UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE CIENCIAS

SOBRE LA CONVERGENCIA DE LOS PROCESOS ESTOCASTICOS

T E S I S

QUE PARA OBTENER EL TITULO DE

A C T U A R I O

P R E S E N T A

LUIS I. ISLAS TAPIA

México, D.F. 1989

INDICE .-

INTRODUCCION .-

En Teoria de la Probabilidad se estudia, entre otras cosas, el concepto de "Convergencia. En el trabajo presente se hace mención de la así llamada "Convergencia Débil". Dicho concepto es la interpretación de la convergencia de una sucesión de Funciones de Distribución de Variables Aleatorias a la Función de Distribución de cierta Variable Aleatoria; sin embargo, a nivel Licenciatura no se estudia con detalle y es por esto que nos decidimos a presentarlo.

La forma de presentar el concepto de Convergencia Débil será estudiando los Procesos Estocásticos, ya que se cuenta ahora con una interpretación aceptada en forma general.

En 1956, Yuri V. Prokhorov, en el artículo "Convergencia de Procesos Estocásticos y Teoremas Límite en Teoria de Probabilidad" expone la interpretación con la cual se estudia actualmente (puede considerarse el punto de partida de lo que hoy llamamos "Teoria de la Invarianza" y ha dado lugar a diversos textos clásicos).

La Teoria General de los Procesos Estocásticos se considera como la Teoria de las Medidas de Probabilidad en Espacios Métricos Separables Completos logrando así, un conjunto de Teoremas que pueden tomarse como la generalización de los Teoremas Límite estudiados en Probabilidad.

Es importante destacar los antecedentes del artículo de Prokhorov, en los cuales se trabajó siguendo la línea mencionada en el párrafo anterior. A continuación reseñamos dos de dichos artículos :

[3] Donsker, 1951 :

En la introducción se mencina al espacio de Wiener (C, W) donde C es el espacio de las funciones continuas definidas en el intervalo $[0,1]$, y W es la medida de Wiener impuesta en C . (Ver apéndice [E]).

El propósito del artículo es mostrar que si $\{X_n\}$ es una familia de variables aleatorías independientes e idénticamente distribuidas con media cero y varianza uno, tales que X_n es una función de $S_1, S_2, \ldots$ ($S_1, S_2, \ldots$ son sumas parciales de variables aleatorias) entonces, bajo restricciones sumamente débiles la distribución límite de Y_n es independiente de la distribución de las variables aleatorias que definen a S_n.

[12] MARUYAMA, 1955 :

Se mencionan las condiciones dadas por Kolmogorov y Feller para que un Proceso Estocástico sea de trayectorias continuas. (Ver apéndice [E]).

Como una aplicación se estudia el problema planteado por S. Bernstein cuya solución se basa en el método de Ito de la Teoría de la Integral Estocástica con respecto a los Procesos Estocásticos fundamentales : El Movimiento Browniano y El Proceso Poisson.

Es interesante observar que al tratar la convergencia de Procesos Estocásticos se utiliza la convergencia en L^1 y la distancia llamada de Fréchet; esto implica el uso de un Espacio Métrico. Al dar ejemplos concretos de Convergencia de los Procesos Estocásticos se emplean las trayectorias de tales Procesos y se verifica la convergencia de las distribuciones, es por esta razón que se emplea al Espacio de las Funciones y se introduce una métrica para cada caso.

Como se puede apreciar, las definiciones de Convergencia de los Procesos Estocásticos son distintas en cada uno de los artículos mencionados; sin embargo el Análisis muestra que en todos los casos ésta convergencia puede interpretarse como "Convergencia Débil" de las distribuciones en un espacio de funciones seleccionado debidamente.

Cabe resaltar que los problemas técnicos son muchos, y que precisar y construir ha dado lugar a la Teoria de la Convergencia Débil de Medidas en Espacios Métricos.

El objetivo del trabajo es ofrecer la Teoria autocontenida; es decir, los resultados fundamentales del artículo de Prokhorov. Motivado por el hecho de que apesar de su importancia y de la trascendencia que ha tenido en el medio matemático, no hay material que se adapte (no solo en cuanto al idioma) en el sentido de facilitar que un mayor número de estudiantes tengan acceso a él a través de una redacción que incorpore técnicas de desarrollo más claras y recientes, logrando una exposición de los temas y resultados en la dirección de la versión de Prokhorov lo más inmediatamente posible.

Este trabajo se divide en tres capítulos con lo cual se cubre la Teoria general y se estudia una de las aplicaciones más interesantes. Se cuenta además con cinco apéndices que tratan en forma separada los temas necesarios para la comprensión de lo expuesto en los capítulos reseñados a continuación :

CAPITULO 1.-

Se introduce el comcepto de Medida y la definición de Convergencia Débil de medidas, dando además, teoremas que permiten establecer criterios de convergencia débil. Para terminar se define al espacio $\mathfrak{D}(\mathfrak{R})$ por medio de la definición de una métrica especial y se dan las condiciones para que un subconjunto de dicho espacio sea compacto.

CAPITULO 2.-

Se expone el Teorema de Prokhorov y se trabaja con el espacio de las funciones continuas . Además se introducen las funciones de distribución finito dimensionales en el espacio enunciando y demostrando el Teorema central de éste trabajo.

CAPITULO 3.-

Se da un ejemplo, que como ya se mencionó, es uno de los más importantes de la Teoría desarrollada.

CAPITULO 1.- MEDIDAS, CONVERGENCIA.

SECCION 1.1.- CONVERGENCIA DEBIL DE MEDIDAS.

En lo que sigue, denotaremos a un espacio métrico separable completo, mediante el símbolo $\mathcal{R}$; y con $C(\mathcal{R})$ denotaremos a la totalidad de funciones continuas y acotadas sobre $\mathcal{R}$.

DEFINICION 1.1.1.- Llamaremos MEDIDA en $\mathcal{R}$ a cualquier función de conjuntos μ cuando :

(μ.1) Todos los conjuntos cerrados del espacio $\mathcal{R}$ están contenidos en $\mathfrak{M}_\mu$.

(μ.2) Su dominio $\mathfrak{M}_\mu$ es el campo σ de subconjuntos de $\mathcal{R}$ con la unidad $\mathcal{R}$.

(μ.3) μ es aditiva numerable

(μ.4) μ es completa

(μ.5) La medida $\mu(A)$ de cualquier conjunto $A \in \mathfrak{M}_\mu$ es la cota superior de las medidas $\mu(F)$ de los conjuntos $F \subset A$ cerrados.

Para definir la convergencia débil de medidas, se dará una definición que incluye el concepto de "Integral de Lebesgue". (Ver apéndice [C])

DEFINICION 1.1.2.- La sucesión $\{\mu_n\}$ de medidas CONVERGE DEBILMENTE a la medida μ si, para cualquier función $f \in C(\mathcal{R})$

se tiene $\int_{\mathcal{R}} f(x)d\mu_n \longrightarrow \int_{\mathcal{R}} f(x)d\mu.$

La denotaremos por $\mu_n \Rightarrow \mu$.

Es importante señalar el hecho que en algunos casos es difícil determinar la convergencia apartir de la definición por lo cual nos vemos en la necesidad de enunciar criterios alternativos que permitan establecer condiciones necesarias y suficientes para la convergencia antes definida.

Expondremos tres teoremas que incluyen definiciones de un tipo especial de conjuntos y una medida que se relaciona con el concepto de "Función de Distribución" (Ver [D])

Teorema 1.1.1.- Para la convergencia débil $\mu_n \Rightarrow \mu$ es necesario y suficiente que las siguientes condiciones se cumplan :

i) $\lim_n \mu_n(\mathcal{R}) = \mu(\mathcal{R})$, $\underline{\lim}_n \mu_n(G) \geq \mu(G)$

para cualquier conjunto $G \subset \mathcal{R}$ abierto;

ó en forma alternativa, las condiciones

ii) $\lim_n \mu_n(\mathcal{R}) = \mu(\mathcal{R})$, $\overline{\lim}_n \mu_n(F) \leq \mu(F)$

para cualquier conjunto $F \subset \mathcal{R}$ cerrado.

DEMOSTRACION .- Necesidad: Primero demostraremos la parte (ii) : Como $\mu_n \Rightarrow \mu$ ent. tomemos la función indicadora $I \in C(\mathcal{R})$ y resulta que

$$\lim_n \mu_n(\mathcal{R}) = \lim_n \int_{\mathcal{R}} I\, d\mu_n = \int_{\mathcal{R}} I\, d\mu = \mu(\mathcal{R})$$

la segunda implicación se tiene de la condición $(\mu.5)$ de la definición de medida; esto es, sea $A \in \mathfrak{M}_\mu$

$$\mu(A) = \sup \{ \mu(F) \; ; \; F \subset A, \; F \text{ cerrado} \}$$

de aquí tenemos

$$\overline{\lim}_n \, \mu_n(F) \leq \sup \{ \mu_n(F) \} \leq \mu(F) \text{ con } F \text{ cerrado.}$$

Para demostrar (i) observemos que $\overline{\lim}_n \mu_n(F) \leq \mu(F)$ con F un conjunto cerrado; sea G un conjunto abierto, entonces

$$\begin{aligned}
& \overline{\lim}_n \, \mu_n(\mathcal{R} - G) \leq \mu(\mathcal{R} - G) \\
\Longrightarrow \; & \overline{\lim}_n [\mu_n(\mathcal{R}) - \mu_n(G)] \leq \mu(\mathcal{R}) - \mu(G) \\
\Longrightarrow \; & \overline{\lim} \, \mu_n(\mathcal{R}) + \overline{\lim}_n (-\mu_n(G)) \leq \mu(\mathcal{R}) - \mu(G) \\
\Longrightarrow \; & \mu(\mathcal{R}) + \overline{\lim} (-\mu_n(G)) \leq \mu(\mathcal{R}) - \mu(G) \\
\Longrightarrow \; & -\overline{\lim}_n (-\mu_n(G)) \geqslant \mu(G) \\
\Longrightarrow \; & \underline{\lim}_n \, \mu_n(G) \geqslant \mu(G) \quad \text{con } G \text{ abierto.}
\end{aligned}$$

Suficiencia: Como $\overline{\lim}_n \mu_n(F) \leq \mu(F)$ $F \subset \mathcal{R}$ cerrado, resulta que $\lim_n \mu_n$ si es que existe, está definido en la forma deseada (cumple μ.1 - 5). Garantizar la existencia de dicho límite equivale a demostrar la convergencia. Pero sabemos que $\lim_n \mu_n(\mathcal{R}) = \mu(\mathcal{R})$ donde

$$\lim_n \mu_n(\mathcal{R}) = \lim_n \int_{\mathcal{R}} \mathbb{I} \, d\mu_n = \int_{\mathcal{R}} \mathbb{I} \, d\mu \quad \text{con } \mathbb{I} \in C(\mathcal{R})$$

lo que significa que $\mu_n \Rightarrow \mu$.

Para enunciar el segundo teorema es necesaria la siguiente definición.

DEFINICION 1.1.3.- Un conjunto $A \subset \mathcal{R}$ es llamado DE CONTINUIDAD si, la medida μ de su frontera es igual a cero.

Nota.- En lo sucesivo, a menos que se diga lo contrario, tendremos en mente conjuntos de continuidad de Borel.

TEOREMA 1.1.2.- Para la convergencia $\mu_n \Rightarrow \mu$ es necesario y suficiente que para cualquier conjunto A de continuidad, para la medida μ, se tenga

$$\mu_n(A) \longrightarrow \mu(A)$$

DEMOSTRACION.- Sabemos que $\lim_n \mu_n(A) = \mu(A)$ con A un conjunto de continuidad. Pero $\mu_n(A) = \mu_n(A^\circ) + \mu_n(\text{front}(A))$ donde A° es abierto $\Rightarrow \mu_n(A) = \mu_n(A^\circ)$
$\Rightarrow \underline{\lim}_n \mu(A^\circ) = \mu(A^\circ)$ con A° abierto y además, dada la convergencia, resulta que

$$\lim_n \mu_n(\mathcal{R}) = \mu(\mathcal{R}).$$

Por lo tanto, en virtud del teorema (1.1.1) resulta que $\mu_n \Rightarrow \mu$.

Con estos dos criterios (Teo 1.1.1 y 1.1.2) podemos establecer, operacionalmente, la convergencia débil; sin embargo, presentaremos otro resultado de mayor utilidad, ya que las medidas que vamos a utilizar son de probabilidad.

Tomemos una función ξ cuyo dominio es $\mathcal{R}$ y su contradominio, el espacio métrico $\mathcal{R}'$; esta función así definida da lugar en $\mathcal{R}'$ a una función que denotaremos mediante μ^{ξ} donde $\mu^{\xi}(A') = \mu(\xi^{-1}(A'))$, la cual es aditiva numerable. El dominio $\mathfrak{M}_{\mu^{\xi}}$ consiste de todos los conjuntos $A' \subset \mathcal{R}'$ cuya imagen inversa $\xi^{-1}(A')$ está en $\mathfrak{M}_{\mu}$.

TEOREMA 1.1.3.- Toda función ξ, continua casi en todas partes, del espacio $\mathcal{R}$ al espacio métrico $\mathcal{R}'$, da lugar en $\mathcal{R}'$ a una medida que denotaremos mediante el símbolo μ^{ξ}.

DEMOSTRACION .- Tenemos que verificar que la función μ^{ξ} cumpla las condiciones (μ.1 - 5).

(μ.1) El conjunto $C \subset \mathcal{R}$ de puntos en los cuales ξ es continua tiene la medida $\mu(C) = \mu(\mathcal{R})$ y puede ser representada de la forma $C = (\bigcup_m F_m) \cup C_1$ donde F_m son cerrados y $\mu(C_1) = 0$. En consecuencia $\mathcal{R} = (\bigcup_m F_m) \cup C_2$ con $\mu(C_2) = 0$. De conformidad con ésto, la imagen inversa $\xi^{-1}(F')$ de los conjuntos cerrados $F' \subset \mathcal{R}'$ permite la representación

$$\xi^{-1}(F') = \bigcup_m (\xi^{-1}(F') \cap F_m) \cup (C_2 \cap \xi^{-1}(F'))$$

donde el conjunto $C_2 \cap \xi^{-1}(F')$ es Medible μ con medida cero y todos los conjuntos restantes son cerrados, entonces el conjunto $\xi^{-1}(F')$ es medible.

De aquí resulta que la función ξ es medible. Más aún, para un conjunto arbitrario $A' \in \mathfrak{M}_{\mu^{\xi}}$ es posible encontrar un conjunto compacto K_{ε} tal que satisfaga las condiciones $K_{\varepsilon} \subset C \cap \xi^{-1}(A')$, $\mu(\xi^{-1}(A') - K_{\varepsilon}) \leq \varepsilon$.

El conjunto $\xi(K_{\varepsilon}) \subset A'$ es compacto, por ser la imagen continua de un compacto, y $\mu^{\xi}(A' - \xi(K_{\varepsilon})) = \mu(\xi^{-1}(A') - K_{\varepsilon}) \leq \varepsilon$. De esto tenemos que en todo conjunto $A' \in \mathfrak{M}_{\mu^{\xi}}$ podemos encontrar un conjunto compacto cuya medida esté tan cercana a la medida de A', como se requiera. Esto implica (μ.5).

(μ.2) Su dominio es el campo σ de los subconjuntos de $\mathcal{R}$ con la unidad $\mathcal{R}$ ya que $\xi^{-1}(\mathcal{R}') = \mathcal{R}$.

(μ.3) μ^{ξ} es aditiva numerable, puesto que μ lo es.

(μ.4) μ^{ξ} es completa en virtud de que μ lo es.

El teorema queda demostrado.

TEOREMA 1.1.4.- Para la convergencia $\mu_n \Rightarrow \mu$ es necesario y suficiente que para cualquier función ξ continua casi en todas partes $\mu_n^{\xi} \Rightarrow \mu^{\xi}$.

DEMOSTRACION .- Suficiencia: Sea $\xi \in C(\mathcal{R})$ tal que $\sup_{x \in \mathcal{R}} |\xi(x)| < a$.

Supongamos que $g(y) = y$ cuando $|y| \leq a$, $g(y) = a$ (ó $-a$) para $y > a$ (respectivamente $y < -a$). Las igualdades válidas para cualquier medida μ en $\mathcal{R}$ $\int_{\mathcal{R}} g(x)\, d\mu = \int_{|y|\leq a} y\, d\mu^{g} = \int g(y) d\mu^{g}$ y las condiciones del teorema, implican que $\mu_n \overset{\mathcal{R}'}{\Rightarrow} \mu$.

Necesidad: Tomemos un conjunto C de puntos donde g sea continua. Por hipótesis $\mu(\mathcal{R}-C)=0$. Para la medida μ de la frontera del conjunto $E_\alpha=\{x;\, g(x)<\alpha\}$ obtenemos $\text{front}(E_\alpha) \subset \{x;\, g(x)=\alpha\} \cup (\mathcal{R}-C)$ la estimación $\mu(\text{front}(E_\alpha)) \leq \mu\{x;\, g(x)=\alpha\}$ donde esta cantidad puede ser distinta de cero para a lo más una cantidad numerable de valores de α. En consecuencia, para un conjunto denso de números reales α, tenemos que $\mu(\text{front}(E_\alpha))=0$. Para tal α, las funciones de distribución $F_n(\alpha)=\mu_n(E_\alpha)=\mu_n^{g}(\{y;\, y<\alpha\})$ convergen a la función $F(\alpha)=\mu(E_\alpha)=\mu^{g}(\{y;\, y<\alpha\})$. En suma $F_n(\alpha) \rightarrow F(\alpha)$. Por consiguiente $\mu_n^{g} \Rightarrow \mu^{g}$ y la demostración termina.

Para terminar esta sección, introduciremos un resultado general que está relacionado con el teorema anterior pero utilizando funciones continuas; cabe señalar que el siguiente teorema cobrará gran importancia al introducir medidas de probabilidad. Observemos además que por primera vez estableceremos la convergencia utilizando el concepto de "Distancia" entre términos de la sucesión.

TEOREMA 1.1.5.- Si $\mu_n \Rightarrow \mu$ y si $\{\zeta_n(x)\}$ es una sucesión de funciones continuas, del espacio $\mathcal{R}$ en el espacio métrico $\mathcal{R}'$, convergente uniforme a la función $\zeta(x)$ en todo conjunto $K \subset \mathcal{R}$ compacto, entonces $\mu_n^{\zeta} \Rightarrow \mu^{\zeta}$.

DEMOSTRACION .- Sea $g(y) \in C(\mathcal{R}')$ y $\varepsilon > 0$ un número dado. Escojamos un conjunto K compacto para el cual

$$\mu(\mathcal{R} - K) < \varepsilon/4M \quad y \quad \mu_n(\mathcal{R} - K) < \varepsilon/4M \ , \quad n \geq 1.$$

donde $M = \sup_{y \in \mathcal{R}'} |g(y)|$. La imagen de K bajo la función $\zeta(x)$ es un conjunto compacto $K' \subset \mathcal{R}'$. Podemos encontrar un número $\delta > 0$ tan pequeño que $|g(y^*) - g(y)| < \frac{\varepsilon}{4}$ siempre que $\rho'(y^*, y) < \delta$ $y \in K'$, y tomemos n_0 tan grande que para $n > n_0$ $\sup_{x \in K} \rho'(\zeta_n(x), \zeta(x)) < \delta$

y $\left| \int_{\mathcal{R}} g(\zeta(x)) d\mu_n - \int_{\mathcal{R}} g(\zeta(x)) d\mu \right| < \frac{\varepsilon}{4}$. Entonces, para $n > n_0$

$\left| \int_{\mathcal{R}} g(\zeta_n(x)) d\mu_n - \int_{\mathcal{R}} g(\zeta(x)) d\mu \right| < \frac{\varepsilon}{4}$ lo cual demuestra el teorema.

SECCION 2.- EL ESPACIO $\mathcal{Q}(\mathcal{R})$.

Utilizaremos ahora los conceptos referentes a espacios métricos; esto es; daremos una métrica en el espacio de medidas para así, de este modo, hablar de convergencia determinando la "cercanía" entre medidas.

Tomando un espacio métrico separable completo $\mathcal{R}$, denotamos mediante el símbolo $\mathcal{M}$ al conjunto de medidas definidas en $\mathcal{R}$. De este modo, introducimos una distancia entre medidas como sigue :

Tomemos un conjunto $F \subset \mathcal{R}$ cerrado. Denotaremos mediante F^{ε} a la bola abierta $S(F,\varepsilon)$ y definimos los números $\varepsilon_{1,2}$ y $\varepsilon_{2,1}$ como sigue.

$\varepsilon_{1,2}$ es la cota inferior de aquellos $\varepsilon>0$ para los cuales, para todo conjunto cerrado $\mu_2(F) < \mu_1(F^{\varepsilon})+\varepsilon$ con $\mu_1, \mu_2 \in \mathcal{M}$ respectivamente, $\varepsilon_{2,1}$ es la cota inferior de aquellos $\varepsilon>0$ para los cuales, para todo conjunto cerrado

$$\mu_1(F) < \mu_2(F^{\varepsilon}) + \varepsilon \quad \text{con } \mu_1, \mu_2 \in \mathcal{M} .$$

DEFINICION 1.2.1.- Definimos a la función distancia entre dos medidas como sigue

$$L(\mu_1,\mu_2) = \max \{ \varepsilon_{1,2},\ \varepsilon_{2,1} \}.$$

Hay que verificar que la distancia $L(\mu_1,\mu_2)$ así definida cumple : i) $L(\mu_1,\mu_2) \geq 0$, $L(\mu_1,\mu_1) = 0$

ii) $L(\mu_1,\mu_2) = L(\mu_2,\mu_1)$

iii) $L(\mu_1,\mu_3) \leq L(\mu_1,\mu_2) + L(\mu_2,\mu_3)$

Finalmente, si $L(\mu_1,\mu_2)$ es igual a cero, entonces las desigualdades donde se involucran $\varepsilon_{1,2}$ y $\varepsilon_{2,1}$ se satisfacen para todo $\varepsilon>0$.

Obs.- En el límite cuando $\varepsilon \to 0$ obtenemos $\mu_2(F) \leq \mu_1(F)$ y $\mu_1(F) \leq \mu_2(F)$, los valores de las medidas μ_1 y μ_2 coinciden para cualquier conjunto cerrado. Por la propiedad (μ.5) (def. 1.1.1) reulta que $\mu_1 \equiv \mu_2$.

DEFINICION 1.2.2.- El conjunto $\mathfrak{N}$ de medidas en $\mathcal{R}$, junto con la función $L(\mu_1, \mu_2)$ definiendo una distancia entre medidas, forman un espacio métrico que denotaremos mediante el símbolo $\mathfrak{D}(\mathcal{R}) = (\mathfrak{N}, L)$.

Veremos un resultado que describe la estructura del espacio $\mathfrak{D}(\mathcal{R})$ y además deja ver la importancia del mismo; pero antes se dará una definición que tiene que ver con el concepto de medidas sobre $\mathcal{R}$.

DEFINICION 1.2.3.- Se dice que un conjunto $\mathfrak{N}$ de medidas en $\mathcal{R}$ satisface la condición (k) si :

(k.1) $\mu(\mathcal{R}) < \infty$ para toda $\mu \in \mathfrak{N}$.

(k.2) Para todo $\varepsilon > 0$ dado existe un conjunto K_ε compacto tal que, para toda $\mu \in \mathfrak{N}$ tenemos $\mu(\mathcal{R} - K_\varepsilon) < \varepsilon$.

Nota .- La condición (k) no solo servirá para la demostración del teorema central de la presente sección, sino que será de gran ayuda en los capítulos siguentes; cabe hacer notar que los conceptos tratados en el artículo de Prokhorov se han asimilado poco a poco, esto viene a colación ya que la condición (k) es conocida actualmente como "la condición (o teorema) de Prokhorov".

Enunciaremos el teorema central del capítulo 1 :

TEOREMA 1.2.1.- El espacio $\mathcal{D}(\mathcal{R})$ es un espacio métrico separable, completo y la Convergencia L es equivalente a la convergencia débil de medidas.

DEMOSTRACION .- La demostración se divide en tres partes; y para ella se requieren algunos lemas.

LEMA 1.2.1.- Si para una función f medible, del espacio $\mathcal{R}$ en sí mismo, el conjunto $H=\{x;\rho(f(x),x)<\delta\}$ tiene una medida μ mayor que $1-\varepsilon$, entonces la distancia entre μ^f y μ no es mayor que $\lambda = \max(\varepsilon,\delta)$.

Demostración.- Sea F un conjunto cerrado y $F^\delta = S(F,\delta)$; en estas condiciones podemos deducir de las siguientes inclusiones $F\cap H \subset f^{-1}(F^\delta)$ y $f^{-1}(F)\cap H \subset F^\delta$

las desigualdades :

$$\mu(F) < \mu(F\cap H)+\varepsilon \le \mu(f^{-1}(F^\delta))+\varepsilon = \mu^f(F^\delta)+\varepsilon$$

$$\text{y} \qquad \mu(f^{-1}(F)) < \mu(f^{-1}(F)\cap H)+\varepsilon \le \mu(F^\delta)+\varepsilon$$

esto implica que

$$\mu(F) \le \mu^f(F^\delta)+\varepsilon \quad \text{y} \quad \mu(f^{-1}(F)) \le \mu(F^\delta)+\varepsilon$$

por lo que

$$L(\mu,\mu^f) < \max(\varepsilon,\delta).$$

Sean ahora $x_0, x_1, \dots x_N$ puntos en $\mathcal{R}$, $X_j=\{x_1,\dots x_j\}$.
definimos a la función $\Psi = \Psi(x_0, X_N, \delta)$ de la siguiente forma:

$$\Psi(x) = x_0 \quad \text{para } x_0 \in T_0 = \mathcal{R} - S(x_N, \delta)$$
$$\Psi(x) = x_j \quad \text{para } x \in T_j = S(x_j, \delta) - S(x_{j-1}, \delta),\ 1 \leq j \leq N$$

Corolario al lema .- Si $\mu(\mathcal{R}-S(x_N,\delta)) \leq \varepsilon$ entonces, para Ψ

$$L(\overset{\Psi}{\mu}, \mu) \leq \lambda = \max(\varepsilon, \delta)$$

PARTE "A" : Convergencia L es equivalente a la convergencia Débil.

LEMA 1.2.2.- Cualquier sucesión de medidas que converge débilmente (a alguna medida en $\mathcal{R}$) satisface la condición (k).

Demostración .- Sea $\{\mu_n\}$ una sucesión de medidas tal que $\mu_n \Rightarrow \mu$ con μ en $\mathcal{R}$. Tomando la función indicadora tenemos

$$\int_{\mathcal{R}} I \, d\mu_n = \mu_n(\mathcal{R}) \longrightarrow \mu(\mathcal{R}) < \infty \quad \text{y de ahí que } \mu_n(\mathcal{R}) < \infty$$

para toda medida en la sucesión. Por lo tanto se cumple (k.1). Para (k.2), observemos el hecho de que $\{\mu_n\}$ es una sucesión de medidas; y en virtud del teorema (1.1.5), para cada $\varepsilon > 0$ existe un conjunto K_ε compacto tal que $\mu_n(\mathcal{R} - K_\varepsilon) < \varepsilon$, y esto se da para toda medida en la sucesión. Queda demostrado el lema.

Para demostrar "A" : Sea $L(\mu_n, \mu) \to 0$ cuando $n \to \infty$, y sea F un conjunto cerrado. Entonces, para cualquier $\varepsilon > 0$

$$\overline{\lim}_n \mu_n(F) \leq \mu(F^\varepsilon) + \varepsilon$$

ésta desigualdad, en el límite cuando $\varepsilon \to 0$, se transforma en $\overline{\lim}_n \mu_n(F) \leq \mu(F)$.
Esto también sucede cuando el conjunto a tratar es $\mathcal{R}$; es decir, $\mu_n(\mathcal{R}) \to \mu(\mathcal{R})$ y por el teorema (1.1.1) $\mu_n \Rightarrow \mu$.

De forma similar admitamos que $\mu_n \Rightarrow \mu$. En virtud del lema (1.2.2) es posible encontrar un conjunto finito de puntos $X = \{x_1, \ldots, x_N\}$ en el cual, para $\delta' = \frac{\varepsilon}{5}$ y para todo $n \geq 0$, la medida del conjunto $\mathcal{R} - S(X, \delta')$ es menor que $\frac{2\varepsilon}{5}$. Mediante una elección conveniente del número δ en el intervalo $(\delta', 2\delta')$ es posible arreglar que todas las esferas $S(x_j, \delta)$, $1 \leq j \leq N$, sean conjuntos de continuidad. Completando el conjuto X con el punto arbitrario $x_0 \in \mathcal{R}$, debemos construir una función Ψ del espacio $\mathcal{R}$ en sí mismo, como se indica en el corolario al lema (1.1.1) Entonces, en primer lugar, los conjuntos T_j que están en la definición de Ψ deben ser conjuntos de continuidad y, en consecuencia es posible escoger n_0 tan grande que para $n \geq n_0$ cada una de las diferencias $\mu_n(T_j) - \mu(T_j)$ sea en valor absoluto menor que $\frac{\varepsilon}{5(N+1)}$. En segundo lugar, para toda $n \geq 0$ $L(\mu^{\Psi}, \mu) < \frac{2\varepsilon}{5}$. Debemos observar que la distancia entre las medidas $\tilde{\mu}_n$ y μ^{Ψ} concentrada en los puntos x_j no debe exceder la suma sobre j , de 0 a N, de las cantidades $|\overset{\vee}{\mu}_n(x_j)-\overset{\vee}{\mu}(x_j)| = |\mu_n(T_j)-\mu(T_j)|$ esto es, para $n \geq n_0$ no excede $\frac{\varepsilon}{5}$. Por lo tanto, para $n \leq n_0$

$$L(\mu_n,\mu) \leq L(\mu_n,\overset{\vee}{\mu}_n) + L(\overset{\vee}{\mu}_n,\overset{\vee}{\mu}) + L(\overset{\vee}{\mu},\mu) < \varepsilon$$

Con lo que queda demostrada la parte "A".

PARTE "B" : El Espacio $\mathfrak{D}(\mathcal{R})$ es separable.

Denotaremos mediante $\mathfrak{D}_\Gamma$ a la totalidad de medidas en $\mathcal{R}$ las cuales se definen de la siguiente forma :

De algún conjunto Γ numerable denso en $\mathcal{R}$, escogemos una cantidad finita de elementos $x_{(1)}, x_{(2)}, \ldots, x_{(s)}$; a cada punto $x_{(j)}$ le asignamos un número racional $r(j)>0$ y definimos $\mu'(A) = \sum_{x_{(j)}\in A} r(j)$.

Debemos probar que el conjunto $\mathfrak{D}_\Gamma$ (que es numerable) es denso en $\mathfrak{D}(\mathcal{R})$.

Sea $\mu \in \mathfrak{D}(\mathcal{R})$ una medida y sea $\varepsilon>0$ fijo. Escogeremos un subconjunto finito de puntos que pertenecen a Γ de tal forma que la medida de sus $\frac{\varepsilon}{2}$ -vecindades sea mayor que $\mu(\mathcal{R})-\frac{\varepsilon}{2}$. Completando este subconjunto con un punto arbitrario $x_0 \in \Gamma$, definiremos la función Ψ tal como en el corolario al lema (1.2.1). Entonces, la distancia entre la medida μ y la medida $\overset{\vee}{\mu}$ correspondiente, concentrada en un número finito de puntos de Γ , no será mayor que $\frac{\varepsilon}{2}$. Como se observa, existe una medida $\mu' \in \mathfrak{D}_\Gamma$

concentrada en los mismos puntos que μ^{ν} para la cual, $L(\mu^{\nu},\mu)\leq\frac{\varepsilon}{2}$. Para la medida μ' $L(\mu,\mu')\leq\varepsilon$. Con lo que se demuestra la parte "B".

PARTE "C" : La completez del espacio $\mathfrak{D}(\mathcal{R})$.

LEMA 1.2.3.- Cualquier sucesión $\{\mu_n\}$ Fundamental L satisface la condición (k).

Demostración .- Bajo las hipótesis, la sucesión de números $\mu_n(\mathcal{R})$ es fundamental y, por lo tanto acotada; (k.1) está demostrado. Ahora sean $\varepsilon>0$ y $\delta>0$ fijos. Debemos determinar n_0 de la desigualdad

$$\sup_{r>0} L(\mu_{n_0},\mu_{n_0+r})<\lambda=\min\left(\frac{\varepsilon}{3},\frac{\delta}{2}\right)$$

y escoger un conjunto X finito de puntos de $\mathcal{R}$ para el cual $\mu_{n_0}(\mathcal{R})-S(X,\frac{\delta}{2})<\frac{\varepsilon}{3}$.

Para $r>0$ la cantidad $\mu_{n_0}(\mathcal{R})+r$ no excede $\mu_{n_0}(\mathcal{R})+\lambda$.

Pero $\mu_{n_0}(\mathcal{R})+\lambda\leq\mu_{n_0+r}\left(\left[S(X,\frac{\delta}{2})\right]^{\lambda}\right)+\lambda+\lambda+\frac{\varepsilon}{3}\cdots(*)$

Las relaciones $\left[S(X,\frac{\delta}{2})\right]^{\lambda}\subset S(X,\delta)$, $2\lambda+\frac{\varepsilon}{3}<\varepsilon$, nos permiten reemplazar (*) por

$$\mu_{n_0+r}(\mathcal{R})<\mu_{n_0+r}(S(X,\delta))+\varepsilon$$

Vemos que (k.2) se cumple. El lema está demostrado.

LEMA 1.2.4.- Si la sucesión $\{\mu_n\}$ es Fundamental L entonces, la sucesión $\{\mu_n^f\}$ de medidas en la línea recta es fundamental L, cualquiera que sea la función continua f.

Demostración .- Sea $h>0$ un número fijo. De acuerdo al lema(1.2.3) debemos escoger un conjunto compacto $K_{\frac{h}{2}}$ tal que para toda n , $\mu_n(R - K_{h/2}) < \frac{h}{2}$. Es posible tomar δ_0 tan pequeño que, la diferncia entre los valores de $f(x)$ a cualesquiera dos puntos $x, x' \in K_{h/2}^{\delta_0}$ $\rho(x,x') < \delta_0$, debe ser menor o igual que $\frac{h}{2}$ en valor absoluto. Entonces, para todo α real y para n suficientemente grande ($n \geq n_0$) tales que

$$\sup_{n_1, n_2 \geq n} L(\mu_{n_1}, \mu_{n_2}) < \delta_1 = \min(\delta_0, \tfrac{h}{2})$$

las siguentes estimaciones son válidas

$$G_\alpha = \{x ; f(x) < \alpha\}, \; F_\alpha = \overline{G_\alpha}$$

$$\mu_{n_1}(F_\alpha) \leq \mu_{n_1}(F_\alpha \cap K_{h/2}) + \frac{h}{2} \leq \mu_{n_2}[(F_\alpha \cap K_{h/2})^{\delta_1}] + \delta_1 + \frac{h}{2} \leq \mu_{n_2}(G_{\alpha+h}$$

En consecuencia, la distancia L entre μ_1^f y μ_2^f para $n_1, n_2 \geq n_0$, no excede el valor de $h>0$ dado. Lo que demuestra el lema.

DEMOSTRACION DE "C" : De la completez del espacio de medidas finitas sobre la recta, tenemos (en la métrica L) que para toda función f continua, existe una medida μ^f sobre la recta, tal que $\mu_n^f \Rightarrow \mu^f$. De ésta última relación podemos concluir que para toda función f continua y acotada, el siguiente límite existe

$$\lim_n \int_R f(x)\, d\mu_n$$

Puesto que $\lim_n \int_{\mathcal{R}} f(x)\, d\mu_n = \int_{\mathcal{R}} f(x)\, d\mu$, $\mu_n \Rightarrow \mu$, $L(\mu_n, \mu) \to 0$

La completez del espacio está demostrada.

Con esto termia la demostración del teorema.

Por último, estableceremos las condiciones para la compacidad de los subconjuntos de $\mathcal{D}(\mathcal{R})$.

TEOREMA 1.2.2.- La condición (k) es necesaria y suficiente para que el conjunto $\mathfrak{N} \subset \mathcal{D}(\mathcal{R})$ de medidas sea compacto.

DEMOSTRACION .- Necesidad: Sean $\varepsilon > 0$ y $\delta > 0$ números fijos. Definamos a $2\lambda = \min\left(\frac{\delta}{2}, \frac{\varepsilon}{3}\right)$ y escojamos en $\mathfrak{N}$ una λ-red que llamaremos $\mathfrak{N}'$. Además, escogeremos un conjunto X finito de puntos tal que $\mu'(S(X, \frac{\delta}{2})) \geq \mu'(\mathcal{R}) - \lambda$ para toda medida $\mu' \in \mathfrak{N}'$. Para cualquier medida $\mu \in \mathfrak{N}$ existe una $\mu' \in \mathfrak{N}'$ con $L(\mu, \mu') \leq \lambda$. Para μ' las siguientes desigualdades se cumplen : $\mu'(\mathcal{R}) - \lambda \leq \mu'(S(X, \frac{\delta}{2})) < \mu(S(X, \frac{\delta}{2} + 2\lambda)) + 2\lambda$

y $\mu'(\mathcal{R}) > \mu(\mathcal{R}) - 2\lambda$

De éstas desigualdades resulta que, cualquiera que sea la medida $\mu \in \mathfrak{N}$ $\mu(S(X, \delta)) > \mu(\mathcal{R}) - \varepsilon$.

Suficiencia: Completando al conjunto $X_{\varepsilon/2}$ de puntos con el punto $x_0 \in \mathcal{R}$, construimos una función $\Psi = \Psi(x_0, X_{\varepsilon/2}, \xi)$ (ver corolario al lema (1.2.1)). Entonces para cualquier $\mu \in \mathfrak{N}$ obtenemos $L(\mu, \mu^{\Psi}) < \frac{\varepsilon}{2}$. En vitrud de la condición (k)

$$\sup_{\mu\in\mathfrak{N}} \mu^{\gamma}(\Omega) = \sup_{\mu\in\mathfrak{N}} \mu(\Omega) < \infty$$

De la última relación y del hecho de que todas las medidas μ^{γ} se concentran en los mismos puntos, se puede deducir que la familia $\mathfrak{N}^{\gamma} = \{\mu^{\gamma}\}, \mu \in \mathfrak{N}$ resulta ser totalmente acotada. Escogiendo en $\mathfrak{N}^{\gamma}$ una $\frac{\varepsilon}{2}$ -red, veremos que sirve, al mismo tiempo, como ε -red en $\mathfrak{N}$. Como $\mathfrak{D}(\Omega)$ es completo, éstas observaciones implican la compacidad de $\mathfrak{N}$.

CAPITULO 2 .- PROCESOS Y FUNCIONES ALEATORIAS.

SECCION 2.1.- TEOREMA DE PROKHOROV.

Los procesos estocásticos que vamos a estudiar tienen sus trayectorias continuas; por ésta razón es que debemos referirinos al espacio métrico de las funciones continuas C (ver apéndice [B]) ,cuando hablemos de algún conjunto de medidas de probabilidad y así, podamos determinar cuándo dicho cunjunto es relativamente compacto. Esto es con el fin de poder determinar la convergencia débil de los procesos a estudiar.

Definiendo al espacio medible (C , $\mathcal{B}_C$), donde $\mathcal{B}_C$ es el álgebra σ generado por los conjuntos abiertos de C, tenemos que una medida de probabilidad P en dicho espacio se caracteriza por susu distribuciones finito dimensionales (ver apéndice [E]). Además, resulta que para una familia $H \subset C$ de funciones, las siguientes aseveraciones son equivalentes :

a) H es relativamente compacta en C .

b) H es equicontinua y equiacotada (ver[E])

c) Cualquier sucesión en H tiene una subsucesión uniformemente convergente.

De esta equivalencia, se tiene la importancia de poder determinar la compacidad relativa de los subconjuntos del espacio $\mathcal{C}$, ya que de la convergencia uniforme se tiene la convergencia débil.

Veremos un concepto, estudiado para medidas en general, pero que se aplica de forma especial en las medidas de probabilidad. Dicho concepto ayuda a establecer condiciones para determinar la convergencia débil.

Designemos con $\mathcal{X}$ a un espacio métrico separable completo; denotaremos mediante Γ a una familia de medidas de probabilidad en $\mathcal{X}$.

DEFINICION 2.1.1.- Una familia Γ de medidas de probabilidad en el espacio métrico $\mathcal{X}$ se llama TENSA si, para todo $\varepsilon>0$ existe un conjunto compacto $K_\varepsilon \subset \mathcal{X}$ tal que

$$\underline{P}(K_\varepsilon) > 1-\varepsilon \qquad \text{para toda } \underline{P} \in \Gamma .$$

Enunciaremos dos teoremas que en conjunto forman el así llamado "TEOREMA DE PROKHOROV". (Billingsley [1]).

TEOREMA 2.1.1.- Si el espacio métrico $\mathcal{X}$ es separable y completo entonces, la familia Γ de medidas de probabilidad es Tensa.

DEMOSTRACION .- Denotemos mediante $S(x,\varepsilon)$ a la bola abierta con centro en X y radio ε . Como $\mathcal{R}$ es separable, existe una sucesión $\{x_n\}$ densa en $\mathcal{R}$ tal que para toda $r \in \mathbb{N}$ y toda $\varepsilon > 0$ $\bigcup_{n=1}^{\infty} \bar{S}(x_n, \frac{1}{r}) = \mathcal{R}$.

Ahora tomemos un conjunto $I_r \subset \mathbb{N}$ finito tal que

$$P\left[\mathcal{R} - \bigcup_{n \in I_r} \bar{S}(x_n, \tfrac{1}{r})\right] \leq \frac{\varepsilon}{2^r} .$$

Sea $K_\varepsilon = \bigcap_r \bigcup_{n \in I_r} \bar{S}(x_n, \frac{1}{r})$. Este conjunto es totalmente acotado y cerrado, por lo tanto compacto; Ahora, de la última desigualdad tenemos que

$$P\left[\left(\bigcup_{n \in I_r} \bar{S}(x_n, \tfrac{1}{r})\right)^c\right] \leq \frac{\varepsilon}{2^r} \quad \text{de donde}$$

$$P(K_\varepsilon^c) \leq \sum_{r=1}^{\infty} \frac{\varepsilon}{2^r} = \varepsilon \quad \therefore \quad P(K_\varepsilon) > 1 - \varepsilon .$$

TEOREMA 2.1.2.- Si la familia Π es tensa entonces, es relativamente compacta.

DEMOSTRACION .- La demostración de este teorema es equivalente a la del teorema (1.2.2) donde haciamos mención de la condición (k); no debe haber confusión ya que se mencionó que la condición (k) es ahora conocida como la condición de Prokhorov.

En base al Teorema de Prokhorov es que podemos enunciar un criterio para la convergencia débil.

PROPOSICION 2.1.1.- Si $\mathcal{R}$ es un espacio métrico separable, de cualquier sucesión $\{P_n\}$ tensa, se puede extraer una subsucesión que converge débilmente (es decir, lamilia $\{P_n\}$ es relativamente compacta). Si la sucesión no converge débilmente, se pueden encontrar dos subsucesiones que converegen débilmente hacia dos probabilidades distintas.

DEMOSTRACION .- Para probar que $\{P_n\}$ es relativamente compacta debemos mostrar que toda subsucesión $\{P_{n_k}\}$ converge. Sea P el límite de $\{P_n\}$. Debe existir una sucesión $n_1 < n_2 < \cdots$ tal que $P_{n_k} \Rightarrow P$ para toda $P_n \in \{P_n\}$. Entonces en virtud del teorema (1.1.1) podemos concluir que $P(K_r) \geq \overline{\lim_{k\to\infty}} P_{n_k}(K_r) \geq 1-\frac{1}{r}$. Escribiendo $E = \bigcup_r K_r$ observamos que $E \subset R$ es un conjunto tal que $P(E) = 1$. Por lo tanto existe una medida P' en R tal que $P' = P$. Sea ahora C un conjunto cerrado de R. Entonces $\overline{\lim}_k P_{n_k}(C) \leq P(C)$ En virtud del teorema(1.1.2) $P_{n_k} \Rightarrow P$. Esto muestra que $\{P_n\}$ tiene una subsucesión convergente.

SECCION 2 .- CONDICIONES PARA LA COMPACIDAD DE FAMILIAS DE DISTRIBUCIONES EN C.

A partir del teorema(B.1) y (B.2) del apéndice [B], obtenemos un criterio para determinar la compacidad requerida.

LEMA 2.2.1.- Sea Γ un conjunto de medidas de probabilidad definidas en $\mathcal{B}_{C}$. A fin de que $\overline{\Gamma}$ sea compacta es necesario y suficiente que para toda $\varepsilon>0$ exista una constante M_ε y una función $g(\delta,\varepsilon)$ la cual decrece a cero cuando $\delta\downarrow 0$, tal que : i) $P(\{x., |x(0)|\leq M_\varepsilon\})>1-\frac{\varepsilon}{2}$

y ii) $P(\{x., \omega_x(\delta)\leq g(\delta,\varepsilon) \text{ para toda } \delta\})>1-\frac{\varepsilon}{2}$

para toda $P\in\Gamma$.

Demostración .- Supongamos que Γ es un conjunto que satisface las condiciones del teorema. Para $\varepsilon>0$, definimos

$$A_\varepsilon=\{x., |x(0)|\leq M_\varepsilon\}$$
$$B_\varepsilon=\{x., \omega_x(\delta)\leq g(\delta,\varepsilon) \text{ para toda } \delta\}$$

Haciendo $K_\varepsilon=A_\varepsilon\cap B_\varepsilon$, resulta ser un conjunto cerrado y en virtud del teorema de Arzelá-Ascoli, K_ε es un conjunto compacto. Claramente $P(K_\varepsilon)>1-\varepsilon$ para toda $P\in\Gamma$; es decir, Γ es tensa, y en virtud del teorema de Prokhorov resulta que Γ es relativamente compacta.

De forma inversa, sea $\overline{\Gamma}$ compacta. Entonces existe un conjunto compacto tal que $P(K_\varepsilon)>1-\frac{\varepsilon}{2}$ para toda $P\in\Gamma$.

Sean $M_\varepsilon=\sup\limits_{x\in K_\varepsilon}|x(0)|$ y $g(\delta,\varepsilon)=\sup\limits_{x\in K_\varepsilon}\omega_x(\delta)$

En virtud del teorema de Arzelá-Ascoli resulta que $M_\varepsilon<\infty$ y $g(\delta,\varepsilon)\downarrow 0$ cuando $\delta\downarrow 0$. Por lo tanto

$$P(\{x., |x(0)|\leq M_\varepsilon\})\geq P(K_\varepsilon)>1-\frac{\varepsilon}{2}$$

y $\quad P(\{x., W_x(\delta) \leq g(\delta,\varepsilon) \text{ para toda } \delta\}) \geq P(K_\varepsilon) > 1-\frac{\varepsilon}{2}$,

para toda $P \in \Gamma$.

En base al lema anterior tenemos el siguiente.

TEOREMA 2.2.1.- A fin de que Γ sea relativamente compacta es necesario y suficiente que las siguientes condiciones se cumplan :

i) Para cada $\varepsilon>0$ debe existir una constante M_ε tal que $P(\{x., |x(0)| \leq M_\varepsilon\}) > 1-\frac{\varepsilon}{2}$

y ii) Para cada $\varepsilon>0$ y $\delta>0$ debe existir una constante $\eta = \eta(\varepsilon,\delta) > 0$ tal que

$$P(\{x., W_x(\eta) \leq \delta\}) > 1-\frac{\varepsilon}{2}$$

para toda $P \in \Gamma$.

DEMOSTRACION .- El que (i) sea necesario se demostró en el lema anterior. Ahora para (ii), notemos que existe una función $g(\delta,\varepsilon) \downarrow 0$ cuando $\delta \downarrow 0$ tal que $P(\{x., W_x(\delta) \leq g(\delta,\varepsilon)\ \forall \delta\}) > 1-\frac{\varepsilon}{2}$ para toda $P \in \Gamma$. Podemos encontrar una $\eta = \eta(\varepsilon,\delta)$ tal que $g(\eta,\varepsilon) \leq \delta$. Entonces resulta que

$$P(\{x., W_x(\eta) \leq \delta\}) > 1-\frac{\varepsilon}{2} \text{ para toda } P \in \Gamma .$$

Para demostrar la suficiencia procederemos como sigue: Para cada entero $n=1,2,\ldots$ y una $\varepsilon>0$ fija, existe una $\eta_{(\varepsilon,n)}$ tal que si definimos al conjunto $E_{\varepsilon,n} = \{x., W_x(\eta_{(\varepsilon,n)}) \leq \frac{1}{n}\}$ entonces $P(E_{\varepsilon,n}) > 1 - \frac{\varepsilon}{2^{n+1}}$ para toda $P \in \Gamma$.

Haciendo $K_\epsilon = \{x, |x(0)| \leq M_\epsilon\} \cap \bigcap_{n=1}^{\infty} E_{\epsilon,n}$,

resulta que $P(K_\epsilon) > 1-\epsilon$ para toda $P \in \Pi$.

Si $x \in K_\epsilon$, entonces para cada n tenemos que $x \in E_{\epsilon,n}$, así que existe una $\eta_{\epsilon,n}$ que posee la propiedad de que $w_x(\eta_{\epsilon,n}) \leq \frac{1}{n}$. En otras palabras, $\sup_{x \in K_\epsilon} w_x(\eta) \to 0$ cuando $\eta \to 0$. Como $\sup_{x \in K_\epsilon} |x(0)| \leq M_\epsilon < \infty$, resulta que K_ϵ es compacto. En virtud del teorema de Prokhorov, Π es relativamente compacta.

SECCION 3 .- PROCESOS ALEATORIOS Y DISTRIBUCIONES EN ESPACIOS DE FUNCIONES.

Denotemos con C al espacio de las funciones continuas en el intervalo $[0,1]$, sea $x \in C$ y definimos a las funciones π^t como $\pi^t : x \to x(t)$ con $t \in [0,1]$.

Denotemos con $\mathcal{B}_C$ a los subconjuntos de Borel de C . Y sean μ y ν dos medidas definidas en $\mathcal{B}_C$. En estas condiciones tenemos el siguiente.

TEOREMA 2.3.1.- La clase $\mathcal{B}_C$ coincide con la álgebra σ mínima de subconjuntos de C con respecto a los cuales las funciones π^t son medibles para todo $t \in [0,1]$. Una condición necesaria y suficiente para que $\mu = \nu$ es que

$\mu^{t_1,\dots,t_k} = \nu^{t_1,\dots,t_k}$ para todo k y $t_1, t_2, \dots, t_k \in [0,1]$, donde $\mu^{t_1,\dots,t_k}$ y $\nu^{t_1,\dots,t_k}$ son medidas en el espacio Euclideano $\mathbb{R}^k$, inducidas por μ y ν, respectivamente, por medio de la función $\pi^{t_1,\dots,t_k} : x \longrightarrow (x(t_1),\dots,x(t_k))$.

DEMOSTRACION .- Sea $\mathcal{A}$ la mínima álgebra σ de subconjuntos de C con respecto a los cuales las funciones π^t son medibles para toda t. Como cada π^t es continua en C, $\mathcal{A} \subset \mathcal{B}_C$. Para probar que $\mathcal{B}_C \subset \mathcal{A}$, es suficiente mostrar que $\mathcal{A}$ contiene a todos los conjuntos abiertos. Como es separable, todo abierto es una unión de bolas cerradas y entonces basta demostrar que $\mathcal{A}$ contiene a todas las bolas cerradas. Si ∂ es cualquier número real positivo, $x_0 \in C$ y $r_1, r_2, \dots$ es un arreglo de racionales en $[0,1]$ tenemos que $\{x, \|x - x_0\| \leq \partial\} = \bigcap_{n=1}^{\infty} \{x, |x(r_n) - x_0(r_n)| \leq \partial\}$. Como la intersección en el miembro derecho pertenece a $\mathcal{A}$ resulta que la bola $\{x, \|x - x_0\| \leq \partial\} \in \mathcal{A}$. Por lo tanto $\mathcal{B}_C = \mathcal{A}$.

Para la segunda parte escribamos $\mathcal{A}_{t_1,\dots,t_k} = \{(\pi^{t_1,\dots,t_k})^{-1}(A), A \subset \mathbb{R}^k\}$ A un conjunto de Borel, $\bigcup_{k,t_1,\dots,t_k} \mathcal{A}_{t_1,\dots,t_k} = \mathcal{F}$. Obviamente $\mu^{t_1,\dots,t_k} = \nu^{t_1,\dots,t_k}$ para toda k y $t_1,\dots,t_k \in [0,1]$ implica que $\mu(E) = \nu(E)$ para todo $E \in \mathcal{F}$. $\mathcal{F}$ es un álgebra generada por $\mathcal{A}$ así resulta que $\mu(E) = \nu(E)$ para todo $E \in \mathcal{A}$. Esto prueba que $\mu = \nu$.

Como el espacio C es un espacio métrico separable completo, resulta que toda medida en $\mathcal{B}_C$ es Tensa.

Una vez examinadas las propiedades de las medidas en $\mathcal{B}_C$, surge una pregunta de modo natural :

¿ Cuáles son aquellos procesos estocásticos cuyas distribuciones correspondientes están definidas en $\mathcal{B}_C$?

De forma precisa, sea $\{\xi_t, 0 \le t \le 1\}$ un proceso estocástico y $P^{t_1,\ldots,t_k}$ su distribucuón finito dimensional; la pregunta se formula de la siguiente forma :

¿ Qué condición adicional debe satisfacerse a fin de que exista una medida μ en $\mathcal{B}_C$ tal que $\mu^{t_1,\ldots,t_k} = P^{t_1,\ldots,t_k}$ para toda K y $t \in [0,1]$?

En virtud del teorema (2.3.1), de existir esa μ, sería única. La siguiente es una interesante condición suficiente.

TEOREMA 2.3.2.- Sea $\{\xi_t, 0 \le t \le 1\}$ un proceso estocástico y $P^{t_1,\ldots,t_k}$ la distribución de probabilidad en $\mathbb{R}^k$ del vector $(\xi_{t_1},\ldots,\xi_{t_k})$. Si existen constantes $\alpha, \delta, K > 0$ tales que

$$E\,|\xi_{t_1}-\xi_{t_2}|^{\alpha} = \iint_{\mathbb{R}^2} |u-v|^{\alpha}\, dP^{t_1,t_2}(u,v) \le K\,|t_1-t_2|^{1+\delta}$$

para toda $t_1,t_2 \in [0,1]$, entonces existe una única medida μ en $\mathcal{B}_C$ tal que $P^{t_1,\ldots,t_k} = \mu^{t_1,\ldots,t_k}$ para toda K y para toda $t_1,\ldots,t_k \in [0,1]$.

DEMOSTRACION ,- Denotemos mediante $\mathcal{R}$ al espacio de todas las funciones de valor real definidas sobre el intervalo $[0,1]$ y $\mathcal{B}_R$ la mínima álgebra σ de subconjuntos de $\mathcal{R}$ respecto a los cuales las funciones $\pi^t : x \to x(t)$ de $\mathcal{R}$ a la recta real son todas medibles. Entonces existe una medida $\mathbb{P}$ en $\mathcal{B}_R$ tal que

$$P^{t_1,\ldots,t_k}(A) = \mathbb{P}(\{x, \pi^{t_1,\ldots,t_k}(x) \in A\})$$

para todo $k, t_1,\ldots,t_k \in [0,1]$ y los conjuntos de Borel $A \subset \mathbb{R}^k$.

Para cualquier $x \in \mathcal{R}$ y cualquier natural $n=1,2,\ldots$ escribimos $x_n^* = \pi_n(x)$ como la única función continua sobre $[0,1]$ que iguala x en los puntos $0, \frac{1}{2^n}, \frac{2}{2^n}, \ldots, 1$ y es lineal entre parejas sucesivas de tales puntos. $\pi_n : x \to x_n^*$ es, para una n fija, una función de $\mathcal{R}$ sobre C . Es posible mostrar que es medible. Vemos que

$$\max_{0\le t\le 1} |\pi_n(x) - \pi_{n+1}(x)| \leq \max_{1\le i\le 2^n} |x(\tfrac{i}{2^n}) - x(\tfrac{i-1}{2^n})|.$$

Debemos probar ahora que $\pi_n(x)$ converge en C para casi toda $x \in \mathcal{R}$. Como C es un espacio métrico completo es suficiente mostrar que $\{\pi_n(x)\}$ es una sucesión fundamental. Para éste fin notemos que para cualquier $\delta > 0$

$$\mathbb{P}(\{x, \|\pi_n(x) - \pi_{n+1}(x)\| > \delta\})$$
$$\leq \mathbb{P}(\{x, \max_i |x(\tfrac{i}{2^n}) - x(\tfrac{i-1}{2^n})| > \delta\})$$
$$\leq \sum_{i=1}^{2^n} \mathbb{P}(\{x, |x(\tfrac{i}{2^n}) - x(\tfrac{i-1}{2^n})| > \delta\})$$
$$\leq \delta^{-\alpha}\left[\sum_{i=1}^{2^n} E|\xi_{\frac{i}{2^n}} - \xi_{\frac{i-1}{2^n}}|^\alpha\right] \leq K\delta^{-\alpha}2^{-n\delta}.$$

Aquí usamos la desigualdad de Tchebyshev y la condición del teorema. Sea θ tal que $0<\theta<\frac{\delta}{\alpha}$. Entonces, haciendo $\partial = 2^{-n\theta}$ tenemos que

$$P\left(\{x, \|\pi_n(x)-\pi_{n-1}(x)\| > 2^{-n\theta}\}\right) \leq K \cdot 2^{n(\alpha\theta-\delta)} .$$

Como $\sum_1^\infty 2^{-n(\delta-\theta\alpha)} < \infty$ resulta, en virtud del lema de Borel-Cantelli, que para casi toda x , $\|\pi_n(x)-\pi_{n-1}(x)\| \leq 2^{-n\theta}$ para toda n suficientemente grande. En consecuencia, para casi toda x , la serie $\sum_n \|\pi_n(x)-\pi_{n-1}(x)\| < \infty$. Como para $m<n$ $\|\pi_m(x)-\pi_n(x)\| \leq \sum_{j=m+1}^{n} \|\pi_j(x)-\pi_{j-1}(x)\|$ tenemos que $\lim_{n,m\to\infty} \|\pi_n(x)-\pi_m(x)\| = 0$ para casi toda $x\in R$. Sea $R_0 \subset R$ el conjunto de todas las x tales que su límite es cero. Para cualquiera $x\in R_0$, existe una $x^* \in C$ tal que $\|\pi_n(x)-x^*\| \to 0$ cuando $n\to\infty$. Definiendo

$$\pi^*(x) = \begin{cases} x^* & \text{si } x \in R_0 \\ 0 & \text{si } x \in R - R_0 \end{cases}$$

π^* es una función medible de R sobre C . Sea μ la medida en C definida como

$$\mu(E) = P\left(\{x, \pi^*(x) \in E\}\right)$$

para todo $E \in \mathcal{B}_C$.

Observemos que para cada $t\in[0,1]$ fija, $x^*(t) = x(t)$ para casi toda x . Esto es claro si t es de la forma $\frac{k}{2^n}$ como en el caso $(\pi_n(x))(t) = x(t)$ para toda n suficientemente grande. Si t no es de la forma $\frac{k}{2^n}$, sean $t_1, t_2, \ldots$ números de la forma $\frac{k}{2^n}$, $t_j \to t$ cuando $j\to\infty$.

$x^*(t_j)=x(t_j)$ para casi toda x. Como $E|\xi_{t_j}-\xi_t|^{\alpha}\to 0$ cuando $j\to\infty$, es claro que para una subsucesión $j_1, j_2, \ldots \to\infty$, $x(t_{j_\nu})\to x(t)$ cuando $\nu\to\infty$ para casi toda x. Como $x^*(t_{j_\nu})\to x^*(t)$ para toda $x\in\mathbb{R}$, resulta que $x(t)=x^*(t)$ para casi toda x.

La demostración se completa mostrando que $\mu^{t_1,\ldots,t_k} = P^{t_1,\ldots,t_k}$ para toda $k, t_1,\ldots,t_k\in[0,1]$. De hecho, como $x^*(t)=x(t)$ para cada t para csi toda x, para cualquier conjunto de Borel $A\subset\mathbb{R}^k$

$$\begin{aligned}\mu^{t_1,\ldots,t_k}(A) &= \mu(\{y \,/\, y\in C, (y(t_1),\ldots,y(t_k))\in A\}) \\ &= P(\{x \,/\, (x^*(t_1),\ldots,x^*(t_k))\in A\}) \\ &= P(\{x \,/\, (x(t_1),\ldots,x(t_k))\in A\}) \\ &= P^{t_1,\ldots,t_k}(A).\end{aligned}$$

Esto completa la demostración.

OBS.- La condición enunciada en el teorema (2.3.2) es conocida como "CONDICION DE KOLMOGOROV". (ver [E]).

CAPITULO 3.- UN EJEMPLO : CONVERGENCIA A LA DISTRIBUCION DE WIENER.

Sea C el espacio de las funciones continuas en el intervalo $[0,1]$. Sea $X_t(\omega), \omega \in C$, el valor de ω en t . Sea $\mathcal{A}$ el álgebra σ mínimo de suboonjuntos de C tal que toda X_t sea medible. En étas condiciones, puede probarse que existe una única medida P de probabilidad sobre $(C, \mathcal{A})$ tal que $\{X_t, 0 \leq t \leq 1\}$ sea un movimiento Browniano. (Ver [E])

Interpretando el teorema (2.3.2) resulta que, toda familia de funciones de distribución finito dimensionales que satisface la condición de Kolmogórov, define una medida μ sobre $(C, \mathcal{A})$, la cual es aditiva σ .

OBS.- En el caso del movimiento Browniano, la medida P resultante es la llamada medida de Wiener. (Ver [E]).

Enunciaremos unteorema en el cual se caracterizan los conceptos mencionados.

TEOREMA 3.1.1.- Existe una única medida W sobre $(C, \mathcal{B}_C)$ con las siguientes propiedades :

i) $W(\{x, x(0)=0\}) = 1$

ii) Si $0 \leq t_1 < t_2 < \cdots < t_n \leq 1$, las variables aleatorias $U_1, U_2, \ldots, U_n$ donde $U_1(x) = x(t_1)$, $U_j(x) = x(t_j) - x(t_{j-1})$ para $1 \leq j \leq n$ son independientes en el espacio de probabilidad $(C, \mathcal{B}_C, W)$.

iii) Si $0 \leq s \leq t \leq 1$, la variable aleatoria V donde $V(x)=x(t)-x(s)$ es normal con media cero y varianza $t-s$.

OBS.- El proceso estocástico formado con las variables aleatorias V_t es el llamado "Movimiento Browniano" y la medida W sobre C es la llamada "Medida de Wiener".

DEMOSTRACION .- Mediante una verificación de la consistencia de los requerimientos concluimos que existe un proceso estocástico $\{\xi_t, 0 \leq t \leq 1\}$ tal que :

i) $\xi_t - \xi_s$ tienen una distribución Normal con media cero y varianza $t-s$ donde $0 \leq s \leq t \leq 1$ y

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

Como $E|\xi_t-\xi_s|^4 = 3(E|\xi_t-\xi_s|^2)^2 = 3|t-s|^2$ resulta, en virtud del teorema (2.3.2) que existe una única medida W en C que satisface (ii) y (iii) . (i) es una consecuencia del hecho que $\xi_{\frac{1}{n}} \to 0$ con probabilidad 1 cuando $n \to \infty$.

Examinaremos una relación muy interesante entre la medida de Wiener en C y las sumas de variables aleatorias que obedecen el teorema límite central. (Ver [E]).

Sean $\xi_{n1}, \xi_{n2}, \ldots, \xi_{nk_n}$ variables aleatorias de una sucesión. En este caso, la distribución de $\xi_{n_1}+\cdots+\xi_{nk_n}$ converge a la distribución Normal.

Usaremos la teoría desarrollada hasta aquí para obtener un teorema general en esa dirección.

LEMA 3.1.1.- Sean $\xi_1, \xi_2, \ldots, \xi_n$ variables aleatorias independientes definidas sobre un espacio de probabilidad $(X, \mathcal{B}, P)$ y definimos $\zeta_k = \sum_{i=1}^{k} \xi_i,\ k=1,2,\ldots,n;\ \zeta_0 = 0$.

Entonces $P(\{\max_{1\le k\le n} |\zeta_k| > 2t\}) \le \alpha^{-1} P(\{|\zeta_n| > t\})$

cualquiera que sea $P(\{|\zeta_n - \zeta_k| \le t\}) \ge \alpha$ para $k=0,1,2,\ldots,n$.

Demostración.- Definamos a los eventos A_k y B_k de la siguiente forma : $A_k = \{|\zeta_1| \le 2t, \ldots, |\zeta_{k-1}| \le 2t, |\zeta_k| > 2t\}$

$B_k = \{|\zeta_n - \zeta_k| \le t\}$, $k=0,1,\ldots,n$.

Entonces $E = \{|\zeta_n| > t\} \supset \bigcup_{k=1}^{n} (A_k \cap B_k)$

$A_1, A_2, \ldots, A_n$ son ajenos y para cada k fija, A_k y B_k son independientes. Por lo tanto

$$P(\{|\zeta_n| > t\}) \ge P\{\bigcup_{k=1}^{n}(A_k \cap B_k)\}$$
$$= \sum_{k=1}^{n} P(A_k)P(B_k) \ge \alpha \sum_{k=1}^{n} P(A_k)$$
$$= \alpha P(\{\max_{1\le k\le n} |\zeta_k| > 2t\}).$$

Esto completa la demostración.

Sea ahora $\{\xi_{ni}\}_{i=1,2,\ldots}$ una sucesión de variables aleatorias definidas sobre el espacio de probabilidad $(X, \mathcal{B}, P)$, que satisface las siguientes condiciones :

i) Para cada n fija, las variables aleatorias $\{\xi_{ni}\}$ son independientes.

ii) $E(\xi_{ni}) = 0$, $V(\xi_{ni}) = b_{ni}$, $\sum_{i=1}^{k_n} b_{ni} = 1$.

iii) Las funciones de distribución F_{ni} de las variables aleatorias ξ_{ni}, $i=1,2,\ldots,k_n$ satisfacen la condición de Lindeberg : $\lim_{n\to\infty} \sum_{i=1}^{k_n} \int_{|v|>\varepsilon} v^2 \, dF_{ni}(v) = 0$.

Sea $\xi_n(t), t \in [0,1]$ una función aleatoria definida de la siguiente forma :

$$\xi_n(t) = S_{nk} + \frac{t - t_{nk}}{t_{nk+1} - t_{nk}} [S_{nk+1} - S_{nk}] , \quad t \in [t_{nk}, t_{nk+1}]$$

donde $t_{n0} = 0$, $S_{n0} = 0$, $t_{nk} = \sum_{i=1}^{k} b_{ni}$, $S_{nk} = \sum_{i=1}^{k} \xi_{ni}$, $k \geq 1$.

Debemos observar que $\xi_n(t) = S_{nk}$ siempre que $t = t_{nk}$ y $\xi_n(t)$ es la línea recta que une a (t_{nk}, S_{nk}) con (t_{nk+1}, S_{nk+1}) en el intervalo $[t_{nk}, t_{nk+1}]$. Por lo tanto, $\xi_n(t)$ es continua con probabilidad uno. Por tanto, le corresponde una medida P_n en el espacio $(C, \mathcal{B}_C)$ conforme a la cual se distribuye el proceso estocástico $\xi_n(t), t \in [0,1]$.

TEOREMA 3.1.2.- Sea $\xi_n(t)$ el proceso descrito arriba y P_n la distribución de $\xi_n(t)$ en el espacio $(C, \mathcal{B}_C)$. Entonces $P_n \Rightarrow W$ cuando $n\to\infty$ donde W es la medida de Wiener en $(C, \mathcal{B}_C)$.

DEMOSTRACION .- Denotemos mediante $\omega(t)$ al proceso del movimiento Browniano y W su distribución en $\mathcal{B}_C$. Debemos establecer primero que las distribuciones finito dimensionales del proceso $\xi_n(t)$ convergen débilmente a las distribuciones finito dimensionales correspondientes al proceso $\omega(t)$. Definamos

$$\xi_n'(t) = \sum_{t_{ni} \leq t} \xi_{ni} \quad , \quad t \in [0,1] \quad , \quad \xi_n'(0) = 0$$

Entonces, para cualquier $\alpha > 0$

$$\mathbb{P}(\{|\xi_n'(t) - \xi_n(t)| > \alpha\}) \leq \mathbb{P}(\{\sup_i |\xi_{ni}| > \alpha\})$$
$$\leq \sum_{i=1}^{k} \mathbb{P}(\{|\xi_{ni}| > \alpha\}) = \sum_{i=1}^{k_n} \int_{|v|>\alpha} dF_{ni}(v)$$
$$\leq \alpha^{-2} \sum_{i=1}^{k_n} \int_{|v|>\alpha} v^2 \, dF_{ni}$$

En virtud de la condición (iii), el último término del lado derecho de la desigualdad anterior tiende a cero cuando $n \to \infty$. Por lo tanto, a fin de probar la convergencia de las distribuciones finito dimensionales de $\xi_n(t)$ es suficiente hacer lo mismo para $\xi_n'(t)$. Como $\omega(t)$ y $\xi_n'(t)$ son procesos con incrementos independientes y $\omega(0) = \xi_n'(0) = 0$ basta con probar la convergencia de la distribución de $\xi_n'(t'') - \xi_n'(t')$ a la distribución de $\omega(t'') - \omega(t')$ para toda $0 \leq t' \leq t'' \leq 1$.

Notemos que

$$\xi_n'(t'') - \xi_n'(t') = \sum_{t' \leq t_{ni} \leq t''} \xi_{ni}$$

que es una suma de variables aleatorias que cumplen la condición de Lindeberg. Entonces, en virtud del teorema límite central (Ver [13] pp. 101-103) la distribución de $\xi_n'(t'') - \xi_n'(t')$ converge a una distribución Normal con media cero y varianza igual a

$$\lim_{n \to \infty} \sum_{t' \leq t_{ni} \leq t''} V(\xi_{ni}) = t'' - t' .$$

En verdad, para cualquier $\varepsilon > 0$

$$\Big|\sum_{t' \le t_{ni} \le t''} V(\xi_{ni}) - (t'' - t')\Big| \leq \max_i b_{ni} \leq \varepsilon^2 + \max_i \int_{|v|>\varepsilon} v^2 \, dF_{ni}(v)$$

$$\leq \varepsilon^2 + \sum_{i=1}^{k_n} \int_{|v|>\varepsilon} v^2 \, dF_{ni}(v)$$

y el último término tiende a cero cuando $n \to \infty$. Esto completa la demostración de la convergencia de las distribuciones finito dimensionales de $\xi_n(t)$.

Debemos mostrar ahora que la sucesión de medidas $\{P_n\}$ en $\mathcal{B}_C$ es relativamente compacta. Sea $\Gamma = \{P_n;\ n = 1, 2, \ldots\}$ Esta Γ satisface la condición (i) del teorema(2.2.1) ya que $P\{\xi_n(0) = 0\} = 1$.

A fin de probar la condición (ii) del teorema mencionado es equivalente demostrar que para todo $\varepsilon > 0$

$$\lim_{h \to 0} \overline{\lim_{n \to \infty}} \; P\Big\{\sup_{|t''-t'| \le h} |\xi_n(t'') - \xi_n(t')| > \varepsilon\Big\} = 0 \quad \cdots \quad (3.1)$$

Como $\sup\limits_{|t''-t'| \le h} |\xi_n(t'') - \xi_n(t')| \leq 2 \sup\limits_k \sup\limits_{kh \le t \le (k+1)h} |\xi_n(t) - \xi_n(kh)|$

$$\leq 4 \sup_k \sup_{kh \le t \le (k+1)h} |\xi_n(t) - \xi_n(kh)|$$

resulta que

$$P\Big\{\sup_{|t''-t'| \le h} |\xi_n(t'') - \xi_n(t')| > \varepsilon\Big\} \leq \sum_{kh < 1} P\Big\{\sup_{kh \le t \le (k+1)h} |\xi_n(t) - \xi_n(kh)| > \frac{\varepsilon}{4}\Big\} \cdots (3.2)$$

Pero $\sup\limits_{kh \le t \le (k+1)h} |\xi_n(t) - \xi_n(kh)| \leq 2 \sup\limits_{j_{nk} \le r \le j_{n,k+1}} \Big|\sum_{i=j_{nk}}^{r} \xi_{nj}\Big| \cdots (3.3)$ donde $j_{nk} = \max\{j, t_n$

Usando la desigualdad de Tchebyshev tenemos

$$P\Big\{\Big|\sum_{j=j_{nk}}^{j_{n,k+1}} \xi_{nj}\Big| > \varepsilon/8\Big\} \leq \frac{64}{\varepsilon^2} \sum_{j=j_{nk}}^{j_{n,k+1}} V(\xi_{nj}) \longrightarrow \frac{64h}{\varepsilon^2} \quad \cdots (3.4)$$

cuando $n \to \infty$. Las ecuaciones (3.3) y (3.4) y el lema(3.1.1) implican $\overline{\lim\limits_{n \to \infty}} \; P\Big\{\sup\limits_{kh \le t \le (k+1)h} |\xi_n(t) - \xi_n(kh)| > \frac{\varepsilon}{4}\Big\}$

$$\leq \overline{\lim_{n\to\infty}} \frac{1}{(1-64h/\varepsilon^2)} \sup_k P\left\{\Big|\sum_{j=j_{nk}}^{j_{n,k+1}} \xi_{nj}\Big| > \frac{\varepsilon}{8}\right\}$$

$$= \overline{\lim_{n\to\infty}} \frac{1}{(1-64h/\varepsilon^2)} \sup_k P\left\{|\xi_n(t_{nj_{n,k+1}}) - \xi_n(t_{nj_{nk}})| > \frac{\varepsilon}{8}\right\}.$$

De la convergencia de las distribuciones finito dimensionales tenemos

$$\overline{\lim_{n\to\infty}} P\left\{|\xi_n(t_{nj_{n,k+1}}) - \xi_n(t_{nj_{nk}})| > \frac{\varepsilon}{8}\right\} = \frac{1}{\sqrt{2\pi h}} \int_{|v|>\frac{\varepsilon}{8}} \exp\left[-\frac{v^2}{2h}\right] dv$$

para toda K . En virtud de (3.2) obtenemos

$$\overline{\lim_{n\to\infty}} P\left\{\sup_{|t''-t'|\leq h} |\xi_n(t'') - \xi_n(t')| > \varepsilon\right\}$$

$$\leq \sum_{kh<1} \frac{1}{(1-64h/\varepsilon^2)} \cdot \frac{1}{\sqrt{2\pi}} \int_{|v|>\varepsilon/8\sqrt{h}} \exp\left[-\frac{v^2}{2}\right] dv$$

$$\leq \frac{1}{\sqrt{6\pi}} \cdot \frac{1}{(1-64h/\varepsilon^2)} \cdot \frac{1}{h} \int_{|v|>\varepsilon/8\sqrt{h}} \exp\left[-\frac{v^2}{2}\right] dv$$

Como $\frac{1}{h}\int_{|v|>\varepsilon/8\sqrt{h}} \exp\left[-\frac{v^2}{2}\right] dv \leq \frac{1}{\varepsilon^2}\int_{|v|>\varepsilon/8\sqrt{h}} v^2 \exp\left[-\frac{v^2}{2}\right] dv \to 0$

cuando $h\to 0$, hemos probado (3.1). En virud del teorema (2.2.1), Γ es relativamente compacta. Por lo tanto $P_n \Rightarrow W$ en $(C, \mathcal{B}_C)$.

F I N .

APENDICE [A] .- ESPACIOS METRICOS.

Dado un conjunto R , tomemos dos elementos $x, y \in R$; en estas condiciones tenemos la siguiente :

DEFINICION A.1.- Llamamos DISTANCIA a una función real no-negativa que tiene como dominio al producto cartesiano $R \times R$ y que verifica las tres siguientes condiciones :

i) $\rho(x,y) = 0$ si, y solo si $x = y$

ii) (Axioma de simetría) $\rho(x,y) = \rho(y,x)$

iii) (Axioma triangular) $\rho(x,z) \leq \rho(x,y) + \rho(y,z)$ con $z \in R$

DEFINICION A.2.- Un ESPACIO METRICO es la pareja (R , ρ), compuesta por un conjunto R no vacío y de una distancia ρ definida para elementos del conjunto R .

Al espacio métrico (R , ρ) lo denotaremos solamente con la letra R .

Designemos por R y R' dos espacios métricos y tomemos una función f del espacio R al R' .

DEFINICION A.3.- La función f se dice <u>Continua</u> en el punto $x_0 \in R$, si para cada $\epsilon > 0$ existe una $\delta > 0$ tal que para todo $x \in R$ que cumpla la desigualdad $\rho(x_0,x)<\delta$, se verifica la desigualdad $\rho'(f(x_0), f(x))<\epsilon$.

Nota : Si la función f es continua en todos los puntos del espacio R , decimos que la funcion es continua sobre R .

Definiremos algunos conceptos relacionados con los espacios métricos, que serán utilizados en forma amplia.

DEFINICION A.4.- Una BOLA ABIERTA $S(x_0,\varepsilon)$ en el espacio métrico R, es el conjunto de puntos $x \in R$ que verifican la condición $\rho(x,x_0) < \varepsilon$. (El punto fijo x_0 se llama <u>Centro</u>, y el número ε <u>Radio</u> de la bola).

Nota : Si la condición cumplida es $\rho(x,x_0) \leq \varepsilon$, llamamos BOLA CERRADA y la denotaremos como $S[x_0,\varepsilon]$.

Para efectos de este trabajo , uno de los conceptos principales es el de LIMITE. Dicho concepto será introducido a partir de sucesiones.

Tomemos un espacio métrico R; tenemos la siguiente :

DEFINICION A.5.- Una SUCESION de puntos en R , es una función cuyo dominio es el conjunto de los números naturales y su contradominio es un subconjunto de R .

Sea $\{x_n\}$ una sucesión de puntos en R .

DEFINICION A.6.- Decimos que $\{x_n\}$ CONVERGE al punto x cuando, toda vecindad $S(x,\varepsilon)$ del punto x contiene todos los puntos x_n empezando desde alguno; es decir, cuando a cada $\varepsilon>0$ le corresponde un índice N_ε tal que $S(x,\varepsilon)$ contiene a todos los puntos x_n para $n>N_\varepsilon$. El punto x se llama LIMITE.

En base a la definición (A.6) tenemos que :

La sucesión $\{x_n\}$ converge a x si, y solo si $\lim_{n\to\infty} \rho(x,x_n)=0$.

Obs.- i) Ninguna sucesión puede tener dos límites distintos. ii) Si la sucesión $\{x_n\}$ converge al punto x , toda subsucesión $\{x_{n_k}\}$ de la sucesión $\{x_n\}$ converge al mismo punto.

Dejaremos los conceptos de limite y convergencia para secciones posteriores donde cobrarán la importancia requerida para los objetivos del trabajo. Por ahora continuamos con definiciones de conjuntos en espacios métricos.

Denotando con R a un espacio métrico arbitrario, sean A y B dos conjuntos que pertenecen a dicho espacio.

DEFINICION A.7.- Un punto $a \in A$ se denomina PUNTO DE ADHERENCIA del conjunto A , si cualquier vecindad suya contiene al menos un punto de A . La totalidad de los puntos de adherencia del conjunto A se denota mediante $\bar{A}$ y se llama la CERRADURA del conjunto A .

DEFINICION A.8.- El conjunto A se llama CERRADO cuando coincide con su cerradura: $A = \bar{A}$. En otras palabras, el conjunto se llama cerrado si contiene todos sus puntos de adherencia.

DEFINICION A.9.- Un punto $x \in A$ se llama PUNTO INTERIOR del conjunto A , cuando hay una vecindad $S(x,\epsilon)$ de este punto, totalmente contenida en A . Un conjunto, tal que todos sus puntos son interiores, se llama ABIERTO.

DEFINICION A.10.- El conjunto A se dice DENSO en B, cuando $\bar{A} \supset B$.

Nota : En particular, el conjunto A es SIEMPRE DENSO (en el espacio R) cuando su cerradura $\bar{A}$ coincide con todo el espacio R.

DEFINICION A.11.- Un espacio métrico que tiene un conjunto numerable siempre denso, se llama SEPARABLE.

Ahora estudiaremos la propiedad de COMPLETEZ de un espacio métrico, la cual juega un papel importante en cuanto al concepto de convergencia se refiere.

Sea R un espacio métrico y $\{x_n\}$ una sucesión en R.

DEFINICION A.12.- Se dice que $\{x_n\}$ es una sucesión FUNDAMENTAL, cuando para cualquier $\varepsilon>0$ exista un número N_ε tal que $\rho(x_n, x_m)<\varepsilon$ para cualesquiera $n,m>N_\varepsilon$.

DEFINICION A.13.- Un espacio métrico se llama COMPLETO cuando toda sucesión fundamental en él, converge.

Cabe destacar el hecho de que si un espacio métrico no es completo puede ser incluido de cierta manera (de manera única) en un espacio completo.

DEFINICION A.14.- Un espacio métrico completo R^* se llama COMPLETEZ del espacio R si :

i) R es un subespacio de R^*.

ii) R es siempre denso en R^*; esto es, $R^* = \overline{R}$.

TEOREMA A.1.- Todo espacio métrico R posee una completez, la cual es única; o bien, cualquier otra es una transformación isométrica de dicha completez.

Nota : Para demostración, ver [9] pp. 76-79.

Para terminar, estudiaremos el concepto de Compacidad.

En el caso de los espacios métricos, la compacidad está estrechamente ligada al concepto de acotación total que introducimos a continuación :

Sea A un conjunto del espacio métrico R , y ε un número positivo; en estas condiciones tenemos:

DEFINICION A.15.- Se dice que el conjunto $B \subset R$ es una ε-red de A , cuando para todo punto $x \in A$ existe al menos un punto $b \in B$ tal que $\rho(x,b) < \varepsilon$.

DEFINICION A.16.- Un conjunto A se llama TOTALMENTE ACOTADO, cuando cualquiera que sea $\varepsilon > 0$ existe una ε-red finita.

De la definición (A.16) se desprende que todo espacio métrico totalmente acotado, es separable. En efecto, construyamos para toda n una $\frac{1}{n}$ -red finita de R . La unión, respecto a n , de todas esas redes representa un conjunto numerable siempre denso en R .

DEFINICION A.17.- Una clase $\{K_i\}$ de conjuntos se llama CUBIERTA del espacio métrico R, cuando $R \subset \bigcup_i K_i$.

Si una parte $\{K_\alpha\}$ de la cubierta $\{K_i\}$ tambien constituye una cubierta del espacio R , se dice que $\{K_\alpha\}$ es una <u>subcubierta</u> de la cubierta $\{K_i\}$.

DEFINICION A.18.- Un espacio métrico R se llama COMPACTO, cuando cualquier cubierta abierta suya, contiene una subcubierta finita.

DEFINICION A.19.- Un conjunto A , que pertenece al espacio métrico R , se llama RELATIVAMENTE COMPACTO cuando su cerradura es compacta.

Destacaremos tres resultados, importantes en cuanto a compacidad en espacios métricos se refiere. La demostración de los mismos puede verse en [9] .

TEOREMA A.2.- Todo espacio métrico R compacto numerable, es totalmente acotado.

TEOREMA A .3.- Para que un espacio métrico R sea compacto es necesario y suficiente que sea :

i) Totalmente acotado.

Y

ii) Completo.

TEOREMA A.4.- Para que un conjunto $A \subset R$,(con R completo) sea relativamente compacto, es necesario y suficiente que sea totalmente acotado.

APENDICE [B] .- EL ESPACIO $C_{[0,1]}$.

Dentro del esquema que hemos mencionado desde el capítulo introductorio, las funciones que consideraremos y que además satisfacen las propiedades que determinan al espacio $C_{[0,1]}$ representan un caso de interés especial para la teoría de Procesos Estocásticos. Como adelanto, baste mencionar que las trayectorias del movimiento Browniano en el intervalo $[0,1]$ pertenecen al espacio $C_{[0,1]}$. Entre otras, las razones anteriores nos llevan a realizar un estudio (desde el punto de vista Topológico) detallado de dicho espacio.

DEFINICION B.1.- El conjunto de todas las funciones reales continuas, definidas en el intervalo $[0,1]$ forman el conjunto $C_{[0,1]}$.

DEFINICION B.2.- En el conjunto $C_{[0,1]}$ se define la DISTANCIA entre funciones como sigue :

Sean $f,g \in C_{[0,1]}$, entonces

$$\rho(f,g) = \max_{t\in[0,1]} |g(t) - f(t)|.$$

OBS.- Los axiomas (i), (ii), (iii) de distancia se comprueban directamente.

DEFINICON B.3.- El conjunto $C_{[0,1]}$ junto con la distancia definida en (B.2), forman el Espacio Métrico llamado $C_{[0,1]}$. (Por comodidad lo denotaremos simplemente como C).

Veamos algunos conceptos definidos en C .

--Una bola cerrada representa un CONJUNTO CERRADO.

En particular, sea $t\in[0,1]$ entonces, se tiene que el conjunto de funciones f del espacio C , que verifican la condición $|f(t)|\leq K$ (K un número fijo) es cerrado.

El conjunto de funciones $f\in C$ que verifican la condición $|f(t)|<K$ $(t\in[0,1]$) no es cerrado; su cerradura es el conjunto de funcines arriba señalado.

El conjunto de funciones continuas sobre $[0,1]$ que verifican la condición $f(t)<g(t)$, donde g es una función continua determinada, representa un subconjunto abierto del espacio C .

PROPOSICION B.1.- El espacio C es separable.

DEMOSTRACION.- Se tiene que mostrar un conjunto numerable siempre denso en C ; para tal efecto, basta contar con el conjunto de todos los polinomios de coeficientes racionales.

PROPOSICION B.2.- El espacio C es completo.

DEMOSTRACION.- Sea $\{X_n(t)\}$ una sucesión fundamental de elementos de C . Esto significa que, para todo $\varepsilon>0$ existe un

N tal que $|x_n(t)-x_m(t)|<\varepsilon$ siempre que $n,m>N$ y para todo $t\in[0,1]$. De aquí se deduce que la sucesión $\{x_n(t)\}$ converge uniformemente. En este caso, como se sabe, su límite $x(t)$ será una función continua. Haciendo tender $m\to\infty$ en la desigualdad anterior, obtendremos $|x_n(t)-x(t)|<\varepsilon$ para todo $t\in[0,1]$ y todo $n>N$, esto significa precisamente que $\{x_n(t)\}$ converge a $x(t)$ en el sentido de la métrica del espacio C .

La demostración de la compacidad de un conjunto de un espacio métrico, es un problema que encontramos con bastante frecuencia en el Análisis. Para los conjuntos situados en un espacio concreto, se pueden dar criterios especiales de compacidad que resulten más cómodos para la aplicación práctica.

Para los subconjuntos del espacio C , un criterio importante (empleado con frecuencia) de compacidad relativa lo ofrece el así llamado TEOREMA DE ARZELÁ.

Para enunciarlo, necesitamos definir dos conceptos.

DEFINICION B.4.- Una familia Φ de funciones definidas en el intervalo $[0,1]$ se llama EQUIACOTADA cuando existe un número K tal que para todo $t\in[0,1]$ y toda $f\in\Phi$, $|f(t)|<K$.

DEFINICION B.5.- Una familia Φ de funciones definidas en el intervalo $[0,1]$ se llama EQUICONTINUA cuando para cada

$\varepsilon>0$ hay una $\delta>0$ tal que para toda $g\in\Phi$ y para todo par $t_1,t_2\in[0,1]$ tal que $\rho(t_1,t_2)<\delta$, sucede que $|g(t_1)-g(t_2)|<\varepsilon$.

TEOREMA B.1.- (DE ARZELÁ) Para que una familia Φ de funciones continuas, definidas en el intervalo $[0,1]$ sea relativamente compacta en C , es necesario y suficiente que esta familia sea equiacotada y equicontinua.

Nota : La demostración está en [14] .

Por último, introduciremos un concepto que será de gran ayuda para efectos del trabajo.

DEFINICION B.6.- EL MÓDULO DE CONTINUIDAD de un elemento $x\in C$ se define como

$$\omega_x(\delta)=\omega(x,\delta)=\sup_{|s-t|<\delta}|x(s)-x(t)| \quad , \; 0<\delta<1 .$$

Tomando en cuenta que $|\omega_x(\delta)-\omega_y(\delta)|\leq 2\rho(x(t),y(t))$, $\omega_x(\delta)$ es, para una δ fija, continua en x , además, como con elementos de C es uniformemente continua, resulta que

$$\lim_{\delta\to 0}\omega_x(\delta)=0 \quad , \; x\in C .$$

Usando el módulo de continuidad definido en (B.6), tenemos un criterio para determinar la compacidad relativa de los subconjuntos de C .

TEOREMA B.2.- (ARZELÁ-ASCOLI) Un subconjunto A de C tiene cerradura compacta si, y solo si

$$\sup_{x\in A}|x(0)|<\infty \quad \text{y} \quad \lim_{\delta\to 0}\sup_{x\in A}\omega_x(\delta)=0 .$$

Nota : Para demostración ver [13] .

APENDICE [C] .- INTEGRACION, INTEGRAL DE LEBESGUE.

El concepto de la integral de Riemann que se maneja en los cursos de Cálculo y de Análisis se aplica solamente a funciones continuas y en muy pocas ocasiones a funciones discontinuas, siendo éstas últimas de modo tal quela cantidad de discontinuidades es "pequeña". Para funciones medibles que pueden ser dicontinuas en todo punto donde estén definidas, la construcción de Riemann de la integral no es válida (recordemos que estas funciones pueden estar definidas en un conjunto abstracto donde el concepto de continuidad carece de sentido).

La idea principal de la integral de Lebesgue, consiste en que los puntos X se agrupan deacuerdo a la proximidad de los valores de la función en dichos puntos. Esto ofrece la posibilidad de extender el concepto de integral a una clase muy amplia de funciones.

En lo que sigue se considerará, mientras no se diga lo contrario, una medida μ aditiva σ definida sobre un álgebra $\mathcal{B}$ boreliana de conjuntos con unidad Ω . Todos los conjuntos $A \subset \Omega$ considerados, se supondrán medibles al igual que las funciones $f(x)$ definidas para $x \in \Omega$.

I.- Integral de Lebesgue para funciones simples.

Introduciremos primero el concepto para las funciones que se conocen como simples.

Sea f una función simple que toma los valores $y_0, y_1, \dots, y_n, \dots$ $y_i \neq y_j$ para $i \neq j$; es decir

$f = \sum y_i I_{A_i}$ sobre un espacio de medida $(\Omega, \mathcal{B}, \mu)$.

Es natural definir la integral de la función f en el conjunto A , mediante la igualdad

$$\int_A f(x)\, d\mu = \sum_n y_n \mu(A_n) \quad \text{donde} \quad A_n = \{x \cdot, x \in A, f(x) = y_n\}$$

si la serie que figura en el miembro derecho converge. De esta forma llegamos a la siguiente definición (en la que se postula de antemano la convergencia absoluta de esta serie).

DEFINICION C.1.- Una función simple se llama INTEGRABLE (respecto a la medida μ) en un conjunto A , cuando la serie $\sum y_n \mu(A_n)$ converge absolutamente.

Si f es integrable, a la suma de la serie referida se le le llama la integral de f en el conjunto A . El ejemplo típico de una función simple integrable es la función indicadora de un conjunto E medible de medida finita; resulta que $\int I_E \, d\mu = \int_E d\mu = \mu(E)$.

Indicaremos algunas propiedades de la integral de Lebesgue de funciones simples :

PROPIEDAD (A) .- $\int_A f(x)\, d\mu + \int_A g(x)\, d\mu = \int_A [f(x)+g(x)]\, d\mu$,

con la particularidad de que la existencia de las integrales del miembro izquierdo implica la existencia de la integral del miembro derecho.

PROPIEDAD (B) .- Para cualquier constante K :

$$K \int_A f(x)\, d\mu = \int_A [K f(x)]\, d\mu$$

donde la existencia de la integral del miembro izquierdo implica la existencia de la integral del miembro derecho.

PROPIEDAD (C) .- Una función simple, acotada en un conjunto A, es integrable en A y además si $|f(x)| \leq M$ en A, se tiene que $\left| \int_A f(x)\, d\mu \right| \leq M \cdot \mu(A)$.

II.- Integral de Lebesgue en conjuntos de medida finita.

DEFINICION C.2.- Una función $f(x)$ se llama INTEGRABLE en un cojunto A cuando exista una sucesión de funciones simples integrables en A convergente uniformemente hacia f .

El límite $\lim_{n\to\infty} \int_A f_n(x)\, d\mu$ se denotará mediante el símbolo $\int_A f(x)\, d\mu$, y llamado INTEGRAL de la función f en el conjunto A .

Establezcamos las propiedades fundamentales de la integral de Lebesgue.

Una consecuencia directa de la definición es que

(1) $$\int_A 1 \cdot d\mu = \mu(A).$$

Para cualquier constante k

(2) $$\int_A \{k\, f(x)\}\, d\mu = k \int_A f(x)\, d\mu$$

donde la existencia de la integral del miembro izquierdo implica la existencia de la integral del miembro derecho.

La siguiente propiedad se deduce del paso al límite de la propiedad (B) para las integrales de funciones simples.

(3) $$\int_A f(x)\, d\mu + \int_A g(x)\, d\mu = \int_A \{f(x) + g(x)\}\, d\mu$$

(4) Una función f acotada en A, es integrable en A.

(5) Si $f(x) \geq 0$, se tiene que $\int_A f(x)\, d\mu \geq 0$ (Sup. que la integral existe).

(6) Si $f(x) \geq g(x)$ se tiene que $\int_A f(x)\, d\mu \geq \int_A g(x)\, d\mu$.

(7) Si $m \leq f \leq M$ para todos (o casi todos) los $x \in A$, tendremos

$$m \cdot \mu(A) \leq \int_A f(x)\, d\mu \leq M \cdot \mu(A)$$

(8) Si $\mu(A) = 0$, se tiene que $\int_A f(x)\, d\mu = 0$.

APENDICE [D] .- CONCEPTOS GENERALES DE LA TEORIA DE PROBABILIDAD.

I.- INTRODUCCION.

El estudio de un fenómeno real intenta llegar a una explicación satisfactoria acerca de cómo este se presenta y comporta. Usualmente no se intenta explicar "porqué" ocurre tal o cual fenómeno, lo que tendría un interés de índole tal vez filosófico, sino más bien "cómo ocurre"; es decir, trata de encontrar algún sistema de reglas que gobiernen el comportamiento del fenómeno bajo estudio y que permitan entonces, predecir de manera confiable lo que sucederá cuando se le observe en determinadas circunstancias.

El estudio de los fenómenos aleatorios sigue la misma linea. No interesa por lo general la naturaleza intrínseca del azar, sino la forma en que este se manifiesta en situaciones concretas. Como quiera que sea, podemos observar que un fenómeno aleatorio se presenta en forma de cierto resultado cuando lo que se persige es descubrir algún conjunto de reglas que permitan formular predicciones relativamente confiables respecto a lo que se puede presentar al observar dicho fenómeno en el futuro.

La teoria de Probabilidad se ocupa de establecer las reglas que gobiernan los fenómenos aleatorios. La teoria de Probabilidad no proporciona recetas mágicas para adivinar el curso de acontecimientos aleatorios, sin embargo, es un intento bastante exitoso de sistematizar las características de los fenómenos aleatoriosy, obtener así toda la información posible sobre su realización.

II.- ESPACIOS DE PROBABILIDAD.

Los conceptos fundamentales de la teoria de Probabilidad son los de EVENTO y PROBABILIDAD. Si consideramos a Ω como el conjunto cuyos puntos corresponden a los resultados posibles de un experimento, ciertos subconjuntos de Ω se llamarán Eventos y les asignaremos su Probabilidad de ocurrencia.

Intuitivamente, $A \subset \Omega$ es un evento si la pregunta ¿$\omega \in A$? puede ser contestada en forma afirmativa o negativa cuando el experimento ha sido realizado. Ahora bien, si es que podemos contestar a la pregunta ¿$\omega \in A$?, nos gustaría tambien contestar ¿$\omega \notin A$? es decir ¿$\omega \in A^c$?; y si para cada $i=1,2,\dots,n$ podemos responder a la pregunta ¿$\omega \in A_i$?, podremos contestar ¿$\omega \in \bigcup_{i=1}^{n} A_i$? , y de forma similar ¿$\omega \in \bigcap_{i=1}^{n} A_i$?, y como la pregunta ¿$\omega \in \Omega$? es siempre positiva, resulta que Ω es también un evento.

Así, hemos visto que los eventos tienen cierta estructura conjuntista, la cual motiva al empleo de la estructura definida anteriormente como ALGEBRA σ ; de forma precisa, utilizaremos el álgebra σ $\mathfrak{F}$ generada por el conjunto Ω , formando un espacio medible $(\Omega, \mathfrak{F})$ asociado al fenómeno aleatorio cuyos resultados posibles son los elementos de Ω .

Sea μ una medido definida sobre $\mathfrak{F}$; en estas condiciones tenemos la siguiente :

DEFINICION D.1.- La medida μ es una <u>medida de probabilidad</u> si para todo elemento de $\mathfrak{F}$ se tiene que

i) $\mu(\Omega) = 1$.

ii) $\mu(A) \geqslant 0$.

iii) $\mu(A \cup B) = \mu(A) + \mu(B)$ si $A \cap B = \phi$.

Nota.- Cuando hagamos mención de una medida de probabilidad la denotaremos mediante el símbolo $\mathfrak{P}$.

De ahí que $\mathfrak{P}$ sea una función cuyo dominio es el álgebra σ $\mathfrak{F}$ y su contradominio el intervalo $[0,1]$.

Con los tres axiomas de probabilidad (definición D.1) podemos calcular las probabilidades de algunos eventos. Este es un punto importante porque en la práctica significa que necesitamos comenzar con las probabilidades para una subcolección de eventos y, encontrar el resto haciendo uso de los axiomas.

Propiedades de P :

i) $P(A) = 1 - P(A^c)$

ii) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

iii) $P(\emptyset) = 0$

iv) Si $A \subset B$ ent. $P(A) \leq P(B)$.

Sea A un evento con $P(A) > 0$. Para cualquier evento tenemos la siguiente :

DEFINICION D.2.- La PROBABILIDAD CONDICIONAL de B "dado" A es $P(B|A) = \frac{P(B \cap A)}{P(A)}$.

La interpretación de $P(B|A)$ es que la ocurrencia del evento B está condicionada a la información de la ocurrencia del evento A .

DEFINICION D.3.- Decimos que dos eventos A y B son INDEPENDIENTES si $P(B|A) = P(B)$; es decir, si $P(A \cap B) = P(A) \cdot P(B)$.

DEFINICION D.4.- Una colección finita $\{A_i\}_{i=1}^{n}$ de eventos se dice INDEPENDIENTE si toda subcolección $\{A_{k_1}, \ldots, A_{k_m}\}$ tiene la propiedad $P\left(\bigcap_{j=1}^{m} A_{k_j} \right) = \prod_{j=1}^{m} P(A_{k_j})$.

Nota : Los eventos en una colección independiente se llaman MUTUAMENTE INDEPENDIENTES.

Una vez definida la medida de probabilidad, podemos definir el concepto de espacio de probabilidad ya que, como veremos, todo fenómeno aleatorio tiene asociado un espacio de probabilidad.

DEFINICION D.5.- Un ESPACIO DE PROBABILIDAD es la terna $(\Omega, \mathfrak{F}, \mathbb{P})$ donde : i) $\Omega \neq \phi$

ii) $\mathfrak{F}$ es un álgebra σ de conjuntos de Ω

iii) $\mathbb{P}$ es una medida de probabilidad.

III.- VARIABLE ALEATORIA.

Si $(\Omega, \mathfrak{F})$ es un espacio medible y X es una función medible de $(\Omega, \mathfrak{F})$ sobre $(\mathbb{R}, \mathfrak{B})$, decimos que X es una función medible $\mathfrak{F}$ de valor real.

DEFINICION D.6.- Si una medida $\mathbb{P}$ de probabilidad es definida sobre $(\Omega, \mathfrak{F})$, decimos entonces que X es una VARIABLE ALEATORIA de valor real definida sobre el espacio $(\Omega, \mathfrak{F}, \mathbb{P})$.

OBS.- Como $\mathfrak{B}$ está generado por la colección de los intervalos semicerrados de la recta, para que X sea variable aleatoria es suficiente que todo conjunto de la forma $\{\omega; X(\omega) \leq a\}$ esté en $\mathfrak{F}$; esto es, que sea un evento.

La probabilidad del evento $\{\omega; X(\omega) \leq a\}$ define una función no-negativa $\mathbb{P}_X(a)$, $-\infty < a < \infty$.

DEFINICION D.7.- La función $\mathbb{P}_X(a)$ es llamada FUNCION DE DISTRIBUCION DE PROBABILIDAD de la variable aleatoria X .

La función de distribución tiene dos propiedades importantes : 1) $\mathbb{P}_X$ es una función monótona no-decreciente con

$$\lim_{a \to \infty} \mathbb{P}_X(a) = 1 \quad \text{y} \quad \lim_{a \to -\infty} \mathbb{P}_X(a) = 0 .$$

2) P_X es continua por la derecha; es decir,

$$\lim_{\varepsilon \to 0} P_X(a+\varepsilon) = P_X(a).$$

En forma concreta, cualquier pregunta probabilista concerniente a la variable aleatoria X puede responderse en forma directa una vez que se conozca la función de distribución P_X .

Existen preguntas que no son estrictamente probabilistas concernientes a la variable aleatoria X y su respuesta es un número que revela información acerca del fenómeno aleatorio; dichas respuestas se llaman Características numéricas, las cuales definiremos a continuación.

DEFINICION D. 8.- la ESPERANZA de una variable aleatoria se define como $E[X] = \int_{\Omega} x \, dP_X(x)$.

De esta definición (D.8) observamos que $E[X]$ es un operador lineal.

DEFINICION D.9.- El k-ésimo momento de una variable aleatoria X se define como $E[X^k]$.

OBS.- Los momentos más útiles son el primero y el segundo.

DEFINICION D.10.- Sea $\mu = E[X]$. Al segundo momento de $X-\mu$ denotado como $E[(X-\mu)^2] = \sigma^2$ lo llamamos la VARIANZA.

DEFINICION D.11.- La FUNCION CARACTERISTICA de una variable aleatoria X se define como $\varphi_X(t) = E(e^{itx})$.

Si X tiene función de distribución P_X , resulta que

$$\varphi_X(t) = \int_\Omega e^{itx}\, dP_X$$

PROPOSICION D.1.- La FORMULA DE INVERSION de la teoria integral de Fourier da lugar a $P_X = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\varphi_X(t)\, dt$

OBS.- Con esto resulta que la función de distribución de probabilidad P_X está determinada completamente por la función característica, y es de gran utilidad como veremos más adelante.

Supongamos que $X_1, \dots, X_n$ son n variables aleatorias definidas sobre el mismo espacio de probabilidad, y definimos a $\bar{X} = (X_1, X_2, \dots, X_n)$ como una función medible de $(\Omega, \mathfrak{F})$ sobre $(\mathbb{R}^n, \mathfrak{B}^n)$ que recibe el nombre de VECTOR ALEATORIO.

DEFINICION D.12.- La función $P_{\bar{X}}(\bar{a}) = P_{\bar{X}}(a_1, a_2, \dots, a_n) = P(\{\omega;\, X_i(\omega) \leq a_i\})\ \forall i$ es llamada FUNCION DE DISTRIBUCION CONJUNTA DE $\bar{X}$.

OBS.- Al mismo tiempo, la relación $P_{\bar{X}}(A) = P(\{\omega;\, \bar{X}(\omega) \in A\})\ A \subset \mathbb{R}^n$ define una medida de probabilidad que llamaremos "de Borel".

DEFINICION D.13.- De acuerdo a la definición (D.3) y (D.4) decimos que las variables $X_1, \dots, X_n$ son INDEPENDIENTES si su función de distribución conjunta tiene la forma :

$$P_{\bar{X}}(a_1, \dots, a_n) = \prod_{i=1}^{n} P_{X_i}(a_i)$$

PROPOSICION D.2.- Si $X_1, \dots, X_n$ son variables aleatorias independientes entonces : i) $E\left(\prod_{i=1}^{n} X_i\right) = \prod_{i=1}^{n} E(X_i)$

ii) $\varphi_{\sum X_i}(t) = \prod_{i=1}^{n} \varphi_{X_i}(t)$.

Daremos a conocer un tipo importante de variable aleatoria :

DEFINICIØN D.14.- Una variable aleatoria se dice GAUSSIANA si su función característica tiene la forma

$$\varphi_X(t) = \exp\left[it\mu - \tfrac{1}{2}\sigma^2 t^2\right]$$

donde $\mu \in \mathbb{R}$ y $\sigma^2 > 0$; si $\sigma^2=0$ ent. $\varphi_X(t) = e^{it\mu}$ y se tiene que $X(\omega)=\mu$ c.t.p.

La razón por la cual introducimos este tipo de variable aleatoria es porque la suma de variables aleatorias que tienen la misma distribución, resulta ser una variable aleatoria gaussiana. Como en la práctica muchos fenómenos aleatorios resultan de la suma de un numero grande de fluctuaciones independientes, esperamos variables de tipo Gaussianas.(Teorema del Límite Central).

Veamos ahora un tipo especial de sucesiones, las sucesiones de variables aleatorias $\{X_n\}$.

Las propiedades de tales sucesiones están especificadas por la función de distribución :

$$P_n(x_1,x_2,\dots,x_n) = P\{X_1 \le x_1, X_2 \le x_2, \dots, X_n \le x_n\} \text{ para toda } n.$$

DEFINICION D.15.- Las DISTRIBUCIONES FINITO DIMENSINALES de la sucesión $\{X_n\}$, resultan ser la familia de funciones de distribución $\{P_n\}$.

OBS.- En el caso de que toda variable aleatoria de la sucesión $\{X_n\}$ tenga la misma distribición P uno-dimensional, y las variables aleatorias sean independientes, se tiene :

$$P_n(x_1,\ldots,x_n) = \prod_{i=1}^{n} P(x_i) .$$

IV.- CONVERGENCIA EN PROBABILIDAD.

Los conceptos de convergencia asociados a variables aleatorias son importantes no solo en el cálculo de probabilidades sino, que desempeñan el papel de conectar la estructura axiomática de probabilidad con las observaciones empíricas. Por una variable aleatoria $X(\omega)$ $\omega\in\Omega$, entendemos que solo podemos observar un valor de X en el punto, digamos $X(\omega_0)$. A fin de obtener propiedades estadísticas relacionadas a X tenemos que repetir las observaciones sobre una sucesión de variables aleatorias que comparten atributos con X .

Sea $\{X_n\}$ una sucesión de variables aleatorias definidas sobre $(\Omega, \mathcal{F}, P)$.

DEFINICION D.16.- Se dice que la sucesión $\{X_n\}$ converge en PROBABILIDAD a la variavle aleatoria X si :

$$\lim_{n\to\infty} P(\{\omega, |X_n(\omega) - X(\omega)| \geq \varepsilon\}) \to 0$$

Este tipo de convergencia, en teoria de la medida, recibe el nombre de CONVERGENCIA EN MEDIDA.

Los distintos tipos de convergencia enunciados para funciones medibles, se aplican a variables aleatorias. Distinguimos los diferentes tipos de convergencia para $\{X_n\}$ como se muestra en la tabla siguiente :

$\{X_n\}$ CONVERGE :

TERMINOLOGIA	DEFINICION	NOTACION
Casi Seguramente (ó con probabilidad uno)	Existe un evento A tal que $P(A)=0$ y, para todo $\omega \notin A$, $\{X_n(\omega)\}$ converge a $X(\omega)$	$X_n \xrightarrow{c.s.} X$
En Probabilidad (ó en media)	$P(\lvert X_n - X\rvert \geq \varepsilon) \xrightarrow[n\to\infty]{} 0$ para todo $\varepsilon > 0$.	$X_n \xrightarrow{P} X$
En Media i-ésima $i>0$	$E\lvert X_n\rvert^i < \infty$ y $E\lvert X_n - X\rvert^i \xrightarrow[n\to\infty]{} 0$	$X_n \xrightarrow{L^i} X$
En Distribución	$\lim\limits_{n\to\infty} P_{X_n} = P_X$	$X_n \xrightarrow{\mathcal{D}} X$

En cada tipo de convergencia, podemos usar el criterio visto en la sección III de este apéndice; es decir, determinando cuáles son sucesiones fundamentales, podemos establecer la convergencia, y esto es posible ya que no se menciona quién es el límite. Con esto verificamos la existencia.

DEFINICION D.17.- Una sucesión $\{X_n\}$ de variables aleatorias se dice que es FUNDAMENTAL EN PROBABILIDAD si :

$$P(\{\omega; |X_n(\omega)-X_m(\omega)| \geq \varepsilon\}) \xrightarrow[n,m\to\infty]{} 0 .$$

APENDICE [E] .- PROCESOS ESTOCÁSTICOS.

DEFINICION E.1.- Un PROCESO ESTOCÁSTICO $\{X_t, t\in T\}$ es una familia de variables aleatorias definidas sobre un mismo espacio de probabilidad $(\Omega, \mathcal{F}, P)$ e indexada por los puntos t que pertenecen a un subconjunto T del la recta real.

DEFINICION E.2.- El conjunto T es llamado el ESPACIO PARAMÉTRICO de el proceso.

OBS.- Usualmente se asocia al parámetro t con el tiempo. Si T es un conjunto numerable, digamos $T=\{t_1,t_2,\ldots\}$ entonces el proceso $\{X_t, t\in T\}$ resulta ser una sucesión de variables aleatorias.

Estaremos más interesados en los casos en que T sea un intervalo, los casos más comunes serán $T=[0,1]$, $T=[0,\infty]$, $T=(-\infty,\infty)$.

Un proceso estocástico es una función medible de dos variables; esto es, $X(\omega,t)$, $\omega\in\Omega$, $t\in T$. El único contraste es que para todo t y todo número real a, $\{\omega ; X(\omega,t)\leq a\}\in\mathcal{F}$.

DEFINICION E.3.- Para una ω fija, la función $X(\omega,\cdot)$ definida sobre T se llamará la TRAYECTORIA del proceso.

Para todo subconjunto finito $\{t_1,t_2,\ldots,t_n\}$ de T denotamos a la función de distribución de $X_{t_1},X_{t_2},\ldots X_{t_n}$ mediante $P^{t_1,t_2,\ldots,t_n}$. De ahí que $P^{t_1,t_2,\ldots,t_n}(x_1,\ldots,x_n) = P(\{X_{t_i} \leq x_i,\ i=1,2,\ldots,n\})$

DEFINICION E.4.- La colección $P^{t_1,t_2,\ldots,t_n}$ para todo subconjunto finito $\{t_1,\ldots,t_n\}$ de T, es llamada LA FAMILIA DE DISTRIBUCIONES FINITO DIMENSIONALES para el proceso $\{X_t,\ t \in T\}$

Una vez determinada la familia de distribuciones finito dimensionales, podremos encontrar la probabilidad de todo evento de la forma $\{\omega;\ X_{t_1}(\omega), X_{t_2}(\omega), \ldots, X_{t_n}(\omega) \in A\}$ donde A es un conjunto en el espacio n-dimensional R^n. Tomando límite, podemos encontrar la probabilidad de el límite de una sucesión de tales conjuntos.

Ejemplo.- Sea Ω = la colección de todas las funciones continuas en $[0,1]$. Sea $T = [0,1]$. Sea $X_t(\omega)$ = El valor de la función ω en t. Como toda ω es continua,

$$[\omega;\ a \leq X_t(\omega) \leq b \text{ para toda } t \in [0,1]\} = \{\omega;\ a \leq X_t(\omega) \leq b, \text{ para todo racional } r \in [0,1]\}$$

En este ejemplo, el conjunto $\{\omega;\ a \leq X_t(\omega) \leq b \text{ para todo } t \in [0,1]\}$ es un evento y su probabilidad puede ser calculada de las distribuciones finito dimensionales, ya que el conjunto de números racionales en el intervalo $[0,1]$ es numerable.

Los conjuntos que envuelven una cantidad no numerable de elementos t deben examinarse de modo individual. Las distribuciones finito dimensionales pueden o no, contener toda la información necesaria para la obtención de las probabilidades de tales conjuntos; sin embargo, la situación en la realidad no es tan rígida. Normalmente obtenemos, por medio de una combinación de hipótesis y medidas, una familia de distribuciones finito dimensionales para el fenómeno aleatorio de interés. Somos entonces, "libres" para construir $(\Omega, \mathfrak{F}, P)$ y $\{X_t, t \in T\}$ mientras que si para toda $\varepsilon > 0$ y toda $t \in T$:

$$P\left[|X_s - X_t| \geq \varepsilon\right] = \int_{|x-y| \geq \varepsilon} d\, F_{t,s}(x,y) \xrightarrow[s \to t]{} 0 \qquad \ldots \ (e, 5.1)$$

entonces el proceso puede ser construido de tal forma que para todo intervalo $[a,b]$, el conjunto $\{\omega; a \leq X_t(\omega) \leq b, \text{ para toda } t \in T\} \in \mathfrak{F}$ y $P\left[\{\omega; a \leq X_t(\omega) \leq b \text{ para toda } t \in T\}\right] = P\left[\{\omega; a \leq X_t(\omega) \leq b, \text{ para toda racional } t \in T\}\right]$ ... (e, 5.2)

Con probabilidad 1, toda trayectoria $X(\omega, \cdot)$ se comporta suficientemente bien que las integrales $\int X(\omega, t)\, dt$ pueden definirse. En particular si $\int_a^b E|X_t|\, dt < \infty$ entonces $Z(\omega) = \int_a^b X(\omega, t)\, dt$ es una variable aleatoria y

$$E[Z] = \int_a^b E[X_t]\, dt \quad \ldots \ (e, 5.3)$$

DEFINICION E.5.- La condición (e5.1) es conocida como CONTINUIDAD EN PROBABILIDAD.

OBS.- Esta condición no se satisface siempre; sin embargo, si se cumple y las distribuciones finito dimensionales son todas las que hemos dado, podemos asumir que el proceso posee las propiedades (e5.2) y (e5.3) deseadas.

En lo que sigue, intro-duciremos algunos conceptos que contribuyen a la simplificación del análisis. Comenzaremos con algunas propiedades de continuidad.

DEFINICION E.6.- Un proceso $\{X_t, t \in T\}$ se dice <u>continuo</u> :

i) EN PROBABILIDAD si $P(|X_s - X_t| \geq \varepsilon) \xrightarrow[s \to t]{} 0$ para toda $\varepsilon > 0$.

ii) EN MEDIA I-ESIMA si $E|X_t - X_s|^i \longrightarrow 0$ cuando $s \to t$.

iii) CASI SEGURAMENTE si $P(\lim_{s \to t} |X_s - X_t| = 0) = 1$.

DEFINICION E.7.- Un proceso $\{X_t ; t \in T\}$ se dice DE TRAYECTORIAS CONTINUAS si, con probabilidad 1, toda trayectoria $X(\omega, \cdot)$ es una función continua en T.

En base a esta definición (E.7) enunciamos la llamada CONDICION DE KOLMOGOROV para determinar cuándo un proceso es de trayectorias continuas.

PROPOSICION E.1.- Si un proceso es tal que satisface la condición (e5.2) y si existen constantes positivas α, ε y K tales que $$E|X_{t+h} - X_t|^{\alpha} \leq K h^{1+\varepsilon}$$
para h suficientemente pequeña y para toda t, entonces el proceso es de trayectorias continuas.

Dada una familia de distribuciones finito dimensionales $\{P_{T_n}, T_n \in T$ finita, toda $T_n\}$, podemos encontrar un espacio de probabilidad (Ω, Σ, P) y una familia de variables aleatorias $\{X_t, t \in T\}$ con las distribuciones finito dimensionales dadas.

La demostración se hace por construcción.:

Sea $\Omega = \mathbb{R}^T = \{$ conjunto de las funciones reales sobre $T\}$

Sea $X_t(\omega)$ el valor de ω en t .

Sean $\mathcal{B}_X$ y $\mathcal{A}_X$, la mínima álgebra y la mínima álgebra σ respectivamente, con la cual toda X_t es medible.

Ahora bien, todo conjunto $\mathcal{B}_X$ es de la forma

$$\{\omega;\ (X_{t_1}(\omega), X_{t_2}(\omega), \ldots, X_{t_n}(\omega)) \in B\} \quad \ldots \ (e, 5.4)$$

donde B es un conjunto de Borel n-dimensional.

Dada una familia de distribuciones finito dimensionales $\{P^{T_n}\}$, resulta :

$$P(\{\omega;\ (X_{t_1}(\omega), X_{t_2}(\omega), \ldots, X_{t_n}(\omega)) \in B\}) = \int_B dP^{t_1 \ldots t_n}(x_1, \ldots, x_n) \quad \ldots \ (e, 5.5)$$

Esto define una medida P de probabilidad sobre $(\Omega, \mathcal{B}_X)$. Además, se puede demostrar que P , definida de éste modo, no es solo aditiva finitamente, sino también aditiva σ . Esto significa que P es una medida de probabilidad que puede extenderse de manera única a $\mathcal{A}_X$. Para probar que P es aditiva σ y no solo aditiva, es complicado, por lo que omitiremos la demostración (Ver Neveu, 1965, pp. 82-83).

Resumiendo :

Si tomamos $\Omega = \mathbb{R}^T$ y definimos

$$X_t(\omega) = \text{al valor de } \omega \text{ en } t \quad \ldots \ (e, 5.6)$$

y $\mathcal{A}_X$ como el mínimo álgebra σ generado por $\{X_t, t \in T\}$, la medida de probabilidad se define mediante (e5.5). Así de este modo, el proceso $\{X_t, t \in T\}$ tiene las funciones finito dimensionales preescritas. Notemos que $\{X_t(\omega), \omega \in R^T, t \in T\}$ definido mediante (e5.6) es la llamada FUNCION COORDENADA, porque $X_t(\omega)$ es la t-ésima coordenada de ω.

DEFINICION E.8.- Los conjuntos de la forma (e5.4) se llaman CILINDROS.

II.- SEPARABILIDAD Y MEDIBILIDAD DE PROCESOS.

La definición de un proceso estocástico X_t, requiere que sea una función de ω, medible $\mathcal{A}$ para cada t; sin dar lugar a condiciones sobre él, como una función de t. Cuando T es un intervalo, las preguntas de naturaleza analítica concernientes a funciones de trayectorias continuas del proceso, envuelven una cantidad no numerable de eventos y variables aleatorias, que no siempre pueden contestarse sin antes añadir algunos conceptos como condiciones. Por ejemplo, el conjunto $\{\omega; X_t(\omega) \geq 0 \text{ para toda } t \in T\} = \bigcap_{t \in T} \{\omega; X_t(\omega) \geq 0\}$ puede no ser un evento cuando se tiene una intersección no nimerable de eventos. De la misma forma, $Y = \sup_t X_t$ puede no ser una variable aleatoria porque conjuntos de la forma

$$\{\omega; Y(\omega) \leq y\} = \bigcap_{t \in T} \{\omega; X_t(\omega) \leq y\}$$

pueden no ser eventos.

Estas restricciones motivaron la introducción de los conceptos de Proceso Separable y Proceso Medible (Doob, 1953, pp 50)

DEFINICION E.9.- Un proceso se llama SEPARABLE si existe un conjunto S numerable, $S \subset T$ y en evento Δ nulo, tal que para todo conjunto K cerrado, $K \subset [-\infty,\infty]$, y cualquier intervalo I abierto, los dos conjuntos

$$\{\omega ; X_t(\omega) \in K, t \in I \cap T\} \text{ y } \{\omega ; X_t(\omega) \in K, t \in I \cap S\}$$

difieran por un subconjunto de Δ .

DEFINICION E.10.- El conjunto S numerable es llamado el SEPARANTE.

Si el proceso $\{X_t, t \in T\}$ es separable, y $\omega \notin \Delta$ entonces, para cualquier intervalo I abierto, $X_t(\omega) \leq a, t \in I \cap S$ implica que $X_t(\omega) \leq a$ para $t \in I \cap T$. De este modo :

$$\sup_{t \in I \cap S} X_t(\omega) \geq \sup_{t \in I \cap T} X_t(\omega)$$

y como $S \subset T$, la desigualdad contraria también tiene lugar. De aquí que para cualquier intervalo I abierto

$$\sup_{t \in I \cap S} X_t(\omega) = \sup_{t \in I \cap T} X_t(\omega) \quad \text{para } \omega \notin \Delta .$$

Como S es numerable, $\sup_{t \in I \cap S} X_t(\omega)$ es una variable aleatoria c.t.p. Si el espacio de probabilidad es completo, entonces $\sup_{t \in I \cap T} X_t(\omega)$ es una variable aleatoria. Y tenemos un resultado similar para $\inf_{t \in I \cap T} X_t(\omega)$, y de ahí, también para

$$\lim_{s \to t} \sup X_s = \lim_{n \to \infty} \sup_{|s-t| < \frac{1}{n}} X_s \quad \text{y} \quad \lim_{s \to t} \inf X_s = \lim_{n \to \infty} \inf_{|s-t| < \frac{1}{n}} X_s$$

De aquí en adelante, supondremos que el espacio de probabilidad es completo.

Dado un espacio de probabilidad $(\Omega, \mathcal{A}, P)$ y un proceso $\{X_t, t\in T\}$ definido sobre él, dicho proceso puede ser separable o no, y no podemos hacer nada para cambiarlo. Cuando el proceso es separable, existe un recurso para encontrar al separante.

PROPOSICION E.2.- Sea $\{X_t, t\in T\}$ un proceso separable el cual es continuo en probabilidad. Entonces, todo conjunto numerable denso en T es un conjunto separante.

Como vimos anteriormente, es también deseable poder definir integrales de la forma $\int_a^b X_t(\omega)dt$. Si dicha integral se interpreta como integral de Lebesgue de funciones con trayectorias continuas, la integrabilidad de casi toda función de este tipo es una necesidad. Aún cuando casi toda función con trayectorias continuas $\{X_t, t\in T\}$ es integrable a la Lebesgue, la integral $\int_a^b X_t(\omega)dt$ resultante, puede no ser una variable aleatoria. Lo que se necesita es que $X_t(\omega)$ defina una función de (t,ω) medible con respecto a $\mathcal{L} \times \mathcal{A}$, donde $\mathcal{L}$ denota el álgebra σ de los conjuntos medibles a la Lebesgue en T . Recordemos que $\mathcal{L}$ es el mínimo álgebra σ que contiene a todos los intervalos y es completa con respecto a la medida que asigna su longitud a tales intervalos.

DEFINICION E.11.- Un proceso $\{X_t, t \in T\}$ con un conjunto T medible a la Lebesgue se llama PROCESO MEDIBLE si $X_t(\omega)$ es una función de (t,ω) medible con respecto a $\mathcal{L} \times \mathcal{A}$.

$\mathcal{A}$ es el álgebra σ de eventos en el espacio de probabilidad definido; es decir, para toda $x \in (-\infty,\infty)$ $\{(t,\omega) ; X_t(\omega) \leq x\} \in \mathcal{L} \times \mathcal{A}$

Notemos que la separabilidad no impone restricción sobre las distribuciones finito dimensionales; sin embargo, la medida lo hace. Por ejemplo, puede mostrarse que si T es un intervalo, y si $\{X_t, t \in T\}$ representa una colección de variables aleatorias independientes e idénticamente distribuidas, entonces no puede ser medible a menos que la distribución de X_t esté concentrada en un punto.

La siguiente proposición ofrece una condición suficiente sobre la distribución finito dimensional de un proceso medible, y también reúne algunos resultados precedentes, (ver DOOB, 1953, pp. 61-62).

PROPOSICION E.3.- Sea $\{X_t, t \in T\}$ un proceso continuo en probabilidad y sea T un intervalo. Entonces, existe un proceso $\{\tilde{X}_t : t \in T\}$ definido sobre el mismo espacio de probabilidad tal que

i) $P(X_t = \tilde{X}_t) = 1$ para todo $t \in T$

ii) $\{\tilde{X}_t, t \in T\}$ es separable

iii) Cualquier conjunto numerable denso en T, es un separante para $\{\tilde{X}_t, t \in T\}$.

DEFINICION E.12.- El proceso $\{\tilde{X}_t, t\in T\}$ se llama MODIFICACION SEPARABLE y MEDIBLE de $\{X_t, t\in T\}$.

OBS.- Siempre que un proceso sea continuo en probabilidad podemos asumir que tenemos listo un reemplazo por medio de una modificación separable y medible.

Finalmente, observemos que si A es un conjunto de la recta real, medible a la Lebesgue, y si $\int_A E|X_t|dt < \infty$ entonces, casi toda trayectoria continua es integrable a la Lebesgue sobre A, y $\int_A X_t dt$ define una variable aleatoria.

III.- PROCESOS GAUSSIANOS.

Como mencionamos, muchos fenómenos aleatorios están bien aproximados por procesos estocásticos que son llamados Gaussianos. Esto es de gran ventaja, ya que dichos procesos tienen distribuciones que gozan de gran simplicidad analítica.

Sea Z una variable aleatoria tal que $E[Z^2]<\infty$. Sea $\mu = E[Z]$ y $\sigma^2 = E[(Z-\mu)^2]$. En estas condiciones tenemos la siguiente

DEFINICION E.13.- La variable aleatoria Z se dice GAUSSIANA, ya sea si $\sigma^2=0$, en cuyo caso Z es igual a μ con probabilidad 1, o bien si $P(Z\leq a) = \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{z-\mu}{\sigma}\right)^2\right] dz$ siempre que $\sigma^2>0$.

Podemos calcular la función característica y encontramos que

$$\varphi_z(u) = \exp\left[i\mu u - \frac{1}{2}\sigma^2 u^2\right]$$

entonces, como la distribución está determinada de manera única por la función característica, una condición necesaria y suficiente para que z sea Gaussiana, es que satisfaga

$$\varphi_z(u) = E\left[e^{iuz}\right] = \exp\left[iuE(z) - \frac{1}{2}E\left[(z-E(z))^2\right]u^2\right]$$

DEFINICION E.14.- Un proceso estocástico $\{X_t, t\in T\}$ se llama GAUSSIANO si toda combinación lineal de la forma

$$z = \sum_{i=1}^{N} \alpha_i X_{t_i}$$

es una variable aleatoria Gaussiana.

OBS.- Para que $\{X_t, t\in T\}$ sea un proceso Gaussiano, es necesario que para cada t, X_t sea una variable aleatoria Gaussiana. Pero esto no es suficiente. Una condición necesaria y suficiente está dada en la siguiente

PROPOSICION E.4.- Un proceso $\{X_t, t\in T\}$ es Gaussiano si, y solo si

i) $E\left[X_t^2\right] < \infty$ para cada $t \in T$

ii) Para toda colección finita $(t_1, \cdots, t_N) \subset T$

$$E\left[\exp\left(i\sum_{k=1}^{N} u_k X_{t_k}\right)\right] = \exp\left[i\sum_{k=1}^{N} u_k \mu(t_k) - \frac{1}{2}\sum_{k,m=1}^{N} u_k u_m B(t_k, t_m)\right]$$

donde $\mu(t) = E[X_t]$ y $B(t,s) = E\left[(X_t - \mu(t))(X_s - \mu(s))\right] \ldots$ (e, 5.7)

OBS.- Supongamos que $\{X_n\}$ es una sucesión de variables aleatorias Gaussianas que converge a una variable aleatoria X. Entonces

$$\mu_n = E[X_n] \xrightarrow[n\to\infty]{L^2} \mu = E[X] \quad y \quad \sigma_n^2 = E\left[X_n - E[X_n]\right]^2 \xrightarrow[n\to\infty]{} \sigma^2 = E\left[X - E(X)\right]^2$$

Por lo expuesto en el punto (V) sobre los tipos de convergencia, y dada la forma de la distribución de cada X_n , la función de distribución de X debe ser de la forma

$$P(X \leq a) = \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right) dx$$

En otras palabras, el límite de una sucesión convergente en media cuadrática de variables aleatorias Gaussianas, es siempre una variable Gaussiana.

DEFINICION E.15.- La función $\mu(t)$ es llamada la función MEDIA de $\{X_t, t \in T\}$.

DEFINICION E.16.- La función $R(t,s)$ definida como (e5.7) es llamada la función COVARIANZA.

Estas consideraciones muestran que todas las funciones finito dimensionales de un proceso Gaussiano están determinadas completamente, una vez especificadas las funciones $\mu(t)$ (Media) y $R(t,s)$ (Covarianza). Esta es realmente la forma más simple de especificar a las distribuciones finito dimensionales de un proceso Gaussiano.

Como ejemplo, consideremos al proceso llamado MOVIMIENTO BROWNIANO $\{X_t, t \geq 0\}$ definido por

i) $\{X_t, t \geq 0\}$ es un proceso Gaussiano

ii) $E[X_t] = 0 \qquad E[X_t; X_s] = \min(t,s)$

Si $0 < t_1 < t_2 < \ldots < t_N$, la matriz $R = [\min(t_m, t_k)]$ es positiva. Por lo tanto, la función de distribución puede expresarse como $P(x_1,t_1; x_2,t_2; \dots ; x_N t_N) = \prod_{i=1}^{N} \int_{-\infty}^{x_i} \frac{1}{\sqrt{2\pi(t_i - t_{i-1})}} \exp\left[-\frac{1}{2}\frac{(\xi_i - \xi_{i-1})^2}{(t_i - t_{i-1})}\right] d\xi$ donde $t_0 = 0 = x_0$. Esta ecuación muestra que $\{X_{t_1}, X_{t_2} - X_{t_1}, \dots, X_{t_N} - X_{t_{N-1}}\}$ es una colección de variables aleatorias independientes para toda colección creciente de puntos $(t_1, \dots, t_N)$. Todo proceso que posea esta propiedad es llamado PROCESO CON INCREMENTOS INDEPENDIENTES. Por lo tanto, un movimiento Browniano es un proceso con incrementos independientes.

Enunciaremos una proposición la cual no demostraremos; sin embargo, será de gran ayuda para entender un tipo de medida de probabilidad empleada en el artículo de Prokhorov.

PROPOSICION E.5- Con probabilidad 1, toda trayectoria continua de un movimiento Browniano separable, es uniformemente continua sobre todo intervalo finito.

La propiedad de trayectorias continuas de un movimiento Browniano separable, puede interpretarse de la siguiente forma:

Sea C el espacio de todas las funciones continuas de valor real sobre $[0,\infty)$. Sea $X_t(\omega)$, $\omega \in C$ el valor de ω en t. Sea $\mathcal{A}$ el álgebra σ mínimo de subconjuntos de C tal que toda X_t sea medible. En estas circunstancias, puede probarse que existe una única medida P de probabilidad sobre $(C, \mathcal{A})$ tal que $\{X_t, 0 \le t < \infty\}$ es un movimiento Browniano. Así construido, $\{X_t, t \geq 0\}$ es necesariamente separable. La continuidad de las trayectorias, en este contexto, no dice más que $P(C) = 1$.

DEFINICION E.17.- La medida P sobre $(C, \mathcal{A})$ se llama MEDIDA DE WIENER.

Sea $\mathcal{B}$ la mínima álgebra, respecto de la cual todo X_t es medible. $\mathcal{B}$ no es álgebra σ. Ahora bien, toda familia de distribuciones finito dimensionales para $\{X_t, t \in [0,1]\}$ define una medida P' sobre $(C_{[0,1]}, \mathcal{B})$. Aunque P' es, en general, aditiva finitamente, si P' es aditiva σ entonces puede extenderse a una medida P de probabilidad sobre $(C_{[0,1]}, \mathcal{A})$. Lo que resulta es un proceso $\{X_t, t \in [0,1]\}$ definido sobre $(C_{[0,1]}, \mathcal{A}, P)$. Como $P(C_{[0,1]}) = 1$ toda trayectoria de X_t es continua.

Podemos interpretar la condición de Kolmogorov como sigue:

Toda familia de funciones de distribución finito dimensionales que satisface la condición de Kolmogorov, define una medida P' sobre $(C, \mathcal{B})$, la cual es aditiva σ.

En el caso del movimiento Browniano, la medida P resultante es la llamada Medida de Wiener.

Debemos notar que todo proceso $\{X_t, t \in T\}$ definido sobre $(C, \mathcal{A}, P)$ de ésta forma, es necesariamente separable y medible.

Con los conceptos definidos hasta aquí, podemos comenzar con el análisis del artículo de Prokhorov (tema central de el trabajo presente).

BIBLIOGRAFIA :

1 BILLINGSLEY, PATRICK

"Convergence of Probability Measures"

2 CHOQUET, GUSTAVE

"Lectures on Analisys"

3 DONSKER, MONROE

"An invariance principle for certain probability limit teorems"

4 FELLER, WILLIAM

"Introducción a la teoria de probabilidades y sus aplicaciones"

5 GARZA TOMAS

"Elementos de cálculo de probabilidades"

6 HALMOS, PAUL

"Measure theory"

7 HALMOS, PAUL

"Teoria intuitiva de los conjuntos"

8 KOLMOGOROV, A. N.

"Foundations of probability"

9 KOLMOGOROV, A. N. ; FOMIN, S. V.
"Elementos de la teoria de funciones y del análisis funcional"

10 KOLMOGOROV, A. N. ; GNEDENKO
"Limit distributions for sums of independent random variables"

11 LOEVE, M.
"Probability theory"

12 MURAYAMA, GISIRO
"Continuos Markov processes and stochastic equations"

13 PARTHASARATHY, K. R.
"Probability measures on metric spaces"

14 RUIZ DE CHAVEZ, JUAN
"Notas de clase"

15 WONG, EUGENE
"Intoduction to random processes"

16 WONG, EUGENE
"Stochastic processes in information and Dynamical systems"