# TEORÍA DE LA PROBABILIDAD

## Procesos Estocásticos

### Desarrollo histórico

El estudio de los procesos estocásticos se inició a principios del siglo XX, en forma paralela a la conclusión del estudio de las sucesiones de variables aleatorias independientes. Sus principales precursores fueron Andrei Andreyevich Markov, A. N. Kolmogorov, A.Ya. Khintchine, N. Wiener y P. Lévy.

Desde la publicación del teorema de Bernoulli, en 1713, el motor de desarrollo de la teoría de la probabilidad fue la búsqueda de resultados que permitieran mejorar y generalizar ese teorema. Ese proceso culminó en la década de los 30's del siglo XX con la formulación acabada de los teoremas límite para sucesiones de variables aleatorias independientes.

El teorema de Bernoulli establece que si $\mathcal{E}$ es un experimento aleatorio y $A$ un evento relativo a ese experimento, de probabilidad igual a $p$, y consideramos un nuevo experimento aleatorio consistente en la repetición indefinida del experimento $\mathcal{E}$, de tal manera que cada repetición es independiente de las otras, entonces, llamando $Y_n$ al número de veces que ocurre el evento $A$ en las primeras $n$ repeticiones del experimento, la sucesión de variables aleatorias $\left(\frac{Y_n}{n}\right)_{n\in\mathbb{N}}$ converge a $p$ en probabilidad. Formulado de otra manera, el resultado dice que si $(X_n)_{n\in\mathbb{N}}$ es una sucesión de variables aleatorias independientes, todas con distribución Bernoulli de parámetro $p$, entonces la sucesión $\left(\frac{X_1+X_2+\cdots+X_n}{n}\right)_{n\in\mathbb{N}}$ converge a $p$ en probabilidad.

A la forma general del resultado de Bernoulli se le conoce como ley débil de los grandes números. En una primera versión, el resultado general fue demostrado por Chebyshev, iniciador de la llamada "escuela rusa"en la teoría de la probabilidad. En el año 1867 Chebyshev demostró (utilizando la que ahora se conoce como "desigualdad de Chebyshev") que si $(X_n)_{n\in\mathbb{N}}$ es una sucesión de variables aleatorias discretas, independientes, idénticamente distribuidas y de varianza finita, entonces la sucesión $\left(\frac{X_1+X_2+\cdots+X_n}{n}\right)_{n\in\mathbb{N}}$ converge en probabilidad a la esperanza común de las variables aleatorias $X_n$.

En el año 1928, el resultado de Chebyshev fue mejorado por Khintchine, quien logró demostrar la ley débil sin la condición de que las variables aleatorias $X_n$ tengan varianza finita. Para esto utilizó lo que ahora se conoce como el "método de truncación", el cual había sido introducido por Markov en el año 1913. Este método resultó muy fructífero para el estudio de los teoremas límite y se extendió después, con nuevos elementos, para el estudio de los procesos estocásticos. Con el método de truncación se puede reducir un problema donde las varianzas pueden ser infinitas a uno donde éstas sean finitas y está basado en el siguiente resultado: Si $(X_n)_{n\in\mathbb{N}}$ es una sucesión de variables aleatorias independientes, idénticamente distribuidas y de esperanza finita $\mu$, $(a_n)_{n\in\mathbb{N}}$ una sucesión creciente de números reales positivos tal que $\lim_{n\to\infty} a_n = \infty$ y, para $n, k \in \mathbb{N}$, definimos:

$$Y_k^n = \begin{cases} X_k & \text{si } |X_k| \leq a_n \\ 0 & \text{en otro caso} \end{cases}$$

Entonces, fijando $n$, las variables aleatorias $Y_1^n, Y_2^n, \ldots$ tienen la misma distribución y varianza finita. Además,si $\mu_n$ es la esperanza común de $Y_1^n, Y_2^n, \ldots$, entonces $\lim_{n\to\infty} \mu_n = \mu$.

De igual manera, en ese periodo se estableció la ley fuerte de los grandes números a partir del resultado de Borel del año 1909, siendo Kolmogorov quien obtuvo el resultado general en el año 1930.

El teorema central del límite, demostrado en el año 1733 por Abraham de Moivre para el caso particular de una sucesión de variables aleatorias independientes, todas con distribución Bernoulli de igual parámetro, fue demostrado de manera general por Lindeberg en el año 1922.

Dentro de este contexto del estudio de los teoremas límite, Markov introdujo en 1906 una nueva línea de investigación al demostrar la ley débil de los grandes números para una sucesión de variables aleatorias no independientes, sino, de acuerdo con el término que él utilizó, encadenadas. Surgió así un tipo de proceso estocástico que ahora es llamado de Markov. De manera general e informal, se dice que un proceso estocástico es de Markov si dado cualquier tiempo $t \in \Gamma$, las familias de variables aleatorias $\{X_s : s < t\}$ y $\{X_s : s > t\}$ son independientes, propiedad que en ocasiones se enuncia diciendo que, dado el presente, el pasado y el futuro son independientes. Por esta propiedad, también se dice que un proceso de este tipo es un proceso sin memoria.

Markov introdujo lo que ahora llamamos una cadena de Markov en tiempo discreto, la cual está formada por una familia de variables aleatorias $\{X_n : n \in \mathbb{Z}^+\}$ las cuales toman valores en un conjunto finito o infinito numerable $D \subset \mathbb{R}$, es decir, son variables aleatorias discretas, y satisfacen la siguiente propiedad, para cualquier $x \in D$ y cualquier $n \in \mathbb{Z}^+$:

$$P\left[X_{n+1} = x \mid X_{0,}X_1, \ldots, X_n\right] = P\left[X_{n+1} = x \mid X_n\right]$$

Un gran avance en el estudio de los procesos estocásticos lo hizo Wiener entre 1921 y 1923 al construir un modelo matemático del movimiento browniano, el cual consiste en el movimiento de un grano de polen que se coloca sobre agua. En el año 1827, al estudiar el proceso de fertilización de las flores de varias plantas, Robert Brown observó que los granos de polen se movían: “al examinar la forma de estas partículas inmersas en agua, observé muchas de ellas muy evidentemente en movimiento; éste consistía no solamente en un cambio de lugar en el fluido, manifestado por alteraciones en sus posiciones relativas, sino que también, con no poca frecuencia, por un cambio en la forma de la partícula misma ... Estos movimientos eran tales que me convencieron, después de observaciones repetidas frecuentemente, de que no surgían de corrientes en el fluido, ni de su gradual evaporación, sino que pertenecían a la partícula misma." Este movimiento ya había sido observado por otros investigadores anteriormente a Brown, sin embargo fue él quien realizó un estudio más detallado, aunque no completo del mismo. En su honor se le conoce como movimiento browniano.

Recordemos que también en los primeros 30 años del siglo XX se desarrolló la teoría de la medida a partir de los trabajos de Borel y Lebesgue de principios de siglo. Paralelamente, esta teoría comenzó a utilizarse en el Cálculo de Probabilidades y poco a poco se fue convirtiendo en la base de la Teoría de la Probabilidad, culminando este proceso con el trabajo de Kolmogorov de 1933, donde se fundamenta el considerar a la probabilidad como una medida.

Mientras tanto, la escuela rusa, que surgió con los trabajos de P. L. Chebyshev, realizó avances significativos generalizando el trabajo de Markov. Entre 1931 y 1937, Kolmogorov realizó un estudio a fondo de las cadenas de Markov en tiempo discreto y en tiempo continuo; también estudio los procesos de difusión, los cuales son procesos de Markov con espacios de estados continuos. En esos trabajos introdujo lo que ahora se conoce como la ecuación de Chapman-Kolmogorov y las ecuaciones backward y forward que satisfacen los procesos de difusión. Por su parte, en el año 1934, Khintchine introdujo un tipo de procesos que pueden ser no markovianos: los procesos estacionarios.

Otro gran avance en el desarrollo de la teoría de los procesos estocásticos, lo realizó Paul Lévy en su libro *Théorie de l'addition des variables aléatoires*, cuya primera edición se publicó en 1937. Ahí planteó que una extensión natural de las sumas o series de variables aleatorias independientes es el siguiente: "En lugar de un índice entero $n$, introduciremos un parámetro $t$ que varía de una manera continua; las sumas sucesivas $S_n$ (se refiere a $S_n = \sum_{k=1}^{n} X_k$, donde $X_1, X_2, \ldots, X_n$ son variables aleatorias independientes) serán entonces reemplazadas por una función aleatoria $X(t)$ de ese parámetro, y la condición de independencia de los diferentes elementos cuya suma es $X(t)$ debe expresarse como sigue: si $t_0 < t_1$, el crecimiento $X(t_1) - X(t_0)$ es una variable aleatoria independiente de $X(t_0)$ y del conjunto de valores de $X(t)$ para $t \leq t_0$."Tomando $t \in \mathbb{R}^+$, lo anterior implica que si $t_1, t_2, \ldots, t_n$ son números reales no negativos tales que $t_1 < t_2 < \ldots < t_n$, entonces las variables aleatorias $X(t_2) - X(t_1)$, $X(t_3) - X(t_2)$, $\ldots$, $X(t_n) - X(t_{n-1})$ son independientes.

Lévy estaba definiendo ahí lo que se conoce ahora como un proceso con incrementos independientes. Aclaraba que algunos casos particulares de ese tipo de funciones aleatorias habían sido consideradas previamente por Bachelier, en 1913, Wiener, en 1923 y Kolmogorov, en 1932.

Obsérvese que si $(Y_n)_{n\in\mathbb{Z}^+}$ es un proceso con incrementos independientes, entonces, definiendo, para $k \in \mathbb{N}$, $X_n = Y_n - Y_{n-1}$, $(X_n)_{n\in\mathbb{N}}$ es una sucesión de variables aleatorias independientes y, para cualquier $n \in \mathbb{N}$, se tiene: $Y_n - Y_0 = \sum_{k=1}^{n} X_k$. De aquí que, para estudiar el proceso $(Y_n)_{n\in\mathbb{Z}^+}$, se puedan utilizar los resultados que se tienen para las sumas de variables aleatorias independientes.

El estudio que hizo Lévy para este tipo de procesos resultó ser de gran importancia en la teoría de los procesos estocásticos. Para enunciar algunos de sus resultados, requerimos de las siguientes definiciones:

Asumimos que tenemos definido un espacio de probabilidad $(\Omega, \Im, P)$.

**Definición 1.** *Diremos que una variable aleatoria $X$ es infinitamente divisible si, para cualquier $n \in \mathbb{N}$, existen $n$ variables aleatorias, $X_1, X_2, \ldots, X_n$, independientes e idénticamente distribuidas tales que $X = \sum_{k=1}^{n} X_k$.*

**Definición 2.** *Diremos que un proceso $(X_t)_{t\in\mathbb{R}^+}$ es fuertemente continuo si existe un conjunto $C \in \Im$ de probabilidad $1$, tal que, para cualquier $\omega \in C$, la función $t \to X_t(\omega)$, definida sobre $\mathbb{R}^+$, es continua.*

**Definición 3.** *Diremos que un proceso $(X_t)_{t\in\mathbb{R}^+}$ es débilmente continuo si:*

*i) Para cada $t_0 \in \mathbb{R}^+$, existe un conjunto $C_{t_0} \in \Im$, de probabilidad $1$, tal que, para cualquier $\omega \in C_{t_0}$, la función $t \to X_t(\omega)$, definida sobre $\mathbb{R}^+$, es continua en $t = t_0$.*

*ii) Existe un conjunto $C \in \Im$, de probabilidad $1$, tal que, para cualquier $\omega \in C$, la función $t \to X_t(\omega)$, definida sobre $\mathbb{R}^+$, tiene límites por la derecha y por la izquierda en cualquier punto.*

**Definición 4.** *Sea $(X_t)_{t\in\mathbb{R}^+}$ un proceso estocástico. Diremos que $t_0 \in \mathbb{R}$ es un punto de discontinuidad fijo de $(X_t)_{t\in\mathbb{R}^+}$ si existe un conjunto $D_{t_0} \in \Im$, de probabilidad positiva, tal que, para cualquier $\omega \in D_{t_0}$, la función $t \to X_t(\omega)$, definida sobre $\mathbb{R}^+$, es discontinua en $t = t_0$.*

Algunos de los resultados más importantes que demostró Lévy son los siguientes:

**Teorema 1.** *Si $(X_t)_{t\in\mathbb{R}^+}$ es un proceso con incrementos independientes y débilmente continuo, entonces, para cualesquiera $s, t \in \mathbb{R}^+$ tales que $s < t$, la variable aleatoria $X_t - X_s$ es infinitamente divisible.*

**Teorema 2.** *Si $(X_t)_{t\in\mathbb{R}^+}$ es un proceso con incrementos independientes, entonces $(X_t)_{t\in\mathbb{R}^+}$ es fuertemente continuo si y sólo si, para cualesquiera $s, t \in \mathbb{R}^+$ tales que $s < t$, la variable aleatoria $X_t - X_s$ tiene distribución normal.*

**Teorema 3.** *Si* $(X_t)_{t\in\mathbb{R}^+}$ *es un proceso con incrementos independientes, entonces existe una función* $f:\mathbb{R}^+\to\mathbb{R}$ *tal que el conjunto de discontinuidades fijas del proceso* $(X_t - f(t))_{t\in\mathbb{R}^+}$ *es a lo más infinito numerable.*

**Teorema 4.** *Todo proceso con incrementos independientes* $(X_t)_{t\in\mathbb{R}^+}$ *se puede descomponer de la siguiente manera:*

$$X_t = f(t) + X_t^{(1)} + X_t^{(2)} + X_t^{(3)}$$

*donde:*

*i)* $f:\mathbb{R}^+\to\mathbb{R}$ *es una función tal que el conjunto de discontinuidades fijas del proceso* $(X_t - f(t))_{t\in\mathbb{R}^+}$ *es a lo más infinito numerable.*

*ii)* $\left(X_t^{(1)}\right)_{t\in\mathbb{R}^+}$ *es un proceso con incrementos independientes para el cual existe un conjunto* $D\subset\mathbb{R}^+$*, finito o infinito numerable, y un conjunto* $C\in\Im$*, de probabilidad* $1$*, tales que, para cualquier* $\omega\in C$*, la función* $t\to X_t(\omega)$*, definida sobre* $\mathbb{R}^+$*, es continua en cualquier* $t\in D^c$*.*

*iii)* $\left(X_t^{(2)}\right)_{t\in\mathbb{R}^+}$ *es un proceso con incrementos independientes fuertemente continuo.*

*iv)* $\left(X_t^{(3)}\right)_{t\in\mathbb{R}^+}$ *es un proceso con incrementos independientes débilmente continuo para el cual existe un conjunto* $C\in\Im$*, de probabilidad* $1$*, tal que, para cualquier* $\omega\in C$*, la función* $t\to X_t(\omega)$*, definida sobre* $\mathbb{R}^+$*, crece o decrece únicamente mediante saltos.*

Los ejemplos básicos, de procesos con incrementos independientes, que menciona Lévy en su libro de 1937 son el proceso de Wiener, el cual es fuertemente continuo, y el proceso de Poisson, el cual es débilmente continuo, sin discontinuidades fijas:

En ese libro, mostró que las propiedades que caracterizan al proceso de Wiener, $(W_t)_{t\in\mathbb{R}^+}$, son las siguientes:

i) $W_0 = 0$.

ii) Si $0 < t_1 < \cdots < t_n$, entonces las variables aleatorias $W_{t_1}$, $W_{t_2} - W_{t_1}$, …, $W_{t_n} - W_{t_{n-1}}$ son independientes.

iii) Si $0 \le s < t$, entonces la variable aleatoria $W_t - W_s$ tiene distribución normal de parámetros $\mu = 0$ y $\sigma^2 = t - s$.

iv) $(W_t)_{t\in\mathbb{R}^+}$ es un proceso fuertemente continuo.

Los resultados de la teoría de los procesos estocásticos obtenidos durante la primera mitad del siglo XX fueron recogidos, sistematizados y complementados con nuevos resultados, en dos libros que marcaron la pauta para continuar desarrollando esta teoría por lo menos durante los 30 años que les siguieron: El libro de Paul Lévy *Processus stochastiques et mouvement brownien*, publicado en 1948 y el libro de Joseph Leo Doob *Stochastic processes*, publicado en 1953.

Lévy comienza su libro dando una nueva construcción del proceso de Wiener y pasa después a considerar el problema general de la construcción de procesos estocásticos. En los capítulos siguientes expone la teoría de los procesos de Markov y de los procesos estacionarios. Después expone una versión más completa de su teoría de los procesos con incrementos independientes, llamados ahí procesos aditivos. Tal vez la parte más importante de su libro es el estudio a fondo que hace del movimiento browniano, tanto en la recta real, como en el plano y en otros espacios.

El libro de J. L. Doob fue fuente de inspiración para muchos matemáticos ya que ahí plantea varias líneas de investigación para diferentes problemas. Doob sigue lo que podríamos llamar el enfoque moderno: se parte de los axiomas de Kolmogorov y la construcción de procesos se realiza vía el teorema de consistencia de Kolmogorov. Comienza su libro estudiando propiedades generales de los procesos estocásticos, en particular, la existencia de versiones de un proceso con trayectorias que son regulares en algún sentido. Realiza un estudio a fondo de los procesos de Markov en tiempo discreto y en tiempo continuo. Trata también el estudio de los procesos con incrementos independientes y de los procesos estacionarios. Tal vez las principales aportaciones de Doob en ese libro son, por una parte, el estudio que hace de los procesos conocidos como martingalas y, por otra parte, la invención del concepto de tiempo de paro. Paul Meyer, en su libro *Probabilités et Potentiel* (1975) comenta: "Il a sans doute fallu autant de génie aux créateurs du calcul infinitésimal pour expliciter la notion si simple de derivée, qu'à leurs successeurs pour faire toute le reste. L'invention des temps d'arrêt par Doob est tout a fait comparable" ("Sin duda, se requirió tanta genialidad de los creadores del calculo infinitesimal para explicitar la noción tan simple de derivada, que de sus sucesores para hacer el resto. La invención de tiempos de paro por Doob es completamente comparable").

Es interesesante la diferencia que existió entre Paul Lévy y J. L. Dob en cuanto a la forma en que pensaban un proceso estocástico. El libro de Doob se publicó en 1954 y en 1955 se publicó una segunda edición del libro de Lévy *Théorie de l'addition des variables aléatoires*, de manera que Lévy pudo leer el libro de Doob antes de que se imprimiera esta segunda edición y, después de hacerlo, decidió agregar una nota al final de su libro. En ella, dice: "Un proceso estocástico es en principio un fenómeno en cuya evolución el azar interviene a cada instante. Para Doob, un proceso estocástico es simplemente una función aleatoria $X(t)$ de una variable $t$ la cual puede uno imaginar que representa el tiempo. Cualquiera que sea el conjunto $\mathfrak{E}$ de determinaciones posibles de $X(t)$, se puede asociar a cada elemento de $\mathfrak{E}$ un valor de una variable simbólica $\omega$; eso le permite a Doob considerar $X(t)$ como una función cierta $x_t(\omega)$ de $t$ y de una variable aleatoria $\omega$ cuya elección resume todas las intervenciones sucesivas del azar. Es una notación a menudo cómoda, aunque dé como un todo, nacido en un instante, lo que, para mí, es esencialmente un perpetuo devenir."

La teoría que más se desarrolló en los años posteriores a la publicación de los libros de Lévy y de Doob fue la de los procesos de Markov, en particular, la de los procesos de difusión. Uno de los problemas centrales que se atacó fue el de la construcción de procesos con trayectorias regulares en algún sentido.

La escuela rusa, con Kolmogorov a la cabeza, hizo las más importantes contribuciones, desarrollando ampliamente la teoría de los procesos de Markov. Uno de los más destacados en este periodo fue E. B. Dynkin, quien vinculó la teoría de los semigrupos de operadores con la de los procesos de Markov.

No se puede dejar de mencionar a G. A. Hunt por sus trabajos en la teoría de los procesos de Markov y su relación con la teoría del potencial y a K. L. Chung por sus aportaciones a la teoría de los procesos de Markov y por su incansable labor de difusión de la teoría de la probabilidad y de los procesos estocásticos en varios libros.

En los 60's comenzó a desarrollarse la escuela francesa, con Paul Meyer a la cabeza. En el año 1966 se publicó el primero de sus libros, titulado *Probabilités et Potentiel*, en el cual inició el desarrollo de lo que se llamaría después la teoría general de los procesos estocásticos. En 1967 se publicó otro libro suyo, titulado *Processus de Markov*, en el cual presenta algunos de los capítulos que incluiría la segunda parte de su primer libro. Sin embargo, el rumbo de las investigaciones cambió y no hubo tal segunda parte; en su lugar, en la segunda edición de su primer libro, publicada en 1975 y teniendo como coautor a Claude Dellacherie, hubo una amplia reformulación, a tal grado que más bien esa segunda edición puede verse como un libro distinto. De esa segunda edición sí hubo una segunda parte, titulada *Théorie des martingales*.

En su libro de 1967 Meyer expuso la teoría general de los procesos de Markov en tiempo continuo, sistematizando gran parte de los resultados que se tenían hasta ese momento.

Como lo mencionamos antes, uno de los problemas en la teoría de los procesos estocásticos es el de su construcción. En el caso de los procesos de Markov, el camino que se siguió fue el de construirlos a partir del generador infinitesimal de un semigrupo de operadores, el cual consiste de una familia de operadores $\{P_t : b(\mathbb{R},\mathbb{R}) \to b(\mathbb{R},\mathbb{R})\}$, donde $t \in \mathbb{R}^+$ y $b(\mathbb{R},\mathbb{R})$ es el conjunto de funciones $f : \mathbb{R} \to \mathbb{R}$ que son medibles y acotadas, tomando en $\mathbb{R}$ la $\sigma$-álgebra de Borel $\mathcal{B}(\mathbb{R})$.

**Definición 5.** *Se dice que esa familia de operadores forma un semigrupo si, para cualquier pareja de números reales no negativos $s$ y $t$, se satisface la siguiente propiedad:*

$$P_{t+s} = P_t \circ P_s$$

El problema que se presentaba entonces era el de determinar bajo que condiciones sobre el semigrupo de operadores se puede construir un proceso de Markov con trayectorias suficientemente regulares que nos permitan trabajar más fácilmente con el proceso. Este camino se siguió durante varios años; sin embargo, al desarrollarse el Cálculo Estocástico en su forma general, este camino quedó medio abandonado ya que se pueden construir procesos como soluciones de ecuaciones diferenciales estocásticas.

Son interesantes los comentarios de Meyer y Dellacherie en el prefacio de la segunda edición de *Probabilités et Potentiel.* Para empezar, hacen la aclaración de que, a pesar del título, el libro contiene poco de probabilidad y nada de teoría del potencial. La razón para ir en otra dirección fue, como ellos lo dicen, la evolución tan rápida de la teoría, con lo cual se crearon nuevos conceptos y algunas cosas que parecían muy importantes antes, ya no lo eran tanto, mientras que aspectos que parecían sin interés, se volvieron importantes. Como enseñanza, dicen los autores: "Hemos tratado en particular de liberarnos de la actitud consistente en creer que un texto matemático presenta, para nuestra contemplación, verdades eternas, salidas del mundo de las ideas, y cuyo valor no puede sufrir alguna inflación... Pero las verdades existen no por el hecho que se impriman sobre buen papel, sino por el interés que se les da. Muchas verdades inmutables de 1966, están muertas, mientras que pequeños comentarios de 1966 aclaran ahora grandes partes de la teoría."

Esto es un ejemplo de cómo la matemática se va creando como producto social y no mediante "descubrimientos" de "verdades" que se encuentran escondidas en alguna parte, las cuales hay que ir encontrando.

En su libro *Processus stochastiques et mouvement brownien* (1948), Lévy trató el problema de la construcción de procesos mediante la solución de ecuaciones diferenciales estocásticas. La idea que formuló es que el incremento infinitesimal de un proceso $(X_t)_{t\in\mathbb{R}^+}$ en un tiempo $t$, lo cual denotaba por $dX(t)$ podría estar dado por una función que depende de $t$, $dt$, $X(t)$ y una o varias variables aleatorias de tipo familiar, como, por ejemplo, variables aleatorias independientes con distribución normal. En el caso del proceso de Wiener $(W_t)_{t\in\mathbb{R}^+}$, como el incremento $W_t - W_s$, para $s, t \in t \in \mathbb{R}^+$ tales que $s < t$, tiene distribución normal con esperanza cero y varianza $t-s$, entonces en un incremento $dt$ del tiempo, el incremento $dW(t)$ tiene distribución normal con esperanza cero y varianza $dt$, de manera que, denotando por $\xi$ a una variable aleatoria con distribución normal estándar, el proceso de Wiener puede definirse por la ecuación diferencial estocástica $dW(t) = \xi\sqrt{dt}$.

Analizó algunos casos particulares de ecuaciones diferenciales estocásticas y planteó el problema general de resolver una ecuación diferencial estocástica del tipo siguiente:

$$dX(t) = f(t, X, Y, Z)\, dY(t) + g(t, X, Y, Z)\, dZ(t)$$

donde $(Y_t)_{t\in\mathbb{R}^+}$ y $(Z_t)_{t\in\mathbb{R}^+}$ son procesos estocásticos definidos previamente.

Hacía énfasis en que hay una dificultad para resolver una ecuación de este tipo cuando $Y(t)$ y $Z(t)$ no son funciones de variación acotada, por ejemplo cuando $Y$ y $Z$ representan las coordenadas de un movimiento browniano en el plano. Mostró entonces que una integral del tipo $\int_{t_0}^{\tau} f(t, X, Y, Z)\, dY(t)$, a la cual denominaba integral estocástica, puede evaluarse no siguiendo el método clásico de la integral de Riemann, sino considerando números $\tau_1, \tau_2, \ldots$, elegidos al azar en el intervalo $(t_0, \tau)$ y tomando después, para cada $n \in \mathbb{N}$, los primeros $n$ de esos números, ordenados del menor al mayor, para formar con ellos una suma de tipo Riemann. Consideró el caso del cálculo de áreas estocásticas, es decir, áreas de curvas definidas por un proceso estocástico y mostró que las integrales estocásticas que resultan quedan bien definidas con el método mencionado.

Siguiendo un trabajo de S. Bernstein, consideró los procesos de Markov que son solución a ecuaciones diferenciales estocásticas de la forma:

$$dX(t) = A[t, X(t)]\, dt + B[t, X(t)]\ \xi\sqrt{dt}$$

donde $\xi$ es una variable aleatoria auxiliar que cumple con determinadas condiciones.

Obsérvese que en el caso en que $\xi$ sea una variable aleatoria con distribución normal estándar, la ecuación anterior puede escribirse en la siguiente forma:

$$dX(t) = A[t, X(t)]\, dt + B[t, X(ty)]\ dW(t)$$

Planteó entonces el problema general de la determinación de todos los procesos de Markov como soluciones de ecuaciones diferenciales estocásticas. Señalaba que ese problema no estaba resuelto de manera general, pero dió la solución para el caso de los procesos con incrementos independientes.

Lévy no hizo referencia en su libro al trabajo de K. Ito, quien en el año 1944 publicó un artículo titulado *Stochastic Integral*, en el cual formuló una definición de una integral de la forma $\int_0^t X_s dW_s$ donde $(X_t)_{t\in\mathbb{R}^+}$ es un proceso estocástico y $(W_t)_{t\in\mathbb{R}^+}$ es un proceso de Wiener.

En su forma general, la teoría de integración estocástica fue desarrollada por la escuela francesa, con Paul Meyer a la cabeza. Ésta quedó plasmada en los dos libros mencionados con anterioridad: 1. *Probabilités et Potentiel* (1966), en el cual inició el desarrollo de lo que se llamaría después la teoría general de los procesos estocásticos; 2. *Théorie des martingales* (1975).

Hacia finales de los70's del siglo pasado la teoría general de integración estocástica, lo que ahora se puede denominar como Cálculo Estocástico Clásico, quedó prácticamente concluída. Después esa teoría se fue ampliando y se desarrolló en varias direcciones.

Ahora bien, al igual que ocurrió con el Cálculo Diferencial e Integral que conocemos, la teoría de Integración Estocástica se desarrolló para poderla aplicar en la solución de cierto tipo de problemas de interés. De manera específica, el objetivo es poder resolver las llamadas ecuaciones diferenciales estocásticas, las cuales también son en realidad ecuaciones integrales estocásticas. En el Cálculo Estocástico Clásico no hay derivadas; es decir, no hay un Cálculo Diferencial Estocástico ya que, en general, se trabaja con funciones que no son diferenciables. A su vez, el resolver ecuaciones diferenciales estocásticas tiene por objetivo el poder construir procesos estocásticos que modelen fenómenos de interés en diversas áreas. Una de las aplicaciones más desarrolladas es la referente a las Finanzas, pero actualmente las aplicaciones están más diversificadas. Por otra parte, si bien, como se mencionó antes, la teoría del Cálculo Estocástico Clásico quedó bien establecida en los 70's del siglo pasado, no fue así en lo que se refiere a las ecuaciones diferenciales estocásticas, tema que actualmente aún sigue en desarrollo.