

# REPORTE TÉCNICO Y ESTRATÉGICO

Sistema Estocástico de Gestión de Riesgo y Capacidad de  
Proveedores

Temporada Alta Demanda: Julio – Diciembre 2026

---

Johanna Camila Willis Ruiz

## 1. Resumen Ejecutivo y Planteamiento del Problema

Este proyecto desarrolla un motor matemático de simulación estocástica para proyectar la demanda, cuantificar el riesgo de desabasto y recomendar niveles óptimos de inventario de seguridad para cada proveedor de la red durante la temporada julio–diciembre 2026.

El dataset de entrada contiene 745,340 registros transaccionales de actividad de ventas (no de inventario) correspondientes a julio 2023 – diciembre 2025, cubriendo 44 proveedores en 5 sucursales para la categoría de calzado (botas). Cada registro especifica: mes, año, proveedor, modelo, talla, sucursal, unidades vendidas y unidades negadas.

### 1.1 Por qué no se usó Machine Learning o Series de Tiempo Clásicas

El reto analítico central es la escasez de profundidad temporal. Aunque el volumen total de registros es alto, al aislar el comportamiento de un proveedor específico en un mes concreto, por ejemplo PROV15 en noviembre, la muestra se reduce a los períodos homólogos de 2023, 2024 y 2025. El tamaño real de la serie es  $N = 3$ .

Bajo esta restricción:

- SARIMA / ARIMAX requieren mínimo 30–50 observaciones para estimación estable de parámetros.
- Redes neuronales y gradient boosting incurrirían en sobreajuste masivo (overfitting) al memorizar los 3 puntos sin capacidad de generalización.
- El Método Monte Carlo, en cambio, usa esas 3 observaciones para ajustar una distribución de probabilidad y genera 10,000 escenarios sintéticos, cuantificando la incertidumbre de manera honesta y explícita.

Esta elección no es una limitación, sino la decisión metodológicamente correcta para el tamaño de muestra disponible.

### 1.2 Por qué Monte Carlo y no otros métodos alternativos

La elección del Método Monte Carlo no fue arbitraria. Se evaluaron tres alternativas antes de seleccionarlo:

Alternativa 1 — Bootstrap Paramétrico:

- ¿Qué hace? Remuestrea con reemplazo los datos históricos existentes miles de veces para construir una distribución empírica del exceso de demanda.
- ¿Por qué no se usó? Con  $N = 3$  observaciones por mes, el Bootstrap solo puede generar combinaciones de esos mismos 3 puntos. No puede simular escenarios peores que el peor año observado. Para recomendaciones conservadoras de inventario de seguridad, esa limitación es crítica: si 2026 resulta ser un año de demanda excepcionalmente alta, el Bootstrap no lo capturaría.
- Ventaja de Monte Carlo: al ajustar una distribución paramétrica (Gamma/Normal) y muestrear de ella, el modelo puede generar escenarios fuera del rango histórico observado, lo que produce recomendaciones más robustas.

Alternativa 2 — Redes Bayesianas:

- ¿Qué hacen? Incorporan conocimiento previo (prior) sobre la distribución de la demanda y lo actualizan con los datos observados para obtener una distribución posterior más informada.
- ¿Por qué no se usaron? Son el método más sofisticado para muestras pequeñas, pero requieren definir priors explícitos para cada proveedor (información que no está

disponible), su implementación es considerablemente más compleja, y su interpretación operativa para equipos de Supply Chain no familiarizados con probabilidad bayesiana presenta barreras prácticas.

- Es la alternativa más potente teóricamente, pero su complejidad supera el beneficio marginal dado el contexto del proyecto.

Alternativa 3 — SARIMA / Modelos de Series de Tiempo:

- Requieren mínimo 30–50 observaciones para estimación estable de parámetros estacionales. Con  $N = 3$  incurirían en sobreajuste masivo (overfitting), memorizando los 3 puntos sin capacidad real de predicción.

El Método Monte Carlo representa el balance óptimo entre rigor estadístico, interpretabilidad operativa y viabilidad con los datos disponibles.

### 1.3 Escalabilidad del Modelo con más Datos

Una pregunta natural es si el modelo mejoraría con más años de historia. La respuesta es sí, y de manera significativa, sin necesidad de cambiar el método:

- Con 5–7 años ( $N = 5–7$  por mes): la distribución ajustada sería más fiel a la demanda real, los intervalos P5–P95 se cerrarían considerablemente y la probabilidad de desabasto sería más diferenciada entre proveedores en lugar de concentrarse en el rango 50–70%.
- Con 8–10 años: el Cross-Validation Walk-Forward tendría suficientes folds para estimar el alpha con mayor precisión estadística. También permitiría detectar ciclos plurianuales de demanda (e.g., booms económicos que aumentan ventas cada 3–4 años).
- El pipeline está diseñado para escalar automáticamente: agregar años de datos al CSV y re-ejecutar el código produce predicciones actualizadas sin modificación alguna del código.

Esta escalabilidad convierte el modelo en una herramienta de uso continuo, no en un análisis de una sola vez. La recomendación es actualizar las simulaciones trimestralmente conforme se acumula más historia transaccional.

## 2. Arquitectura del Pipeline Analítico

El sistema se divide en ocho módulos funcionales ejecutados secuencialmente. Cada módulo recibe un input definido, aplica una transformación matemática y entrega un output concreto.

Módulo	Función Principal	Input	Output
cargar_datos()	Ingesta y limpieza	CSV transaccional	DataFrame filtrado Jul–Dic
calcular_metricas()	Radiografía histórica	DataFrame limpio + $\alpha$	Tabla de riesgo con score por proveedor
cross_validate_pesos()	Optimización Walk-Forward	DataFrame + rango de $\alpha$	$\alpha$ óptimo y $\rho$ de Spearman
ajustar_distribucion()	Selección de distribución MLE + AIC	Serie numérica ( $N \geq 1$ )	Nombre y parámetros de distribución óptima
simular_proveedor_estacional()	Motor Monte Carlo	Datos por proveedor y mes	10,000 escenarios de exceso por mes
simular_transferencia_volumen()	Análisis what-if estratégico	Escenarios del proveedor problemático + datos del estrella	Reducción de riesgo y volumen salvado
plot_*	Visualización	Resultados de simulación	Gráficos PNG (heatmaps, dashboards, barras)
generar_tablas()	Exportación	Todos los resultados	3 archivos CSV entregables

## 3. Fundamentos Matemáticos y Estadísticos

### 3.1 Reconstrucción de la Demanda Real (Factor de Negados)

Los datos transaccionales registran ventas concretadas, no la demanda total del mercado. Cuando no hay inventario disponible, la venta no ocurre aunque el cliente la haya solicitado. Esta diferencia se conoce como demanda censurada y debe corregirse antes de cualquier análisis.

Para estimar la demanda latente verdadera se integra el parámetro FACTOR\_NEGADOS = 0.25, derivado del conocimiento del negocio (domain knowledge):

$$\text{Demanda\_Real} = \text{Ventas} + (\text{Negados} \times 0.25)$$

La justificación operativa es la siguiente: ante un quiebre de stock (una unidad negada), el 75% de los clientes adquieren un producto sustituto dentro de la misma tienda, por lo que esa venta no se pierde para la empresa. El 25% restante abandona la tienda sin comprar, representando una pérdida neta de demanda que sí debe ser considerada en el modelo de capacidad del proveedor.

Este ajuste es crítico porque sin él se subestimaría sistemáticamente la demanda real, generando recomendaciones de inventario de seguridad insuficientes.

### 3.2 Score de Riesgo por Proveedor

Para clasificar a los proveedores según su historial de confiabilidad se construye un score compuesto que combina dos señales:

Componente 1 — Tasa de Negados (Fulfillment Gap):

$$\text{Tasa\_Negados} = 1 - (\text{Ventas\_Total} / \text{Demanda\_Total\_Ajustada})$$

Esta métrica mide qué fracción de la demanda real quedó insatisfecha. Un valor de 0.09 significa que el proveedor no pudo cubrir el 9% de lo que el mercado pidió. Es el indicador más directo de riesgo de desabasto.

Componente 2 — Coeficiente de Variación (Volatilidad de la demanda):

$$CV = \sigma_{\text{demanda}} / \mu_{\text{demanda}}$$

Mide la dispersión relativa de la demanda. Un proveedor con demanda muy irregular mes a mes es más difícil de planear aunque su cumplimiento promedio sea alto.

Score de Riesgo compuesto:

$$\text{Score\_Riesgo} = \alpha \times \text{Tasa\_Negados} + (1 - \alpha) \times CV$$

El parámetro  $\alpha$  controla el peso relativo de cada componente. El valor óptimo de  $\alpha$  se determina mediante validación cruzada (Sección 3.3). En este proyecto el modelo convergió a  $\alpha = 1.0$ , lo que significa que la Tasa de Negados por sí sola es el predictor más informativo del riesgo futuro.

Clasificación en tres niveles de riesgo mediante percentiles adaptativos:

$$\text{BAJO} \rightarrow \text{Score} \leq P_{33} \quad | \quad \text{MEDIO} \rightarrow P_{33} < \text{Score} \leq P_{66} \quad | \quad \text{ALTO} \rightarrow \text{Score} > P_{66}$$

Al usar los percentiles 33 y 66 de la distribución real de scores (en lugar de umbrales fijos), la clasificación se adapta automáticamente a la dispersión del portafolio analizado, garantizando siempre una distribución equilibrada de los tres niveles independientemente de la escala de los datos.

### 3.3 Validación Cruzada Walk-Forward y Coeficiente de Spearman

Para determinar el  $\alpha$  óptimo se implementó un esquema de Walk-Forward Cross-Validation, el único método válido para datos temporales. El K-Fold clásico aleatoriza el tiempo y violaría la causalidad: usaría datos futuros para predecir el pasado.

Esquema Walk-Forward con  $N = 3$  años:

- Fold 1: Entrena con 2023, valida con 2024.
- Fold 2: Entrena con 2023–2024, valida con 2025.

En cada fold se evalúa si el Score\_Riesgo calculado en el periodo de entrenamiento predice correctamente la Tasa\_Negados del periodo de validación. La métrica de evaluación es el Coeficiente de Correlación de Rangos de Spearman:

$$\rho = 1 - (6 \times \sum d_i^2) / (n \times (n^2 - 1))$$

donde  $d_i$  es la diferencia de rangos entre el score histórico y la tasa futura para el proveedor  $i$ . Se usa Spearman (y no Pearson) porque el objetivo no es predecir el valor exacto de la falla futura, sino garantizar que el ranking de riesgo se preserve: el proveedor más riesgoso hoy debe seguir siendo el más riesgoso mañana.

Resultado: el modelo convergió a  $\alpha = 1.0$  con  $\rho \approx 0.30$ . En entornos de alta volatilidad con  $N = 3$ , una correlación de rangos de 0.30 representa una señal direccional válida y estadísticamente superior al azar puro ( $\rho = 0$ ), lo que valida que la Tasa de Negados es un predictor robusto del riesgo futuro.

### 3.4 Ajuste Dinámico de Distribuciones por Criterio AIC

Para simular demanda futura es necesario definir cómo se distribuyen estadísticamente los datos históricos de cada proveedor en cada mes. El sistema evalúa dos familias candidatas y elige la óptima automáticamente.

Se aplica Estimación por Máxima Verosimilitud (MLE) para ajustar los parámetros de cada distribución a los datos empíricos, y se selecciona la distribución que minimiza el Criterio de Información de Akaike:

$$AIC = 2k - 2 \times \ln(\hat{L})$$

donde  $k$  es el número de parámetros del modelo y  $\hat{L}$  es la verosimilitud máxima. El AIC penaliza la complejidad del modelo, favoreciendo distribuciones más parsimoniosas cuando la diferencia de ajuste no es significativa.

Las dos distribuciones candidatas son:

- Distribución Normal  $N(\mu, \sigma)$ : adecuada cuando los datos son simétricos y la varianza no depende de la magnitud. Parámetros: media  $\mu$  y desviación estándar  $\sigma$ .
- Distribución Gamma  $\Gamma(k, \theta)$ : estándar en gestión de inventarios estocásticos. Su dominio es estrictamente positivo ( $x > 0$ ), evita demandas negativas imposibles y modela naturalmente la asimetría positiva de los picos de compra. Se fuerza el parámetro de locación a cero ( $floc = 0$ ) para garantizar este dominio. Parámetros: forma  $k$  y escala  $\theta$ .

La distribución ganadora y sus parámetros se documentan por proveedor y mes en la columna DIST\_DEMANDA de la tabla de predicción.

### 3.5 Motor de Simulación Monte Carlo Estacional

El componente central del sistema. Para cada proveedor  $p$  y cada mes  $m \in \{\text{Jul, Ago, Sep, Oct, Nov, Dic}\}$ , el algoritmo:

### Paso 1 — Construcción de series históricas:

$$\text{Serie_Demanda}(p,m) = \{\text{Demanda_Real}(p,m,2023), \text{Demanda_Real}(p,m,2024), \text{Demanda_Real}(p,m,2025)\}$$

$$\text{Serie_Capacidad}(p,m) = \{\text{Ventas}(p,m,t) : \text{Negados}(p,m,t) = 0, t \in \{2023, 2024, 2025\}\}$$

La capacidad del proveedor se estima exclusivamente de los períodos sin negados: si no hubo quiebre de stock, las ventas representan la demanda satisfecha sin restricción de capacidad. Si no existe ningún mes sin negados (proveedor en desabasto crónico), se aplica un buffer del 5% sobre las ventas máximas históricas (ver Sección 4.2).

Paso 2 — Ajuste de distribuciones (Sección 3.4) para demanda y capacidad.

Paso 3 — Generación de 10,000 escenarios independientes:

$$\begin{aligned} D_{sim}[i] &\sim \text{Dist_Demanda}(\text{params\_dem}) \quad \forall i \in \{1, \dots, 10000\} \\ C_{sim}[i] &\sim \text{Dist_Capacidad}(\text{params\_cap}) \quad \forall i \in \{1, \dots, 10000\} \\ \text{Exceso}[i] &= \max(0, D_{sim}[i] - C_{sim}[i]) \end{aligned}$$

El exceso representa la demanda que el proveedor no puede cubrir en el escenario  $i$ . Cuando  $\text{Exceso} = 0$ , el proveedor abastece toda la demanda; cuando  $\text{Exceso} > 0$ , hay desabasto.

Paso 4 — Métricas de salida por mes:

$$\begin{aligned} \text{Demanda_Esperada} &= E[D_{sim}] = (1/10000) \times \sum D_{sim}[i] \\ \text{Prob_Desabasto} &= P(\text{Exceso} > 0) = \#\{\text{Exceso}[i] > 0\} / 10000 \\ \text{Stock_Seguridad_P95} &= \text{Percentil}_{95}(\text{Exceso}) \end{aligned}$$

El Stock de Seguridad en el percentil 95 significa que si la empresa mantiene ese nivel de inventario adicional, tendrá cobertura suficiente en el 95% de los 10,000 universos posibles simulados. Calcular la media del exceso solo garantizaría cobertura el 50% de las veces.

### 3.6 Análisis de Tendencia Interanual

Para capturar el crecimiento o declive secular de la demanda de cada proveedor, se ajusta una regresión lineal simple sobre los tres puntos históricos:

$$\text{Tendencia}(p,m) = \beta_1 \quad \text{donde } \beta_1 \text{ es la pendiente de OLS sobre } \{(2023, D_{2023}), (2024, D_{2024}), (2025, D_{2025})\}$$

Para evitar proyecciones explosivas derivadas de variaciones extremas entre solo 3 puntos (e.g., un crecimiento de 100 a 400 unidades generaría una pendiente irreal), se aplica un clipping al  $\pm 20\%$  de la media histórica mensual del proveedor:

$$\text{Tendencia_Acotada} = \text{clip}(\text{Tendencia_cruda}, -0.20 \times \mu_{hist}, +0.20 \times \mu_{hist})$$

## 4. Heurísticas y Reglas de Control

Al modelar procesos estocásticos con muestras muy pequeñas, la matemática pura puede generar valores incontrolables en las colas de las distribuciones. Se implementaron tres reglas de control para anclar el modelo a la realidad operativa.

### 4.1 Cap de Cola Larga Gamma (Límite 5x)

La distribución Gamma con alta varianza y  $N = 3$  puede generar, en la cola derecha, valores decenas de veces superiores a la media histórica. Para que un escenario extremo entre 10,000 no distorsione el percentil 95:

```
Exceso_capado[i] = min(Exceso[i], 5 * μ_demanda_histórica_mensual)
```

Este límite establece que ningún escenario de exceso puede superar 5 veces la demanda media histórica del proveedor en ese mes. Es equivalente a asumir que el mercado no puede crecer más de 5 veces su nivel promedio en un horizonte de un año, lo cual es un supuesto conservador razonable.

### 4.2 Buffer de Capacidad Latente (+5%)

Algunos proveedores presentan historial de desabasto crónico: nunca tuvieron un mes sin negados, por lo que la serie de capacidad está vacía. Este fenómeno se conoce estadísticamente como datos censurados por la derecha: se sabe que la capacidad real es mayor o igual a las ventas observadas, pero no se conoce el límite superior real.

Si el algoritmo asumiera que la capacidad máxima es exactamente igual a las ventas históricas, sobreestimaría el riesgo de manera artificiosa. En su lugar, se otorga un buffer de elasticidad operativa:

```
Capacidad_latente = Ventas_históricas × 1.05
```

El +5% representa la holgura mínima que cualquier proveedor puede activar ante un aumento de demanda (horas extra, optimización de línea), sin requerir inversión de capital adicional.

### 4.3 Semilla Aleatoria Determinista

Para garantizar reproducibilidad total de los resultados —requisito fundamental en cualquier análisis científico— el generador de números pseudoaleatorios se inicializa con una semilla fija:

```
rng = np.random.default_rng(seed=42)
```

Esto asegura que cada ejecución del pipeline produce exactamente los mismos 10,000 escenarios, los mismos porcentajes de desabasto y las mismas recomendaciones de stock de seguridad. Sin semilla fija, los resultados variarían en cada corrida, haciendo imposible la auditoría y comparación de versiones.

## 5. Hallazgos y Análisis Operativo

### 5.1 Radiografía Histórica de Riesgo (2023–2025)

La tabla de riesgo histórico clasifica los 44 proveedores en tres niveles. Los 15 proveedores de nivel ALTO concentran las tasas de negados más elevadas del portafolio:

Proveedor	Fulfillment Rate	Tasa Negados	Nivel Riesgo
PROV16	84.6%	15.4%	ALTO
PROV15	90.9%	9.1%	ALTO
PROV48	91.1%	8.9%	ALTO
PROV32	91.7%	8.3%	ALTO
PROV10	96.0%	4.0%	ALTO
...	...	...	ALTO
PROV1	99.4%	0.6%	BAJO
PROV33	99.5%	0.5%	BAJO
PROV6	99.3%	0.7%	BAJO

El Fulfillment Rate de los tres mejores proveedores supera el 99.2%, frente al 84.6%–91.7% de los cinco peores. Esta diferencia de 8 a 15 puntos porcentuales tiene impacto directo en ventas perdidas.

### 5.2 El Hallazgo Estratégico: Complejidad vs. Especialización

Al cruzar los outputs de riesgo con la dimensión MODELO del dataset, emerge el insight más crítico del estudio: la causa raíz del desabasto sistémico no es la falta de capacidad bruta, sino la complejidad de manufactura.

Proveedores de Alto Riesgo (15 proveedores, PROV15, PROV48, PROV32, etc.) fabrican simultáneamente entre 8 y 19 modelos diferentes de calzado. Su ineficiencia proviene de los costos de complejidad:

- Setup times: tiempos de preparación al cambiar entre modelos en la línea de producción.
- Rotación de moldes/hormas: cada modelo requiere herramiental específico.
- Fragmentación de corridas de producción: lotes más pequeños por modelo reducen la eficiencia de escala.

Pool Estrella (PROV33, PROV6, PROV1) demuestran hiper-especialización productiva:

- PROV33: 1 modelo exclusivo. Tasa de negados: 0.52%.
- PROV6: 1 modelo exclusivo. Tasa de negados: 0.74%.
- PROV1: 4 modelos. Tasa de negados: 0.56%.

El análisis de intersección de SKUs confirma que ninguno de los tres proveedores estrella comparte moldes o modelos con el resto de la red. Su excelencia operativa está directamente vinculada a mantener líneas de producción continuas y especializadas.

### 5.3 Predicción Monte Carlo Jul–Dic 2026

Las probabilidades de desabasto proyectadas reflejan el nivel de incertidumbre real del sistema dado N = 3 observaciones por mes. Los meses críticos identificados por proveedor son:

Proveedor	Mes Crítico	Prob. Desabasto	SS Recomendado (P95)
PROV15	Nov	73.6%	17,694,591 u
PROV48	Dic	93.2%	14,249,567 u
PROV41	Oct	73.7%	12,638,434 u
PROV32	Dic	93.6%	3,265,981 u
PROV10	Oct	82.5%	3,585,537 u
PROV44	Dic	96.3%	2,694,203 u
PROV1	Dic	83.2%	1,809,759 u
PROV33	Nov	39.1%	61,089 u
PROV6	Dic	44.5%	8,360 u

PROV33 y PROV6, los dos proveedores con 1 modelo exclusivo, presentan probabilidades de desabasto en el rango 36–45% — significativamente inferiores al promedio del portafolio. Sus bajos requerimientos de inventario de seguridad confirman su estabilidad operativa.

#### 5.4 Casos de Negocio: Simulación de Transferencia de Volumen

El simulador what-if cuantifica matemáticamente cuánto riesgo se reduciría y cuántas ventas se recuperarían si parte de la demanda actualmente asignada a proveedores de alto riesgo fuera transferida al pool estrella. Los 5 casos de negocio de mayor impacto son:

Proveedor Origen	Transferir a	Volumen Salvado Jul–Dic 2026 (u)
PROV10	PROV1	999,722
PROV48	PROV1	902,621
PROV15	PROV1	866,114
PROV10	PROV33	787,200
PROV48	PROV33	777,229
PROV32	PROV33	720,478

El 'volumen salvado' representa la diferencia entre el exceso de demanda no cubierto con el proveedor original y el exceso residual después de que el proveedor estrella absorba parte del déficit. Es el proxy cuantitativo del ROI de la negociación estratégica.

## 6. Recomendaciones Tácticas y Estratégicas

### 6.1 Acciones Tácticas Inmediatas (Jul–Dic 2026)

La tabla\_prediccion\_2026.csv funge como lista de compras cuantificada. Define mes a mes y proveedor por proveedor el Stock de Seguridad adicional a negociar con cada proveedor. Los tres proveedores que requieren atención prioritaria inmediata son PROV15, PROV48 y PROV41, que concentran más del 60% del SS total recomendado para la temporada.

Nota sobre el costo de esta táctica: mantener inventario de seguridad de manera permanente inmoviliza capital de trabajo y genera costos de almacenamiento. Esta estrategia es necesaria en el corto plazo pero no sostenible como modelo permanente de gestión, lo que justifica la estrategia de mediano plazo.

### 6.2 Plan Estratégico de Mediano Plazo (Transferencia de Volumen)

La solución estructural al problema de desabasto crónico es diversificar la asignación de volumen hacia proveedores especializados. La tabla\_transferencia\_estrategica.csv provee los argumentos cuantitativos exactos para cada negociación.

Plan de acción propuesto:

- Abrir mesa de Desarrollo de Proveedores con PROV33, PROV6 y PROV1, presentando los volúmenes proyectados Jul–Dic 2026 como base de negociación de capacidad adicional.
- Identificar cuáles modelos de mayor rotación están asignados a PROV15, PROV48 y PROV32, y evaluar si PROV33/PROV6/PROV1 pueden absorberlos sin comprometer su especialización actual.
- Implementar contratos de desempeño con SLA (Service Level Agreement) basados en Fulfillment Rate mínimo del 97%, respaldados por los datos históricos del modelo.
- Establecer un comité de revisión trimestral que actualice las simulaciones con los datos de ventas más recientes, ajustando el SS recomendado conforme avanza el año.

### 6.3 Limitación Operativa de la Recomendación de Transferencia

Es importante señalar que la simulación de transferencia de volumen asume que el pool estrella tiene capacidad instalada suficiente para absorber el déficit de los proveedores de alto riesgo. Los datos transaccionales no incluyen información sobre capacidad máxima instalada ni costos de expansión de línea.

Por lo tanto, la viabilidad operativa de cada transferencia debe validarse mediante una evaluación de capacidad directa con cada proveedor antes de comprometer el volumen. La tabla de transferencia provee los casos de negocio (el 'qué'), pero la ingeniería de manufactura debe confirmar la factibilidad (el 'cómo').

## 7. Limitaciones del Estudio y Líneas de Mejora Futura

### 7.1 Limitaciones Actuales

- Muestra temporal:  $N = 3$  años por mes obliga al uso de distribuciones paramétricas estándar. Un horizonte de 5–10 años permitiría el uso de Mixture Models o distribuciones no paramétricas con mayor fidelidad empírica.
- Datos censurados por tipo de negado: el dataset no especifica si un 'negado' fue por falta de stock del proveedor o por error de pronóstico interno en la colocación de la orden de compra. El modelo asume que toda la culpa logística recae en la capacidad del proveedor.
- Ausencia de regresores exógenos: el modelo no incorpora variables macroeconómicas (inflación, tipo de cambio), climáticas (severidad del invierno) ni de marketing (campañas que distorsionen la demanda estacional), los cuales introducirían varianza no capturada en el histórico.
- Capacidad instalada del pool estrella no observable: la simulación de transferencia asume capacidad elástica en PROV33, PROV6 y PROV1, supuesto que requiere validación operativa directa.

### 7.2 Líneas de Mejora Futura

- Incorporar datos de 2026 conforme se generen para actualizar la simulación trimestralmente (pipeline rolling).
- Añadir la dimensión de costo unitario por proveedor para transformar el SS recomendado en unidades a SS recomendado en pesos, habilitando análisis de ROI directo.
- Implementar un modelo de clasificación de causas de negado (falta de stock vs. error de OC) para desagregar la responsabilidad entre proveedor y área de compras.
- Explorar simulación de Monte Carlo con distribuciones de cola pesada (Pareto, Log-Normal) para proveedores con alta varianza y tendencia a picos extremos.

## 8. Glosario de Términos Técnicos

Término	Definición
AIC	Criterio de Información de Akaike. Métrica para comparar modelos estadísticos que penaliza la complejidad.
Datos censurados	Observaciones donde el valor real está truncado por una restricción externa (stock agotado).
Distribución Gamma	Familia de distribuciones continuas, dominio positivo, usada para modelar tiempos y volúmenes con asimetría positiva.
Fulfillment Rate	Porcentaje de la demanda total que fue efectivamente cubierta por el proveedor.
MLE	Estimación por Máxima Verosimilitud. Método para estimar parámetros de distribuciones estadísticas.
Monte Carlo	Método de simulación que genera miles de escenarios aleatorios para cuantificar la incertidumbre de un sistema.

N_sim = 10,000	Número de iteraciones de simulación. Garantiza estabilidad estadística de los percentiles estimados.
Percentil 95 (P95)	El valor que 9,500 de los 10,000 escenarios no superan. Umbral de cobertura para el Stock de Seguridad.
Score de Riesgo	Métrica compuesta (Tasa de Negados + CV ponderados por $\alpha$ ) que clasifica proveedores por riesgo.
Spearman $\rho$	Coeficiente de correlación de rangos. Mide si el orden de riesgo se preserva entre períodos.
Stock de Seguridad	Inventario adicional recomendado para cubrir la variabilidad entre demanda y capacidad del proveedor.
Walk-Forward CV	Esquema de validación cruzada que respeta el orden temporal, entrenando siempre en el pasado.

---

*Todos los gráficos generados por el pipeline (heatmaps, dashboards por proveedor, curvas de cross-validation) se incluyen como Anexos al final del documento.*