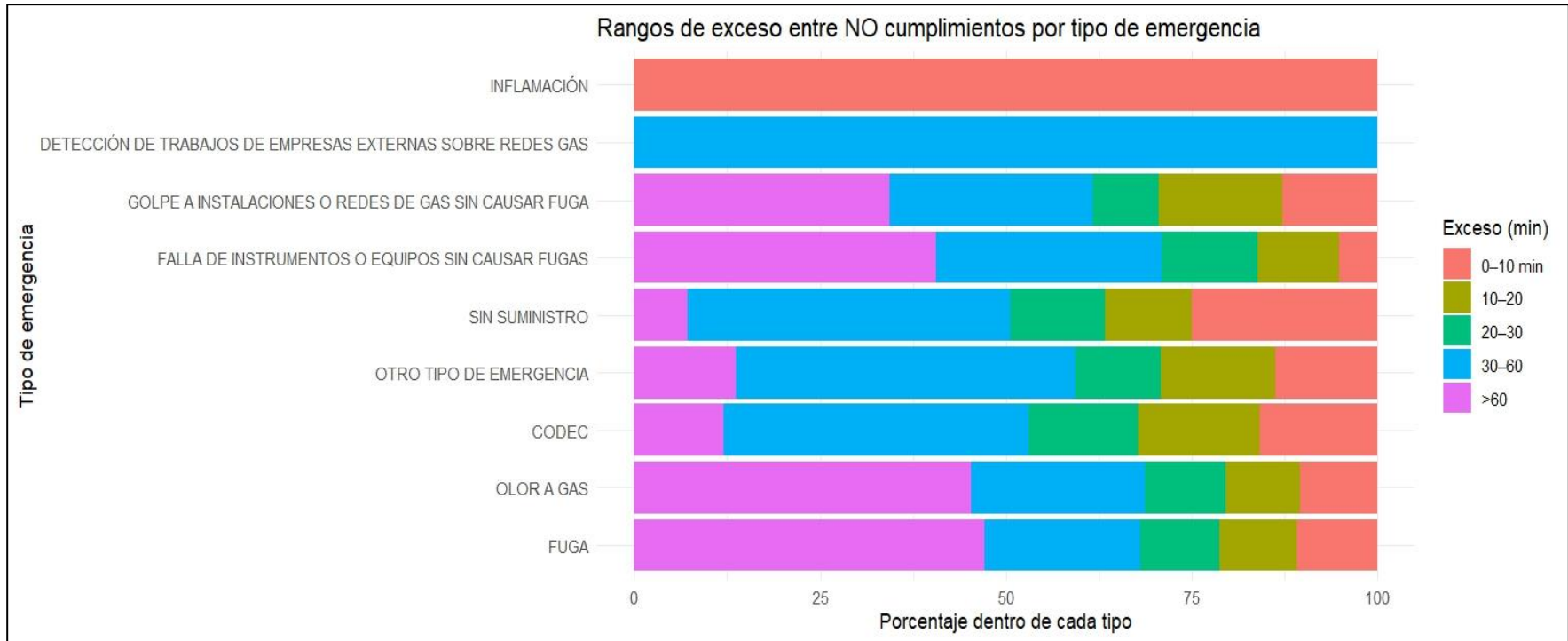


Nota. Cumplimiento de tiempos de respuesta (global). Elaboración propia a partir de base de datos.

La Figura 9 presenta la distribución de los incumplimientos según el exceso de tiempo registrado. Si bien existe un grupo cercano al 11% de llamadas que superan el límite en un rango corto (0–10 minutos), un 40% de los casos presentan excesos mayores a 60 minutos, lo que implica una brecha significativa entre la meta de atención y la realidad operativa. Estos valores extremos no solo comprometen el nivel de servicio, sino también la seguridad y satisfacción de los clientes.

Figura 9

Distribución de los excesos en minutos

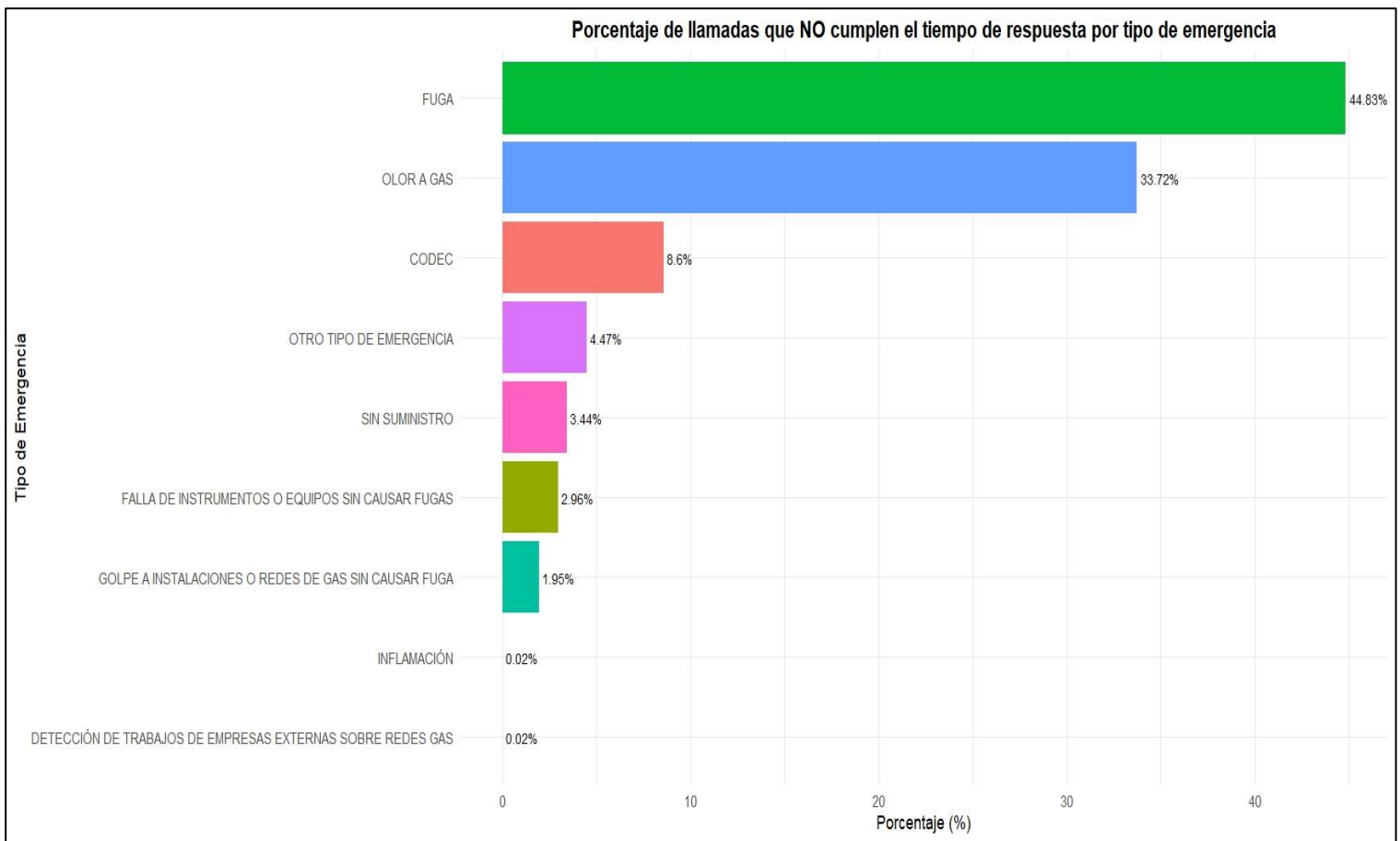


Nota. Rangos de exceso de tiempo en llamadas no cumplidas. Elaboración propia en programa RStudio a partir de base de datos.

En la Figura 10 se observa que las emergencias más críticas (Fuga y Olor a gas) concentran la mayor parte de los incumplimientos (44,8% y 33,7%, respectivamente). Esto resulta especialmente relevante, ya que estos tipos de emergencia demandan una respuesta rápida para evitar riesgos mayores. El hecho de que sean justamente las categorías más incumplidas refuerza la necesidad de evaluar escenarios de mayor dotación de móviles.

Figura 10

Porcentaje de llamadas no cumplidas según tipo de emergencia



Nota. Elaboración propia en programa RStudio a partir de base de datos.

La Tabla 2 presenta la distribución de las llamadas que no cumplieron con el tiempo de respuesta, clasificadas según el rango de exceso en minutos. Se observa que un 11,6% de los casos superó el límite por un margen reducido de hasta 10 minutos, mientras que un 25,8% lo hizo en un rango de entre 30 y 60 minutos. Sin embargo, el grupo más crítico corresponde a los casos con excesos superiores a 60 minutos, que representan un 40,2% del total de incumplimientos. Este resultado evidencia que, además de la alta proporción de llamadas que no cumplen con el estándar, una parte significativa presenta demoras prolongadas que pueden comprometer tanto la eficiencia operativa como la seguridad del servicio.

Tabla 2

Distribución de subconjuntos de llamadas incumplidas según exceso de tiempo de respuesta

Rango de exceso de tiempo	n	Porcentaje (%)
0–10 min	606	11.6
10–20 min	585	11.2
20–30 min	588	11.2
30–60 min	1351	25.8
>60 min	2105	40.2

Nota. Elaboración propia a partir de base de datos.

En la Figura 11 se muestra la tasa de no cumplimiento de los tiempos de respuesta por comuna. Se identifican casos críticos en comunas como Rancagua y La Higuera, donde el incumplimiento alcanza el 100% pero solo porque la cantidad de llamada fue 1 y no se cumplió. Les siguen localidades intermedias como Llay Llay, Calera y Villa Alemana, con tasas sobre el 70%. En cambio, comunas como Catemu, Curauma y Los Andes exhiben menores niveles de incumplimiento (50% o menos). Este patrón revela que la variabilidad territorial es un factor relevante, lo que refuerza la necesidad de ajustar los recursos de acuerdo con la demanda.

Figura 11

Porcentaje de no cumplimiento de tiempos de respuesta por comuna

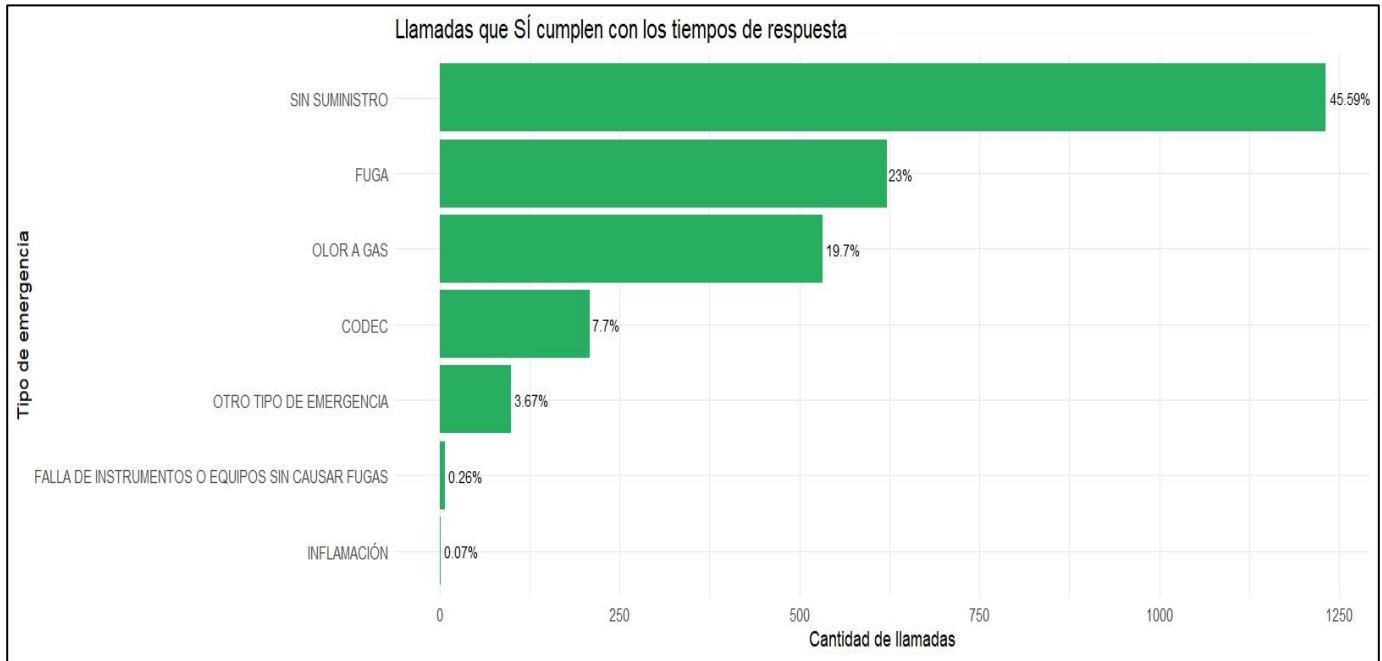


Nota. Elaboración propia a partir de base de datos.

En la Figura 12 se observa la distribución de las llamadas que sí cumplen con los tiempos de respuesta según el tipo de emergencia. Se aprecia que los eventos de “sin suministro” concentran la mayor proporción de cumplimientos (45,6%), esto se explica, en parte, porque el tiempo de respuesta para este tipo de llamadas es mayor (180 minutos), seguidos por las fugas (23%) y los olores a gas (19,7%). En contraste, emergencias menos frecuentes como falla de instrumentos, inflamaciones u otros tipos presentan niveles marginales de cumplimiento. Estos resultados sugieren que el cumplimiento tiende a concentrarse en emergencias de mayor frecuencia relativa, posiblemente debido a procedimientos más estandarizados de atención.

Figura 12

Llamadas que si cumplen con los tiempos de respuesta según tipo de emergencia



Nota. Elaboración propia a partir de base de datos.

En la Tabla 3 se presenta la distribución de llamadas según su focalización. Se observa que la mayor parte de los eventos corresponden a emergencias en interior (42%) y en exterior o medidor (40,1%), mientras que el 17,9% restante se asocia a otras categorías. Este patrón confirma que la demanda se concentra principalmente en espacios residenciales e instalaciones de conexión, lo cual refuerza la necesidad de evaluar con detalle los tiempos de atención en estas áreas críticas. Complementariamente, la Tabla 4 resume las estadísticas descriptivas de los tiempos de respuesta. El promedio alcanzó los 93,9 minutos, con una mediana de 80,6 minutos y una desviación estándar de 57,7 minutos, lo que evidencia una alta dispersión en los datos.

El 50% central de las observaciones se sitúa entre 48,6 y 125 minutos, aunque se registraron valores extremos que van desde un mínimo de 3 minutos hasta un máximo de 917 minutos (más de 15 horas).

Estos resultados ponen de manifiesto que, además de una elevada proporción de emergencias en interior y exterior, los tiempos de respuesta presentan una marcada variabilidad que debe ser considerada en el diseño de escenarios de capacidad.

Tabla 3

Distribución de llamadas por focalización

Focalización	n	Porcentaje (%)
Interior	3331	42,0
Exterior/Medidor	3177	40,1
Otras	1415	17,9

Nota. Elaboración propia a partir de base de datos.

Tabla 4

Estadísticas descriptivas de tiempos de respuesta

n	Media (min)	Mediana (min)	Varianza	Desv. est.	Mínimo	P25	P75	Máximo
7936	93,9	80,6	3326	57,7	3	48,6	125	917

Nota. Elaboración propia a partir de base de datos.

La Figura 13 y la Tabla 5 se muestran las diferencias en los niveles de cumplimiento de tiempos de respuesta entre los distintos grupos de comunas. Se aprecia que el Grupo 5 (Talca) alcanza el porcentaje más alto de cumplimiento (42,2%), seguido por el Grupo 4 (La Serena y Coquimbo) con un 39,1%.

En contraste, los grupos con mayor carga de emergencias presentan porcentajes considerablemente más bajos: el Grupo 1, que concentra el 72% de todas las llamadas, exhibe un cumplimiento de apenas 34,0%, mientras que el Grupo 3 (Llay Llay y Los Andes) evidencia el nivel más crítico, con solo un 26,8%.

El Grupo 2 (Quillota, Calera y La Cruz) se ubica en un nivel intermedio, con un cumplimiento de 29,9%.

Estos resultados permiten identificar que, aunque algunos grupos logran desempeños relativamente más favorables, el impacto global del sistema se ve determinado casi por completo por el comportamiento del Grupo 1, debido a su altísima concentración de demanda. En otras palabras, mejorar los tiempos de respuesta en este grupo tendría un efecto desproporcionado en la eficiencia global del sistema, mucho más significativo que el que podría alcanzarse al intervenir en grupos de menor volumen.

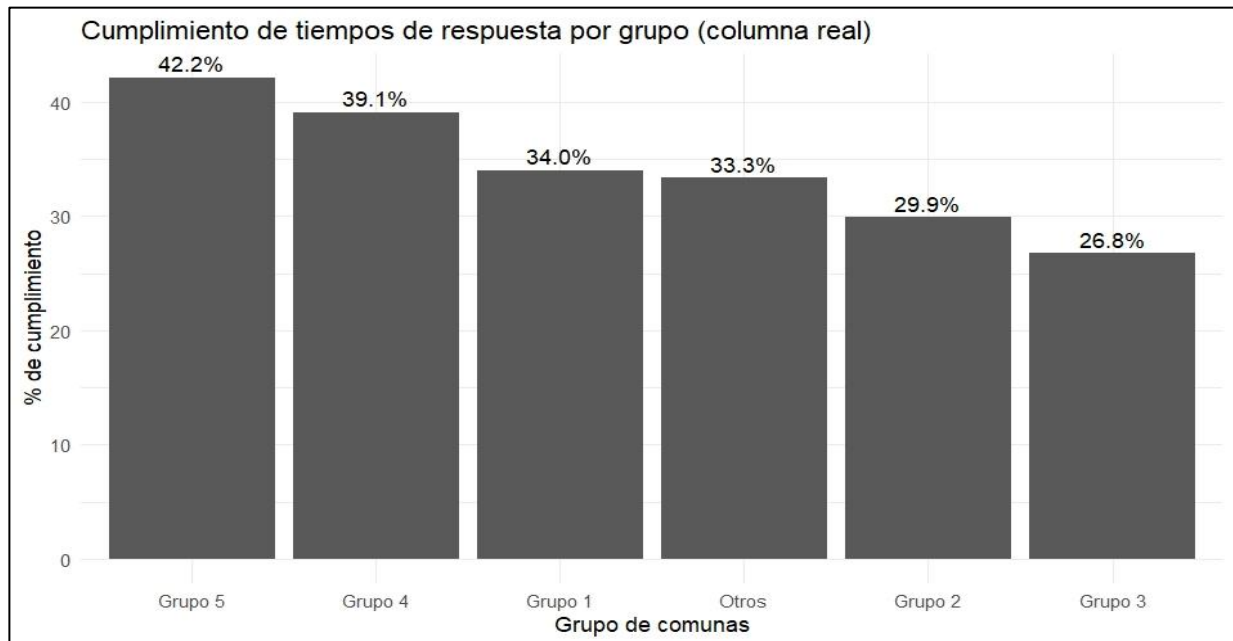
En este sentido, los esfuerzos de planificación de capacidad y asignación de móviles deben priorizar estrategias focalizadas en el Grupo 1, ya que allí se concentran las mayores brechas de cumplimiento y, al mismo tiempo, se genera el mayor número absoluto de emergencias. Asimismo, resulta relevante considerar que los grupos de menor volumen (como el Grupo 3) también presentan tasas críticas de incumplimiento, lo que pone de manifiesto la necesidad de buscar un equilibrio entre eficiencia global y equidad territorial.

En conclusión, la evidencia sugiere que cualquier mejora sustantiva en el desempeño del sistema pasará necesariamente por el refuerzo operativo en las comunas que conforman el Grupo 1, al mismo tiempo que se evalúan estrategias complementarias para los grupos con menores volúmenes, pero que muestran rezagos importantes en el cumplimiento de tiempos de respuesta.

Tabla 5*Distribución de llamadas y nivel de cumplimiento por grupo de comunas*

Grupo	Llamadas	Llamadas/día	Cumplimiento (%)
Grupo 1 (Viña, Valpo, Quilpué, Villa Alemana, Concón, Casablanca, Curauma)	6.222 (72%)	8,23	34,0
Grupo 2 (Quillota, Calera, La Cruz)	866 (10%)	1,85	29,9
Grupo 3 (Llay Llay, Los Andes)	142 (2%)	1,09	26,8
Grupo 4 (La Serena, Coquimbo)	238 (3%)	1,25	39,1
Grupo 5 (Talca)	465 (5%)	1,60	42,2
Otros (marginal)	3	≈0	33,3

Nota. Elaboración propia a partir de base de datos.

Figura 13*Nivel de cumplimiento por grupo de comunas*

Nota. Cumplimiento de tiempos ordenados. Elaboración propia a partir de base de datos.

El análisis de simultaneidad de eventos permitió estimar la probabilidad de que ocurran múltiples llamadas dentro de un mismo intervalo temporal. Este procedimiento, descrito previamente en la sección metodológica, se basó en la comparación de los tiempos de ingreso de las llamadas dentro de ventanas móviles de distinta duración (10, 30, 60 y 120 minutos), identificando aquellas que se superponen en el tiempo.

La Tabla 6 muestra la distribución de llamadas analizadas bajo distintos tamaños de ventana. El número total de eventos ($n = 7.936$) corresponde al total de registros válidos evaluados, mientras que la columna “N° simultáneos” indica cuántos de esos eventos presentaron coincidencia temporal con al menos otra llamada dentro del intervalo considerado. Por tanto, no significa que existan miles de eventos ocurriendo exactamente al mismo tiempo, sino que en una ventana de, por ejemplo, 120 minutos, el 74,6% de las llamadas coincidió temporalmente con otra dentro de ese lapso de tiempo.

Como se observa, la probabilidad de simultaneidad aumenta progresivamente a medida que se amplía la ventana temporal: un 17,7% de las llamadas presenta superposición en una ventana de 10 minutos, un 55,4% en una de 60 minutos y un 74,6% en una de 120 minutos.

Estos resultados ponen de manifiesto que la simultaneidad constituye un factor crítico en la planificación de recursos, pues incrementa la presión sobre la capacidad de respuesta. En la práctica, implica que disponer de un número limitado de móviles puede no ser suficiente para atender los incidentes dentro de los tiempos establecidos, especialmente en horarios o comunas donde la ocurrencia de llamados es más densa.

Tabla 6*Distribución de llamadas por simultaneidad según ventanas de tiempo*

Ventana (min)	N° eventos	N° simultáneos	Probabilidad de simultaneidad
10	7.936	1.407	0,177
20	7.936	2.245	0,283
30	7.936	2.937	0,370
40	7.936	3.490	0,440
50	7.936	4.016	0,506
60	7.936	4.394	0,554
70	7.936	4.750	0,599
80	7.936	5.032	0,634
90	7.936	5.289	0,666
100	7.936	5.531	0,697
110	7.936	5.747	0,724
120	7.936	5.918	0,746

Nota. Elaboración propia a partir de base de datos.

La probabilidad de simultaneidad de llamadas de emergencia en todas las comunas incluidas en la base de datos se presenta en la Tabla 7, considerando intervalos de 30, 60 y 120 minutos. Los resultados permiten distinguir con claridad la diferencia entre comunas de alta, media y baja demanda.

En comunas con gran volumen de llamados, como Viña del Mar y Valparaíso, la probabilidad de simultaneidad se ubica en torno al 19% y 15% en 30 minutos, aumentando de manera sostenida hasta alcanzar un 28% y 32% respectivamente en 120 minutos. De forma similar, en comunas intermedias como Quilpué, Quillota, Concón y Talca, las probabilidades se mantienen en niveles moderados (entre 9% y 10% en 30 minutos), pero crecen hacia valores cercanos al 20 a 25% en ventanas de 120 minutos.

En contraste, comunas de menor tamaño o baja incidencia, como Calera, La Cruz, Casablanca o La Serena, muestran probabilidades reducidas de simultaneidad, que rara vez superan el 10% incluso en horizontes de 120 minutos. Más extremo aún es el caso de Catemu, La Higuera, Los Andes y Rancagua, donde los registros de simultaneidad son prácticamente inexistentes debido al bajo número de eventos observados.

Estos hallazgos refuerzan la idea de que la simultaneidad es un fenómeno estructural en las comunas de mayor densidad de emergencias, donde el dimensionamiento de móviles debe considerar que atender múltiples llamadas en paralelo no es una excepción, sino una situación recurrente. En contraste, para comunas periféricas o con baja carga, la simultaneidad representa un riesgo marginal, aunque no completamente descartable en ventanas temporales más largas.

Tabla 7

Probabilidad de simultaneidad de llamadas por comuna y tiempos

Comuna	Llamadas (n)	30 min	60 min	120 min
Viña del Mar	2.493	0,1890	0,2760	0,4130
Valparaíso	1.558	0,1550	0,2180	0,3220
Quilpué	962	0,0977	0,1600	0,2520
Villa Alemana	623	0,0514	0,0915	0,1410
Quillota	515	0,0951	0,1110	0,1570
Concón	454	0,0947	0,1190	0,1610
Talca	465	0,0989	0,1480	0,2280
La Serena	208	0,0577	0,0721	0,1150
Calera	243	0,0247	0,0370	0,0658
La Cruz	108	0,0556	0,0648	0,0926
Casablanca	118	0,0169	0,0254	0,0508
Curauma	14	0,1430	0,1430	0,1430
Coquimbo	30	0,0333	0,0667	0,0667
Llay Llay	116	0,0086	0,0172	0,0259
Los Andes	26	0,0000	0,0000	0,0000
La Higuera	1	0,0000	0,0000	0,0000
Rancagua	1	0,0000	0,0000	0,0000
Catemu	1	0,0000	0,0000	0,0000

Nota. Elaboración propia a partir de base de datos.

Tabla 8*Resumen de cumplimientos de tiempos de respuesta por comuna y focalización*

Comuna	Focalización	Llamadas	Promedio (min)	Estándar (min)	Diferencia (Prom. – Estándar)
Viña del Mar	Interior	1118	89,9	60	+29,9
	Medidor	639	88,9	30	+58,9
	Exterior	254	92,2	30	+62,2
	Sin Suministro	482	108,0	—	—
Valparaíso	Interior	674	87,9	60	+27,9
	Medidor	393	92,2	30	+62,2
	Exterior	227	90,0	30	+60,0
	Sin Suministro	264	111,0	—	—
Quilpué	Interior	369	91,6	60	+31,6
	Medidor	280	87,7	30	+57,7
	Exterior	186	96,1	30	+66,1
	Sin Suministro	127	109,0	—	—
Villa Alemana	Interior	226	91,0	60	+31,0
	Medidor	181	90,4	30	+60,4
	Exterior	136	95,5	30	+65,5
	Sin Suministro	80	112,0	—	—
Talca	Interior	228	90,5	60	+30,5
	Medidor	75	93,1	30	+63,1
	Exterior	39	88,2	30	+58,2
	Sin Suministro	123	101,0	—	—
Quillota	Interior	210	86,1	60	+26,1
	Medidor	144	87,7	30	+57,7
	Exterior	80	102,0	30	+72,0
	Sin Suministro	81	93,5	—	—

Concón	Interior	177	92,8	60	+32,8
	Medidor	121	95,1	30	+65,1
	Exterior	48	95,3	30	+65,3
	Sin Suministro	108	100,0	—	—
La Serena	Interior	99	88,2	60	+28,2
	Medidor	46	81,0	30	+51,0
	Exterior	15	104,0	30	+74,0
	Sin Suministro	48	107,0	—	—
Calera	Interior	91	96,4	60	+36,4
	Medidor	79	99,9	30	+69,9
	Exterior	45	112,0	30	+82,0
	Sin Suministro	28	127,0	—	—
La Cruz	Interior	30	89,1	60	+29,1
	Medidor	43	95,9	30	+65,9
	Exterior	7	99,7	30	+69,7
	Sin Suministro	28	123,0	—	—
Llay Llay	Interior	38	94,1	60	+34,1
	Medidor	45	94,2	30	+64,2
	Exterior	21	75,2	30	+45,2
	Sin Suministro	12	96,6	—	—
Los Andes	Interior	6	108,0	60	+48,0
	Medidor	7	123,0	30	+93,0
	Exterior	4	75,8	30	+45,8
	Sin Suministro	9	107,0	—	—
Casablanca	Interior	47	91,9	60	+31,9
	Medidor	35	99,3	30	+69,3
	Exterior	19	81,3	30	+51,3
	Sin Suministro	17	131,0	—	—
Curauma	Interior	6	68,8	60	+8,8
	Medidor	5	126,0	30	+96,0

	Exterior	1	17,0	30	-13,0
	Sin Suministro	2	99,5	—	—
Coquimbo	Interior	21	97,0	60	+37,0
	Medidor	4	116,0	30	+86,0
	Sin Suministro	5	97,0	—	—
Otros	Catemu – Sin Suministro	1	115,0	—	—
	La Higuera – Exterior	1	30,8	30	+0,8
	Rancagua – Interior	1	176,0	60	+116,0

Nota. Elaboración propia a partir de base de datos.

8.3 Análisis Econométrico

Con el objetivo de complementar la evidencia obtenida en el análisis exploratorio, se estimaron modelos econométricos Logit y Probit orientados a explicar dos aspectos centrales: la ocurrencia de emergencias y el cumplimiento de los tiempos de respuesta.

8.3.1 Modelos de Clasificación Binaria

8.3.1.1 Modelos para Explicar la Ocurrencia de Emergencias Reales.

i. Modelo Logit

El modelo Logit se estimó con el objetivo de explicar la probabilidad de que una llamada corresponda a una emergencia real, utilizando como variables explicativas la comuna, el tipo de emergencia, la focalización y la hora de recepción. Los resultados indican que varias comunas presentan un efecto estadísticamente significativo sobre la probabilidad de ocurrencia.

En particular, Concón, La Serena, Talca, Valparaíso, Villa Alemana y Viña del Mar destacan como comunas en las que la probabilidad de emergencia real difiere significativamente respecto de la comuna base (Calera).

En cambio, las variables tipo de emergencia y focalización no resultaron significativas. Esto puede atribuirse a fenómenos de concentración en algunas categorías, que generan problemas de separación cuasi-perfecta en la estimación. La hora de recepción tampoco fue significativa, aunque se observó una leve tendencia negativa. En términos de ajuste, el modelo alcanzó un AIC de 2026, con una reducción importante de la deviance residual (1970) respecto a la nula (10347), lo que evidencia que el modelo logra explicar una fracción considerable de la variabilidad observada.

ii. Modelo Probit

El modelo Probit se estimó bajo las mismas condiciones que el Logit, y arrojó resultados muy similares como se pueden apreciar en el capítulo de anexos, en las dos primeras imágenes del apéndice C. Al igual que en el modelo anterior, las comunas Concón, La Serena, Talca, Valparaíso, Villa Alemana y Viña del Mar se identifican como significativas en la explicación de la probabilidad de ocurrencia de emergencias reales. Del mismo modo, las variables de tipo de emergencia, focalización y hora de recepción no mostraron significancia estadística.

La calidad de ajuste también fue comparable, con un AIC de 2026.2 y deviance residual de 1970.2, prácticamente idénticos a los del modelo Logit. Esto confirma la consistencia de los resultados independientemente de la especificación de la función de enlace.

iii. Comparación entre modelos

Ambos modelos Logit y Probit fueron comparados, donde entregan resultados equivalentes como se aprecia en la Tabla 9 en cuanto a significancia de variables y ajuste global.

La diferencia principal radica en la interpretación: mientras el Probit ofrece una aproximación basada en una distribución normal acumulada, el Logit facilita la interpretación en términos de ratios, lo que resulta más práctico en contextos aplicados.

Tabla 9

Comparativa de métricas para modelos Logit y Probit

Métrica	Logit	Probit
Null deviance	10.347	10.347
Residual deviance	1.970	1.970
AIC	2026,0	2026,2
Variables significativas	Algunas comunas	Mismas comunas
Tipo de emergencia (NS)	No significativa	No significativa
Focalización (NS)	No significativa	No significativa

Nota. Elaboración propia a partir de base de datos.

8.3.1.1 Modelos para Evaluar la Capacidad Predictiva

i. Modelo Logit

La matriz de confusión del modelo Logit muestra que este clasifica correctamente la gran mayoría de las observaciones:

- VP (Verdaderos Positivos): 5085 llamadas reales correctamente clasificadas.
- FP (Falsos Positivos): 548 llamadas clasificadas como emergencias sin serlo.
- FN (Falsos Negativos): 1 llamada de emergencia no identificada.

- VN (Verdaderos Negativos): 2299 llamadas no emergencias correctamente detectadas.

Con base en estos resultados, las métricas de desempeño son:

- Precisión global (accuracy) = 93,1%.
- Sensibilidad (recall) = 99,98%.
- Especificidad = 80,8%.
- Precisión positiva (precision) = 90,3%.

Esto indica que el modelo tiene una capacidad sobresaliente para detectar emergencias reales, con un riesgo prácticamente nulo de dejar sin clasificar un evento crítico (sensibilidad casi perfecta). No obstante, la especificidad relativamente más baja evidencia que aún clasifica como emergencias a un número importante de llamadas que no lo eran, lo cual implica un nivel de falsas alarmas que puede movilizar recursos innecesariamente.

ii. Modelo Probit

La matriz de confusión del modelo Probit entrega resultados prácticamente idénticos a los del Logit:

- VP (Verdaderos Positivos): 5083 llamadas reales correctamente clasificadas.
- FP (Falsos Positivos): 549 llamadas clasificadas como emergencias sin serlo.
- FN (Falsos Negativos): 3 llamadas de emergencia no identificadas.
- VN (Verdaderos Negativos): 2298 llamadas no emergencias correctamente detectadas.

A partir de estas cifras, las métricas obtenidas son equivalentes a las del modelo Logit:

- Precisión global (accuracy): 93,1%.
- Sensibilidad (recall): 99,98%.
- Especificidad: 80,8%
- Precisión positiva (precision): 90,3%.

En este caso, también se confirma que la principal fortaleza del modelo radica en su capacidad para identificar emergencias reales, reduciendo al mínimo los falsos negativos. Al igual que el Logit, el costo de este buen desempeño es una menor capacidad para discriminar llamadas que no corresponden a emergencias, lo que genera un nivel considerable de falsos positivos.

iii. Comparación entre modelos

Tanto el modelo Logit como el modelo Probit presentan un desempeño prácticamente idéntico en términos de clasificación y métricas de evaluación. Ambos logran una sensibilidad casi perfecta, lo que garantiza que las emergencias reales sean detectadas con alta seguridad. Sin embargo, ambos comparten también la limitación de una especificidad más baja, lo que significa que clasifican erróneamente como emergencias a un número no despreciable de llamadas que no lo son. En términos prácticos, se opta por privilegiar el modelo Logit como base del análisis, dado que permite interpretar los coeficientes en términos de ratios, lo que facilita la aplicación en escenarios operativos. No obstante, la similitud de resultados entre ambos modelos valida la robustez del análisis.

8.3.1.2 Modelos para Explicar el Cumplimiento de Tiempos de Respuesta

i. Modelo Logit

El modelo Logit ajustado para explicar la probabilidad de cumplir los tiempos de respuesta muestra varios predictores estadísticamente significativos. En cuanto a las comunas, destacan Curauma ($p=0,03$) y Valparaíso ($p=0,026$), ambas con efecto positivo sobre la probabilidad de cumplimiento.

Además, Quillota, Quilpué y Viña del Mar se encuentran en el límite de la significancia ($p<0,10$), sugiriendo un posible efecto positivo. El resto de las comunas no presenta influencia

estadísticamente relevante. Entre los tipos de emergencia, las fallas de instrumentos o equipos sin causar fugas muestran un efecto negativo y significativo ($p=0,03$), reduciendo la probabilidad de cumplir los tiempos. En tanto, las emergencias clasificadas como fuga quedan en el borde de la significancia ($p=0,06$), con un efecto positivo.

Respecto a la focalización, la categoría Interior resulta altamente significativa ($p<0,001$), aumentando considerablemente la probabilidad de cumplimiento. Por su parte, Medidor muestra un efecto negativo cercano al límite de significancia ($p=0,07$), mientras que Sin Suministro no presenta relevancia estadística. La variable hora de recepción presenta un coeficiente negativo, aunque sin significancia estadística.

En conjunto, el modelo Logit ($AIC = 6927$) destaca la importancia de la localización geográfica y de la focalización del evento como factores clave en el cumplimiento de los tiempos de respuesta

ii. Modelo Probit

El modelo Probit entrega resultados muy similares a los del Logit, reforzando la consistencia de los hallazgos. En este caso, nuevamente Curauma ($p=0,026$) y Valparaíso ($p=0,018$) presentan un efecto positivo significativo, mientras que Quillota, Quilpué y Viña del Mar se mantienen cercanas al umbral de significancia ($p=0,06-0,08$).

En relación con los tipos de emergencia, las fallas de instrumentos mantienen un efecto negativo y significativo ($p=0,016$). Sin embargo, a diferencia del Logit, aquí las fugas sí aparecen como significativas y con efecto positivo ($p=0,046$).

El resto de las categorías no presenta efectos relevantes. En cuanto a la focalización, la categoría Interior vuelve a ser altamente significativa y positiva ($p<0,001$), confirmándose como el

predictor más influyente. Medidor muestra un efecto negativo cercano al límite ($p=0,08$), mientras que Sin Suministro no resulta significativo. La hora de recepción mantiene un efecto negativo, aunque no significativo.

El ajuste del Probit (AIC = 6923) es levemente mejor que el del Logit, aunque sin implicar una diferencia práctica relevante como se aprecia en los resultados ilustrados en el capítulo de anexos, en la tercera y cuarta imagen del apéndice C.

iii. Comparación entre modelos

Ambos modelos Logit y Probit se comparan según Tabla 10, se presentan patrones según la base de datos: la focalización Interior y las comunas Curauma y Valparaíso están asociadas a una mayor probabilidad de cumplir los tiempos de respuesta, mientras que las emergencias por fallas de instrumentos reducen significativamente esta probabilidad. El Probit aporta la diferencia de identificar como significativa la categoría fuga, mientras que en el Logit se mantiene en el límite de la significancia.

Dado que ambos modelos ofrecen resultados prácticamente equivalentes y con buena capacidad explicativa, se privilegia el uso del Logit como modelo principal.

Tabla 10*Comparativa de métricas para modelos Logit y Probit*

Métrica	Logit	Probit
Null deviance	10.196,7	10.196,7
Residual deviance	6.873,3	6.869,8
AIC	6.927,3	6.923,8
Comunas significativas	Curauma, Valparaíso	Curauma, Valparaíso
Comunas cercanas ($p < 0,10$)	Quillota, Quilpué, Viña del Mar	Quillota, Quilpué, Viña del Mar
Tipo de emergencia	Falla de instrumentos	Falla de instrumentos, Fuga
Focalización	Interior, Medidor	Interior, Medidor
Hora de recepción	No significativa	No significativa

Nota. Elaboración propia a partir de base de datos.

8.3.1.2.1 Modelos para Evaluar la Capacidad Predictiva

i. Modelo Logit

La matriz de confusión para el modelo Logit muestra el siguiente desempeño:

- Verdaderos positivos (TP): 1251
- Falsos positivos (FP): 172
- Falsos negativos (FN): 1466
- Verdaderos negativos (TN): 5044

A partir de estos valores, las métricas de desempeño son:

- Precisión global (Accuracy): 0,794, el modelo clasifica correctamente cerca del 79% de las observaciones.
- Sensibilidad (Recall): 0,460, logra identificar solo el 46% de los casos que efectivamente cumplieron con el tiempo, es decir, pierde más de la mitad de los positivos.
- Especificidad: 0,967, clasifica correctamente el 96,7% de los casos que no cumplieron con el tiempo, mostrando gran capacidad para identificar la clase mayoritaria.
- Precisión positiva (Precision): 0,879, cuando predice que se cumple el tiempo, acierta en casi un 88% de los casos.
- F1 Score: 0,606, balance moderado entre precisión y sensibilidad.

En síntesis, el Logit ofrece un desempeño aceptable en términos generales, pero su fortaleza está en reconocer los incumplimientos (alta especificidad). En cambio, su debilidad es detectar los cumplimientos, lo cual se explica por el desbalance de la variable dependiente (66% “no cumple” vs 34% “cumple”).

ii. Modelo Probit

La evaluación del modelo Probit, bajo las mismas condiciones (sin balanceo y con $\tau=0,5$), entrega resultados idénticos al Logit en este escenario:

- Verdaderos positivos (TP): 1251
- Falsos positivos (FP): 172
- Falsos negativos (FN): 1466
- Verdaderos negativos (TN): 5044

Con las métricas:

- Precisión global 0,794
- Sensibilidad: 0,460
- Especificidad: 0,967
- Precisión Positiva: 0,879
- F1 Score: 0,606

La interpretación es equivalente: el modelo detecta con gran confiabilidad los casos que no cumplen, y cuando predice cumplimiento suele acertar. Sin embargo, su sensibilidad es baja, por lo que se pierden más de la mitad de los casos positivos.

Tanto el modelo Logit como el Probit muestran un desempeño general satisfactorio (79% de precisión global), pero comparten una limitación crítica: la baja capacidad de detectar todos los cumplimientos (sensibilidad 46%). En cambio, destacan por su alta especificidad (97%), reflejando un sesgo hacia la clase mayoritaria (“no cumple”). Esto sugiere que, sin técnicas adicionales (balanceo o ajuste de umbral), ambos modelos tienden a favorecer la detección de incumplimientos sobre la de cumplimientos.

8.3.1.2.2 Modelos para Evaluar la Capacidad Predictiva (Reducción de Umbral a 0,3).

i. Modelo Logit

Al reducir el umbral de decisión a 0,30, la matriz de confusión cambia sustancialmente, y con ello las métricas:

- Precisión global (Accuracy): 0,700, la proporción de observaciones clasificadas correctamente cae al 70%.
- Sensibilidad (Recall): 0,917, el modelo ahora detecta la gran mayoría de los casos en que efectivamente se cumplió el tiempo.
- Especificidad: 0,587, la capacidad para identificar los incumplimientos disminuye, confundiendo más casos “no cumple” como si hubieran cumplido.
- Precisión positiva (Precision): 0,536, cuando el modelo predice cumplimiento, solo acierta en poco más de la mitad de las veces.
- F1 Score: 0,677, refleja un mejor balance entre precisión y sensibilidad que en el escenario base, gracias al aumento significativo del recall

En resumen, el Logit ajustado con umbral bajo es mucho más eficaz para detectar los cumplimientos, pero a costa de generar una mayor cantidad de falsos positivos.

ii. Modelo Probit

El Probit, bajo el mismo ajuste ($\tau = 0,30$), arroja métricas prácticamente idénticas:

- Precisión Global: 0,700
- Sensibilidad: 0,917
- Especificidad: 0,586
- Precisión Positiva: 0,536
- F1 Score: 0,677.

Esto confirma que, al igual que el Logit, el Probit mejora notablemente la sensibilidad, asegurando que casi todos los cumplimientos sean identificados. No obstante, se sacrifica la especificidad y la precisión de las predicciones positivas.

iii. Conclusión del escenario con umbral reducido

Ambos modelos (Logit y Probit) presentan el mismo patrón: alta sensibilidad (92%) y bajo desempeño en especificidad y precisión positiva (59% y 54%). En términos prácticos, esta configuración es útil si el interés principal es no dejar pasar cumplimientos, aceptando como costo un mayor número de falsas alarma.

8.3.1.2.3 Modelos para Evaluar la Capacidad Predictiva (Con Balance de Datos de 1/3 y Bajo Umbral).

i. Modelo Logit

La matriz de confusión del modelo Logit balanceado y con umbral optimizado muestra:

- Verdaderos negativos (TN): 985
- Falsos positivos (FP): 186
- Falsos negativos (FN): 1754
- Verdaderos positivos (TP): 2531

De estas cifras se derivan las siguientes métricas:

- Accuracy (precisión global): 0,789, el 79% de las observaciones fueron clasificadas correctamente.
- Recall (sensibilidad, clase 1): 0,931, detecta el 93% de los casos que sí cumplen con el tiempo.
- Specificity (clase 0): 0,566, reconoce bien al 56% de los que no cumplen.

- Precision (PPV): 0,770, cuando predice que se cumple, acierta el 77% de las veces.
- F1 Score: 0,843, refleja un excelente balance entre precisión y sensibilidad.

Esto significa que el modelo Logit logra capturar prácticamente todos los cumplimientos, con una precisión positiva aceptable y un equilibrio superior al de los escenarios previos.

ii. Modelo Probit

Los resultados del Probit bajo el mismo escenario son idénticos, dado que ambos modelos convergen al mismo patrón de clasificación con el balanceo aplicado y el umbral optimizado. Se obtienen exactamente las mismas métricas:

- Precisión Global: 0,789
- Sensibilidad: 0,931
- Especificidad: 0,566
- Precisión Positiva: 0,770
- F1 Score: 0,843

Tabla 11

Comparativa de métricas para modelos Logit y Probit con balanceo de datos

Métrica	Original (Logit/Probit, $\tau=0,5$)	Balanceado (1/3, $\tau=0,3$)
Accuracy	0,794	0,789
Recall	0,462	0,932
Specificity	0,967	0,566
Precision	0,878	0,770
F1 Score	0,606	0,843

Nota. Elaboración propia a partir de base de datos.

iii. Conclusión del escenario

La comparación entre el modelo sin balanceo, $\tau=0,5$ y el modelo balanceado con umbral optimizado ($\tau=0,3$) muestra diferencias en la capacidad predictiva. El modelo original estaba fuertemente sesgado hacia predecir no cumple (0), logrando una especificidad muy alta (97%), pero con una sensibilidad baja (46%), lo que significa que se perdía más de la mitad de los cumplimientos.

Con el balanceo (1/3 de ceros) y el ajuste de umbral, el comportamiento cambia: la sensibilidad aumenta (93%), asegurando que casi todos los cumplimientos sean detectados. Aunque se pierde especificidad (56%), el F1 mejora notablemente a 0,843, lo que refleja un equilibrio mucho más sano entre precisión y sensibilidad. Las ventajas y desventajas de solo bajar el umbral (0,15 a 0,30):

- Ventaja: utiliza todo el dataset y mejora el recall (ejemplo 0,88).
- Desventaja: la precisión se desploma (0,36 en la prueba), y el desbalance de clases se mantiene.

Las ventajas y desventajas del balanceo más reducción de umbral (0,3):

- Ventaja: el modelo actúa en un entorno menos sesgado, logrando un mejor equilibrio entre detectar cumplimientos y acertar en sus predicciones positivas.
- Desventaja: se descarta parte de la información (2/3 de la clase negativa), y se depende de que la distribución real no sea mucho más desbalanceada.

En síntesis, si el objetivo es exploratorio y se busca usar todos los datos, basta con bajar el umbral. Sin embargo, lo buscado es un modelo predictivo útil para la práctica operativa donde lo más importante es detectar cumplimientos de tiempo sin que la precisión caiga drásticamente, la estrategia de balanceo más ajuste de umbral es claramente superior.

Este escenario representa el mejor desempeño global entre todos los evaluados. Tanto el modelo Logit como el Probit logran una sensibilidad muy alta (93%), asegurando que casi todos los cumplimientos sean detectados, al mismo tiempo que mantienen una precisión positiva adecuada (77%). Aunque la especificidad cae al 56%, el equilibrio global ($F1 = 0,843$) lo convierte en la opción más adecuada para la planificación operativa, donde es preferible aceptar algunos falsos positivos antes que dejar sin identificar los cumplimientos de tiempo.

8.4 Estimación Distribución de Probabilidad

El análisis de la demanda de llamadas se realizó desde diferentes perspectivas: intervalos entre llamadas, conteos horarios y diarios, y ajuste de modelos de probabilidad, con el fin de caracterizar tanto el promedio como la variabilidad del proceso.

i. Resultados de intervalos entre llamadas

El estudio de los tiempos transcurridos entre llamadas sucesivas evidenció una marcada dispersión: el mínimo fue de 0,0167 minutos (1 segundo) y el máximo de 257.964 minutos (179 días), esto sugiere datos atípicos. La mediana se situó en 51,4 minutos, mientras que la media alcanzó 175,2 minutos, reflejando la influencia de valores atípicos muy altos. La desviación estándar fue de 2.944 minutos, con asimetría positiva (87) y curtosis extremadamente alta (7.621), confirmando la presencia de concentración de datos en valores pequeños y outliers extremos. Entre las distribuciones evaluadas, la Lognormal resultó ser la que mejor describió este comportamiento con un $AIC = 89.016$, superando a la gamma con un $AIC = 91.520,48$ y descartando la exponencial con un $AIC = 95.188,02$.

La Lognormal es la que mejor ajusta los tiempos entre llegadas de llamadas. Esto concuerda con la asimetría y colas largas de espera observadas.

ii. Resultados por horas

En el análisis por hora de 22.536 observaciones, la media fue de 0,35 llamadas/hora, equivalente a 8–9 llamadas diarias. Cuando se excluyen las horas sin llamadas, la media asciende a 1,53, aunque con pérdida de representatividad. Se identificó sobre dispersión clara, dado que la varianza supera ampliamente a la media, por lo cual la distribución Binomial Negativa mostró un ajuste más adecuado con un $AIC = 34.516$, versus la distribución Poisson con un $AIC = 37.410$.

iii. Resultados por días

En los 940 días observados, la media fue de 8,43 llamadas/día, con una mediana de 10 llamadas. La varianza alcanzó 41,23, confirmando nuevamente la sobre dispersión. El mínimo fue de 1 llamada y el máximo de 34, con un rango intercuartílico entre 7 y 13 llamadas diarias. En cuanto a modelos, la Binomial Negativa presentó un ajuste significativamente mejor con un $AIC = 5.993$, frente al Poisson que presentó un $AIC = 8.628$, consolidándose como el modelo más adecuado para representar los conteos de llamadas.

Por lo determinado y relacionado respecto a los datos observados, es que los resultados diarios presentan mayor representatividad respecto al proceso de atención de emergencias.

iv. Pronóstico probabilístico

El modelo identificado y ajustado permitió estimar la probabilidad de ocurrencia de distintos escenarios de demanda diaria:

- 47% de los días, 5 o menos llamadas (baja demanda).
- 23,5% de los días, entre 6 y 10 llamadas (demanda típica).
- 18% de los días, 15 o más llamadas (alta demanda).

Estos resultados reflejan que la mayoría de los días corresponden a escenarios de baja o demanda media, pero que casi 1 de cada 5 días es de alta exigencia, lo que implica la necesidad de capacidad operativa adicional

v. Aplicación práctica

La estimación de la demanda muestra un promedio de 9 llamadas diarias, con una variabilidad considerable que justifica el uso de modelos avanzados como la Binomial Negativa en lugar del Poisson. La evidencia demuestra que:

- Los días de baja demanda (<5 llamadas) representan casi la mitad, donde basta la dotación mínima.
- Los días típicos (6 a 10 llamadas) concentran una cuarta parte de los casos, reflejando la carga más habitual.
- Los días de alta demanda (>15 llamadas) alcanzan el 18% de la muestra, una frecuencia suficientemente alta como para ser considerada en la planificación operativa.

Estos hallazgos ofrecen una base cuantitativa sólida para el dimensionamiento de recursos móviles, ya que permiten anticipar no solo el promedio esperado, sino también los escenarios extremos que requieren estrategias de refuerzo y flexibilidad en la asignación de recursos.

8.5 Simulación de Demanda

Para complementar el análisis probabilístico de la demanda, se implementó una simulación Monte Carlo sobre la distribución Binomial Negativa ajustada. El objetivo fue generar escenarios representativos de la variabilidad diaria en el número de llamadas, y a partir de ellos establecer criterios de dimensionamiento de la capacidad operativa.

Se utilizaron los parámetros de la Binomial Negativa diaria ($n = 940$ días, media = 8,43 llamadas/día, varianza = 41,23), que previamente había mostrado un ajuste significativamente mejor que el modelo de Poisson (AIC = 5.993 vs. 8.628). Con estos parámetros se generaron 10.000 réplicas de “días futuros” de llamadas mediante la función “`rnbinom()`” en RStudio.

8.5.1 Resultados de la simulación

Los escenarios obtenidos fueron consistentes con los cálculos analíticos, confirmando la validez del modelo:

i. Probabilidades de ocurrencia

- ≤ 5 llamadas: 0,47
- 6–10 llamadas: 0,235 a 0,236
- 11–14 llamadas: 0,104
- ≥ 15 llamadas: 0,182

ii. Percentiles simulados

- P50 = 6 llamadas/día, en la mitad de los días la demanda no supera 6 llamadas.
- P90 = 19 llamadas/día, en el 90% de los días la demanda no supera 19 llamadas.
- P95 = 25 llamadas/día, en el 95% de los días la demanda no supera 25 llamadas.

Estos percentiles permiten traducir la incertidumbre estadística en escenarios operativos concretos.

Como se aprecia en la Tabla 12, si se desea cubrir el 90% de los días, la capacidad mínima requerida sería de 19 llamadas/día, mientras que una política más conservadora (95%) exigiría capacidad para 25 llamadas/día.

Tabla 12*Capacidad simulada por nivel de servicio*

Nivel de Servicio	Llamadas / Día
80 %	14
85 %	16
90 %	19
95 %	25

Nota. Elaboración propia a partir de datos simulados en RStudio.

8.5.2 Conexión Entre la Simulación y la Planificación

i. Lo que mostraron los modelos econométricos (Logit/Probit)

- Que cuando hay más llamadas concentradas en un mismo periodo, aumenta la probabilidad de incumplir los tiempos de respuesta.
- Es decir, no solo importa cuántas llamadas llegan en promedio, sino también los picos de demanda.

ii. Lo que mostró la estimación de la demanda (Binomial Negativa)

- Que la media diaria es de 8 a 9 llamadas, pero existen días con hasta 34 llamadas.
- La distribución tiene sobre dispersión, lo que significa que hay más variabilidad de lo que una distribución Poisson clásico predice.

iii. Lo que aporta la simulación Monte Carlo

Al simular 10.000 “días futuros”, se generan escenarios realistas de días típicos y días extremos. Los percentiles resumen esa variabilidad:

- P50 = 6 llamadas, escenario típico.

- P90 = 19 llamadas, escenario exigente.
- P95 = 25 llamadas, escenario extremo.

En base a los resultados y para esta investigación se decidió continuar para la simulación de capacidad de respuesta operativa con el escenario de 19 llamadas diarias (percentil 90). Esta elección responde a varios criterios:

iv. Equilibrio entre exigencia y realismo

- El percentil 90 refleja un escenario exigente pero no extremo: permite la mayoría de los días observados, garantizando un nivel de servicio robusto en la mayoría de las jornadas.
- Al mismo tiempo, evita sobredimensionar la capacidad para casos muy poco frecuentes (percentil 95 o superior), que implicarían costos operativos que pueden ser desproporcionados.

v. Consistencia con la evidencia de la investigación

- El análisis econométrico previo mostró que el riesgo de incumplimiento aumenta en jornadas con mayor concentración de llamadas.
- Tomar como referencia 19 llamadas se alinea con este hallazgo, ya que precisamente esos escenarios de alta demanda son los que más comprometen el cumplimiento de los tiempos de respuesta.

vi. Criterio operativo y de planificación

- Dimensionar con base en 19 llamadas asegura que se pueda atender la gran mayoría de los días sin considerar los días con exigencias excepcionales.

- Esto permite diseñar un sistema eficiente en condiciones normales y, al mismo tiempo, prever mecanismos de refuerzo específicos para los días extremos, en lugar de sobredimensionar la operación diaria.

8.6 Simulación y Cálculo de capacidad

El análisis de capacidad constituye la etapa final de los resultados, donde se integran los hallazgos obtenidos en los apartados anteriores. A partir de la demanda proyectada mediante la distribución Binomial Negativa y la simulación Monte Carlo, y considerando los resultados de los modelos Logit y Probit respecto a la ocurrencia y cumplimiento de tiempos de respuesta, resulta una simulación operativa orientada a determinar la cantidad de móviles necesaria para distintos escenarios de carga.

8.6.1 Demanda Proyectada por Grupos de Comunas

A partir del escenario de referencia definido en la sección anterior ($P90 = 19$ llamadas diarias), se realizó la distribución de la demanda según la proporción histórica de ocurrencia en los distintos grupos de comunas. Esta aproximación integra dos resultados previos:

- Por un lado, el pronóstico de demanda obtenido a través del ajuste de la distribución Binomial Negativa y su simulación mediante Monte Carlo, que permitió estimar escenarios representativos de carga diaria.
- Por otro, la evidencia de los modelos Logit y Probit, los cuales mostraron que la probabilidad de ocurrencia y de incumplimiento de tiempos de respuesta se concentra en determinadas comunas, principalmente en el Grupo 1 (Viña del Mar, Valparaíso y comunas adyacentes). Con base en estas dos fuentes de información, la proyección de llamadas por grupo de comunas para un día típico en el percentil 90 es la siguiente:

- **Grupo 1:** 14 llamadas/día (72%).
- **Grupo 2:** 2 llamadas/día.
- **Grupo 3:** 0 a 1 llamadas/día.
- **Grupo 4:** 1 llamada/día.
- **Grupo 5:** 1 llamada/día.

En términos prácticos, esta distribución implica que el Grupo 1 concentra la mayor parte de la carga proyectada. Esto coincide tanto con los resultados econométricos como con el análisis histórico, y refuerza la idea de que es en este grupo donde se deben enfocar los esfuerzos principales de dimensionamiento de la capacidad operativa.

8.6.2 Simulación de Capacidad Operativa

La simulación de escenarios se realizó mediante un modelo de colas M/M/c considerando con los tiempos de respuesta observados bajo la dotación actual de tres móviles en el Grupo 1. Este enfoque permitió separar el tiempo de traslado del tiempo de espera en cola, el cual depende de la simultaneidad de llamadas y, por lo tanto, disminuye al incrementar el número de móviles disponibles. En la Tabla 13 siguiente se presentan los resultados simulados para las principales focalizaciones bajo la demanda proyectada de 14 llamadas diarias en el Grupo 1 (P90):

Tabla 13

Capacidad simulada por focalización, cantidad de móviles y tiempos de respuesta

Focalización	Móviles	Tiempo total (min)	Estándar (min)	Exceso/Holgura (min)	Resultado
Interior	3	90.0	60	+30.0	Incumple
	4	49.9	60	-10.1	Cumple
	5	42.1	60	-17.9	Cumple
	6	40.4	60	-19.6	Cumple

Exterior/Med.	3	91.9	30	+61.9	Incumple
	4	50.2	30	+20.2	Incumple
	5	42.2	30	+12.2	Incumple
	6	40.5	30	+10.5	Incumple
Sin suministro	3	108.0	180	-72.0	Cumple
	4	53.4	180	-126.6	Cumple
	5	43.0	180	-137.0	Cumple
	6	40.6	180	-139.4	Cumple

Nota. Elaboración propia a partir de datos simulados en RStudio.

Los resultados de la simulación muestran con claridad que el desempeño del sistema depende fuertemente del número de móviles disponibles en el Grupo 1.

- **Emergencias interiores:** con tres móviles el tiempo promedio de atención se ubica en 90 minutos, superando en 30 minutos el estándar de 60 minutos. Al incorporar un cuarto móvil, el promedio cae a 49.9 minutos, situándose dentro del umbral exigido. Con cinco y seis móviles, la reducción es aún mayor (42.1 y 40.4 minutos), garantizando holgura mayor.
- **Emergencias de medidor y exterior:** constituyen el principal cuello de botella. Con tres móviles, los tiempos promedian 91.9 minutos, más del triple del límite de 30 minutos. Aunque la dotación de cuatro, cinco y seis móviles reduce los promedios a 50.2, 42.2 y 40.5 minutos respectivamente, aún permanecen por encima del estándar, lo que evidencia la dificultad estructural de alcanzar dicho nivel de exigencia.
- **Emergencias sin suministro:** cumplen de manera consistente con el estándar de 180 minutos incluso en la situación base de tres móviles, y la incorporación de móviles adicionales solo mejora aún más la holgura en este tipo de casos.

En conjunto, los resultados indican que el incremento de tres a cuatro móviles genera el mayor impacto, al permitir cumplir el estándar en emergencias interiores y elevar de manera importante el nivel global de cumplimiento. Sin embargo, las emergencias de medidor y exterior continúan mostrando incumplimiento, incluso en escenarios con cinco y seis móviles, lo que sugiere que el problema no se resuelve únicamente con mayor capacidad, sino que requiere estrategias adicionales de gestión operativa.

Dado que el cumplimiento histórico global del Grupo 1 es 34%, se estima que en un día típico con 14 llamadas (escenario P90) unas 5 cumplirían y 9 no cumplirían los tiempos de respuesta. Distribuyendo este resultado según la proporción de llamadas por focalización (42% Interior, 40% Exterior/Medidor, 18% Sin suministro), se proyecta de la siguiente forma:

- Llamadas/día (P90): 14
- Mezcla por focalización: Interior = 6; Exterior/Medidor = 6; Sin suministro = 2.

La Tabla 14 muestra que, el porcentaje global de casos que cumplen con los tiempos de respuesta pasa de 36% con tres móviles a 57% con cuatro y cinco móviles. El estancamiento en 57% se explica porque las llamadas de Exterior/Medidor, que representan cerca del 40% de la demanda del Grupo 1, manteniendo un tiempo promedio superior al estándar de 30 minutos, incluso con cinco y seis móviles.

Sin embargo, de la Tabla 15 es importante destacar que en la focalización los tiempos proyectados mejoran de 91.9 minutos con tres móviles a 40.5 minutos con seis móviles, lo que significa una reducción de la brecha frente al estándar de un 67% con cuatro móviles, un 80% con cinco móviles y un 83% con seis móviles. En términos operativos, esto implica que, aunque no se alcance el umbral de cumplimiento, sí existe una mejora sustantiva en los tiempos de espera, la cual no queda reflejada en el indicador de cumplimiento global.

Tabla 14*Cantidad de móviles y porcentaje de cumplimiento para escenarios simulados*

Escenario	Cumplen	No cumplen	% Global
3 móviles	5	9	36%
4 móviles	8	6	57%
5 móviles	8	6	57%
6 móviles	8	6	57%

Nota. Elaboración propia a partir de datos simulados en RStudio.

Tabla 15*Variación de tiempos de respuesta en escenarios simulados*

Focalización	3 móviles	4 móviles	5 móviles	6 móviles
Interior	90.0 min	49.9 min	42.1 min	40.4 min
Exterior/Medidor	91.9 min	50.2 min	42.2 min	40.5 min
Sin suministro	108.0 min	53.4 min	43.0 min	40.6 min

Nota. Los valores corresponden a tiempos promedio de respuesta proyectados mediante simulación en el Grupo 1, considerando la distribución de 14 llamadas diarias según focalización.

8.6.2.1 Simulación con Variación de Tiempos de Respuesta. Con el fin de evaluar la robustez de los resultados obtenidos en la simulación, se desarrolló un ejercicio de sensibilidad focalizado en la categoría Exterior/Medidor. Esta focalización fue seleccionada por constituir el principal cuello de botella del sistema, ya que concentra un volumen relevante de llamadas y, al mismo tiempo, presenta los mayores niveles de incumplimiento respecto a los tiempos.

En este contexto, se exploró la hipótesis de modificar el umbral de cumplimiento desde 30 a 40 minutos y de 30 a 45 minutos, con el objetivo de observar cómo variaría el desempeño operativo y el impacto potencial sobre los indicadores globales.

8.6.2.1.1 Simulación con Variación de Tiempo de Respuesta a 40 minutos. Aun elevando el umbral de 30 a 40 min, Exterior/Medidor no alcanza a cumplir en ninguno de los escenarios; con 5 móviles queda a solo 2.2 min del nuevo estándar.

En Interior (60 min) y Sin Suministro (180 min) no se modifica el criterio, por lo que sus conclusiones se mantienen: Interior cumple con 4–5 móviles y Sin Suministro cumple siempre.

Como las 6 llamadas diarias de Exterior/Medidor del Grupo 1 siguen clasificando como “no cumple” con el umbral de 40 min, los porcentajes globales no cambian respecto a la tabla principal:

- 3 móviles: 5/14, 36%
- 4 móviles: 8/14, 57%
- 5 móviles: 8/14, 57%
- 6 móviles: 8/14, 57%.

Aunque el porcentaje global no varía con el umbral de 40 minutos, la brecha promedio en Exterior/Medidor se reduce sustantivamente (de +61.9 a +20.2, +12.2 y +10.5 minutos con 4, 5 y 6 móviles), lo que indica una mejora operativa real que el porcentaje global no captura. Si se adopta un umbral alternativo de 40 minutos, la reducción de la brecha frente al estándar alcanza aproximadamente un 80% con cuatro móviles, un 96% con cinco móviles y un 98% con seis móviles. Esto evidencia que, aunque el indicador global se mantenga en 0%, la focalización Exterior/Medidor presenta mejoras operativas sustantivas, reflejadas en una disminución significativa de los tiempos de respuesta promedio.

8.6.2.1.2 Simulación con Variación de Tiempo de Respuesta a 45 minutos. Al considerar un umbral alternativo de 45 minutos para la focalización Exterior/Medidor, los resultados muestran un cambio sustantivo en el cumplimiento global. Al incorporar un cuarto móvil, si bien Exterior/Medidor aún se ubica levemente por sobre el umbral (50.2 min), la brecha respecto al estándar se reduce de 46.9 minutos sobre el estándar a solo 5.2 minutos sobre el estándar, lo que equivale a una mejora relativa cercana al 89% en los tiempos promedio.

Finalmente, al dimensionar con cinco móviles, el tiempo promedio de Exterior/Medidor desciende a 42.2 minutos, con lo cual no solo se elimina la brecha, sino que se alcanza el cumplimiento pleno. En este escenario, el cumplimiento global asciende a 100%, lo que implica que todas las focalizaciones logran responder dentro de los tiempos establecidos. Este análisis de sensibilidad confirma que el principal cuello de botella identificado en Exterior/Medidor puede mitigarse significativamente mediante el aumento de capacidad: con cuatro móviles se obtiene una mejora operativa cercana al estándar, y con cinco móviles se logra un cumplimiento total de los tiempos definidos bajo el umbral de 45 minutos.

Tabla 16

Cantidad de móviles y porcentaje de cumplimientos con tiempos de respuesta con 45 minutos para focalización Exterior y Medidor

Escenario	Cumplen	No cumplen	% Global
3 móviles	5	9	36%
4 móviles	8	6	57%
5 móviles	14	0	100%
6 móviles	14	0	100%

Nota. Los resultados corresponden a un análisis de sensibilidad en el que se ajusta el umbral de cumplimiento para la focalización Exterior/Medidor desde 30 a 45 minutos.

8.6.3 Capacidad Operativa Para Grupo de Comunas con Baja Demanda

En los grupos de comunas distintos al Grupo 1, los resultados deben interpretarse de manera diferenciada. Dado que las tasas de llamadas son significativamente más bajas 1 y 2 llamadas diarias, como se mostró en la simulación de distribución de la demanda y en el análisis de simultaneidad de llamadas, los elevados tiempos de respuesta observados no reflejan un problema de capacidad ni de colas, sino más bien están asociados a factores operativos, principalmente los tiempos de traslado y la ausencia de móviles dedicados en dichas zonas.

Por esta razón, se vuelve necesario actualizar los parámetros de tiempo de respuesta en función de la realidad territorial (traslados) y considerar la asignación de al menos un móvil que, además de atender las emergencias como su principal tarea, pueda complementar sus labores diarias con otras actividades operativas que permitan flexibilizar su utilización, priorizando siempre la atención de llamados de emergencia. Esta estrategia permite optimizar recursos y, al mismo tiempo, garantizar la atención oportuna en comunas con menor volumen de llamados, manteniendo la prioridad sobre los casos críticos.

8.7 Análisis Costo-Beneficio

Con el fin de complementar el análisis de capacidad operativa, se realizó una evaluación costo-beneficio de la incorporación de móviles adicionales en el Grupo 1 de acuerdo con la simulación de capacidad realizada. Este ejercicio permite contrastar los resultados obtenidos en la simulación de tiempos de respuesta con el costo económico asociado a la dotación de recursos. La motivación surge de que, si bien el aumento de móviles mejora los niveles de cumplimiento y reduce sustantivamente los tiempos de espera, estas mejoras deben ser analizadas en relación con la inversión requerida.

La Tabla 17 presenta el costo anual estimado de incorporar un móvil adicional en el escenario base, desglosado en los principales ítems asociados a su operación. En el caso de la remuneración y otros por técnico calificado, se incluyen no solo el sueldo base, sino también componentes como cotizaciones sociales, gratificaciones, movilización, colación, etc. El ítem de arriendo de vehículo corresponde al contrato necesario para disponer de un móvil equipado, mientras que telefonía refleja los gastos de comunicación asociados al servicio.

Por su parte, los gastos de operación agrupan consumos menores como papelería, fotocopias y otros insumos, y los gastos de mantención consideran reparaciones y servicios preventivos requeridos para asegurar la disponibilidad del móvil.

Es importante destacar que estos valores corresponden a datos reales de una empresa de distribución de gas, los cuales fueron utilizados con fines académicos. Por motivos de confidencialidad, las cifras se presentan de manera resumida y sin exponer información sensible de la organización.

Tabla 17

Costos anuales estimados por incorporación de un móvil para atención de emergencias

Resumen	Real anual
Remuneración y otros por técnico calificado	\$ 18.972.086
Arriendo de vehículo	\$ 2.825.000
Telefonía	\$ 1.473.116
Gastos de operación	\$ 1.520.000
Gastos de mantención	\$ 15.172.922
Total	\$ 39.963.124

Nota. Los valores corresponden a datos reales de una empresa de distribución de gas.

8.7.1 Evaluación de Aumento de Móviles

La Tabla 18 muestra la variación en el porcentaje de cumplimiento global al aumentar progresivamente la dotación de móviles en el Grupo 1. Con el estándar de 30 minutos para los casos de exterior y medidor, el escenario de tres móviles alcanza solo un 36% de cumplimiento, mientras que al pasar a cuatro móviles se observa un aumento significativo hasta un 57%, equivalente a +21 puntos porcentuales. Esto implica un costo incremental de aproximadamente \$40 millones anuales, lo que se traduce en un costo de \$1,90 millones por cada punto porcentual adicional de cumplimiento.

Los escenarios de cinco y seis móviles no generan un aumento en el porcentaje de cumplimiento. Sin embargo, estos escenarios sí contribuyen a la reducción de los tiempos promedio de respuesta en la focalización Exterior/Medidor, lo que representa una mejora operativa que no queda reflejada en la métrica binaria.

Tabla 18

Análisis costo-beneficio al aumentar móviles con estándar base de tiempos de respuesta

Escenario	% Global	Variación puntos vs. Anterior	Costo incremental	Costo por punto (%)
3 móviles	36%	—	—	—
4 móviles	57%	+21 %	\$40 MM/año	$\$40/21 = \$1,90\text{MM por } \%$
5 móviles	57%	+0 %	\$80 MM/año	— (no hay ganancia)
6 móviles	57%	+0 %	\$120 MM/año	— (no hay ganancia)

Nota. El costo por punto se calcula como el cociente entre el costo incremental y la variación de cumplimiento global respecto al escenario anterior.

8.7.1.1 Evaluación de Aumento de Móviles y Tiempos de Respuesta. Es importante mencionar que también fue considerado un umbral alternativo de 40 minutos para los casos de Exterior/Medidor, donde no se observan cambios en el porcentaje de cumplimiento global respecto del escenario base, aunque los tiempos promedio se reducen de manera importante.

Los resultados del análisis según la Tabla 19 considerando un estándar de 45 minutos para la focalización en Exterior/Medidor muestran un cambio sustantivo en la relación costo–beneficio de incorporar móviles adicionales. Con tres móviles, el cumplimiento global alcanza solo un 36%, mientras que al aumentar a cuatro móviles se eleva a 57%. Este incremento implica un costo de \$40 millones anuales, con una eficiencia de \$1,90 millones por punto de cumplimiento.

Lo más relevante ocurre al pasar de cuatro a cinco móviles, ya que en este escenario se alcanza el 100% de cumplimiento global. El costo asociado es de \$80 millones anuales, lo que se traduce en aproximadamente \$0,93 millones por punto, evidenciando una relación costo–efectividad mucho más favorable que en el escenario anterior. El aumento a seis móviles no genera un incremento en el cumplimiento, En consecuencia, el sexto móvil representa un costo adicional sin beneficios tangibles en términos de porcentaje de cumplimiento, lo que indica que el óptimo en este análisis se encuentra en cinco móviles.

Tabla 19

Análisis costo-beneficio con variación a 45 minutos para focalización en Exterior y Medidor

Escenario	% Global	Variación puntos vs. anterior	Costo incremental	Costo por punto (%)
3 móviles	36%	—	—	—
4 móviles	57%	+21 %	\$40 MM/año	\$1,90MM/%
5 móviles	100%	+43 %	\$80 MM/año	\$40/43 = \$0,93MM/%
6 móviles	100%	+0 %	\$120 MM/año	—

Nota. La tabla presenta los resultados del análisis de sensibilidad con 45 minutos.

8.7.1.2 Beneficio de Mejora en Brechas de Tiempos de Respuesta. El análisis de cumplimiento binario (cumple/no cumple) no refleja todas las mejoras obtenidas con el aumento de móviles, especialmente en la focalización Exterior/Medidor, donde pequeñas reducciones en los tiempos promedio implican mejoras sustantivas en la operación. Por ello, se evaluó la reducción de la brecha frente al estándar, tanto en valores absolutos (minutos) como en términos relativos (porcentaje de mejora).

i. Estándar 30 minutos (escenario base):

- 3 móviles: tiempo total 91,9 min. Con una brecha de +61,9 min.
- 4 móviles: 50,2 min. Con una brecha de +20,2 min, lo que representa una reducción de -67% respecto a 3 móviles.
- 5 móviles: 42,2 min. Con una brecha de +12,2 min, reducción acumulada de -80%. Sin embargo, el crecimiento marginal es bajo: el salto de 3→4 móviles entrega 67% de reducción, mientras que de 4 a 5 móviles solo se logran un 13% adicional.
- 6 móviles: 40,5 min. Con una brecha de +10,5 min, reducción acumulada de -83%. El aporte marginal es aún menor, con apenas 3% adicionales frente al escenario de cinco móviles.

ii. Estándar 40 minutos (sensibilidad):

- 3 móviles: brecha de +51,9 min.
- 4 móviles: +10,2 min. Con un -80% de reducción.
- 5 móviles: +2,2 min. Con un -96% de reducción y un crecimiento marginal de 16% respecto al salto anterior.
- 6 móviles: +0,5 min. Con un -99% de reducción, lo que implica una ganancia marginal mínima (3% adicional frente al escenario de cinco móviles).

iii. Estándar 45 minutos (sensibilidad):

- 3 móviles: brecha de +46,9 min.
- 4 móviles: +5,2 min. Con una reducción de -89%.
- 5 móviles: -2,8 min (es decir, por debajo del estándar). Implica una reducción acumulada de -106%. El salto de 4 a 5 móviles entrega la mejora más significativa, con un crecimiento marginal cercano a 17% adicionales.
- 6 móviles: -4,5 min. Con una reducción acumulada de -110%, lo que implica un aporte marginal muy reducido (aproximadamente 4% adicionales).

iv. Minutos ahorrados/día (6 llamadas diarias en Ext/Med):

- 3 a 4: $(91,9 - 50,2) \times 6 = 250$ min/día.
- 4 a 5: $(50,2 - 42,2) \times 6 = 48$ min/días adicionales (298 min acumulados).
- 5 a 6: $(42,2 - 40,5) \times 6 = 10$ min/días adicionales (308 min acumulados).

El beneficio incremental es muy alto al pasar de 3 a 4 móviles, pues se logra una reducción de entre 67% y 89% de la brecha, dependiendo del estándar considerado. El paso de 4 a 5 móviles también aporta mejoras relevantes (13 – 17% adicionales), mientras que el salto a 6 móviles ofrece un beneficio marginal muy bajo (3 – 4% adicionales). Esto refuerza la idea de que el análisis binario subestima los avances intermedios y que la eficiencia marginal decrece a medida que se incorporan más recursos.

9. Conclusiones y Recomendaciones

9.1 Conclusiones

El presente trabajo tuvo como propósito analizar la dotación óptima de móviles de emergencia en una empresa de distribución de gas, considerando tanto criterios de cumplimiento como eficiencia económica y operativa. A través de la combinación de análisis de datos, métodos econométricos, simulación estocástica y análisis de costo–beneficio, fue posible obtener una visión integral sobre el desempeño del sistema de atención frente a distintos escenarios de demanda y estándares de tiempo de respuesta.

En primer lugar, el análisis exploratorio de datos, desarrollado como primera etapa de la investigación, permitió comprender en detalle el comportamiento histórico de las llamadas de emergencia. A través de medidas descriptivas y visualizaciones se identificaron patrones de ocurrencia, distribución por focalización, tiempos de respuesta y niveles de cumplimiento. Estos hallazgos iniciales entregaron una visión clara de la situación de base, evidenciando las principales brechas operativas y orientando las etapas posteriores de modelación y simulación.

En segundo lugar, el análisis econométrico permitió caracterizar los factores asociados a la ocurrencia y asignación de emergencias, evidenciando la relevancia de la focalización territorial y el tipo de incidente (interior, exterior/medidor y sin suministro) en los tiempos de atención. Estos resultados constituyeron la base para la simulación de escenarios de demanda futura mediante Monte Carlo, que permitió estimar la distribución probable de llamadas diarias y proyectar el uso esperado de recursos móviles.

En tercer lugar, los resultados de la simulación de capacidad mostraron diferencias significativas según la focalización. En el caso de emergencias en interior, se observó que a partir de cuatro móviles se logra cumplir consistentemente con el estándar normativo de 60 minutos.

Para las emergencias de sin suministro, el cumplimiento se alcanza incluso con tres móviles, presentando holgura considerable a medida que se incrementa la dotación. Sin embargo, en exterior/medidor se identificó el principal cuello de botella: aun con seis móviles, el tiempo promedio de respuesta supera ligeramente el umbral de 30 minutos. No obstante, al analizar la reducción de la brecha respecto al estándar, se constató una mejora progresiva y significativa: la brecha se reduce en un 67% al pasar de tres a cuatro móviles, en un 80% con cinco móviles y hasta un 83% con seis móviles. Con umbrales alternativos (40 y 45 minutos), el cumplimiento aumenta de manera considerable, alcanzando el 100% global en el escenario de cinco móviles con estándar de 45 minutos.

En cuarto lugar, el análisis económico evidenció que el salto de tres a cuatro móviles constituye la decisión más costo–efectiva, con un costo aproximado de \$1,90 millones anuales por punto porcentual de mejora en cumplimiento. El quinto y sexto móvil, en cambio, no generan un aumento en el porcentaje general de cumplimiento bajo el estándar de 30 minutos, aunque sí aportan mejoras operativas al reducir la brecha de tiempos de respuesta. El beneficio incremental del sexto móvil resultó marginal (10 minutos/días adicionales), confirmando rendimientos decrecientes en la ampliación de la flota.

Finalmente, se incorporó un análisis de sensibilidad y riesgo respecto al aumento de los estándares de respuesta (de 30 a 40 y 45 minutos). Si bien esta opción reduce la presión sobre la operación y permite alcanzar mayores niveles de cumplimiento general, implica un aumento en el riesgo regulatorio, reputacional y de seguridad.

Por ello, resulta fundamental evaluar de manera conjunta el costo de incorporar móviles adicionales y el costo esperado de eventuales incidentes o incumplimientos.

En síntesis, este estudio aporta un marco metodológico replicable que combina econometría, simulación y análisis económico, ofreciendo a la organización una herramienta concreta para balancear eficiencia operativa, cumplimiento y costos en la gestión de emergencias.

9.2 Recomendaciones

Con base en los resultados obtenidos, se formulan las siguientes recomendaciones:

- **Actualización periódica de los tiempos de traslado.** Se recomienda revisar y actualizar de forma periódica los tiempos de desplazamiento considerados en la planificación operativa.
- **Dotación operativa base.** Mantener una dotación de cuatro móviles como escenario estándar para el Grupo 1, dado que representa la mejor relación costo–beneficio bajo el estándar de 40 minutos para la focalización exterior/medidor siendo esta la más exigente.
- **Ajuste operacional para los otros grupos de comunas.** Para los demás grupos de comunas, se recomienda implementar un esquema de ajuste operacional flexible, en el cual cada móvil tenga como función prioritaria la atención de emergencias, pero pueda también desempeñar tareas operativas complementarias en períodos de menor demanda. Este enfoque permitiría optimizar el uso de recursos, mantener la disponibilidad necesaria para responder a emergencias y mejorar los niveles de cumplimiento,
- **Uso táctico de un quinto móvil.** Considerar la incorporación de un quinto móvil en períodos de alta demanda, anticipándose mediante los patrones históricos de llamadas y las probabilidades de simultaneidad estimadas, o en caso de que la organización adopte un estándar de 45 minutos para focalización exterior/medidor.

-
- **No se recomienda la incorporación de un sexto móvil de manera permanente.** Si bien reduce marginalmente los tiempos de espera, el beneficio incremental no justifica el costo adicional. El sexto móvil debería reservarse como capacidad de contingencia frente a eventos extraordinarios de demanda.
 - **Monitoreo continuo.** Implementar un sistema de seguimiento trimestral o semestral que permita evaluar la simultaneidad de llamadas, la variación en la demanda y el cumplimiento de tiempos de respuesta. Esto permitiría ajustar dinámicamente la dotación de móviles y anticipar necesidades futuras.
 - **Gestión del riesgo.** Incorporar explícitamente en la planificación un análisis de riesgo asociado a la flexibilización de estándares de respuesta. Este análisis debe considerar tanto posibles multas regulatorias como impactos en la satisfacción de clientes y en la seguridad de las instalaciones.
 - **Indicadores complementarios.** No limitar la evaluación únicamente al cumplimiento global. Se recomienda incluir indicadores de brecha promedio y minutos ahorrados por día como métricas oficiales de gestión, ya que reflejan mejor las mejoras operativas y la eficiencia marginal de las decisiones.
 - **Líneas futuras de investigación.** Ampliar el horizonte temporal de la base de datos para robustecer los modelos econométricos; explorar metodologías de localización óptima de móviles que consideren la geografía urbana y tiempos de desplazamiento, así evaluar con mayor precisión la variabilidad en la atención de emergencias.

10. Discusión

Los resultados obtenidos en este estudio confirman lo planteado en la literatura sobre gestión de operaciones y teoría de colas: el aumento de la capacidad de respuesta mejora los niveles de servicio, pero con rendimientos decrecientes a medida que se incorporan más recursos.

En concordancia con estudios previos, se identificó que el principal cuello de botella se concentra en las emergencias de exterior/medidor, lo que coincide con la evidencia internacional sobre focos críticos en la distribución de gas. Sin embargo, un aporte distintivo de esta investigación fue la incorporación de métricas de reducción de brecha de tiempos de respuesta, que permiten visibilizar mejoras operativas no reflejadas en el cumplimiento porcentual general.

Asimismo, la aplicación conjunta de modelos econométricos, simulación y análisis de costo–beneficio constituye un enfoque integral que extiende lo discutido en el marco teórico y lo adapta al contexto particular del tipo de proceso estudiado

11. Limitaciones

Entre las limitaciones del estudio se encuentra la disponibilidad de datos históricos, restringida a un número acotado de años, lo que impide proyectar tendencias de más largo plazo.

Asimismo, los modelos de simulación se basaron en supuestos simplificados sobre tiempos promedio de atención y distribución de la demanda, lo que, si bien resulta metodológicamente válido, no captura toda la variabilidad de la operación real. El análisis económico se sustentó en datos de costos actuales, sin contemplar escenarios de variación futura por inflación, reajustes laborales o cambios en tarifas de servicios asociados. Finalmente, el trabajo se centró en un grupo de comunas y no incorporó la dimensión espacial de la localización de móviles, aspecto que podría enriquecer investigaciones posteriores.

Además, debe señalarse que los resultados de las simulaciones están condicionados por el contexto en que se desarrollaron. Si bien se utilizaron parámetros derivados de datos y hechos históricos de la empresa, estos representan una aproximación promedio que no contempla todos los escenarios posibles. En la práctica, los tiempos de respuesta y la demanda de emergencias pueden variar significativamente frente a eventos puntuales de carácter regional, nacional o internacional, tales como desastres naturales, crisis energéticas o cambios regulatorios. En consecuencia, los resultados de este estudio deben interpretarse como un marco de referencia robusto, pero no como predicciones deterministas, lo que refuerza la importancia de actualizar periódicamente los modelos con nueva información empírica.

12. Referencias

- Asociación Gas Natural de Chile. (2022, julio). *Diferencias entre gas licuado de petróleo y gas natural*.
<https://www.agnchile.cl/2022/07/diferencias-entre-gas-licuado-de-petroleo-y-gas-natural/>
- Asociación Gas Natural de Chile. (s.f.). *Gas natural en Chile*. Consultado el 6 de mayo de 2025.
<https://www.agnchile.cl/gas-natural/>
- Aprende con Energía. (s.f.). *Origen y características del gas natural*. Consultado el 6 de mayo de 2025.
<https://www.aprendeconenergia.cl/fuentes-de-energia/origen-y-caracteristicas-del-gas-natural>
- Burnham, K. P., & Anderson, D. R. (2004). *Model selection and multimodel inference: A practical information-theoretic approach*. Springer.
- Cameron, A. C., & Trivedi, P. K. (2013). *Regression Analysis of Count Data* (2nd ed.). Cambridge University Press.
- Comisión Nacional de Energía. (2023). *Anuario sector energético*.
<https://www.cne.cl/nuestros-servicios/reportes/informacion-y-estadisticas/>
- El Orden Mundial. (2022). *Los principales productores de gas del mundo*.
<https://elordenmundial.com/mapas-y-graficos/principales-productores-gas-mundo/>
- Energía y Sociedad. (s.f.). *Reservas, extracción y producción del gas natural*. Consultado el 6 de mayo de 2025.

<https://www.energiaysociedad.es/manual-de-la-energia/3-2-reservas-extraccion-y-produccion/>

Escalona, P., & Ramírez, D. (2012). *Expansión de capacidad para un proceso, múltiples ítems y mezcla de tecnologías*.

https://41jaiio.sadio.org.ar/sites/default/files/1_SII_2012.pdf

Espinosa, F. (2018). *Modelo de gestión de emergencias mayores en la industria del gas y petróleo en Ecuador* (Magister en gestión ambiental). Universidad Internacional SEK. Quito, Ecuador.

Gasco Educa. (s.f.). *Gas natural*. Consultado 5 de mayo de 2025.

http://www.gascoeduca.cl/Maqueta/gas_04.html

Gasco Educa. (s.f.). *Historia del gas*. Consultado 5 de mayo de 2025.

http://www.gascoeduca.cl/Maqueta/historia_06.html#p0

Gobierno de Chile. (2020). *Contribución determinada a nivel nacional (NDC) actualizada*.

Ministerio del Medio Ambiente.

<https://www4.unfccc.int/sites/ndcstaging/PublishedDocuments/Chile%20First/Contribución%20Nacional%20Chile%202020.pdf>

Gross, D., & Harris, C. M. (1998). *Fundamentals of Queueing Theory* (3rd ed.). Wiley.

Gujarati, D. N. (2010). *Econometría* (5ª ed.). McGraw-Hill.

<https://fvela.wordpress.com/wp-content/uploads/2012/10/econometria-damodar-n-gujarati-5ta-ed.pdf>

Haddow, G. D., Bullock, J. A., & Coppola, D. P. (2016). *Introduction to emergency management* (5th ed.). Butterworth-Heinemann.

[https://books.google.cl/books?hl=es&lr=&id=xW4IBgAAQBAJ&oi=fnd&pg=PP1&dq=Haddow,+G.,+Bullock,+J.,+%26+Coppola,+D.+\(2016\).+Introduction+to+Emergency+Management.+Butterworth-Heinemann.&ots=dJ_UuR-K0f&sig=akWwrdJGdQfQpo11aTAn-8TeBNc#v=onepage&q&f=false](https://books.google.cl/books?hl=es&lr=&id=xW4IBgAAQBAJ&oi=fnd&pg=PP1&dq=Haddow,+G.,+Bullock,+J.,+%26+Coppola,+D.+(2016).+Introduction+to+Emergency+Management.+Butterworth-Heinemann.&ots=dJ_UuR-K0f&sig=akWwrdJGdQfQpo11aTAn-8TeBNc#v=onepage&q&f=false)

Hilbe, J. M. (2011). *Negative Binomial Regression* (2nd ed.). Cambridge University Press.

International Energy Agency. (2023a). *Natural gas*.

<https://www.iea.org/energy-system/fossil-fuels/natural-gas#tracking>

International Energy Agency. (2023b). *World Energy Outlook 2023*.

<https://www.iea.org/reports/world-energy-outlook-2023>

Krajewski, L. J., & Ritzman, L. P. (s.f.). *Administración de operaciones*. Pearson Educación.

https://www.gob.mx/cms/uploads/attachment/file/566458/Administracion_De_Operaciones_-_LEE_J._K-comprimido.pdf

Kroese, D. P., Brereton, T., Taimre, T., & Botev, Z. I. (2014). *Why the Monte Carlo method is so important today*. *Wiley Interdisciplinary Reviews: Computational Statistics*, 6(6), 386–392.

Lara, D. (2022). *Implementación de un modelo de optimización para aumentar el nivel de cumplimiento en los tiempos de atención de emergencias de la empresa de gas natural de Lima y Callao S.A.* (Trabajo de suficiencia profesional para optar el título profesional Ingeniero Industrial). Universidad Privada del Norte. Lima, Perú.

-
- Lawless, J. F. (1987). Negative binomial and mixed Poisson regression. *The Canadian Journal of Statistics*, 15(3), 209–225.
- Limpert, E., Stahel, W. A., & Abbt, M. (2001). *Log-normal distributions across the sciences: Keys and clues*. *BioScience*, 51(5), 341–352.
- Ministerio de Economía, Fomento y Reconstrucción. (2000). *Decreto Supremo Número 280: Reglamento de instalaciones interiores de gas*.
<https://www.bcn.cl>
- Ministerio de Economía, Fomento y Reconstrucción. (2007). *Decreto Supremo Número 66: Reglamento de seguridad de instalaciones de consumo de gas*.
<https://www.bcn.cl>
- Ministerio de Energía. (2019). *Informe resumen: Ruta energética hacia la carbono neutralidad 2050*. Gobierno de Chile.
https://energia.gob.cl/sites/default/files/paginabasica/informe_resumen_cn_2019_v07.pdf
- Muñiz, M. (2019). *Diseño de una metodología para colocar puntos de referencia para la disminución de los tiempos de respuesta en la atención de incidentes reportados a una central de emergencias en la ciudad de Guayaquil* (Magister en estadísticas con mención en gestión de la calidad y productividad). Escuela Superior Politécnica del Litoral. Guayaquil, Ecuador.
- Prieto, R. y Simonin, I. (2014). Muchas emergencias y aún más llamadas. *Miscelánea Matemática*, 59, 41–50.
- Ross, S. M. (2014). *Introduction to Probability Models* (11th ed.). Academic Press.
-

Superintendencia de Electricidad y Combustibles. (26 de mayo de 2014). *Derechos de los clientes de servicio eléctrico y de combustibles*.

https://sec.custhelp.com/app/answers/detail/a_id/440/kw/atencion%20emergencias

Swissinfo.ch. (2018). *Fuerte explosión de gas deja 3 muertos y unos 50 heridos en clínica del sur de Chile*.

<https://www.swissinfo.ch/spa/fuerte-explosi%C3%B3n-de-gas-deja-3-muertos-y-unos-50-heridos-en-cl%C3%ADnica-del-sur-de-chile/44065016>.

Toregas, C., Swain, R., ReVelle, C., & Bergman, L. (1971). *The location of emergency service facilities*. *Operations Research*, 19(6), 1363–1373.

http://yalma.fime.uanl.mx/~roger/work/teaching/class_tso/docs/1971-or-Toregas%20etal-The%20Location%20of%20Emergency%20Service%20Facilities.pdf

Wooldridge, J. M. (2009). *Introductory econometrics: A modern approach* (4th ed.). South-Western Cengage Learning.

<https://herioscarlanda.wordpress.com/wp-content/uploads/2018/10/wooldridge-2009-introduccion-a-la-econometria-un-enfoque-moderno.pdf>

Yergin, D. (2009). *The prize: The epic quest for oil, money & power*. Free Press

<https://www.iwp.edu/wp-content/uploads/2020/06/The-Prize-The-Epic-Quest-for-Oil-Money-and-Power-by-Daniel-Yergin.pdf>

13. Apéndices

Apéndice A: Códigos en RStudio Para el Análisis Exploratorio de Datos

```
library(readxl)
datos <- read_excel(file.choose())

# Instalar paquetes si es necesario
install.packages(c("readxl", "dplyr", "ggplot2", "lubridate"))

# Cargar librerías
library(readxl)
library(dplyr)
library(ggplot2)
library(lubridate)

colnames(datos)

# 2. Extraer hora desde "Hora de Recep."
datos <- datos %>%
  mutate(
    hora_recep = hour(hms(format(`Hora de Recep.`, "%H:%M:%S"))),
    tiempo_min = as.numeric(hms(format(`Tiempos de respuesta 2`, "%H:%M:%S")) / 60),
    comuna = as.factor(`Comuna...9`),
    tipo_emergencia = as.factor(`Tipo Emergencia...12`),
    focalizacion = as.factor(Focalización)
  )

# 3. Distribución de llamadas por hora
ggplot(datos, aes(x = hora_recep)) +
  geom_histogram(binwidth = 1, fill = "#377eb8", color = "white") +
  labs(title = "Distribución de llamadas por hora",
       x = "Hora del día", y = "Número de llamadas") +
  theme_minimal()

# 4. Llamadas por comuna
datos %>%
  count(Comuna...9) %>%
  arrange(desc(n)) %>%
  ggplot(aes(x = reorder(Comuna...9, n), y = n)) +
  geom_bar(stat = "identity", fill = "#4daf4a") +
  coord_flip() +
  labs(title = "Número de llamadas por comuna", x = "Comuna", y = "Cantidad") +
  theme_minimal()

# 5. Llamadas por tipo de emergencia
datos %>%
  count(`Tipo Emergencia...12`) %>%
  arrange(desc(n)) %>%
  ggplot(aes(x = reorder(`Tipo Emergencia...12`, n), y = n)) +
  geom_bar(stat = "identity", fill = "#984ea3") +
  coord_flip() +
  labs(title = "Distribución por tipo de emergencia", x = "Tipo", y = "Cantidad") +
  theme_minimal()

library(lubridate)

datos <- datos %>%
  mutate(
    tiempo_min = suppressWarnings(as.numeric(hms(format(`Tiempos de respuesta 2`,
"%H:%M:%S")) / 60)
  )

datos <- datos %>% filter(!is.na(tiempo_min), tiempo_min >= 0, tiempo_min <= 240)

library(ggplot2)

ggplot(datos, aes(x = "", y = tiempo_min)) +
  geom_boxplot(fill = "#ff7f00") +
  labs(title = "Tiempos de respuesta (minutos)", x = "", y = "Minutos") +
  theme_minimal()

ggplot(datos, aes(x = tipo_emergencia, y = tiempo_min)) +
  geom_boxplot(fill = "#FF9933") +
  coord_flip() +
  labs(title = "Tiempos de respuesta por tipo de emergencia", x = "Tipo de emergencia",
       y = "Minutos") +
  theme_minimal()
```

```
library(readxl)
datos <- read_excel(file.choose())

library(dplyr)
library(lubridate)

col <- "Cumple tiempos de respuesta"

# No cumplen tiempos de respuesta
resumen_ceros <- datos %>%
  transmute(valor = suppressWarnings(as.numeric(.data[[col]]))) %>%
  summarise(
    total_no_na = sum(!is.na(valor)),
    conteo_ceros = sum(valor == 0, na.rm = TRUE),
    porcentaje_ceros = round(100 * conteo_ceros / total_no_na, 2)
  )

resumen_ceros

library(ggplot2)

# Datos para gráfico de torta
df_torta <- data.frame(
  categoria = c("No cumplen", "Cumplen"),
  cantidad = c(5235, 7936 - 5235)
)

# Calcular porcentajes
df_torta$porcentaje <- round(df_torta$cantidad / sum(df_torta$cantidad) * 100, 1)

# Crear gráfico
ggplot(df_torta, aes(x = "", y = cantidad, fill = categoria)) +
  geom_col(width = 1) +
  coord_polar(theta = "y") +
  geom_text(aes(label = paste0(porcentaje, "%")),
            position = position_stack(vjust = 0.5), size = 5, color = "white") +
  labs(title = "Cumplimiento de tiempos de respuesta",
       fill = "Categoría") +
  theme_void() +
  theme(
    plot.title = element_text(hjust = 0.5, face = "bold"),
    legend.position = "bottom"
  )

# Desglose de no cumple tiempos de respuesta por tipo de emergencia
col_tipo <- "Tipo Emergencia...12"

desglose_tipo <- datos %>%
  filter(`Cumple tiempos de respuesta` == 0) %>%
  group_by(`Tipo Emergencia...12`) %>%
  summarise(
    total_no_cumplen = n(),
    .groups = "drop"
  ) %>%
  mutate(
    porcentaje = round(100 * total_no_cumplen / sum(total_no_cumplen), 2)
  )

desglose_tipo

# Grafico segun desglose
ggplot(desglose_tipo, aes(x = reorder(`Tipo Emergencia...12`, porcentaje),
                          y = porcentaje, fill = `Tipo Emergencia...12`)) +
  geom_col() +
  geom_text(aes(label = paste0(porcentaje, "%")),
            hjust = -0.1, size = 3) +
  coord_flip() +
```

```
labs(
  title = "Porcentaje de llamadas que NO cumplen el tiempo de respuesta por tipo de
emergencia",
  x = "Tipo de Emergencia",
  y = "Porcentaje (%)"
) +
theme_minimal() +
theme(
  plot.title = element_text(hjust = 0.5, face = "bold"),
  legend.position = "none"
)

# Rangos de no cumplimiento de tiempos de respuesta
library(dplyr)
library(lubridate)
library(ggplot2)

# Ajusta los nombres
col_cumple <- "Cumple tiempos de respuesta" # 0 = NO cumple
col_focal <- "Focalización"
col_tipo <- "Tipo Emergencia...12" # o ...11
col_time <- "Tiempos de respuesta 2" # o "Tiempo de respuesta"

# 1) Asegurar tiempo_respuesta en minutos
if (!"tiempo_respuesta" %in% names(datos)) {
  datos <- datos %>%
    mutate(tiempo_respuesta = as.numeric(hms(format(.data[[col_time]], "%H:%M:%S")))) /
    60)
}

# 2) Filtrar NO cumplen y calcular exceso en minutos (incluye 180 min para Sin
Suministro)
no_cumplen <- datos %>%
  filter(.data[[col_cumple]] == 0) %>%
  mutate(
    # normaliza focalización para evitar problemas de espacios/mayúsculas
    focal_norm = tolower(trimws(.data[[col_focal]])),
    tiempo_max = case_when(
      focal_norm %in% c("exterior", "medidor") ~ 30,
      focal_norm == "interior" ~ 60,
      focal_norm == "sin suministro" ~ 180,
      TRUE ~ NA_real_
    ),
    exceso_min = pmax(tiempo_respuesta - tiempo_max, 0)
  )

# 3) Definir RANGOS de exceso en minutos
cortes <- c(0, 10, 20, 30, 60, Inf)
etiquetas <- c("0-10 min", "10-20", "20-30", "30-60", ">60")

no_cumplen <- no_cumplen %>%
  mutate(rango_min = cut(exceso_min, breaks = cortes, labels = etiquetas, right =
FALSE))

# 4) Resumen (conteo y %)
rangos_global <- no_cumplen %>%
  filter(!is.na(rango_min)) %>%
  count(rango_min, name = "n") %>%
  mutate(porcentaje = round(100 * n / sum(n), 2)) %>%
  arrange(rango_min)

print(rangos_global)

# (Opcional) gráfico rápido
ggplot(rangos_global, aes(x = rango_min, y = porcentaje)) +
  geom_col() +
  geom_text(aes(label = paste0(porcentaje, "%"), vjust = -0.2, size = 3)) +
  labs(title = "Exceso de tiempo (min) entre NO cumplimientos",
  x = "Rango de exceso (min)", y = "Porcentaje (%)") +
```

```
theme_minimal()

library(dplyr)
library(ggplot2)
library(tidyr)

# Cruce por tipo de emergencia y rango
niveles_rangos <- c("0-10 min", "10-20", "20-30", "30-60", ">60")

# Cruce por tipo de emergencia y rango
rangos_por_tipo <- no_cumplen %>%
  mutate(
    rango_min = factor(rango_min, levels = niveles_rangos),
    tipo_emergencia = .data[[col_tipo]]
  ) %>%
  count(tipo_emergencia, rango_min, name = "n") %>%
  complete(tipo_emergencia, rango_min, fill = list(n = 0)) %>%
  group_by(tipo_emergencia) %>%
  mutate(
    total_tipo = sum(n),
    porcentaje = ifelse(total_tipo > 0, round(100 * n / total_tipo, 2), 0)
  ) %>%
  ungroup()

# Ver tabla
print(rangos_por_tipo, n = 100)

# Grafico
ggplot(rangos_por_tipo, aes(x = reorder(tipo_emergencia, -total_tipo),
                             y = porcentaje, fill = rango_min)) +
  geom_col() +
  coord_flip() +
  labs(
    title = "Rangos de exceso entre NO cumplimientos por tipo de emergencia",
    x = "Tipo de emergencia",
    y = "Porcentaje dentro de cada tipo",
    fill = "Exceso (min)"
  ) +
  theme_minimal() +
  theme(axis.text.y = element_text(size = 9))

# Tasa de no cumplimiento por comuna

col_cumple <- "Cumple tiempos de respuesta" # 0 = NO cumple, 1 = cumple
col_comuna <- "Comuna...9"

install.packages("stringi")
library(stringi)

# 1) Normalizar nombres de comuna
datos <- datos %>%
  mutate(comuna_norm = toupper(stri_trans_general(as.character(.data[[col_comuna]]),
"Latin-ASCII")))

# 2) Calcular tasa de NO cumplimiento por comuna
tasa_no_cumple_comuna <- datos %>%
  filter(!is.na(.data[[col_cumple]])) %>%
  group_by(comuna_norm) %>%
  summarise(
    total_llamadas = n(),
    no_cumplen = sum(.data[[col_cumple]] == 0),
    tasa_no_cumple = round(100 * no_cumplen / total_llamadas, 2)
  ) %>%
  arrange(desc(tasa_no_cumple), desc(no_cumplen))

print(tasa_no_cumple_comuna)
```

```
# Grafico
library(ggplot2)

ggplot(tasa_no_cumple_comuna,
       aes(x = reorder(comuna_norm, tasa_no_cumple), y = tasa_no_cumple)) +
  geom_col() +
  coord_flip() +
  geom_text(aes(label = paste0(tasa_no_cumple, "%")), hjust = -0.1, size = 3) +
  labs(title = "Tasa de no cumplimiento por comuna",
       x = "Comuna",
       y = "Tasa de no cumplimiento (%)") +
  theme_minimal() +
  theme(axis.text.y = element_text(size = 9),
        plot.margin = margin(5.5, 35, 5.5, 5.5))

# Si cumplen tiempos de respuesta por tipo de emergencia

library(dplyr)

col_cumple <- "Cumple tiempos de respuesta"
col_tipo   <- "Tipo Emergencia...12"

# Filtrar las que SI cumplen y agrupar por tipo de emergencia
desglose_cumplen <- datos %>%
  filter(.data[[col_cumple]] == 1) %>%
  group_by(tipo_emergencia = .data[[col_tipo]]) %>%
  summarise(
    total_cumplen = n(),
    .groups = "drop"
  ) %>%
  mutate(
    porcentaje = round(100 * total_cumplen / sum(total_cumplen), 2)
  ) %>%
  arrange(desc(total_cumplen))

print(desglose_cumplen)

# Gráfico
ggplot(desglose_cumplen, aes(x = reorder(tipo_emergencia, total_cumplen),
                              y = total_cumplen)) +
  geom_col(fill = "#27AE60") +
  geom_text(aes(label = paste0(porcentaje, "%")),
            hjust = -0.1, size = 3) +
  coord_flip() +
  labs(title = "Llamadas que Sí cumplen con los tiempos de respuesta",
       x = "Tipo de emergencia",
       y = "Cantidad de llamadas") +
  theme_minimal()
```

```
library(readxl)
datos <- read_excel(file.choose())

library(dplyr)
library(stringi)
library(lubridate)
library(ggplot2)

# 0) Asegurar columna 'comuna' (si no la creaste antes)
if (!"comuna" %in% names(datos)) {
  # Ajusta el nombre si tu columna válida es `Comuna...10`
  datos <- datos %>% mutate(comuna = as.character(`Comuna...9`))
} else {
  datos <- datos %>% mutate(comuna = as.character(comuna))
}

# 1) Normalizar nombres de comuna para el match (mayúsculas y sin acentos)
datos <- datos %>%
  mutate(
    comuna_norm = toupper(stri_trans_general(comuna, "Latin-ASCII"))
  )

# 2) Definir tus grupos (también normalizados para evitar problemas de acentos)
grupos <- list(
  "Grupo 1" = c("VINA DEL MAR", "VALPARAISO", "QUILPUE", "VILLA
ALEMANA", "CASABLANCA", "CONCON", "CON CON", "CURAUMA"),
  "Grupo 2" = c("QUILLOTA", "CALERA", "LA CRUZ"),
  "Grupo 3" = c("LOS ANDES", "LLAY LLAY"),
  "Grupo 4" = c("LA SERENA", "COQUIMBO"),
  "Grupo 5" = c("TALCA")
)

# 3) Asignar grupo a cada registro según comuna
datos <- datos %>%
  mutate(
    grupo = case_when(
      comuna_norm %in% grupos[["Grupo 1"]] ~ "Grupo 1",
      comuna_norm %in% grupos[["Grupo 2"]] ~ "Grupo 2",
      comuna_norm %in% grupos[["Grupo 3"]] ~ "Grupo 3",
      comuna_norm %in% grupos[["Grupo 4"]] ~ "Grupo 4",
      comuna_norm %in% grupos[["Grupo 5"]] ~ "Grupo 5",
      TRUE ~ "Otros"
    )
  )

# 4) Usar tu columna REAL de cumplimiento
# Aseguramos que sea 0/1 numérico (por si viene como texto/factor)
datos <- datos %>%
  mutate(
    cumple_real = suppressWarnings(as.numeric(`Cumple tiempos de respuesta`))
  )

# 5) Resumen por GRUPO (llamadas y % cumplimiento)
tabla_grupos <- datos %>%
  group_by(grupo) %>%
  summarise(
    llamadas = n(),
    cumplimiento_pct = mean(cumple_real, na.rm = TRUE) * 100
  ) %>%
  arrange(desc(llamadas))

print(tabla_grupos)

# 6) Resumen por GRUPO y COMUNA para ver el detalle interno
tabla_grupo_comuna <- datos %>%
  group_by(grupo, comuna) %>%
  summarise(
```

```
    llamadas = n(),
    cumplimiento_pct = mean(cumple_real, na.rm = TRUE) * 100,
    .groups = "drop"
  ) %>%
  arrange(grupo, desc(llamadas))

# Mira las 15 comunas con más llamadas (para foco)
print(head(tabla_grupo_comuna, 20))

# 7) Promedio de llamadas por día en cada grupo (si tienes columna Fecha)
if ("Fecha" %in% names(datos)) {
  llamadas_por_dia_grupo <- datos %>%
    mutate(fecha = as.Date(Fecha)) %>%
    group_by(grupo, fecha) %>% summarise(llamadas = n(), .groups = "drop") %>%
    group_by(grupo) %>% summarise(llamadas_dia_prom = mean(llamadas), .groups = "drop")
  %>%
    arrange(desc(llamadas_dia_prom))
  print(llamadas_por_dia_grupo)
}

# 8) Gráfico para la tesis
ggplot(tabla_grupos, aes(x = reorder(grupo, -cumplimiento_pct), y = cumplimiento_pct))
+
  geom_col() +
  geom_text(aes(label = sprintf("%.1f%%", cumplimiento_pct)), vjust = -0.4, size = 3.8)
+
  labs(title = "Cumplimiento de tiempos de respuesta por grupo (columna real)",
       x = "Grupo de comunas", y = "% de cumplimiento") +
  theme_minimal()
```

```
# Probabilidad que existan emergencias simultaneas
library(readxl)
datos <- read_excel(file.choose())

library(dplyr)
library(lubridate)
library(stringi)

# 1) Timestamp a partir de Fecha + Hora de Recep.
col_fecha <- "Fecha"
col_hora <- "Hora de Recep."
col_comuna <- "Comuna...9"

datos_ts <- datos %>%
  mutate(
    # normaliza comuna (opcional, pero ayuda a agrupar)
    comuna = toupper(stri_trans_general(as.character(.data[[col_comuna]]), "Latin-
ASCII")),
    # arma timestamp; intenta parsear la hora robustamente
    hora_txt = format(.data[[col_hora]], "%H:%M:%S"),
    ts = ymd_hms(paste(as.Date(.data[[col_fecha]]), hora_txt), tz = "America/Santiago")
  ) %>%
  filter(!is.na(ts)) %>%
  arrange(ts)

# Vector de ventanas en minutos
ventanas <- seq(10, 120, by = 10)

# 2) Función que calcula prob. de simultaneidad para una tabla ordenada por ts
prob_simult <- function(ts_vec, windows_min) {
  # Convierte a segundos y pre-calcula índices de "hasta dónde llega la ventana"
  tnum <- as.numeric(ts_vec) # segundos UNIX
  n <- length(tnum)
  base_idx <- seq_len(n)
  # contadores por ventana: para cada evento i, cuántos siguientes caen en [ti, ti + w]
  res <- lapply(windows_min, function(wm) {
    wsec <- wm * 60
    # findInterval devuelve, para cada límite (ti + w), el índice del último <= a ese
    límite
    last_in_window <- findInterval(tnum + wsec, tnum)
    # número de eventos futuros dentro de la ventana
    cnt_future <- pmax(last_in_window - base_idx, 0L)
    # simultáneo si hay al menos 1 en la ventana
    simult_flag <- as.integer(cnt_future > 0)
    tibble(
      ventana_min = wm,
      n_eventos = n,
      n_simultaneos = sum(simult_flag),
      prob_simult = n_simultaneos / n_eventos
    )
  })
  bind_rows(res)
}

# 3) Probabilidad GLOBAL
resultado_global <- prob_simult(datos_ts$ts, ventanas)
print(resultado_global)
# columnas: ventana_min, n_eventos, n_simultaneos, prob_simult

# 4) Probabilidad por comuna
resultado_por_comuna <- datos_ts %>%
  group_by(comuna) %>%
  group_modify(~ prob_simult(.x$ts, ventanas)) %>%
  ungroup()

# Tabla resumida por ventana (promedio simple de comunas)
resumen_comunal <- resultado_por_comuna %>%
  group_by(ventana_min) %>%
  summarise(
```

```
    comunas = n_distinct(comuna),
    prob_media = mean(prob_simult, na.rm = TRUE),
    prob_mediana = median(prob_simult, na.rm = TRUE),
    .groups = "drop"
  )

print(resumen_comunal)

# Global en porcentaje
tabla_global <- resultado_global %>%
mutate(porcentaje = round(100 * prob_simult, 2)) %>%
select(ventana_min, n_eventos, n_simultaneos, porcentaje)

print(tabla_global)

# Top comunas con mayor prob. de simultaneidad a 10, 30 y 60 min (ejemplo)
topc <- resultado_por_comuna %>%
  filter(ventana_min %in% c(10, 20, 30, 40, 50, 60)) %>%
  group_by(ventana_min) %>%
  arrange(desc(prob_simult), .by_group = TRUE) %>%
  slice_head(n = 10) %>%
  ungroup()

print(topc)

print(resultado_por_comuna, n = 228)
```

Apéndice B: Códigos en RStudio Para el Análisis Econométrico (Modelos Logit y Probit)

```
library(readxl)
datos <- read_excel(file.choose())

install.packages(c("dplyr", "lubridate", "pscl")) # Ejecuta solo una vez
library(dplyr)
library(lubridate)
library(pscl)

colnames(datos)

library(dplyr)
library(lubridate)
install.packages("stringi")

datos <- datos %>%
  mutate(
    hora_recep = hour(hms(format(`Hora de Recep.`, "%H:%M:%S"))),
    # Puedes aplicar lo mismo para otras variables si las necesitas más adelante:
    hora_checkin = hour(hms(format(`Hora de Check in`, "%H:%M:%S"))),
    tiempo_respuesta = as.numeric(hms(format(`Tiempo de respuesta`, "%H:%M:%S"))),

    comuna = as.factor(`Comuna...9`),
    tipo_emergencia = as.factor(`Tipo Emergencia...12`),
    focalizacion = as.factor(Focalización),
    emergencia = as.numeric(`Emergencia Ocurrida`)
  ) %>%
  filter(!is.na(emergencia), !is.na(hora_recep))

# Modelo LOGIT
modelo_logit <- glm(emergencia ~ comuna + tipo_emergencia + focalizacion + hora_recep,
  data = datos, family = binomial("logit"))
```

```
table(datos$comuna)
table(datos$tipo_emergencia)
table(datos$focalizacion)

# 1. Convertir a mayúsculas y quitar tildes
library(stringi)

datos <- datos %>%
  mutate(
    comuna = stri_trans_general(as.character(comuna), "Latin-ASCII"), # quita tildes
    comuna = toupper(comuna) # pone todo en mayúsculas
  )

table(datos$comuna)

comunas_validas <- names(table(datos$comuna)[table(datos$comuna) >= 10])
datos <- datos %>% filter(comuna %in% comunas_validas)

modelo_logit <- glm(emergencia ~ comuna + tipo_emergencia + focalizacion + hora_recep,
  data = datos, family = binomial("logit"))

summary(modelo_logit)

modelo_probit <- glm(emergencia ~ comuna + tipo_emergencia + focalizacion + hora_recep,
  data = datos, family = binomial("probit"))

summary(modelo_probit)

datos_modelo <- datos %>%
  mutate(
    comuna = stri_trans_general(as.character(Comuna...9), "Latin-ASCII"),
    comuna = toupper(Comuna...9)
  ) %>%
  filter(Comuna...9 %in% comunas_validas)

# Modelo con esta versión
modelo_logit <- glm(emergencia ~ comuna + tipo_emergencia + focalizacion + hora_recep,
  data = datos_modelo, family = binomial("logit"))

# Predicciones con los mismos datos
predicciones_prob <- predict(modelo_logit, newdata = datos_modelo, type = "response")
predicciones_bin <- ifelse(predicciones_prob >= 0.5, 1, 0)

# Revisa longitud
length(predicciones_bin) == length(datos_modelo$emergencia) # Esto debe dar TRUE

# Matriz de confusión
matriz_confusion <- table(Predicho = predicciones_bin, Real = datos_modelo$emergencia)
print(matriz_confusion)

# Metricas

VP <- 5085
FP <- 548
FN <- 1
VN <- 2299

# Precisión global (Accuracy)
accuracy <- (VP + VN) / (VP + FP + FN + VN)

# Sensibilidad (Recall o Tasa de verdaderos positivos)
sensibilidad <- VP / (VP + FN)

# Especificidad (Tasa de verdaderos negativos)
especificidad <- VN / (VN + FP)

# Precisión positiva (Precision)
precision <- VP / (VP + FP)

# Mostrar resultados
cat("Precisión global (Accuracy):", round(accuracy, 4), "\n")
cat("Sensibilidad (Recall):", round(sensibilidad, 4), "\n")
cat("Especificidad:", round(especificidad, 4), "\n")
cat("Precisión positiva (Precision):", round(precision, 4), "\n")
```

```
# Modelo Probit

datos_modelo <- datos %>%
  mutate(
    comuna = stri_trans_general(as.character(comuna), "Latin-ASCII"),
    comuna = toupper(comuna)
  ) %>%
  filter(comuna %in% comunas_validas)

# Modelo con esta versión
modelo_probit <- glm(emergencia ~ comuna + tipo_emergencia + focalizacion + hora_recep,
  data = datos_modelo, family = binomial("probit"))

# Predicciones con los mismos datos
predicciones_prob <- predict(modelo_probit, newdata = datos_modelo, type = "response")
predicciones_bin <- ifelse(predicciones_prob >= 0.5, 1, 0)

# Revisa longitud
length(predicciones_bin) == length(datos_modelo$emergencia) # Esto debe dar TRUE

# Matriz de confusión
matriz_confusion <- table(Predicho = predicciones_bin, Real = datos_modelo$emergencia)
print(matriz_confusion)

# Métricas
TP <- 5083 # Verdaderos positivos
TN <- 2298 # Verdaderos negativos
FP <- 549  # Falsos positivos

FN <- 3    # Falsos negativos

# Calcular métricas
accuracy <- (TP + TN) / (TP + TN + FP + FN)
sensibilidad <- TP / (TP + FN)
especificidad <- TN / (TN + FP)
precision <- TP / (TP + FP)

# Mostrar resultados
cat("Precisión global (Accuracy):", round(accuracy, 4), "\n")
cat("Sensibilidad (Recall):", round(sensibilidad, 4), "\n")
cat("Especificidad:", round(especificidad, 4), "\n")
cat("Precisión positiva (Precision):", round(precision, 4), "\n")
```

```
# install.packages(c("readxl", "dplyr", "lubridate", "ggplot2", "stringi"))
library(readxl)
library(dplyr)
library(lubridate)
library(ggplot2)
suppressWarnings(library(stringi)) # si no está, más abajo usamos fallback con iconv

# ==== 1) Carga de datos ====
# datos <- read_excel(file.choose()) # ya lo tienes cargado arriba

# ==== 2) Nombres de columnas (usa exactamente como aparecen en tu Excel) ====
col_hora_recep <- "Hora de Recep."
col_hora_check <- "Hora de Check in"
col_time_1 <- "Tiempo de respuesta"
col_time_2 <- "Tiempos de respuesta 2"
col_comuna_raw <- "Comuna...9"
col_tipo_raw <- "Tipo Emergencia...12"
col_focal_raw <- "Focalización"

# Elegir columna de tiempo válida (prefiere la 2 si existe)
col_time <- if ("Tiempos de respuesta 2" %in% names(datos)) col_time_2 else col_time_1

# ==== 3) Transformaciones base (SIN filtrar filas todavía) ====
datos <- datos %>%
  mutate(
    hora_recep = hour(hms(format(.data[[col_hora_recep]], "%H:%M:%S"))),
    hora_checkin = hour(hms(format(.data[[col_hora_check]], "%H:%M:%S"))),

    # tiempo_respuesta en minutos
    tiempo_respuesta = as.numeric(hms(format(.data[[col_time]], "%H:%M:%S"))) / 60,

    # Normalizar comuna (quitar tildes y pasar a MAYÚSCULAS)
    comuna = if ("stringi" %in% .packages()) {
      toupper(stri_trans_general(as.character(.data[[col_comuna_raw]]), "Latin-ASCII"))
    } else {
      toupper(iconv(as.character(.data[[col_comuna_raw]]), from = "UTF-8", to =
"ASCII//TRANSLIT"))
    },

    # Normalizar tipo de emergencia (opcional pero recomendado)
    tipo_emergencia = if ("stringi" %in% .packages()) {
      toupper(stri_trans_general(as.character(.data[[col_tipo_raw]]), "Latin-ASCII"))
    } else {
      toupper(iconv(as.character(.data[[col_tipo_raw]]), from = "UTF-8", to =
"ASCII//TRANSLIT"))
    },

    # Normalizar focalización a 4 etiquetas canónicas
    focalizacion = case_when(
      tolower(trimws(.data[[col_focal_raw]])) %in% "exterior" ~ "Exterior",
      tolower(trimws(.data[[col_focal_raw]])) %in% "medidor" ~ "Medidor",
      tolower(trimws(.data[[col_focal_raw]])) %in% "interior" ~ "Interior",
      grepl("sin\\s*suministro", tolower(.data[[col_focal_raw]])) ~ "Sin Suministro",
      TRUE ~ as.character(.data[[col_focal_raw]])
    ),

    # Variable dependiente: CUMPLE (1) / NO CUMPLE (0)
    cumple_tiempo = case_when(
      focalizacion %in% c("Exterior", "Medidor") & tiempo_respuesta <= 30 ~ 1,
      focalizacion == "Interior" & tiempo_respuesta <= 60 ~ 1,
      focalizacion == "Sin Suministro" & tiempo_respuesta <= 180 ~ 1, # si también
quieres aplicar 180'
      TRUE ~ 0
    )
  )

# Chequeos rápidos
cat("Total filas tras transformaciones:", nrow(datos), "\n")
cat("NA en tiempo_respuesta:", sum(is.na(datos$tiempo_respuesta)), "\n")
```

```
cat("NA en cumple_tiempo:", sum(is.na(datos$cumple_tiempo)), "\n")

# ==== 4) Filtrado mínimo (opcional): comunas con al menos 10 casos ====
comunas_validas <- names(table(datos$comuna)[table(datos$comuna) >= 10])
datos <- datos %>% filter(comuna %in% comunas_validas)

# Confirmar tamaño
cat("Filas tras filtrar comunas >=10 casos:", nrow(datos), "\n")

# ==== 5) Modelos (Logit y Probit) sobre el mismo dataset 'datos' ====
modelo_logit_cumple <- glm(cumple_tiempo ~ comuna + tipo_emergencia + focalizacion +
hora_recep,
                          data = datos, family = binomial("logit"))
modelo_probit_cumple <- glm(cumple_tiempo ~ comuna + tipo_emergencia + focalizacion +
hora_recep,
                           data = datos, family = binomial("probit"))

summary(modelo_logit_cumple)
summary(modelo_probit_cumple)

# 6) Habilidad (matriz de confusión y métricas)
# LOGIT
pred_logit_prob <- predict(modelo_logit_cumple, newdata = datos, type = "response")
pred_logit_bin <- ifelse(pred_logit_prob >= 0.5, 1, 0)

matriz_logit <- table(Predicho = pred_logit_bin, Real = datos$cumple_tiempo)
print(matriz_logit)
cat("Total LOGIT en matriz:", sum(matriz_logit), " | Debe ser:", nrow(datos), "\n")

TP <- as.numeric(matriz_logit["1","1"]); FP <- as.numeric(matriz_logit["1","0"])
FN <- as.numeric(matriz_logit["0","1"]); TN <- as.numeric(matriz_logit["0","0"])

accuracy <- (TP + TN) / (TP + TN + FP + FN)
recall <- TP / (TP + FN) # Sensibilidad (clase 1)
specific <- TN / (TN + FP) # Especificidad (clase 0)
precision <- TP / (TP + FP)

cat("\n[LOGIT] Accuracy:", round(accuracy,4),
    "| Recall:", round(recall,4),
    "| Specificity:", round(specific,4),
    "| Precision:", round(precision,4), "\n")

# PROBIT
pred_probit_prob <- predict(modelo_probit_cumple, newdata = datos, type = "response")
pred_probit_bin <- ifelse(pred_probit_prob >= 0.5, 1, 0)

matriz_probit <- table(Predicho = pred_probit_bin, Real = datos$cumple_tiempo)
print(matriz_probit)
cat("Total PROBIT en matriz:", sum(matriz_probit), " | Debe ser:", nrow(datos), "\n")

TP <- as.numeric(matriz_probit["1","1"]); FP <- as.numeric(matriz_probit["1","0"])
FN <- as.numeric(matriz_probit["0","1"]); TN <- as.numeric(matriz_probit["0","0"])

accuracy <- (TP + TN) / (TP + TN + FP + FN)
recall <- TP / (TP + FN)
specific <- TN / (TN + FP)
precision <- TP / (TP + FP)

cat("\n[PROBIT] Accuracy:", round(accuracy,4),
    "| Recall:", round(recall,4),
    "| Specificity:", round(specific,4),
    "| Precision:", round(precision,4), "\n")
```

```
# install.packages(c("readxl","dplyr","lubridate","ggplot2","stringi"))
library(readxl)
library(dplyr)
library(lubridate)
library(ggplot2)
suppressWarnings(library(stringi)) # si no está, más abajo usamos fallback con iconv

# ==== 1) Carga de datos ====
# datos <- read_excel(file.choose()) # ya lo tienes cargado arriba

# ==== 2) Nombres de columnas (usa exactamente como aparecen en tu Excel) ====
col_hora_recep <- "Hora de Recep."
col_hora_check <- "Hora de Check in"
col_time_1 <- "Tiempo de respuesta"
col_time_2 <- "Tiempos de respuesta 2"
col_comuna_raw <- "Comuna...9"
col_tipo_raw <- "Tipo Emergencia...12"
col_focal_raw <- "Focalización"

# Elegir columna de tiempo válida (prefiere la 2 si existe)
col_time <- if ("Tiempos de respuesta 2" %in% names(datos)) col_time_2 else col_time_1

# ==== 3) Transformaciones base (SIN filtrar filas todavía) ====
datos <- datos %>%
  mutate(
    hora_recep = hour(hms(format(.data[[col_hora_recep]], "%H:%M:%S"))),
    hora_checkin = hour(hms(format(.data[[col_hora_check]], "%H:%M:%S"))),

    # tiempo_respuesta en minutos
    tiempo_respuesta = as.numeric(hms(format(.data[[col_time]], "%H:%M:%S"))) / 60,

    # Normalizar comuna (quitar tildes y pasar a MAYÚSCULAS)
    comuna = if ("stringi" %in% .packages()) {
      toupper(stri_trans_general(as.character(.data[[col_comuna_raw]]), "Latin-ASCII"))
    } else {
      toupper(iconv(as.character(.data[[col_comuna_raw]]), from = "UTF-8", to =
"ASCII//TRANSLIT"))
    },

    # Normalizar tipo de emergencia (opcional pero recomendado)
    tipo_emergencia = if ("stringi" %in% .packages()) {
      toupper(stri_trans_general(as.character(.data[[col_tipo_raw]]), "Latin-ASCII"))
    } else {
      toupper(iconv(as.character(.data[[col_tipo_raw]]), from = "UTF-8", to =
"ASCII//TRANSLIT"))
    },

    # Normalizar focalización a 4 etiquetas canónicas
    focalizacion = case_when(
      tolower(trimws(.data[[col_focal_raw]])) %in% "exterior" ~ "Exterior",
      tolower(trimws(.data[[col_focal_raw]])) %in% "medidor" ~ "Medidor",
      tolower(trimws(.data[[col_focal_raw]])) %in% "interior" ~ "Interior",
      grepl("sin\\s*suministro", tolower(.data[[col_focal_raw]])) ~ "Sin Suministro",
      TRUE ~ as.character(.data[[col_focal_raw]])
    ),

    # Variable dependiente: CUMPLE (1) / NO CUMPLE (0)
    cumple_tiempo = case_when(
      focalizacion %in% c("Exterior", "Medidor") & tiempo_respuesta <= 30 ~ 1,
      focalizacion == "Interior" & tiempo_respuesta <= 60 ~ 1,
      focalizacion == "Sin Suministro" & tiempo_respuesta <= 180 ~ 1, # si también
quieres aplicar 180'
      TRUE ~ 0
    )
  )

# Chequeos rápidos
cat("Total filas tras transformaciones:", nrow(datos), "\n")
cat("NA en tiempo_respuesta:", sum(is.na(datos$tiempo_respuesta)), "\n")
```

```
cat("NA en cumple_tiempo:", sum(is.na(datos$cumple_tiempo)), "\n")

# ==== 4) Filtrado mínimo (opcional): comunas con al menos 10 casos ====
comunas_validas <- names(table(datos$comuna)[table(datos$comuna) >= 10])
datos <- datos %>% filter(comuna %in% comunas_validas)

# Confirmar tamaño
cat("Filas tras filtrar comunas >=10 casos:", nrow(datos), "\n")

# ==== 5) Modelos (Logit y Probit) sobre el mismo dataset 'datos' ====
modelo_logit_cumple <- glm(cumple_tiempo ~ comuna + tipo_emergencia + focalizacion +
hora_recep,
                          data = datos, family = binomial("logit"))
modelo_probit_cumple <- glm(cumple_tiempo ~ comuna + tipo_emergencia + focalizacion +
hora_recep,
                           data = datos, family = binomial("probit"))

summary(modelo_logit_cumple)
summary(modelo_probit_cumple)

# 6) Habilidad (matriz de confusión y métricas)
# LOGIT
pred_logit_prob <- predict(modelo_logit_cumple, newdata = datos, type = "response")
pred_logit_bin <- ifelse(pred_logit_prob >= 0.5, 1, 0)

matriz_logit <- table(Predicho = pred_logit_bin, Real = datos$cumple_tiempo)
print(matriz_logit)
cat("Total LOGIT en matriz:", sum(matriz_logit), " | Debe ser:", nrow(datos), "\n")

TP <- as.numeric(matriz_logit["1","1"]); FP <- as.numeric(matriz_logit["1","0"])
FN <- as.numeric(matriz_logit["0","1"]); TN <- as.numeric(matriz_logit["0","0"])

accuracy <- (TP + TN) / (TP + TN + FP + FN)
recall <- TP / (TP + FN) # Sensibilidad (clase 1)
specific <- TN / (TN + FP) # Especificidad (clase 0)
precision <- TP / (TP + FP)

cat("\n[LOGIT] Accuracy:", round(accuracy,4),
    "| Recall:", round(recall,4),
    "| Specificity:", round(specific,4),
    "| Precision:", round(precision,4), "\n")

# PROBIT
pred_probit_prob <- predict(modelo_probit_cumple, newdata = datos, type = "response")
pred_probit_bin <- ifelse(pred_probit_prob >= 0.5, 1, 0)

matriz_probit <- table(Predicho = pred_probit_bin, Real = datos$cumple_tiempo)
print(matriz_probit)
cat("Total PROBIT en matriz:", sum(matriz_probit), " | Debe ser:", nrow(datos), "\n")

TP <- as.numeric(matriz_probit["1","1"]); FP <- as.numeric(matriz_probit["1","0"])
FN <- as.numeric(matriz_probit["0","1"]); TN <- as.numeric(matriz_probit["0","0"])

accuracy <- (TP + TN) / (TP + TN + FP + FN)
recall <- TP / (TP + FN)
specific <- TN / (TN + FP)
precision <- TP / (TP + FP)

cat("\n[PROBIT] Accuracy:", round(accuracy,4),
    "| Recall:", round(recall,4),
    "| Specificity:", round(specific,4),
    "| Precision:", round(precision,4), "\n")

# Bajar el umbral

# --- MÉTRICAS Y BUSCADOR DE UMBRAL ---
conf_metrics <- function(y_true, y_pred_bin) {
```

```
TP <- sum(y_pred_bin == 1 & y_true == 1)
FP <- sum(y_pred_bin == 1 & y_true == 0)
FN <- sum(y_pred_bin == 0 & y_true == 1)
TN <- sum(y_pred_bin == 0 & y_true == 0)

accuracy <- (TP + TN) / (TP + TN + FP + FN)
recall <- ifelse((TP + FN) > 0, TP / (TP + FN), NA)
spec <- ifelse((TN + FP) > 0, TN / (TN + FP), NA)
precision<- ifelse((TP + FP) > 0, TP / (TP + FP), NA)
f1 <- ifelse(!is.na(precision) & !is.na(recall) & (precision+recall)>0,
            2 * precision * recall / (precision + recall), NA)
list(TP=TP, FP=FP, FN=FN, TN=TN,
     accuracy=accuracy, recall=recall, specificity=spec,
     precision=precision, F1=f1)
}

eval_thresholds <- function(probs, y, grid = seq(0.1, 0.9, by=0.05)) {
  out <- lapply(grid, function(t) {
    pred_bin <- ifelse(probs >= t, 1, 0)
    m <- conf_metrics(y, pred_bin)
    youden <- m$recall - (1 - m$specificity) # TPR - FPR
    c(threshold=t, accuracy=m$accuracy, recall=m$recall,
      specificity=m$specificity, precision=m$precision,
      F1=m$F1, youden=youden)
  })
  as.data.frame(do.call(rbind, out))
}

pretty_print <- function(title, cm, met) {
  cat("\n", title, "\n", sep = "")
  print(cm)
  cat(sprintf("Accuracy: %.4f | Recall: %.4f | Specificity: %.4f | Precision: %.4f |
  F1: %.4f\n",
             met$accuracy, met$recall, met$specificity, met$precision, met$F1))
}

# =====
# USAR SIEMPRE 'datos' COMPLETO
# =====

# 1) Entrena los modelos sobre 'datos'
modelo_logit <- glm(cumple_tiempo ~ comuna + tipo_emergencia + focalizacion +
hora_recep,
                  data = datos, family = binomial("logit"))
modelo_probit <- glm(cumple_tiempo ~ comuna + tipo_emergencia + focalizacion +
hora_recep,
                  data = datos, family = binomial("probit"))

# 2) Predicciones sobre las MISMAS filas usadas por glm (no pases newdata)
p_logit <- predict(modelo_logit, type = "response")
p_probit <- predict(modelo_probit, type = "response")
y_true <- datos$cumple_tiempo

# 3) Buscar mejor umbral por F1 (puedes mirar también 'youden')
grid <- seq(0.1, 0.9, by = 0.05)
tab_logit <- eval_thresholds(p_logit, y_true, grid)
tab_probit <- eval_thresholds(p_probit, y_true, grid)

best_logit <- tab_logit [which.max(tab_logit$F1), ]
best_probit <- tab_probit[which.max(tab_probit$F1), ]

cat("\n[ORIGINAL] Mejor umbral por F1\n")
print(best_logit)
print(best_probit)

# 4) Matrices y métricas con esos umbrales
predL <- ifelse(p_logit >= best_logit$threshold, 1, 0)
predP <- ifelse(p_probit >= best_probit$threshold, 1, 0)
```

```
cmL <- table(Predicho = predL, Real = y_true)
cmP <- table(Predicho = predP, Real = y_true)

metL <- conf_metrics(y_true, predL)
metP <- conf_metrics(y_true, predP)

pretty_print(sprintf("[ORIGINAL] LOGIT (thr=%.2f)", best_logit$threshold), cmL, metL)
pretty_print(sprintf("[ORIGINAL] PROBIT (thr=%.2f)", best_probit$threshold), cmP, metP)

# 5) Chequeos: deben sumar igual a nrow(datos)
cat("\nTotal LOGIT en matriz:", sum(cmL), " | Debe ser:", nrow(datos), "\n")
cat("Total PROBIT en matriz:", sum(cmP), " | Debe ser:", nrow(datos), "\n")

# 6) (Opcional) Mostrar cuántas observaciones usó glm:
cat("\nObs usadas por LOGIT (glm):", modelo_logit$df.null + 1, "\n")
cat("Obs usadas por PROBIT (glm):", modelo_probit$df.null + 1, "\n")

# =====
# ESCENARIO A - Balancear: 1/3 de los "0", todos los "1"
# =====

set.seed(123) # reproducibilidad

# Índices de positivos (1) y negativos (0)
idx1 <- which(datos$cumple_tiempo == 1)
idx0 <- which(datos$cumple_tiempo == 0)

# Tomar solo 1/3 de los ceros
keep0 <- sample(idx0, size = round(length(idx0) * (1/3)), replace = FALSE)

# Nuevo dataset balanceado
datos_balance <- datos[c(idx1, keep0), ]

# Revisar proporciones
cat("\nProporción original:\n")
print(prop.table(table(datos$cumple_tiempo)))

cat("\nProporción balanceada:\n")
print(prop.table(table(datos_balance$cumple_tiempo)))

# =====
# Reentrenar y evaluar modelos LOGIT y PROBIT balanceados
# =====

modelo_logit_bal <- glm(cumple_tiempo ~ comuna + tipo_emergencia + focalizacion +
hora_recep,
                        data = datos_balance, family = binomial("logit"))
modelo_probit_bal <- glm(cumple_tiempo ~ comuna + tipo_emergencia + focalizacion +
hora_recep,
                        data = datos_balance, family = binomial("probit"))

# Predicciones
p_logit_bal <- predict(modelo_logit_bal, type = "response")
p_probit_bal <- predict(modelo_probit_bal, type = "response")
y_true_bal <- datos_balance$cumple_tiempo

# Buscar mejor umbral (por F1)
grid <- seq(0.1, 0.9, by = 0.05)
tab_logit_bal <- eval_thresholds(p_logit_bal, y_true_bal, grid)
tab_probit_bal <- eval_thresholds(p_probit_bal, y_true_bal, grid)

best_logit_bal <- tab_logit_bal [which.max(tab_logit_bal$F1), ]
best_probit_bal <- tab_probit_bal [which.max(tab_probit_bal$F1), ]

cat("\n[ESCENARIO A] Mejor umbral por F1\n")
print(best_logit_bal)
print(best_probit_bal)
```

```
# Matrices y métricas
predL_bal <- ifelse(p_logit_bal >= best_logit_bal$threshold, 1, 0)
predP_bal <- ifelse(p_probit_bal >= best_probit_bal$threshold, 1, 0)

cmL_bal <- table(Predicho = predL_bal, Real = y_true_bal)
cmP_bal <- table(Predicho = predP_bal, Real = y_true_bal)

metL_bal <- conf_metrics(y_true_bal, predL_bal)
metP_bal <- conf_metrics(y_true_bal, predP_bal)

pretty_print(sprintf("[BALANCEADO] LOGIT (thr=%.2f)", best_logit_bal$threshold),
cmL_bal, metL_bal)
pretty_print(sprintf("[BALANCEADO] PROBIT (thr=%.2f)", best_probit_bal$threshold),
cmP_bal, metP_bal)
```

Apéndice C: Resultados de Modelos Logit y Probit de Ocurrencia de Emergencias y Cumplimiento de Tiempos de Respuesta

```
modelo_logit <- glm(emergencia ~ comuna + tipo_emergencia + focalizacion + hora_recep,
+
data = datos, family = binomial("logit"))
Warning message:
glm.fit: fitted probabilities numerically 0 or 1 occurred
> summary(modelo_logit)

Call:
glm(formula = emergencia ~ comuna + tipo_emergencia + focalizacion +
hora_recep, family = binomial("logit"), data = datos)

Coefficients:
(Intercept)
ComunaCASABLANCA
ComunaCONCON
ComunaCOQUIMBO
ComunaCURAUMA
ComunaLA CALERA
ComunaLA CRUZ
ComunaLA SERENA
ComunaLLAY LLAY
ComunaLOS ANDES
ComunaQUILLOTA
ComunaQUILPUE
ComunaTALCA
ComunaVALPARAISO
ComunaVILLA ALEMANA
ComunaVINA DEL MAR
tipo_emergenciaDETECCIÓN DE TRABAJOS DE EMPRESAS EXTERNAS SOBRE REDES GAS
tipo_emergenciaFALLA DE INSTRUMENTOS O EQUIPOS SIN CAUSAR FUGAS
tipo_emergenciaFUGA
tipo_emergenciaGOLPE A INSTALACIONES O REDES DE GAS SIN CAUSAR FUGA
tipo_emergenciaINFLAMACIÓN
tipo_emergenciaOLOR A GAS
tipo_emergenciaOTRO TIPO DE EMERGENCIA
tipo_emergenciaSIN SUMINISTRO
focalizacionInterior
focalizacionMedidor
focalizacionSin Suministro
hora_recep
z value Pr(>|z|)

(Intercept)
ComunaCASABLANCA
ComunaCONCON
ComunaCOQUIMBO
ComunaCURAUMA
ComunaLA CALERA
ComunaLA CRUZ
ComunaLA SERENA
ComunaLLAY LLAY
ComunaLOS ANDES
ComunaQUILLOTA
ComunaQUILPUE
ComunaTALCA
ComunaVALPARAISO
ComunaVILLA ALEMANA
ComunaVINA DEL MAR
tipo_emergenciaDETECCIÓN DE TRABAJOS DE EMPRESAS EXTERNAS SOBRE REDES GAS
tipo_emergenciaFALLA DE INSTRUMENTOS O EQUIPOS SIN CAUSAR FUGAS
tipo_emergenciaFUGA
tipo_emergenciaGOLPE A INSTALACIONES O REDES DE GAS SIN CAUSAR FUGA
tipo_emergenciaINFLAMACIÓN
tipo_emergenciaOLOR A GAS
tipo_emergenciaOTRO TIPO DE EMERGENCIA
tipo_emergenciaSIN SUMINISTRO
focalizacionInterior
focalizacionMedidor
focalizacionSin Suministro
hora_recep
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 10347 on 7926 degrees of freedom
Residual deviance: 1970 on 7899 degrees of freedom
(32 observations deleted due to missingness)
AIC: 2026

Number of Fisher Scoring iterations: 21
```

```
> modelo_probit <- glm(emergencia ~ comuna + tipo_emergencia + focalizacion + hora_recep,
+                       data = datos, family = binomial("probit"))
Warning message:
glm.fit: fitted probabilities numerically 0 or 1 occurred
> summary(modelo_probit)

Call:
glm(formula = emergencia ~ comuna + tipo_emergencia + focalizacion +
     hora_recep, family = binomial("probit"), data = datos)

Coefficients:
                Estimate Std. Error
(Intercept)      1.138e+01  4.346e+02
comunaCASABLANCA -7.467e-01  4.051e-01
comunaCONCON     -8.400e-01  3.126e-01
comunaCOQUIMBO  -9.907e-01  5.269e-01
comunaCURAUMA   -1.105e+00  8.374e-01
comunaLA CALERA -7.196e+00  1.335e+02
comunaLA CRUZ   -3.793e-01  3.852e-01
comunaLA SERENA -8.322e-01  3.378e-01
comunaLLAY LLAY -5.534e-01  4.438e-01
comunaLOS ANDES -5.759e-01  5.035e-01
comunaQUILLOTA -3.143e-01  3.256e-01
comunaQUILPUE  -6.320e-01  3.098e-01
comunaTALCA    -7.546e-01  3.081e-01
comunaVALPARAISO -7.270e-01  2.994e-01
comunaVILLA ALEMANA -9.294e-01  3.175e-01
comunaVINA DEL MAR -8.593e-01  2.947e-01
tipo_emergenciaDETECCIÓN DE TRABAJOS DE EMPRESAS EXTERNAS SOBRE REDES GAS -3.913e+00  1.071e+04
tipo_emergenciaFALLA DE INSTRUMENTOS O EQUIPOS SIN CAUSAR FUGAS 1.564e-01  9.022e+02
tipo_emergenciaFUGA 5.300e+00  4.457e+02
tipo_emergenciaGOLPE A INSTALACIONES O REDES DE GAS SIN CAUSAR FUGA 2.359e-01  1.102e+03
tipo_emergenciaINFLAMACIÓN -2.043e-01  6.172e+03
tipo_emergenciaOLOR A GAS -1.668e+01  4.686e+02
tipo_emergenciaOTRO TIPO DE EMERGENCIA -4.983e+00  4.059e+02
tipo_emergenciaSIN SUMINISTRO -5.500e+00  4.059e+02
focalizacionInterior -3.816e+00  1.554e+02
focalizacionMedidor -4.241e+00  1.554e+02
focalizacionSin Suministro -4.646e+00  1.554e+02
hora_recep -1.242e-02  8.489e-03
z value Pr(>|z|)
(Intercept) 0.026 0.97911
comunaCASABLANCA -1.843 0.06533 .
comunaCONCON -2.687 0.00721 **
comunaCOQUIMBO -1.880 0.06008 .
comunaCURAUMA -1.320 0.18686
comunaLA CALERA -0.054 0.95702
comunaLA CRUZ -0.985 0.32477
comunaLA SERENA -2.464 0.01376 *
comunaLLAY LLAY -1.247 0.21244
comunaLOS ANDES -1.144 0.25270
comunaQUILLOTA -0.965 0.33437
comunaQUILPUE -2.040 0.04132 *
comunaTALCA -2.449 0.01434 *
comunaVALPARAISO -2.428 0.01519 *
comunaVILLA ALEMANA -2.927 0.00342 **
comunaVINA DEL MAR -2.915 0.00355 **
tipo_emergenciaDETECCIÓN DE TRABAJOS DE EMPRESAS EXTERNAS SOBRE REDES GAS 0.000 0.99971
tipo_emergenciaFALLA DE INSTRUMENTOS O EQUIPOS SIN CAUSAR FUGAS 0.000 0.99986
tipo_emergenciaFUGA 0.012 0.99051
tipo_emergenciaGOLPE A INSTALACIONES O REDES DE GAS SIN CAUSAR FUGA 0.000 0.99983
tipo_emergenciaINFLAMACIÓN 0.000 0.99997
tipo_emergenciaOLOR A GAS -0.036 0.97160
tipo_emergenciaOTRO TIPO DE EMERGENCIA -0.012 0.99021
tipo_emergenciaSIN SUMINISTRO -0.014 0.98919
focalizacionInterior -0.025 0.98041
focalizacionMedidor -0.027 0.97823
focalizacionSin Suministro -0.030 0.97615
hora_recep -1.463 0.14356
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 10346.9 on 7926 degrees of freedom
Residual deviance: 1970.2 on 7899 degrees of freedom
(32 observations deleted due to missingness)
AIC: 2026.2

Number of Fisher Scoring iterations: 21
```

```
summary(modelo_logit_cumple)

Call:
glm(formula = cumple_tiempo ~ comuna + tipo_emergencia + focalizacion +
     hora_recep, family = binomial("logit"), data = datos_tesis)

Coefficients:
                Estimate Std. Error
(Intercept)      -2.868e+00  2.677e-01
comunaCASABLANCA  4.556e-01  3.082e-01
comunaCONCON      2.123e-01  2.277e-01
comunaCOQUIMBO    4.312e-02  4.747e-01
comunaCURAUMA     1.501e+00  6.891e-01
comunaLA CRUZ     -1.481e-01  3.472e-01
comunaLA SERENA   3.133e-01  2.593e-01
comunaLLAY LLAY   1.873e-01  3.334e-01
comunaLOS ANDES   3.681e-01  6.114e-01
comunaQUILLOTA    3.842e-01  2.215e-01
comunaQUILPUE     3.591e-01  2.070e-01
comunaTALCA       2.980e-01  2.222e-01
comunaVALPARAISO  4.427e-01  1.994e-01
comunaVILLA ALEMANA 3.077e-01  2.180e-01
comunaVINA DEL MAR 3.285e-01  1.953e-01
tipo_emergenciaDETECCION DE TRABAJOS DE EMPRESAS EXTERNAS SOBRE REDES GAS -1.289e+01  1.455e+03
tipo_emergenciaFALLA DE INSTRUMENTOS O EQUIPOS SIN CAUSAR FUGAS -9.062e-01  4.153e-01
tipo_emergenciaFUGA 1.960e-01  1.041e-01
tipo_emergenciaGOLPE A INSTALACIONES O REDES DE GAS SIN CAUSAR FUGA -1.264e+01  1.438e+02
tipo_emergenciaINFLAMACION 1.259e+00  1.231e+00
tipo_emergenciaOLOR A GAS 1.017e-01  1.044e-01
tipo_emergenciaOTRO TIPO DE EMERGENCIA 7.729e-02  1.596e-01
tipo_emergenciaSIN SUMINISTRO -1.340e+01  8.401e+02
focalizacionInterior 2.087e+00  1.237e-01
focalizacionMedidor -2.725e-01  1.526e-01
focalizacionSin Suministro 1.809e+01  8.401e+02
hora_recep        -1.031e-02  7.603e-03

z value Pr(>|z|)
(Intercept)      -10.716 <2e-16 ***
comunaCASABLANCA  1.478  0.1393
comunaCONCON      0.933  0.3510
comunaCOQUIMBO    0.091  0.9276
comunaCURAUMA     2.178  0.0294 *
comunaLA CRUZ     -0.427  0.6697
comunaLA SERENA   1.208  0.2270
comunaLLAY LLAY   0.562  0.5742
comunaLOS ANDES   0.602  0.5471
comunaQUILLOTA    1.735  0.0828 .
comunaQUILPUE     1.735  0.0827 .
comunaTALCA       1.342  0.1797
comunaVALPARAISO  2.221  0.0264 *
comunaVILLA ALEMANA 1.412  0.1580
comunaVINA DEL MAR 1.682  0.0926 .
tipo_emergenciaDETECCION DE TRABAJOS DE EMPRESAS EXTERNAS SOBRE REDES GAS -0.009  0.9929
tipo_emergenciaFALLA DE INSTRUMENTOS O EQUIPOS SIN CAUSAR FUGAS -2.182  0.0291 *
tipo_emergenciaFUGA 1.882  0.0598 .
tipo_emergenciaGOLPE A INSTALACIONES O REDES DE GAS SIN CAUSAR FUGA -0.088  0.9300
tipo_emergenciaINFLAMACION 1.023  0.3064
tipo_emergenciaOLOR A GAS 0.974  0.3298
tipo_emergenciaOTRO TIPO DE EMERGENCIA 0.484  0.6282
tipo_emergenciaSIN SUMINISTRO -0.016  0.9873
focalizacionInterior 16.871 <2e-16 ***
focalizacionMedidor -1.786  0.0741 .
focalizacionSin Suministro 0.022  0.9828
hora_recep        -1.356  0.1752

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 10196.7 on 7932 degrees of freedom
Residual deviance:  6873.3 on 7906 degrees of freedom
AIC: 6927.3

Number of Fisher Scoring iterations: 14
```

```
> summary(modelo_probit_cumple)

Call:
glm(formula = cumple_tiempo ~ comuna + tipo_emergencia + focalizacion +
     hora_recep, family = binomial("probit"), data = datos_tesis)

Coefficients:
(Intercept)                                Estimate Std. Error
comunaCASABLANCA                          0.273011  0.175743
comunaCONCON                               0.143864  0.129221
comunaCOQUIMBO                             0.031654  0.279495
comunaCURAUMA                              0.881995  0.396211
comunaLA CRUZ                             -0.079079  0.194413
comunaLA SERENA                            0.178735  0.148661
comunaLLAY LLAY                            0.076914  0.188611
comunaLOS ANDES                            0.265331  0.337937
comunaQUILLOTA                             0.230326  0.125863
comunaQUILPUE                              0.215264  0.117265
comunaTALCA                                0.183526  0.126764
comunaVALPARAISO                          0.266451  0.113063
comunaVILLA ALEMANA                       0.182911  0.123456
comunaVINA DEL MAR                        0.194952  0.110745
tipo_emergenciaDETECCION DE TRABAJOS DE EMPRESAS EXTERNAS SOBRE REDES GAS -3.659679 376.753572
tipo_emergenciaFALLA DE INSTRUMENTOS O EQUIPOS SIN CAUSAR FUGAS -0.520560 0.215962
tipo_emergenciaFUGA                       0.125699  0.062835
tipo_emergenciaGOLPE A INSTALACIONES O REDES DE GAS SIN CAUSAR FUGA -3.554652 37.186759
tipo_emergenciaINFLAMACION                0.779008  0.751207
tipo_emergenciaOLOR A GAS                 0.075498  0.062942
tipo_emergenciaOTRO TIPO DE EMERGENCIA    0.041683  0.096071
tipo_emergenciaSIN SUMINISTRO             -3.943692 217.429002
focalizacionInterior                       1.175807  0.062654
focalizacionMedidor                       -0.127886  0.073773
focalizacionSin Suministro                 6.676144 217.428998
hora_recep                                -0.005428  0.004343

z value Pr(>|z|)
(Intercept) -11.179 <2e-16 ***
comunaCASABLANCA 1.553 0.1203
comunaCONCON 1.113 0.2656
comunaCOQUIMBO 0.113 0.9098
comunaCURAUMA 2.226 0.0260 *
comunaLA CRUZ -0.407 0.6842
comunaLA SERENA 1.202 0.2292
comunaLLAY LLAY 0.408 0.6834
comunaLOS ANDES 0.785 0.4324
comunaQUILLOTA 1.830 0.0673 .
comunaQUILPUE 1.836 0.0664 .
comunaTALCA 1.448 0.1477
comunaVALPARAISO 2.357 0.0184 *
comunaVILLA ALEMANA 1.482 0.1384
comunaVINA DEL MAR 1.760 0.0783 .
tipo_emergenciaDETECCION DE TRABAJOS DE EMPRESAS EXTERNAS SOBRE REDES GAS -0.010 0.9922
tipo_emergenciaFALLA DE INSTRUMENTOS O EQUIPOS SIN CAUSAR FUGAS -2.410 0.0159 *
tipo_emergenciaFUGA 2.000 0.0455 *
tipo_emergenciaGOLPE A INSTALACIONES O REDES DE GAS SIN CAUSAR FUGA -0.096 0.9238
tipo_emergenciaINFLAMACION 1.037 0.2997
tipo_emergenciaOLOR A GAS 1.199 0.2303
tipo_emergenciaOTRO TIPO DE EMERGENCIA 0.434 0.6644
tipo_emergenciaSIN SUMINISTRO -0.018 0.9855
focalizacionInterior 18.767 <2e-16 ***
focalizacionMedidor -1.734 0.0830 .
focalizacionSin Suministro 0.031 0.9755
hora_recep -1.250 0.2113

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 10196.7 on 7932 degrees of freedom
Residual deviance: 6869.8 on 7906 degrees of freedom
AIC: 6923.8

Number of Fisher Scoring iterations: 14
```

Apéndice D: Códigos en RStudio Para Estimación de Distribución de Probabilidad

```
# --- Paquetes ---
pkgs <- c("readxl", "dplyr", "tidyr", "lubridate", "MASS")
to_install <- setdiff(pkgs, rownames(installed.packages()))
if (length(to_install)) install.packages(to_install, quiet = TRUE)
invisible(lapply(pkgs, library, character.only = TRUE))

# --- 1) Cargar y construir timestamp (robusto) ---
df <- readxl::read_excel("C:/Users/diego/OneDrive/Desktop/datos_tesis.xlsx")
col_fecha <- "Fecha"; col_hora <- "Hora de Recep."

fecha <- as.Date(df[[col_fecha]])
hraw <- df[[col_hora]]
hora_txt <- tryCatch(format(hraw, "%H:%M:%S"), error = function(e) as.character(hraw))
hora_secs <- suppressWarnings(as.numeric(lubridate::hms(hora_txt)))
if (all(is.na(hora_secs)) && is.numeric(hraw)) hora_secs <- round(hraw * 24 * 3600)

ts <- lubridate::as_datetime(fecha, tz = "America/Santiago") +
lubridate::seconds(hora_secs)

df <- df %>% mutate(ts = ts) %>% filter(!is.na(ts)) %>% arrange(ts)

# --- 2) Conteos DIARIOS (incluyendo ceros entre el primer y último día observado) ---
cont_dia_raw <- df %>%
  mutate(dia = as_date(ts)) %>%
  count(dia, name = "y")

idx <- tibble(dia = seq(min(cont_dia_raw$dia), max(cont_dia_raw$dia), by = "1 day"))
cont_dia_full <- idx %>%
  left_join(cont_dia_raw, by = "dia") %>%
  mutate(y = replace_na(y, 0L))

y <- cont_dia_full$y
mu <- mean(y); va <- var(y)
cat(sprintf("[Diagnóstico DIARIO con ceros] n=%d media=%.3f var=%.3f\n",
  length(y), mu, va, ifelse(mu>0, va/mu, NA_real_)))

# --- 3) Ajuste y AIC: Poisson vs Binomial Negativa ---
# Poisson
loglik_pois <- sum(dpois(y, lambda = mu, log = TRUE))
AIC_pois <- -2*loglik_pois + 2*1

# NB (MLE con MASS::fitdistr)
fit_nb <- suppressWarnings(MASS::fitdistr(y, densfun = "Negative Binomial"))
size_hat <- fit_nb$estimate[["size"]]
mu_hat <- fit_nb$estimate[["mu"]]
loglik_nb <- sum(dnbinom(y, size = size_hat, mu = mu_hat, log = TRUE))
AIC_nb <- -2*loglik_nb + 2*2

aics_day <- dplyr::arrange(
  data.frame(dist = c("Poisson", "BinomialNegativa"),
    AIC = c(AIC_pois, AIC_nb)),

# --- 4) Pronóstico probabilístico DIARIO (NB MLE) ---
p_le5 <- pnbinom(5, size = size_hat, mu = mu_hat) #
P(X ≤ 5)
p_6_10 <- pnbinom(10, size = size_hat, mu = mu_hat) - pnbinom(5, size = size_hat, mu =
mu_hat) # P(6..10)
p_ge15 <- 1 - pnbinom(14, size = size_hat, mu = mu_hat) #
P(X ≥ 15)

escenarios_full <- tibble::tibble(
  Intervalo = c("Día", "Día", "Día"),
  Escenario = c("<=5 llamadas", "6-10 llamadas", ">=15 llamadas"),
  Probabilidad = round(c(p_le5, p_6_10, p_ge15), 3)
)
print(escenarios_full)
```

Apéndice E: Códigos en RStudio Para Simulación de Demanda

```
# --- Paquetes ---
pkgs <- c("readxl","dplyr","tidyr","lubridate","MASS")
to_install <- setdiff(pkgs, rownames(installed.packages()))
if (length(to_install)) install.packages(to_install, quiet = TRUE)
invisible(lapply(pkgs, library, character.only = TRUE))

# --- 1) Cargar y construir timestamp (robusto) ---
df <- readxl::read_excel("C:/Users/diego/OneDrive/Desktop/datos_tesis.xlsx")
col_fecha <- "Fecha"; col_hora <- "Hora de Recep."

fecha <- as.Date(df[[col_fecha]])
hraw <- df[[col_hora]]
hora_txt <- tryCatch(format(hraw, "%H:%M:%S"), error = function(e) as.character(hraw))
hora_secs <- suppressWarnings(as.numeric(lubridate::hms(hora_txt)))
if (all(is.na(hora_secs)) && is.numeric(hraw)) hora_secs <- round(hraw * 24 * 3600)

ts <- lubridate::as_datetime(fecha, tz = "America/Santiago") +
lubridate::seconds(hora_secs)

df <- df %>% mutate(ts = ts) %>% filter(!is.na(ts)) %>% arrange(ts)

# --- 2) Conteos DIARIOS (incluyendo ceros entre el primer y último día observado) ---
cont_dia_raw <- df %>%
  mutate(dia = as_date(ts)) %>%
  count(dia, name = "y")

idx <- tibble(dia = seq(min(cont_dia_raw$dia), max(cont_dia_raw$dia), by = "1 day"))
cont_dia_full <- idx %>%
  left_join(cont_dia_raw, by = "dia") %>%
  mutate(y = replace_na(y, 0L))

y <- cont_dia_full$y
mu <- mean(y); va <- var(y)
cat(sprintf("[Diagnóstico DIARIO con ceros] n=%d media=%.3f var=%.3f\n",
var/mean=%.3f\n",
length(y), mu, va, ifelse(mu>0, va/mu, NA_real_)))

# --- 3) Ajuste y AIC: Poisson vs Binomial Negativa ---
# Poisson
loglik_pois <- sum(dpois(y, lambda = mu, log = TRUE))
AIC_pois <- -2*loglik_pois + 2*1

# NB (MLE con MASS::fitdistr)
fit_nb <- suppressWarnings(MASS::fitdistr(y, densfun = "Negative Binomial"))
size_hat <- fit_nb$estimate[["size"]]
mu_hat <- fit_nb$estimate[["mu"]]
loglik_nb <- sum(dnbinom(y, size = size_hat, mu = mu_hat, log = TRUE))
AIC_nb <- -2*loglik_nb + 2*2

aics_day <- dplyr::arrange(
  data.frame(dist = c("Poisson","BinomialNegativa"),
            AIC = c(AIC_pois, AIC_nb)),
  AIC
)
print(aics_day)
P(X ≤ 5)
p_6_10 <- pnbinom(10, size = size_hat, mu = mu_hat) - pnbinom(5, size = size_hat, mu =
mu_hat) # P(6..10)
p_ge15 <- 1 - pnbinom(14, size = size_hat, mu = mu_hat) #
P(X ≥ 15)

escenarios_full <- tibble::tibble(
  Intervalo = c("Día","Día","Día"),
  Escenario = c("≤5 llamadas","6-10 llamadas","≥15 llamadas"),
  Probabilidad = round(c(p_le5, p_6_10, p_ge15), 3)
```

```
)
print(escenarios_full)

# --- 5) Simulación de escenarios diarios (Monte Carlo) ---
set.seed(123) # reproducible
n_sims <- 10000
sim_dias <- rnbinom(n_sims, size = size_hat, mu = mu_hat)

# Resumen y percentiles
resumen_sim <- summary(sim_dias)
percentiles <- quantile(sim_dias, probs = c(0.05, 0.10, 0.50, 0.90, 0.95))
print(resumen_sim)
print(percentiles)

# --- 5.1) Tabla de escenarios (ajusta umbrales si quieres) ---
escenarios_sim <- tibble::tibble(
  Escenario = c("<= 5 llamadas", "6-10 llamadas", "11-14 llamadas", "> 15
llamadas"),
  Probabilidad_simulada = c(
    mean(sim_dias <= 5),
    mean(sim_dias >= 6 & sim_dias <= 10),
    mean(sim_dias >= 11 & sim_dias <= 14),
    mean(sim_dias >= 15)
  )
) |> dplyr::mutate(Probabilidad_simulada = round(Probabilidad_simulada, 3))

print(escenarios_sim)

# --- 5.2) Niveles de capacidad sugeridos por percentil ---
capacidad_sugerida <- tibble::tibble(
  Nivel_servicio = c("80%", "85%", "90%", "95%"),
  Llamadas_dia = as.integer(quantile(sim_dias, probs = c(0.80, 0.85, 0.90, 0.95)))
)
print(capacidad_sugerida)

# --- 5.3) Gráfico histograma de la simulación ---
hist(sim_dias, breaks = 30, col = "skyblue",
  main = "Simulación de llamadas diarias (NB)",
  xlab = "Llamadas por día")
abline(v = percentiles["90%"], lwd = 2) # marca P90
text(x = percentiles["90%"], y = par("usr")[4]*0.9,
  labels = paste0("P90 = ", as.integer(percentiles["90%"])),
  pos = 4)
```

Apéndice F: Códigos en RStudio Para Simulación de Escenarios

```
library(readxl)
library(dplyr)
library(lubridate)
library(stringi)

# 1) Cargar datos
datos <- read_excel(file.choose())

# ===== ENTRADAS =====
llamadas_G1 <- 14
cumplen_G1 <- 5
no_cumplen_G1 <- 9

# Proporciones históricas de focalización en Grupo 1
prop_focalizacion <- c(
  Interior = 0.42,
  Medidor = 0.401,
  Exterior = 0.179/2,          # si quieres separar Exterior y Medidor ajusta aquí
  Sin_Suministro = 0.179/2   # aquí otra parte si quieres distinguirlos
)

# Tasa de cumplimiento histórica en Grupo 1 (según tu tabla)
tasa_cumpl_foc_G1 <- c(
  Interior = 0.387,
  Medidor = 0.0586,
  Exterior = 0.0769,
  Sin_Suministro = 0.869
)

# ===== CÁLCULO =====
# Distribuir las 14 llamadas según proporciones
raw <- llamadas_G1 * prop_focalizacion
base <- floor(raw)
resto <- llamadas_G1 - sum(base)
if(resto > 0){
  order_idx <- order(raw - base, decreasing = TRUE)
  base[order_idx[seq_len(resto)]] <- base[order_idx[seq_len(resto)]] + 1
}
llamadas_por_foc <- base

# Calcular cumplen/no cumplen según tasa de cumplimiento de cada focalización
cumplen <- round(llamadas_por_foc * tasa_cumpl_foc_G1)
no_cumplen <- llamadas_por_foc - cumplen

resultado_G1 <- data.frame(
  Focalizacion = names(llamadas_por_foc),
  Llamadas = as.integer(llamadas_por_foc),
  Cumplen = cumplen,
  No_Cumplen = no_cumplen,
  Tasa_Cumplimiento = round(cumplen / llamadas_por_foc, 3)
)

# ===== RESULTADO =====
resultado_G1

# ===== ENTRADAS =====
llamadas_G1 <- 14
cumplen_actual <- 5
no_cumplen_actual <- 9

# Probabilidad de simultaneidad (ejemplo: ventana de 30 min del archivo)
p_simult <- 0.37
```

```
# Supuesto: el nuevo móvil resuelve todos los choques de 2 llamadas (f=1)
# pero no resuelve choques de 3+ llamadas. Ajustable.
f <- 1

# ===== CÁLCULO =====
recuperados <- round(no_cumplen_actual * p_simult * f)
no_cumplen_nuevo <- no_cumplen_actual - recuperados
cumplen_nuevo <- llamadas_G1 - no_cumplen_nuevo

resultado <- data.frame(
  Escenario = c("Actual (3 móviles)", "Con +1 móvil"),
  Cumplen = c(cumplen_actual, cumplen_nuevo),
  No_Cumplen = c(no_cumplen_actual, no_cumplen_nuevo),
  Tasa_Cumplimiento = c(cumplen_actual/llamadas_G1, cumplen_nuevo/llamadas_G1)
)

print(resultado)

library(dplyr)

# ===== ENTRADAS =====
N_G1 <- 14 # llamadas futuras (Grupo 1)
target_cumplen <- 5 # de tus modelos (3 móviles)
target_nocumplen <- N_G1 - target_cumplen

# Proporciones proyectadas (3 categorías)
prop <- c(Interior = 0.42, ExtMed = 0.401, Otras = 0.179)

# Tasas de cumplimiento históricas (Grupo 1)
tasa <- c(
  Interior = 0.387, # 38.7%
  ExtMed = 0.065, # promedio ponderado Medidor+Exterior ≈ 6.5%
  Otras = 0.869 # 86.9%
)

# Probabilidad de simultaneidad (ej. ventana 30 min)
p_simult <- 0.37

# Factores de sensibilidad f por categoría
f <- c(Interior = 0.4, ExtMed = 0.9, Otras = 0.0)

# ===== UTILIDAD: asignación entera =====
allocate_integer <- function(p, total){
  raw <- p * total
  base <- floor(raw)
  k <- total - sum(base)
  if(k > 0){
    add_idx <- order(raw - base, decreasing = TRUE)[seq_len(k)]
    base[add_idx] <- base[add_idx] + 1
  }
  base
}

# ===== 1) DISTRIBUCIÓN DE LLAMADAS =====
llam <- allocate_integer(prop, N_G1)

# ===== 2) ESCENARIO BASE (3 móviles) =====
raw_cumplen <- llam * tasa
cumplen_base <- floor(raw_cumplen)
# Calibrar al total=5 cumplimientos
delta <- target_cumplen - sum(cumplen_base)
if(delta > 0){
  restos <- raw_cumplen - cumplen_base
  add_idx <- order(restos, decreasing = TRUE)[seq_len(delta)]
  cumplen_base[add_idx] <- cumplen_base[add_idx] + 1
}
```

```
}
nocumplen_base <- llam - cumplen_base

base_tbl <- data.frame(
  Categoria = names(llam),
  Llamadas = as.integer(llam),
  Cumplen = as.integer(cumplen_base),
  No_Cumplen = as.integer(nocumplen_base),
  Tasa_Cumplimiento = round(cumplen_base / pmax(llam,1), 3)
)

# ===== 3) ESCENARIO +1 MÓVIL (4 móviles) =====
recuperados <- round(nocumplen_base * p_simult * f)
nocumplen_new <- pmax(nocumplen_base - recuperados, 0)
cumplen_new <- llam - nocumplen_new

plus1_tbl <- data.frame(
  Categoria = names(llam),
  Llamadas = as.integer(llam),
  Cumplen_new = as.integer(cumplen_new),
  No_Cumplen_new = as.integer(nocumplen_new),
  Tasa_Cumplimiento_new = round(cumplen_new / pmax(llam,1), 3)
)

# ===== 4) RESUMEN GLOBAL =====
resumen <- data.frame(
  Escenario = c("Base (3 móviles)", "+1 móvil (4 móviles)"),
  Cumplen = c(sum(base_tbl$Cumplen), sum(plus1_tbl$Cumplen_new)),
  No_Cumplen = c(sum(base_tbl$No_Cumplen), sum(plus1_tbl$No_Cumplen_new))
)
resumen$Tasa_Cumplimiento <- round(resumen$Cumplen / N_Gl, 3)

# Mostrar resultados
base_tbl
plus1_tbl
resumen

library(dplyr)

# ---- Erlang C: espera promedio Wq (horas) ----
erlangC_Wq <- function(lambda, mu, c){
  rho <- lambda / (c * mu)
  if (rho >= 1) return(Inf)
  # sumatorio 0..c-1
  k <- 0:(c-1)
  sum_term <- sum((c*rho)^k / factorial(k))
  last_term <- (c*rho)^c / (factorial(c) * (1 - rho))
  Pw <- last_term / (sum_term + last_term) # prob de esperar (Erlang C)
  Wq <- Pw / (c*mu - lambda) # espera promedio en cola (horas)
  return(Wq)
}

# ---- Calibrar mu para que Wq(c=3) = Wq_obs ----
calibrar_mu <- function(lambda, c, Wq_obs_h, mu_low = 1e-4, mu_high = 5){
  # mu debe ser > lambda/c para estabilidad; ajustamos el límite inferior:
  mu_min <- lambda/c + 1e-6
  lower <- max(mu_low, mu_min)
  f <- function(mu) erlangC_Wq(lambda, mu, c) - Wq_obs_h
  # expandir hasta encontrar signos opuestos (por si acaso)
  a <- lower; b <- mu_high
  fa <- f(a); fb <- f(b)
  tries <- 0
  while (is.finite(fa) && is.finite(fb) && fa*fb > 0 && tries < 20){
    # si no hay cruce, abre el intervalo hacia arriba
    b <- b*1.5
    fb <- f(b)
    tries <- tries + 1
  }
}
```

```
if (!is.finite(fa) || !is.finite(fb) || fa*fb > 0){
  return(NA_real_) # no se pudo calibrar con estos supuestos
}
uniroot(f, c(a,b))$root
}

# ===== ENTRADAS (Grupo 1) =====
lambda <- 14/24 # llamadas/hora (P90 = 14 llamadas/día en Grupo 1)

# Tiempos observados (promedios) con 3 móviles, MINUTOS:
T_total_obs_min <- c(Interior = 90.0, ExtMed = 91.9, SinSuministro = 108.0)

# Traslado asumido por focalización, MINUTOS:
T_traslado_min <- c(Interior = 40, ExtMed = 40, SinSuministro = 40)

# Espera observada con 3 móviles (minutos y horas):
Wq_obs_min <- pmax(T_total_obs_min - T_traslado_min, 0) # no negativa
Wq_obs_h <- Wq_obs_min / 60

# ===== CALIBRACIÓN DE  $\mu$  (por focalización) =====
mu_cal <- sapply(names(T_total_obs_min), function(cat){
  calibrar_mu(lambda = lambda, c = 3, Wq_obs_h = Wq_obs_h[cat])
})
mu_cal
# Nota: si algún  $\mu$  sale NA, significa que con esos supuestos ( $\lambda$ ,  $Wq\_obs$ ) no hay
solución estable.
# En ese caso, revisa traslado o  $\lambda$ .

# ===== CÁLCULO DE ESCENARIOS c=3,4,5 =====
escenarios <- expand.grid(Focalizacion = names(T_total_obs_min), Moviles = 3:6) %>%
  arrange(Focalizacion, Moviles)

result <- escenarios %>%
  rowwise() %>%
  mutate(
    mu = mu_cal[Focalizacion],
    Wq_h = erlangC_Wq(lambda, mu, Moviles),
    Wq_m = Wq_h * 60,
    T_total_m = T_traslado_min[Focalizacion] + Wq_m
  ) %>%
  ungroup()

print(result)

# (Opcional) Comparar con estándar normativo:
estandares <- c(Interior = 60, ExtMed = 30, SinSuministro = 180)
result_comp <- result %>%
  mutate(
    Estandar = estandares[Focalizacion],
    Exceso_min = T_total_m - Estandar
  )
print(result_comp)

# GRUPO 2

library(dplyr)

# -----
# Función Erlang-C (Wq en horas)
# -----
erlangC_Wq <- function(lambda, mu, c){
  rho <- lambda / (c * mu)
  if (rho >= 1) return(Inf)

  k <- 0:(c-1)
  sum_term <- sum((c*rho)^k / factorial(k))
  last_term <- (c*rho)^c / (factorial(c) * (1 - rho))
  Pw <- last_term / (sum_term + last_term) # prob de esperar
```

```
Wq <- Pw / (c*mu - lambda) # espera promedio (horas)
return(Wq)
}

# -----
# Calibración de mu (como hicimos en Grupo 1)
# -----
calibrar_mu <- function(lambda, c, Wq_obs_h, mu_low = 1e-4, mu_high = 5){
  mu_min <- lambda/c + 1e-6
  lower <- max(mu_low, mu_min)
  f <- function(mu) erlangC_Wq(lambda, mu, c) - Wq_obs_h
  a <- lower; b <- mu_high
  fa <- f(a); fb <- f(b)
  tries <- 0
  while (is.finite(fa) && is.finite(fb) && fa*fb > 0 && tries < 20){
    b <- b*1.5
    fb <- f(b)
    tries <- tries + 1
  }
  if (!is.finite(fa) || !is.finite(fb) || fa*fb > 0){
    return(NA_real_)
  }
  uniroot(f, c(a,b))$root
}

# ===== ENTRADAS (Grupo 2) =====
lambda <- 2/24 # llamadas/hora

# Tiempos observados promedio con 2 móviles (tomados de tu tabla global)
T_total_obs_min <- c(Interior = 90.0, ExtMed = 91.9, SinSuministro = 108.0)

# Supuesto de traslado (minutos)
T_traslado_min <- c(Interior = 40, ExtMed = 40, SinSuministro = 40)

# Espera observada (min y horas)
Wq_obs_min <- pmax(T_total_obs_min - T_traslado_min, 0)
Wq_obs_h <- Wq_obs_min / 60

# ===== Calibrar mu con c=2 (dotación base) =====
mu_cal <- sapply(names(T_total_obs_min), function(cat){
  calibrar_mu(lambda = lambda, c = 2, Wq_obs_h = Wq_obs_h[cat])
})
mu_cal

# ===== Escenarios con c = 1,2,3 =====
escenarios <- expand.grid(Focalizacion = names(T_total_obs_min), Moviles = 1:3) %>%
  arrange(Focalizacion, Moviles)

result <- escenarios %>%
  rowwise() %>%
  mutate(
    mu = mu_cal[Focalizacion],
    Wq_h = erlangC_Wq(lambda, mu, Moviles),
    Wq_m = Wq_h * 60,
    T_total_m = T_traslado_min[Focalizacion] + Wq_m
  ) %>%
  ungroup()

print(result)

# Comparar con estándar normativo
estandares <- c(Interior = 60, ExtMed = 30, SinSuministro = 180)
result_comp <- result %>%
  mutate(
    Estandar = estandares[Focalizacion],
    Exceso_min = T_total_m - Estandar
  )
print(result_comp)
```