

Estadística aplicada a la Ingeniería Logística Digital y Gestión de la Cadena de Suministro

Frida Karem Rivas Moreno
frida.rivas@correo.buap.mx

Al pensar en grandes compañías que entregan sus productos a domicilio es inevitable hacerse las siguientes preguntas:

¿Cómo consigue una compañía entregar millones de paquetes en pocos días? ¿Cómo sabe una tienda cuánto producto solicitar antes de que se agote? ¿Cómo plataformas como Shein, Mercado Libre o FedEx anticipan la demanda y organizan millares de rutas todos los días?

Es inevitable concluir que esto no lo hacen de forma manual, debe existir una herramienta que registre la información y la procese para que los ingenieros de datos o los directores de operaciones puedan utilizarla, y aunque muchas personas consideran que estas capacidades únicamente las poseen las nuevas tecnologías digitales, existe un elemento que pocos observan, pero que es de suma importancia en cualquier nivel: **la estadística**.

En la actualidad, la ingeniería logística digital y la gestión de la cadena de suministro elaboran inmensas cantidades de información debido al uso de tecnologías como el Internet de las Cosas (IoT), sistemas ERP, RFID, plataformas de administración de almacenes (WMS), sistemas de transporte (TMS) y herramientas de análisis de Big Data. No obstante, tener grandes volúmenes de datos no garantiza que se van a tomar mejores decisiones. La ventaja competitiva real proviene cuando estos datos se interpretan de forma adecuada para convertirlos en conocimiento útil.

Los datos cuentan historias

La estadística ayuda a descubrir patrones de comportamiento y decisiones ocultos en la información que diariamente generan las operaciones logísticas.

A través de ella es posible responder a preguntas como:

- ¿Cuál es el tiempo promedio que se necesita para entregar un pedido?
- ¿Qué tan constantes son los tiempos que se utilizan en el transporte?
- ¿Cuáles son los productos que tienen una mayor demanda?
- ¿Cuáles son las rutas que reportan más retrasos?
- ¿Qué proveedores brindan un desempeño más confiable?

Para responder preguntas como estas se utilizan herramientas conocidas como **medidas de tendencia central**, entre ellas la media o promedio, la mediana y la moda. Estas permiten resumir miles e incluso millones de registros en indicadores fáciles de interpretar. Un ejemplo de ello es cuando una empresa tiene la información respecto al tiempo promedio que tarda en entregar un paquete o identificar cuál es el horario en donde se concentran más pedidos.

Sin embargo, saber únicamente el promedio puede resultar un dato realmente engañoso. Dos compañías pueden tener el mismo tiempo promedio de entrega, pero una puede ser más puntual que la otra. Aquí es donde las **medidas de dispersión**, como la desviación estándar y la varianza, sirven para mostrar qué tan estables o variables son los procesos. Mientras menor sea la variabilidad, mayor será la confiabilidad del servicio.

De esta forma, la estadística analiza el modo en que se distribuyen los datos. Este tipo de estudio permite revelar si existen retrasos poco frecuentes pero muy severos, comportamientos atípicos o riesgos que podrían afectar el funcionamiento de toda la cadena de suministro.

Tomar decisiones cuando el futuro es incierto

La logística enfrenta diariamente situaciones impredecibles: variaciones en la demanda, accidentes carreteros, marítimos o aéreos, retrasos aduanales, condiciones climáticas adversas o fallas en los sistemas de transporte.

En este contexto, la **probabilidad** se convierte en una herramienta necesaria para estimar la posibilidad de que ocurran ciertos eventos. Por ejemplo, una empresa puede estimar la probabilidad de que un pedido llegue a tiempo, que un producto no se agote en el inventario o que una entrega sea rechazada por algún cliente.

Además, existen modelos matemáticos que permiten actualizar continuamente estas probabilidades conforme aparece nueva información. Gracias a ello, las organizaciones pueden reaccionar rápidamente ante cambios inesperados y reducir los riesgos operativos.

No todos los datos se comportan de la misma manera

Imagina que la cadena de suministro global es un gigantesco casino. Cada vez que compras un producto por internet, se activa una ruleta con preguntas como: ¿está disponible?, ¿llegará a tiempo?

A diferencia de los juegos de azar, en la logística digital las empresas no entregan su futuro a la suerte. Utilizan “modelos matemáticos predictivos”, es decir, las distribuciones de probabilidad, que permiten adelantarse a los problemas, ahorrar millones de pesos y asegurar que el envío llegue a tu puerta.

Cuando se analizan eventos que se pueden contar solo mediante números enteros, como el número de pedidos recibidos por hora, las fallas de una máquina o los productos defectuosos en un lote, se utilizan **distribuciones discretas**.

La **distribución binomial** es la “moneda” del éxito o fracaso, cuando las cosas se reducen a dos opciones: entregado o no entregado, a tiempo o retrasado, aceptado o rechazado. Cuando un proceso solo presenta dos resultados posibles, los ingenieros de datos usan la probabilidad binomial.

Cuando un simple “sí” o “no” no es suficiente aparece la **probabilidad multinomial**. Este es el caso de una compra en un comercio electrónico, desde que das clic en “comprar” hasta que llega a tus manos, pasa por múltiples estados.

Cuando se debe evaluar la calidad de un producto alimenticio o de fármacos, no se puede

sacar del empaque y analizar uno por uno y luego regresarlos a su envoltura. Al inspeccionar un lote de mercancías, las piezas revisadas se apartan. Esto se conoce como proceso sin reemplazo. Cada pieza que se extrae cambia las reglas del juego para la siguiente. A esto se le conoce como **distribución hipergeométrica**.

El semáforo de las filas de espera es la **distribución de Poisson** la cual modela cuántas cosas ocurren en un periodo determinado, utilizando únicamente un promedio histórico.

Cuando las variables representan mediciones continuas, es decir, que se pueden contar en fracciones o decimales, como el tiempo de entrega, la distancia recorrida o el costo del transporte, se emplean **distribuciones continuas**, donde destaca la distribución normal como la más utilizada debido a que muchos procesos logísticos presentan este comportamiento.

Estos modelos permiten anticipar escenarios futuros y planificar con mayor precisión los recursos obligatorios para mantener un servicio con mayor eficiencia.

La información que arroja una muestra pequeña

Analizar cada uno de los millones de datos que genera una empresa puede resultar prácticamente imposible. Es por ello, que la estadística se apoya del **muestreo**, que implica el estudiar tan solo una parte que sea representativa, de todo el universo de información.

A partir de estas muestras es posible estimar indicadores como la demanda futura, los tiempos promedio de entrega o la calidad de los productos, con un nivel conocido de confianza.

Este enfoque permite reducir costos, ahorrar tiempo y obtener resultados suficientemente precisos para apoyar la toma de decisiones estratégicas.

¿Las mejoras funcionan en realidad?

Existen muchas organizaciones que implementan tecnologías nuevas con la esperanza de que sus resultados mejoren, pero ¿cómo se comprueba que realmente funcionaron?

Las **pruebas de significación estadística** pueden brindar la respuesta a esta incógnita. Estas pruebas

permiten distinguir entre un cambio real y uno que simplemente ocurrió por mera casualidad.

Un ejemplo de ello es cuando después de instalar un nuevo programa de computadora para optimizar las rutas de traslado de los vehículos de transporte, una empresa puede hacer comparaciones entre los tiempos de entrega antes y después de su implementación. Si las diferencias obtenidas son estadísticamente significativas, existe evidencia de que la mejora fue efectiva.

Cuando se desea comparar a varios proveedores, a distintos centros de distribución o a diferentes métodos de transporte, herramientas como el **Análisis de Varianza (ANOVA)** se utilizan para identificar cuál alternativa ofrece mejores resultados.

Los datos ayudan a predecir el futuro

Una de las herramientas más valiosas de la estadística es la elaboración de pronósticos, cuya finalidad es anticipar escenarios futuros y disminuir la incertidumbre. Los pronósticos analizan datos históricos, tendencias y lo que el consumidor prefiere entre cada estación, mediante el uso de modelos matemáticos, con la finalidad de optimizar la planificación de recursos, la gestión de los inventarios y la mejora en la toma de decisiones estratégicas.

Mediante técnicas como la **correlación**, la **regresión** y el análisis de **series de tiempo**, las empresas pueden predecir la demanda, estimar costos de transporte, identificar tendencias de consumo y preparar con anticipación sus operaciones.

La correlación sirve para medir la dirección de la relación lineal entre dos o más variables, cuantificando si al cambiar el valor de una variable, la otra presenta la misma tendencia o por el contrario disminuye de forma predecible. El análisis de regresión se utiliza para modelar la relación entre la variable que queremos medir y los factores que influyen sobre ella. Los modelos de regresión relacionan una variable de interés con factores externos o causales (como descuentos o promociones) para la estimación de diferentes escenarios. A su vez, las series de tiempo analizan el comportamiento histórico a lo largo del tiempo para predecir eventos futuros,

como las ventas semanales. Involucran componentes como la tendencia, los ciclos, las estacionalidades y las irregularidades.

Esto resulta especialmente útil durante temporadas de alta demanda, como el Buen Fin, el *Hot Sale*, Navidad o el regreso a clases, así cuando una planeación se elabora de manera adecuada puede evitar pérdidas económicas ocasionadas por el desabasto o el exceso de inventario.

Asimismo, los **números índices** permiten la comparación del comportamiento de variables como los costos logísticos, el precio de los combustibles o el volumen de inventarios a lo largo del tiempo, haciendo más fácil la evaluación del desempeño organizacional.

Algo más que sólo números

Con mucha frecuencia se cree que la estadística se utiliza únicamente para realizar cálculos complejos o para construir gráficas. No obstante, su verdadero valor radica en ayudar a comprender la realidad y respaldar decisiones con evidencia.

En un ambiente en donde la logística depende cada vez más de la digitalización, la automatización y la inteligencia artificial, la estadística funciona como el puente que conecta los datos con el conocimiento. Debido a ella, las organizaciones pueden reducir sus costos, mejorar el nivel del servicio que ofrecen, optimizar sus inventarios, anticipar los riesgos y robustecer su competitividad.

En otras palabras, detrás de cada entrega puntual, cada ruta optimizada y cada decisión logística basada en datos, existe un conjunto de herramientas estadísticas trabajando silenciosamente para hacer que la cadena de suministro funcione de manera más eficiente.

La logística del futuro no dependerá únicamente de generar más datos, sino de interpretarlos inteligentemente. Y en ese desafío, la estadística seguirá siendo una de las aliadas más poderosas de la ingeniería logística digital.

La habilidad de no asumir la “normalidad”

Ahora imaginemos que lideramos la logística de un gran almacén en línea y los vehículos repartidores se enfrentan a diferentes imprevistos todos los días, como el clima, el tránsito local o en carretera, manifestaciones o retrasos en el almacén. Al intentar obtener el promedio de los tiempos de entrega para hacer predicciones, la estimación resultaría errónea, ¿cuál es la razón? El mundo real es un caos, la información no es perfecta ni simétrica.

En este punto entra al rescate la **estadística no paramétrica**, herramienta flexible que no exige datos que sigan reglas ideales.

En la estadística tradicional, se asume que los datos se comportan de forma “normal” (haciendo referencia a la curva de Gauss, más conocida como curva normal, donde la mayoría de los eventos sucede al centro). Sin embargo, en la logística digital, si 99 llegan puntualmente pero sólo uno se pierde o retrasa, ese único dato rompe el promedio.

La estadística no paramétrica no se deja engañar por esos valores extremos, por lo que, en lugar de promedios, utiliza las medianas y el orden de los datos, mejor conocidos en la jerga estadística como rangos. Esto es como calificar una competencia de clavados: el puntaje exacto que proporcionan los jueces no es el importante, sino quien quedó en primero, segundo o tercer lugar.

Las compañías de logística digital usan estas pruebas matemáticas para la toma de decisiones estratégicas con base en datos reales. Por ejemplo, si se quiere evaluar qué proveedor es más rápido, pero sus tiempos de entrega varían por inclemencias del tiempo, la **prueba de Mann-Whitney U** compara el rendimiento real de los proveedores sin importar los días de tormenta extrema.

Para medir si un algoritmo nuevo reduce los tiempos de traslado, se emplea la **prueba de Wilcoxon**. Herramienta que compara el antes y el después de los mismos conductores, esta herramienta es ideal para datos con subidas y bajadas muy drásticas.

Cuando el éxito de un producto depende de variables que no son números exactos, como el nivel de satisfacción de los clientes o la intensidad de una campaña en las redes sociales, las **correlaciones de Spearman o Kendall** descubren qué estrategia mueve realmente el indicador de la demanda.

Otro caso muy común lo presentan las encuestas de satisfacción donde se evalúa un producto o servicio, con indicadores como “excelente”, “bueno”, “regular”, “malo”, que no se pueden sumar ni promediar. Para ello sale al rescate la **prueba de Friedman**, la cual ayuda a los analistas a descifrar analíticamente en qué parte de la entrega se pierde la lealtad del usuario.

REFERENCIAS

- Aguilar, J. (2025). Estadística inferencial: Regresión y correlación. Series de tiempo–Gráficos de control. Ediciones de la U.
- Ayyub, B. M., & McCuen, R. H. (2025). Probability, statistics, and reliability for engineers and scientists. CRC press.
- Durivage, M. A. (2022). Practical engineering, process, and reliability statistics. Quality Press.
- Escudero, R. E. S. (2022). Logística estadística. Gestión e indicadores en la cadena de suministro-1ra edición: Gestión e indicadores en la cadena de suministro. Ecoe Ediciones.
- Karyshev, M. Y. (2023). To the Question of the Relevance of Statistics of Information and Communication Technologies in the Context of the Digital Transformation of the Economy. Statistics and Economics, 20(1), 53-63. <https://doi.org/10.21686/2500-3925-2023-1-53-63>
- Pham, H. (Ed.). (2023). Springer handbook of engineering statistics. Springer Nature.
- Ross, S. M. (2020). Introduction to probability and statistics for engineers and scientists. Academic press.
- Schiefer, H., & Schiefer, F. (2021). Statistics for engineers. Wiesbaden, Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-32397-4>.
- Siregar, T. (2025). Descriptive statistics: Concepts, data collection methods, and sampling techniques in quantitative research. Jurnal Aplikasi Matematika Dan Statistik.

