APLICACIONES MEDIBLES. VARIABLES ALEATORIAS

APLICACIONES MEDIBLES. PROPIEDADES

Definición.- Sean $(\Omega, \mathcal{A})$ y $(\mathcal{X}, \beta)$ dos espacios medibles. Sea h una aplicación de Ω en $\mathcal{X}$, decimos que h es una aplicación medible si

$$\forall B \in \beta, \quad h^{-1}(B) \in \mathcal{A}.$$

Veamos a continuación algunas propiedades de interés.

PROPOSICION I.- Sea h una aplicación medible de $(\Omega, \mathcal{A})$ en $(\mathcal{X}, \beta)$ y g una aplicación medible de $(\mathcal{X}, \beta)$ en $(Z, \mathcal{C})$, entonces $g \circ h$ es una aplicación medible

Demostración.- En efecto, $\forall C \in \mathcal{C}$, tenemos $g^{-1}(C) \in \beta$ y por tanto $h^{-1}[g^{-1}(C)] \in \mathcal{A}$, pero

$$[g \circ h]^{-1}(C) = h^{-1}[g^{-1}(C)]$$ y de aquí el resultado.

PROPOSICION II.- Sea h una aplicación medible entre los espacios medibles $(\Omega, \mathcal{A})$ y $(\mathcal{X}, \beta)$, entonces

a) Definimos $\mathcal{A}_h = \{B,\ B \subset \mathcal{X} \,/\, h^{-1}(B) \in \mathcal{A}\}$. Entonces $\mathcal{A}_h$ es una σ-algebra, a la que llamaremos σ-algebra imagen de $\mathcal{A}$ por h y que tiene la propiedad de ser la mayor σ-algebra (para la relación de orden de inclusión) que hace medible la función h (entre los espacios medibles $(\Omega, \mathcal{A})$ y $(\mathcal{X}, \mathcal{A}_h)$).

b) Definimos $\beta_h = \{A,\ A \subset \Omega \,/\, A = h^{-1}(B) \text{ con } B \in \beta\}$. β_h es una σ-algebra, la σ-algebra engendrada por h a partir de β. Tiene la propiedad de ser la menor σ-algebra con la que podemos dotar a Ω para que h sea medible (entre (Ω, β_h) y $(\mathcal{X}, \beta)$).

Demostración.- Es fácil comprobar que tanto $\mathcal{A}_h$ como β_h son σ-algebras, por ejemplo, si $\{B_n\}_{n\geq 1} \subset \mathcal{A}_h$, entonces $h^{-1}(B_n) \in \mathcal{A}$, $\forall n$ y de aquí $\bigcup_{n\geq 1} h^{-1}(B_n) \in \mathcal{A}$

pero

$$\bigcup_{n\geq 1} h^{-1}(B_n) = h^{-1}\Big(\bigcup_{n\geq 1} B_n\Big), \text{ luego } \bigcup_{n\geq 1} B_n \in \mathcal{A}_h.$$

Análogamente se comprueba para β_h y también las otras propiedades que definen el concepto de σ-algebra. Por otra parte, de acuerdo con las definiciones de aplicación medible y la de $\mathcal{A}_h$, tenemos que

$$h : (\Omega, \mathcal{A}) \longrightarrow (\mathcal{X}, \mathcal{A}_h) \quad \text{es medible}$$

pero si β es tal que $h : (\Omega, \mathcal{A}) \to (\mathcal{X}, \beta)$ es medible, entonces $h^{-1}(B) \in \mathcal{A}$, $\forall B \in \beta$ luego $\beta \subset \mathcal{A}_h$ y ésta es la mayor de cuantas hacen medible a h ($(\Omega, \mathcal{A})$ fijo).

Razonando de forma análoga comprobaríamos que β_h es la menor de las que hace medible h, cuando actúa entre $(\Omega, \beta_h) \to (\mathcal{X}, \beta)$, y este último espacio medible es fijo.

PROPOSICION III.- 1) Sean $(\Omega, \mathcal{A})$ y $(\mathcal{X}, \beta)$ dos espacios medibles. Sea $\mathcal{C}$ una familia de partes de $\mathcal{X}$ tal que $\beta = \sigma(\mathcal{C})$. Una aplicación h entre Ω y $\mathcal{X}$ es medible si y sólo si,

$$\forall C \in \mathcal{C}, \quad h^{-1}(C) \in \mathcal{A}. \qquad [\sigma(\mathcal{C}), \text{ σ-algebra engendrada por } \mathcal{C}]$$

2) En particular, si $\mathcal{X}$ es un espacio topológico y si β es su tribu de Borel, h es medible si la antiimagen de cualquier abierto es un elemento de $\mathcal{A}$.

3) Para espacios topológicos dotados de sus sistemas de abiertos y de sus tribus de Borel respectivas, toda aplicación continua es medible.

Demostración.- 1) Si h es medible obviamente $h^{-1}(C) \in \mathcal{A}$, $\forall C \in \mathcal{C}$.
Recíprocamente, si $h^{-1}(C) \in \mathcal{A}$, $\forall C \in \mathcal{C}$, ello supone que $\mathcal{A}_h$, de la proposición anterior, contiene a $\mathcal{C}$ y por tanto contendrá a la σ-algebra que éste engendra, es decir, $\beta \subset \mathcal{A}_h$, pero esto supone que $h^{-1}(B) \in \mathcal{A}$, $\forall B \in \beta$, y h será medible.

2) Se deriva inmediatamente de 1)

3) Se deriva del hecho de que las aplicaciones continuas verifican que las antiimagenes de abiertos son abiertos.

FUNCIONES MEDIBLES. PROPIEDADES

Definición.- Una función medible es una aplicación medible, f, entre el espacio medible $(\Omega, \mathcal{A})$ y (R, β), donde R es la recta real y β la tribu de Borel correspondiente.

Enunciaremos a continuación algunas propiedades cuyas demostraciones omitiremos por escapar de las posibilidades y pretensiones de este curso. En cualquier caso convendrá tenerlas presente por el uso posterior que de ellas hagamos.

Propiedades. 1) La suma de funciones medibles, si está definida, es también medible

2) El producto de funciones medibles es medible, si está definido.

3) Los supremos e ínfimos de funciones medibles son medibles

En estas 3 propiedades el número de funciones al que hacemos referencia es finito y, obviamente, todas ellas están definidas sobre el mismo espacio medible $(\Omega, \mathcal{A})$.

4) El $\underline{\lim}$, $\overline{\lim}$ y $\lim$ (cuando existe) de una sucesión de funciones medibles es también una función medible.

Las funciones medibles admiten una caracterización al margen de la definición dada y es:

Caracterización de las funciones medibles: Una función f, de $(\Omega, \mathcal{A})$ en (R, β) es medible si y sólo si los conjuntos de la forma $\{w : f(w) \leq a\} \in \mathcal{A}$, $\forall a \in R$.

ALGUNAS FUNCIONES MEDIBLES DE INTERES

1) Función constante. Sea $f: \Omega \to R$ tal que $f(\omega) = k$, $\forall \omega \in \Omega$. Esta función es medible. Observemos que los conjuntos de la forma $\{\omega;\ f(\omega) \le a\}$ son

$$\{\omega:\ f(\omega) \le a\} = \begin{cases} \phi, & \text{si } a < k \\ \Omega, & \text{si } a \ge k \end{cases}$$

y por tanto, como ϕ y Ω son elementos de $\mathcal{A}$, f será medible.

2) Función característica de un conjunto. Sea $A \subset \Omega$, definimos $f_A: \Omega \to R$ mediante

$$f_A(\omega) = \begin{cases} 1, & \text{si } \omega \in A \\ 0, & \text{si } \omega \notin A. \end{cases}$$

La función así definida no siempre será medible si dotamos a Ω de una σ-algebra, su medibilidad va ligada al conjunto al cual caracteriza. Señalemos, antes de comprobar la afirmación que acabamos de hacer, que se designan como conjuntos medibles a los elementos de la σ-algebra del espacio medible. Es decir, $(\Omega, \mathcal{A})$ medible, $A \in \mathcal{A}$, entonces A es un conjunto medible.

Proposición.- La función característica de un conjunto medible es medible.

En efecto, si A es medible, entonces

$$\{\omega : f_A(\omega) \le a\} = \begin{cases} \phi, & a < 0 \\ A^c, & 0 \le a < 1 \\ \Omega, & a \ge 1 \end{cases}$$

y como todos ellos son medibles, entonces f_A es medible.

3) Función simple. Una función simple es una función medible con rango finito. Es decir

$$\forall \omega \in \Omega,\ f(\omega) \in \{a_1, \dots, a_n\}$$

La medibilidad de la función supone que $f^{-1}(\{a_i\}) = A_i \in \mathcal{A}$, $i = 1, \dots, n$ y además los A_i constituyen una partición de Ω. A saber

$$\bigcup_{i=1}^{n} A_i = \Omega, \quad A_i \cap A_j = \phi,\ i \ne j \quad i,j = 1, \dots, n.$$

Esta propiedad sugiere una representación de f mediante las funciones características de los A_i. En efecto, fácilmente se comprueba que toda función simple puede representarse de la forma

$$f = \sum_{i=1}^{n} a_i f_{A_i}.$$

Pero por otra parte observemos que una función, s, definida mediante la expresión $\sum_{i=1}^{n} a_i f_{A_i}$ con $A_i = s^{-1}(\{a_i\})$ medible, es una función simple. Por eso en ocasiones encontramos la siguiente definición para las funciones simples

Definición.- Sea s una aplicación de $\Omega \to R$, decimos que s es una función simple si
$s = \sum_{i=1}^{n} a_i f_{A_i}$, donde los A_i constituyen una partición medible de Ω y $a_i \in R$, $i = 1, \dots, n$.

La importancia de las funciones simples quedará puesta de manifiesto más adelante. Ocupemonos ahora de unas funciones medibles que son una generalización inmediata de las simples.

4) Función discreta. Una función d, definida del espacio medible $(\Omega, \mathcal{A})$ en (R, β) decimos que es discreta si

$$d = \sum_{n \ge 1} a_n f_{A_n},$$ donde los $a_n \in R$ y los A_n constituyen una partición medible numerable de Ω.

Obviamente si la partición numerable la tomamos finita aparece la función simple como caso particular de las funciones discretas.

VARIABLES ALEATORIAS

Cuando estamos estudiando el experimento consistente en lanzar un dado representamos el suceso A_i, la cara que nos muestra el dado presenta i puntos, simplemente mediante el número i, $i = 1, \dots, 6$. Nos limitamos a decir, ha salido un 6, o ha salido un 2, etc... Quizás no seamos conscientes de ello, pero estamos pasando del mundo abstracto de los sucesos al mundo concreto de los números, mediante una aplicación que a cada cara le hace corresponder el número de puntos que contiene. Esta actuación es muy frecuente y va de hecho implícita en el tratamiento que damos a todos los procesos aleatorios. Nos permite, en definitiva, abandonar las abstracciones que los sucesos suponen y actuar en la recta real, medio en el que nos desenvolvemos con facilidad y cuyas propiedades hemos estudiado exhaustivamente.

No obstante este cambio de escenario no puede llevarse a cabo de cualquier manera. Parece lógico pensar que la aplicación que nos traslada de Ω a R goce de unas mínimas propiedades que garanticen el manejo de las imágenes de los sucesos de la misma forma en que manejabamos a estos. Sería bueno que dicha aplicación conservara las estructuras que poseíamos sobre Ω, por ejemplo, y puesto que la única estructura que allí teníamos es la σ-algebra de sucesos, en realidad estamos pidiendo que la aplicación sea una función medible. De aquí la definición

Definición.- Una variable aleatoria es una función medible definida sobre un espacio de probabilidad.

Es costumbre designar las variables aleatorias mediante las letras mayúsculas X, Y, Z, etc... Así pues, si X es una variable aleatoria (v.a.), tenemos

$$X: \Omega \longrightarrow R$$ y además $X^{-1}(B) \in \mathcal{A}$, con $B \in \beta$, $\mathcal{A}$ la tribu de sucesos.

Así, cada suceso se ha convertido, mediante X, en la antiimagen de un conjunto de Borel y, en definitiva en un subconjunto de R.

Ejemplo.- Consideremos el experimento consistente en lanzar una moneda. El espacio muestral Ω está constituido por

$$\Omega = \{\text{cara}, \text{cruz}\}$$

podemos definir $X: \Omega \to R$ mediante $X(\{\text{cara}\}) = 1$, $X(\{\text{cruz}\}) = 0$. X es una variable aleatoria.

Ejemplo.- Lanzamos un dado dos veces consecutivas y estamos interesados en la suma de puntos que sus caras presentan. El espacio muestral está constituido por

$$\Omega = \{(i,j),\ i = 1, \dots, 6,\ j = 1, \dots, 6\}$$ i, j representan los puntos que muestran la primera y la segunda cara, respectivamente

Definimos sobre Ω la aplicación

$$\mathbb{X}(\omega) = i+j \qquad , \omega \in \Omega, \quad \omega = (i,j), \ i,j = 1, \dots, 6$$

Esta aplicación es una aplicación medible y por tanto una variable aleatoria. Más concretamente se trata de una función simple por cuanto su rango es finito, a saber $\{2, 3, \dots, 12\}$, las posibles sumas que pueden obtenerse. Si denominamos S_d, el conjunto de puntos que constituyen el rango de $\mathbb{X}$, tendremos que para cualquier conjunto de Borel, B

$$\mathbb{X}^{-1}(B) = \phi \quad , \text{ si } \quad B \cap S_d = \phi$$

Si $\quad B \cap S_d = \{a_1, \dots, a_k\}$, entonces

$$\mathbb{X}^{-1}(B) = \{(i,j) = \omega \in \Omega \ / \ i+j = a_l, \ l = 1, \dots, k\}$$

conjunto que evidentemente es un suceso por cuanto recordemos que lo usual es dotar a los espacios muestrales finitos de la σ-algebra de sus partes.

Por ejemplo, si $B \cap S_d = \{8\}$, tendremos $\mathbb{X}^{-1}(B) = \{(2,6), (6,2), (3,5), (5,3), (4,4)\}$.

Probabilidad inducida. Sobre el espacio de probabilidad $(\Omega, \mathcal{A}, P)$ tenemos una medida de probabilidad P que nos permite determinar la probabilidad de cualquier suceso. Mediante la variable aleatoria $\mathbb{X}$ hemos trasladado nuestra acción a (R, β) y requeriremos de una medida de probabilidad que nos permita determinar la probabilidad de cualquier conjunto de Borel, en tanto en cuanto representa la transformación de algún suceso o conjunto de sucesos del espacio muestral. Veamos como hacerlo.

Sea $B \in \beta$, sabemos que $\mathbb{X}^{-1}(B) \in \mathcal{A}$, vamos a definir sobre β la siguiente aplicación

$$\forall B \in \beta, \ P'(B) = P(\mathbb{X}^{-1}(B)).$$

Esta aplicación está correctamente definida sobre todos los elementos de β y además:

I) $P'(\phi) = P(\mathbb{X}^{-1}(\phi)) = P(\phi) = 0.$

II) $P'(B) \geq 0, \ \forall B \in \beta.$

III) $\{B_n\}_{n \geq 1} \subset \beta \quad , \underset{n \neq m}{B_n \cap B_m = \phi} \quad , \quad P'\left(\bigcup_{n\geq 1} B_n\right) = P\left(\mathbb{X}^{-1}\left(\bigcup_{n \geq 1} B_n\right)\right) = P\left(\bigcup_{n \geq 1} \mathbb{X}^{-1}(B_n)\right)$

y por las propiedades de P y de la imagen inversa,

$$P'\left(\bigcup_{n \geq 1} B_n\right) = \bigcup_{n \geq 1} P(\mathbb{X}^{-1}(B_n)) = \bigcup_{n \geq 1} P'(B_n)$$

IV) $P'(R) = P(\mathbb{X}^{-1}(R)) = P(\Omega) = 1$

P' es pues una probabilidad sobre R, más exactamente sobre la tribu de Borel, a la que denominamos probabilidad inducida y que actua trasladando P sobre β mediante $\mathbb{X}$.

Ejemplo.- En el ejemplo anterior, si $B \cap S_d = \{8\}$, $P'(B) = P(\mathbb{X}^{-1}(B)) = P(\{(2,6), (6,2), (3,5), (5,3), (4,4)\})$ y si el dado es equilibrado, la probabilidad se obtendrá mediante la formula de Laplace y de aquí

$$P'(B) = \frac{5}{36}.$$

(5) Observación.- Observese que la probabilidad inducida es algo característico de la variable aleatoria que la define. (6) No es de extrañar que en ocasiones se la designe mediante $P'_{\mathbb{X}}$, sobre todo si existe riesgo de confusión por el manejo simultaneo de varias variables aleatorias. Partiendo de un mismo espacio de probabilidad, distintas variables aleatorias dan lugar a distintas probabilidades inducidas. Veamoslo mediante un ejemplo.

Sobre el espacio muestral $\Omega = \{(i,j), \ i,j = 1, \dots, 6\}$ de los posibles resultados obtenidos al lanzar dos veces un dado (o una vez dos dados), definimos la v.a. $\mathbb{X}$ como antes lo hicimos y una nueva v.a. $\mathbb{Y}$, como sigue

$$\mathbb{Y}(\omega) = |i-j| \quad , \omega = (i,j), \ i,j = 1, \dots, 6.$$

Consideremos un mismo conjunto de Borel, $B = \{2\}$, entonces

$$P'_{\mathbb{X}}(B) = P(\mathbb{X}^{-1}(\{2\})) = P(\{(1,1)\}) = \frac{1}{36}$$

$$P'_{\mathbb{Y}}(B) = P(\mathbb{Y}^{-1}(\{2\})) = P(\{(1,3), (3,1), (2,4), (4,2), (3,5), (5,3), (4,6), (6,4)\}) = \frac{8}{36}.$$

Función de distribución de una variable aleatoria. Mediante P', probabilidad inducida por $\mathbb{X}$, podemos definir una función de R en R con unas propiedades muy interesantes.

para $x \in R$, definimos $F_{\mathbb{X}}(x) = P'(]-\infty, x]) = P(\mathbb{X}^{-1}\{]-\infty, x]\}) = P(\mathbb{X} \leq x)$.

Esta función tiene las siguientes propiedades

F.1) Si $x_1 \leq x_2$, entonces $]-\infty, x_1] \subset]-\infty, x_2]$ y por la monotonía de P'

$$P'(]-\infty, x_1]) \leq P'(]-\infty, x_2]) \rightarrow F_{\mathbb{X}}(x_1) \leq F_{\mathbb{X}}(x_2)$$

la función es monotona creciente.

F.2) Sea x un punto cualquiera de R. Consideremos la siguiente familia de intervalos

$\{]-\infty, x_n], \ n \geq 1,$ [redacted] $x_n \geq x_{n+1}, \forall n\}$

que constituye una familia monotona decreciente de intervalos. Supongamos además que la sucesión de los x_n converge a x, entonces $\lim_n \{]-\infty, x_n]\} = \bigcap_{n \geq 1} \{]-\infty, x_n]\} =]-\infty, x]$.

Por las propiedades de continuidad de P', tenemos

$$P'(]-\infty, x]) = P'\left(\bigcap_{n \geq 1} \{]-\infty, x_n]\}\right) = \lim_n P'(]-\infty, x_n])$$

y en terminos de la función que acabamos de definir

$$F_{\mathbb{X}}(x) = \lim_n F_{\mathbb{X}}(x_n) \qquad x_n \downarrow x,$$

y dada la monotonía de F esto supone que F es continua por la derecha.

Estas dos propiedades caracterizan a F como una función de distribución, pero posee otras dos que la hacen especialmente interesante.

F.3) Si $\{x_n\}_{n\geq 1}$ es una sucesión de números reales que ~~[illegible]~~ tiende a $-\infty$ y además $x_n \geq x_{n+1}$ entonces los $I_n =]-\infty, x_n]$ constituyen una sucesión monótona decreciente tal que

$$\bigcap_{n\geq 1} I_n = \phi$$

aplicando la continuidad en ϕ de P',

$$\lim_n P'(I_n) = 0 \quad , \text{pero} \quad P'(I_n) = F_{\mathbb{X}}(x_n).$$

En definitiva, dada la monotonía de F, podemos afirmar

$$\lim_{x\to -\infty} F_{\mathbb{X}}(x) = 0.$$

F.4) Razonando análogamente para una sucesión monótona de x_n que tiende a $+\infty$, podemos afirmar que

$$\lim_{x\to +\infty} F_{\mathbb{X}}(x) = 1.$$

Se trata pues de una función de distribución acotada por la unidad a la que se conoce como <u>función de distribución de probabilidad de la variable aleatoria $\mathbb{X}$</u>.

Siendo P' una característica de $\mathbb{X}$ y estando $F_{\mathbb{X}}$ derivada a partir de P', obviamente, será también una característica de la variable aleatoria que la origina. El interés de P' y $F_{\mathbb{X}}$ reside en que ambos conceptos permiten <u>conocer</u> la variable aleatoria que los ha originado. En este contexto "conocer" significa poder determinar la <u>probabilidad de que la variable aleatoria tome valores en cualquier conjunto de Borel</u>. Conocer esto de una variable aleatoria es saberlo todo de la misma.

<u>Continuemos estudiando las propiedades de $F_{\mathbb{X}}$</u>. Hemos afirmado que la función era continua por la derecha, veamos ahora que ocurre cuando tratamos de comprobar su continuidad por la izquierda. Dada la monotonía de $F_{\mathbb{X}}$ bastará que trabajemos con sucesiones monótonas.

Sea $\{x_n\}_{n\geq 1}$ una sucesión monótona creciente que converge a x, ello supone que los $I_n =]-\infty, x_n]$ constituyen una sucesión monótona creciente de intervalos tales que

$$\lim_n I_n = \bigcup_n I_n =]-\infty, x[,$$

definamos una nueva familia de intervalos, $I_n' = I - I_n =]x_n, x]$, estos constituyen una familia monótona decreciente cuyo límite es

$$\lim_n I_n' = \bigcap_n I_n = \{x\},$$

tomando probabilidades,

$$P'(\lim_n I_n') = P'(\{x\}) = \lim_n P'(I_n') = \lim_n P'(I - I_n) = \lim_n \{P'(I) - P'(I_n)\} =$$

$$= \lim_n [F_{\mathbb{X}}(x) - F_{\mathbb{X}}(x_n)] = F_{\mathbb{X}}(x) - \lim_n F_{\mathbb{X}}(x_n).$$

Podemos pues afirmar que

$$F_{\mathbb{X}}(x) - F_{\mathbb{X}}(x-0) = P'(\{x\}) \quad , \forall x \in R.$$

La continuidad por la izquierda no está garantizada, a menos que $P'(\{x\}) = 0$. Directamente relacionada con la discontinuidad que esto supone podemos enunciar el siguiente teorema

<u>TEOREMA</u>.- El número de discontinuidades (saltos) que presenta una función de distribución de probabilidad es a lo sumo numerable.

Demostración.- Sabemos que $0 \leq F_{\mathbb{X}}(x) \leq 1$, $\forall x \in R$. Acabamos de comprobar que

$$F_{\mathbb{X}}(x+0) - F_{\mathbb{X}}(x-0) = P'(\{x\}),$$

siendo $P'(\{x\})$ la magnitud de la discontinuidad.

$F_{\mathbb{X}}$ no puede tener más de un salto (discontinuidad) con magnitud mayor que $\frac{1}{2}$, ni más de tres cuyas magnitudes estén comprendidas entre, $1/4 < P'(\{x\}) \leq \frac{1}{2}$. En general el número de saltos cuyas magnitudes estén comprendidas entre $1/2^n < P'(\{x\}) \leq 1/2^{n-1}$, no puede superar a $2^n - 1$. De acuerdo con ello su número total (unión numerable de cantidades finitas) será a lo sumo numerable.

El conjunto de puntos de discontinuidad de la función de distribución de una variable aleatoria juega un importante papel en la caracterización de ésta y sobre ello volveremos más tarde.

<u>Señalemos finalmente</u> la utilidad de la función de distribución de probabilidad a la hora de determinar la probabilidad de un intervalo de la forma $]x_1, x_2]$. En efecto, dada su definición

$$F_{\mathbb{X}}(x) = P'(]-\infty, x])$$

y teniendo en cuenta que

$$]x_1, x_2] =]-\infty, x_2] -]-\infty, x_1]$$

y que

$$P'(]x_1, x_2]) = P'(]-\infty, x_2] -]-\infty, x_1]) = P'(]-\infty, x_2]) - P'(]-\infty, x_1]) = F_{\mathbb{X}}(x_2) - F_{\mathbb{X}}(x_1).$$

<u>Obtención de una medida de probabilidad a partir de una función de distribución de probabilidad.</u>

En el párrafo precedente hemos visto que partiendo del espacio de probabilidad (R, β, P') se define una función cuyas propiedades se han puesto de manifiesto y cuya principal característica es la de ser una función de distribución tal y como se entiende en el Análisis Matemático clásico.

Observemos la validez del desarrollo independientemente del origen de P'. Es decir, con cualquier medida de probabilidad definida inicialmente sobre los borelianos de R hubiéramos llegado a una función de características similares.

Un importante aspecto de este proceso, quizás el más relevante, es que puede invertirse. A saber, partiendo de una función de distribución debidamente normalizada podemos definir sobre β una medida de probabilidad. Veamos como.

Sea F una función de distribución, es decir

P.1) Si $x_1 \leq x_2 \rightarrow F(x_1) \leq F(x_2)$

P.2) F es continua por la derecha

además verifica

P.3) $\lim_{x\to -\infty} F(x) = 0$

P.4) $\lim_{x\to +\infty} F(x) = 1$,

lo que se conoce como función de distribución de probabilidad.

Consideremos ahora la siguiente familia de intervalos

$$\mathcal{Z} = \{\phi, R,]-\infty, x],]x,y],]x,+\infty[\}$$

y definamos sobre ella la siguiente función de conjunto

$$P(\phi)=0\ ,\ P(]x,y]) = F(y)-F(x)\ ,\ P(]-\infty,x]) = F(x) - \lim_{x\to-\infty} F(x) = F(x)$$

$$P(]x,+\infty[) = \lim_{x\to+\infty} F(x) - F(x) = 1-F(x)\ ,\ P(R) = \lim_{x\to+\infty}F(x) - \lim_{y\to-\infty}F(y) = 1.$$

La familia $\mathcal{Z}$ posee además, dos interesantes propiedades:

Z.1) la σ-algebra que engendra, $\mathcal{B}(\mathcal{Z})$ es la tribu de Borel

Z.2) El algebra que engendra, $\mathcal{A}(\mathcal{Z})$, verifica

$$\forall T\in \mathcal{A}(\mathcal{Z}),\ T=\bigcup_{i=1}^n I_i\ ,\ I_i\in \mathcal{Z},\ I_i\cap I_j = \phi \ (i\neq j).$$

Esta última propiedad nos permite extender P de $\mathcal{Z}$ a $\mathcal{A}(\mathcal{Z})$ de forma inmediata,

$$P(T) = \sum_{i=1}^n P(I_i)$$

con lo que P sobre $\mathcal{A}(\mathcal{Z})$ verifica

a) Es no negativa, $P(T)\geq 0,\ \forall T\in \mathcal{A}(\mathcal{Z})$

b) Es aditiva, $\forall T_1,T_2 \in \mathcal{A}(\mathcal{Z})$, con $T_1\cap T_2=\phi$, $P(T_1\cup T_2)=P(T_1)+P(T_2)$.

En estas condiciones podemos aplicar el teorema de extensión de Hahn-Caratheodory y definir sobre β una medida, de probabilidad, pues $P(R)=1$, que resulta ser la extensión, única, de la inicialmente definida sobre $\mathcal{Z}$ y $\mathcal{A}(\mathcal{Z})$.

Queda así puesta en evidencia la correspondencia biunívoca existente entre la medida de probabilidad que caracteriza a una variable aleatoria y la función de distribución de probabilidad de la misma. La equivalencia que ello supone hace que sea indistinto para nosotros trabajar con una u otra, y solheremos hacerlo con la más sencilla de manejar, la función de distribución.

Ejemplo.- Volvamos nuevamente al espacio muestral $\Omega = \{(i \text{ puntos en la cara del dado}), i=1,\dots,6\}$. Mediante la aplicación

$$X\{(i \text{ puntos en la cara del dado})\} = i$$

definimos una variable aleatoria y a partir de ella, P', en la forma ya conocida, y siendo P la probabilidad clásica sobre Ω. Estudiemos con un poco de detalle la función de distribución de X, F_X.

Sea $S_d=\{1,2,3,4,5,6\}$, el rango de X. Tendremos

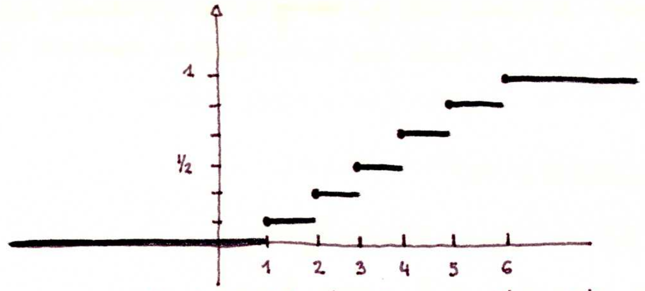
$$F_X(x)=0,\ \forall x<1$$

para $x=i,\ i\in S_d$, tendremos

$$F_X(i) = P(X\leq i) = P\{(j \text{ puntos en la cara del dado}), j\leq i\} = \sum_{j=1}^i \frac{1}{6} = \frac{i}{6}$$

para $x>6$, tendremos

$$F_X(x)=P(X\leq x)=P(\Omega)=1$$



La función de distribución obtenida en el lanzamiento de un dado, mediante la variable aleatoria que a cada cara le hace corresponder el número de puntos que contiene.

Si el punto x está comprendido entre $[i,i+1[$, $i\in S_d$, entonces

$$F_X(x) = P(X\leq x) = P\{(j \text{ puntos en la cara del dado}), j\leq x\} = \sum_{j\leq x} \frac{1}{6} = \sum_{j\leq i} \frac{1}{6} = \frac{i}{6}.$$

De acuerdo con esto la gráfica de la función de distribución es la que presentamos. Es una gráfica en escalera, cuyos peldaños (tramos constantes) aparecen entre dos valores consecutivos de S_d, tal como hemos visto en la obtención de F_X.

Digamos finalmente algo sobre la continuidad de F_X. Recordemos que los puntos de discontinuidad de F_X son aquellos reales para los que,

$$x\in R,\ P'(\{x\}) = P(X=x) \neq 0.$$

En el caso que estudiamos ello sólo ocurre para $x\in S_d$, es decir, F_X presenta discontinuidades en $x\in\{1,2,3,4,5,6\}$, todas ellas de igual magnitud, a saber

$$F_X(x+0) - F_X(x-0) = P(X=x) = \frac{1}{6},\quad x=1,2,3,4,5,6.$$

La función, como se pone de manifiesto en la gráfica, es continua en el complementario de S_d.

VARIABLES ALEATORIAS DISCRETAS

Definición.- Sea $(\Omega, \mathcal{A}, P)$ un espacio de probabilidad, definimos sobre él una variable aleatoria X. Decimos que X es una v.a. discreta si el conjunto de valores de su rango, S_d, es a lo sumo numerable.

Una consecuencia inmediata de esta definición es que la variable X será de la forma

$$X = \sum_{i\geq 1} x_i \cdot 1_{A_i}\ ,\ \text{con } x_i\in S_d, \forall i \ \text{ y } A_i = X^{-1}(\{x_i\}) \in \mathcal{A}, \forall i,\ \cup A_i = \Omega,\ A_i\cap A_j=\phi \ (i\neq j)$$

es decir, se trata de lo que definimos en su momento como una función discreta.

Por otra parte la función de distribución de X, F_X, será una función de salto, o en escalera, por cuanto

$$F_X(x) = 0,\ \forall x < \min S_d = x_1,\ [S_d = \{x_1, x_2, \dots, x_n, \dots\} \text{ lo suponemos ordenado}].$$

$$F_X(x) = P(X\leq x) = P(X\in\{x_1,\dots,x_i\}) = \sum_{j=1}^i P(X=x_j) = \sum_{j=1}^i P(A_j),\quad x\in [x_i, x_{i+1}[.$$

Los puntos de discontinuidad de F_X son los de S_d y la magnitud de la discontinuidad en x_i es $P(A_i)$.

La función de cuantía. Podemos definir una nueva función, mediante X, que será muy útil para estudiar las peculiaridades de X.

Sea $f_X: R\to R$ definida mediante

$$f_X(x) = P(X=x) = P'(\{x\})$$

a esta función se la denomina función de cuantía de la variable aleatoria X. En ocasiones se la denomina también función de probabilidad, puesto que proporciona, en cada punto, la probabilidad de que la v.a. X tome dicho valor. ¿Cuáles son las propiedades de f_X? ¿cuál es su relación, si la hay, con F_X?

<u>Propiedades de la función de cuantía</u>: (11)

1) $f_X(x) \geq 0 \quad \forall x \in R$

En efecto, si $x \in S_d$, $f_X(x) = f_X(x_i) = P(A_i)$, $x = x_i$

si $x \in S_d^c$, $f_X(x) = P(X=x) = 0$.

2) $\sum_{x_i \in S_d} f_X(x_i) = 1$

$$\sum_{x_i \in S_d} f_X(x_i) = \sum_{x_i \in S_d} P(X=x_i) = \sum_{x_i \in S_d} P(A_i) = P\left(\bigcup_i A_i\right) = P(\Omega) = 1$$

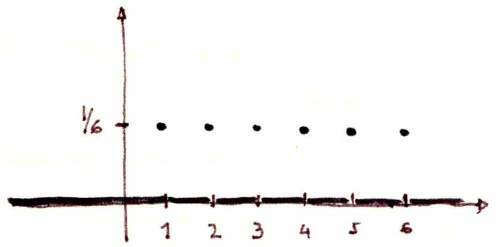
3) $f_X(x) = P(X=x) = P'(\{x\}) = F_X(x+0) - F_X(x-0) = F_X(x) - F_X(x-0)$.

La tercera propiedad establece la relación existente entre ambas funciones y más exactamente <u>permite obtener la función de cuantía a partir de la función de distribución</u> de la v.a. X. Pero el conocimiento de <u>f_X permite también proceder en sentido contrario</u>, veremos para ello que

$$F_X(x) = P'(]-\infty, x]) = P(X \leq x) = P(X \in S_d \cap]-\infty, x]) = P(X \in \{x_1, \dots, x_n\}, x_n \leq x) =$$
$$= \sum_{i=1}^{n} P(X=x_i) = \sum_{i=1}^{n} f_X(x_i).$$

lo que permite construir F_X a partir de f_X. Esto supone que las características probabilísticas de X, que conocíamos a través de su distribución de probabilidad, es decir, F_X, pueden también conocerse mediante f_X y en ello estriba precisamente la importancia y utilidad de la función de cuantía.

<u>Ejemplo</u>.- Considerando nuevamente el ejemplo del dado.



La función de cuantía valdrá cero en toda la recta real excepto en los puntos de $S_d = \{1,2,3,4,5,6\}$. En dichos puntos la función vale

$$f_X(i) = P(X=i) = \frac{1}{6}, \quad i=1,2,3,4,5,6.$$

La representación gráfica de $f_X(x)$ es la que se presenta.

<u>Ejemplo</u>.- Supongamos una familia con 4 hijos y supongamos también que las probabilidades de nacimiento de varón o hembra son iguales e iguales a $1/2$. El espacio muestral de las posibles combinaciones de sexos entre los hijos de la familia está constituido por 16 puntos, a saber:

$$\Omega = \{VVVV, VVVH, VVHV, VHVV, HVVV, VVHH, \dots\dots\dots, HHHH\}$$

V-varón

H-hembra

Dotamos a Ω de la σ-álgebra de sus partes y definimos sobre ella la medida de probabilidad clásica, de manera que $P(A) = \frac{n}{16}$, n=número de puntos de A, $A \in \mathcal{A}$.

Sea ahora X una variable aleatoria definida mediante, $X: \Omega \to R$, $X(w)$ = nº de varones en w. La variable así definida será discreta y $S_d = \{0,1,2,3,4\}$. Veamos cómo son sus funciones de cuantía, f_X, y de distribución, F_X.

<u>Función de cuantía</u>: ya sabemos que $f_X(x) = 0$, si $x \notin S_d$. Para $x=n$, $n=0,1,2,3,4$, tendremos (12)

$$f_X(n) = P(X=n) = P(\text{haya } n \text{ varones en la familia}) = P(A_n)$$

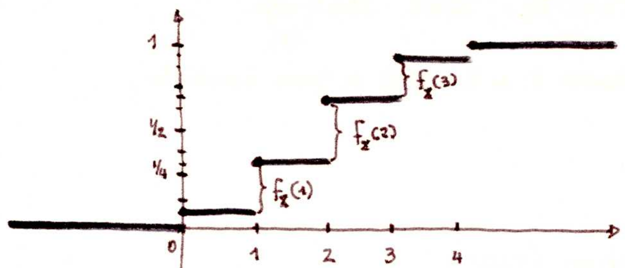
el número de casos posibles es 16, mientras que los casos favorables serán las posibles elecciones de n nacimientos (etiquetados para varones) entre los 4 que han ocurrido, es decir, $\binom{4}{n}$. Así pues

$$f_X(n) = P(A_n) = \binom{4}{n} \cdot \frac{1}{16} = \binom{4}{n}\left(\frac{1}{2}\right)^n \cdot \left(\frac{1}{2}\right)^{4-n}, \quad n=0,1,2,3,4$$

concretamente

$$f_X(0) = \frac{1}{16}, \quad f_X(1) = \frac{1}{4}, \quad f_X(2) = \frac{6}{16}, \quad f_X(3) = \frac{1}{4}, \quad f_X(4) = \frac{1}{16}.$$

<u>Función de distribución</u>: Utilizando la relación existente entre f_X y F_X, tendremos:



$F_X(x) = \sum_{n \leq x} f_X(n)$, así

$F_X(x) = 0$, si $x < 0$

$F_X(x) = \frac{1}{16}$, si $0 \leq x < 1$

$F_X(x) = \frac{5}{16}$, si $1 \leq x < 2$

$F_X(x) = \frac{11}{16}$, si $2 \leq x < 3$

$F_X(x) = \frac{15}{16}$, si $3 \leq x < 4$

$F_X(x) = 1$, si $x \geq 4$.

Esta variable aleatoria es un caso particular de una familia de variables aleatorias a las que se conoce generalmente como <u>variable aleatoria binomial</u> y son muy útiles en el estudio de fenómenos que obedecen a un determinado esquema, <u>el modelo binomial</u>. De una y otra nos ocuparemos a continuación cuando estudiemos algunos casos concretos de variables aleatorias discretas.